





OPEN

DATA DESCRIPTOR

A Chinese soil conservation dataset preventing soil water erosion from 1992 to 2019

Jialei Li^{1,2} , Hongbin He¹, Qinghua Zeng¹, Liding Chen^{1,2} & Ranhao Sun^{1,2} 

Soil conservation service (SC) is defined as the ability of terrestrial ecosystems to control soil erosion and protect soil function. A long-term and high-resolution estimation of SC is urgent for ecological assessment and land management on a large scale. Here, a 300-m resolution Chinese soil conservation dataset (CSCD) from 1992 to 2019, for the first time, is established based on the Revised Universal Soil Loss Equation (RUSLE) model. The RUSLE modelling was conducted based on five key parameters, including the rainfall erosivity (interpolation of daily rainfall), land cover management (provincial data), conservation practices (weighted by terrain and crop types), topography (30 m), and soil properties (250 m). The dataset agrees with previous measurements in all basins ($R^2 > 0.5$) and other regional simulations. Compared with current studies, the dataset has long-term, large-scale, and relatively high-resolution characteristics. This dataset will serve as a base to open out the mechanism of SC variations in China and could help assess the ecological effects of land management policies.


Background & Summary

Excessive soil erosion can negatively impact crop production, carbon transfer, soil organisms, and soil nutrition^{1–4}. To prevent soil erosion, many countries have attempted to improve soil conservation services, including China. China governments have made a lot of plans and policies to control soil erosion for recent years, for example, legislation for soil and water conservation, building agricultural terraces^{5,6}, 3-North Shelter Forest Program (since 1978)⁷, and Conversion of Farmland to Forests and Grass (Grain-for-Green) (since 1999)⁸. Recent research shows that China has led in a quarter of the global increase in greening after 2000⁹. However, China supports one-fifth of the world population with 9% of the total global cropland¹⁰, which means that the reduction of cropland has a potential to affect China grain security^{11,12}.

Over the past few decades, China has experienced an expansion of farmland, a massive growing total population (from 987 million people in 1980 to 1.413 billion people in 2020), and a large urbanization process^{13–15}. Although these increasing trends have slowed down in recent years, anthropogenic activities, along with climate change, have brought about non-negligible changes in soil erosion risk and soil conservation service. To support the huge population, agriculture may face the challenges to be expanded or intensified^{16,17}. Either approach puts pressure on the land, including soil fertility and water conservation¹⁸. To weight the relationship between human and the land, a reliable dataset of soil conservation service in China is urgent.

High-precision, long-term soil conservation service data are requested in China. However, accurate and high-precision soil conservation service datasets based on field measurement are unfeasible and money-consuming on a national scale. The Chinese government has decided to conduct the third national soil census from 2022 to 2025¹⁹; however, this would be an arduous task and hard to reflect long-term changes. Modelled soil conservation service data could meet these demands and be more feasible, which were usually based on modelling soil water erosion prevented by vegetation and practice measures²⁰.

Development of soil erosion models based on remote-sensing data and GIS technology makes it possible to assess soil conservation captivity and its dynamic changes. By far the most widely used empirical model to simulate soil erosion is the Revised Universal Soil Loss Equation (RUSLE) model, which predicts longtime average annual soil loss in specified experimental fields^{21,22}. They describe the relationship between soil loss rate and the following effects: soil properties, topography, vegetation cover, land management, and rainfall and runoff²³. In order to apply the USLE-family model on a large scale, many researchers have made improvements

¹Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China. ²University of Chinese Academy of Sciences, Beijing, 100190, China.  e-mail: rhsun@rcees.ac.cn

and enhancements to the parameters of the model^{23–29}. Yang *et al.*³⁰ develop a GIS-based RUSLE model to offer a pioneering overview of global soil erosion on cell grids. Now the USLE-based models have been used to estimate soil erosion on regional and global scales^{31–34}. They were also embedded in the InVEST model to compute the Sediment Retention Index³⁵. Although there were some uncertainties in the accuracy of global soil erosion estimation based on USLE-family models^{36–38}, considering the availability of data, model mechanism, and China's geographical conditions, we finally chose the RUSLE model. Currently, soil conservation and soil erosion data in China are mostly based on regional or watershed scales^{39–42}, only few studies are based on national scales^{43,44}. These national-scale studies were only for one year or average year data, lacking long-term series data.

Based on high-precision and long-term series data, this study calibrated the input factors and obtained the database of soil conservation service and soil erosion rate in China from 1992 to 2019. These results can provide data support for assessing soil erosion risk in China and drive analysis of past changes to enhance regional management. The main goal of this data is to analyze the interannual changes in soil conservation in China and provide a basis for identifying potential soil erosion hotspots. In addition, this dataset can be regarded as a comparison of soil conservation for other studies on different scales in the future. The results can provide a reference for further improvement of model parameters and lay a foundation for identifying driving factors of soil erosion changes.

Methods

We used the RUSLE (Revised Universal Soil Loss Equation) model to estimate the soil erosion rate in China during 1992–2019, and certain adjustments were made in the factor calculation according to the actual backgrounds in China. The model equations are as follows:

$$SE_p = R \cdot L \cdot S \cdot K \quad (1)$$

$$SE_a = R \cdot L \cdot S \cdot K \cdot C \cdot P \quad (2)$$

where SE_p is the predicted potential annual soil erosion on bare land ($t\ ha^{-1}\ a^{-1}$), SE_a is the predicted actual annual soil erosion ($t\ ha^{-1}\ a^{-1}$) on land with vegetation cover and erosion control practices, R is the rainfall erosivity factor ($MJ\ mm\ ha^{-1}\ h^{-1}\ a^{-1}$), LS is the topographic factor (dimensionless) with L being the slope length factor and S being the slope factor, K is the soil erodibility factor ($t\ h\ MJ^{-1}\ mm^{-1}$), C is the vegetation cover and management factor (dimensionless), and P is the support practice factor (dimensionless). Every factor was calculated according to the original resolution of the input data (for example, DEM data for $30\ m \times 30\ m$, soil properties data for $250\ m \times 250\ m$), then all the factors were resampled to a resolution of $300\ m \times 300\ m$ by bilinear interpolation and multiplied to obtain the soil erosion rate in each year.

In this study, soil conservation service (SC , $t\ ha^{-1}\ a^{-1}$) is defined as soil retention, which is soil water erosion prevented by vegetation and practice measures, equal to potential soil erosion minus actual soil erosion:

$$SC = SE_p - SE_a = R \cdot L \cdot S \cdot K \cdot (1 - C \cdot P) \quad (3)$$

The list of the input data is shown in Table 1, and the spatial distribution of the raw data are shown in Fig. 1. The framework to develop the dataset is shown in Fig. 2.

