



FROGs: a daily $1^{\circ} \times 1^{\circ}$ gridded precipitation database of rain gauge, satellite and reanalysis products

5 Rémy Roca¹, Lisa V. Alexander^{2,3}, Gerald Potter⁴, Margot Bador^{2,3}, Rômulo Jucá⁵, Steefan Contractor^{2,6}
Michael G. Bosilovich⁴ and Sophie Cloché⁷

¹Laboratoire d'Etudes Géophysiques et d'Océanographie Spatiales, Toulouse, France

²Climate Change Research Centre, UNSW Sydney, Australia

10 ³ARC Centre of Excellence for Climate Extremes, UNSW Sydney, Australia

⁴NASA Goddard Space Flight Center, Greenbelt, Maryland

⁵Geoscience Environnement Toulouse, Toulouse, France

⁶School of Mathematics and Statistics, UNSW Sydney, Australia

⁷IPSL, Palaiseau, France

15

Correspondence to: Rémy Roca (remy.roca@legos.obs-mip.fr)

20



Abstract.

We introduce the Frequent Rainfall Observations on GridS (FROGS) database (Roca et al. 2019). It is composed of gridded daily precipitation products on a common $1^{\circ} \times 1^{\circ}$ grid to ease intercomparison and assessment exercises. The database includes satellite, ground-based and reanalysis products. As most
5 of the satellite products rely on rain gauges for calibration, unadjusted versions of satellite products are also provided where available. Each product is provided over its length of record and up to 2017 if available. Quasi-global, quasi-global land only, ocean only as well as tropical only and regional products (over continental Africa and South America) are included. All products are provided on a common netCDF format that is compliant with CF and ACDD standards. Preliminary investigations of
10 this large ensemble indicate that while a lot of features appear robust across the products, the characterization of precipitation extremes exhibit a large spread calling for careful selection of the products used for scientific applications. All datasets are freely available via an ftp server and identified thanks to the DOI: 10.14768/06337394-73A9-407C-9997-0E380DAC5598.



1. Introduction

Precipitation is a key element of the water and energy cycle. Observational efforts to document precipitation have a long history (Park et al., 2017) and have matured rapidly in recent decades.

5 Historically available *in situ* archives and associated gridded products (e.g., Becker et al. 2013) are being complemented by the burgeoning capability of satellite observations (Levizzani et al., 2018).

Reanalyses precipitation is not only derived from the model physics, but from a short term forecast, and reflect observed atmospheric variability (Bosilovich et al. 2008). While monthly and/or large spatial scale observations have been available for some time, recent progress permits documentation of

10 precipitation at finer space and time scales that consequently allow us to address new challenges such as extreme precipitation (Westra et al., 2013).

Indeed, recent processing of global scale ground-based archives has led to new datasets at the $1^\circ \times 1^\circ$ spatial resolution and daily frequency (Ziese et al. 2018, Contractor et al. 2019). A large variety of

15 satellite-based estimates of precipitation at various scales and over various record lengths and spatial coverage have been or are being produced by a growing number of agencies, laboratories and

consortiums as monitored and summarized within the International Precipitation Working Group (IPWG) (Levizzani et al., 2018). Finally, the reanalysis community is sustaining a large set of

reanalyses, offering products at the $1^\circ \times 1^\circ$ daily scale (Potter et al., 2018). This large, unique collection

20 of gridded observational datasets opens the possibility of building an ensemble of products similar to the approach taken by the coupled climate model community. This avoids the need to choose only one product for research applications such as model evaluation. However, it is also important to assess the robustness of the various data sources and understand their differences in order to appreciate the uncertainties within the ensemble and guide its formulation.

25



Recent GEWEX led assessments have paved the way for efficient and useful coordinated inter comparison/validation exercises: the Cloud Assessment (Stubenrauch et al., 2013) and the Water Vapor Assessment (Schröder et al., 2016, 2019). One of reasons for these successes lies in the making of a common gridded database, facilitating the handling of many datasets that were originally available on various grids, resolutions, formats, data types, etc. The Cloud Assessment database encompasses 23 products, mainly remotely sensed, and is now being extended in time and in the number of products available (Stubenrauch, 2018 personal communication). Within the Water Vapor Assessment, precipitable water spans 22 products, originating from *in situ* observations, satellite measurements and reanalysis output (Schröder et al., 2018). In both assessments though, a common publicly available database has been provided for further analysis after the main assessment effort. Here we follow the legacy of the GEWEX Data Analysis Panel (GDAP) assessments by building a database of daily precipitation with data originating from rain gauges, satellite and reanalysis products. The database, Frequent Rainfall Observations on Grids (FROGs), is being released at the beginning of the assessment. This will help to integrate new investigations and its overall assessment recently initiated under the auspices of GEWEX/GDAP and IPWG (Haddad and Roca, 2017). This also includes a dedicated effort to analyse extreme events and assess their characteristics under the joint WCRP Grand Challenge on Weather and Climate Extremes/GEWEX GDAP project on “Precipitation Extremes” (Alexander et al., 2018), with a Special Issue being developed on the topic.

The aim of this paper is to introduce the FROGs database that includes ground-based, satellite and reanalysis products, gridded to a common $1^{\circ} \times 1^{\circ}$ daily resolution format in support of the above activities. FROGs is first presented by type of observational source. For each precipitation product we summarize its main characteristics and the regridding methodology used along with some known pros and cons. Section 3 presents some illustrations of the large ensemble of precipitation from FROGs. Section 4 gives the technical details of the database files including how to access the data and how to reference the data using DOI. We end with a conclusion and outlook section.



2. The database

2.2 The ground-based products

5 In situ products are seen by many as offering a 'ground truth' but this is not necessarily the case since they are gridded estimates based on various interpolation algorithms which use incomplete station networks of varying quality and density. All products have their pros and cons especially when it comes to extreme precipitation estimates and all are limited in data coverage over certain land regions e.g. Africa. Some include error estimates or quality masks while others do not. Here we have chosen to
10 include products in FROGs which offer daily global or quasi-global land estimates. All products are either already available on a $1^\circ \times 1^\circ$ grid or have been re-interpolated onto that resolution using simple averaging. The list of products is summarized in Table 1 and below we present an overview of each individual product.

15 2.2.1 NOAA CPC Global Daily analysis

The NOAA CPC Unified gauge-based analysis of global daily precipitation (Chen et al. 2008; Xie et al., 2010) interpolates data from over 30,000 stations over land onto a regular grid with dedicated quality control (QC) and accounting for topographic biases (Xie et al. 2007). The QC is performed through
20 historical record comparisons and neighbour checks, concurrent radar/satellite observations and also utilises numerical model forecasts. The gauge reports come from multiple sources including GTS, COOP and other national and international agencies. The daily analysis is constructed on a $0.125^\circ \times 0.125^\circ$ grid over the entire global land area but is released on a $0.5^\circ \times 0.5^\circ$ grid for the period since 1979. The high resolution and station density are a key strength of the dataset but the quality of
25 the gauge-based analysis is poor in data sparse regions and accumulations can differ from country to country based on time of observation leading to potential discontinuities across national boundaries. A



possible inhomogeneity in the early part of the record has also been recorded with a large decrease in total precipitation and number of wet days between 1981 and 1982 (Herold et al. 2016). There are two versions of the dataset: (i) a "retrospective version" available from 1979-2005 which uses all the available ~30,000 stations and (ii) a "real-time version" which uses ~17000 stations and spans 2006-5 present. The main changes being over the CONUS region. The two versions remain temporally homogeneous (Xie P-P, Personal communication, 2019) For FROGS, the two 0.5°x0.5° resolution products are merged and are then averaged onto the common 1°x1° grid using a simple arithmetical mean calculation.

