

Reviewer's original comment in black

Our response to each comment in blue

In the submitted paper, authors investigated relationship between gauge data and gridded daily precipitation datasets. Multiple variables were used for the comparison with focus on the precipitation and rainfall erosivity. The topic of the paper could potentially be of interest for readers of this journal. However, there are several drawbacks related to the submitted manuscript that should be improved before further evaluation of this manuscript.

Thank you for the review and most valuable comments and suggestions. We have carefully considered each of the comments and outlined, following each comment, how we will address these comments and how the manuscript will be revised.

Most importantly, based on the presented results (Table 5) and second aim of the study (i.e. develop a correction factors) I think that authors should perform additional investigations in order to fulfil this second aim since according to Table 5, the developed correction factors do not lead to improved results (at least not for all cases).

Additionally, there are multiple parts that should be either better explained or enhanced (some specific comments are listed below). More specifically, I am missing a take home message that could be useful for the international readership. Authors state that correction factors need to be applied but the factors they developed have local characteristics and do not even improve the results in all investigated cases.

In the revised manuscript, we will highlight the inherent underestimation of rainfall erosivity when gridded precipitation products are used to compute this, and provide the bias correction factors for areas where gridded precipitation products remain the only data source. For regions where point (gauge) data are available to compute rainfall erosivity as in the case of China, bias correction factors are not needed.

We will revise the objectives of this study in the revised version to make these clearer as we have found that the objectives were not clearly stated in the manuscript as follows: (1) to contrast the gridded daily precipitation products with gauge data in terms of PDFs and extreme precipitation amounts, and to evaluate the smoothing effect of interpolation when areal precipitation for grid cells were generated using point (gauge) observations;

(2) to evaluate the magnitude of underestimation of rainfall erosivity calculated using

gridded daily precipitation products compared with that produced by spatial interpolation of rainfall erosivity computed using point (gauge) observations;

(3) to establish bias correction factors to improve the accuracy of rainfall erosivity maps where only gridded precipitation products were available for estimating rainfall erosivity over large areas.

We will also highlight the take home messages in the Discussions and Conclusions sections as follows:

For rainfall erosivity estimation in other parts of the world:

- (1) since gridded daily precipitation products are average values over grid cells, they are inherently less extreme compared with gauge observations, the R-factor calculated from gridded products are systematically underestimated by 10-40% compared with that generated by spatial interpolation of the gauge calculated values, and the magnitude of underestimation is larger (30-50%) for the 1-in-10-year event EI_{30} ;
- (2) it is better to use gauge observations to estimate rainfall erosivity where gauge data are available. When only gridded daily precipitation products are available, a bias correction factor ranging from 1.4~2.0 could be used to improve the accuracy of estimated rainfall erosivity. Without bias correction, direct application of gridded precipitation products would lead to systematic underestimation of rainfall erosivity in most parts of the world.

For wider research community:

- (1) because the gridded precipitation products (represent areal average precipitation over a grid cell) are generally unable to capture the extreme precipitation amount, compared to gauge observations, care needs to be taken to avoid systematic bias when using gauge-based, satellite-based gridded precipitation products, reanalysis products and outputs of climate models as inputs to nonlinear models often found in hydrology and agriculture;
- (2) when empirical models, such as the daily rainfall erosivity model considered in this study, are developed based on point observations, and gridded precipitation is the only available data source, bias correction is absolutely needed, or alternatively the model needs to be re-calibrated at a commensurate spatial scale.

Moreover, it should be noted that there are already quite some things done in relation to the rainfall erosivity assessment, even at global scale (for example Global rainfall erosivity assessment based on high-temporal resolution rainfall records by Panagos et al., 2017) using high-temporal resolution data. Thus, this assessment of the erosivity

using daily data (either gridded or point-observed) should be well justified. Also in China you have a nice network of high-temporal resolution data. Additionally, there are also some satellite products already developed that have sub-daily temporal resolution. Thus, I am missing a better justification of using of daily data because also for example Yue et al. 2020b map is based on the hourly data. Thus, why would one need to estimate the erosivity based on daily data if a map based on hourly data is already developed and available? Why dealing with daily data since such estimates of erosivity (based on daily data) should only be used in cases without hourly or sub-hourly data because they are less accurate.