Estimation of the R-factor. The calculation of the R-factor is essential in the RUSLE model because it can reflect the impact of natural rainfall on soil erosion, which is especially sensitive to climate change. According to the USLE manual, the R-factor calculation requires the storm kinetic energy (E) and maximum 30-min rainfall intensity (I_{30}). A recent study uses hourly rainfall data to estimate the R-factor⁴⁵, which contains average rainfall erosivity data from 1951 to 2018 in China. However, it is difficult to obtain continuous, high-precision, and full-coverage rainfall records in long-term estimations of the R-factor on a national scale. Considering the feasibility, large-scale studies generally use monthly or annual rainfall data to calculate the R-factor; however, this method can bring many uncertainties and inaccuracies⁴⁶. Here, to better describe China's long-term annual rainfall erosivity, we made a trade-off and used a widely accepted daily rainfall erosivity model developed according to climate characteristics in China⁴⁷:

$$R = \sum_{j=1}^{24} R_j \quad (4)$$

$$R_j = \alpha \sum_{j=1}^k (P_j)^\beta \quad (5)$$

$$\beta = 0.8363 + \frac{18.77}{P_{d12}} + \frac{24.455}{P_{y12}} \quad (6)$$

$$\alpha = 21.586\beta^{-7.1892} \quad (7)$$

where R is the annual rainfall erosivity, and R_j is the half monthly rainfall erosivity. P_j is the daily erosive rainfall amount on the j -th day during the half-month (only select the days with $P_j \geq 12\ mm$, which is the threshold of a rainfall erosivity event in China⁴⁸). P_{d12} represents the average daily rainfall (mm) with a daily rainfall of 12 mm or more, and P_{y12} represents the average annual rainfall with a daily rainfall of 12 mm or more.

Data name	Spatial resolution	Format	Temporal period	Source
NASA SRTM digital elevation	30 m	grid	Static (2000)	Shuttle Radar Topography Mission digital elevation model (SRTM DEM) data
Rainfall data	Station	txt	Daily 1992–2019	China Meteorological Data Service Centre (CMDSC) (http://data.cma.cn/en)
Normalized difference vegetation index (NDVI)	8 km	tiff	Half-monthly 1992–2015	Global Agricultural Monitoring System (GIMMS) AVHRR NDVI3g datasets
	0.05°	tiff	Half-monthly 2016–2019	Terra MODIS NDVI data (MOD13C2 V6)
Land temperature	0.5°	tiff	Monthly 1992–2019	Climatic Research Unit (CRU TS)
Land cover	300 m	tiff	Annual 1992–2019	European Space Agency Climate Change Initiative Land Cover (ESA CCI LC)
Soli properties	250 m	tiff	Static	International Soil Reference and Information Centre (ISRIC) (http://data.isric.org/geonetwork/srv/chi/catalog.search?sessionId=A887E5B4)
Sown areas of major farm crops	Provincial	csv	Annual 1992–2019	National Bureau of Statistics of China (NBS) Database (https://data.stats.gov.cn)

Table 1. Input data.

Although the above model is the most widely used method in China, there are some improved models that are more applicable in some specific regions^{48,49}. Especially it overestimated rainfall erosion in the karst areas of southwest China due to the special geological background there⁴⁹. Therefore, we searched for a more suitable model for the karst region and finally selected a power law equation with a sinusoidal relationship reflecting seasonal variations of rainfall based on daily data⁵⁰. It has proven to be that the performance of this model is good in karst areas of China⁴⁹. The equation is as follows:

$$R_d = 0.2686 \left[1 + 0.5412 \cos \left(\frac{\pi}{6}j - \frac{7\pi}{6} \right) \right] P_d^{1.7263} \quad (8)$$

where R_d is the daily rainfall erosivity in the month j ($\text{MJ}\cdot\text{mm}\cdot\text{ha}^{-1}\cdot\text{h}^{-1}$), P_d is the daily rainfall (mm) in the day d , and the sum of the R_d in a year is the annual rainfall erosivity.

The source of the daily rainfall data and the distribution of meteorological stations are shown in Table 1 and Fig. 1a, respectively. Annual values of the R -factor were firstly calculated for each station, then the values were interpolated to erosivity maps employing the method of Universal Kriging, which has proven to be an effective method for the spatial interpolation of the rainfall erosivity in China^{44,51}.

Modification of the C-factor. The C -factor is closely related to the types of vegetation and crops²⁵; therefore, the values of the C -factor in arable and non-arable land in China were calculated separately. Some adjustments were made to the classification of crop types and the reference C -values of these crop types in arable land^{25,52} according to the actual agricultural conditions in China (Table 2). The released crops were classified into ten categories according to the main crop types in each province issued by the National Bureau of Statistics. The values of the C -factor in arable land (C_A) were calculated by the following equation:

$$C_A = \sum_{n=1}^{10} C_{cropn} \times Region_{Cropn} \quad (9)$$

where C_{cropn} represents the value of the C -factor of the crop type n , and $Region_{Cropn}$ represents the share of the sown areas of this crop type n in the total agricultural land in a province.

The values of the C -factor in non-arable land (C_{NA}) were estimated by vegetation coverage (based on the NDVI data) and reference values of the C -factor of various land cover types. The GIMMS NDVI products were the main source of the NDVI data from 1992 to 2015 for data integrity; however, they were unavailable after 2015. To form a temporally continuous SC dataset covering 1992–2019, we searched for other monthly NDVI data covering 2016–2019. We compared the MODIS NDVI with GIMMS NDVI with the temporal span overlapping from January 2010 to December 2015. We found that the temporal fluctuations of the monthly NDVI were consistent, and their correlation was obvious ($R^2 = 0.95$, $p < 0.01$) (Fig. 3). Therefore, the MODIS NDVI was selected to supplement the NDVI data after 2015. The monthly NDVI was then normalized using the following equation⁵²:

$$F_{cover} = \frac{NDVI_i - \min(NDVI_n)}{\max(NDVI_n) - \min(NDVI_n)} \quad (10)$$

where F_{cover} is the vegetation coverage, which is the monthly average value of the NDVI during the growing season (the months with a mean surface air temperature $\geq 0^\circ\text{C}$) (Fig. 1b)⁵³. $NDVI_i$ means the value of the NDVI in grid cell i , and $\max(NDVI_n)$ and $\min(NDVI_n)$ represent the maximum and minimum values of the NDVI in year n , respectively. The values of the C_{NA} were calculated by^{25,52}:

$$C_{NA} = \min(C_{NA_n}) + [\max(C_{NA_n}) - \min(C_{NA_n})] \times (1 - F_{cover}) \quad (11)$$

where C_{NA_n} is the reference values of the C -factor in land cover type n (Table 3) (Fig. 1c), and $\max(C_{NA_n})$ and $\min(C_{NA_n})$ are the maximum and minimum value of the C_{NA_n} , respectively.

