



OPEN

DATA DESCRIPTOR

GloSEM: High-resolution global estimates of present and future soil displacement in croplands by water erosion

Pasquale Borrelli ^{1,2,3} , Cristiano Ballabio ⁴, Jae E. Yang³, David A. Robinson ⁵ & Panos Panagos ⁴

Healthy soil is the foundation underpinning global agriculture and food security. Soil erosion is currently the most serious threat to soil health, leading to yield decline, ecosystem degradation and economic impacts. Here, we provide high-resolution (ca. 100 × 100 m) global estimates of soil displacement by water erosion obtained using the Revised-Universal-Soil-Loss-Equation-based Global Soil Erosion Modelling (GloSEM) platform under present (2019) and future (2070) climate scenarios (i.e. Shared Socioeconomic Pathway [SSP]1–Representative Concentration Pathway [RCP]2.6, SSP2–RCP4.5 and SSP5–RCP8.5). GloSEM is the first global modelling platform to take into account regional farming systems, the mitigation effects of conservation agriculture (CA), and climate change projections. We provide a set of data, maps and descriptive statistics to support researchers and decision-makers in exploring the extent and geography of soil erosion, identifying probable hotspots, and exploring (with stakeholders) appropriate actions for mitigating impacts. In this regard, we have also provided an Excel spreadsheet that can provide useful insights into the potential mitigating effects of present and future alternative CA scenarios at the country level.

Background & Summary

Ploughing activities and unsuitable agricultural practices are the primary causes of accelerated rates of soil displacement (or soil loss¹) by water erosion, which is recognised as having detrimental effects on agricultural production and reducing ecosystem functioning². Because we live on a cultivated planet, where agriculture covers ~38% of the Earth's ice-free land³, accelerated water erosion represents a major socioeconomic and environmental threat through its several on-site and off-site effects, such as pollution in our waterways, dam siltation, eutrophication phenomena and the contamination of coastal and marine ecosystems^{4,5}. In the context of soil being globally poorly governed⁶, with conservation farming still playing a marginal role (estimated to cover 9% to 15% of croplands worldwide, equal to ca. 122–215 Mha⁷), seeking to gain a better understanding of global soil erosion dynamics is important in order to: (i) help decision-makers in developing new interventions and farming practices that are better for the environment and reduce the global footprint of agriculture on ecosystems⁸; and (ii) strengthen knowledge-sharing with earth-system modellers working on the soil erosion implications of sediment, nutrient and carbon cycling⁹.

The US Department of Agriculture (USDA) developed the Revised Universal Soil-Loss Equation (RUSLE)¹⁰, a deterministic and empirically based prediction approach that developed from a statistical analysis of more than 10,000 plot-years of basic runoff and soil displacement data covering a large variety of North American landscape conditions. Today, RUSLE-based algorithms have been applied in over 100 countries worldwide and are by far the most widely applied soil-erosion prediction models, globally (~41% of the total, according to the global review and statistical analysis of Borrelli *et al.*¹¹). The estimates from Global Soil Erosion Modelling

¹Department of Science, Roma Tre University, 00146, Rome, Italy. ²Department of Earth and Environmental Sciences, University of Pavia, 27100, Pavia, Italy. ³Department of Biological Environment, Kangwon National University, Chuncheon, 24341, Republic of Korea. ⁴European Commission, Joint Research Centre (JRC), Ispra, Italy. ⁵UK Centre for Ecology and Hydrology, Environment Centre Wales, Bangor, LL57 2UW, United Kingdom. e-mail: pasquale.borrelli@uniroma3.it; panos.panagos@ec.europa.eu

