Geoderma 446 (2024) 116905

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

High-resolution soil erosion mapping in croplands via Sentinel-2 bare soil imaging and a two-step classification approach

Lulu Qi [a,b], Yue Zhou [b], Kristof Van Oost [b], Jiamin Ma [a], Bas van Wesemael [b], Pu Shi [a,b,*]

[a] College of Earth Sciences, Jilin University, Changchun 130061, China
[b] Earth and Life Institute, UCLouvain, 1348 Louvain-la-Neuve, Belgium

ARTICLE INFO

Handling Editor: B. Minasny

Keywords:
Soil erosion
Sentinel-2
Bare soil composite
Soil spectroscopy
Google Earth
Unsupervised classification

ABSTRACT

Erosion-induced lateral soil redistribution leads to spatially heterogenous soil composition, which can be captured through the distinctive spectral reflectance of soils under varying levels of erosion influence. This points to the potential of using remotely sensed soil spectra to detect severe erosion and deposition hotspots in exposed croplands and, importantly, further differentiate the intra-class spectral variability of moderate erosion that often occupies the largest proportion. Here, we aim to develop a two-step erosion classification and mapping approach based on multitemporal compositing of Sentinel-2 bare soil images of a typical agricultural region (11,500 km^2) of northeast China. A random forest classifier was firstly trained against the ground-truth data derived from very high resolution (VHR) imagery in Google Earth, with an overall accuracy of 91 % that allowed for clear delineation of severe erosion and deposition areas based on their distinct topographic and spectral features particularly in the red and red-edge bands. In the second step, the remaining area of moderate erosion (60.30 %) was further differentiated using Iterative Self-Organizing cluster unsupervised classification to yield a five-class soil erosion map at 10 m spatial resolution. The accuracy of the predicted map was successfully validated by independent Caesium-137 (^{137}Cs) and soil organic carbon observations at catchment and regional scales, as revealed by significant inter-class differences in soil redistribution rates estimated from ^{137}Cs inventory. The severe erosion class had a soil loss rate of 5.5 mm yr^{-1}, suggesting that previous assessments have underestimated erosion severity. The spatial accordance of crop growth with soil erosion intensity, particularly in localized settings, further highlighted the potential of bare soil imaging for mapping the spatiotemporal development of soil erosion and its response to targeted sustainable cropland management efforts.

1. Introduction

Soil erosion has become an increasingly serious environmental and food security concern around the world (Wuepper et al., 2019), as intensifying human activities has led to rates of erosion that exceed those of natural soil formation by several orders of magnitude (Evans et al., 2020; García-Ruiz et al., 2015). Accelerated soil erosion not only causes significant loss of the fertile topsoil resources that are essential to sustaining soil productivity (Bhandari et al., 2021), but also poses a range of ecological challenges, such as biodiversity loss, water pollution and river sedimentation (Mathieu et al., 2007; Vagen and Winowiecki, 2019). In this context, it is important to have an accurate and up-to-date spatial representation of soil erosion processes and patterns, to pinpoint severely degraded areas and to allow assessments on the countermeasures to alleviate their on-site and off-site impacts (Borrelli et al., 2018; Poesen, 2018). This is further highlighted by the United Nations Sustainable Development Goals to achieve land degradation neutrality by 2030, for which innovative soil erosion assessment tools are required for more effective quantification and monitoring of land degradation (Lorenz et al., 2019; Sims et al. (2021)).

Spatially explicit analysis of soil erosion and degradation has progressed from approaches based on expert judgments to a more comprehensive framework of mathematical modelling and remote sensing applications, driven by an unprecedented flow of Earth Observation (EO) data streams (Efthimiou et al., 2020; Giuliani et al., 2020). A prime example is the widely-used of USLE-type soil erosion models that capitalize on the high-quality soil, land use and cover, and management data available from the new generation of satellite imaging and spatial interpolation techniques at regional (Al-Mamari et al., 2023), national (Liu et al., 2020; Prasuhn et al., 2013), and global scales (Alewell et al.,

* Corresponding author at: College of Earth Sciences, Jilin University, Changchun 130061, China.
E-mail address: shipu@jlu.edu.cn (P. Shi).

https://doi.org/10.1016/j.geoderma.2024.116905
Received 7 August 2023; Received in revised form 29 April 2024; Accepted 3 May 2024
Available online 7 May 2024

2019; Benavidez et al., 2018; Borrelli et al., 2022). However, the empirical USLE-based algorithms remain limited in two important aspects: (1) they lack any consideration of other forms of erosion (e.g. wind, tillage) than water erosion; and (2) they have a compromised spatial predictive capability due to insufficient observation and validation (Alewell et al., 2019; Jetten et al., 2003). Aside from soil erosion modelling, researchers have exploited the spatial heterogeneity of vegetation metrics to indicate the intensity of soil erosion. In particular, Öttl et al. (2021) employed the Enhanced Vegetation Index (EVI), which was calculated from high-resolution RAPIDEYE satellite imagery, to demonstrate that intra-field variation in crop biomass was primarily driven by tillage erosion. Vagen and Winowiecki (2019) developed predictive models to map soil erosion prevalence in the global tropics utilizing the spectral characteristics of annual MODIS composites. In addition, based on consistent, remotely sensed Normalized Difference Vegetation Index (NDVI) datasets, Bai et al. (2008) and Le et al. (2016) used the temporal deviation of NDVI from the norm as a proxy to identify land degradation hotspots at a global scale, and suggested that such approaches are more suitable for large-scale investigations that cover a gradient of agroecosystems and biophysical conditions.

In recent years, significant progress has been made in spectral imaging of bare earth surfaces for enhanced digital soil mapping and other surveying purposes (Chabrillat et al., 2019; Roberts et al., 2019), particularly since the new generation of multispectral satellites, such as Sentinel-2, which offer improved spectral, temporal, and spatial resolutions (Drusch et al., 2012). This is particularly relevant for soil erosion mapping in croplands because erosion-induced lateral mobilization of soil materials can lead to spatial variations in soil composition, which displays distinct absorption features in the visible to short-wave infrared (Vis-SWIR) spectral region. Based on this principle, numerous studies have combined the spectral response of bare soils provided by Sentinel-2 with machine learning to effectively predict key soil properties such as soil organic carbon (SOC) at high spatial resolutions (see reviews by Vaudour et al. (2022) or Tziolas et al. (2021)). In brief, it has become well-established that accurate predictive mapping of soil properties requires (1) that pure bare soil spectra are extracted with minimal interference from crop residue and soil moisture, usually by applying spectral index thresholding to remove contaminated pixels (Dvorakova et al., 2023; Heiden et al., 2022); and (2) that multitemporal composites are created from a stack of processed bare soil images to obtain stable and representative spectral signatures and more spatially continuous coverage (Demattê et al., 2018; Shi et al., 2022; Zepp et al., 2021).

Building on these advancements, several pilot studies have demonstrated that bare soil imaging can be used to map soil erosion in croplands, differentiating between soils experiencing different erosion intensities based on their respective spectral responses. For instance, Schmid et al. (2016) and Bracken et al. (2019) used airborne and spaceborne hyperspectral data to classify soils of different erosion degrees in a Mediterranean setting. Their investigations revealed that the disparate soil horizons, formed by the heterogenous processes of erosion and deposition, exhibited discernible spectral characteristics that corresponded to varying soil properties, including iron oxides and SOC. Furthermore, Žížala et al. (2019) extracted bare soil spectra from a Sentinel-2 time-series and used the ISODATA unsupervised classification model to successfully identify severely eroded areas at an accuracy of 80.9 %, which was further improved through visual refinement using orthoimages, in a Chernozem region of the Czech Republic. Similarly, using a Sentinel-2 multitemporal bare soil composite in the Phaeozem and Chernozem region of northeast China, Qi et al. (2023) developed a soil erosion classification scheme consisting of two distinctive spectral indexes, which related to soil albedo and biochemical composition influenced primarily by erosion-induced SOC mobilization. However, while evidence that supports the potential use of remotely-sensed bare soil spectra for mapping soil erosion is growing, most studies to date have been conducted on a limited scale and have focused on detecting hotspots of excessive removal or accumulation of topsoil materials. Less attention is paid to the less extreme levels of erosion, while treating the erosion stage in-between as a lumped class, which often occupied the majority of the investigated areas (Bracken et al., 2019; Qi et al., 2023). There is thus a need to develop a more detailed and comprehensive erosion classification framework that is applicable to larger regions.

One important obstacle to developing a spectra-based erosion classification model at large scale is constructing an adequate training dataset, which would require an unrealistic level of investment in conventional field observation and/or fallout radionuclide tracing (Manić et al., 2022; Pradhan et al., 2012). In this regard, the increasing availability of very high resolution (VHR) imagery on platforms such as Google Earth offers new opportunities to visually identify clear erosional and depositional features (Boardman, 2016) and use this evidence for model calibration. Thaler et al. (2021) used a collection of VHR images (e.g. Worldview, Quickbird) of bare soil fields in the corn belt of the United States of America to identify SOC-depleted areas, the spectral signatures of which were then combined with terrain attributes to develop a binary classification model for mapping areas with no remaining A-horizon. Šarapatka and Bednář (2021) also used Google Earth VHR imagery to delineate erosion-affected localities and depositional areas by identifying visibly lighter and darker areas, highlighting the capability of VHR imagery to facilitate the creation of a ground-truth dataset for model development.