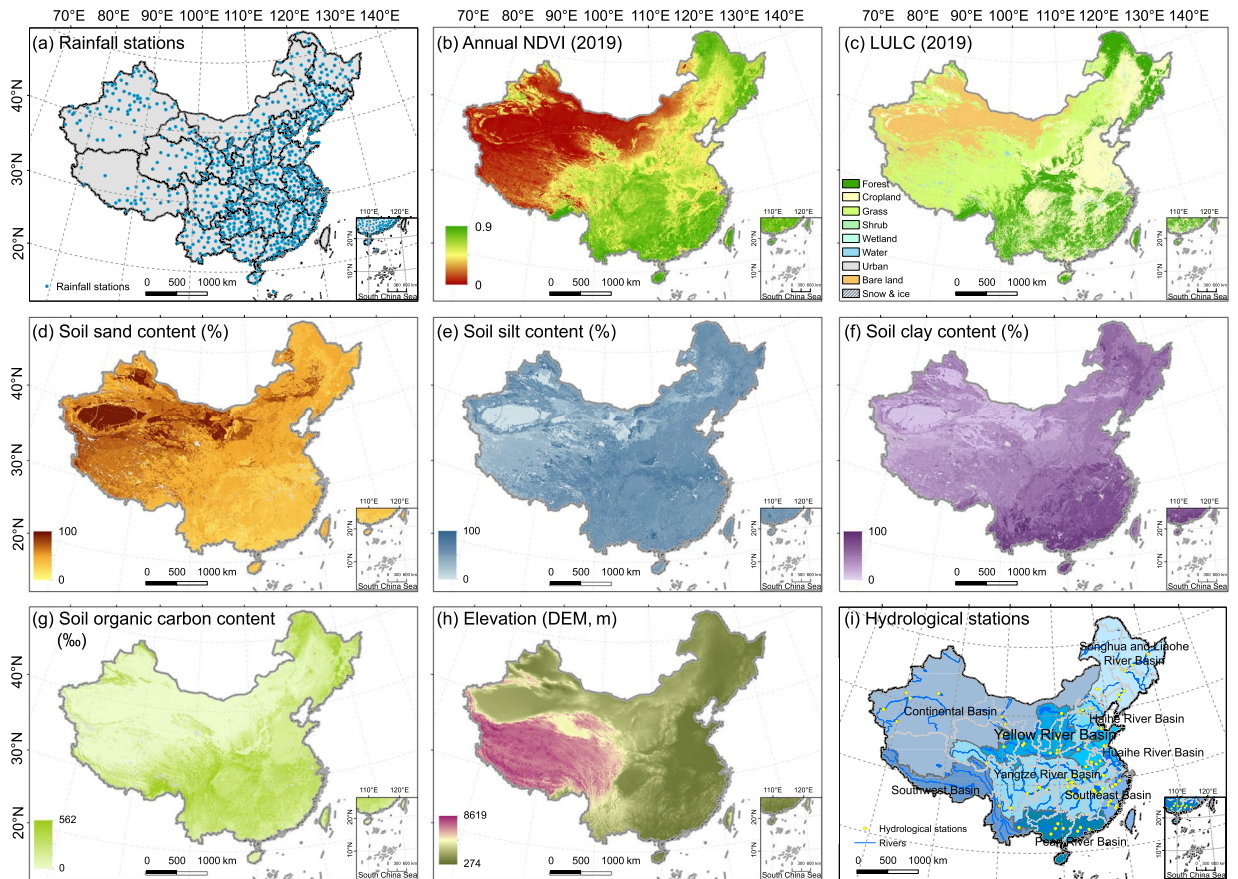


Fig. 1 Raw data to develop the soil conservation dataset. NDVI: Normalized difference vegetation index. LULC: Land use and land cover types.

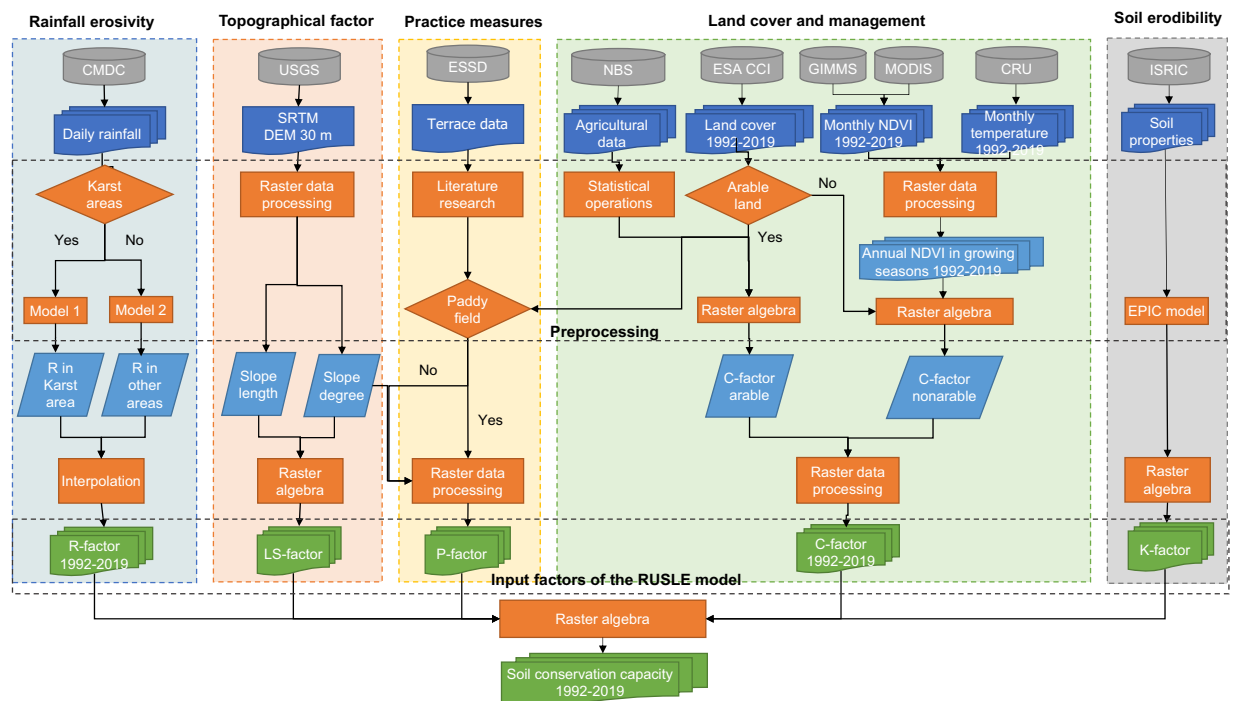


Fig. 2 Flow chart of the methods and data sources.

Farm crops		n	C_{crop}	
Grain crops	Cereal	Rice	1	0.15
		Corn	2	0.38
		Other cereals	3	0.2
	Beans	4	0.32	
Tubers		5	0.34	
Sugar crops	Sugarcane and Beetroots	6	0.15	
Oil-bearing crops		7	0.25	
Fiber crops	Cotton	8	0.4	
	Other fiber crops	9	0.28	
Tobacco		10	0.5	
Vegetables		11	0.25	
Medicinal materials		12	0.15	
Succulence		13	0.1	
Orchards	Grapes	14	0.35	
	Melons and fruits orchards	15	0.25	
	Other orchards	16	0.32	
Other farm crops		17	0.15	

Table 2. The values of the C -factor in cropland (C_{crop}).