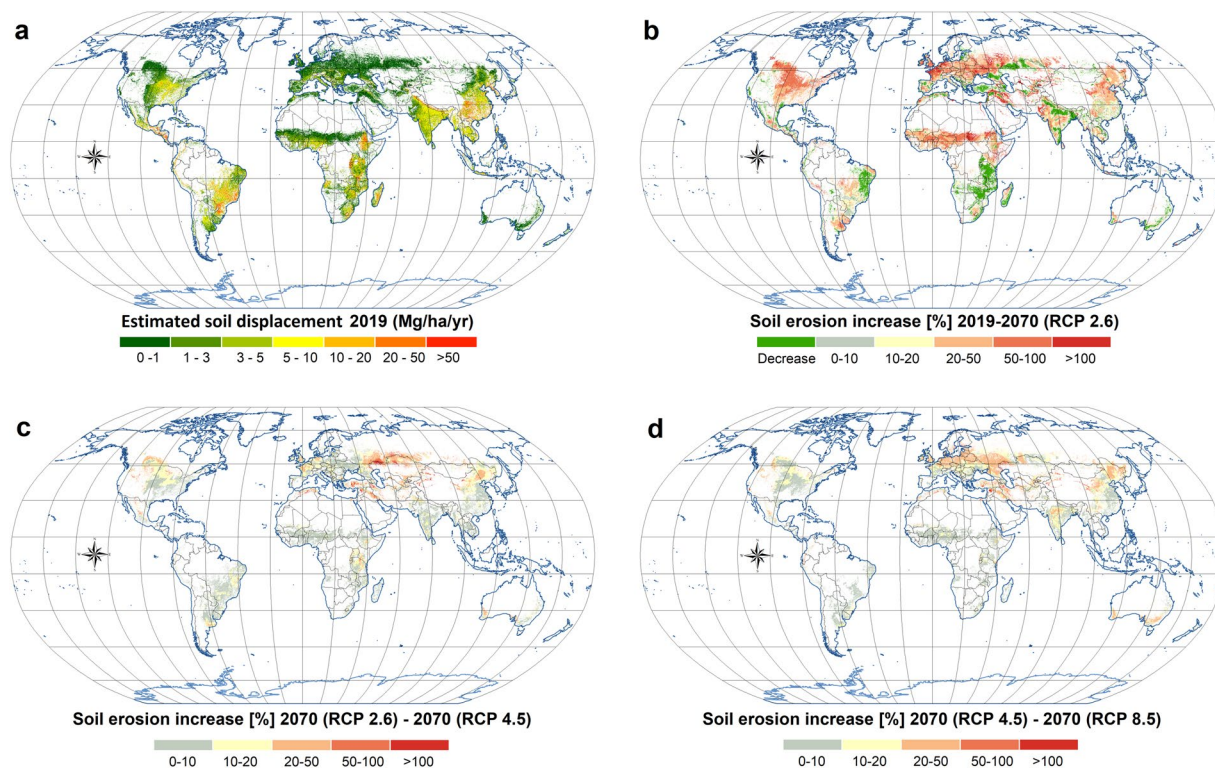


Fig. 1 Soil erosion estimates predicted through GloSEM 1.3 for global croplands: (a) soil erosion rates divided into seven classes, according to the European Soil Bureau classification; (b–d) changes in annual average soil erosion between 2019 and 2070 for three distinct RCP greenhouse-gas trajectories. The changes exclusively refer to effects from the climate change projections. For these simulations, the 2019 croplands were used. (b–d) share the same legend.

(GloSEM) rest on the RUSLE prediction approach implemented in a GIS environment by means of statistical operations and advanced spatial interpolation techniques. We recognise that using an empirically based prediction tool, underpinned by data-driven assumptions, in areas outside the range of the original estimates (e.g. tropical, sub-Arctic and tundra regions) may considerably reduce the local accuracy of the model. However, considering the proven capacity of RUSLE-based models to overcome their empirical origins¹², the current lack of better-performing models^{13,14}, and the need to be able to predict the possible impacts of global change on soil erosion, we argue that this physically plausible empirical method for predicting soil erosion through water erosion represents a legitimate approach to narrowing the current gap in our knowledge and supporting targeted soil-conservation efforts aimed at mitigating soil erosion. Studies detailed in the literature describe the methodological development of the GloSEM framework under present¹⁵ and future⁹ land-use and climate conditions.

However, GloSEM is not the only modelling approach presented in the literature that quantitatively estimates soil displacement by water erosion on a global scale. Several global modelling approaches have been developed, all of them using RUSLE-based algorithms^{16–23}. While these approaches differ in terms of degree of complexity, they all tend to have lumped and coarse-resolution (ca. 10–60-km cell size) modelling schemes that lack the ability to statistically represent the conditions of local farming systems and the mitigation potential that can derive from conservation agriculture (CA) (further details are given in Borrelli *et al.*¹⁵). This limits the suitability of most water-erosion prediction models to address policy questions, and to support policy-makers in exploring the geographical extent of the problem, deal with hotspots, and work with stakeholders to mitigate on-site and off-site impacts.

The outcomes stemming from GloSEM have been used by researchers in adjacent disciplines to map the world's free-flowing rivers²⁴, quantify the negative cascading effects of soil erosion on nutrient loss²⁵, create a link between soil erosion and economic models²⁶, measure land degradation^{27,28}, and consider alternative strategies for coral-reef conservation^{29,30}, among other things. In response to a rising demand for GloSEM data-sharing, we are hereby releasing the latest GloSEM 1.3 estimates for global croplands. We present a gridded geo-dataset (GeoTIFF format) of soil displacement by water erosion, expressed as a mass of soil per unit area and time ($\text{Mg ha}^{-1} \text{ yr}^{-1}$), obtained by running GloSEM coupled with the cropland map from the global land-cover layers released by the European Space Agency (ESA) and the Copernicus Global Land Service land cover (CGLS-LC100–100 spatial resolution) Collection 3 product based on PROBA-V and Sentinel-2 satellite imagery³¹. Present (2019) (Fig. 1a) and future (2070) (Fig. 1b–d) soil displacement under different climate scenarios (Shared Socioeconomic Pathway [SSP]1–Representative Concentration Pathway [RCP]2.6, SSP2–RCP4.5 and SSP5–RCP8.5) are being released on a 100×100 -m grid-cell basis, covering the croplands of 199 countries (~1.4 billion ha). The descriptive statistics are provided in Table 1 and S1, while Fig. 2 illustrates the present and

	Area	Total displacement				Displacement rate			
		2019	RCP 2.6	RCP 4.5	RCP 8.5	2019	RCP 2.6	RCP 4.5	RCP 8.5
	Million ha	Pg yr ⁻¹				Mg ha ⁻¹ yr ⁻¹			
Africa	248.2	4.3	4.9	5.1	5.2	17.1	19.7	20.5	20.9
Asia	500.9	6.4	7.7	8.0	8.5	12.8	15.3	16.0	17.1
Europe	251.5	0.7	0.9	0.9	1.0	2.6	3.5	3.7	4.1
North America	205.2	2.3	2.9	3.0	3.1	11.0	14.3	14.7	15.2
Oceania	34.2	0.1	0.1	0.1	0.1	2.7	3.1	3.2	3.5
South America	158.0	3.5	3.9	4.0	4.1	22.0	24.5	25.6	25.7
Total	1398.0	17.2	20.4	21.2	22.1	12.3	14.6	15.2	15.8

Table 1. Descriptive statistics of continental and global soil displacement estimates for the 2019 and 2070 scenarios.

future rainfall erosivity dynamics. Figure 3 illustrates the detailed modelling results for three areas in Italy, US and China.