The objective of this study was to establish a spectral-based classification approach for accurate and high-resolution mapping of various soil erosion stages, ranging from severe erosion to deposition, in a representative agricultural region of northeast China. More specifically, bare soil spectral data derived from a Sentinel-2 multitemporal composite were integrated with erosion-related vegetation and topographic variables, in an attempt to form a two-step classification process that includes (1) developing of a Random Forest (RF) classifier based on a ground-truth dataset containing three classes (severe erosion, moderate erosion, and deposition) as derived from Google Earth VHR imagery; and (2) subsequently differentiating the moderate erosion class in detail, based on an unsupervised classification method using the same collection of explanatory variables. Finally, observational datasets which included Caesium-137 (^{137}Cs) and SOC measurements at catchment and regional scales were used to validate the predicted soil erosion map.

2. Study region

The black soil region of northeast China, which is mostly characterized by loess-derived Phaeozems and Chernozems (IUSS Working Group WRB, 2015), is known as an important breadbasket for the country. However, prolonged intensive cultivation has caused severe erosion and degradation of the fertile croplands in this region. Recent reports indicate that it is among the most severely affected regions in China in terms of soil erosion (Liu et al., 2020). The selected study area (43°50′–45°15′ N, 125°14′–127°05′ E) comprises three administrative counties (Yushu, Dehui,and Jiutai) which occupy an area of around 11,500 km^2 in central Jilin Province of northeast China (Fig. 1). This area was selected because it represents a typical maize production region that has been largely under conventional ridge tillage for decades. Its representativeness is also reflected in the homogenous distribution of soil types (i.e. Phaeozem and Chernozem) and land cultivation history. Land reclamation activities in this area started in the 1880s following large-scale human migration from the North China Plain to northeast China (Ye et al., 2009).

The study area has a temperate continental monsoon climate, with average annual precipitation of 565.2 mm and an average annual temperature of 4 °C. Rainfall is typically concentrated into the summer months (June-August) covering the entire vegetative stage of maize growth. The topography of the region is characterized by long and gentle slopes that often display localized erosion and deposition hotspots with distinct soil biochemical properties (Hu et al., 2020). Croplands occupy almost 70 % of the region, with maize being the dominant crop in a

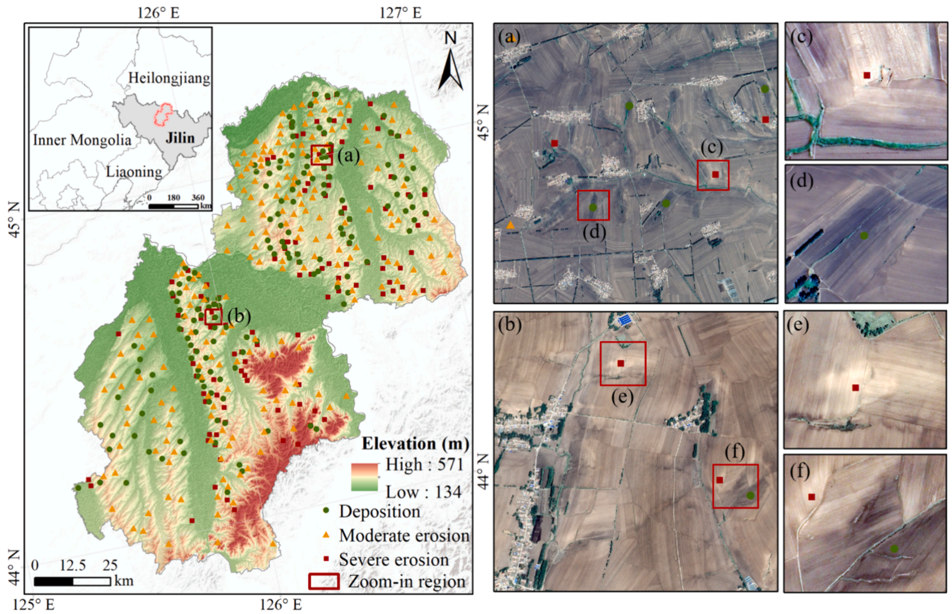

Fig. 1. Spatial distribution of the selected training data points in the study region. Three categories representing severe erosion, moderate erosion, and deposition were selected according to the very high resolution (VHR) bare soil images (CNES/Airbus) in Google Earth. The VHR images on the right show the visual distinctions between areas of severe erosion and deposition.

largely monocultural farming system. Maize is sown in April and May, during which months vast areas of croplands are exposed as bare fields, ideal for multispectral remote sensing to identify different erosion features.

3. Methodology

Fig. 2 depicts the workflow of the two-step soil erosion classification methodology. Firstly, an RF model was trained to classify croplands into three categories (severe erosion, moderate erosion, deposition), primarily based on remotely sensed bare soil spectral data; secondly, the Iterative Self-Organizing (ISO) cluster unsupervised classification was conducted to further separate moderate erosion into three sub-classes, thus producing a high-resolution soil erosion map with five classes.

3.1. Ground-truth dataset

A categorical ground-truth dataset, consisting of severe erosion, moderate erosion, and deposition classes, was established by selecting representative sites with identifiable erosion and deposition features from Google Earth VHR images (CNES/Airbus and Maxar Technologies) at sub-meter resolution. To ensure the quality of the ground-truth data, only cloud- and shadow-free VHR images from April and May of 2017–2022 were used, maximizing the coverage of bare croplands. In cases where multiple images were available for a site, only the most recent image was used.

Fig. 1 shows the spatial distribution of the selected points in the study region. The selection procedure was conducted on the basis of previous findings that, in this region, severe erosion would result in the exposure of light-colored subsoils due to significant truncation of the A-horizon, and that the redistribution of eroded materials in colluvial settings would show up as a darker color (Qi et al., 2023), as illustrated

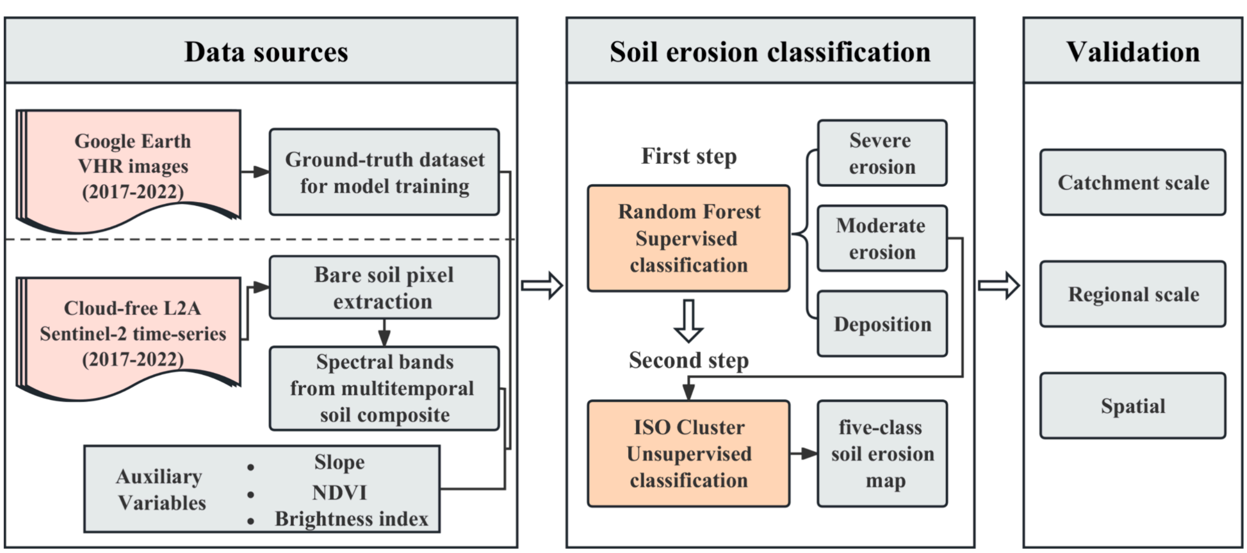

Fig. 2. Flowchart of the development and validation of the two-step soil erosion classification approach.

by the visible differences between severe erosion and deposition in the zoomed-in images. Having selected the severe erosion and deposition sites, moderate erosion sites were chosen between those two extreme conditions. It is important to note that this study assumed that spatial variability of soil color and biochemical composition was primarily driven by soil erosion processes, because other factors, such as land cultivation history and soil type, were homogenous throughout the region (Hu et al., 2020; Zhu et al., 2021). The visual selection of training data points was independently carried out by one person, then judged by a second person without prior knowledge of their assigned classes. Only the mutually agreed points were kept, and their geographic coordinates extracted, resulting in a total of 471 points (severe erosion,123; moderate erosion,199; and deposition, 149).