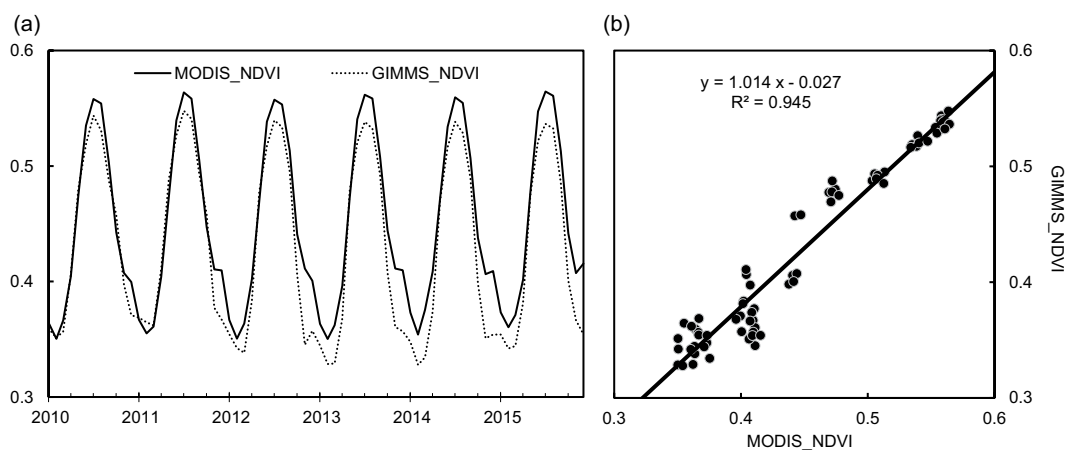


Fig. 3 Comparison of the NDVI from GIMMS with MODIS based on temporal fluctuations (a) and correlative relationship (b).

Improvement of the P -factor. The P -factor reflects the efficacy of support practices that humans made to control soil erosion with a range of 0 (no soil water erosion) to 1 (no support practices)^{21,22}. However, the P -factor is difficult to quantify in large-scale modelling of soil erosion⁵⁴. At present, the large-scale quantification of the P -factor is mainly based on literature analysis^{54–56}, which can only show statistical data without time changing. In this study, we have mapped and refined the distribution of the P -factor by an unprecedented method. We separated paddy and terraced fields from other land use types. Then we assigned different P -factor values according to different terrains based on a meta-analysis of P values of support practices⁵⁴. The horizontal paddy field was assigned a value of 0.2. P -factor values in other arable areas were assigned according to the slope (Table 4), which was only applied on the terraces because terraces are one of the most common and identifiable support practices in mountainous regions in China^{5,54}. For other non-arable lands, the value of the P -factor is 1 because support practices are mainly built on croplands in China⁵⁶.

Calculation of the LS - and K -factor. The LS -factor reflects the natural geographic influence of topography on soil erosion. The LS -factor consists of a slope length (L) factor and a slope steepness (S) factor, which are calculated by DEM data in GIS-based RUSLE model^{21,22}. The resolution of the DEM data influences the accuracy of the simulation of the LS -factor, as it has significant scaling effect^{57–59}. Considering the trade-off between the accuracy of simulations and feasibility of calculation, we used the DEM data with a resolution of $30\text{ m} \times 30\text{ m}$ (Fig. 1h). The formulas for calculating the L -factor are as follows^{21,60}:

n	Land cover type	C_{NAn} (maximum-minimum)
1	Broadleaf evergreen forest	0.0001–0.003
2	Broadleaf deciduous forest	0.0001–0.003
3	Needleleaf evergreen forest	0.0001–0.003
4	Needleleaf deciduous forest	0.0001–0.003
5	Mixed forest	0.0001–0.003
6	Tree open	0.01–0.15
7	Shrub	0.01–0.15
8	Herbaceous	0.01–0.15
9	Herbaceous with sparse tree/shrub	0.01–0.15
10	Sparse vegetation	0.1–0.5
11	Wetland	No data
12	Bare land	0.1–0.5
13	Urban	No data
14	Snow and ice	No data

Table 3. The values of the C-factor in non-cropland (C_{NAn}).

Slope gradient	The values of the P-factor
<4.6°	0.45
4.6° ~ 7°	0.52
7° ~ 9°	0.61
9° ~ 14°	0.70
14° ~ 24°	0.92
>24°	1

Table 4. The values of the P-factor for cropland on terraces.

$$L = \left(\frac{l_i}{22.13} \right)^m, \quad (12)$$

$$l_i = \sum_1^i \frac{D_i}{\cos\theta_i} - \sum_1^{i-1} \frac{D_i}{\cos\theta_i} = \frac{D_i}{\cos\theta_i}, \quad (13)$$

$$m = \begin{cases} 0.2, & \beta < 1\% \\ 0.3, & 1\% \leq \beta < 3\% \\ 0.4, & 3\% \leq \beta < 5\% \\ 0.5, & \beta \geq 5\% \end{cases}, \quad (14)$$

where m is the slope length exponent, l_i is the slope length of a grid i , D_i is the horizontal projection distance of the slope length of each grid along the runoff direction, θ_i is the angle of the slope of a grid i , and β is the slope gradients in %.

The S-factor was calculated by following the method in the CSLE model, which is improved according to the different slope degree in China⁶¹:

$$S = \begin{cases} 10.80 \sin\theta + 0.03, & \theta < 5^\circ \\ 16.80 \sin\theta - 0.5, & 5^\circ \leq \theta < 10^\circ \\ 21.91 \sin\theta - 0.96, & \theta \geq 10^\circ \end{cases} \quad (15)$$

where θ is the slope angle in degrees.

Soil erodibility (K -factor) reflects the soil's resistance to both detachment and transportation, which is originally measured by establishing unit plot for each soil type²¹. Sharpley and Williams⁶² establish regression equations between the measured data of plot experiments and soil properties in the EPIC model (Sharpley and Williams 1990), which make it feasible to calculate the K -factor on a large scale^{30,63,64}. The formula is as follows:

$$K = \left\{ 0.2 + 0.3 \exp \left[-0.0256 * SAN \left(1 - \frac{SIL}{100} \right) \right] \right\}$$

$$\left(\frac{SIL}{CLA + SIL} \right)^{0.3} \left(1 - \frac{0.25 \cdot OC}{OC + \exp(3.72 - 2.95 \cdot OC)} \right) \quad (16)$$

$$\left(1 - \frac{0.7SN1}{SN1 + \exp(-5.51 + 22.9SN1)} \right)$$

$$SN1 = 1 - \frac{SAN}{100} \quad (17)$$

where *SAN*, *SIL*, *CLA* and *OC* are the percentage content of sand, silt, clay and organic carbon from topsoil, respectively (Fig. 1d–g).