10 2.2.2 The Global Precipitation Climatology Centre (GPCC) daily products

The Global Precipitation Climatology Centre (GPCC) builds a suite of gridded precipitation products based on rain gauge measurements and comprehensive quality control (Becker et al., 2013). In particular a full global 1°x1° daily analysis is available as well as a first guess analysis product 15 (Schamm et al., 2014). GPCC products have the advantage that they can access more data than available to other products but due to data restrictions they cannot share raw station data (Becker et al., 2013). Here, the GPCC Full Data Daily V1 (GPCC-FDDv1) product (Schamm et al., 2015) is made available up to December 2013. In GPCC-FDDv1 station data were interpolated using ordinary block Kriging, a stochastic interpolation method which accounts for the statistical structure of precipitation in terms of a 20 distance-weighted spatial autocorrelation function. The daily precipitation estimates represent area averages which results in estimates directly comparable to other forms of data that produce area average estimates such as satellite products. Also included in FROGS is the recent Full Data Daily V2018 (GPCC-V2018) product, which was released in June 2018 and covers 1982 to 2016 (Ziese et al., 2018). In addition to more data and a more advanced quality control, the main difference between GPCC- 25 V2018 compared to GPCC-FDDv1 is that a modified SPHEREMAP interpolation scheme (Becker et al., 2013) is used rather than ordinary block Kriging to align with the method that is applied to GPCC's



monthly products. GPCP recommend the GPCP-V2018 product for analyses of extreme events and related statistics at daily resolution however to date there have been no independent analyses to confirm this. Finally, the first guess product (Schamm et al., 2015), based on NRT stations is also used for the recent period from 2009-2016.

5 2.2.3 REGEN datasets

REGEN is the name given to a set of daily land-based precipitation datasets created through a collaboration with the University of New South Wales (UNSW), GPCP and NOAA's National Center for Environmental Information (NCEI). There are two related datasets that are currently available on a
10 $1^\circ \times 1^\circ$ daily grid resolution (Contractor et al. 2019): (i) "REGEN-All" which interpolates all available station data and (ii) "REGEN-Long" which only considers stations that have a minimum of 40 years of data (long term stations only). The REGEN products combines the GPCP *in situ* data used to create the $1^\circ \times 1^\circ$ GPCP-FDDv1 product described above with data from the Global Historical Climatology Network-Daily (GHCN-D) and other sources which results in unprecedented station density and length
15 of record (since 1950) compared to other existing products. The REGEN-All dataset contains an average of over 50,000 station records per day. The two REGEN products use the ordinary block Kriging algorithm described in Schamm et al. 2015 to interpolate the data onto a $1^\circ \times 1^\circ$ grid resolution. The gridded fields are also supplemented with metadata including number of observations, standard deviation, Kriging error and data quality mask. The addition of metadata is a key strength of REGEN-
20 All along with the number of observations utilised. However, the dataset does suffer from a varying station network over time (e.g. stations per day double in North America after 2000, decrease substantially in South America from the late 1990s and in India since 1970s). However as this is a very new dataset testing has been limited and so the effect of network changes has not been fully explored. The version V1-2019 of REGEN products is included in the FROGS database.

25

2.3 The satellite-based products



Most of the “satellite” precipitation estimation products make use of ancillary, non-satellite data in their estimations and these enriched products are regarded as the best estimate. Nevertheless, for most of these products, an unadjusted version is also available and is included here. Hardly any dataset is truly global and we present below the quasi-global land and ocean datasets as well as the ocean only, the land
5 only quasi-global data products currently available in FROGS. The database is completed with some regional products including tropical land and oceans or ones covering only continental Africa and South America. The list of products is summarized in Table 2 as well as information regarding their respective spatial/temporal coverage.

10 Most of the products are available at the daily scale and so this is the version we use. The day is defined over the 00Z-00Z time frame for each of the datasets unless otherwise specified. The daily data are then regridded onto a $1^{\circ} \times 1^{\circ}$ regular grid covering the whole earth. We currently use a simple arithmetical averaging (no interpolation) going from the small scale in degrees to the large scale in degrees that allows a conservative average of precipitation at $1^{\circ} \times 1^{\circ}$.

15

2.3.1 The global and quasi-global land and ocean datasets

3B42 v7.0 by NASA

20 The 3B42 v7 product (Huffman et al. 2009; Huffman et al., 2011) is a reference product in various previous studies of tropical rainfall distribution (Maggioni et al., 2016). It also exemplifies what a dataset that is highly geared towards microwave data (imagers and 183GHz sounders) can provide in terms of daily accumulation. It also puts to perspective the use of scattering based retrieval over land for instantaneous retrieval (Gopalan et al., 2010). The combined TRMM radar-imager product (Haddad et
25 al., 1997) serves as a reference for other microwave instruments prior to merging. Geostationary based IR imagery is also used in the algorithm to ensure observations when no low earth observing satellites



are available. The technique also relies on a sophisticated bias correction approach that relies on GPCP monthly analysis over land. While this popular product has been evaluated over a very large number of small regions and catchments as well as over the whole tropics for certain metrics (Sun et al., 2017), it still lacks systematic intercomparison with the whole suite of products presented here. As a consequence, it is included here even though NASA has announced the discontinuation of its production post 2018. As a complement to the gauges adjusted product, the microwave-calibrated IR estimates and the microwave only estimates are also provided. Along with this reference product, NASA has also been releasing a low latency version called 3B42RT where RT is for real-time for which no gauges data are used. Two versions are provided here, the actual products and the uncalibrated one (Huffman et al., 2009). The suite of 3B42 products is at the core of the constellation-based family of satellite rainfall products.

GsMAP v6 with and without gauges by JAXA

The Global Satellite Mapping of Precipitation product provides high resolution precipitation estimation using satellite observations from multiple platforms (Kubota et al., 2007; Mega et al., 2018). This product is mainly based on microwave estimation of rainfall for a suite of microwave imagers and sounders. The suite of GsMAP products hence belongs to the constellation-based family of satellite rainfall products. The microwave instantaneous rain rate estimates (Aonashi et al. 2009) are propagated based on Cloud Motion Wind vectors originally derived from IR geostationary imagery to yield to a gridded high resolution precipitation product (Ushio et al. 2009). GsMAP belongs to the morphing based microwave algorithm, like CMORPH (Joyce et al., 2004) or SearchLight (Bellerby, 2013). To complement the satellite-only estimation, the product is further scaled to rain gauges estimates to correct for some bias over land. Two sets of products are included in the database: unadjusted and adjusted. Two versions of the GsMAP products are provided here: the so-called reanalysis and the near real time versions of the products that differs in the amount of data they used in the processing. Owing



to the production schedule, the homogeneous processing of the reanalysis data has been performed from mid-2000 up to April 2014. As a consequence, we have restricted our use to the processed full years from 2001 to 2013 inclusive. While the original product is offered at two daily average range (00-00Z and 12-12Z), here for the sake of homogeneity, only the 00-00Z daily average is provided in the
5 ensemble database. The near real time dataset extends up to 2017.

