

Diverse estimates of annual maxima daily precipitation in 22 state-of-the-art quasi-global land observation datasets

Margot Bador, Lisa V. Alexander, Steefan Contractor, Rémy Roca

▶ **To cite this version:**

Margot Bador, Lisa V. Alexander, Steefan Contractor, Rémy Roca. Diverse estimates of annual maxima daily precipitation in 22 state-of-the-art quasi-global land observation datasets. Environmental Research Letters, 2020, 15 (3), pp.035005. 10.1088/1748-9326/ab6a22 . hal-03035441

HAL Id: hal-03035441
https://cnrs.hal.science/hal-03035441

Submitted on 2 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Environmental Research Letters

LETTER • **OPEN ACCESS**

Diverse estimates of annual maxima daily precipitation in 22 state-of-the-art quasi-global land observation datasets

To cite this article: Margot Bador *et al* 2020 *Environ. Res. Lett.* **15** 035005

View the article online for updates and enhancements.

This content was downloaded from IP address 195.220.93.240 on 04/03/2020 at 13:41

Environmental Research Letters

LETTER

Diverse estimates of annual maxima daily precipitation in 22 state-of-the-art quasi-global land observation datasets

Margot Bador[1,2,5], **Lisa V Alexander**[1,2], **Steefan Contractor**[3] and **Remy Roca**[4]

[1] CLEX ARC Centre of Excellence for Climate Extremes, Sydney, Australia
[2] Climate Change Research Centre, School of BEES, UNSW Sydney, Australia
[3] School of Mathematics and Statistics, Faculty of Science, UNSW Sydney, Australia
[4] Laboratoire d'Etudes Géophysiques et d'Océanographie Spatiales, Toulouse, France
[5] Author to whom any correspondence should be addressed.

E-mail: m.bador@unsw.edu.au

OPEN ACCESS

RECEIVED
4 April 2019

REVISED
17 December 2019

ACCEPTED FOR PUBLICATION
10 January 2020

PUBLISHED
21 February 2020

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Keywords: intercomparison of observational products, global distribution of precipitation extremes, annual extremes of daily precipitation, *in situ* and satellite precipitation observations

Supplementary material for this article is available online

Abstract

Observational evidence of precipitation extremes is vital to better understand how these events might change in a future warmer climate. Over the terrestrial regions of a quasi-global domain, we assess the representation of annual maxima of daily precipitation (Rx1day) in 22 observational products gridded at $1^\circ \times 1^\circ$ resolution and clustered into four categories: station-based *in situ*, satellite observations with or without a correction to rain gauges, and reanalyses (5, 8, 4 and 5 datasets, respectively). We also evaluate the interproduct spread across the ensemble and within the four clusters, as a measure of observational uncertainty. We find that reanalyses present a heterogeneous representation of Rx1day in particular over the tropics, and their interproduct spread is the highest compared to any other cluster. Extreme precipitation in satellite data broadly compares well with *in situ*-based data. We find a general better agreement with *in situ*-based observations and less interproduct spread for the satellite products with a correction to rain gauges compared to the uncorrected products. Given the level of uncertainties associated with the estimation of Rx1day in the observations, none of the datasets can be thought of as the best estimate. Our recommendation is to avoid using reanalyses as observational evidence and to consider *in situ* and satellite data (the corrected version preferably) in an ensemble of products for a better estimation of precipitation extremes and their observational uncertainties. Based on this we choose a subsample of 10 datasets to reduce the interproduct spread in both the representation of Rx1day and its timing throughout the year, compared to all 22 datasets. We emphasize that the recommendations and selection of datasets given here may not be relevant for different precipitation indices, and other grid resolutions and time scales.

1. Introduction

Precipitation is heterogeneous in space and time and its measurement is further complicated by the heterogeneity of the ground-based measurement network. While most of the mid-latitudinal terrestrial regions are well monitored, in the tropics the station density is sparse, of lower quality and data availability is limited (Alexander 2016). Overall, the observation of extreme precipitation from ground-based instruments is challenging. Better spatial coverage is achieved by interpolating *in situ*-based data onto a grid, however, gridded datasets represent an area-averaged measure of precipitation and are therefore intrinsically different to a station-based measure (Chen and Knutson 2008, Gervais *et al* 2014a). Uncertainties arise due to the different gridding methods applied (Dunn *et al* 2014, Avila *et al* 2015), but the advantage of gridded datasets is that they allow comparison with other observational products (such as satellite datasets) and are useful for model validation.

Observational datasets created from satellite retrievals generally provide better spatial coverage

© 2020 The Author(s). Published by IOP Publishing Ltd

compared to station-based gridded products, and this allows the assessment of precipitation in data-sparse areas of the globe. However, as the satellite era is relatively new, there is little guidance on the reliability of these observations to the study of long-term changes in precipitation extremes. Reanalysis data provide an alternative to ground-based and satellite-only observations. Using a model constrained by assimilated observations, they allow a complete spatio-temporal characterization of precipitation and other atmospheric fields to investigate the mechanisms at play. However, precipitation is generally not assimilated and is a result of the model physics although guided by the other assimilated variables. There is generally less agreement between different reanalyses datasets than between *in situ* datasets (Bosilovich *et al* 2009, Donat *et al* 2014, Sun *et al* 2018). While we recognise that reanalyses are not strictly observations, they are used widely enough in the literature as a proxy for precipitation observations especially in trend analyses and model evaluation studies (e.g. Kharin *et al* 2013, IPCC 2013, Sillmann *et al* 2013, Donat *et al* 2014, Sun *et al* 2018). Furthermore, studies focusing on mechanisms often opt for reanalyses because they provide a consistent multi-variate framework that is harder to obtain from observations alone. Therefore, reanalyses are incorporated in this study for intercomparison purposes.

In order to estimate how diverse the representation of extreme precipitation is across observational products, an ensemble of datasets can be considered. Most studies intercomparing extreme precipitation in observational products are regional in scale and often in regions of high data-density (e.g. Gervais *et al* 2014b, Yin *et al* 2015, Timmermans *et al* 2019) while less effort has been made to characterize regions of various spatial coverage at the global scale (e.g. Donat *et al* 2014, Herold *et al* 2017). At present, a large number of observational datasets exist for precipitation, from a variety of observational sources. This is mainly explained by an increasing number of satellite and reanalyses datasets as well as new product versions being released regularly. This study aims to comprehensively examine the representation of a measure of precipitation extremes in a large number of observational datasets commonly used in the climate community.