Methods

The GloSEM platform follows the same principles of most RUSLE-type models or more-complex catchment-scale process-based models, with a driving force (erosivity of the climate [R-factor], expressed in MJ mm h⁻¹ ha⁻¹ yr⁻¹), a resistance term (erodability of the soil [K-factor], expressed in Mg h MJ⁻¹ mm⁻¹), and other factors representing the farming choice, such as topographical conformation of the field (dimensionless LS-factor), cropping system (dimensionless C-factor) and soil conservation practices (dimensionless P-factor). The long-term annual average soil displacement (Mg ha⁻¹ yr⁻¹) estimates rest on a multiplicative equation (Eq. [1]) of these environmental factors:

$$A = R \cdot L \cdot S \cdot K \cdot C \cdot P \quad (1)$$

Using a GIS-based gridded modelling approach applied to a RUSLE-type algorithm, such as GloSEM, corresponds to the hypothesis that each cell is independent of the others with respect to soil displacement. The provided estimates of soil displacement by water erosion refer to the amount of sediment displaced within each cell. The model does not, in any way, include areas of slope that experience net deposition over the long term¹. This is because the displaced soil amount is not routed downslope across each cell, from hill slope to sink area or riverine system, through a transport/deposition capacity module. This RUSLE-based soil displacement prediction scheme is preferred to other process-based physical models because the latter require a large number of input data that are not yet mature enough for global-scale applications.

Model parametrisation. The six environmental factors described in Eq. (1) were computed as follows.

Land-cover and management factor. The (dimensionless) C-factor measures the combined effect of the interrelated soil-cover and management variables on the soil erosion process. Here, for its spatial computation, we followed the pathway developed in previous GloSEM applications^{9,32}. However, in this case, the built-in GloSEM probabilistic land-use-allocation module was substituted with an external cropland layer (reference year 2019)—the CGLS-LC100 Collection 3, developed by the ESA and the CGLS³¹. The CGLS-LC100 Collection 3 comprises a set of annual global land-cover maps (2015–2019), with an overall 80% accuracy, obtained using PROBA-V and Sentinel-2 satellite imagery, and more than 150,000 crowd-sourced training points. The decision to use this new ESA product to represent the global cropland in this new version of GloSEM 1.3 rested on: (i) the high mapping accuracy of the new ESA product; (ii) the high level of agreement with the FAOSTAT cropland information at the global and continental scale; and (iii) its capacity to describe cropland-cover fractions (%) and, in turn, define spatial locations where cropland is mixed with pasture and natural vegetation. The original approach of GloSEM to describe crop rotations was not changed, and it still relies on the 12 years (2001–2012) of crop yields obtained from the FAOSTAT database (<https://www.fao.org/faostat/en/#data>) for each considered country. The cropland C-factor (C_{CROP}) values were computed at the sub-national level for each of 3,252 administrative regions, as follows:

$$C_{CROP} = \sum_{n=1}^{13} C_{CROPn} \cdot [\%] \text{Region}_{CROPn} \quad (2)$$

Where, C_{CROP} represents the overall C-factor estimated for each sub-country administrative unit layers (FAO GAUL), C_{CROPn} is the individual C-factor for each of the 170 crops considered (FAOSTAT), and $[\%] \text{Region}_{CROPn}$ represents the share of each crop in the arable land of the given FAO GAUL administrative unit. The $[\%] \text{Region}_{CROPn}$ values were assessed through a statistical downscaling of the national crop statistics using the spatially explicit harvested areas proposed by Monfreda *et al.*³³, combining national-, state- and county-level census statistics with remote-sensing data (for 15,990 administrative units)³⁴. The global cropland obtained from the global land-cover map for the reference year 2019 was also applied to the 2070 scenarios. Accordingly, the extent of the global cropland in both temporal scenarios overlapped precisely and consistently, and was equal to ~1.4 billion ha. The regional C_{CROP} values were also kept the same for both temporal scenarios. However, in

