

Response to anonymous referee #2

This paper presents a new global gridded rainfall dataset that combines existing gauge datasets, quality controls the data and produces a useful new gridded daily rainfall product. Rainfall data at a daily timestep are difficult to access and in parts of the world, not even collected. The work presented here represents a substantial effort and contribution to both the literature on rainfall data and in providing a new resource for hydrologists amongst others. Well done!

The paper is well written and the methods and analysis are sound. I think the paper only requires some minor revisions to be acceptable for publishing.

We thank the reviewer for the encouraging review. The suggestions and points raised help clarify any misleading information and provide additional detailed documentation of pertinent information related to REGEN.

We have responded (in red) to each reviewer comment below. Page and line numbers refer to the original unmodified text.

P2 L9: ALL measurements of PPT have errors including gauges. Please clarify this and add in a short discussion of the errors in gauge data. (see McMillan et al 2012 as a starting point)

The sentence on P2 L9 will be replaced with the following paragraph.

“All observations have errors, for example, gauge-based precipitation measurements are subject to undercatch, wind related errors, evaporation loss, wetting loss, splash in/out errors and tipping errors (see McMillan et al 2012 for details). However, alternatives to gauge-based measurements such as satellite observations, model reanalysis products and radar-based observations have additional limitations.”

P2 L10: Who thinks that reanalysis products represent the 'true state of the system'? Either cite or say that reanalysis data are sometimes misused as observations

The sentence will be modified as follows:

“Reanalyses are often misused as observations but in fact inherit...”

P2 L13 NASA MERRA2 assimilates precipitation

(<https://gmao.gsfc.nasa.gov/pubs/docs/McCarty885.pdf> Table 1)

The sentence will be modified as follows:

“Furthermore, none of the reanalysis products assimilate gauge-based precipitation observations (MERRA2 however incorporates satellite infrared and microwave measurements) and as such...”

P2 L16 There are significant errors associated with rainfall measurements from RADAR. I would not call them 'Highly accurate' (e.g. Krajewski et al (2010) gives a good summary of radar-rainfall uncertainties, focusing on improvements since the key paper by Wilson and Brandes in 1979. A more recent paper which focuses more on the applications in urban hydrology is Thorndahl et al (2017) lists a lot of uncertainties. Attempts to model the uncertainties are given in Villarani et al (2014), Rico-Ramirez et al (2015), Bong-Chul Seo and Krajewski (2015) and Cecinati et al (2017).)

Thank you for the literature on radar uncertainties. The sentence will be modified as follows:

“Radar estimates provide high spatial and temporal resolution estimates of rainfall over local regions, however these estimates can be inaccurate compared to rain gauges (Krajewski et al.

2010, Villarini and Krajewsky 2010 and McKee and Bins 2015), and very few national networks of radar observations exist.”

P2 L19 Replace 'global if not quasi-global' with 'global/quasi-global'

The change will be implemented.

P2 P3 I think that GPM should be mentioned as this is the most cutting edge satellite measurement of rainfall, Also mention blended datasets:

MSWEP (<https://journals.ametsoc.org/doi/full/10.1175/BAMS-D-17-0138.1>), Blended gauge/satellite (<https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD020686>) etc.

The following text will be added on P2 L30 before the sentence beginning with “The biggest limitation...”:

“New satellite missions and technology will be able to overcome these shortcomings over time. For example, the recently launched Global Precipitation Measurement (GPM) mission is an international satellite mission that aims to improve the detection of light rain and snowfall as well as provide quantitative estimates of precipitation particle size distribution (Hou 2014).”

A note on blended datasets will also be added at the end of the paragraph (P2 L34):

“Very recently, datasets that blend together precipitation estimates from multiple sources such as gauge observations, satellite observation and even reanalyses have become available. Examples include MSWEP V2 (Beck et al. 2019), CHIRPS (Funk et al. 2015) and Shen et al. (2014). These datasets offer very high temporal and spatial resolution data with a reasonably long temporal record. However, these datasets may exhibit increased temporal variability due to the incorporation of various observational sources over time and do not include as many in situ station observations as the gauge-only datasets.”