3.2. Sentinel-2 multitemporal bare soil composite

A multitemporal Sentinel-2 bare soil composite was created as the core input into the classification model. For this purpose, Sentinel-2 top-of-atmosphere Level-1C products covering five tiles (T51TXK, T51TYL, T51TYK, T51TYJ and T52TCQ) were downloaded using an optimal time window filter as designated by Shi et al. (2022), including a cloud filter that only retained images with less than 10 % cloud cover. This was done to ensure that only images from April and May, when bare soil coverage is at its maximum, were included in the time-series from 2017 to 2022. As a result, 146 images were downloaded, atmospherically corrected to Level-2A using the Sen2Cor processor (standalone version 2.10), and resampled to a 10 m resolution. Then, each image went through a processing chain, including (1) masking clouds, thin cirrus, and shadows according to the scene classification layer output from the Sen2Cor algorithm; (2) masking croplands using the parcel map from the 2019 Land Use Survey in Jilin Province, to remove other land use types (Fig. 3); and (3) spectral index thresholding using a combined NDVI (0.1–0.25) and normalized difference tillage index (NDTI: 0–0.075) to retain only bare soil pixels with minimal disturbances. These thresholds for NDVI and NDTI were adopted from a recent study by Shi et al. (2022), who demonstrated the effectiveness of removing pixels influenced by green vegetation, crop residue, and soil moisture in the same study region. Finally, all the processed images were composited, and pixel-wise calculation of mean reflectance for each band was conducted to create a multitemporal bare soil composite. In addition, we extracted the spectra of the training data points and performed principal component analysis (PCA), with the goal of examining whether soil spectral characteristics differed under various erosion influences. This procedure served as the basis for conducting the spectra-based soil erosion classification.

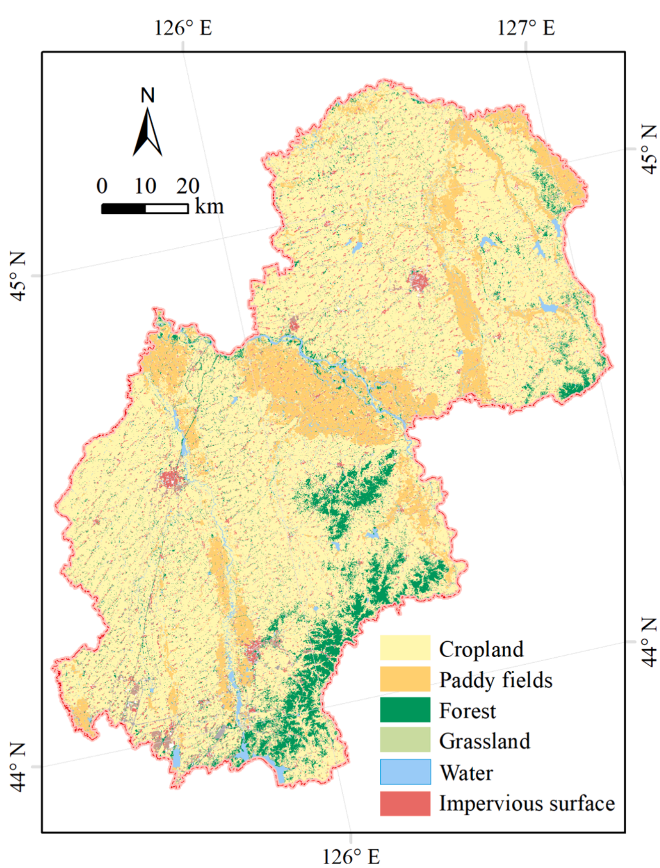

Fig. 3. Land use distribution in the study area. The spatial extent of croplands was used as the cropland mask to remove other land use types.

3.3. Other explanatory variables

In addition to the raw data from 10 Sentinel-2 bands that encompass visible (B2, B3, B4), red-edge (B5, B6, B7), near-infrared (B8, B8A), and SWIR (B11 and B12) regions, additional explanatory variables, including slope gradient, NDVI, and brightness index (BI), were incorporated into the classification model to account for potential influences of vegetation and topography on soil erosion (Fig. 2). Specifically, slope gradient is known to greatly affect erosion prevalence and serves as an important input in almost all erosion models. Slope gradient was calculated using the SRTM one arc-second digital elevation model (DEM) using ArcGIS 10.0. It is worth mentioning that slope curvature and slope length were initially included in the model as additional terrain attributes, but a recursive feature elimination analysis showed that these variables did not improve the model performance and they were therefore removed from subsequent development of the model. NDVI was intended to represent the response of crop growth to erosion, and average NDVI during 2019–2022 was obtained from Sentinel-2 data using the Google Earth Engine platform. It should be noted that only images from between May and September of each year were included, corresponding to the crop growth period that is most indicative of any erosion influence. Lastly, Eq. (1) was applied to Sentinel-2 B3 and B4 bands to calculate BI, which has been reported to be related to soil organic matter and salt content (Escadafal et al., 1989).

$$BI = \sqrt{(Red)^2 + (Green)^2} \quad (1)$$

3.4. Two-step soil erosion classification

3.4.1. Random forest classifier

Using the collection of explanatory variables outlined above, a supervised RF classification model was developed to predict the spatial distribution of severe erosion, moderate erosion, and deposition. RF is an ensemble learning algorithm that combines the predictions of multiple decision trees to achieve a final output (Breiman, 2001). The decision trees are developed based on a procedure known as bagging, which essentially creates multiple subsets of training data through bootstrap resampling with replacement. In this study, we performed additional data partitioning prior to RF modelling, meaning that the entire training dataset of 471 observations was split into 10 folds while maintaining the relative proportions of the three erosion classes. We then performed 10-fold cross-validation to enable a more robust assessment of prediction accuracy than the out-of-bag error using the entire dataset.

The overall performance of the model was evaluated by generating a confusion matrix based on the prediction results from cross-validation. The confusion matrix categorizes the predictions into true positive (TP), true negative (TN), false positive (FP), and false negative (FN), from which the overall accuracy (OA) and kappa coefficient were calculated using Eqs. (2) and (3).

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Kappa coefficient} = \frac{p_o - p_e}{1 - p_e} \quad (3)$$

where: p_o is the observed proportionate agreement; p_e is the expected proportionate agreement.

Sensitivity (true positive rate), Specificity (true negative rate), and F1 score (Eqs. (4)–(7) were also included as complementary indicators, where high values indicate higher classification accuracy for each individual class.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$F1 = \frac{2^*(Precision^*Sensitivity)}{Precision + Sensitivity} \quad (7)$$

To identify the important variables that contributed to the RF classification model, a variable importance plot was created using the ‘importance’ function of the ‘randomForest’ (Liaw and Wiener, 2002) package in R (R Core Team, 2022), ranking the mean decreases in model accuracy and Gini coefficient if a particular variable was excluded. Finally, pixel-wise soil erosion mapping for the study region was conducted using the developed RF model. The modelling and mapping procedure was carried out using the ‘caret’ (Kuhn, 2008) and ‘randomForest’ packages.

3.4.2. Unsupervised classification of moderate erosion

The spatial extent of moderate erosion areas was extracted from the three-class erosion map above, and ISO cluster unsupervised classification was then performed to further classify moderate erosion into three separate classes. This method was chosen because it allows for the identification of spectrally similar clusters without training data, and is computationally inexpensive, thus can be applied to large regions. The same collection of explanatory variables was normalized and subjected to ISO clustering analysis, which is an iterative self-organizing process to assign pixels into specific clusters, based on minimum Euclidean distance to cluster means in the multidimensional feature space of the input variables. Cluster means were updated during each iteration, and the maximum number of iterations was set at 30. The resultant signature file, containing the mean value and the variance–covariance matrix of each identified cluster, was then treated as pseudo training data, and passed through a maximum likelihood classification algorithm to refine a map of moderate erosion with three classes (M1, M2, M3), which was eventually merged with severe erosion and deposition classes to produce a five-class soil erosion map.

3.5. Model validation

3.5.1. Catchment scale

To validate the predicted soil erosion map, two observational datasets were constructed at catchment and regional scales respectively. For catchment-scale validation, the fallout radionuclide ^{137}Cs inventory was quantified at pre-designated locations, as shown in Fig. 4, and used to estimate the rate of soil redistribution. These data served not only as direct evidence to assess whether there were significant differences among predicted erosion classes, but also enable each class to be assigned a quantitative erosion rate. Soil cores were collected from 72 locations in 2021, following a stratified sampling design as detailed in Qi et al. (2023). ^{137}Cs activity (Bq kg^{-1}) was analyzed using a hyperpure coaxial Ge detector coupled to a multichannel analyzer at 662 keV peak (Li et al., 2006) and converted to inventory (Bq m^{-2}) based on the sampling area and the weight of the soil cores. Finally, ^{137}Cs inventory was used to estimate the erosion (E, m yr^{-1}) and deposition (D, m yr^{-1}) rates, together referred to as the soil redistribution rate, according to the formula proposed by Van Oost et al. (2007).