Extreme precipitation is defined here as the annual maxima of daily precipitation, and we focus on quasi-global land areas (50 °S:50 °N, 130 °W:180 °E). The observational products used in this study are available through the Frequent Rainfall Observations on Grids (FROGS) database (Roca *et al* 2019a), which provides a variety of gridded observational precipitation datasets (at 1° × 1° resolution). We consider 22 datasets from different sources: station-based *in situ* data, satellite retrievals with or without a correction to rain gauges, and reanalyses. We define observational uncertainty as the spread across the large ensemble of observational products, and the respective uncertainty for each data source. We further aim to give guidance on the use of these different observational products specifically as it relates to annual maxima of daily precipitation at the (quasi) global scale and over land. The observational datasets used in this study are presented in section 2 with the definition of extreme precipitation. Section 3 describes the results. The findings are discussed in section 4 and conclusions are given in section 5.

2. Data and indices for observed extreme precipitation

2.1. Observational datasets of daily precipitation

We consider 22 datasets with a quasi-global coverage (50 °S:50 °N; 130 °W:180 °E) and we focus on land only. All datasets were gathered and reformatted onto 1° × 1° daily grids for the FROGS database (Roca *et al* 2019a) with a common land-sea mask (from REGEN_ALL_v2019) applied. Roca *et al* (2019a) describe the estimation of daily precipitation accumulation in all these datasets (see table 1). In addition to analysing each dataset separately with every other dataset, we also form product 'clusters': *in situ*-based (5 datasets), satellite with (8 datasets) or without (4 datasets) a correction to rain gauges and reanalyses (5 datasets).

2.2. Precipitation extremes

We define extreme precipitation as the annual maximum 1 day precipitation amount (in mm), or Rx1day (Zhang *et al* 2011). This gives information on the magnitude rather than frequency of extreme precipitation events, but we compare the timing of these annual maxima throughout the year between the datasets in figure 3 (and related text). Annual extreme precipitation is compared to annual total wet-day precipitation (i.e. total from days with precipitation >1 mm), or the prcptot index. Both precipitation indices were characterized and are recommended by the WMO/WCRP/JCOMM Expert Team on Climate Change and Detection Indices (ETCCDI, Zhang *et al* 2011). They are calculated using the ClimPACT software (Alexander and Herold 2015, https://climpact-sci.org/) from daily precipitation fields. This ensures consistency in the calculation of the indices across all datasets, and in particular how missing values are treated.

By definition, Rx1day spans one value per year (wettest day of year), which is a common criticism of this index in that it can miss other 'extremes'. Indices based on the exceedance of a threshold such as the 95th or the 99th percentile (i.e. R95p and R99p) might be preferable but the drawback with these percentile-based indices is the requirement of a base period for the percentile calculation. Base periods are recommended to be at least several decades long (for example, WMO uses a 30 year standard—currently 1961–1990). However, the longest overlapping period for the 22 datasets used here is 13 years (2001–2013),

Table 1. List of the 22 observational datasets of daily precipitation used in this study and grouped in 4 clusters with regards to their type of data. See Roca *et al* (2019a) for a detailed description of each product.

Type of data		Dataset	Temporal coverage	Spatial coverage	Native resolution prior to regridding onto a 1° × 1° grid	References
In situ-based (5)	1	REGEN_ALL_v2019	1950–2016	60 °S-90 °N	1° × 1°	Contractor *et al* (2019)
	2	REGEN_LONG_v2019	1950–2016	60 °S-90 °N	1° × 1°	
	3	GPCC_FDD_v1.0	1988–2013	60 °S-90 °N	1° × 1°	Becker *et al* 2013
	4	GPCC_FDD_2018	1982–2016	60 °S-90 °N	1° × 1°	
	5	CPC_v1.0	1979–2017	90 °S-90 °N	0.5° × 0.5°	Xie *et al* (2010)
Satellites with correction to *in situ* (8)	6	GPCP_CDR_v1.3	1997–2017	90 °S-90 °N	1° × 1°	Huffman *et al* (2001)
	7	CMORPH_v1.0_CRT	1998–2017	60 °S-60 °N	8 km × 8 km	Xie *et al* (2017)
	8	CHIRPS_v2.0	1981–2016	50 °S-50 °N	0.05° × 0.05°	Funk *et al* (2015)
	9	3B42_v7.0	1998–2016	50 °S-50 °N	0.25° × 0.25°	Huffman *et al* (2007)
	10	3B42RT_v7.0	2000–2016	50 °S-50 °N	0.25° × 0.25°	Huffman *et al* (2007)
	11	GSMAP-gauges-RNLv6.0	2001–2013	60 °S-60 °N	0.1° × 0.1°	Kubota *et al* (2007)
	12	GSMAP-gauges-NRT-v6.0	2001–2017	60 °S-60 °N	0.1° × 0.1°	Kubota *et al* (2007)
	13	PERSIANN_v1_r1	1983–2017	60 °S-60 °N	0.25° × 0.25°	Ashouri *et al* (2015), Sorooshian *et al* (2014)
Satellites uncorrected (4)	14	CMORPH_v1.0_RAW	1998–2017	60 °S-60 °N	8 km × 8 km	Xie *et al* (2017)
	15	CHIRP_V2	1981–2016	50 °S-50 °N	0.05° × 0.05°	Funk *et al* (2015)
	16	3B42_IR_v7.0	1998–2016	50 °S-50 °N	0.25° × 0.25°	Huffman *et al* (2007)
	17	3B42RT_UNCAL_v7.0	2000–2016	50 °S-50 °N	0.25° × 0.25°	Huffman *et al* (2007)
Reanalyses (5)	18	CFSR	1979–2017	90 °S-90 °N	T382 (~38 km)	Saha et al (2010)
	19	ERAi	1979–2017	90 °S-90 °N	T255 (~79 km)	Dee *et al* (2011)
	20	JRA-55	1958–2017	90 °S-90 °N	TL319 (~55 km)	Kobayashi *et al* (2015)
	21	MERRA1	1979–2015	90 °S-90 °N	0.66° × 0.5°	Rienecker *et al* (2011)
	22	MERRA2	1980–2017	90 °S-90 °N	0.625° × 0.5°	Gelaro *et al* (2017)