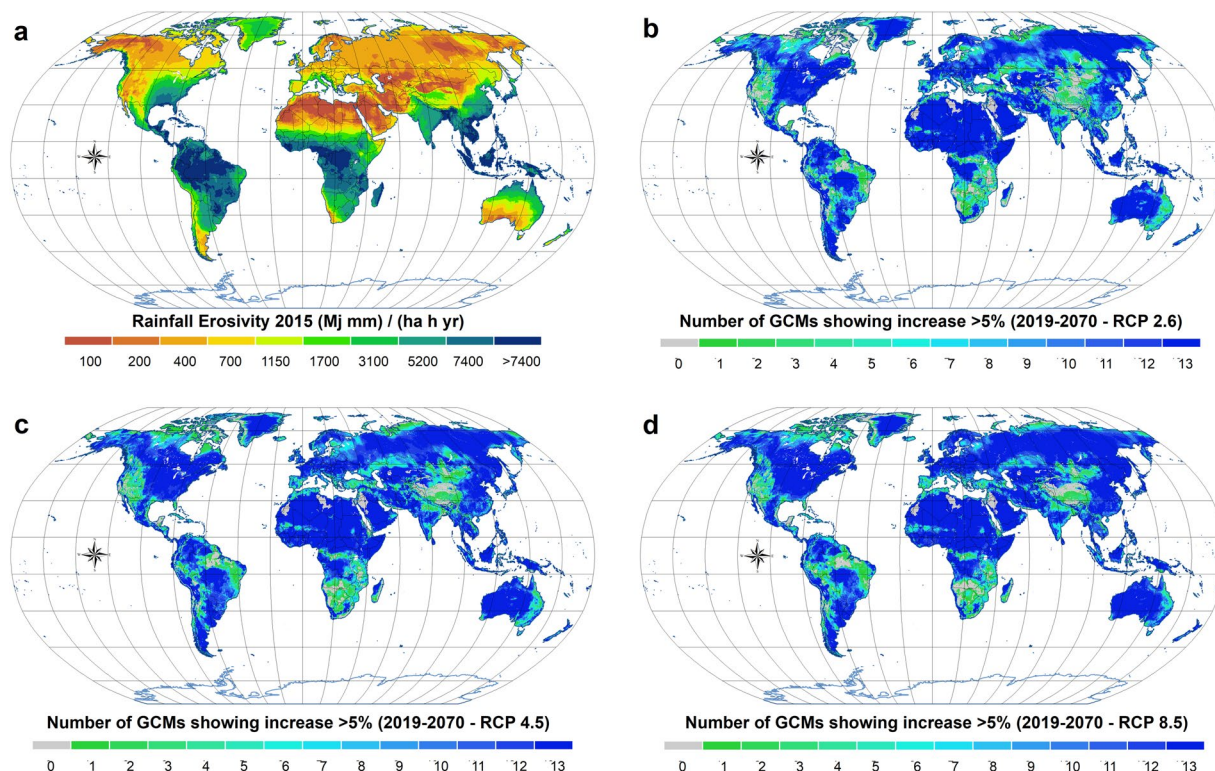


Fig. 2 Global rainfall erosivity: (a) erosivity classes subdivided according to quantiles; (b–d) number of GCMs showing changes greater than 5% between 2019 and 2070.

the Excel spreadsheet we provide further data concerning the country-scale extension of cropland, and estimates computed according to the future land-use/cover scenarios from the integrated assessment model given in the Land-Use Harmonization (LUH2) database³⁵—that is, SSP1–RCP2.6 derived from IMAGE, SSP2–RCP4.5 derived from MESSAGE-GLOBIOM and SSP5–RCP8.5 derived from REMIND-MAGPIE. While in the present study future land-use change is not considered, readers can use the Excel spreadsheet (Appendix A) to perform country-scale estimates of soil displacement under alternative scenarios (Table S1). In the future, we plan to release refined GloSEM input data with the purpose of allowing readers to perform their own spatially explicit estimates.

Rainfall-runoff erosivity. The average annual rainfall-runoff erosivity (R-factor, $\text{MJ mm h}^{-1} \text{ha}^{-1} \text{yr}^{-1}$) is spatially defined using a Gaussian process regression (GPR) geostatistical model for both present³⁶ and future⁹ climate conditions. The GPR model established a statistical relationship between the R-factor point values (from 3,625 meteorological stations, associated with the Global Rainfall Erosivity Database [GloREDa], in 63 countries that provide sub-hourly [61%] and hourly [39%] pluviographic data) and a set of WorldClim³⁷ climatic data (version 2.1, 30 seconds spatial resolution), acting as spatially exhaustive covariates. The R-factor values for every erosive rainstorm³⁶ were calculated adopting the methodology suggested in *USDA Handbook No. 703*¹⁰, while the USDA's Rainfall Intensity Summarization Tool was used to facilitate the computation phase of the individual erosive rainstorms. The R-factor values ($\text{MJ mm ha}^{-1} \text{h}^{-1} \text{yr}^{-1}$) from each of the 3,625 meteorological stations resulted from the following equation:

$$R = \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^{mj} (EI_{30})_k \quad (3)$$

Where, EI_{30} is the rainfall-runoff erosivity of a single event, k ; n expresses the number of years observed; and m_j expresses the erosive events during a given year, j . The WorldClim climatic data, acting as spatially exhaustive covariates in both the present and future climate scenarios, were: (i) average monthly precipitation; (ii) average minimum and maximum monthly precipitation; (iii) average monthly temperature; (iv) precipitation in the wettest month; (v) precipitation in the driest month; and (vi) precipitation seasonality. It is worth mentioning that, for the future (2070) scenarios, the R-factor values were estimated using the climate projections from 14 general circulation models (GCMs) and considering three different SSPs and RCPs (i.e. SSP1–RCP2.6, SSP1–RCP4.5 and SSP5–RCP8.5). This yielded a total of 43 possible future scenarios.

Soil erodability. The K-factor ($\text{Mg h MJ}^{-1} \text{mm}^{-1}$) spatially defines the susceptibility of the soil to be eroded. Here, it was algebraically defined (Eq. [4]), based on certain intrinsic soil properties, according to the methodology proposed by Wischmeier *et al.*³⁸. Some of the required soil properties (i.e. texture, organic matter and coarse