Change detection. A linear regression was employed to quantify long-term changes of the *R*-factor, *C*-factor, and *SC* from 1992 to 2019. The least squares were applied to calculate annual change rates, and the equation is as follows:

$$\theta_{slope} = \frac{n \times \sum_{i=1}^n i \times A_i - \sum_{i=1}^n i \sum_{i=1}^n A_i}{n \times \sum_{i=1}^n i^2 - (\sum_{i=1}^n i)^2}, \quad (18)$$

where θ_{slope} is the regression coefficient (slope of the linear regression), representing the annual changing rate of the data. A_i represents the data in the year of i .

Data Records

The dataset of soil conservation service preventing soil water erosion in China (1992–2019) is available at Science Data Bank (<https://cstr.cn/31253.11.sciencedb.07135>). Data <https://doi.org/10.57760/sciencedb.0713565>. This dataset includes nine zip files (“.rar”). All the data in the zip files are raster data (“.tif”). The details about every zip file were as follows.

- “C1992–2019.rar”: It contains the mean value and changing rate of the *C*-factor from 1992 to 2019, which were named “C_mean.tif” and “C_slope.tif”, respectively.
- “C_year.rar”: The detailed data of the *C*-factor in two-year increments from 1992 to 2019. The raster data were named “Cyyyy_300.tif”; “yyyy” is the year and “300” is the resolution of the data (300 m × 300 m resolution).
- “K_300.rar”: It contains the *K*-factor data, named “K_300.tif” (300 m × 300 m resolution).
- “LS_300.rar”: It contains the *LS*-factor data, named “LS_300.tif” (300 m × 300 m resolution).
- “P_300.rar”: It contains the *P*-factor data, named “P_300.tif” (300 m × 300 m resolution).
- “R1992–2019.rar”: The mean value and changing rate of the *R*-factor in China from 1992 to 2019, named “R_mean.tif” and “R_slope.tif”, respectively. The resolution of the data is 1 km × 1 km.
- “R_year.rar”: The *R*-factor data in two-year increments from 1992 to 2019 (1 km × 1 km resolution). The naming of the data is as “Ryyyy.tif”, and “yyyy” is the year.
- “SC1992–2019.rar”: The mean value and changing rate of soil conservation service (300 m × 300 m resolution).
- “SC_year.rar”: *SC* data in two-year increments from 1992 to 2019 (300 m × 300 m resolution).

Technical Validation

Since soil conservation service is calculated based on soil erosion rates estimated by the RUSLE model, its reliability can be indirectly validated by verifying simulated soil erosion^{39,42,43}. However, it is difficult to obtain abundant measured soil erosion rates from plot experiments for large-scale validation of modelled soil erosion, so we applied a cross-comparison of the assessments with substituted observations and other regional simulations of soil erosion⁵². Substituted observations refer to observed sediment runoff data for hydrologic stations in eight major river basins in China from 2010 to 2019 (Fig. 1i)⁶⁶, which were compared with modelled soil erosion rates by computing the correlation coefficients (R^2) between the mean values of them in every basin. High values of R^2 usually mean satisfactory estimations of soil erosion^{67–69}. We also compared assessments of soil erosion rates with other regional simulations using the USLE-based models, which can verify whether our results are in a reliable range. We also presented a correlation analysis between our estimated soil conservation service and other regional estimations that applying USLE-based models. The sources of collected studies are in the supplementary Tables S1, S2.

Furthermore, because the development of the RUSLE model is based on multiple regression of soil loss from the unit plot data, it was necessary to compare our results with the plot measurements. We used measured data representing 957 plot years in China from published literature to verify our results⁶⁹. The distribution of the collected plots is shown in Fig. 4. To compare the results, we grouped the measured data according to climate zones and land cover types. The land use types used included cropland, grassland, shrubland, and forest land, and the climate zones included Cfa, Cwa, Cwb, Dwa, and Dwb (the naming of the climate zones follows

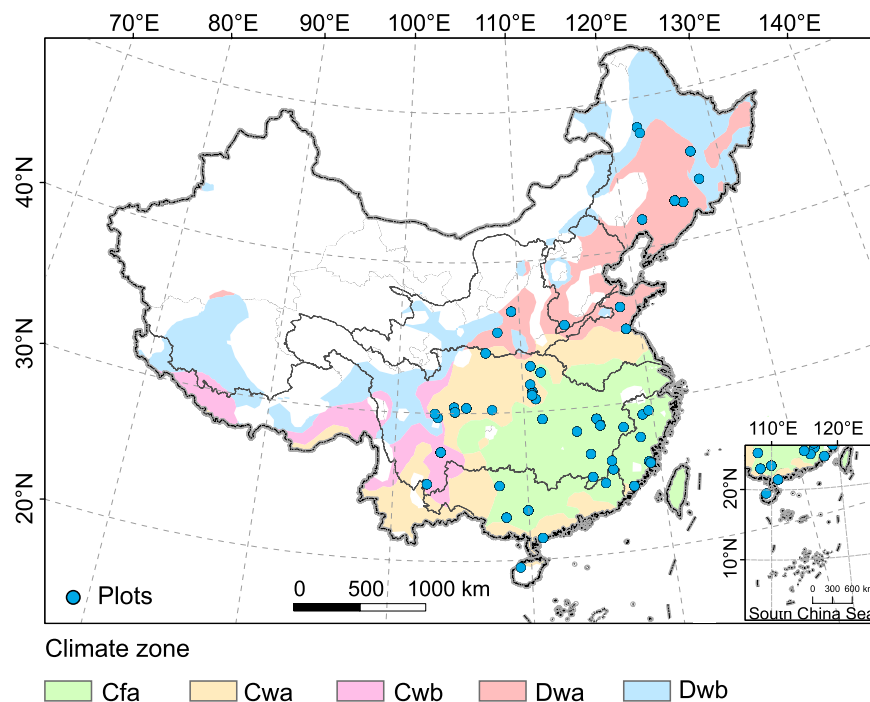


Fig. 4 Distribution of the collected plot data and climate zones in China. C: temperate, D: cold, f: without dry season, w: dry winter, a: hot summer, b: warm summer.

the Köppen-Geiger climate classification⁷⁰). Then the grouped measures data were compared with the average annual simulated data in these grouped areas. Because the principle of acquiring measured and simulated data was different, they are unsuitable for direct comparison. Therefore, if data distributions of the two types of results were similar, it could be considered that the simulated data can reflect the distribution difference of the actual data.

The comparison between the plot data and the modelled data (Fig. 5a) shows that the simulation of soil erosion can be accepted (all $R^2 \geq 0.5$). The simulation in the Pearl Basin has the highest correlation with the sediment runoff data ($R^2 = 0.91$), while the correlation is lowest in the Haihe River Basin ($R^2 = 0.50$). In addition, modelled soil erosion results in this study were compared with those in other studies representing seven basins in China (Fig. 5b and Table S1), which indicates that our simulations of soil erosion are within a reliable range. However, our modelled soil erosion rates are overestimated compared with those from collected literature in the Pearl River Basin. We also collected some SC data from other studies on a regional scale (Table S2). Results of comparison with other regional SC (Fig. 5c) indicates that our SC data in these regions are consistent with these results. The model uncertainty is mainly derived from different algorithms of the parameters and data sources⁷¹, which is also the reason for the differences between our results and other results.

