Rainfall erosivity mapping in mainland China using 1-minute precipitation data from densely distributed weather stations

Yue Li Chen1, Yun Xie2, Xingwu Duan3, Minghu Ding1

1State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, 100081, China
2College of Arts and Sciences, Beijing Normal University at Zhuhai, Zhuhai, 519087, China
3Institute of International Rivers and Eco-security, Yunnan University, Kunming, 650091, China

Correspondence to: Yue Li Chen (chenylechina@yeah.net) and Yun Xie (xieyun@bnu.edu.cn)

Abstract. The risk of water erosion in mainland China is intensifying due to climate change. A high-precision rainfall erosivity dataset is crucial for revealing the spatiotemporal patterns of rainfall erosivity and identifying key areas of water erosion. However, due to the insufficient spatiotemporal resolution of historical precipitation data, there are certain biases in the estimation of rainfall erosivity in China, especially in regions with complex terrain and climatic conditions. Over the past decade, the China Meteorological Administration has continuously improved its ground-based meteorological observation capabilities, forming a dense network of ground-based observation stations. These high-precision precipitation data provide a solid foundation for quantifying the patterns of rainfall erosivity in China. In this study, we first performed rigorous quality control on the 1-minute ground observation precipitation data from nearly 70,000 stations nationwide from 2014 to 2022, ultimately selecting 60,129 available stations. Using the precipitation data from these stations, we calculated event rainfall erosivity and generated a national mean annual rainfall erosivity dataset with a spatial resolution of 0.25°. This dataset shows that the mean annual rainfall erosivity in mainland China is approximately 1241 MJ·mm·ha⁻¹·h⁻¹·yr⁻¹, with areas exceeding 4000 MJ·mm·ha⁻¹·h⁻¹·yr⁻¹ mainly concentrated in the southern China and southern Tibetan Plateau. Compared to our study, previously released datasets overestimate China’s mean annual rainfall erosivity by 31%–65%, and there are significant differences in performance across different river basins. In summary, the release of this dataset facilitates a more accurate assessment of the current water erosion intensity in China. The dataset is available from the National Tibetan Plateau/Third Pole Environment Data Center (https://doi.org/10.11888/Terre.tpdc.301206; Chen, 2024).

1 Introduction

Rainfall-induced soil erosion stands as the primary contributor to global soil loss, as highlighted by the Intergovernmental Panel on Climate Change (IPCC, 2019), posing a significant threat to soil functionality. This phenomenon jeopardizes various crucial aspects including food security, water quality, and climate change mitigation (FAO and IPTS, 2015; Panagos et al. 2020). Precipitation serves as the principal driver of erosion processes, influencing soil particle detachment, aggregate breakdown, and particle transport via runoff (Wischmeier and Smith, 1965, 1978). In this context, the rainfall erosivity index has been introduced to delineate rainfall's potential in causing soil loss, particularly through the Universal Soil Loss Equation...
(USLE), which links rainfall characteristics to soil loss based on extensive data from thousands of plot-years of natural rainfall and runoff (Nearing et al., 2017).

Rainfall erosivity is commonly assessed by multiplying the accumulated kinetic energy ($E$) by the maximum 30-minute rainfall intensity ($I_{30}$) of a rainfall event. From a dynamic standpoint, this index encapsulates the comprehensive effects on soil particle detachment and transport processes. The $E$ of a rainfall event can be quantified utilizing raindrop microphysics, which includes physical parameters such as raindrop size and falling velocity measured using distrometers. However, quantifying these parameters at a large spatial scale poses a considerable challenge due to the unrealistic for a dense observational network of expensive distrometers. To simply the calculation, substantial efforts have been made to detect the empirical relations between $E$ and rainfall intensity ($I$) (hereafter $E$-$I$ relation). Various $E$-$I$ models have been developed, employing linear, polynomial, exponential, logarithmic, and power-law functions. It is notable that the accuracy of $E$ is determined not only by the models used, but also the temporal resolution of the in-situ precipitation observations. Studies have indicated that $E$ values derived from 1-hourly in-situ precipitation data tend to underestimate those obtained from 1-minute data by approximately 10% (Agnese et al., 2006; Yin et al., 2007). In 2023, Dai et al. (2023) introduced the first global rainfall microphysics-based $E$, utilizing parameters retrieved from radar reflectivity. Compared to the radar remote sensing-based $E$ values, the multi-year averaged annual rainfall kinetic energy calculating using $E$-$I$ method was smaller with biases ranging from -6.17% to -12.5% across distinct regions worldwide.