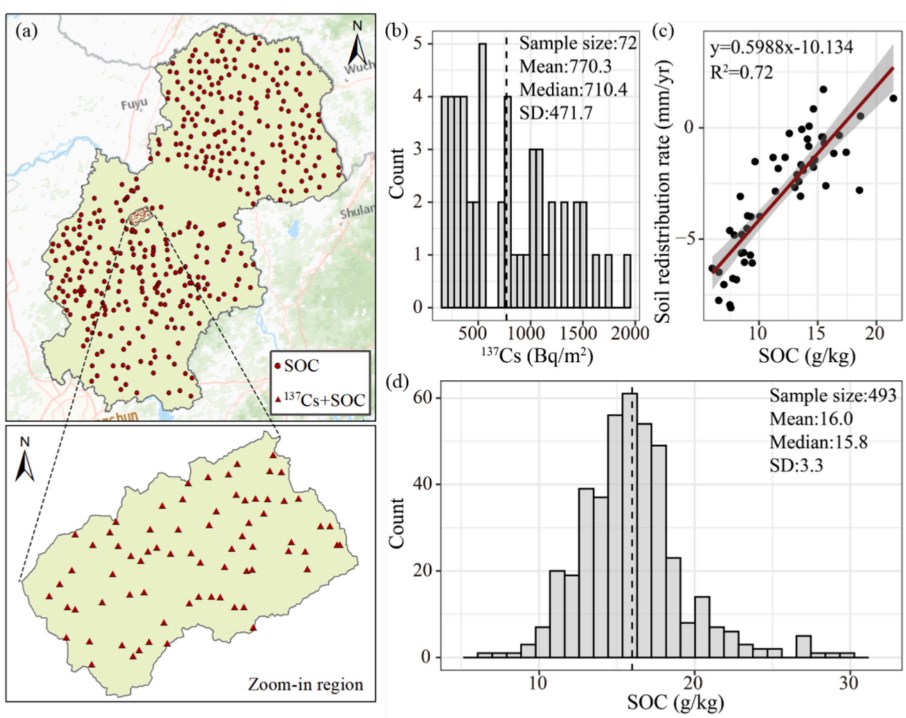

Fig. 4. (a) Spatial distribution of the observational datapoints at catchment and regional scale; (b) histogram of ^{137}Cs activity in the catchment; (c) linear regression analysis between SOC and soil redistribution rate, which was estimated using ^{137}Cs inventory; (d) histogram of the SOC data at regional scale. SD: standard deviation.

$$E = P \times [1 - (\frac{Cs}{Cs_{ref}})^{\frac{1}{ts}}] \quad (8)$$

$$D = P \times \frac{Cs - Cs_{ref}}{ts \times Cs_e} \quad (9)$$

where P is the depth (m) of plough layer, Cs and Cs_{ref} (Bq m^{-2}) represent the ^{137}Cs inventory of the sampling locations and the reference site, the latter of which was taken from a site in close proximity to the catchment (Yan and Tang, 2004), ts is the period between the sampling year and reference year of 1954, and Cs_e is the average ^{137}Cs inventory of the locations experiencing erosion.

In addition, topsoil (0–20 cm) samples for SOC measurement were collected from the same locations. Air-dried and finely ground (<150 μm) soil samples were analyzed for total carbon content by means of dry combustion (VarioMax CN Analyzer Elementar GmbH, Langenselbold, Hesse, Germany). For samples that showed clear reactions under 10 % HCl treatment, inorganic C was determined using a pressure-calcimeter method (Sherrod et al., 2002). SOC was then calculated as the difference between total and inorganic carbon. Fig. 4 (b) and (c) show the distribution of ^{137}Cs activity and the estimated soil redistribution rate in relation to SOC content. Apart from the wide range of variation in soil redistribution rate from around 2 mm yr^{-1} (net deposition) to more than −7 mm yr^{-1} (net erosion), a linear regression analysis demonstrated the significant correlation between the two variables ($R^2 = 0.72$).

3.5.2. Regional scale

For regional-scale validation, given that SOC correlated significantly with soil redistribution rate and that the relationship between the two variables was found to be stable in areas under the same land use and climatic conditions (Wang et al., 2023), an independent SOC dataset covering the entire region was used to evaluate the inter-class differences by means of pair-wise Wilcox tests at $P < 0.05$. Only SOC data was determined because ^{137}Cs tracing is resource-intensive and not suitable for estimating erosion rates at large spatial scales (Zhang et al., 2019). In total, 493 topsoil (0–20 cm) samples were randomly collected (Fig. 4) and analyzed for SOC content following the same procedure described above. Fig. 4(d) shows the distribution of SOC content, with a mean value of 16.0 g kg^{-1} and a standard deviation of 3.3 g kg^{-1}.

Lastly, the predicted soil erosion map was confronted against an NDVI map derived from a Sentinel-2 image acquired on June 22, 2021 during the vegetative stage of maize growth which is most sensitive to the influence of soil erosion (Qi et al., 2023). The NDVI data were grouped and their distributions compared between the different erosion classes. Two zoom-in areas were also selected for further investigation into the localized distribution of varying erosion intensities and their correspondence with the NDVI pattern.

4. Results

4.1. Bare soil spectral characteristics of three erosion classes

Spectra-based classification of soil erosion is based on the premise that bare soil pixels extracted from the Sentinel-2 multitemporal composite would exhibit distinct spectral responses according to erosion intensity. To assess this, we first analyzed the spectral characteristics of the training dataset with severe erosion, moderate erosion, and deposition classes. As depicted in Fig. 5, the spectral curves for the three classes followed similar shapes across the Vis-SWIR region, but differed in absolute band-wise reflectance levels, which increased with erosion intensity. The inter-class spectral difference could be largely attributed to the albedo shifts caused by erosion-induced changes in soil color, as described in Chabrillat et al. (2019). This was further confirmed by PCA analysis, which revealed that the first and second principal components (PC1 and PC2) accounted for 98.86 % of the total variance (Fig. 5(b)). Clear spectral separability was observed between severe erosion and deposition classes, as the PC1 score of the severe erosion classes was significant higher ($P < 0.05$) than that of the deposition class. Furthermore, there was some overlap between moderate erosion and the other two classes, which could be attributed to the broader spectral range encompassed by the moderate erosion class. In contrast to severe erosion and deposition areas, that visibly differ in VHR images, soils experiencing moderate erosion are likely to exhibit more intricate gradients of variations in soil properties and be less clear-cut in their spectral characteristics. This is precisely why a more detailed classification of this class is needed.

4.2. Evaluation of the spectra-based RF classification model

The RF classification model was developed using bare soil spectra as the core inputs and NDVI, BI and slope as auxiliary variables. The confusion matrix based on cross-validation revealed that, in general, the RF model produced excellent classification accuracy, with the OA at 91

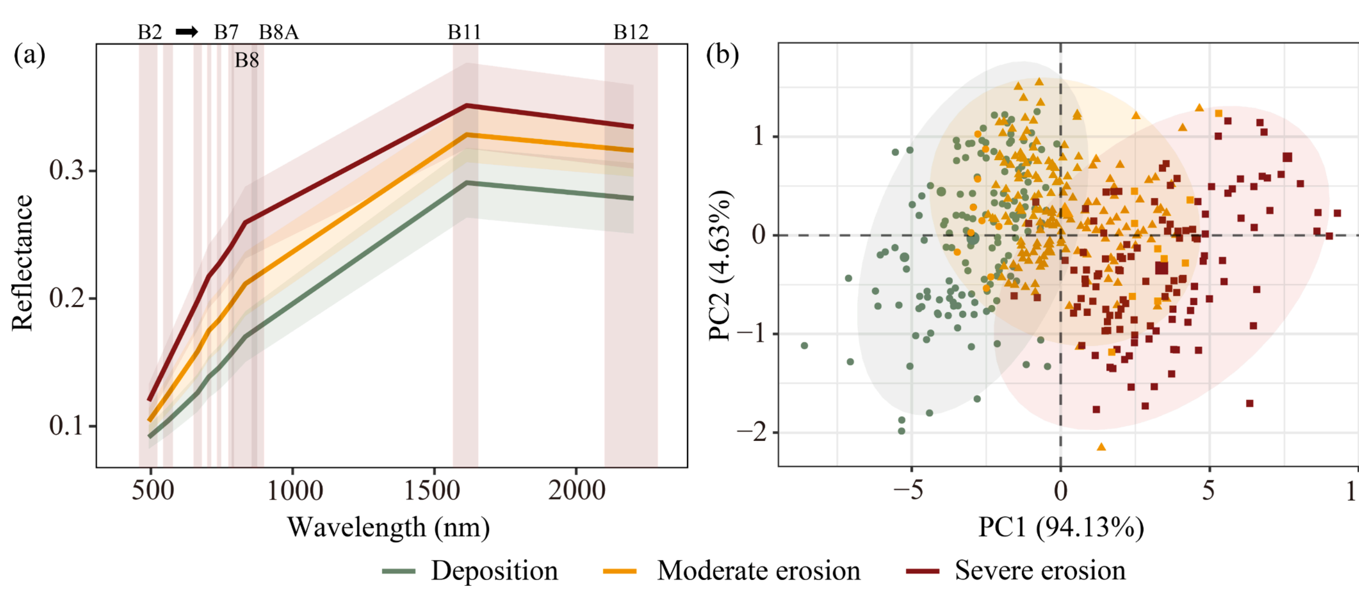

Fig. 5. (a) Bare soil spectral data of three erosion classes extracted from the Sentinel-2 multitemporal composite; and (b) principal component analysis (PCA) of bare soil spectra. The mis-classified moderate erosion points in Table 1 are indicated with different symbols.