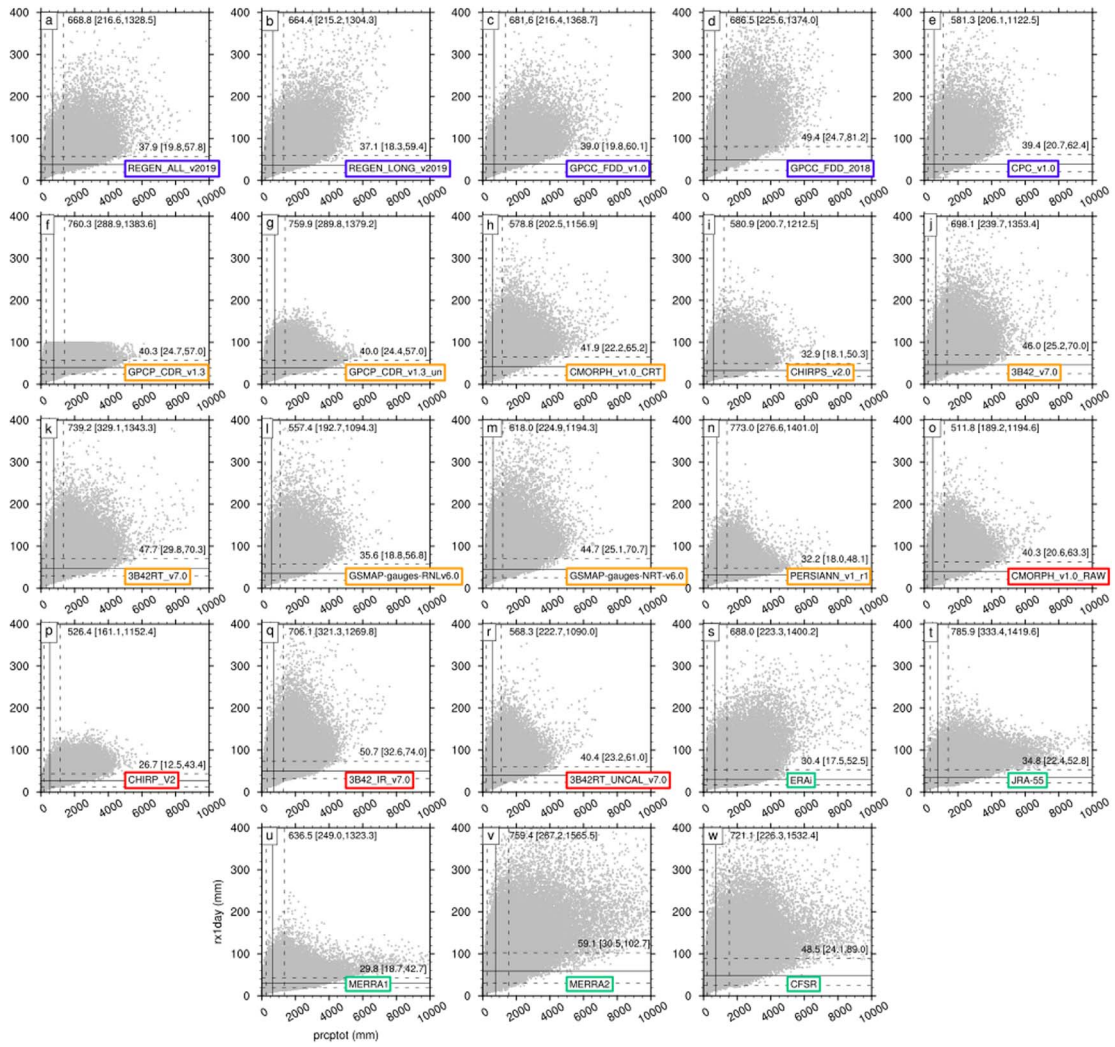


Figure 1. Scatterplot of Rx1day values versus prcptot values (both in mm) in an ensemble of 22 observational datasets (see on panels and table 1). Values from all years between 2001 and 2013 and all grid cells of the quasi-global land domain are considered. The datasets are clustered into 4 groups: *in situ*-based data (blue label), satellite data with or without a correction to rain gauges (orange or red label), and reanalyses (green label). Horizontal and vertical solid (dashed) lines indicate the median value (the 25th and 75th percentiles) of the distribution of Rx1day and prcptot, respectively, and their values are written on each panel. Note that a common range is applied to *X*- and *Y*-axes for an easier intercomparison and that it can mask highest data of some datasets.

which is therefore probably too short to use percentile-based indices. As an illustration, we compare the mean global R99p values in two datasets (REGEN_ALL_v2019 and JRA-55) but using three base periods of different lengths (15, 30 and 50 years) and we show large differences in particular after 1990 (up to above 12 and 17 mm respectively; see supplementary figure 1 is available online at stacks.iop.org/ERL/15/035005/mmedia). This shows that the sensitivity to the choice of the base period is amplified in the presence of a trend, i.e. when the base period is calculated prior to the start of a trend, the trend estimate will be higher compared to if the base period spans years when there is a trend.

3. Results

The global distribution of Rx1day values is first compared to the global distribution of annual total wet-day precipitation (prcptot) for all years during 2001–2013 (figure 1). We start by making a point on the use of GPCP_CDR_v1.3 dataset. This satellite product is provided with a 'valid range attribute' included in the file, and we compare here Rx1day and prcptot distributions with (panel f) and without (panel g) this valid range applied to the data. This shows that the use of valid range masks all values above 100 mm. It is worth noting that this is not perceivable from the distribution of prcptot values, as they consist in the sum of daily precipitation amounts in a year. It is also not perceivable if the climatological Rx1day values are considered instead of the values of all years over the 2001–2013 period (supplementary figure 2 versus figure 1). In addition, comparing 50% of the global distribution of Rx1day (through the 25th and 75th percentiles; horizontal dashed lines on figure 1) across the datasets shows that GPCP_CDR_v1.3 has among the narrowest distribution (MERRA1 has the

narrowest). However, all other datasets have values above 100 mm and it thus seems reasonable to conclude that GPCP_CDR_v1.3 data should be used without applying the valid range as it hinders the study of extreme precipitation by excluding the most extreme values over the globe. In the rest of this study, we therefore only consider the raw data of GPCP_CDR_v1.3.

The scatterplots form different 'data clouds' for each datasets. Some indicate that the largest values of Rx1day are found for the largest values of prcptot (e.g. most of the *in situ*-based datasets), while others indicate the largest values of Rx1day occur around the 75th percentile of the prcptot distribution (e.g. n, o, q, r, t, u; rightmost vertical dashed line). The median of the Rx1day distributions (horizontal solid lines in figure 1) indicates that the driest datasets are (in decreasing order) CHIRP_V2, MERRA1, ERAi, PERSIANN_v1_r1, CHIRPS_v2.0 and the wettest are (in decreasing order) MERRA2, 3B42_IR_v7.0, GPCC_FDD_v2018, CFSR, 3B42RT_v7.0. The median values range from 26.7 to 59.1 mm, showing large interproduct spread, which is even larger for the 25th (from 12.5 to 32.6 mm) and 75th (from 42.7 to 102.7 mm) percentile values. On the contrary, prcptot values compare much better across the 22 datasets. We find a better comparison between the datasets for the median and 25th and 75th percentiles of the global distributions of prcptot (vertical lines in figure 1), in agreement with the findings of Alexander *et al* (2020) and Roca (2019b).