The $I_{30}$ value for a rainfall event is derived from precipitation process data, including sub-hourly in-situ and remote sensing-based precipitation data. When the in-situ data is used, $I_{30}$ tend to be increasingly underestimated with increasing time intervals of precipitation data. It has reported that $I_{30}$ value is derived by multiplying maximum hourly precipitation intensity derived from hourly records by 1.668 (Yin et al., 2007). Consequently, the in-situ precipitation data with 1-minute temporal resolution are the best suitable data for deriving $I_{30}$ of an rain event. Recent years, there occurs some gridded precipitation datasets with high temporal resolution. However, it should be caution when the gridded data are directly used to calculate $I_{30}$, because large underestimation in $I_{30}$ has been widely reported. For example, ERA5, the state-of-the-art reanalysis precipitation data has the underestimation of $I_{30}$ by over 80% in the Tibetan Plateau (Chen et al., 2022). Satellite-based products, such as the Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (GPM, IMERG) dataset, have also shown significant underestimations in precipitation intensity, suggesting potential challenges in accurately identifying $I_{30}$ (Freitas et al., 2020).

Based on the analysis presented, the following conclusion can be drawn: The bias in estimating the $I_{30}$ of individual rainfall events is significantly larger than that for estimating $E$ under the latest available datasets. The estimation error of $I_{30}$ is the most crucial source of inaccuracies in determining rainfall erosivity. Strategies to achieve accurate $I_{30}$ values may include advancements in satellite remote sensing technologies to identify sub-daily rainfall events and their properties, coupled with leveraging in-situ observations from densely spaced weather station networks. Due to the limited availability of high temporal-spatial resolution precipitation data, rainfall erosivity in China has historically been estimated using precipitation data with coarser temporal resolutions, such as hourly, daily, monthly, and...
yearly scales from limited in-situ records and gridded precipitation data (Yin et al., 2015; Panagos et al., 2017; Liu et al., 2020; Yue et al., 2022; Chen et al., 2022, 2023). Consequently, the existing maps of the multi-year averaged annual rainfall erosivity (hereafter named $R$ factor) may contain uncertainties, particularly in regions with complex terrain and climate conditions (Chen et al., 2022). Since 2012, the China Meteorological Administration (CMA) has been developing a dense network of automatic weather stations, providing 1-minute in-situ precipitation records nationwide. This development presents opportunities to identify sub-daily rainfall events and their characteristics. By leveraging this dataset, we can derive accurate $I_{50}$ values and corresponding event rainfall erosivity for different precipitation events. Thus, this study aims to develop a high-quality $R$ factor map to enhance our understanding of water erosion across mainland China using the latest 1-minute in-situ precipitation dataset.

2 Data and methods

2.1 Data

2.1.1 Precipitation data

Over the past decade, approximately 70,000 weather stations have been gradually established by the CMA. These stations have maintained nearly 10 years of 1-minute interval precipitation observation records. Each station's data integrity was assessed using quality control codes at 1-minute intervals annually. Records with an integrity level exceeding 90% for a given year were deemed suitable for calculating annual rainfall at the respective station. Ultimately, data in 2014-2022 from 60,129 stations across the mainland China were utilized in this study. The entirety of mainland China was divided into 16,167 grids with a spatial resolution of 0.25°. Based on this grid division, we analyzed the density of weather stations and the coverage time length of available data at the grid scale (Figure 1). It was evident that the density of stations and coverage time length in the southeastern regions of China significantly exceeded those in the northwest. On a national scale, approximately 57% of all grids in mainland China had in-situ precipitation observations. Among these observed grids, there were an average of 6.7 stations per grid, with a data coverage spanning 5.2 years.
A monthly gridded precipitation dataset, released by the National Meteorological Information Center (NMIC) of the CMA, was also employed to analyze the spatial characteristics of precipitation across China (hereafter named CMA gridded precipitation data). This dataset was particularly valuable for regions like northwestern China and areas with complex terrain, where in-situ observations from weather stations was insufficient. The gridded precipitation dataset is based on precipitation data from national weather stations, and then interpolated spatially into 0.5° grids by using the Thin Plate Spline method. This study made use of data spanning from 2014 to 2022 to detect the spatial characteristics of multi-year averaged annual precipitation across China.
2.1.2 Rainfall erosivity maps in previous studies