The comparison between measured and simulated results indicated that modelled soil erosion was generally higher than measured soil erosion (Fig. 6). The simulation results in cropland were closer to the measured values than those in natural vegetation (grassland, forest, and shrubland). However, the gap between the simulated and the measured results of each land cover type in the Cwa climate zone was larger than in other climate zones. The main reason may be that the Cwa climate zone is mainly in regions with extreme rainfall and changeable meteorological conditions, such as China's southern and southwestern edges, where the measured data may be insufficient and cannot represent annual soil erosion well. Although there were some differences between the simulated and the measured results, the trend of data distribution was similar among different land cover types and climatic zones. The comparisons demonstrated that the dataset could reflect the difference in soil erosion and conservation among different climatic zones and land cover types in China. The prediction of the variation trend in soil erosion and conservation was close to the actual situation. These areas usually lack sufficient measured data for model-fitting research^{38,69}, so there was a need for more references for model correction.

Although these data have been improved by using relatively high-resolution input data and calibrating model factors, some limitations still remain. First, there are certain difficulties in the trade-off between the feasibility of calculation and the accuracy of the simulations in large-scale modelling. The use of the RUSLE model has a spatial scale effect on the precision of the input data⁷², which is one of the reasons that lead to the differences between large-scale model estimations (our study) and small-scale estimations (other studies). The finer input data can produce more reliable results. We will employ finer input data in further improvements; for example, simulations of the *R*-factor could be improved based on 30-min or finer rainfall data. Second, further exploration of factor calculation methods that are applicable to different regions is still needed. For example, the *R*-factor in high mountain areas (such as the Tibetan Plateau) can be improved using more suitable methods⁴⁸.

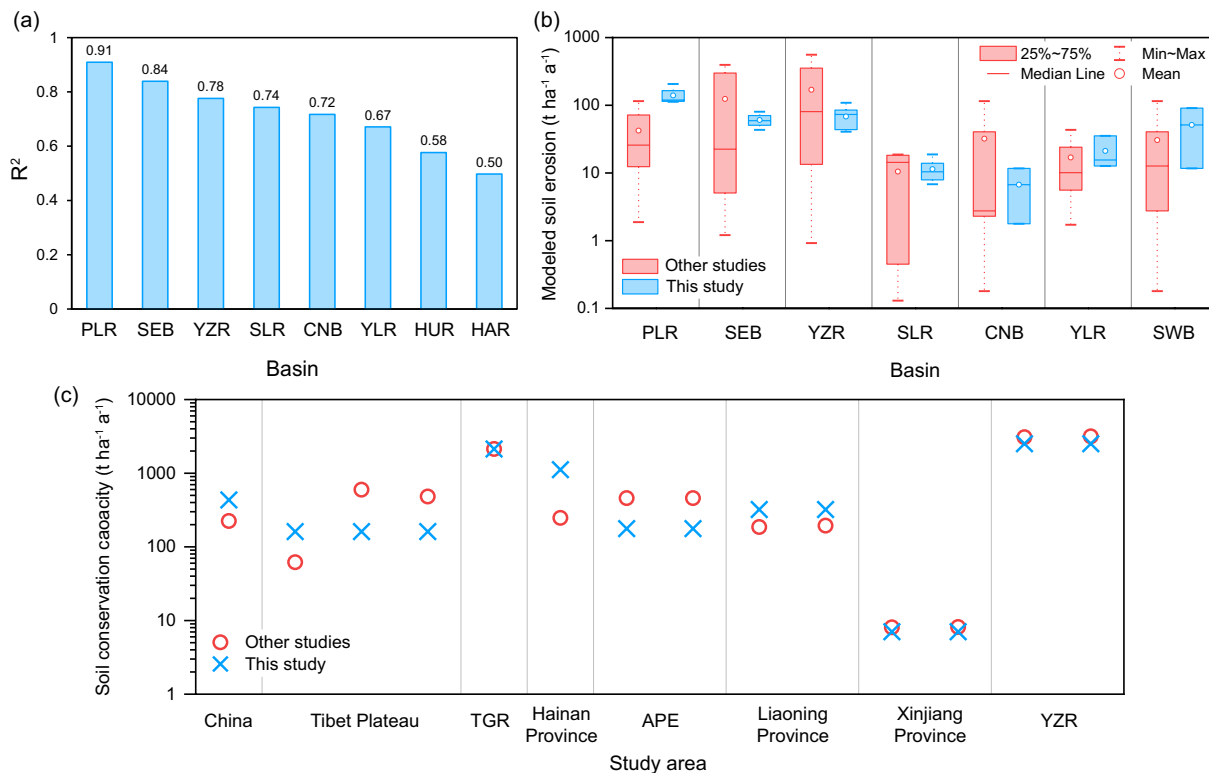


Fig. 5 (a) Correlation (R^2) between measurements of sediment runoff and simulations of soil erosion. (b) Comparison of simulated annual soil water erosion rates among different studies using USLE-based models. PLR: Pearl River Basin, SEB: Southeast Basin, YZR: Yangtze River Basin, SLR: Songhua and Liaohe River Basin, CNB: Continental Basin, YLR: Yellow River Basin, HAR: Haihe River Basin, HUR: Huaihe River Basin, SWB: Southwest Basin. (c) Comparison of soil conservation service from this study with those from other studies. TGR: Three Gorges Reservoir region, APE: Agro-pastoral ecotone of northern China.

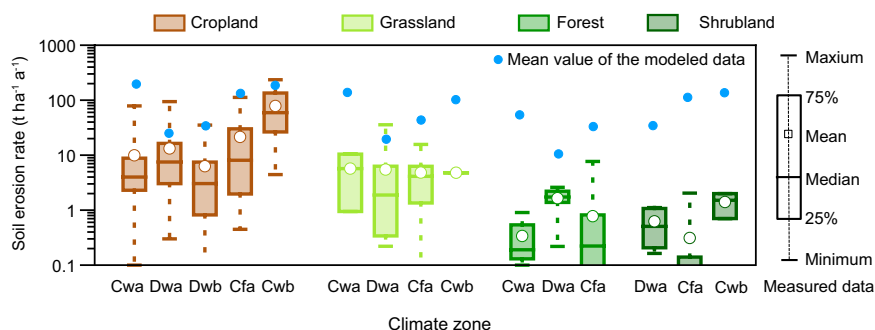


Fig. 6 Comparison between soil erosion rates simulated by the model and those measured on runoff plots.

Lastly, this dataset still needs further validations of every factor; for example, the K-factor estimation based on the EPIC model could be compared with measured data in some specific areas. However, the lack of sufficient measurement data for comparison disturbs the subsequent dataset verification.

Code availability

The computing progress was completed in the ArcGIS 10.6 software. No custom code was used.

