

## ***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 #1**

Received and published: 13 March 2019

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

The paper presents and documents a global daily precipitation dataset at a 1 degree resolution. The dataset is compiled from several data sources. In the paper, discussions on the method to grid daily precipitation and the density of the network, including the consequences this has for the precipitation estimates on the grid, are included. In addition to REGEN, a dataset based only on the long-running stations is presented which will be temporally more homogeneous than the dataset based on all station data.

The study is a very welcome contribution to the field where for daily precipitation now

C1

many national and several regional datasets exist, but (up to now) not a global dataset based on in-situ measurements. The study is well written and clear and as far as I can judge, no problems in the analysis are there. However, the study could do with a more expansive and in-depth comparison against existing (regional) observational datasets - the current comparisons are too ad-hoc and uninformative.

My advise to the editor is to accept the paper with minor adjustments.

### **More serious concerns**

1. Metadata is often a problem (lacking, erroneous etc.). The day shift discussed in sect. 2.3 - a very necessary thing to do - is in the face of poor metadata an action which might be problematic. The give an example, according to the Appendix, the data from the Netherlands are shifted one day backward in time. This is appropriate for the manual rain gauges but not for the 30+ automatic weather stations which measure precip between 0-0 UTC. It could be that only the rain gauges are in the GPCC dataset - but the reader can't tell.

Checking with other NMHSs in Europe, I could confirm the necessity to shift the date, except for Hungary (there is a question to the Hungarian NMHS out now).

There is another confusing part of this date-adjustment. Are you aiming to get 24-hour values coinciding with a day defined by local time or by UTC? The reason for asking is that for Indonesia, my understanding is that measurements are from 7 - 7 local time, which is (nearly) 0 - 0 UTC.

The correlations with CPC (figure 11d) are low in some areas - could it be that an erroneous timeshift or a missed time shift could be related to the low correlation? I guess you have tried to shift the whole dataset back and forth and looked for areas on the globe with increases in correlation?

2. The comparison against regional dataset of daily precipitation (sect. 3.2) is too ad-hoc. In your article, you claim (rightly so) that national and regional datasets

C2

are based on a more extensive dataset. I would like to add that especially national datasets have a far greater detail in the understanding of the metadata. This means that a meaningful comparison can be made between national/regional datasets and the REGEN dataset. This should go beyond simply picking one event of a few days, averaging precip over a region and plot a few timeseries.

Please add a more expansive comparison with regional datasets like Aphrodite etc. Other datasets might be interesting to use as well, like the SA-OBS for Southeast Asia (van den Besselaar et al. 2017. doi:10.1175/JCLI-D-16-0575.1). Comparison you could make easily are the standard deviation of the daily difference, perhaps stratified over different periods, but other comparison metrics are equally useful. Given the particular focus on precipitation extremes by the international community - and of some of the authors of the paper - a dedicated focus of representation of extremes (beyond one example) is required.

I was taken by surprise when reading about the the "Great flood of 1968" in Southeast England and France. The article referred to (Jackson, 1977) never mentions France. Below are a few pictures from the E-OBSv19.0e for 14-16 September 1968 and the area you use to compare REGEN with the regional dataset is somewhat large compared to the area affected by this event.

### Other issues the authors may want to look into

1. page 2, line 16: here is is claimed that radar provides 'highly accurate' estimates of precipitation. It is my understanding that radar can underestimate extreme precipitation by as much as 40% (e.g. <https://journals.ametsoc.org/doi/10.1175/2009BAMS2747.1>)
2. the referencing to figures is a bit curious: The first 3 are referenced chronologically, but then on page 7, you refer to fig. 4b, the next reference (line 31) is to fig. 7b page 9 10 have refs to fig 5 and page 11 has a reference to fig 7.