% and the kappa coefficient at 0.87 (Table 1). This was further supported by the specificity, sensitivity, and F1-score indexes, all of which had high values, particularly for the severe erosion and deposition classes. The moderate erosion class had some misclassifications, with 10 observations from this class erroneously classified as deposition and 10 observations classified as severe erosion (the misclassified points are indicated in Fig. 5), highlighting once again the wide range of variation in its spectral characteristics. It should be noted that, to assess the prediction uncertainty of the model, the RF modelling procedure was repeated 50 times. The results show that the OA ranged from 80 % to 93 %, with a mean OA of 88 %, indicating that the procedure was generally robust.

Further examination of the relative importance of variables in the RF model revealed that the slope factor contributed the most to erosion classification, followed by the red (B4) and red-edge (B5, B6, B7) bands from the bare soil composite, whereas the blue and green bands and, to a lesser extent, the SWIR bands, played a relatively minor role (Fig. 6). Comparing models with and without the slope factor, the addition of slope gradient into the modelling framework increased the OA of the RF model from 77 % to 91 %. The other two additional covariates, BI and NDVI, made drastically different contributions. BI was found to be an important predictor while NDVI was ranked as the least influential variable in the RF model, suggesting that soil erosion did not greatly affect the spatial pattern of multi-year average NDVI values.

4.3. Two-step soil erosion mapping and comparison

Applying the established RF model to the multi-band raster of input variables, a three-class pixel-wise soil erosion map at 10 m resolution revealed the spatial distribution of different erosion classes and their percentages relative to the total cropland area (Fig. 7(a)). Specifically, severe erosion, which occupied 12.55 % of the croplands, was found throughout the study region in the form of localized hotspots, but most commonly in the eastern part of the study region which is characterized by the highest elevation and steepest slopes (Fig. 1). This observation aligns with the finding that slope was the most important predictor. 27.15 % of the croplands were classified as deposition areas while, most importantly, over 60 % were in the moderate erosion class. This justifies the rationale for the two-step classification approach adopted here, in which the second step specifically attempted to achieve a detailed classification of moderate erosion.

ISO cluster unsupervised classification of moderate erosion resulted in a five-class soil erosion map that offered more insights into the extent and degree of soil erosion on croplands (Fig. 7(b)). The 60.30 % moderate erosion area was further divided into three sub-classes: M1, M2, and M3, occupying 25.26 %, 20.13 %, and 14.91 % respectively. One notable insight arising from the five-class map is that it demonstrates how intertwined the erosion–deposition continuum is, particularly in the south west of the study area where the detailed map reveals a heterogeneous pattern of soil erosion intensity which reflects an undulating topography characterized by frequent changes in slope gradient and soil composition, and hence the soil spectral signature.

4.4. Multi-aspect validation of the predicted erosion map

The predicted five-class soil erosion map was validated from three aspects: (1) at catchment scale using the estimated soil redistribution data from $^{137}C_S$ observations; (2) at regional scale with SOC observations based on the positive correlation between soil redistribution rate and SOC; and (3) from a spatial perspective to examine whether NDVI during crop growth varied in line with soil erosion intensity. Firstly, pair-wise significance tests revealed significantly different soil redistribution rates among the five classes, ranging from strikingly low at < -6 mm yr^{-1} in the severe erosion class to positive values in the deposition class (Fig. 8(a)). This alignment between soil redistribution rates and soil erosion classes suggests that the two-step classification was able to accurately predict soil erosion patterns at the catchment scale. However, minor discrepancies were noted in the deposition class, as the estimated soil redistribution rate was associated with some negative values, indicating net erosion scenarios. This may mean that the classification model overestimated the percentage of deposition areas, and/or suggest that the selection of the reference site used to determine net erosion or deposition conditions needs to be further investigated.

Secondly, scaling up to the entire study region, a similar trend was found in the SOC observational dataset. As the intensity of soil erosion reduced, SOC content in the region exhibited a consistent and significant increase, with average values increasing almost two-fold from 10.8 g kg^{-1} for severe erosion to 18.4 g kg^{-1} for deposition. This further proves that the spectra-based erosion classification appraoch is capable of reproducing soil erosion patterns at a large scale.

Thirdly, using NDVI during the maize growth stage as a proxy for crop biomass, it was found that the density distribution of NDVI displayed a slight but consistent shift towards larger values as the intensity of soil erosion lessened, with the net deposition class having the highest NDVI values overall (Fig. 9). However, there were significant inter-class overlaps in NDVI distribution, which perhaps explains why NDVI's contribution to the classification model was insignificant (Fig. 6). Nevertheless, the effect of erosion on NDVI became clearer when selected zoom-in regions were inspected. As shown in Fig. 9(a)–(d), the accuracy in pinpointing erosion and deposition hotspots was supported by the clear spatial agreement between soil erosion intensity and NDVI, with severely eroded areas generally associated with lower NDVI values and vice versa. To further elaborate the spatial agreement between the two variables, correlation analysis was conducted on 200 randomly selected points from each of the two zoom-in regions. The results show significant negative correlation between soil erosion intensity and NDVI, with correlation coefficients at -0.50 and -0.48 respectively. These findings suggest that soil erosion intensity could exert important control on crop growth, particularly in localized settings where the environmental variables remain homogenous, while the impact of land management and pedogenic factors could be more relevant at the regional scale.

Table 1
Classification accuracy of the random forest model based on the confusion matrix.

Prediction	Observation					
	Deposition	Moderate Erosion	Severe Erosion	Specificity	Sensitivity	F1-Score
Deposition	140	10	0	0.97	0.93	0.93
Moderate Erosion	10	170	10	0.93	0.89	0.89
Severe Erosion	0	10	107	0.97	0.91	0.91
Overall accuracy	91 %					
Kappa coefficient	0.87					

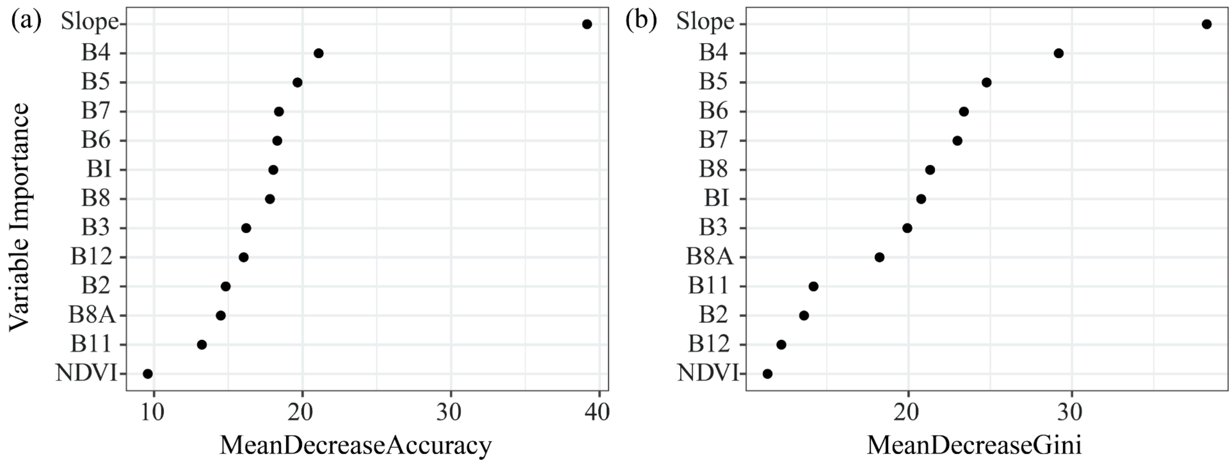

Fig. 6. Variable importance resulting from the random forest classification model.