CMORPH v1.0 RAW and CRT by NOAA

The CMORPH product (Joyce et al., 2004; Xie et al., 2017) belongs to the microwave based morphing
10 algorithms like GsMAP (Kubota et al., 2007). The microwave derived instantaneous rain rate estimates from multiple platforms are propagated using Cloud Motion Wind vectors originally derived from IR geostationary imagery and a Kalman filter (Joyce and Xie, 2011). Such an approach results in a high-resolution precipitation product. The CMORPH products belong to the constellation-based family of satellite rainfall products and both microwave imagers and sounders are used. The instantaneous
15 imagers based rain rates are obtained using GPROF 2004 (Kummerow et al., 2001) while the sounders estimation rely on the algorithm of (Ferraro et al., 2005). Two versions are provided with (CRT) or without (RAW) gauges adjustment. The adjustment is performed using PDF matching. Over land, the daily CPC gauges analysis is used for this correction (Xie et al., 2003) while over ocean, the adjustment is done using the GPCP pentad merged product. The product is thought to perform well overall, with a
20 small bias relative to the gauges; yet experiences difficulty with snow and cold season rainfall (Xi et al., 2017).

GPCP IDD CDR v1.3 by NOAA

25 The Global Precipitation Climatology Product CDR Version 1.3 daily product (Huffman et al. 2001) is another reference product used in various previous studies (e.g. Alder 2017 and references therein). The



GPCP CDR dataset is the only global product in the database. It is adapted from the Geostationary Operational Environmental Satellite Precipitation Index technique with monthly, local adjustments. The approach merges IR imagery from geostationary and polar platforms. It relies on the use of one single microwave platform and the Level 2 retrievals of Kummerow et al. (1996). A bias adjustment scheme is finally used over land that relies on rain gauge data from the GPCC database at the monthly scale. Over the high latitudes, GPCP incorporates IR based precipitation estimations in complement to microwave derived rain rates for the lower latitudes. The original data file from NOAA contains a valid range attribute between 0-100 mm/d. Values beyond 100mm/d are nevertheless found in the dataset. Two versions are provided here: 1) one where the valid range attribute is not enforced so original values extending beyond the valid range are kept in the analysis (Alder, 2019 personal communication) and 2) a version in which the valid range is enforced.

PERSIANN CDR v1 by NOAA

PERSIANN-CDR v1 is a quasi-global IR-based product, trained over radar data in the US, and normalized to GPCP monthly totals (Ashouri et al., 2015). It can be thought of as an alternative daily downscaling of the GPCP monthly data to that of GPCP 1DD CDR. Despite sharing monthly totals, the two products differ substantially in their estimation of the daily precipitation distribution (Sun et al., 2017) and so are included in the database. The product does not rely on passive microwave data and as a consequence extends back further in time than GPCP 1DD. PERSIANN-CDR has been extensively evaluated over various regions and was shown to provide mixed levels of agreement with observations from local rain gauge networks (Miao et al., 2015; Tan and Santo, 2018). PERSIANN-CDR is the climate monitoring oriented product of the PERSIANN family of datasets that can otherwise be accessed at the CHRS data portal (Nguyen et al., 2019).

25



2.3.2 The quasi-global ocean only dataset

HOAPS v4.0 by CMSAF

5 The “Hamburg Ocean-Atmosphere Parameters and Fluxes from Satellite Data” product is described at length in (Andersson et al., 2010). This product relies on recalibrated and inter calibrated measurements from SSM/I and SSMIS passive microwave radiometers (Fennig et al., 2013) to estimate a suite of fresh water budget elements globally (80°S-80°N) over sea-ice free ocean surface, including precipitation. Here version 4 of the product is provided (Andersson et al., 2017) but corresponds to version 3.2 for the
10 precipitation algorithm. The precipitation retrieval is based on a neural network technique that relies on the polarized brightness temperature measurements of the conical scanning imager. The neural network is trained on ECMWF inputs and radiative transfer simulations. Unlike other similar products, HOAPS precipitation appears to detect snowfall well during the cold season (Klepp et al., 2010). The orbit data have been regridded on the common 1°x1° daily grid courtesy of M. Schröder.

15

2.3.2 The quasi-global land only dataset

CHIRP/CHIRPS v2.0 by UCSB

20 The Climate Hazards Infrared Precipitation (CHIRP) and the Climate Hazards Infrared Precipitation with Stations (CHIRPS) are satellite based precipitation product (Funk et al., 2015). While CHIRP is satellite only, CHIRPS also benefits from station data from 5 public data sources (GHCN monthly and daily, Global Summary Of the Day, GTS daily, Southern African Science Service Centre for Climate Change and Adaptative Land Management) as well as private datasets from various countries in the
25 world (see Funk et al., 2015 for the list). As a result, the density of gauges in the final product varies significantly in time as well as space. CHIRP uses the infrared observations from geostationary observations in a GOES GPI modified approach and various ground based and alternative sources (an



in-house climatology, 3B42, Climate Forecasts Systems outputs) for its calibration at the monthly scale.

It is considered as a rain gauges free or satellite only product. Then the CHIRPS estimates are obtained by merging the stations with the CHIRP estimates using a weighted average of the closest stations and CHIRP results for each 0.05° grid point. The unique characteristics of the CHIRPS product are its

5 native high resolution, low latency and long record (the longest of the satellite data in the database, Table 2). The product has been used in many evaluation and process studies and is thought to support hydrological forecasts and trends analysis in Ethiopia for instance (Pricope et al., 2013). CHIRPS is a quasi-global implementation of algorithms and methodology that also have been implemented at regional scales like the TAMSAT products (see below).

10

SM2RAIN-CCI by ESA

While all of the other products are based on indirect measurements more or less, this product actually relies on very indirect evidence of precipitation by relating satellite-based estimations of soil moisture

15 to the precipitation that impacted the surface. This product is based on the SM2RAIN algorithm (Brocca et al., 2013). The algorithm is applied onto the active and passive ESA Climate Change Initiative soil moisture datasets (Ciabatta et al., 2018). It is an alternative way to use indirect satellite-based measurements to estimate rainfall. Note that due to soil moisture data quality issues, a mask is applied to the rainfall products and no estimates are provided over the tropical rainforest areas, frozen and snow
20 covered soil, rainforest areas, and areas with topographic complexity.