Received: 22 March 2023; Accepted: 17 May 2023;
Published online: 26 May 2023

References

- Wang, Z. *et al.* Human-induced erosion has offset one-third of carbon emissions from land cover change. *Nat. Clim. Chang.* **7**, 345–350 (2017).
- FAO. Soil erosion: the greatest challenge to sustainable soil management. (Rome, 2019).
- Zhang, R. H., Li, R. & Zhang, L. L. Soil nutrient variability mediates the effects of erosion on soil microbial communities: results from a modified topsoil removal method in an agricultural field in Yunnan plateau, China. *Environ. Sci. Pollut. Res.* **29**, 3659–3671 (2022).
- Borrelli, P. *et al.* Lateral carbon transfer from erosion in noncroplands matters. *Glob. Chang. Biol.* **24**, 3283–3284 (2018).
- Cao, B. W. *et al.* A 30 m terrace mapping in China using Landsat 8 imagery and digital elevation model based on the Google Earth Engine. *Earth Syst. Sci. Data* **13**, 2437–2456 (2021).
- Chen, D., Wei, W. & Chen, L. D. Effects of terracing practices on water erosion control in China: A meta analysis. *Earth-Sci. Rev.* **173**, 109–121 (2017).
- SC. No.244 Document of State Council [1978]. (State Council of the People's Republic of China, 1978).
- SC. Regulations on Conversion of Farmland to Forests. Decree of the State Council of the People's Republic of China NO.367 (State Council of the People's Republic of China, 2002).
- Chen, C. *et al.* China and India lead in greening of the world through land-use management. *Nat. Sustain.* **2**, 122–129 (2019).
- SCIO. Food Security in China. (State Council Information Office of the People's Republic of China, 2019).
- Lu, Q. S., Xu, B., Liang, F. Y., Gao, Z. Q. & Ning, J. C. Influences of the Grain-for-Green project on grain security in southern China. *Ecol. Indic.* **34**, 616–622 (2013).
- Li, C. *et al.* The hidden risk in China's cropland conversion from the perspective of slope. *Catena* **206** (2021).
- Zhang, X. *et al.* A large but transient carbon sink from urbanization and rural depopulation in China. *Nat. Sustain.* (2022).
- Liu, X. P. *et al.* High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. *Nat. Sustain.* **3**, 564–570 (2020).
- Liu, X. H., Wang, J. F., Liu, M. L. & Meng, B. Spatial heterogeneity of the driving forces of cropland change in China. *Science in China Series D-Earth Sciences* **48**, 2231–2240 (2005).
- d'Amour, C. B. *et al.* Future urban land expansion and implications for global croplands. *Proc. Natl. Acad. Sci. USA* **114**, 8939–8944 (2017).
- Zhang, C., Dong, J. & Ge, Q. Mapping 20 years of irrigated croplands in China using MODIS and statistics and existing irrigation products. *Sci. Data* **9**, 407 (2022).
- Feng, X. M. *et al.* Revegetation in China's Loess Plateau is approaching sustainable water resource limits. *Nat. Clim. Chang.* **6**, 1019–1022 (2016).
- SC. Circular of the State Council on Conducting the Third National Soil Census. State Council Gazette Issue No.7 Serial No.1762 (State Council of the People's Republic of China, 2022).
- Fu, B. *et al.* Assessing the soil erosion control service of ecosystems change in the Loess Plateau of China. *Ecol. Complex.* **8**, 284–293 (2011).
- Wischmeier, W., Smith, D. D., Wischmer, W. H. & Smith, D. D. Predicting rainfall erosion losses: a guide to conservation planning. *U.S. Dep. Agric. Handb. No. 537* 1–69 (1978).
- Renard, K. G., Foster, G. R., Weesies, G. A. & Porter, J. P. RUSLE: revised universal soil loss equation. *J. Soil Water Conserv.* **46**, 30–33 (1991).
- Montgomery, D. R. Soil erosion and agricultural sustainability. *Proc. Natl. Acad. Sci. USA* **104**, 13268–13272 (2007).
- De Asis, A. M. & Omasa, K. Estimation of vegetation parameter for modeling soil erosion using linear Spectral Mixture Analysis of Landsat ETM data. *ISPRS-J. Photogramm. Remote Sens.* **62**, 309–324 (2007).
- Panagos, P. *et al.* Estimating the soil erosion cover-management factor at the European scale. *Land Use Pol.* **48**, 38–50 (2015).
- Naipal, V., Reick, C., Pongratz, J. & Van Oost, K. Improving the global applicability of the RUSLE model - adjustment of the topographical and rainfall erosivity factors. *Geosci. Model Dev.* **8**, 2893–2913 (2015).
- Panagos, P. *et al.* Modelling the effect of support practices (P-factor) on the reduction of soil erosion by water at European scale. *Environ. Sci. Policy* **51**, 23–34 (2015).
- Lufafa, A., Tenywa, M. M., Isabiry, M., Majaliwa, M. J. G. & Woome, P. L. Prediction of soil erosion in a Lake Victoria basin catchment using a GIS-based Universal Soil Loss model. *Agric. Syst.* **76**, 883–894 (2003).
- Barao, L. *et al.* Assessment of promising agricultural management practices. *Sci. Total Environ.* **649**, 610–619 (2019).
- Yang, D. W., Kanae, S., Oki, T., Koike, T. & Musiak, K. Global potential soil erosion with reference to land use and climate changes. *Hydrol. Process.* **17**, 2913–2928 (2003).
- Bamatuze, Y., Mukwaya, P., Oyama, S., Nadhomi, D. & Nsemire, P. Intersecting RUSLE modelled and farmers perceived soil erosion risk in the conservation domain on mountain Elgon in Uganda. *Appl. Geogr.* **126**, 2366–2366 (2021).
- Vallebona, C., Mantino, A. & Bonari, E. Exploring the potential of perennial crops in reducing soil erosion: A GIS-based scenario analysis in southern Tuscany, Italy. *Appl. Geogr.* **66**, 119–131 (2016).
- Borrelli, P., Ballabio, C., Yang, J. E., Robinson, D. A. & Panagos, P. GloSEM: High-resolution global estimates of present and future soil displacement in croplands by water erosion. *Sci. Data* **9**, 406 (2022).
- Hateffard, F. *et al.* CMIP5 climate projections and RUSLE-based soil erosion assessment in the central part of Iran. *Sci Rep* **11**, 7273 (2021).
- Sharp, R. *et al.* *INVEST 3.5.0.post502-h7855734e4db6 User's Guide*. The Natural Capital Project, Stanford University, University of Minnesota, The Nature Conservancy, World Wildlife Fund, (2018).
- FAO & ITPS. Status of the World's Soil Resources (SWSR) – Main Report. (Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, 2015).
- Batista, P. V. G., Davies, J., Silva, M. L. N. & Quinton, J. N. On the evaluation of soil erosion models: Are we doing enough? *Earth-Sci. Rev.* **197**, 17 (2019).
- Quine, T. A. & Oost, V. K. Insights into the future of soil erosion. *Proc. Natl. Acad. Sci. USA* **117**, 23205–23207 (2020).
- Ma, X., Zhu, J., Yan, W. & Zhao, C. Assessment of soil conservation services of four river basins in Central Asia under global warming scenarios. *Geoderma* **375**, 114533 (2020).
- Duan, X. *et al.* Effects of soil conservation measures on soil erosion in the Yunnan Plateau, southwest China. *Journal of Soil and Water Conservation* **75**, 131–142 (2020).
- Kong, L. *et al.* Evaluating indirect and direct effects of eco-restoration policy on soil conservation service in Yangtze River Basin. *Sci. Total Environ.* **631–632**, 887–894 (2018).
- Xiao, Q., Hu, D. & Xiao, Y. Assessing changes in soil conservation ecosystem services and causal factors in the Three Gorges Reservoir region of China. *J. Clean Prod.* **163**, S172–S180 (2017).
- Rao, E., Ouyang, Z., Yu, X. & Xiao, Y. Spatial patterns and impacts of soil conservation service in China. *Geomorphology* **207**, 64–70 (2014).
- Teng, H. F., Hu, J., Zhou, Y., Zhou, L. Q. & Shi, Z. Modelling and mapping soil erosion potential in China. *J. Integr. Agric.* **18**, 251–264 (2019).
- Yue, T., Yin, S., Xie, Y., Yu, B. & Liu, B. Rainfall erosivity mapping over mainland China based on high-density hourly rainfall records. *Earth Syst. Sci. Data* **14**, 665–682 (2022).
- Li, J. L., Sun, R. H. & Chen, L. D. Assessing the accuracy of large-scale rainfall erosivity estimation based on climate zones and rainfall patterns. *Catena* **217**, 106508 (2022).
- Zhang, W., Xie, Y. & Liu, B. Research on the Method of Calculating Rainfall Erosivity Using Daily Rainfall. *SIENTA Geogr. Sin.* **22**, 705–711 (2002).