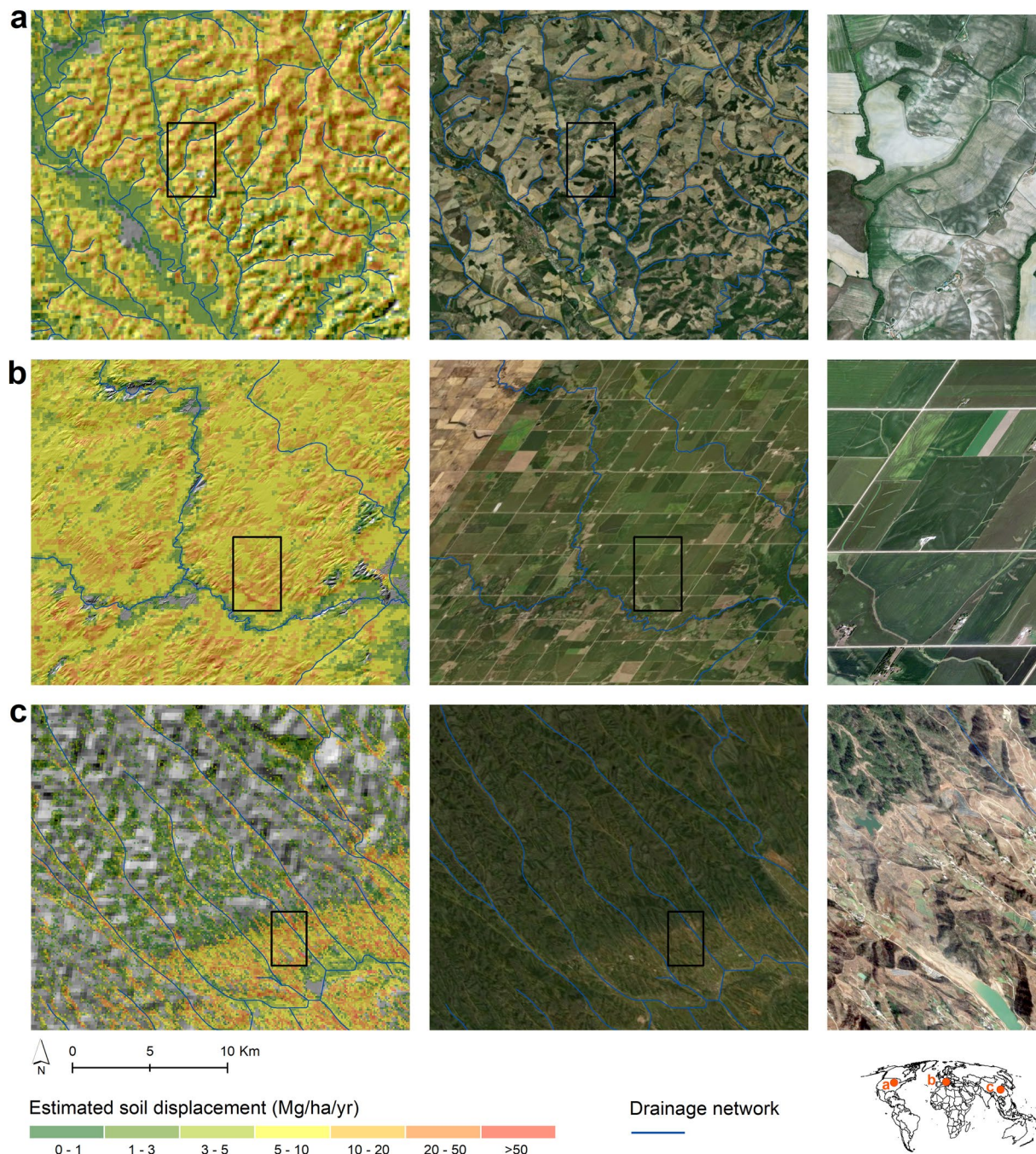


Fig. 3 GloSEM estimates in three locations showing signs of susceptible to soil erosion by water in (a) Italy (Asciano, Tuscany – 11.49E; 43.25 N), (b) USA (Albion, Iowa – 93.09 W; 42.11 N) and (c) China (Fangxian, Hubei – 110.71E; 32.13 N). The figures on the left column report the estimates of soil displacement superimposed on a hillshade. The figures on the central column show aerial images (Google Image) for the same locations already reported in the figures on the left column. The figures on the right column are a magnification of the panels reported in black.

fragmentation) were obtained from the International Soil Reference and Information Centre SoilGrids database at a 1-km spatial resolution³⁹. Further soil properties, such as soil structure and permeability, were derived according to the methodology proposed for the GloSEM soil erodability map³², which combines textural characteristics (M —percentage of the silt-plus-fine-sand fraction, multiplied by 100, minus the clay fraction), organic matter (OM, in %), soil structure (s) and permeability class (p):

$$K = \frac{(2.1 \times 10^{-4} M^{1.14} (12 - OM) + 3.25(s - 2) + 2.5(p - 3))}{100} \cdot 0.137 \quad (4)$$

Slope length and steepness factor. The (dimensionless) LS-factor represents the influence of the terrain on the surface runoff and sediment transport capacity. It was computed using the two-dimensional GIS-based approach proposed by Desmet and Govers⁴⁰. The slope and upslope contributing area were calculated using hole-filled Shuttle Radar Topography Mission (SRTM) and Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) v2 data with a 3-arc-seconds spatial resolution (ca. 90 m). Due to the large amount of processed data, the computational operations were carried out using the Saginaw Area (SAGA) GIS interface in R statistical computing software. We created 870 tiles of ca. 500 × 500 km.

Conservation support-practice factor. The (dimensionless) P-factor expresses the ratio of soil displacement using a conservation support practice, such as contouring, strip-cropping, terracing or subsurface drainage. Values for the support-practice P-factor are difficult to report above the field-scale due to the lack of spatial information, and they are not taken into account in the vast majority of regional- and global-scale assessments. In GloSEM^{9,32}, the possible mitigation effects of soil conservation rely on the CA data provided by 54 countries to the UN Food and Agriculture Organization's (FAO's) AQUASTAT database. These countries regularly communicate the proportion of their cropland managed in accordance with the three FAO CA standards—minimum soil disturbance, organic soil cover and crop rotation/association. Overall, these countries cover 73% of the global cropland. For the remaining 27%, the continental-level CA was considered. For areas under CA, we assumed a reduction in soil erosion ranging between 45% and 70% compared to conventional agriculture. Customised conservation scenarios can be simulated using the provided Excel spreadsheet.