We cannot conclude which of these 'data clouds' is the most realistic but we observe less difference between *in situ*-based datasets than within any other clusters, and in particular between the reanalyses which exhibit the largest differences. *In situ*-based observations (blue labels in figure 1) show a similar representation of Rx1day and prcptot global distribution with the exception of GPCC_FDD_2018 that shows higher values for Rx1day, and among the highest values of the ensemble. Compared to the *in situ* -based cluster, we find more spread (e.g. among the median values) within the satellites clusters, either with or without correction to rain gauges (orange and red labels in figure 1, respectively).

Cluster-averaged 2001–2013 climatological values of Rx1day show differences across the four clusters, or the four types of observations (figures 2(a), (d), (g), (j)). Extreme precipitation intensity over the driest regions varies across the four clusters (e.g. central Asia, Arabian Peninsula, western US), as previously highlighted by Donat *et al* (2014). The intercluster differences are also large over the wettest regions and in the tropics in general (figures 2(a), (d), (g), (j) and (n), (o)). Extreme precipitation intensity on the leeward side of extratropical northern and southern America is contrasted between the clusters, and in particular between the corrected and uncorrected satellite products, with a tendency towards lower estimates in the corrected datasets (i.e. closer to *in situ*-based data). Note that for an easier intercomparison, values are plotted up to 160 mm but cluster-averaged *in situ*-based data indicate higher values (up to 260 mm) compared to reanalyses (up to 180 mm), satellite with correction (up to 160 mm) and satellite without correction (up to 150 mm). This cannot be explained by scaling issues as all data were first interpolated onto a common grid but could be explained by structural differences in the measurement of precipitation and/or by interproduct spread within each cluster, which is further investigated as follows.

We then examine how much interproduct spread is associated to each source of observation using the multiproduct standard deviation and coefficient of variation (i.e. standard deviation normalized by the multiproduct mean of climatological Rx1day; figures 2(b), (e), (h), (k) and (c), (f), (i), (l)). Reanalyses show the largest spread while *in situ*-based data show the smallest (see also supplementary figures 3 and 4), in agreement with Donat *et al* (2014). Satellite data lie in between with generally more spread within the uncorrected than the corrected product clusters (figures 2(e), (f), (h), (i)). Within the reanalyses cluster, the largest interproduct differences are located in the tropics with multiproduct standard deviation values generally over 50 mm (coefficient of variation values above 50%; figures 2(k), (l)) or more for regions like central Africa for instance. In the extra-tropics, uncertainties in reanalyses are lower, yet generally higher than for any other observational source. On the contrary, *in situ*-based data show little interproduct spread with the highest uncertainties in the tropics and in particular in the Sahara desert, where there are few rain gauges (figures 2(b), (c)). The spread within the *in situ*-based cluster is mainly explained by large differences between GPCC_FDD_v2018 and the other datasets (see also supplementary figures 3 and 4). Satellite data show interproduct spread levels closer to *in situ*-based than reanalyses except over a dry region extending from northern Africa to Central Asia where interproduct differences are large (figures 2(f), (i)). Over such semi-arid regions, satellite products tend to show large uncertainty due to low detection skill (Maggioni *et al* 2016). These results agree with those of Herold *et al* (2017), and they are generally in line with those for annual total wet-day precipitation (prcptot; supplementary figures 5–7) yet of weaker intensity compared to annual extreme precipitation.

Finally, the cluster-averaged representation of climatological Rx1day in the *in situ*-based and corrected satellite clusters are the most similar and show reduced interproduct spread compared to the uncorrected satellite data and reanalyses clusters. We further show that selecting only these two groups of data instead of considering all four reduces the uncertainties (bottom panels of figure 2). Some areas still present a relatively high spread, but these can be either dry or data-spare regions, or both. Furthermore, figure 1 highlighted PERSIANN_v1_r1 as one of the driest datasets (see above) and indeed it presents a pronounced