As mentioned in Line 369 in the discussion section, though rainfall erosivity maps based on high-temporal resolution rainfall records are currently available, they are not easy to update in a timely manner since the collection of high-temporal resolution rainfall data is harder. In China the sub-daily precipitation observation datasets are not freely shared with ordinary users. The current rainfall erosivity assessments based on high-temporal resolution rainfall records generally use fewer stations in China, for example, high resolution rainfall records from only 387 stations in China were used in the study of Panagos et al., 2017. Many satellite products or radar rainfall products have sub-daily temporal resolution, but these products tend to have large bias and short time series. In many regions of the world, daily data are still the best choice considering both availability and accuracy. Besides, when predicting the future rainfall erosivity, the daily precipitation outputs from the climate model are much more reliable than the hourly or sub-hourly precipitation outputs.

Some specific comments:

L76: how can a spatial map be highly accurate since no information about the actual R-factor is available. In order to obtain a value that is as close to actual drop-size distribution measurements are needed, which can only performed for specific station.

Accurate R-factor values depend the rainfall intensity-kinetic energy relationship that is site-specific. In the manuscript, 'highly accurate R-factor' was meant to indicate that direct interpolation of gauge-based R-factor values would be more accurate than that using gridded precipitation products generated through interpolation of point rainfall observations. We will revise these sentences when revising the manuscript.

L116: I think that more detailed description of the gauge data should be provided. What is the equipment used, is the data verified, what is the data quality, anything that have

an effect on the results of this study should be included.

Thanks for your advice. We will add more details of the gauge data in the revised version.

Table 1: What is the number of stations in the period 2006-present for the CPC, more than 17000 or more than 700?

Sorry for our mistake. After reviewing the references, we have corrected Table 1 as follows:

Table 1. Basic information on gridded daily precipitation datasets.

Data Source	CPC	GPCC	CN05.1	NMIC
References	Chen et al., 2008	Schamm et al., 2014	Wu et al., 2013	Shen et al., 2010
Spatial resolution	0.5° × 0.5°	1° × 1°	0.25° × 0.25°	0.5° × 0.5°
Interpolation method	Optimal interpolation (OI) with anomalies	Ordinary block kriging with anomalies	Climatology—thin plate smoothing splines; anomaly—angular distance weighting	Optimal interpolation (OI) with anomalies
Coverage	Global land surface	Global land surface	China	China
Period	1979.1.1–present	1982.1.1–present	1961.1.1–present	1957.1.1–present
No. of stations	1979–2005: ~30000; 2006–present: ~17000	6000~8000 in Global Telecommunication System (GTS)	~2400	~2400
No. of stations in China	~200	~200	~2400	~2400

Eq. 2: Why a threshold of 10 mm is used if standard RUSLE threshold is 12.7 mm or 6.25 mm in 15 min?

Xie et al. (2016, Line 535 in the manuscript) evaluated the accuracy of the daily model (Equation 2) and reported that when the erosive daily rainfall threshold was set to 12 mm, average annual erosivity values were 1.1–6.2% lower than measured values defined in the USLE. When the erosive daily rainfall was set to be 9.7 mm, the average relative deviation reached the minimum (–0.1%). When the daily erosive rainfall threshold was set to be 10 mm, the average relative deviation was –0.5%, which has a

negligible difference from that when the threshold was set to be 9.7 mm. Besides, the threshold of 10mm corresponds to the demarcation between light and moderate rain stipulated by the China Meteorological Administration (CMA, 2003). For ease of use, the erosive daily rainfall threshold was determined to be 10 mm.

L158-159: These variables should be better explained and these sentences should be rewritten.

Eq. 3: Can ARF be defined twice using different variables?

Thanks for your advice. We will make these sentences clearer and change the index name, considering your comments on Figure 4.

Table 2: What is the difference between mean annual precipitation from only wet days or from both wet and dry days?

Wet days here is defined as the days when precipitation ≥ 1 mm to correspond to *WD*. We will revise the definition of *PRCPTOT*: Mean annual total precipitation from days when precipitation ≥ 1 mm.

Eq. 9: I do not understand this equation, R_{ref} is used on left and right side? Thus, a can only be 1?

Sorry for our mistake. The equation 9 should be $R_{gri} = a \cdot R_{ref}$. This typo will be corrected in the revised manuscript.

Figure 2: The readability of these figures is too low.