P4 L15 Given that you are taking the GPCC to be the 'best' dataset, I would list that first.

GPCC will be placed first in the list.

P4 L15-18 It would be useful to put in brackets the number of gauges in each data source.

The list will be modified as follows:

1. Global Precipitation Climatology Centre ... (approx. 100,000 stations)
2. Global Historical Climatology Network - Daily ... (103,635 stations)
3. Other: Argentina and Russian stations (approx. 1000 stations)”

P4 L18 Include a table of 'other' data sources that can be referred to here. Currently left wondering for a long time what the other data sources are. Is it just Argentina and Russia? If so, just state that.

We will include the information about “other” data as shown in the modified list above.

P4 L24 Re-phrase to be a call to action for met services to share more data.

This series of sentences had to be picked very carefully in order to not offend the agencies that are currently providing GPCC with in situ data. As such we have to be careful not to use strong language, however, in the interest of increasing the sharing of data we can add the following sentence at the end of the paragraph. “We encourage maintainers and providers of data to advocate for increased and more open sharing of meteorological data within their organisations.”.

P5 L3 Any idea why there was a decline in 2010?

We suspect that due to the manual data acquisition and quality checks employed by GPCC, recent data is slowly acquired and manually incorporated into the high quality archive.

P6 F2 Awesome! I can't really see any of the 'other' points though

As seen in Figure 1b, there are only 43 unique stations that are incorporated in the end by REGEN. As such they are difficult to spot but they can be spotted in Argentina and Russia (only 3 in Russia).

P6 L7 Clarify if you used existing QC code or if you rewrote it based upon Durre et al.

The following sentence will be added after the first sentence of the paragraph.

"Only minor changes to account for different data formats were made to the original QC procedures from Durre et al. (2010) before applying them."

P7 L1 Describe any validation of the QC code that was undertaken to assess correctly/incorrectly flagged values.

The following sentence can be added at the end of the sentence on P7 L2:

"For a thorough account of the validation of each QC check including the respective false-positive rates, see Durre et al. 2010. The total false positive rate based on all checks is 1% (Durre et al. 2010)."

P7 L11 Why did you choose 70% for a threshold?

The following sentence will be added after the line:

"We chose 70% because the same threshold was used by GPCC for creating their daily gridded products (Schamm et al. 2014). Haylock et al. 2009 (E-Obs) also use a similar threshold of 80%."

P7 L13 Why is there a drop in Indian station data in the 70s?

This is because India has not shared any new data since 1970s. The line can be modified as follows:

"The spike in missing month percentage in South Asia is because there are no Indian stations available after 1970."

P8 L12 'Higly' typo

"Higly" will be changed to Highly

P8 L8 1 degree seems like a huge area to consider gauges to be the same over. Why was such a large area used? How many gauges were merged in this way?

The following sentence will be added at the end of the sentence on P8 L10:

"A search radius of 1 degree was necessary to allow for many stations to be compared with each other in order to account for possible inaccuracies regarding stations metadata (coordinates)."

We understand that this may not be the most intuitive approach. As stated on P8 L11 this approach will be modified in the next iteration of REGEN. Unfortunately, we did not record how many gauges were merged in this way.

P8 L15 Please could you elaborate on the merging process. how were records combined? did you use the whole record from the highest quality source? or did you insert, for example, a few days from a GHCND record into periods of missing data from a GPCC record? If so, did you replace all missing days or did you only replace when there were a whole month of values? Do you have any idea of how many values were merged this way and whether it impacted the homogeneity of records?

No completeness criteria was applied before replacing missing values from higher quality sources. All together there were 36,828 station records that were created by merging records from 2 or more sources.

The following description will be added at the end of the paragraph:

“This way if data from a higher quality source was missing, it was replaced with data from a matching station from a lower quality source but not vice versa. Note that this approach may introduce inhomogeneities in the raw station data.”