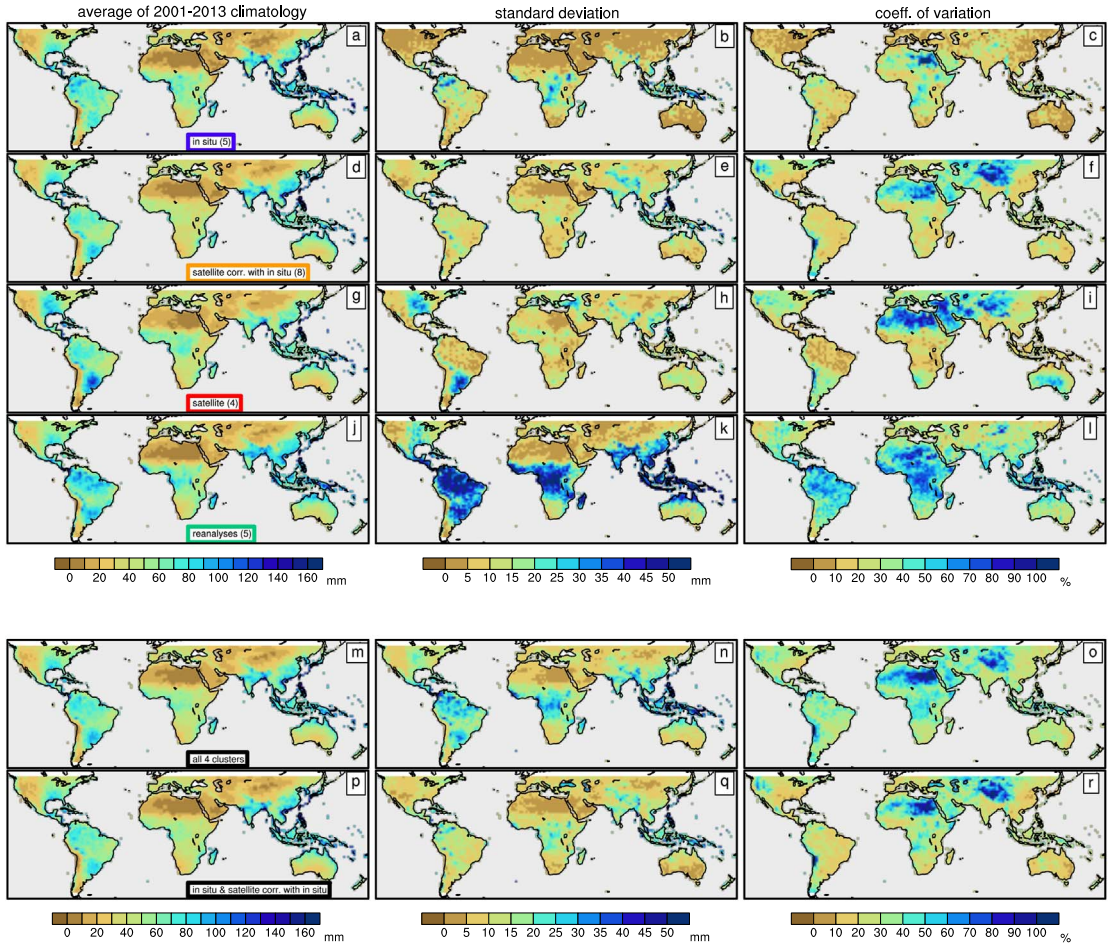


Figure 2. Top 12 panels: average over the datasets of each cluster of 2001–2013 climatological Rx1day (mm; (a), (d), (g), (j)) and its corresponding standard deviation (mm; (b), (e), (h), (k)) and coefficient of variation (standard deviation normalized by the climatology, in %; (c), (f), (I), (l)). The datasets are clustered into 4 groups: *in situ*-based data (blue label; first line), satellite with (orange label; second line) or without (red label; third line) correction to rain gauges, and reanalyses (green label; bottom line). The number of datasets varies between clusters (as indicated on panels (a), (d), (g), (j) and in table 1). Bottom 6 panels: average over the datasets from all clusters (m) and two clusters (*in situ*-based and satellite with correction; (p) of 2001–2013 climatological Rx1day (mm) and its corresponding standard deviation (mm; (n), (q)) and coefficient of variation (%; (o), (r)). Note that by definition, spread measured by standard deviation gives higher emphasis for the wettest regions and spread measured by coefficient of variation gives higher emphasis for the driest regions. A common range is applied and this can mask values higher than 160 mm.

widespread dryness compared to other products (supplementary figures 3 and 4(m)), as well as large regions of missing values (especially at the beginning of the 2001–2013 period). Therefore, in the context of this study, we do not recommend using PERSIANN_v1_r1. Finally, these results point to a subset of 5 *in situ*-based and 7 satellite datasets that present reduced interproduct differences in the estimation of annual daily precipitation maxima over global land compared to the initial 22 products.

We further investigate how the timing of extremes in a year compares across the selected 12 datasets (figure 3). We first intercompare their climatological annual cycle and find that interproduct spread is generally higher in the tropical band (up to 1.1 mm d^{-1}) than in the extra-tropics (up to 0.8 and 0.9 mm d^{-1} in the Southern and Northern hemisphere, respectively; first column of figure 3; note that a 21 d running averaged is applied). The satellite product 3B42RT_v7.0 indicates the highest values of mean daily precipitation during the wet season in the Northern Hemisphere extra-tropics, leading to larger interproduct spread. Except for this, results show a relatively similar annual cycle between the 12 selected datasets, in particular in the extra-tropics, with little differences between the *in situ*-based and satellite clusters. This is however very different when all 22 datasets are considered (supplementary figure 8). The interproduct spread is much higher (up to 1.0 mm d^{-1} in both extra-tropical bands and 2.8 mm d^{-1} in the tropics) and the reanalyses present higher values of mean daily precipitation, in particular in the tropics.

Next, we compare the timing of extreme precipitation through the distribution of the month of occurrence of Rx1day (i.e. distribution of the months when annual maxima of daily precipitation are recorded between 2001 and 2013 for each grid cell; right column of figure 3). These distributions are estimated in latitudinal bands and we find that distributions are relatively comparable across the selected datasets. In the

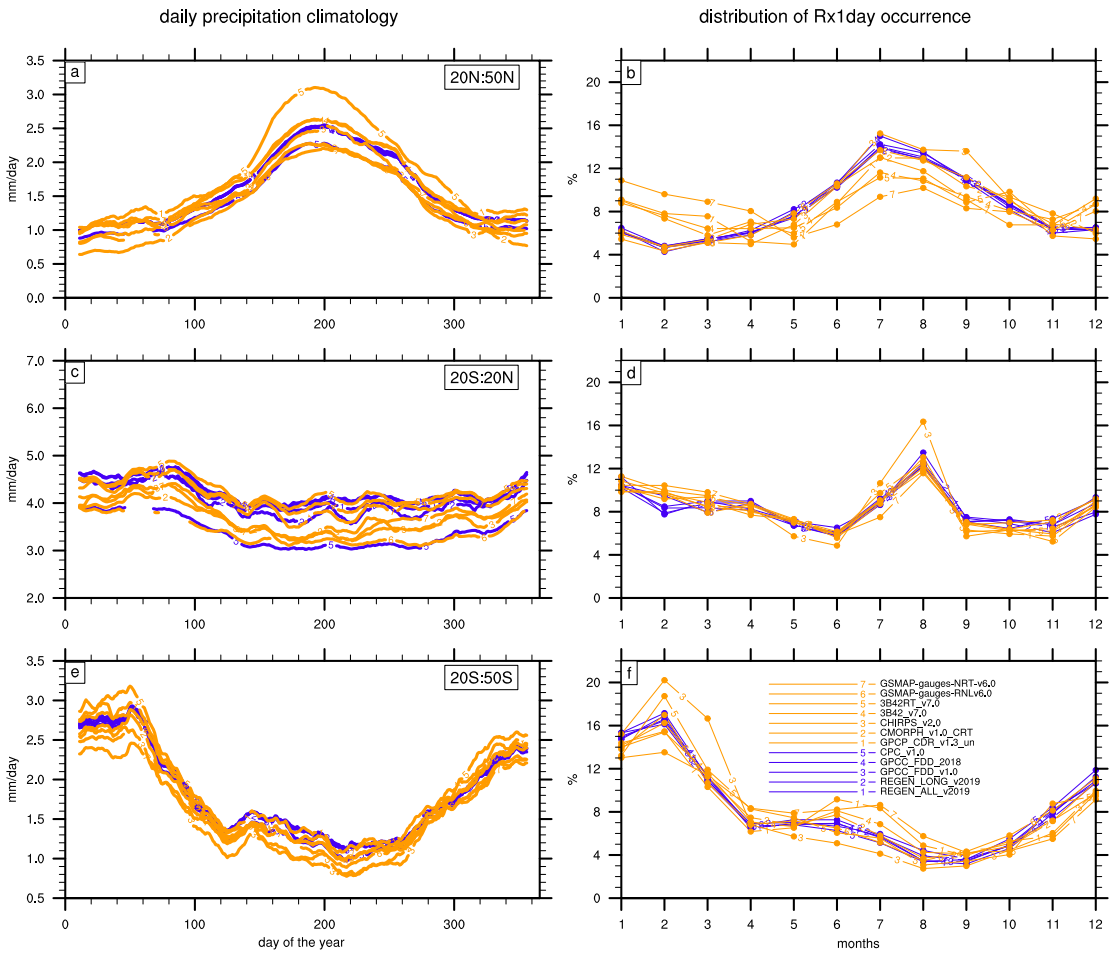


Figure 3. Climatological (2001–2013) annual cycle of daily precipitation (in mm d^{-1}; left panels), and monthly distribution of Rx1day occurrence (i.e. distribution of the months when annual maxima of daily precipitation are recorded between 2001 and 2013 for each grid cell, in %; right panels). Land values in three latitudinal bands are considered: (a)–(b) 20 °N:50 °N, (c)–(d) 20 °S:20 °N and (e)–(f) 20 °S:50 °S. A subset of 12 selected datasets are considered here based on results from figures 1 and 2: 5 *in situ*-based (blue label) and 7 satellite products with correction to *in situ* (orange label; see legend on panel (f)).