In the results section, the newly developed $R$ factor map for mainland China in this study is compared with two existing maps from Panogos et al. (2017) and Yue et al. (2022). The former map is the first global-scale erosivity database (https://esdac.jrc.ec.europa.eu/themes/global-rainfall-erosivity). It is built upon sub-hourly in-situ precipitation records gathered from 3,625 stations spread across 63 countries worldwide. Notably, the distribution of stations varies across continents, with Asia and the Middle East boasting 1,220 stations (34% of the total) distributed among 10 countries, including the Siberian part of the Russian Federation, China, India, and Japan.

The latter map, released by Yue et al. (2022), employed hourly and daily rainfall data from 2381 stations spanning the period 1951-2018 to generate the $R$ factor map across mainland China. This study demonstrated satisfactory performance by comparing the derived values against true rainfall erosivity values calculated using 1-minute rainfall data collected from 62 stations across China.

2.2 Method for calculating rainfall erosivity

Not all rainfall processes lead to soil erosion. It is generally believed that a rainfall event must exceed a certain threshold in precipitation magnitude to cause soil erosion. According to Wischmeier and Smith (1978), a continuous six-hour dry period, without any rainfall, is used to delineate individual rainfall events. This means that if there is an interruption of more than six hours, subsequent rainfall is considered a separate event. Erosive rainfall events are defined as those where the precipitation exceeds 12 mm (Xie et al., 2000). The rainfall erosivity ($EI_{50}$) of an erosive rainfall event is calculated as follows (Brown and Foster, 1987):

$$e_r = 0.29[1 - 0.72\exp \left( -0.05i_r \right)](1)$$

$$E = \sum_{r=1}^{n} \left( e_r \cdot P_r \right)(2)$$

$$r_{event} = E \cdot I_{30} (3)$$

where $E$ (MJ·ha$^{-1}$) is the total energy of the erosive event, and $r_{event}$ (MJ·mm·ha$^{-1}$·h$^{-1}$) is the event rainfall erosivity. For the 1-minute in-situ precipitation data, $i_r$ (mm/h) is the rainfall intensity for the $r^{th}$ minute, $e_r$ (MJ·ha$^{-1}$·mm$^{-1}$) is the unit energy for the $r^{th}$ minute, $P_r$ (mm) is the rainfall amount for the $r^{th}$ minute, $n$ is the rainfall duration, and $I_{30}$ (mm/h) is the maximum contiguous 30-min peak intensity.

Due to occasional observation errors, particularly in low-temperature environments, we conducted a quality check on the calculated erosive event rainfall erosivity. Initially, for each station, we selected the nearest 100 stations and categorized the recorded events at these 101 stations into two distinct classifications: those occurring during the warm season (April to October) and those during the cold season (January to March, November to December). Subsequently, the median and standard deviation of event rainfall erosivity are computed for the warm and cold seasons, respectively. The threshold value for each station was then defined as the sum of the median and three times the standard deviation for different seasons. It is
assumed that any event rainfall erosivity surpassing the threshold within the season is considered an outlier and therefore excluded from the annual rainfall erosivity calculation. Ultimately, the annual rainfall erosivity at each station was aggregated using the quality-checked event rainfall erosivity values, with the $R$ factor representing the mean annual rainfall erosivity. The $R$ factor at one grid is the multi-station averaged value.