P8 L29 Was it ever the case that the daily gauges showed no rainfall for the month but the monthly gauge did? If so, what did you do? Also, what do you do with one or more missing days of data? By disaggregating a monthly rainfall value with an incomplete daily record, you will most likely be increasing the average daily rainfall and reported extremes.

We do not calculate ratios for interpolation based on any monthly gauges. The monthly totals are calculated from the daily records. For example, if for a particular month, there are at least 70% of days with non-missing data, then we simply add them up to generate a monthly total. This is explained on P9 L5, however, we can move it to the start of the paragraph after the first sentence. “The monthly totals for calculating daily ratios in the station timeseries were obtained by summing the daily station data as well. A month was considered complete if it had at least 70% of non-missing days.”

As a result, the sentence beginning on P9 L6 will be modified as follows:

“A disadvantage of interpolating anomalies was that even if a daily record existed, it was not used for interpolation if the monthly total was missing because of the completeness criteria.”

P8 L30 Please clarify how the monthly data is used. Is it the case that you are effectively temporally disaggregating the monthly rainfall to a daily timestep, ultimately preserving the monthly values? Or will the monthly totals end up being slightly different?

Prior to interpolation the daily ratios (fractions) add up to 1 for each month for each individual station. However, because of the interpolation procedure itself the interpolated grid cell estimates of the ratios may not add up to 1 for the month. As such the monthly totals are not necessarily preserved.

P9 L25 I'm confused by point 3- how are the number of observations different to the number of stations? Is it that you may have 5 stations in the grid box but one of them has no data for that day and so the number of observations is 4? Please clarify.

This is because in some grid boxes with low numbers of stations, stations outside the grid box may be used for interpolation. The last sentence of point 3 will be extended to include this explanation: “Note that this is not the number of stations used for interpolation of that grid cell estimate, as stations outside the grid cell may be used for interpolation in some cases where station density is low.”

P9 L27-35 Please add in some descriptive numbers of how 'wrong' the interpolated rainfall can be. The sentence starting on P9 L31 will be modified as follows:

“The largest CoV values (maximum of 2.06) are once again seen in Africa, South America, Greenland and Southeast Asia (figures 5c and 5d). This means that the variance between the grid cell estimate and the observations used for interpolation is around twice as large as the average precipitation in these grids. Grids with CoV greater than 1.9 make up less than 0.1% (18 all together) of the grid cells with the mode of CoV being around 1.”

P10 L9 Replace 'trends' with 'changes' (we should discourage the use of trends in hydrology: <https://doi.org/10.1016/j.advwatres.2017.10.015>)

"Trends" will be replaced with "changes" on P10 L9.

P10 L11 replace "we highly encourage users to" with "Users must"

The sentence will be modified as per the recommendation:

"Users must use a dataset..."

P10 L15-17 This is very unclear. You use 'either' but do not provide an 'or'. Please clarify.