2.3.3 Tropical land and ocean dataset

TAPEER v1.5 by AERIS

25



The recently released TAPEER product is based on the universally adjusted Geostationary Operational Environmental Satellite precipitation index technique (Xu et al. 1999) that merges geostationary infrared imagery with microwave instantaneous rain rates estimates at daily local scales to yield the daily precipitation accumulation (Kidd et al., 2003). The current implementation relies on the BRAIN
5 L2 dataset (Viltard et al., 2006) for a suite of conical microwave imagers and include the SAPHIR data from the Megha-Tropiques mission (Roca et al., 2015) for rainfall detection and is available at $1^\circ \times 1^\circ$ resolution (Roca et al., 2018). Along with the accumulation, an estimation of the sampling uncertainty of the daily accumulation is provided (Roca et al., 2010; Chambon et al., 2012). The TAPEER product has been favourably compared against various datasets over tropical Africa (Guilloteau et al., 2016)
10 (Gosset et al., 2018). Unlike many other operational satellite precipitation products, the TAPEER estimations do not ingest nor are calibrated to any rain gauge datasets. As such it provides, one solution independent from the rain gauge network and with an enhanced tropical sampling thanks to the use of the SAPHIR data from the Megha-Tropiques mission. While the original product is offered at four daily average range (00-00Z, 06-06Z, 12-12Z and 18-18Z), here for the sake of homogeneity, only the 00-00Z
15 is provided in the ensemble database. This product belongs to the constellation-based family of satellite precipitation products.

2.3.3 Africa land only datasets

20 TAMSAT v2.0 and v3.0

The Tropical Applications of Meteorology using SATellite data and ground-based observation (version 2.0 and 3.0 (Maidment et al., 2017), is a product that provides rainfall estimates across Africa based both on geostationary thermal infrared (TIR) images, obtained every 15 min (30 min prior to June
25 2006), and on ground-based observations from the Global Telecommunications System (GTS). The TAMSAT algorithm is based on two primary data inputs: *i*) Meteosat TIR imagery provided by The



European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) and *ii*) rain gauge observations (daily accumulated, 0600-0600 UTC) for calibration. The general procedure follows three steps: a) algorithm calibration - at the decadal (version 2.0) and pentadal (version 3.0) time-step; b) estimation of the pentadal/decadal rainfall estimates; and c) estimation of daily rainfall estimates. The

5 TAMSAT daily rainfall estimates have a native resolution of 0.0375° (about 4 km) and covers all of Africa since January 1983 to the present.

ARC v2.0

10 The African Rainfall Climatology, version 2.0 (ARC2) it is a revision of the first version of the ARC and is consistent with the operational Rainfall Estimation Version 2 (RFE 2.0) (Novella and Thiaw 2013). The product is a composite of *i*) 3-hourly geostationary infrared (IR) data centred over Africa from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), and *ii*) quality-controlled 24-h (0600-0600 UTC) rainfall accumulation records from the Global

15 Telecommunication System (GTS) gauge database. The calibrated IR and the quality controlled GTS gauges are then combined following multiple criteria (i.e., the two-step merging process) to produce the final rainfall estimates. The ARC2 daily data set is updated regularly. The native resolution is $0.1^\circ \times 0.1^\circ$ over a spatial domain of 40°S – 40°N and 20°W – 55°E and over the period from 1 January 1983 to the present.

20

2.3.4 Latin America land only datasets

COSCH

25 The Combined Scheme approach CoSch (Vila et al., 2009) is a gauge-satellite-based precipitation product that provides daily gridded estimates over Latin America. The general procedure for satellite-



gauge merging and data production involves the follow tasks: *i*) obtain and run quality control of global and regional rain gauge data, from GTS and multiple institutions, respectively; *ii*) reprocess the daily accumulated satellite-based rainfall fields, following the same time accumulation as the rain gauges (1200-1200 UTC); and *iii*) apply the additive and multiplicative bias correction schemes for each station

5 on a daily basis. The CoSch actual product-version uses the Real Time TRMM Multisatellite Precipitation Analysis (TMPA-RT, Huffman, 2007) (Version 7) as a high-quality satellite rainfall algorithm. The CoSch daily rainfall estimates database is available from March 2000 to the present and its native spatial resolution is 0.25° over the Latin America land areas.

10 2.4 The reanalysis products

Atmospheric reanalyses blend observed meteorological state fields (temperature, humidity, wind and pressure) with a global weather model through assimilation to provide a continuous representation of not only the state fields, but also the model generated fields. Precipitation is one such model derived but

15 observationally guided field. Typically, reanalysis precipitation is considered to have more uncertainty than the analysed state fields (e.g. Kalnay et al. 1996). However, precipitation is a key quantity in both the reanalyses representations of global water and energy cycles (through the latent heat of condensation) and so should be understood (e.g. Bosilovich et al. 2008). There are few studies inter

comparing many reanalyses daily precipitation, though Bosilovich et al (2009) found distinctly different

20 distributions among a collection of 10 analyses and reanalyses (focusing on gauge data over the United States). Even for a given weather event, the distribution of the precipitation can have large variance. Shiu et al (2102) results suggest that reanalyses can reproduce the temperature-precipitation relationship as temperature warms, but the more recent reanalyses had higher variance than the older generation.

The long-term collection of daily reanalyses precipitation here will help characterize and understand the

25 state of the reanalyses ability to reproduce the high frequency occurrences of extreme precipitation.



The list of products is summarized in Table 3 and below we detail the common grid and present each individual product.

5 ERA-Interim

The ECMWF Interim Reanalysis (ERA-Interim, Dee et al. 2011) was developed to test the recent advancement of the forecast model and assimilation development beyond ERA-40 (Uppala et al. 2005), especially in the representation of the hydrologic cycle. This included advances in the humidity
10 analysis, radiance bias correction, and cloud parameterization, which are crucial for the representation of the water vapour state and generation of precipitation. While the large-scale representation of the precipitation has improved over ERA-40, some differences from observed data can be found (Simmons et al. 2010).

15 MERRA and MERRA 2

Throughout the development of both MERRA (Rienecker et al. 2011) and MERRA-2 (Gelaro et al. 2017), the water cycle representation was identified as a key component to understanding weather and climate. Significant improvements were included in the model (Molod et al. 2015) and the water vapor
20 analysis (Takacs et al. 2016). While the influence of observing system changes is still apparent in MERRA-2 (Bosilovich et al. 2017), and there are some significant regional biases (e.g. tropical land topography overestimates), there is indication that the extreme end of the distribution is significantly improved in MERRA-2 over MERRA for the continental United States (Bosilovich et al. 2015). The observations evaluated here will allow the testing of these improvements in other regions in reanalyses.

25



JRA 55

Kobayahi et al (2015) provide the details of the Japanese 55-year Reanalysis (JRA-55). This version introduced 4D variational analysis extending in time beyond the introduction of satellite data for

5 weather analysis (back to 1958). Wind profile retrievals for tropical cyclones were assimilated, and provide a significant contribution to the analysis of tropical cyclones. While some improvements have been noted in the stability of the precipitation time series and certain water vapour biases, the JRA-55 mean precipitation tends to be high, attributed to a dry model bias and spin-down effect of the forecast following reinitialization.

10

CFSR

The National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis

(CFSR) was developed to provide initial conditions for continuing seasonal predictions, as well as for

15 climate studies (Saha et al. 2010). At a horizontal resolution of 38km, the representation of the modelled precipitation will be the highest resolution source reanalysis data included here. While high resolution should provide improved locations of precipitation events and structural patterns, the CFSR also uses observation corrected precipitation for forcing its land surface model. This was done to provide the best surface forcing and soil moisture for the subsequent forecasts. As with the other reanalyses here, the
20 influence of changing observations, especially the addition of ATOVS radiances, significantly affects the mean precipitation of CFSR (Zhang et al. 2012).