Heterogeneity of input data. GloSEM versions 1.1 and 1.2 came with a spatial resolution of ca. 250 m at the equator. This resolution was set to be consistent with the satellite data of the Moderate Resolution Imaging Spectroradiometer (MODIS), which constituted a key input for the computation of the Land-cover and management factor (C), the spatial interpolation of the rainfall-runoff erosivity factor (R), and the spatial interpolation of the soil erodability factor (K). Accordingly, all these GloSEM input data were spatially optimized and consistently multiples of the base unit of 250 m defined by the MODIS data. The slope length and steepness factor (LS) was computed at 90 m native resolution (SRTM 3 arc-seconds data) and resampled to 250 m. The resampling caused a loss of detail that we consider within the tolerance threshold.

In the new GloSEM version 1.3, the adoption of the 100 m land-use data provided by ESA caused (i) a slight misalignment of the gridded structure and (ii) the loss of the perfect multiples structure guaranteed from the former 250 m MODIS-based spatial resolution. Despite not being optimal, in our view the increase in heterogeneity of input data due to the change in land-use product is overcompensated for, by the relevant increase in geographical representativeness provided by the ESA fractional (%) cropland layer. A future refinement of the other input data of GloSEM to 100 m spatial resolution is to be considered desirable.

Error estimate. A summary of error propagation in the present and future predictions was calculated considering: (i) the uncertainty of the spatial predictions estimated using a Markov Chain Monte Carlo approach; (ii) the uncertainty of estimating the area under CA (considered only for the 2015 scenario); (iii) the uncertainty related to the effectiveness of the CA practices using a Monte Carlo method; and (iv) the uncertainty in regional-rainfall-intensity-kinetic-energy relationships. A further uncertainty (v), related to the variation found in the 14 different GCMs, was also considered. The error propagation is expressed as the square root of the sum of squares of the different uncertainties.

Data Records

The GloSEM 1.3 dataset, containing present and future scenarios, is being made available via the European Soil Data Centre (ESDAC) (<https://esdac.jrc.ec.europa.eu/content/glosem>) and Figshare⁴¹ data repositories. The files are being provided in GeoTIFF format, which is a common GIS extension, storing georeferenced data accessible using GIS open-source software. Concerning the coordinate systems, all data are referenced to the World Geodetic System (WGS84). The cell size is 0.00099 decimal degrees, which is equivalent to ~100 m close to the Equator. A total of five layers is being released. Four of them provide estimated rates ($\text{Mg ha}^{-1} \text{ yr}^{-1}$) of soil displacement by water (SOIL_DISPLACEMENT_ESTIMATE_“scenario”.tif), while the last layer provides the fractional (%) cropland cover (ESA100_Cropland_fraction.tif). Future updated modelling outcomes will also be made available via the same ESDAC repository. Further details about the layers are provided below.

- SOIL_DISPLACEMENT_ESTIMATE_2019.tif—This scenario considers the ESA's global cropland (reference year 2019; 1.4 billion ha) and the recent climate scenarios described by Panagos *et al.*³⁶ and Borrelli *et al.*¹⁵.
- SOIL_DISPLACEMENT_ESTIMATE_2070_SSP1-RCP2.6.tif—This scenario considers the ESA's global cropland (reference year 2019; 1.4 billion ha) and the future (2070) climate scenario obtained from 14 GCMs (WorldClim) for the scenario RCP2.6 (CMIP5) (more details are given in Borrelli *et al.*⁹).
- SOIL_DISPLACEMENT_ESTIMATE_2070_SSP2-RCP4.5.tif—This scenario considers the ESA's global cropland (reference year 2019; 1.4 billion ha) and the future (2070) climate scenario obtained from 14 GCMs (WorldClim) for the scenario RCP4.5 (CMIP5) (more details are given in Borrelli *et al.*⁹).
- SOIL_DISPLACEMENT_ESTIMATE_2070_SSP5-RCP8.5.tif—This scenario considers the ESA's global cropland (reference year 2019; 1.4 billion ha) and the future (2070) climate scenario obtained from 14 GCMs (WorldClim) for the scenario RCP8.5 (CMIP5) (more details are given in Borrelli *et al.*⁹).
- ESA100_Cropland_fraction.tif—Global land-cover map (expressed in %) of the ESA and CGLS (CGLS-LC100 Collection 3) for the reference year 2019.

Technical Validation

As explained by Auerwald *et al.*⁴² and Borrelli *et al.*¹⁵, the *sensu stricto* validation of regional or larger-scale applications of RUSLE-based models, such as GloSEM, is challenging due to the lack of long-term field-scale measurements. Therefore, GloSEM is presented with a set of procedures to provide insights that support the plausibility of the global estimates, including: (i) uncertainty of the spatial predictions; (ii) a cross-comparison of results with continental studies; (iii) a comparison of results against empirical data; (iv) a comparison of results with river sediment load; (v) a comparison of results with past UN global assessments; (vi) a comparison of results from field/plot measurements; (vii) an uncertainty estimate of the climate erosivity; (viii) a GPR forecasting-capacity analysis; and (ix) a sensitivity analysis to determine the influence of the input parameters (see details in Borrelli *et al.*⁹). Here, we focused our attention on the GloSEM cross-comparison with continental-scale studies and the forecasting capacity of the GPR approach to predict future rainfall erosivity.