extra-tropics, selected datasets agree on a higher occurrence of extremes during the wet season (figures 3(b), (f)), and in the tropics they indicate a flatter distribution peaking in August (the Indian monsoon imprint; figure 3(d)). Interestingly, uncertainties are higher in both extra-tropical bands compared to the tropical band, where little interproduct spread exists and only CHIRPS_v2.0 presents a slightly higher peak in August. Previously, we highlighted higher interproduct spread in extreme precipitation intensity in the tropical band. We show here that it is certainly not explained by different timings of extremes across the datasets, i.e. that we are comparing annual maxima of precipitation that generally occur in similar months but have intensity values that differ between products. Over extra-tropical land in the Southern Hemisphere, the occurrence distributions are relatively similar across the datasets however with higher interproduct spread for the satellite than the *in situ*-based cluster, and again CHIRPS_v2.0 showing slightly higher occurrences during the wet season. In the extra-tropical Northern Hemisphere, the satellite products generally indicate a flatter distribution compared to *in situ*-based datasets, and their cluster presents higher interproduct spread than the *in situ*-based cluster.

In extra-tropical Northern Hemisphere (and the Southern Hemisphere to a lesser extent), GSMAP-gauges-NRT-v6.0 shows the least variation over the months in comparison to the other datasets and to GSMAP-gauges-RNLv6.0 whose distribution peaks in summer and early autumn (months 6–10), in agreement with *in situ*-based observations (figure 3(b)). Therefore, for purpose of our study investigating Rx1day at the global scale, we suggest using the RNL preferably over the NRT version for GSMAP-gauges products in order to ensure that the extremes occur in a reasonably similar period throughout the year (compared to other datasets). In the Northern Hemisphere (and the Southern hemisphere to a lesser extent), the timing of extremes is relatively similar between both 3B42 versions, whereas 3B42RT_v7.0 shows higher mean precipitation values than 3B42_v7.0 throughout the year and in particular during the wet season (figures 3(a), (b)). While it is difficult to conclude whether or not the 3B42RT_v7.0 version is overestimating daily mean precipitation, we can suggest the use of the

3B42_v7.0 preferably over the 3B42RT_v7.0 version for the 3B42 products as the former benefits from much more rain gauges information than the latter. Finally, we cannot conclude that there is a best version among the REGEN and GPCC *in situ*-based products but we have previously highlighted more intense precipitation extremes in GPCC_FDD_v2018 compared to the other *in situ*-based products (and to most of the datasets used here). Based on this intercomparison and within the context of this study, we suggest a final selection 10 observational datasets that enable reduced interproduct spread in the estimation of Rx1day over quasi-global land.

4. Discussion

We find a better general agreement within the *in situ* -based cluster compared to any other cluster. This is partly explained by some interproduct dependencies. Out of 5 *in situ*-based datasets, there are two pairs from a common center and indeed the two REGEN datasets are largely similar (see also supplementary figure 4). They are also generally similar to GPCC_FDD_v1.0 as they share the most of their rain gauges but present lower estimates of Rx1day compared to GPCC_FDD_v2018. Differences are likely to be due to additional station data and improved quality control in the later version of the GPCC dataset (pers comms U. Schneider) but also might be related to changes in the gridding method employed. The fifth *in situ*-based dataset (CPC_v1.0) shares less with the four others, but complete independence can never really be achieved with station-based gridded products that need to span as many gauges as possible. Independence for observations in general is very difficult to obtain as most of the satellite products correct their estimates to rain gauges and share the data from the same instruments (radiosondes, satellite observations, etc).

An important limitation to our knowledge and our ability to estimate observed extreme precipitation over the globe is the lack of rain gauges in many regions such as Africa, South-America and South-East Asia. This is an evident limitation for products gridded from *in situ* data but more generally for any type of observational dataset as most of them rely on station-based estimates or use them for validation. The quantity of stations is insufficient and heterogeneously distributed over the globe but coverage is also limited in time (Kidd *et al* 2017). The two versions of the REGEN dataset (Contractor *et al* 2019) allow the evaluation of the impacts of temporal limitations. Indeed, REGEN_LONG_v2019 only considers stations with at least 40 years of data to generate the final gridded product while REGEN_ALL_v2019 considers all available stations. We find higher amounts of annual precipitation extremes in REGEN_LONG_v2019 compared to REGEN_ALL_v2019 (figures 1(a), (b)), with the largest differences over north-western South-America and south-east Asia (supplementary figures 5(a), (b)), but no such impacts for total annual precipitation (supplementary figures 6 et 7(a), (b)). Furthermore, the longest period of overlap among all datasets used in this study is 13 years, which hinders the intercomparison of trends in observed extremes.

Based on this quasi-global intercomparison of observations for Rx1day, we suggest (upon availability) using the version of a satellite product corrected to rain gauges rather than the uncorrected version. However, such suggestion might not hold for all land regions of the globe or other grid resolutions. Over south-east Asia and tropical Africa and South-America for instance, we do not have sufficient stations and the available data quality is doubtful. Hence, it remains to be verified if, over these regions of poor coverage, it is best to consider the corrected or uncorrected version of a satellite product. Interestingly, comparing such pairs of products over south-east Asia for instance shows significant differences between the two versions (e.g. CMORPH_v1.0_CRT versus CMORPH_v1.0_RAW, CHIRPS_v2.0 versus CHIRP_V2 and 3B42RT_v7.0 versus 3B42RT_UNCAL_v7.0; supplementary figure 5(g) versus (o), (h) versus (p), (j) versus (r)), while only a few rain gauges exist over this area. Similarly, we also find important differences for annual total precipitation between the corrected and uncorrected versions of a satellite product over tropical regions poorly sampled on the ground (supplementary figures 7(j), (r)). The accuracy of such correction to rain gauges and its value should then be further investigated region by region and individually for each dataset.