Regridding method

25 All of the data for the reanalyses (MERRA, MERRA-2, CFSR, and JRA-55, ERA-Interim) were obtained from the CREATE service (Potter et al. 2018). These data are identically formatted with one



variable per file for both 6-hour and monthly. The 6-hour was then used to create the daily form and the data time was adjusted to have a 12-hour mid time. The files were also adjusted to have the same longitudinal wrap as GPCP. The files were regridded to $1^{\circ} \times 1^{\circ}$ using a bilinear remapping with the Climate Data Operators (CDO).

5

4 An illustration of the database

Figure 1 shows the annual mean precipitation time series all of the products and indicates the various time spans and spatial coverage of the products. This large ensemble of products is characterized by various trends in their depiction of the average precipitation evolution. Note that the regional products might not be compared directly with the quasi-global ones. Despite this there are some clear outliers and inhomogeneities in the products available. It is recommended that further work aims to understand these differences between the products through a concerted community intercomparison effort.

[Figure 1 here]

5 Data Availability

15

Data format

Files are produced within netCDF-4 format with metadata following the Climate and Forecast (CF) Convention version 1.6 and Attribute Convention for Dataset Discovery (ACDD) version 1.3. An example of the header of a product is provided in the Annex.

20 One file per product, per year, at the resolution $1^{\circ} \times 1^{\circ} \times 1$ -day. Each yearly file contains the following information as a minimum:

- **lon:** The longitude values of the grid, in degrees, ranging between $[-179.5; +179.5]$. The grid is centred, i.e., a value of $+0.5$ corresponds to the degree $[0W, 1W]$.
- **lat:** The latitude values of the grid, in degrees, ranging between $[-89.5; +89.5]$. The grid is centred, i.e., a value of $+0.5$ corresponds to the degree $[0N, 1N]$.
- **time:** The time values of the grid, following the standard CF calendar, in days, centred, since 01-01-1970, midnight.
- **rain:** The precipitation estimate, in $[\text{mm}/\text{day}]$. Missing values are represented with `NaN_F`.

25



For some products, extra information can be found in the files. For instance, the TAPEER product is completed with an estimate of the uncertainty of the daily precipitation.

Accessibility and DOI

The database (Roca et al., 2019) is referenced with the following DOI:

- 5 <http://dx.doi.org/10.14768/06337394-73A9-407C-9997-0E380DAC5598>. The DOI landing page provides the up to date how-to access the database as well as a number of useful references for the users.

3. Conclusions and outlook

- 10 For the first time we offer an easily accessible database of daily precipitation products on a common grid which we hope will prove invaluable for intercomparison, model evaluation (Tapiador et al., 2017, 2018) and other research purposes. In particular, FROGS offers an invaluable resource to study precipitation extremes and to help us understand some of the uncertainties that are inherent across all precipitation products. A few studies based on this database are already under consideration in various
15 journals with a focus on extreme precipitation (Special Issue in ERL) or water budget closure analysis (Blasquez et al., 2019). This is a 'living' database, that is, products will continue to be added with time. This includes the latest release of IMERG (Huffman et al., 2015), MSWEP (Beck et al., 2018) and possibly regional *in situ* products (e.g. APHRODITE, E-OBS, AWAP etc). How the database continues to evolve will be shaped by the needs of the community and the feedback from various on-going
20 assessments. Similarly, efforts will be geared towards updating the database with the most recent years as they become available.

Acknowledgements



We kindly acknowledge the participants the WCRP meeting on Extreme Precipitation at DWD in Offenbach, 2018 who provided a strong incentive for this database to be finalized. Dr A. Guérou's help at an early stage of this effort is appreciated. All the data providers are also acknowledged for making their data freely available. We further thank the GSMAP, GPCP, GPCC, CPC and CHIRPS teams for enriching exchanges about their products. We thank Drs. M Schroder, A. Ramos for helpful discussions on the data.

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(2), 215–234, doi:10.5194/essd-2-215-2010, 2010.
- Andersson, A., Graw, K., Marc, S., Fennig, K., Liman, J., Bakan, S., Hollmann, R. and Klepp, C.: Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data - HOAPS 4.0, Satellite Application Facility on Climate Monitoring, , doi:: DOI:10.5676/EUM_SAF_CM/HOAPS/V002, 2017.
- AONASHI, K., AWAKA, J., HIROSE, M., KOZU, T., KUBOTA, T., LIU, G., SHIGE, S., KIDA, S., SETO, S., TAKAHASHI, N. and TAKAYABU, Y. N.: GSMaP Passive Microwave Precipitation Retrieval Algorithm : Algorithm Description and Validation, *J. Meteorol. Soc. Japan*, 87A(June 2008), 119–136, doi:10.2151/jmsj.87A.119, 2009.
- Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R. and Prat, O. P.: PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies, *Bull. Am. Meteorol. Soc.*, 96(1), 69–83, doi:10.1175/BAMS-D-13-00068.1, 2015.
- Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar, T. R. and Adler, R. F.: MSWEP V2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment, *Bull. Am. Meteorol. Soc.*, doi:10.1175/BAMS-D-17-0138.1, 2018.
- 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(1), 71–99, doi:10.5194/essd-5-71-2013, 2013.
- Bellerby, T. J.: Searchlight: Precipitation advection tracking using multiplatform low-earth-orbiting satellite data, *IEEE Trans. Geosci. Remote Sens.*, 51(4), 2177–2187, doi:10.1109/TGRS.2012.2211604, 2013.
- Brocca, L., Moramarco, T., Melone, F. and Wagner, W.: A new method for rainfall estimation through soil moisture observations, *Geophys. Res. Lett.*, 40(5), 853–858, doi:10.1002/grl.50173, 2013.
- Chen, M., Shi, W., Xie, P., Silva, V. B. S., Kousky, V. E., Higgins, R. W. and Janowiak, J. E.: Assessing objective techniques for gauge-based analyses of global daily precipitation, , 113, 1–13, doi:10.1029/2007JD009132, 2008.
- Ciabatta, L., Massari, C., Brocca, L., Gruber, A., Reimer, C., Hahn, S., Paulik, C., Dorigo, W., Kidd, R. and Wagner, W.: SM2RAIN-CCI: A new global long-term rainfall data set derived from ESA CCI soil moisture, *Earth Syst. Sci. Data*, 10(1), 267–280, doi:10.5194/essd-10-267-2018, 2018.
- Fennig, K., Andersson, A. and Schröder, M.: Fundamental Climate Data Record of SSM/I Brightness Temperatures, , doi:DOI:10.5676/EUM_SAF_CM/FCDR_SSMI/V001, 2013.
- Ferraro, R. R., Weng, F., Grody, N. C., Zhao, L., Meng, H., Kongoli, C., Pellegrino, P., Qiu, S. and Dean, C.: NOAA operational hydrological products derived from the advanced microwave sounding unit, *IEEE Trans. Geosci. Remote Sens.*, 43(5), 1036–1048, doi:10.1109/TGRS.2004.843249, 2005.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelsen, J.: The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes, *Sci. Data*, 2, 1–21, doi:10.1038/sdata.2015.66, 2015.
- Gopalan, K., Wang, N. Y., Ferraro, R. and Liu, C.: Status of the TRMM 2A12 land precipitation