For the figures that are not clearly displayed in this manuscript, we will attach high-resolution images to the comments. If they are still hard to read and understand, we will consider changing the visual presentation in the revised manuscript.

Figure 3: Can you really say that these are PDFs? You are showing number of rainfall events in different bins? Or at least better visual presentation should be made since it is not easy to see which dataset yields better agreement with observed data. Additionally, can you add a summary of these differences between models and observed data? Thus, which model/dataset yields the best fit to the observed data.

We agree that the results in Figure 3 are not PDFs of the daily precipitation amounts, but histograms of the frequency of precipitation in different ranges which can reflect the PDFs. We will revise our wording and Figure 3 to improve the readability of the results. NMIC yields the best fit to the observed data, and it is reasonable to believe that NMIC is the closest to the real situation of areal average precipitation field among the four gridded products. We will add more clarification on this point and add a table including the results of Figure 2 and Figure 3 to make a summary.

Figure 4: I am sorry but I cannot understand this figure since obviously I do not understand correctly what should the ARF be according to your study. Perhaps this is related to the definition in L158-162 and Eq. 3 that should be improved.

What is usually defined as areal reduction factor can be seen here (for example):

<https://www.sciencedirect.com/science/article/abs/pii/S0022169418301999>.

Thus, something different that is shown in Figure 4. For example, the ARF should be a value between 0 and 1. I suggest that authors try to make this a bit easier to understand (what is shown here) for the reader.

We agree that the original definition of ARF is the ratio between the areal average rainfall and the point rainfall, to obtain areal rainfall from point rainfall values of specified durations and return periods. And generally, the ARF varies between 0 and 1. In the manuscript we defined ARF as the ratio of gridded precipitation metrics/rainfall erosivity over gauge-based precipitation metrics/rainfall erosivity, which can be confusing. We intend to change the index name to *ratio* in the revised version.

L289: Any specific reason for such behavior?

As mentioned in the discussion section, the difference between the western and eastern region mainly comes from the station density used for generating the gridded dataset, and the precipitation characteristics. The eastern region has a higher station density (> 4 stations per 10,000 square kilometers) and relatively humid climate (precipitation > 400 mm/y), whereas in the western region the density of stations is quite low and the climate is relative arid (< 1 station per 10,000 square kilometers, precipitation < 400 mm/y). The more stations were used, the more accurate the interpolation are, the better the scale discrepancy between gridded and gauge precipitation can be simulated, and the better the linear relationships are (Line 347-257). Large temporal and spatial variability of precipitation in arid western region also increases the relative error of interpolation, and the relationship between gridded gauge precipitation is more difficult

to describe. Besides, among four gridded products, different number of stations and interpolation methods were used (Table 1), which lead to the different performances of correction factors. In addition to large random errors affecting the effect of linear regression, some correction factors in western region are close to 1, so the correction do not improve the results. We will clarify this more clearly in the Discussion section in the revised version. There are some mistakes about the number of stations used for the generation of the gridded products, and we will correct them in the revised manuscript.

Figure 5: The resolution of this figure is quite low. I would suggest to add the number of points in all plots, since it seems that some cases (h) or (d) have relatively low number of points compared to the size of the investigated area. Or is this already written in the caption where “grids” is used? Moreover, grids or grid cells?

Since it was assumed that the grid cells containing meteorological stations can better represents the relationship between R_{gri} and R_{ref} than those without meteorological stations, only grid cells containing stations were used to establish and evaluate the correction factors. Due to the sparse stations in the western region, the number of points in (b), (d), (f) or (h) is much less than that in the cases of eastern region. Moreover, due to the different spatial resolutions of different gridded datasets, the number of non-repeated grid cells selected in each case is also different. We have written the numbers of used grid cells in the caption, and we will change “grids” to “grid cells”.

Table 5: Correction in some cases leads to worse results? What is then the rationale behind adopting such “corrections” factors if the final result is even worse than without these factors.

Please see the responses to the general comment and the comment on L289.

Figure 7: I suggest to add a map that shows the difference between the erosivity map after applying correction and the Yue map.

We agree.