5. Conclusion

This study focuses on the estimation of the annual maximum 1 day precipitation amount (Rx1day index) in a variety of observations and the assessment of the observational uncertainty (defined as interproduct spread) over land at the (quasi) global scale. We have conducted an intercomparison of 22 gridded products (at $1^\circ \times 1^\circ$ daily resolution) that we have clustered into four groups: *in situ*-based (5), satellite with (8) or without (4) a correction to rain gauges and reanalyses (5). We have compared the climatology of annual maxima 1 day precipitation over the 2001–2013 period (overlapping all datasets) and have evaluated the interproduct spread across the ensemble and within each cluster. Compared to annual total daily precipitation, annual extreme precipitation shows higher interproduct spread. This is not sensitive to the use of Rx1day as Alexander *et al* (2020) find that other extreme precipitation indices (e.g. R99p) show similar levels of interproduct spread (see also Herold *et al* 2017 and Masunaga *et al* 2019).

Reanalyses present a heterogeneous representation of extreme precipitation and in particular over the

tropics with either a widespread wet (MERRA2 and CFSR) or dry (MERRA1 and JRA-55) state -or a spatially contrasted state for ERAi- compared to *in situ* -based data. The interproduct spread is the highest within the reanalyses cluster compared to any other cluster. Our main recommendation is therefore to avoid using reanalyses for observational evidence when investigating extreme precipitation at the global scale. Furthermore, we recommend the use of GPCP_CDR_v1.3 without applying the valid range provided in the file as it disregards all values above 100 mm. Finally, we do not recommend PERSIANN_v1_r1 because of a widespread dryness compared to *in situ*-based data in addition to large areas of missing values for some years. We emphasize that these recommendations are relevant to the context of this study, i.e. the estimation of annual maxima of daily precipitation over quasi-global land.

Extreme precipitation intensity in satellite data broadly compares well with *in situ*-based data. At the quasi-global scale and at a $1^\circ \times 1^\circ$ grid resolution, our results thus indicate that satellite data can be used with *in situ*-based observations when assessing Rx1day. Some satellite data provide different versions, with and without a correction to rainfall gauges. Our work has shown a general better agreement with *in situ*-based observations and less interproduct spread for the corrected datasets, which is therefore preferred to the uncorrected version for broad-scale studies assessing annual maxima of daily precipitation.

Based on the level of observational uncertainty associated with Rx1day over global land, we cannot conclude in any product emerging as the best observational evidence (in agreement with Herold *et al* 2017). We strongly encourage using an ensemble of observations from different sources and centers to estimate precipitation extremes and better assess their associated uncertainties. The interproduct spread in the observations is probably underestimated in many studies focusing on observations or on model evaluation. Herold *et al* (2016) show that this spread is similar to the uncertainties (intermodel and internal variability) of the Coupled Model Intercomparison Project Phase 5 (CMIP5, Taylor *et al* 2012) with regards to a precipitation index representing the mean daily precipitation amount when it rains (i.e. SDII index). Furthermore, Herold *et al* (2017) show that Rx1day is sensitive to different resolution. The interproduct spread in the observations is significantly higher at a resolution of $1^\circ \times 1^\circ$ than $2^\circ \times 2^\circ$, which are, respectively, the resolutions that the next generation of global climate models (i.e. CMIP6) is likely to have and the resolution of the last generation of models (i.e. CMIP5). We therefore encourage model evaluation studies to consider product sensitivity as higher model resolution will certainly continue to be sought after.

For studying annual maximum 1 day precipitation over land at the (quasi) global scale, we suggest using an ensemble of observations from the FROGS database. Indeed, based on our results we recommend a selection of *in situ*-based and satellite products (specifically REGEN_ALL_v2019, REGEN_LONG_v2019, GPCC_FDD_v1.0, GPCC_FDD_v2018, CPC_v1.0 (*in situ*-based) and GPCP_CDR_v1.3 –no valid range applied-, CMORPH_v1.0_CRT, CHIRPS_v2.0, 3B42_v7.0, GSMAP-gauges_RNL_v6.0 (corrected satellite)). We find greater similarity in extreme precipitation intensity and the timing of these extremes (i.e. distribution of month of occurrence of the wettest day in a year) within these selected datasets compared to all 22 products, giving some confidence to the use of this selection of datasets. It is important to acknowledge that this selection is relevant for the purposes of estimating annual precipitation maxima over quasi-global land but might not be relevant to other precipitation indices, grid resolutions or time scales. This reinforces the need for more wide ranging extreme precipitation intercomparisons.

Acknowledgments

The data (FROGS database, Roca *et al* (2019a)) that support the findings of this study are openly available at https://doi.org/10.14768/06337394-73A9-407C-9997-0E380DAC5598. All analyses and graphics have been done using the NCAR Command Language (NCL 2013). MB and SC were supported by Australian Research Council (ARC) Discovery Project DP160103439 and LVA was supported by the ARC Centre of Excellence for Climate Extremes (CE170100022). RR acknowledges the funding of CNES through the Megha-Tropiques program.

ORCID iDs

Margot Bador https://orcid.org/0000-0003-3976-6946
Lisa V Alexander https://orcid.org/0000-0002-5635-2457
Steefan Contractor https://orcid.org/0000-0002-3987-2311
Remy Roca https://orcid.org/0000-0003-1843-0204

References

Alexander L V, Bador M, Roca R, Contractor S and Donat M 2020 Intercomparison of precipitation extremes over global land areas from in situ, space-based and reanalysis products. *Environ. Res. Lett.* (in review)

Alexander L V 2016 Global observed long-term changes in temperature and precipitation extremes: a review of progress and limitations in IPCC assessments and beyond *Weather Clim. Extremes* **11** 4–16

Alexander L and Herold N 2015 ClimPACTv2 Indices and Software. A document prepared on behalf of the Commission for Climatology (CCl) Expert Team on Sector-Specific Climate Indices (ET-SCI) (https://github.com/ARCCSS-extremes/climpact2/blob/master/ClimPACTv2_manual.pdf)

Ashouri H, Hsu K L, Sorooshian S, Braithwaite D K, Knapp K R, Cecil L D, Nelson B R and Prat O P 2015 PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies *Bull. Am. Meteorol. Soc.* **96** 69–83

Avila F B, Dong S, Menang K P, Rajczak J, Renom M, Donat M G and Alexander L V 2015 Systematic investigation of gridding-related scaling effects on annual statistics of daily temperature and precipitation maxima: a case study for south-east Australia *Weather Clim. Extremes* **9** 6–16

Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U and Ziese M 2013 A description of the global land-surface precipitation data products of the global precipitation climatology centre with sample applications including centennial (trend) analysis from 1901–present *Earth Syst. Sci. Data* **5** 71–99

Bosilovich M G, Mocko D, Roads J O and Ruane A 2009 A multimodel analysis for the coordinated enhanced observing period (CEOP) *J. Hydrometeor.* **10** 912–34