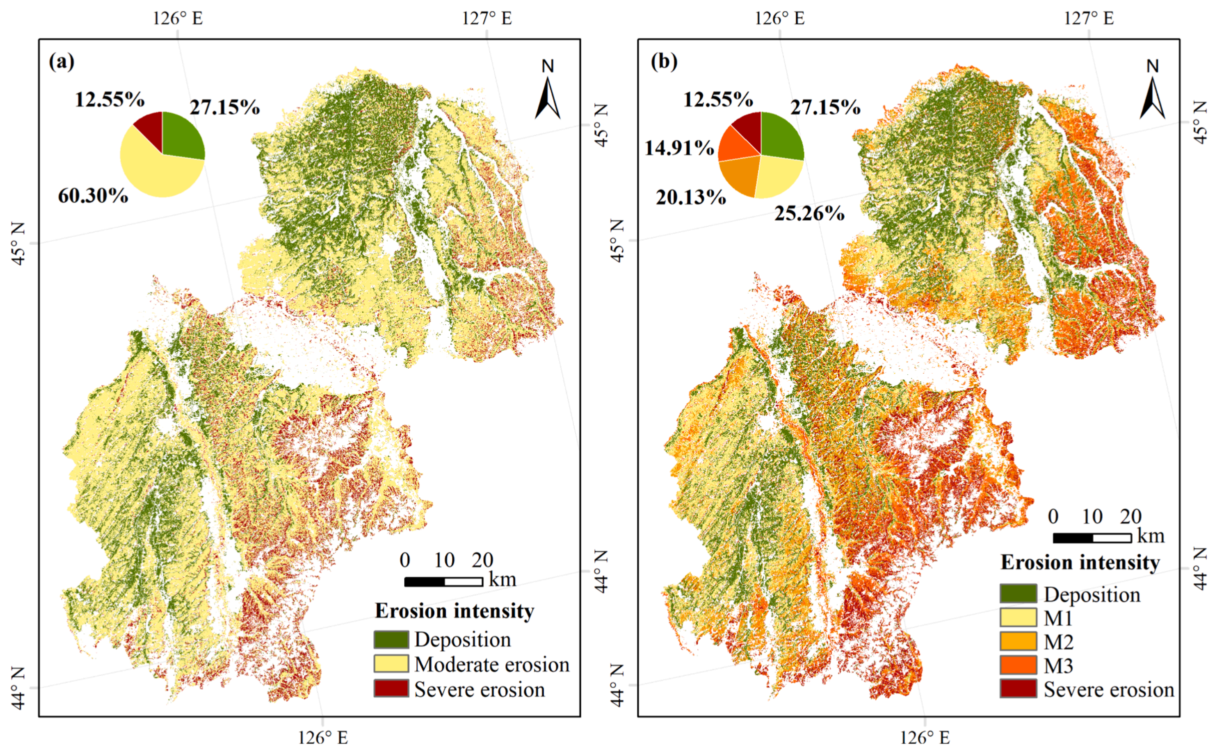

Fig. 7. (a) Three-class soil erosion map at 10 m resolution resulting from the random forest (RF) classification model; (b) Five-class soil erosion map resulting from the combined RF and ISO cluster classification. Pie charts represent the area percentages of various erosion classes.

5. Discussion

5.1. Effectiveness of the two-step soil erosion classification

The key advantage of a spectra-based soil erosion classification was its ability to distinguish severe erosion and deposition areas at a high spatial resolution. This was achieved by using Sentinel-2 bare soil spectral information to differentiate soils with distinct biochemical compositions in relation to erosion-induced lateral soil redistribution (Doetterl et al., 2016; Hu et al., 2020). In the first step, the RF classification model yielded an OA of 91 %, higher than previous spectral imaging studies by Žížala et al. (2017) and Bracken et al. (2019): the former developed four-class erosion maps by fuzzy C-means classification with an OA ranging from 51 % to 82 % depending on the investigated soil types, while the latter employed spectral unmixing to map the fraction of three erosion classes at 70 % OA. The satisfactory performance of the RF classifier developed in this study could be attributed to at least three important aspects: (1) erosion-induced spatial variations in soil properties may be reflected in the spectral signatures of the multi-temporal Sentinel-2 bare soil composite, as already highlighted by numerous studies that compositing multiple bare soil images can increase the representativeness of the extracted bare soil pixels (Gasmi et al., 2021; Rizzo et al., 2020); (2) the addition of slope gradient as a well-documented controlling factor of water and tillage erosion helped improve the classification accuracy (Bryan, 2000; Van Oost et al., 2006); (3) a comprehensive ground-truth dataset, made possible by VHR imagery, allowed for more extensive coverage of the study region, and thus a more robust and accurate classification model. This is particularly useful for large-scale mapping studies, as traditional field investigations and obervations of soil erosion are too resource-demanding to be applied

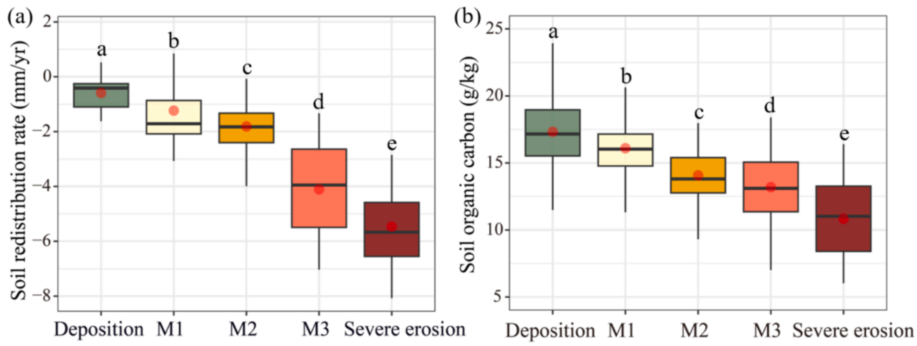

Fig. 8. (a) Comparisons of soil redistribution rates among the five predicted erosion classes at catchment scale; (b) comparisons of soil organic carbon contents among the five predicted erosion classes at regional scale. Soil redistribution rate was estimated from ^{137}Cs inventory data, using the formula proposed by Van Oost et al. (2007).Letters denote significant differences at P < 0.05 using pair-wise Wilcox test.

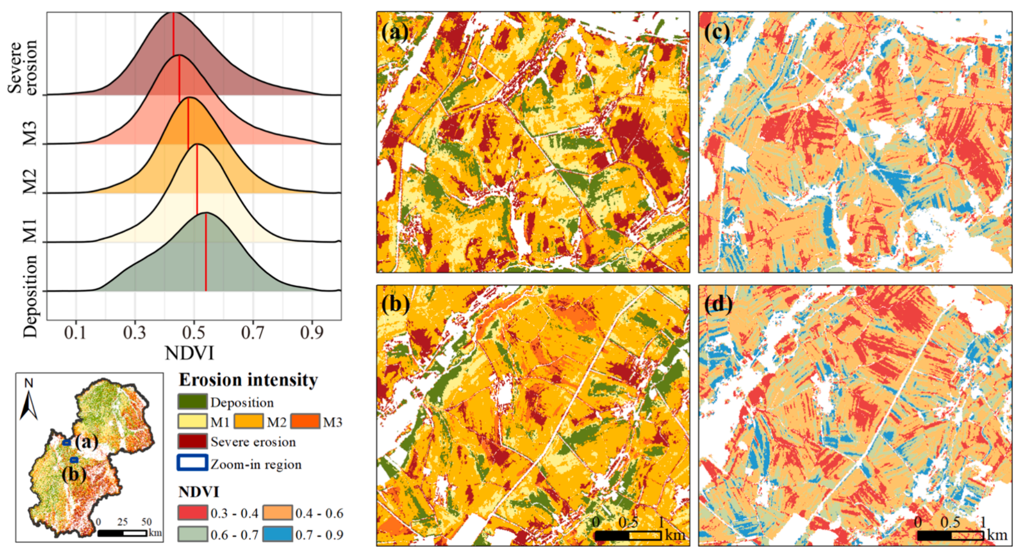

Fig. 9. Comparing NDVI distributions among different erosion classes in the study region, and selected zoom-in regions depicting (a)-(b) localized pattern of soil erosion intensity from the five-class map, and (c)-(d) the NDVI spatial distribution in relation to erosion and deposition hotspots. NDVI data was based on a Sentinel-2 image acquired on June 22, 2021, corresponding to the growth stage of maize.

more broadly (Boardman, 2016; Fischer et al., 2018).

However, the predicted erosion maps resulting from previous studies as well as the RF model used in this study have a common shortcoming, namely the lack of detailed classification of the moderate erosion class that often occupies the majority of the investigated area. More than 60 % of the cropland was classified as moderate erosion by the RF model in this study (Fig. 7), and similar or higher figures have also been reported by others (Bracken et al., 2019; Qi et al., 2023; Žížala et al., 2019). This implies that potential variations within the lumped class of moderate erosion remain largely unaccounted for. One important obstacle that limits a more detailed classification scheme is the difficulty of acquiring sufficient training data to cover the erosion gradient between severe erosion and deposition. The VHR imagery was used to construct a training dataset, primarily by identifying the severe erosion and deposition features that display substantial visual contrast, but areas under the influence of moderate erosion could not be further differentiated based on VHR images alone. This was also reflected in Fig. 5, where relatively wider variations in bare soil spectra were observed in the moderate erosion class.

From this perspective, the ISO cluster unsupervised classification was employed to further separate the moderate erosion class in the second-step. By capitalizing on the dissimilarity matrix of the spectra-based input variables in their multidimensional feature space, the proposed method was proved effective as it ensured a rapid and large-scale classification of moderate erosion into meaningful sub-clusters without the need for training data (Fig. 7). It should be noted that, unlike previous studies that directly applied unsupervised classification against bare soil spectra (Žížala et al., 2019), the two-step classification adopted in this study first ensured the accurate detection of severe erosion and deposition areas. This was necessary because, with the assistance of verifiable ground-truth data and model validation, quantitative knowledge about the extent of erosion hotspots could facilitate a better understanding of erosion-associated carbon losses (Thaler et al., 2021) and more targeted conservation measures to help restore the lost carbon (Amelung et al., 2020).

5.2. Mapping and monitoring of soil erosion

The accuracy of the predicted five-class soil erosion map was successfully validated using the independent ^{137}Cs and SOC datasets at both catchment and regional scales (Fig. 8), as also confirmed by the spatial agreement of crop biomass with soil erosion intensity (Fig. 9). In particular, the erosion maps highlighted the capability of high-resolution Sentinel-2 spectral imaging to unveil soil erosion patterns at a high level of spatial detail. Regrading the ^{137}Cs data, the converted soil redistribution rate from ^{137}Cs inventory offered additional insights, as it assigned each erosion class an actual magnitude of erosion-induced soil displacement. This represents an improvement on previous soil erosion classification research that has typically lacked quantitative assessments. The severe erosion class was found to have an annual average soil redistribution rate of −5.5 mm yr^{-1}, which is slightly higher than that reported in a recent large-scale survey by Wang et al. (2022), who found that the most severely eroded areas in northeast China typically had a soil loss rate of around 4 mm yr^{-1}, similar to that of the M3 class. This difference suggests that, without prior knowledge about the location of severely eroded areas, large-scale soil erosion surveys via radionuclide tracing could lead to underestimation of soil loss rates. This again points to the usefulness of the proposed remote sensing method to accurately detect severely eroded areas.