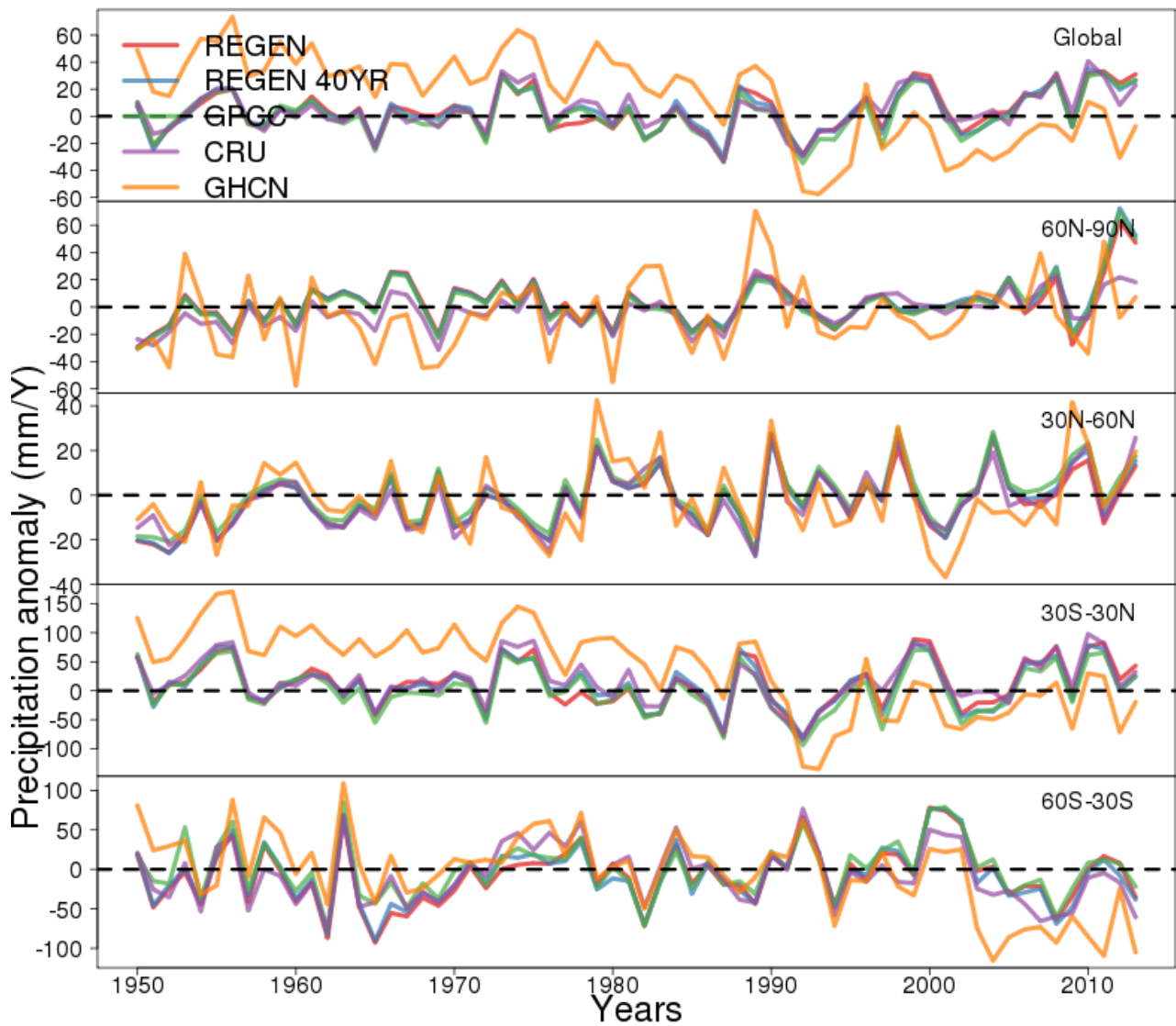
There is an "or" on L16 after "least one station". A comma can be added to make it clearer:

"...either contained at least 60% days in every decade from 1950 to 2013 (7 in total) with at least one station, or both the grid cell coefficient of variation..."

P11 L2 Is this because of the QC applied to GPCP and REGEN?