- algorithm, *J. Atmos. Ocean. Technol.*, 27(8), 1343–1354, doi:10.1175/2010JTECHA1454.1, 2010.
- Guiloteau, C., Roca, R. and Gosset, M.: A Multiscale Evaluation of the Detection Capabilities of High-Resolution Satellite Precipitation Products in West Africa, *J. Hydrometeorol.*, 17(7), 2041–2059, doi:10.1175/jhm-d-15-0148.1, 2016.
- 5 Haddad, Z. S. . E. A. S. C. D. K. T. I. M. R. F. S. L. D. M. A. and W. S. O.: The TRMM ‘Day-1’ Radar/Radiometer Combined Rain-Profiling Algorithm, *J. Meteor. Soc. Japan*, 75, 799–809, 1997.
- Huffman, George J., Adler, Robert F., Morrissey, Mark M., Bolvin, David T., Curtis, Scott, Joyce, Robert, McGavock, Brad and Susskind, Joel: Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations, *J. Hydrometeorol.*, 2(1), 36–50, doi:10.1175/1525-10 7541(2001)002<0036:GPAODD>2.0.CO;2, 2001.
- Huffman, G. J., Adler, R. F., Bolvin, D. T. and Nelkin, E. J.: The TRMM Multi-Satellite Precipitation Analysis (TMPA), *Satell. Rainfall Appl. Surf. Hydrol.*, 3–22, doi:10.1007/978-90-481-2915-7_1, 2009.
- Huffman, G. J., Bolvin, D. T. and Nelkin, E. J.: Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation, NASA/GSFC Code, 612(2015), 47, 2015.
- 15 Joyce, R. J. and Xie, P.: Kalman Filter–Based CMORPH, *J. Hydrometeorol.*, 12(6), 1547–1563, doi:10.1175/JHM-D-11-022.1, 2011.
- Joyce, R. J., Janowiak, J. E., Arkin, P. A. and Xie, P.: CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution, *J. Hydrometeorol.*, 5(3), 487–503, doi:10.1175/1525-20 7541(2004)005<0487:CAMTPG>2.0.CO;2, 2004.
- Kidd, C., Kniveton, D. R., Todd, M. C. and Bellerby, T. J.: Satellite Rainfall Estimation Using Combined Passive Microwave and Infrared Algorithms, *J. Hydrometeorol.*, 4(6), 1088–1104, doi:10.1175/1525-7541(2003)004<1088:sreucp>2.0.co;2, 2003.
- Klepp, C., Bumke, K., Bakan, S. and Bauer, P.: Ground validation of oceanic snowfall detection in



- satellite climatologies during LOFZY, *Tellus, Ser. A Dyn. Meteorol. Oceanogr.*, 62(4), 469–480, doi:10.1111/j.1600-0870.2010.00459.x, 2010.
- Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y. N., Ushio, T., Nakagawa, K., Iwanami, K., Kachi, M. and Okamoto, K.: Global Precipitation Map
- 5 Using Satellite-Borne Microwave Radiometers by the GSMaP Project: Production and Validation, *IEEE Trans. Geosci. Remote Sens.*, 45(7), 2259–2275, doi:10.1109/tgrs.2007.895337, 2007.
- L., X., X., G., S., S., P.A., A. and B., I.: A Microwave Infrared Threshold Technique to Improve the GOES Precipitation Index, *J. Appl. Meteorol.*, 38(5), 569–579, 1999.
- Levizzani, V., Kidd, C., Aonashi, K., Bennartz, R., Ferraro, R. R., Huffman, G. J., Roca, R., Turk, F. J.
- 10 and Wang, N.-Y.: The activities of the International Precipitation Working Group, *Q. J. R. Meteorol. Soc.*, (September 2017), doi:10.1002/qj.3214, 2018.
- Maggioni, V., Meyers, P. C. and Robinson, M. D.: A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era, *J. Hydrometeorol.*, 17(4), 1101–1117, doi:10.1175/jhm-d-15-0190.1, 2016.
- 15 Maidment, R. I., Grimes, D., Black, E., Tarnavsky, E., Young, M., Greatrex, H., Allan, R. P., Stein, T., Nkonde, E., Senkunda, S. and Alcántara, E. M. U.: A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa, *Sci. Data*, 4, 1–19, doi:10.1038/sdata.2017.63, 2017.
- Mega, T., Ushio, T., Matsuda, T., Kubota, T., Kachi, M. and Oki, R.: Gauge-Adjusted Global Satellite Mapping of Precipitation, *IEEE Trans. Geosci. Remote Sens.*, PP, 1–8,
- 20 doi:10.1109/TGRS.2018.2870199, 2018.
- Miao, C., Ashouri, H., Hsu, K.-L., Sorooshian, S. and Duan, Q.: Evaluation of the PERSIANN-CDR Daily Rainfall Estimates in Capturing the Behavior of Extreme Precipitation Events over China, *J. Hydrometeorol.*, 16(3), 1387–1396, doi:10.1175/JHM-D-14-0174.1, 2015.
- Nguyen, P., Shearer, E. J., Tran, H., Ombadi, M., Hayatbini, N., Palacios, T., Huynh, P., Braithwaite,



- D., Updegraff, G., Hsu, K., Kuligowski, B., Logan, W. S. and Sorooshian, S.: The CHRS data portal, an easily accessible public repository for PERSIANN global satellite precipitation data, *Sci. Data*, 6, 1–10, doi:10.1038/sdata.2018.296, 2019.
- P., X., M., C., S., Y., A., Y., T., H., Y., F. and C., L.: A Gauge-Based Analysis of Daily Precipitation
5 over East Asia, *J. Hydrometeor.*, 8(3), 607–626, 2007.
- Park, K. J., Yoshimura, K., Kim, H. and Oki, T.: Chronological development of terrestrial mean precipitation, *Bull. Am. Meteorol. Soc.*, (November), BAMS-D-16-0005.1, doi:10.1175/BAMS-D-16-0005.1, 2017.
- Potter, G. L., Carriere, L., Hertz, J. D., Bosilovich, M., Duffy, D., Lee, T. and Williams, D. N.:
10 Enabling reanalysis research using the collaborative reanalysis technical environment (CREATE), *Bull. Am. Meteorol. Soc.*, 99(4), 677–687, doi:10.1175/BAMS-D-17-0174.1, 2018.
- Pricope, N. G., Husak, G., Lopez-Carr, D., Funk, C. and Michaelsen, J.: The climate-population nexus in the East African Horn: Emerging degradation trends in rangeland and pastoral livelihood zones, *Glob. Environ. Chang.*, 23(6), 1525–1541, doi:10.1016/j.gloenvcha.2013.10.002, 2013.
- 15 Roca, R., Brogniez, H., Chambon, P., Chomette, O., ClochÃ©, S., Gosset, M. E., Mahfouf, J.-F., Raberanto, P. and Viltard, N.: The Megha-Tropiques mission: a review after three years in orbit, *Front. Earth Sci.*, 3, doi:10.3389/feart.2015.00017, 2015.
- Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Schneider, U., Schröder, M. and Stender, P.: Global gridded precipitation over land: A description of the new GPCC First Guess Daily
20 product, *Earth Syst. Sci. Data*, 6(1), 49–60, doi:10.5194/essd-6-49-2014, 2014.
- Schröder, M., Lockhoff, M., Forsythe, J. M., Cronk, H. Q., Haar, T. H. V. and Bennartz, R.: The GEWEX water vapor assessment: Results from intercomparison, trend, and homogeneity analysis of total column water vapor, *J. Appl. Meteorol. Climatol.*, 55(7), 1633–1649, doi:10.1175/jamc-d-15-0304.1, 2016.