3 Results

3.1 Rainfall erosivity map

Based on the methodology described in Section 2.2, the erosive event rainfall erosivity of 60,129 stations across mainland China from 2014 to 2022 was calculated. Subsequently, the station-scale $R$ factor was obtained. The gridded $R$ factor values, with a spatial resolution of 0.25°, represent the average $R$ factor values of the stations within corresponding grids (Figure 2). The averaged $R$ factor in the grids with in-situ observations across mainland China is calculated to be 1917 MJ·mm·ha$^{-1}$·h$^{-1}$·yr$^{-1}$. Overall, the southern region of China exhibits the highest mean annual rainfall erosivity, followed by the northern region. The $R$ factor is minimal in the northwest arid and semi-arid areas, as well as in the Tibetan Plateau. The $R$ factor in the southern region generally exceeds 2000 MJ·mm·ha$^{-1}$·h$^{-1}$·yr$^{-1}$, with the highest values observed along the southeast coast, reaching over 10000 MJ·mm·ha$^{-1}$·h$^{-1}$·yr$^{-1}$. In contrast, the $R$ factor in the northwest arid and semi-arid areas and the Tibetan Plateau region mainly ranges below 500 MJ·mm·ha$^{-1}$·h$^{-1}$·yr$^{-1}$.

![Image](https://doi.org/10.5194/essd-2024-195)

Preprint. Discussion started: 20 June 2024
© Author(s) 2024. CC BY 4.0 License.
On a national scale, this study employs the Kriging method to generate the spatial distribution of the $R$ factor map with a spatial resolution of $0.25^\circ$ across mainland China. Notably, there are extensive areas of ground observation gaps in the northwest and Tibetan Plateau. It is essential to identify the impacts of the spatial interpolation method on the accuracy of the $R$ factor map in these regions. Due to the significant positive correlation between precipitation and rainfall erosivity (Richardson et al., 1983; Renard & Freimund, 1994; Yu et al., 1996; Xie et al., 2016; Chen et al., 2024), the spatial distribution of mean annual precipitation can help assess the potential bias in interpolated rainfall erosivity in observation gap areas. Here, the CMA gridded precipitation data is utilized to detect the spatial characteristics of mean annual precipitation. As shown in Figure 3a, most of the observation gaps in the northwest and Tibetan Plateau have relatively low annual precipitation, with minimal differences compared to annual precipitation in surrounding areas with observations. However, the Dawang-Chayu area in the southern part of the Tibetan Plateau, where famous Yarlung Zangbo River Grand Canyon is located, is an exception. Precipitation in this region is primarily affected by the southwest monsoon, which brings warm and humid airflow to the hinterland of the Tibetan Plateau along the Yarlung Zangbo River Grand Canyon (Chen et al., 2023). The observed mean annual precipitation (exceeding 1800 mm) in this region is much larger than that in its surrounding areas. Thus, using the $R$ factor calculated from surrounding stations to extrapolate rainfall erosivity for this area is unreasonable.

In 2022, we utilized hourly reanalysis precipitation data from the fifth European Centre for Medium-Range Weather Forecasts reanalysis (ERA5) precipitation data, combined with in-situ precipitation records, to generate a gridded dataset of annual rainfall erosivity for the Tibetan Plateau from 1950 to 2020 (Chen et al., 2022). This study used this released gridded dataset to calculate the mean annual rainfall erosivity from 2014 to 2020 for the Dawang-Chayu area, instead of relying on direct spatial interpolation values. Figure 3b illustrates the integrated $R$ factor map across mainland China. Generally, the $R$ factor in mainland China exhibits a decreasing trend from southeast to northwest, with an overall average value of 1241 MJ·mm·ha$^{-1}$·h$^{-1}$·yr$^{-1}$.
Figure 3: (a) The spatial distribution of mean annual precipitation in China. Grids without crossed diagonal lines indicates areas without station records. The area marked using black dashed line is the Dawang-Chayu region. (b) The spatial distribution of R factor across mainland China.
The newly generated $R$ factor map over mainland China is compared with the existing maps. Compared to the map developed by Panagos et al. (2017), there is a good correlation in regions where the mean annual rainfall erosivity is less than 10,000 MJ·mm·ha$^{-1}$·h$^{-1}$·yr$^{-1}$. However, in areas where the mean annual rainfall erosivity exceeds 10,000 MJ·mm·ha$^{-1}$·h$^{-1}$·yr$^{-1}$, our estimates are significantly higher than those of Panagos et al. (2017) (Figure 4a). When compared with the map developed by Yue et al. (2022), the correlation is good overall, but our calculated values are significantly lower than those of Yue et al. (2022). In regions where the mean annual rainfall erosivity exceeds 10,000 MJ·mm·ha$^{-1}$·h$^{-1}$·yr$^{-1}$, the differences are larger, but no clear pattern is observed. In summary, our results show high correlation with existing studies in areas with lower mean annual rainfall erosivity, but significant differences in areas with higher rainfall erosivity.