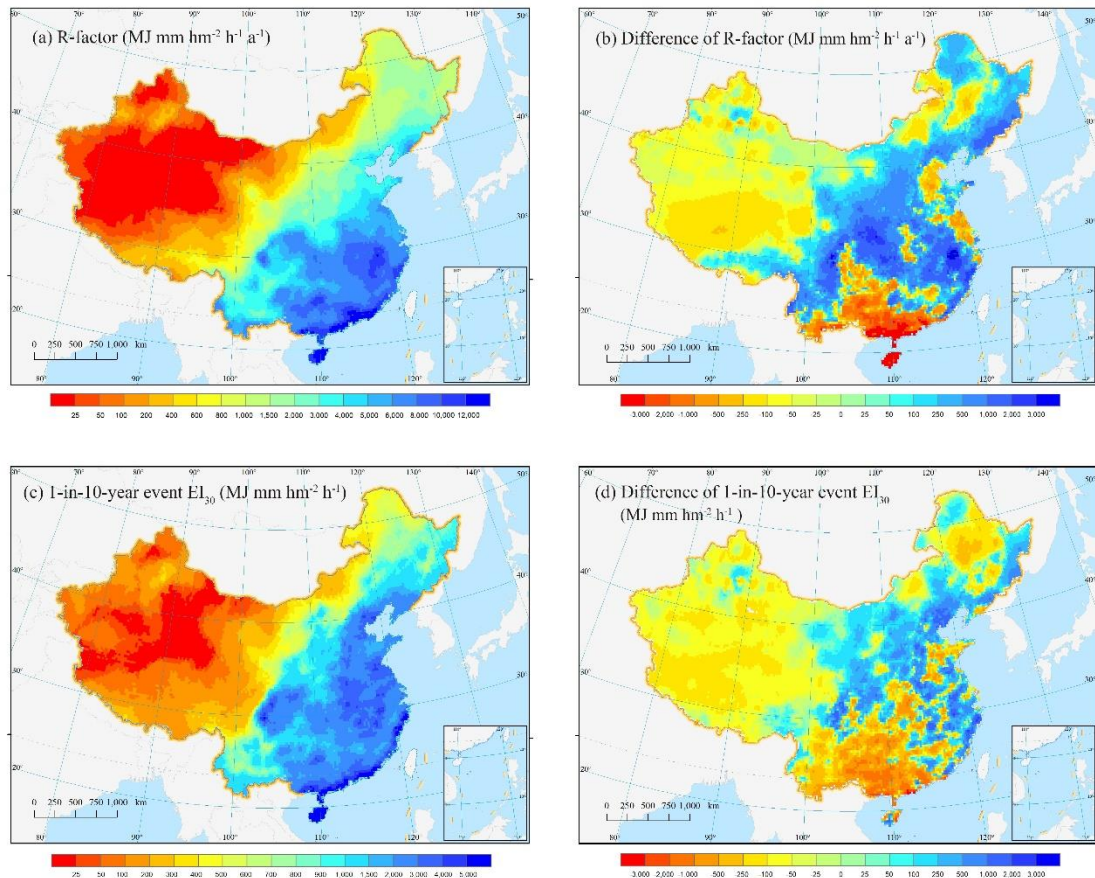


Figure 7. Spatial distribution of the R-factor and 1-in-10-year event EI_{30} using CN05.1 with bias correction factors (a and c), and the difference (b and d) in comparison with the original R-factor map (Yue et al., 2020b).

In Figure 7, (b) and (d) show that the erosivity maps with bias correction factors using CN05.1 still have marked discrepancy in comparison the with the original R-factor map of Yue et al. (2020b). The difference remained because the same correction factor value was applied to all grid cells in the eastern and western regions. However, the accuracy of these maps generated with CN05.1 has been noticeably improved. Using the bias correction factors to reduce the bias is necessary only when gridded precipitation products are only data source for erosivity estimation.

Discussion: What not merging results and discussion since you already have some discussion in the results section? And then perhaps also the results section would be easier to read and understand.

This is a suggestion worth considering. We will carefully think how to better organize the structure of the manuscript.

L385: “Reductions”: gauge data compared to grid data or grid data compared to gauge data?

Thank you for pointing this out. It should be “compared to gauge data, gridded products reduce the no-rain days and heavy precipitation days, and increase light precipitation days.”

L400: What is the purpose of using a correction factor if it does not yield improved performance? Is there any alternative, a better method that should be elaborated?

Please see the response to the comments on Table 5.

Conclusions: What are the practical conclusions of this study that could be useful for people dealing with rainfall erosivity in other parts of the world? What is the main take home message?

Please see the responses to the general comment.