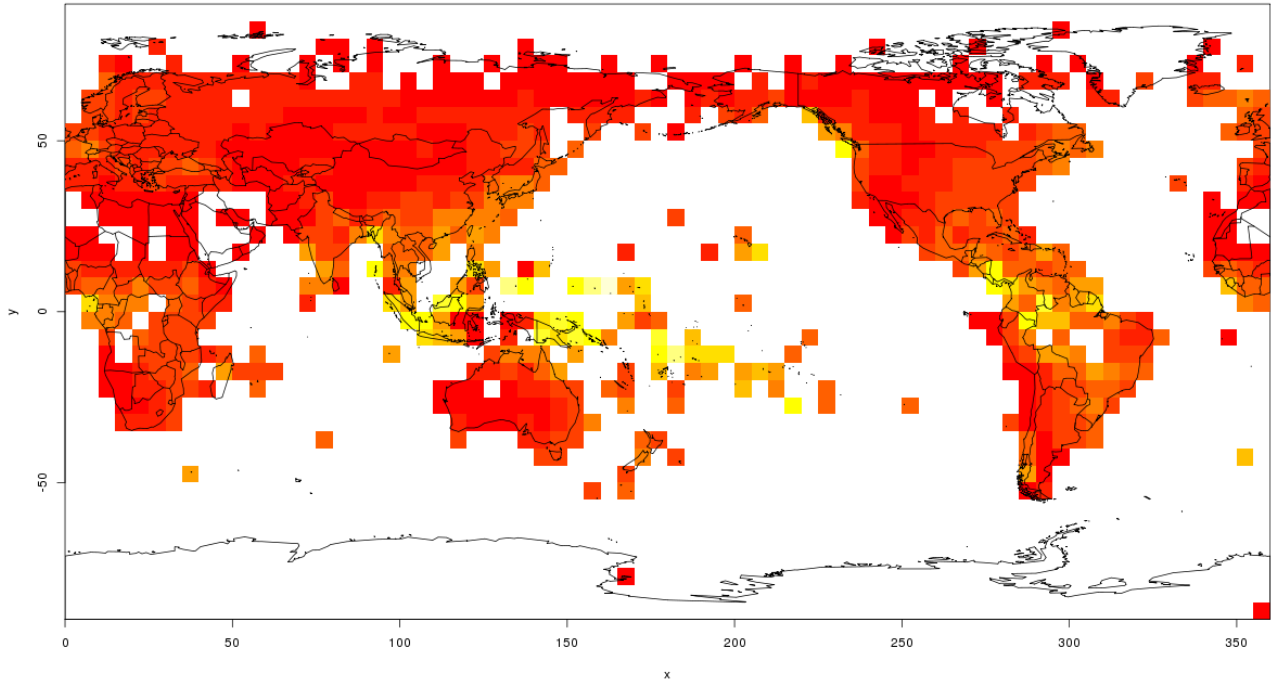
We do not fully understand why this is the case. We have re-downloaded the data and re-plotted the figure with different climatology periods (see plot below which uses a 1981-2000 climatology instead of 1951-2010 climatology as in the main text) however the enhanced variability remains.

GHCM-Monthly is also QC'd albeit slightly differently to REGEN so we do not think that is the reason behind this variability either. We suspect that this may be due to the coarse resolution (5° x 5° compared to 1° or finer for the other datasets) and poor spatial coverage of GHCM-Monthly dataset (see plot of annual totals for an example year, 1981, using GHCM-Monthly below).



Above: Comparison of annual precipitation anomaly timeseries with monthly datasets. Anomalies were calculated relative to the average of daily precipitation totals over 1981-2000 for each dataset.

Below: Annual total precipitation for 1981 based on GHCN-Monthly V2 gridded precipitation shown as an example of spatial coverage and resolution of the dataset.



P11 L9 Can you provide any estimates of how many more stations are used in national datasets compared to those used in REGEN?

The following table will be included in the main text:

Regional Dataset Name	Regional Dataset Stations	All REGEN v1.1 stations	Long term (40yr) REGEN v1.1 stations
APHRODITE	Daily max of 8000+	8551	1539
SA-Obs v1.0	7956	2527	64
E-Obs	17,468	28,338	11,261
CPC CONUS	~28,500	42,229	3940
AWAP	Daily max of ~7500	12,993	1424

P11 L10 State what interpolation method E-Obs used.

We can add a note in the parenthesis after mentioning E-Obs on P11 L10:

“...and both 0.25 degree and 1 degree versions of E-Obs Version 16 (Haylock et al., 2008; note that E-Obs also uses CAI with global Kriging to interpolate the daily anomalies) for the events...”

P12 L19 Comma after Africa

A comma will be added after “Africa”.

P13 L17 Remove ‘running’

Running will be deleted from the sentence.

P13 L28 Replace 'are encouraged to' with 'should'

The sentence will be modified as follows:

"User of REGEN should use the quality mask..."

P14 L25 It is such a shame that REGEN will not be updated.