GloSEM cross-comparison with continental-scale studies. A cross-comparison of the GloSEM results was performed, comparing aggregated regional values of predicted soil erosion against the values obtained by regional soil-erosion assessments. The goal of this analysis was to assess the plausibility of the quantitative soil-erosion estimates. The first regional comparison was made with the cropland of the US. According to the National Resources Inventory (NRI) of the USDA⁴³, the soil displacement due to sheet and rill erosion in US croplands was estimated to be 1.59 Pg yr⁻¹ in 1982 and 0.96 Pg yr⁻¹ in 2012. The GloSEM estimates for soil erosion in US croplands during the same periods (considering the annual conditions of land use and CA) were equal to 1.52 and 0.91 Pg yr⁻¹ in 1982 and 2012, respectively. The deviation between the statistical data of the USDA and our model estimates was, in both cases, below 5%. The second regional comparison was carried out for the cropland of the European Union (EU) using the higher-spatial-resolution (25-m) RUSLE-based model, developed by the European Commission (EC)⁴⁴. Here, the comparison also revealed a consistency in soil erosion estimates for the EU cropland in the selected reference year 2012 (GloSEM = 0.304 Pg yr⁻¹; EC RUSLE2015 = 0.30 Pg yr⁻¹). In this second case, we also noted a minor deviation (below 5%) between the GloSEM and regional-scale estimates. Concerning 2019, the combination of the GloSEM 1.3 soil displacement rates with the country cropland statistics (US = 0.36 billion ha [US census of agriculture 2017]; EU = 0.17 billion ha [Eurostat 2018]) yielded estimates of 0.99 Pg yr⁻¹ for the US and 0.3 Pg yr⁻¹ for the EU croplands. Both of these are in line with the regional estimates provided by the NRI for the US (0.98 Pg yr⁻¹) (USDA⁴⁵) and the EC for the EU28 (0.3 Pg yr⁻¹)⁴⁴. The good agreement between our estimates and the ones provided by independent studies give confidence that the quantitative estimates achieved through the global model are reliable and valid to a level close to the higher-resolution regional assessments.

Evaluation of the GPR forecasting capacity. In order to test the GPR forecasting capacity, a model was fitted on the pre-2000 data using the estimated rainfall erosivity (R)-factor derived from measured data (dependent variable) and WorldClim historical data (averaged over the same timespan). The fitted GPR model was subsequently used to predict the values for the post-2000 period using the WorldClim data for the same time horizon. This procedure constitutes a form of external validation in which the validation dataset belongs to a different period. The procedure resulted in a very good prediction capacity for the pre-2000 training set (0.85 R²) and a good prediction capacity for the post-2000 validation set (0.6 R²). Note, however, that the model tended to be less effective with data with very high R-factor values, which mainly appeared at >10,000 MJ mm h⁻¹ ha⁻¹ yr⁻¹. Here, the model tended to underestimate the predicted R-factor values. Of further note was that the resolution of the WorldClim data is rather coarse, and this might result in a mismatch between station-measured data and WorldClim data, especially in areas where the topography strongly influences precipitation. Moreover, the two datasets are spatially mismatched, meaning that the stations from the pre-2000 dataset do not necessarily overlap those of the post-2000 dataset. This introduced more error because the model lacked full spatial coverage, allowing it to extrapolate to more diverse climatic conditions. Nevertheless, the model showed little bias when predicting the post-2000 R-factor values, which were comparable to those of the training data (pre-2000).

Usage Notes

The provided layers express rates of soil displacement by water erosion, expressed as a mass of soil lost per unit area and time (Mg ha⁻¹ yr⁻¹), given by the local combination of climate, soil, topography, farming and management system. These estimates do not account for the fractional (%) cropland cover of each cell. The total soil displacement of each scenario can be estimated by combining the provided soil erosion rate with the fractional (%) cropland cover layer (ESA100_Cropland_fraction.tif). This operation may require the reprojection of the (WGS84) geographic data to a local or global state-plane projection.

Code availability

No custom code was used to generate or process the data.

Received: 16 February 2022; Accepted: 20 June 2022;

Published online: 13 July 2022

References

1. Nearing, M. A., Yin, S.-Q., Borrelli, P. & Polyakov, V. O. Rainfall erosivity: An historical review. *Catena* **157** (2017).
2. Montgomery, D. R. Soil erosion and agricultural sustainability. *Proc. Natl. Acad. Sci. USA* **104**, 13268–72 (2007).
3. Foley, J. A. *et al.* Solutions for a cultivated planet. *Nature* **478**, 337–342 (2011).
4. Poesen, J. Soil erosion in the Anthropocene: Research needs. *Earth Surf. Process. Landforms* **84**, 64–84 (2018).
5. Amundson, R. *et al.* Soil and human security in the 21st century. *Science* (80-). **348**, 1261071–1261071 (2015).
6. Montanarella, L. Agricultural policy: Govern our soils. *Nature* **528**, 32–33 (2015).