Furthermore, soil erosion classification based on the spatial heterogeneity of soil surface reflectance implies that the resultant map represents a retrospective assessment on the cumulative effect of multiple erosion forms. As the vast majority of soil erosion models only consider water erosion processes, our proposed approach is particularly relevant for areas where tillage and wind erosion are important (Zhao et al., 2018). To illustrate this, we compared the class-specific soil redistribution estimates from this study with existing RUSLE-type estimates provided by Borrelli et al. (2022) (Fig. 10). The SOC validation data was converted to soil redistribution rate based on the linear relationship given in Fig. 4, which was assumed to be stable across scales given that the climate conditions remained constant (Wang et al., 2023). The comparison showed that our proposed approach was better able to capture net deposition scenarios. More importantly, the estimated soil erosion rates from RUSLE only varied within a narrow range and generally underestimated the magnitude of soil loss for M2, M3, and severe erosion classes, likely due to the aforementioned lack of consideration of other erosion forms. However, the exact class-specific gap between two types of estimates should be viewed with caution because of the uncertainties introduced by intra-class variations in estimated soil redistribution rate and the error propagated from the linear regression relationship.

5.3. Limitations and outlook

In this study, the training data used to constrain the RF classifier was obtained by visually interpreting erosion and deposition features from VHR images, based on the assumption that erosion-induced topsoil loss from Phaeozems and Chernozems, or Mollisols in the USDA soil taxonomy, results in the exposure of light-colored subsoils with depleted SOC. This assumption is supported by previous research in the Czech Republic (Žížala et al., 2019), the United States corn belt (Thaler et al., 2021), and northeast China (Qi et al., 2023). However, future work should address two important limitations associated with this approach. Firstly, direct field observations were only used for model validation, but not for the construction of the training data, because quantifying erosion rates using conventional methods is too resource-demanding to carry out at a large scale. This could hinder the predictive capability of such classification models, especially in regions where soil erosion is not the primary driver of localized changes in soil color and SOC. From this perspective, future research should explore new ways to develop large-scale soil erosion observational datasets for model training, as exemplified by Benaud et al. (2020) who consolidated national scale soil erosion observations into a consistent database. This is particularly relevant for regions with limited availability of VHR images (Ardelean et al., 2020), or where heterogenous soil genesis and management practices can confound the erosional control over soil color and composition. A more comprehensive erosion observation dataset could also help improve the classification scheme, because the class of non-eroded soil was not considered in this study based on the assumption that all cropland soils experience varying degrees of erosion influence.

Second, the general applicability of the spectra-based soil erosion assessment approach needs to be further tested. Soil erosion mapping in black soils, which are ideally suited for this approach because of the strong contrast between the organic horizon and subsoils, has significant global implications, because black soils produce 30 % wheat and 46 % barley worldwide but are under increasing threat of soil degradation (FAO, 2022). Schmid et al. (2016) also demonstrated the effectiveness of spectra-based erosion classification in the Mediterranean environment where severe erosion led to the exposure of calcic horizons. Nonetheless, erosion mapping based purely on soil spectral information may be less powerful in soils with weaker depth gradients of key properties such as SOC, iron oxides and soil texture. One possible way forward could be to incorporate environmental covariates into the spectra-based modelling framework to better represent the diversity of soil biophysical conditions and the impact of soil erosion on vegetation cover (Vågen and Winowiecki, 2019).

Finally, the frequent revisit time of Sentinel-2 satellites can reflect the temporal dynamics of soil surface conditions. In theory, this highlights the potential of using spectral imaging to monitor the spatio-temporal development of eroded areas. In the RF model (Fig. 6), the important contributing spectral bands were the red and red-edge bands (B4-B7), which have been reported to be sensitive to SOC changes (Castaldi et al., 2019). This aligned with the well-documented close association between soil erosion and SOC dynamics particularly in this study region (Hu et al., 2020; Zhu et al., 2021). Considering the prospect that future spaceborne hyperspectral sensors will offer more detailed soil spectral information (Demattê et al., 2015), along with current efforts in digital soil mapping to reduce prediction uncertainty (Dvorakova et al., 2023), the development of spectra-based soil erosion mapping at high spatial resolution could support future endeavors to monitor the proportion of degraded soils and fulfill land degradation neutrality goals (Sims et al. (2021)), to assess how and to what extent soil erosion responds to sustainable cropland management (Martínez-Mena et al., 2020), and to help envisage effective measures that promote carbon storage (Amelung et al., 2020).

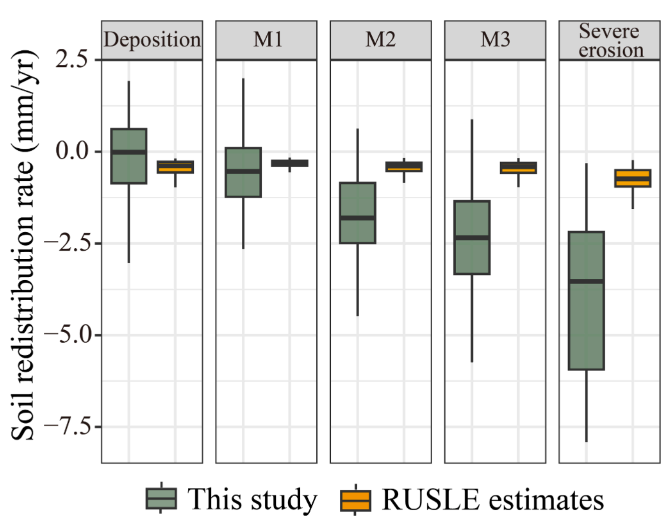

Fig. 10. Comparing the estimated soil redistribution rates of this study to RUSLE estimates using open-Borrelli et al. (2022). Regional SOC data and its linear relationship to soil redistribution, as outlined in , were used to derive the regional-scale estimates (n = 493). The coordinates of the 493 points were then used to extract the RUSLE estimates, assuming a bulk density of 1.2 g cm^{-3}. Source data provided by Fig. 4

6. Conclusions

Remotely sensed bare soil spectra serve as valuable data sources for characterizing the spatial heterogeneity of soil composition in relation to soil erosion processes. This study presented a novel spectra-based, two-step classification approach to performing high-resolution soil erosion mapping in a typical agricultural region (11,500 km^2) of northeast China. Severe erosion and deposition hotspots were accurately detected in the first step by an RF classifier (OA: 91 %) based on ground-truth data derived from VHR imagery in Google Earth. Slope gradient, red and red-edge Sentinel-2 bands were the most important predictors in the RF model, with which more than 60 % of the study area remained in the lumped class of moderate erosion. In the second step, the intra-class variability of moderate erosion was further differentiated via ISO cluster unsupervised classification, thus producing a five-class soil erosion map that revealed the regional soil erosion pattern as well as the distribution of localized hotpots with excessive topsoil loss. The five erosion classes occupied 27.15 %, 25.26 %, 20.13 %, 14.91 %, and 12.55 % of the total cropland area respectively, with the severe erosion class concentrated across the northeast-southeast hilly areas.

The effectiveness of the proposed approach was confirmed through independent catchment and regional scale validations, where significant differences in ^{137}Cs-derived soil redistribution rates existed among the predicted erosion classes. The spatial accordance of crop growth to soil erosion intensity further highlighted the potential of bare soil imaging for mapping the spatiotemporal dynamics of soil erosion, for the purposes of achieving land degradation neutrality and implementing targeted conservation practices to enhance carbon storage.

CRediT authorship contribution statement

Lulu Qi: Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yue Zhou:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kristof Van Oost:** Writing – review & editing, Validation, Supervision, Methodology, Investigation. **Jiamin Ma:** Writing – review & editing, Software, Methodology, Data curation. **Bas van Wesemael:** Writing – review & editing, Supervision, Methodology, Investigation. **Pu Shi:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This research was funded by the Natural Science Foundation of Jilin Province, China (Grant No. 20230101079JC). L. Qi was supported by the China Scholarship Council for a one-year joint Ph.D. study in UC Louvain, Belgium (Grant No. 202106170101). P. Shi is a postdoc fellow of the Fonds de la Recherche Scientifique—FNRS (Belgium).