- Schröder, M., Lockhoff, M., Fell, F., Forsythe, J., Trent, T., Bennartz, R., Borbas, E., Bosilovich, M. G., Castelli, E., Hersbach, H., Kachi, M., Kobayashi, S., Robert Kursinski, E., Loyola, Di., Mears, C., Preusker, R., Rossow, W. B. and Saha, S.: The GEWEX Water Vapor Assessment archive of water vapour products from satellite observations and reanalyses, *Earth Syst. Sci. Data*, 10(2), 1093–1117, doi:10.5194/essd-10-1093-2018, 2018.
- 5
- Schröder, M., Lockhoff, M., Shi, L., August, T., Bennartz, R., Brogniez, H., Calbet, X., Fell, F., Forsythe, J., Gambacorta, A., Ho, S., Kursinski, E., Reale, A., Trent, T. and Yang, Q.: The GEWEX Water Vapor Assessment: Overview and Introduction to Results and Recommendations, *Remote Sens.*, 11(3), 251, doi:10.3390/rs11030251, 2019.
- 10
- Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Di Girolamo, L., Getzewich, B., Guignard, A., Heidinger, A., Maddux, B. C., Menzel, W. P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S. and Zhao, G.: Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX radiation panel, *Bull. Am. Meteorol. Soc.*, 94(7), 1031–1049, doi:10.1175/BAMS-D-12-00117.1, 2013.
- 15
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S. and Hsu, K.-L.: A review of global precipitation datasets: data sources, estimation, and intercomparisons, *Rev. Geophys.*, 1–29, doi:10.1002/2017RG000574, 2017.
- Tan, M. L. and Santo, H.: Comparison of GPM IMERG, TMPA 3B42 and PERSIANN-CDR satellite precipitation products over Malaysia, *Atmos. Res.*, 202(November 2017), 63–76, doi:10.1016/j.atmosres.2017.11.006, 2018.
- 20
- Tapiador, F. J., Navarro, A., Levizzani, V., García-Ortega, E., Huffman, G. J., Kidd, C., Kucera, P. A., Kummerow, C. D., Masunaga, H., Petersen, W. A., Roca, R., Sánchez, J. L., Tao, W. K. and Turk, F. J.: Global precipitation measurements for validating climate models, *Atmos. Res.*, 197, 1–20, doi:10.1016/j.atmosres.2017.06.021, 2017.
- 25
- Tapiador, F. J., Roca, R., Genio, A. Del, Dewitte, B., Petersen, W. and Zhang, F.: Is precipitation a



- good metric for model performance?, *Bull. Am. Meteorol. Soc.*, doi:10.1175/BAMS-D-17-0218.1, 2018.
- USHIO, T., SASASHIGE, K., KUBOTA, T., SHIGE, S., OKAMOTO, K., AONASHI, K., INOUE, T., TAKAHASHI, N., IGUCHI, T., KACHI, M., OKI, R., MORIMOTO, T. and KAWASAKI, Z.-I.: A
5 Kalman Filter Approach to the Global Satellite Mapping of Precipitation (GSMaP) from Combined Passive Microwave and Infrared Radiometric Data, *J. Meteorol. Soc. Japan*, 87A, 137–151, doi:10.2151/jmsj.87a.137, 2009.
- Vila, D. A., de Goncalves, L. G. G., Toll, D. L. and Rozante, J. R.: Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America, *J. Hydrometeorol.*, 10(2), 533–543, doi:10.1175/2008JHM1048.1, 2009.
10
- Viltard, N., Burlaud, C. and Kummerow, C. D.: Rain Retrieval from TMI Brightness Temperature Measurements Using a TRMM PR–Based Database, *J. Appl. Meteorol. Climatol.*, 45(3), 455–466, doi:10.1175/jam2346.1, 2006.
- Westra, S., Alexander, L. V. and Zwiers, F. W.: Global increasing trends in annual maximum daily
15 precipitation, *J. Clim.*, 26(11), 3904–3918, doi:10.1175/JCLI-D-12-00502.1, 2013.
- Xie, P., Janowiak, J. E., Arkin, P. A., Adler, R., Gruber, A., Ferraro, R., Huffman, G. J. and Curtis, S.: GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates, *J. Clim.*, 16(13), 2197–2214, doi:10.1175/2769.1, 2003.
- Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F. and Lin, R.: Reprocessed, Bias-Corrected
20 CMORPH Global High-Resolution Precipitation Estimates from 1998, *J. Hydrometeorol.*, 18(6), 1617–1641, doi:10.1175/JHM-D-16-0168.1, 2017.



Zhang, L., A. Kumar, and W. Wang (2012), Influence of changes in observations on precipitation: A case study for the Climate Forecast System Reanalysis (CFSR), *J. Geophys. Res.*, 117, D08105, doi:[10.1029/2011JD017347](https://doi.org/10.1029/2011JD017347).

Ziese, Markus; Rauthe-Schöch, Armin; Becker, Andreas; Finger, Peter; Meyer-Christoffer, Anja;

5 Schneider, Udo (2018): GPCC Full Data Daily Version.2018 at 1.0°: Daily Land-Surface

Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI:

[10.5676/DWD_GPCC/FD_D_V2018_100](https://doi.org/10.5676/DWD_GPCC/FD_D_V2018_100)



Annex

Example of the header of the netCDF-4 file for the 3B42v7 product.