The performance of these three datasets are further compared across different watersheds in China. Mainland China has been divided into nine watersheds, including the Songhua and Liaohe River Basin, Haihe River Basin, Yellow River Basin, Continental Basin, Huaihe River Basin, Southwest Basin, Yangtze River Basin, Southeast Basin, and Pearl River Basin (Figure 5a). From the perspective of mean and median values, the Haihe River Basin and Huaihe River Basin show the largest differences among the three datasets. Overall, while there are some differences in the performance of these three maps across different watersheds, there is no consistent pattern (Figure 5b).
Figure 5: The box-plots of R factor in various watersheds of mainland China. Grey, red, and blue boxes represent different R factor maps released by this study, Panagos et al. (2017), and Yue et al. (2022), respectively. The lower and upper of the boxes were defined as the 25th percentile of R factor (Q1) and the 75th percentile (Q2). The difference (Q2−Q1) is called the interquartile range (IQR). The lower and upper bounds are calculated using (Q1−3IQR) and (Q2 + 3IQR). The plot and line in the box are the average and median values, respectively.

4 Data availability

The dataset is available from the National Tibetan Plateau/Third Pole Environment Data Center (https://doi.org/10.11888/Terre.tpdc.301206; Chen, 2024).
5 Conclusions

The rainfall erosivity of individual rainfall events is determined by the product of two parameters: the kinetic energy of the event \( E \) and the maximum 30-minute rainfall intensity \( I_{30} \). With current technological conditions, the accuracy of obtaining \( I_{30} \) during individual rainfall events is far lower than that of event rainfall \( E \). In other words, the deviation in estimated rainfall erosivity primarily stems from the deviation in the maximum 30-minute rainfall intensity. At present, using high spatiotemporal resolution ground precipitation observation data yields the highest accuracy in calculating the \( I_{30} \) for rainfall events, making the resulting rainfall erosivity estimates the most reliable. Therefore, this study utilized nearly ten years of 1-minute in-situ precipitation data from 60,129 stations to estimate the \( R \) factor in mainland China. The main results are as follows:

1. The \( R \) factor across the mainland China shows significant spatial variability, with the regional average being approximately 1241 MJ·mm·ha\(^{-1}\)·h\(^{-1}\)·yr\(^{-1}\).

2. Compared to our study, previously released datasets overestimate China's mean annual rainfall erosivity by 31%–65%, and there are significant differences in performance across different river basins.

This newly developed dataset for the average annual rainfall erosivity in mainland China is based on ground precipitation observation data offering the highest spatiotemporal resolution from the recent decade. On one hand, this data provides a foundation for the detailed assessment of current water erosion intensity in China. On the other hand, rainfall erosivity can also be seen as a characteristic of rainfall events, which has certain spatial indications for precipitation-induced disasters in China.

Author contributions

YC designed the study and wrote the manuscript. YX and BL contributed to the methodology. XD and MD contributed to the suggestions for manuscript revision.

Competing Interests

The contact author has declared that neither they nor their co-authors have any competing interests.

Acknowledgments

This research was jointly supported by the National Natural Science Foundation of China (42201156), National Key Research and Development Program of China (2021YFD1500700), and Basic Research Fund of Chinese Academy of Meteorological Sciences (2023Z004, 2023Z025). Thanks for the Resource and Environmental Science Data Platform for supporting the watersheds data in China (https://www.resdc.cn/data.aspx?DATAID=141).
References