Chen C-T and Knutson T 2008 On the verification and comparison of extreme rainfall indices from climate models *J. Clim.* **21** 1605–21

Contractor S, Donat M G, Alexander L V, Ziese M, Meyer-Christoffer A, Schneider U, Rustemeier E, Becker A, Durre I and Vose R S 2019 Rainfall estimates on a gridded network (REGEN)–A global land-based gridded dataset of daily precipitation from 1950–2013 *Hydrol. Earth Syst. Sci. Discuss.* 1–30 (in review)

Dee D P *et al* 2011 The ERA-interim reanalysis: configuration and performance of the data assimilation system *Q. J. R. Meteor. Soc.* **137** 553–97

Donat M G, Sillmann J, Wild S, Alexander L V, Lippmann T and Zwiers F W 2014 Consistency of temperature and precipitation extremes across various global gridded *in situ* and reanalysis datasets *J. Clim.* **27** 5019–35

Dunn R J H, Donat M G and Alexander L V 2014 Investigating uncertainties in global gridded datasets of climate extremes *Clim. Past* **10** 2171–99

Funk C *et al* 2015 The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes *Sci. Data* **2** 1–21

Gelaro R *et al* 2017 The modern-era retrospective analysis for research and applications, Version 2 (MERRA-2) *J. Clim.* **30** 5419–54

Gervais M, Tremblay L B, Gyakum J R and Atallah E 2014a Representing extremes in a daily gridded precipitation analysis over the United States: impacts of station density, resolution, and gridding methods *J. Clim.* **27** 5201–18

Gervais M, Gyakum J R, Atallah E, Tremblay L B and Neale R B 2014b How well are the distribution and extreme values of daily precipitation over North America represented in the community climate system model? A comparison to reanalysis, satellite, and gridded station data *J. Clim.* **27** 5219–29

Herold N, Alexander L V, Donat M G, Contractor S and Becker A 2016 How much does it rain over land? *Geophys. Res. Lett.* **43** 341–8

Herold N, Behrangi A and Alexander L V 2017 Large uncertainties in observed daily precipitation extremes over land *J. Geophys. Res.* **122** 668–81

Huffman G J, Adler R F, Morrissey M M, Bolvin D T, Curtis S, Joyce R, McGavock B and Susskind J 2001 Global precipitation at one-degree daily resolution from multisatellite observations *J. Hydrometeorol.* **2** 36–50

Huffman G J, Bolvin D T, Nelkin E J, Wolff D B, Adler R F, Gu G, Hong Y, Bowman K P and Stocker E F 2007 The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales *J. Hydrometeorol.* **8** 38–55

IPCC 2013 *Summary for Policymakers Climate Change 2013: the Physical Science Basis Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed T F Stocker et al (Cambridge: Cambridge University Press)

Kharin V V, Zwiers F W, Zhang X and Wehner M 2013 Changes in temperature and precipitation extremes in the CMIP5 ensemble *Climatic Change* **119** 345–57

Kidd C, Becker A, Huffman G J, Muller C L, Joe P, Skofronick-Jackson G and Kirschbaum D B 2017 So, how much of the Earth's surface is covered by rain gauges? *Bull. Am. Meteorol. Soc.* **98** 69–78

Kobayashi S *et al* 2015 The JRA-55 reanalysis: general specifications and basic characteristics *J. Meteorol. Soc. Japan.* II **93** 5–48

Kubota T *et al* 2007 Global pecipitation map using satellite-borne microwave radiometers by the GSMaP project: production and validation *IEEE Trans. Geosci. Remote* **45** 2259–75

Maggioni V, Meyers P C and Robinson M D 2016 A review of merged high-resolution satellite precipitation product accuracy during the tropical rainfall measuring mission (TRMM) Era *J. Hydrometeorology* **17** 1101–17

Masunaga H, Schröder M, Furuzawa F A, Kummerow C, Rustemeier E and Schneider U 2019 Inter-product biases in global precipitation extremes *Environ. Res. Lett.* **14** 125016

Rienecker M M *et al* 2011 MERRA: NASA's modern-era retrospective analysis for research and applications *J. Clim.* **24** 3624–48

Roca R, Alexander L V, Potter G, Bador M, Jucá R, Contractor S and Cloché S 2019a FROGS: a daily $1^{\circ} \times 1^{\circ}$ gridded precipitation database of rain gauge, satellite and reanalysis products *Earth Syst. Sci. Data* **11** 1017–35

Roca R 2019b Estimation of extreme daily precipitation thermodynamic scaling using gridded satellite precipitation products over tropical land *Environ. Res. Lett.* **14** 095009

Saha S *et al* 2010 The NCEP climate forecast system reanalysis *Bull. Am. Meteorol. Soc.* **91** 1015–58

Sillmann J, Kharin V V, Zhang X, Zwiers F W and Bronaugh D 2013 Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate *Journal of Geophysical Research: Atmospheres* **118** 1716–33

Sorooshian S, Hsu K, Braithwaite D, Ashouri H and NOAA CDR Program 2014 *NOAA Climate Data Record (CDR) of Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN-CDR)* Version 1 Revision 1 (NOAA's National Centers for Environmental Information) (https://doi.org/10.7289/V51V5BWQ)

Sun Q, Miao C, Duan Q, Ashouri H, Sorooshian S and Hsu K L 2018 A review of global precipitation data sets: data sources, estimation, and intercomparisons *Rev. Geophys.* **56** 79–107

Taylor K E, Stouffer R J and Meehl G A 2012 An overview of CMIP5 and the experiment design *Bull. Am. Meteorol. Soc.* **93** 485–98 article

Timmermans B, Wehner M, Cooley D and Brien T O 2019 An evaluation of the consistency of extremes in gridded precipitation data sets *Clim. Dyn.* **52** 6651–70

Xie P-P, Chen M and Shi W 2010 CPC unified gauge-based analysis of global daily precipitation *24th Conf. on Hydrology (17–21 January 2010(Atlanta, GA: Amer. Meteor. Soc.) 2.3A*

Xie P, Joyce R, Wu S, Yoo S-H, Yarosh Y, Sun F and Lin R : 2017 Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998 *J. Hydrometeorol.* **18** 1617–41

Yin H, Donat M G, Alexander L V and Sun Y 2015 Multi-dataset comparison of gridded observed temperature and precipitation extremes over China *Int. J. Clim.* **35** 2809–27

Zhang X, Alexander L, Hegerl G C, Jones P, Tank A K, Peterson T C and Zwiers F W 2011 Indices for monitoring changes in extremes based on daily temperature and precipitation data *Wiley Interdiscip. Rev. Clim. Change* **2** 851–70