Hydrology and
Earth System
Sciences
Discussions

Interactive comment on "Rainfall Estimates on a Gridded Network (REGEN) – A global land-based gridded dataset of daily precipitation from 1950–2013" *by* Steefan Contractor et al.

Anonymous Referee #2

Received and published: 30 April 2019

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.

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)

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

P2 L13 NASA MERRA2 assimilates precipitation (https://gmao.gsfc.nasa.gov/pubs/docs/McCarty885.pdf Table 1)

P2 L16 There are significant errors associated with rainfall measurements from RADAR. I would not call then '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).)

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

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.

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

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

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.

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

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

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

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

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

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

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

P8 L12 'Higly' typo

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?

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?

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 disaggegating a monthly rainfall value with an incomplete daily record, you will most likely be increasing the average daily rainfall and reported extremes.

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?

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.

P9 L27-35 Please add in some descriptive numbers of how 'wrong' the interpolated rainfall can be.

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)

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

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

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

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

P11 L10 State what interpolation method E-Obs used.

P12 L19 Comma after Africa

P13 L17 Remove 'running'

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

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

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 aplications.

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?

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

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

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

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.

F4.c The lower line is missing.

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

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.

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

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?

_____