C3

3. page 8, lines 4-7. Relocated stations often keep the WMO id and if the relocation is to a site in the vicinity, then your criterion labels the old and new station as the same. This may not be a problem for precipitation, but perhaps it is good to inform the reader about this.
4. page 11, line 10. There is no 1.0 degree version of the E-OBS. I guess you regridded the E-OBS data to the REGEN grid to arrive at the 1 degree resolution?
5. page 14, line 19-20. Here you make the point that there is an ordering in the number of stations used by national, regional and global datasets. The point is very valid, but the example provided is misleading. Herrera used 2756 stations, the E-OBS uses 210 station in Spain (incl. Catalonia) and for the whole of Europe, 15962 series are used. Hardly 'roughly the same number' as claimed.

### Very minor issues

1. page 3, line 4, It is the Climatic Research Unit (not Climate)
2. page 7, line 4, typo in procedures
3. page 9, line 20, perhaps an odd formulation?
4. references, many citations have the http address twice in the citation, e.g. Jackson (1977).
5. Appendix A. you apparently made an effort to make an alphabetical list - but didn't quite succeed. There are duplicates in the list too - like Indonesia.
6. caption fig. 5: fig. c d show what?

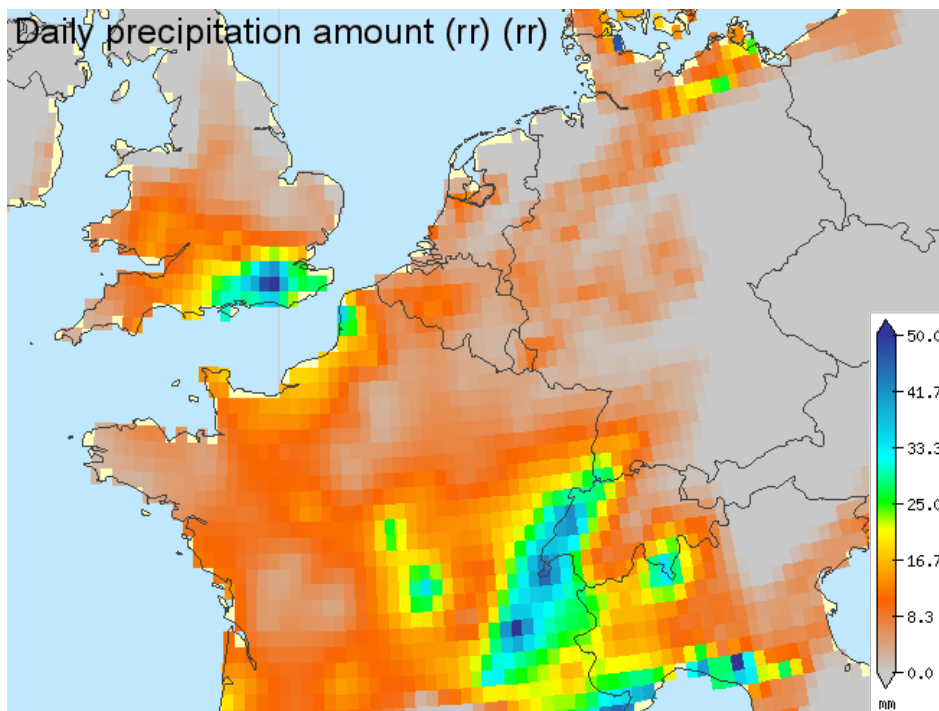


Fig. 1. rainfall for 1968/09/14 based on E-OBSv19.0e

C5

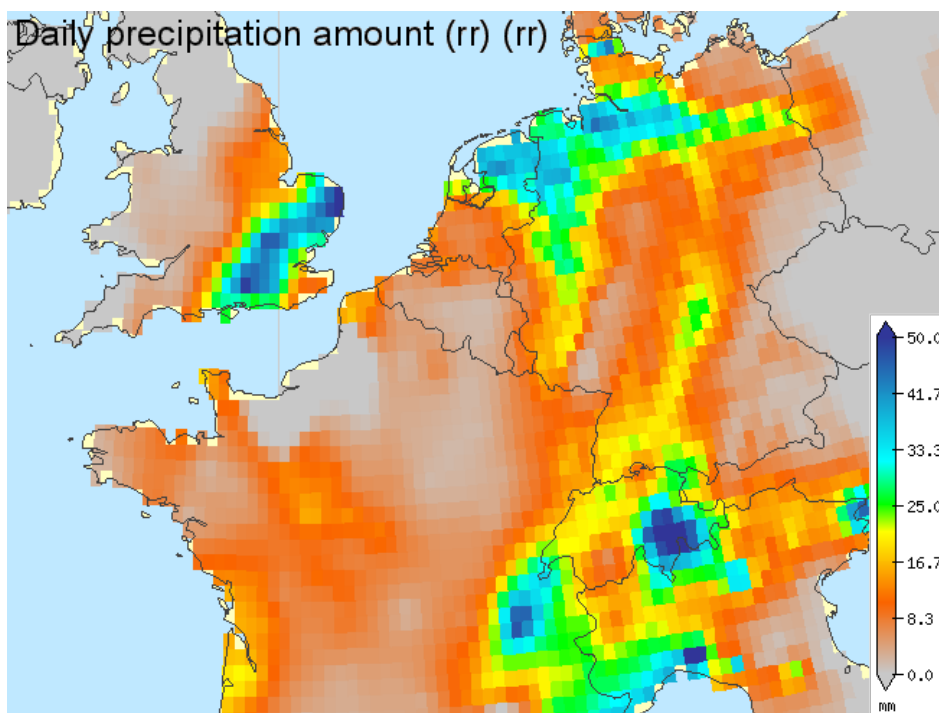
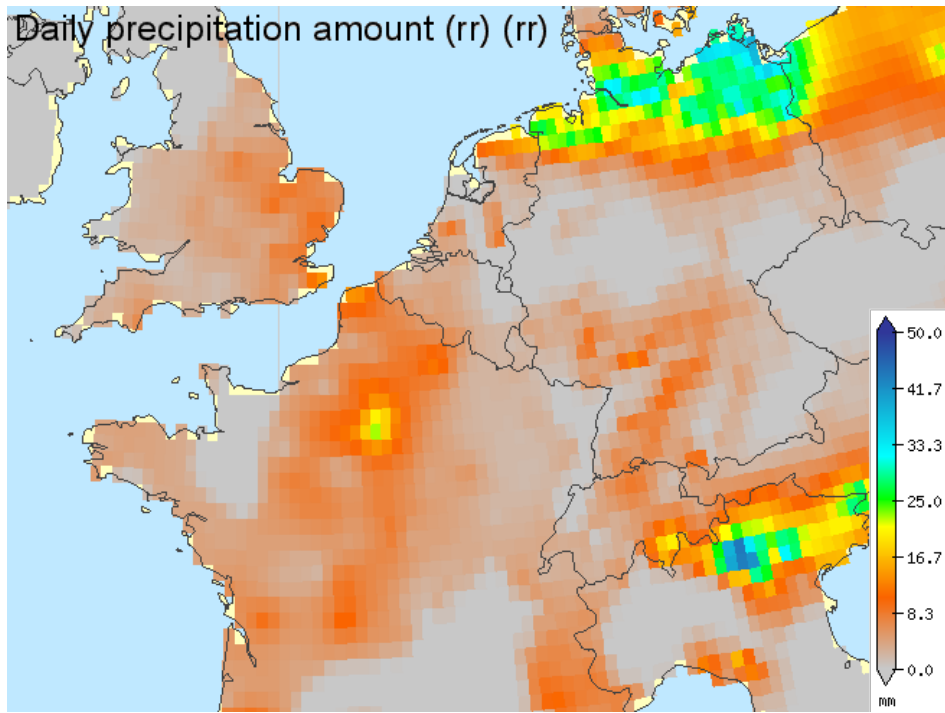


Fig. 2. rainfall for 1968/09/14 based on E-OBSv19.0e

C6



**Fig. 3.** rainfall for 1968/09/14 based on E-OBSv19.0e