variables:

```
double time(time) ;

5         time:standard_name = "time" ;

         time:long_name = "time" ;

         time:units = "days since 1970-01-01 00:00:00 UTC" ;

         time:calendar = "standard" ;

         time:axis = "T" ;

10        time:comment="The time values of the grid, following the standard CF
calendar, in days, centred, since 01-01-1970 00:00:00"

float lon(lon) ;

         lon:standard_name = "longitude" ;

         lon:long_name = "longitude";

15        lon:units = "degrees_east" ;

         lon:axis = "X" ;

         lon:comment="The longitude values of the grid, in degrees, ranging
between [-179.5; +179.5]. The grid is centred, i.e., a value of +0.5 corresponds to
the degree [0W, 1W]"

20        float lat(lat) ;

         lat:standard_name = "latitude" ;

         lat:long_name = "latitude";

         lat:units = "degrees_north" ;

         lat:axis = "Y" ;
```



```
        lat:comment="The latitude values of the grid, in degrees, ranging
between [-89.5; +89.5]. The grid is centred, i.e., a value of +0.5 corresponds to
the degree [0N, 1N]";

        float rain(time, lat, lon) ;

5         rain:standard_name = "Precipitation » ;

        rain:long_name = "3B42 v7 Daily Accumulated Rainfall (0h-0h)" ;

        rain:units = "mm" ;

        rain:missing_value = NaNf ;

        rain:_FillValue = NaNf ;

10        rain:comment="TBD";

// global attributes:

        :conventions= "CF-1.6, ACDD-1.3" ;

        :title="TBD";

15        :project=" TBD ";

        :summary=" TBD "

        :source="Gridded Daily Accumulated Rainfall derived from TRMM/3B42
v7 products "

        :institution="NASA/LEGOS/IPSL"

20        :license="TBD";

        :acknowledgement="TBD";

        :creator_name = "NASA" ;

        :creator_email = "none" ;

        :creator_url = "https://pmm.nasa.gov/TRMM" ;

25        :creator_type="institution";
```



```

:creator_institution="NASA";

:contributor_name = "Rémy Roca" ;

:contributor_role = "data gridding and formatting in netCDF" ;

:contributor_email = "none" ;

5 :contributor_type="person";

:contributor_institution="LEGOS";

:publisher_name = "ESPRI" ;

:publisher_email = "none" ;

:publisher_type="institution";

10 :publisher_url = "http://geonetwork..." ;

:publisher_institution="IPSL"

:product_version="1";

:time_coverage_start = "1988-01-01" ;

:time_coverage_end = "2013-12-31" ;

15 :time_coverage_resolution = "day" ;

:geospatial_lat_min = -89.5f ;

:geospatial_lat_max = 89.5f ;

:geospatial_lon_min = -179.5f;

:geospatial_lon_max = 179.5f ;

20 :geospatial_lat_resolution = 1f ;

:geospatial_lon_resolution = 1f ;

:geospatial_lat_units = "degrees_north" ;

:geospatial_lon_units = "degrees_east" ;

:cdm_data_type="grid";
```



```
      :keywords="GCMD:EARTH SCIENCE,GCMD:ATMOSPHERE,GCMD:PRECIPITATION" ;  
  
      :platform="GCMD:Earth Observation Satellites";  
  
      :instruments="GCMD:Earth Remote Sensing Instruments,GCMD:Passive  
Remote Sensing";  
5      :instrument_vocabulary = "GCMD:GCMD Keywords" ;  
  
      :platform_vocabulary = "GCMD:GCMD Keywords" ;  
  
      :keywords_vocabulary = "GCMD:GCMD Keywords, CF:NetCDF COARDS Climate  
and Forecast Standard Names" ;
```




Tables

Product shortname and version	Period used	Spatial coverage	References
CPC	1979-2017	60°s-90°n	(Xie et al. 2010)
GPCC Full Daily v2018	1982-2016	60°s-90°n	(Schneider et al., 2018)
GPCC Full Daily v1	1982-2013	60°s-90°n	(Becker et al. 2013)
GPCC Fisrt Guess v1	2009-2016	60°s-90°n	(Becker et al. 2013)
REGEN All V1-2019	1950-2016	60°s-90°n	(Contractor et al., 2019)
REGEN Long V1-2019	1950-2016	60°s-90°n	(Contractor et al., 2019)

Table 1 The ground-based datasets



Product shortname and version	Period used	Spatial coverage	Use of rain gauges data	Use of IR satellite data	Use of MW satellite data rainfall estimate	Main Scientific References and ATBD
Satellite based quasi global						
3B42 v7.0	1998-2016	50°s-50°n	Yes	Yes	multiple platforms	(Huffman et al. 2009)
3B42 v7.0 IR	1998-2016	50°s-50°n	No	Yes	No	(Huffman et al. 2009)
3B42 v7.0 MW	1998-2016	50°s-50°n	No	No	Yes	(Huffman et al. 2009)
3B42 RT v7.0	2000-2017	50°s-50°n	Yes	Yes	multiple platforms	(Huffman et al. 2009)
3B42 RT v7.0 uncalibrated	2000-2017	50°s-50°n	No	Yes	multiple platforms	(Huffman et al. 2009)
GSMAP-RNL-gauges v6.0	2001-2013	50°s-50°n	Yes	yes	multiple platforms	(Kubota et al., 2007)
GSMAP-RNL-no gauges v6.0	2001-2013	50°s-50°n	No	yes	multiple platforms	(Kubota et al., 2007)
GSMAP-NRT-gauges v6.0	2001-2017	50°s-50°n	Yes	yes	multiple platforms	(Kubota et al., 2007)
GSMAP-NRT-no gauges v6.0	2001-2017	50°s-50°n	No	yes	multiple platforms	(Kubota et al., 2007)
PERSIANN CDR v1 r1	1983-2017	50°s-50°n	yes	Yes	No	(Ashouri et al., 2015) (Sorooshian et al., 2014)
CMORPH V1.0, RAW	1998-2017	60°s-60°n	No	Yes	multiple platforms	(Xie et al., 2017)
CMORPH V1.0, CRT	1998-2017	60°s-60°n	Yes	Yes	multiple platforms	(Xie et al., 2017)
GPCP 1DD CDR v1.3	1997-2017	90°s-90°n	Yes	Yes	One platform	(Huffman et al. 2001)
Land Only						
CHIRPS v2.0	1981-2016	50°s-50°n Land only	Yes	Yes	No	(Funk et al. 2015)
CHIRP v2.0	1981-2016	50°s-50°n Land only	Yes	Yes	No	(Funk et al. 2015)
SM2RAIN-CCI	1998-2015	Global Land only	No	No	No	(Ciabatta et al., 2018)
Ocean only						
HOAPS	1996-2014	ocean only	no	no	multiple platforms	(Andersson et al., 2017)
Satellite based regional						
TAPEER v1.5	2012-2016	30°s-30°n	No	Yes	multiple platforms	(Roca et al, 2018)
TAMSAT v2	1983-2017	Africa (land only)	Yes	Yes	No	(Maidment et al., 2017)
TAMSAT v3	1983-2017	Africa (land only)	Yes	Yes	No	(Maidment et al., 2017)
ARC v2	1983-2017	Africa (land only)	Yes	Yes	No	Novella, N.S. and W.M. Thiaw, 2013
COSH	2000-2018	60°s-33°n 120°w-30°w	Yes	Yes	Yes	Vila et al., 2009



Table 2: the satellite-based datasets

Product shortname and version	Period used	References
MERRA 1	1979-2015	Rienecker et al. 2011
MERRA 2	1980-2017	Gelaro et al. 2017
JRA-55	1958-2017	Kobayahi et al (2015)
ERAinteri	1979-2017	Dee et al. (2011)
CFSR	1979-2017	Saha et al. 2010

Table 3: the reanalysis datasets



List of figures

Figure 1: Time series of annual total daily precipitation (in mm) averaged over each dataset

domain (regional or global, land or ocean, or both) as shown on the embedded maps in the panels.

The name of the dataset and number of years available are indicated in each panel. Datasets are

5 clustered with regards to their source type as indicated at the bottom of the figure.