We just wanted to clarify that REGEN is not an operational product created by a national meteorological agency but instead created during a PhD project. Resource permitting, we would of course continue maintaining REGEN as long as possible.

P14 Include a discussion of the limitations of 1 degree dataset. Rainfall is highly spatially variable and a 100km² estimate is unlikely to contain the information necessary for many typical rainfall applications.

The start of the paragraph on P14 L16 will be modified as follows:

"Rainfall is highly variable and a 1 degree spatial resolution (roughly a 10,000 sq. km) dataset such as REGEN is unlikely to contain the information necessary for many typical local-to-regional rainfall applications. However, we note that actual rainfall amounts in gridded datasets are subject to large uncertainties anyway (ref. Herold et al 2016, GRL), which likely complicated hydrological applications, whereas estimates of variability are more robust. We therefore believe that REGEN will prove itself valuable for climatological applications including studies of climate variability and long-term changes in daily precipitation intensity and extremes, as it provides long temporal coverage of quasi-global daily precipitation observations."

P14 L31 Why should we expect differences in the total annual precipitation between REGEN and REGEN-40 if they have both been adjusted by the monthly data? Are the monthly totals not necessarily preserved?

As explained with regards to the comment on P8 L30, the monthly totals are not necessarily preserved after interpolation. Since the station distribution between REGEN and REGEN40YR is different, the interpolated fields of ratios will also be different, and hence the monthly totals will in turn be different as well.

P15 L1 Replace 'REGEN has proved itself by providing' with 'REGEN provides'

The sentence will now begin with "REGEN provides"

P15 L4 Remove 'To this note' and include 'therefore' after are.

The sentence will be modified as follows:

"REGEN and its variant REGEN40YR (which minimises station network variability) are therefore accompanied by various uncertainty estimates as well as a quality mask..."

P15 L7 Include a statement about copyright/useage. Can anyone use this dataset freely? Industry? Or just for research?

The following copyright statement will be added:

"Licence & Rights:

Non-Commercial Licence: CC-BY-NC-SA

Creative Commons - Attribution - Non Commercial - No Derivatives 4.0 International

<http://creativecommons.org/licenses/by-nc-sa/4.0/legalcode>

Access to this dataset is free, the users are free to download this dataset and share it with others and adapt it as long as they credit the dataset owners, provide a link to the license, and if changes

were made, indicate it clearly and distribute their contributions under the same license as the original, commercial use is not permitted.”

FIGURES

Please include labels on all of your figures/scale bars. This makes them much easier to interpret than having to refer to the caption.

Every colour bar for all figures will be modified to have a label above it.

F4.c The lower line is missing.

A bottom border will be added to the box around the map in Fig 4c.

F5 Caption: space needed between (KE) and (figures Missing '(figures 5c and 5d)' from the caption

The figure caption will be modified as follows:

“Figure 5. Kriging error (KE) (figures 5a and 5b), Coefficient of variation (CoV) (figures 5c and 5d) defined by...”

F9 Please label the columns and rows on the diagram, it would make it much easier to interpret. I think 9h and 9g are mixed up in the caption. Also, Rx1day is not defined anywhere in the text or figures.

Besides labels describing the individual maps above the colour bars, labels will be added in the margins describing the columns and rows.

“Rx1Day” will be replaced by “annual maxima” in accordance with the main text (P12 L8). Also a reference to figure 9 will be added at the end of the sentence (P12 L10):

“... (Chen and Xie, 2008; Xie et al., 2007; Chen et al., 2008) and GPCP Full Data Daily V1 (GPCP-FDD1) (figure 9).”

F10 Please label the columns and rows on the diagram, it would make it much easier to interpret.

Similar modification to Fig 9 will be made to Fig 10.

Style points: - I found the italicisation of latin terms like 'in situ' distracting , especially in the abstract. - Why are you using 'in situ measurements' as opposed to 'gauge measurements' as your terminology?

“In situ” and “gauge” are equivalent. There was no particular reason behind choosing to use “in situ”. We will continue to use “in situ” but all italicisation of “in situ” will be removed from the text.