7. Prestele, R., Hirsch, A. L., Davin, E. L., Seneviratne, S. I. & Verburg, P. H. A spatially explicit representation of conservation agriculture for application in global change studies. *Glob. Chang. Biol.* **24**, 4038–4053 (2018).
8. Amundson, R. Soil biogeochemistry and the global agricultural footprint. *Soil Secur.* **6**, 100022 (2022).
9. Borrelli, P. *et al.* Land use and climate change impacts on global soil erosion by water (2015–2070). *Proc. Natl. Acad. Sci. USA* **117**, 21994–22001 (2020).
10. Renard, K. G., Yoder, D. C., Lightle, D. T. & Dabney, S. M. Universal Soil Loss Equation and Revised Universal Soil Loss Equation. *Handb. Eros. Model.* 137–167, <https://doi.org/10.1002/9781444328455.ch8> (2010).
11. Borrelli, P. *et al.* Soil erosion modelling: A global review and statistical analysis. *Sci. Total Environ.* **780**, 146494 (2021).
12. Ferro, V. Deducing the USLE mathematical structure by dimensional analysis and self-similarity theory. *Biosyst. Eng.* **106**, 216–220 (2010).
13. Nearing, M. A. Soil Erosion and Conservation. in *Environmental Modelling: Finding Simplicity in Complexity: Second Edition* 365–378, <https://doi.org/10.1002/9781118351475.ch22> (2013).
14. Jetten, V., Govers, G. & Hessel, R. Erosion models: Quality of spatial predictions. *Hydrol. Process.* **17**, 887–900 (2003).
15. Borrelli, P. *et al.* An assessment of the global impact of 21st century land use change on soil erosion. *Nat. Commun.* **8**, 2013 (2017).
16. Doetterl, S., Van Oost, K. & Six, J. Towards constraining the magnitude of global agricultural sediment and soil organic carbon fluxes. *Earth Surf. Process. Landforms* **37**, 642–655 (2012).
17. Van Oost, K. *et al.* The Impact of Agricultural Soil Erosion on the Global Carbon Cycle. *Science (80-)*. **318**, 626–9 (2007).
18. Quinton, J. N., Govers, G., Van Oost, K. & Bardgett, R. D. The impact of agricultural soil erosion on biogeochemical cycling. *Nat. Geosci.* **3**, 311–314 (2010).
19. Chappell, A., Baldock, J. & Sanderman, J. The global significance of omitting soil erosion from soil organic carbon cycling schemes. *Nat. Clim. Chang.* <https://doi.org/10.1038/nclimate2829> (2016).
20. 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).
21. Naipal, V. *et al.* Global soil organic carbon removal by water erosion under climate change and land use change during AD 1850–2005. *Biogeosciences* **15**, 4459–4480 (2018).
22. Ito, A. Simulated impacts of climate and land-cover change on soil erosion and implication for the carbon cycle, 1901 to 2100. *Geophys. Res. Lett.* **34** (2007).
23. Yang, D., Kanae, S., Oki, T., Koike, T. & Musiakke, K. Global potential soil erosion with reference to land use and climate changes. *Hydrol. Process.* **17**, 2913–2928 (2003).
24. Grill, G. *et al.* Mapping the world's free-flowing rivers. *Nature* **569**, 215–221 (2019).
25. Alewell, C. *et al.* Global phosphorus shortage will be aggravated by soil erosion. *Nat. Commun.* **11** (2020).
26. Sartori, M. *et al.* A linkage between the biophysical and the economic: Assessing the global market impacts of soil erosion. *Land use policy* **86**, 299–312 (2019).
27. IPCC. Chapter 4: Land Degradation. In *Climate Change and Land* (eds. Olsson, L. *et al.*) 4–186 (2020).
28. Prävälle, R. *et al.* Arable lands under the pressure of multiple land degradation processes. A global perspective. *Environ. Res.* **194** (2021).
29. Suárez-Castro, A. F. *et al.* Global forest restoration opportunities to foster coral reef conservation. *Glob. Chang. Biol.* **27**, 5238–5252 (2021).
30. Simmonds, J. S. *et al.* Limiting the loss of terrestrial ecosystems to safeguard nature for biodiversity and humanity. *bioRxiv* **4111**, 2021.02.07.428694 (2021).
31. Buchhorn, M. *et al.* Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe. *zenodo* <https://doi.org/10.5281/zenodo.3939050> (2020).
32. Borrelli, P. *et al.* An assessment of the global impact of 21st century land use change on soil erosion. *Nat. Commun.* **8** (2017).
33. Monfreda, C., Ramankutty, N. & Foley, J. A. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochem. Cycles* **22**, 1–19 (2008).
34. Ramankutty, N., Evan, A. T., Monfreda, C. & Foley, J. A. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochem. Cycles* **22** (2008).
35. Hurtt, G. C. *et al.* Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geosci. Model Dev.* **13**, 5425–5464 (2020).
36. Panagos, P. *et al.* Global rainfall erosivity assessment based on high-temporal resolution rainfall records. *Sci. Rep.* **7**, 1–12 (2017).
37. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* <https://doi.org/10.1002/joc.5086> (2017).
38. 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, <https://doi.org/10.1029/TR039i002p00285> (1978).
39. Hengl, T. *et al.* SoilGrids1km — Global Soil Information Based on Automated Mapping. *PLoS One* **9**, e105992 (2014).
40. Desmet, P. & Govers, G. A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units. *J. Soil Water Conserv.* **51**, 427–433 (1996).
41. 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, *figshare*, <https://doi.org/10.6084/m9.figshare.c.5844758.v1> (2022).
42. Auerswald, K., Kainz, M. & Fiener, P. Soil erosion potential of organic versus conventional farming evaluated by USLE modelling of cropping statistics for agricultural districts in Bavaria. *Soil Use Manag.* **19**, 305–311 (2003).
43. USDA. 2012 National Resources Inventory. *2012 National Resources Inventory* 5–8, https://www.nrcs.usda.gov/Internet/NRCS_RCA/reports/nri_wet_md.html (2015).
44. Panagos, P. *et al.* The new assessment of soil loss by water erosion in Europe. *Environ. Sci. Policy* **54** (2015).
45. USDA. National Resources Inventory. <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/technical/nra/nri/> (2017).

Acknowledgements

P.B. and J.E.Y. were funded by the EcoSSoil Project, Korea Environmental Industry & Technology Institute (KEITI), Korea (Grant No. 2019002820004). We thank all GloSEM data providers, i.e., EC-JRC, EC-ESDAC, ESA, NASA, ISRIC, JRC FAOSTAT, AQUASTAT, LUH2, WorldClim, GloREDa.

Author contributions

Data processing, analysis, and methodology: P.B., C.B., P.P. Data interpretation, discussions and writing: P.B., D.A.R., J.E.Y., C.B., P.P. Funding acquisition: J.E.Y., D.A.R.

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/s41597-022-01489-x>.

Correspondence and requests for materials should be addressed to P.B. or P.P.

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) 2022