48. Wang, Y. S. *et al.* Increasing trends in rainfall-runoff erosivity in the Source Region of the Three Rivers, 1961–2012. *Science of the Total Environment* **592**, 639–648 (2017).
49. Zhu, D. Y., Xiong, K. N. & Xiao, H. Multi-time scale variability of rainfall erosivity and erosivity density in the karst region of southern China, 1960–2017. *Catena* **197**, 15 (2021).
50. Xie, Y., Yin, S. Q., Liu, B. Y., Nearing, M. A. & Zhao, Y. Models for estimating daily rainfall erosivity in China. *J. Hydrol.* **535**, 547–558 (2016).
51. Li, L., Jiang, X. & Sun, Y. Geostatistics-based spatial interpolation method for study of rainfall erosivity—a case study of Jiangsu. *Journal of Ecology and Rural Environment* **27**, 88–92 (2011).
52. Borrelli, P. *et al.* An assessment of the global impact of 21st century land use change on soil erosion. *Nat. Commun.* **8**, 1–13 (2017).
53. Pan, N. Q. *et al.* Increasing global vegetation browning hidden in overall vegetation greening: Insights from time-varying trends. *Remote Sens. Environ.* **214**, 59–72 (2018).
54. Xiong, M., Sun, R. & Chen, L. Global analysis of support practices in USLE-based soil erosion modeling. *Prog. Phys. Geogr.* **43**, 391–409 (2019).
55. Jia, L. Z. *et al.* Regional differences in the soil and water conservation efficiency of conservation tillage in China. *Catena* **175**, 18–26 (2019).
56. Zhao, J. L., Yang, Z. Q. & Govers, G. Soil and water conservation measures reduce soil and water losses in China but not down to background levels: Evidence from erosion plot data. *Geoderma* **337**, 729–741 (2019).
57. Wang, C. M. *et al.* Impacts of horizontal resolution and downscaling on the USLE LS factor for different terrains. *Int. Soil Water Conserv. Res.* **8**, 363–372 (2020).
58. Panagos, P., Borrelli, P. & Meusburger, K. A New European Slope Length and Steepness Factor (LS-Factor) for Modeling Soil Erosion by Water. *Geosciences* **5**, 117–126 (2015).
59. Lu, S. J. *et al.* Soil erosion topographic factor (LS): Accuracy calculated from different data sources. *Catena* **187** (2020).
60. Liu, B. Y., Nearing, M. A., Shi, P. J. & Jia, Z. W. Slope Length Effects on Soil Loss for Steep Slopes. *Soil Sci. Soc. Am. J.* **64**, 1759–1763 (2000).
61. Liu, B., Zhang, K. & Xie, Y. in *12th ISCO Conference* (2002).
62. Sharp, A. N. & Williams, J. R. *EPIC-erosion/productivity impact calculator: 1. Model determination*. US Department of Agriculture. (1990).
63. Scherer, L. & Pfister, S. Modelling spatially explicit impacts from phosphorus emissions in agriculture. *Int. J. Life Cycle Assess.* **20**, 785–795 (2015).
64. Pham, T. N., Yang, D., Kanae, S., Oki, T. & Musiak, K. Application of RUSLE Model on Global Soil Erosion Estimate. *J. Hydraul. Eng.-ASCE* **45**, 811–816 (2001).
65. Li, J., He, H., Zeng, Q., Chen, L. & Sun, R. Dataset of Soil Conservation Capacity Preventing Water Erosion in China (1992–2019). *Science Data Bank*, <https://doi.org/10.57760/sciencedb.07135> (2023).
66. MWR. *Bulletin of River Sediment in China 2010–2015*. China Water & Power Press, (2015).
67. Wang, H., Zhao, W., Li, C. & Pereira, P. Vegetation greening partly offsets the water erosion risk in China from 1999 to 2018. *Geoderma* **401**, 115319 (2021).
68. Jin, F. M., Yang, W. C., Fu, J. X. & Li, Z. Effects of vegetation and climate on the changes of soil erosion in the Loess Plateau of China. *Sci. Total Environ.* **773** (2021).
69. Xiong, M., Sun, R. & Chen, L. A global comparison of soil erosion associated with land use and climate type. *Geoderma* **343**, 31–39 (2019).
70. Beck, H. E. *et al.* Present and future Köppen–Geiger climate classification maps at 1-km resolution. *Sci. Data* **5**, 12 (2018).
71. Saltelli, A. *et al.* Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environ. Modell. Softw.* **114**, 29–39 (2019).
72. Alewell, C., Borrelli, P., Meusburger, K. & Panagos, P. Using the USLE: Chances, challenges and limitations of soil erosion modelling. *Int. Soil Water Conserv. Res.* **7**, 203–225 (2019).

Acknowledgements

This work was funded by the National Natural Science Foundation of China (U2102209).

Author contributions

J.L. designed the algorithm and wrote the article. H.H., Q.Z., and J.L. performed the simulations. L.C. offered critical review. R.S. formulated overarching research goals and aims.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-02246-4>.

Correspondence and requests for materials should be addressed to R.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023