References

Alewell, C., Borrelli, P., Meusburger, K., Panagos, P., 2019. Using the USLE: Chances, challenges and limitations of soil erosion modelling. Int. Soil Water Conserv. Res. 7 (3), 203–225.
Al-Mamari, M.M., Kantoush, S.A., Al-Harrasi, T.M., Al-Maktoumi, A., Abdrabo, K.I., Saber, M., Sumi, T., 2023. Assessment of sediment yield and deposition in a dry reservoir using field observations, RUSLE and remote sensing: Wadi Assarin. Oman. J. Hydrol. 617, 128982.
Amelung, W., Bossio, D., de Vries, W., Kogel-Knabner, I., Lehmann, J., Amundson, R., Bol, R., Collins, C., Lal, R., Leifeld, J., Minasny, B., Pan, G., Paustian, K., Rumpel, C., Sanderman, J., van Groenigen, J.W., Mooney, S., van Wesemael, B., Wander, M., Chabbi, A., 2020. Towards a global-scale soil climate mitigation strategy. Nat. Commun. 11 (1), 5427.
Ardelean, F., Onaca, A., Chețan, M.-A., Dornik, A., Georgievski, G., Hagemann, S., Timofte, F., Berzescu, O., 2020. Assessment of spatio-temporal landscape changes from VHR images in three different permafrost areas in the Western Russian Arctic. Remote Sens. 12, 3999.
Bai, Z.G., Dent, D.L., Olsson, L., Schaepman, M.E., 2008. Proxy global assessment of land degradation. Soil Use Manag. 24 (3), 223–234.
Benaud, P., Anderson, K., Evans, M., Farrow, L., Glendell, M., James, M.R., Quine, T.A., Quinton, J.N., Rawlins, B., Jane Rickson, R., Brazier, R.E., 2020. National-scale geodata describe widespread accelerated soil erosion. Geoderma 371, 114378.
Benavidez, R., Jackson, B., Maxwell, D., Norton, K., 2018. A review of the (Revised) Universal Soil Loss Equation ((R)USLE): with a view to increasing its global applicability and improving soil loss estimates. Hydrol. Earth Syst. Sci. 22 (11), 6059–6086.
Bhandari, D., Joshi, R., Regmi, R.R., Awasthi, N., 2021. Assessment of soil erosion and its impact on agricultural productivity by using the RMMF model and local perception: A case study of rangun watershed of mid-hills. Nepal. Appl. Environ. Soil Sci. 2021, 5747138.
Boardman, J., 2016. The value of Google Earth™ for erosion mapping. CATENA 143, 123–127.
Borrelli, P., Van Oost, K., Meusburger, K., Alewell, C., Lugato, E., Panagos, P., 2018. A step towards a holistic assessment of soil degradation in Europe: Coupling on-site erosion with sediment transfer and carbon fluxes. Environ. Res. 161, 291–298.
Borrelli, P., Ballabio, C., Yang, J.E., Robinson, D.A., Panagos, P., 2022. GloSEM: High-resolution global estimates of present and future soil displacement in croplands by water erosion. Sci. Data. 9, 406.
Bracken, A., Coburn, C., Staenz, K., Rochdi, N., Segl, K., Chabrillat, S., Schmid, T., 2019. Detecting soil erosion in semi-arid Mediterranean environments using simulated EnMAP data. Geoderma 340, 164–174.
Breiman, L., 2001. Random Forests. Mach. Learn. 45 (1), 5–32.
Bryan, R.B., 2000. Soil erodibility and processes of water erosion on hillslope. Geomorphology 32 (3), 385–415.
Castaldi, F., Chabrillat, S., Don, A., van Wesemael, B., 2019. Soil organic carbon mapping using LUCAS topsoil database and Sentinel-2 data: An approach to reduce soil moisture and crop residue effects. Remote Sens. 11, 2121.
Chabrillat, S., Ben-Dor, E., Cierniewski, J., Gomez, C., Schmid, T., van Wesemael, B., 2019. Imaging Spectroscopy for Soil Mapping and Monitoring. Surv. Geophys. 40 (3), 361–399.
Demattê, J.A., Morgan, C.L., Chabrillat, S., Rizzo, R., Franceschini, M.H., Terra, F.d.S., Vasques, G.M., Wetterlind, J., 2015. Spectral sensing from ground to space in soil science: state of the art, applications, potential, and perspectives. In: P.S. Thenkabail (Ed.), Land resources monitoring, modeling, and mapping with remote sensing. CRC Press, Boca Raton, pp. 661-732.
Demattê, J.A.M., Fongaro, C.T., Rizzo, R., Safanelli, J.L., 2018. Geospatial Soil Sensing System (GEOS3): A powerful data mining procedure to retrieve soil spectral reflectance from satellite images. Remote Sens. Environ. 212, 161–175.
Doetterl, S., Berhe, A.A., Nadeu, E., Wang, Z., Sommer, M., Fiener, P., 2016. Erosion, deposition and soil carbon: A review of process-level controls, experimental tools and models to address C cycling in dynamic landscapes. Earth-Sci. Rev. 154, 102–122.
Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sens. Environ. 120, 25–36.
Dvorakova, K., Heiden, U., Pepers, K., Staats, G., van Os, G., van Wesemael, B., 2023. Improving soil organic carbon predictions from a Sentinel–2 soil composite by assessing surface conditions and uncertainties. Geoderma 429, 116128.
Efthimiou, N., Psomiadis, E., Panagos, P., 2020. Fire severity and soil erosion susceptibility mapping using multi-temporal Earth Observation data: The case of Mati fatal wildfire in Eastern Attica. Greece. Catena 187, 104320.
Escadafal, R., Girard, M.-C., Courault, D., 1989. Munsell soil color and soil reflectance in the visible spectral bands of landsat MSS and TM data. Remote Sens. Environ. 27 (1), 37–46.
Evans, D.L., Quinton, J.N., Davies, J.A.C., Zhao, J., Govers, G., 2020. Soil lifespans and how they can be extended by land use and management change. Environ. Res. Lett. 15, 0940b0942.
Fao, 2022. Global Map of Black Soils. Italy, Rome.
Fischer, F.K., Kistler, M., Brandhuber, R., Maier, H., Treisch, M., Auerswald, K., 2018. Validation of official erosion modelling based on high-resolution radar rain data by aerial photo erosion classification. Earth Surf. Process. Landf. 43 (1), 187–194.
García-Ruiz, J.M., Beguería, S., Nadal-Romero, E., González-Hidalgo, J.C., Lana-Renault, N., Sanjuán, Y., 2015. A meta-analysis of soil erosion rates across the world. Geomorphology 239, 160–173.
Gasmi, A., Gomez, C., Lagacherie, P., Zouari, H., Laamrani, A., Chehbouni, A., 2021. Mean spectral reflectance from bare soil pixels along a Landsat-TM time series to increase both the prediction accuracy of soil clay content and mapping coverage. Geoderma 388, 114864.
Giuliani, G., Mazzetti, P., Santoro, M., Nativi, S., Van Bemmelen, J., Colangeli, G., Lehmann, A., 2020. Knowledge generation using satellite earth observations to support sustainable development goals (SDG): A use case on Land degradation. Int. J. Appl. Earth Obs. Geoinf. 88, 102068.
Heiden, U., d'Angelo, P., Schwind, P., Karlshöfer, P., Müller, R., Zepp, S., Wiesmeier, M., Reinartz, P., 2022. Soil reflectance composites— improved thresholding and performance evaluation. Remote Sens. 14, 4526.
Hu, Y., Li, X., Guo, S., Gao, X., Ou, X., Liu, B., 2020. On-site soil dislocation and localized CNP degradation: the real erosion risk faced by sloped cropland in northeastern China. Agric. Ecosyst. Environ. 302, 107088.

IUSS Working Group WRB, 2015. World reference base for soil resources 2014, update 2015. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome.
Jetten, V., Govers, G., Hessel, R., 2003. Erosion models: quality of spatial predictions. Hydrol. Process. 17 (5), 887–900.
Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26.
Le, Q.B., Nkonya, E., Mirzabaev, A., 2016. Biomass Productivity-Based Mapping of Global Land Degradation Hotspots. In: Nkonya, E., Mirzabaev, A., von Braun, J. (Eds.), Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development. Springer International Publishing, Cham, pp. 55–84.
Li, Y., Zhang, Q.W., Reicosky, D.C., Bai, L.Y., Lindstrom, M.J., Li, L., 2006. Using 137Cs and 210Pbex for quantifying soil organic carbon redistribution affected by intensive tillage on steep slopes. Soil Tillage Res. 86 (2), 176–184.
Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News. 2, 18–22.
Liu, B., Xie, Y., Li, Z., Liang, Y., Zhang, W., Fu, S., Yin, S., Wei, X., Zhang, K., Wang, Z., Liu, Y., Zhao, Y., Guo, Q., 2020. The assessment of soil loss by water erosion in China. Int. Soil Water Conserv. Res. 8 (4), 430–439.
Lorenz, K., Lal, R., Ehlers, K., 2019. Soil organic carbon stock as an indicator for monitoring land and soil degradation in relation to United Nations' Sustainable Development Goals. Land Degrad. Dev. 30 (7), 824–838.
Manić, M., Đorđević, M., Đokić, M., Dragović, R., Kićović, D., Đorđević, D., Jović, M., Smičiklas, I., Dragović, S., 2022. Remote sensing and nuclear techniques for soil erosion research in forest areas: Case study of the Crveni Potok catchment. Front. Environ. Sci. 10, 897248.
Martínez-Mena, M., Carrillo-López, E., Boix-Fayos, C., Almagro, M., García Franco, N., Díaz-Pereira, E., Montoya, I., de Vente, J., 2020. Long-term effectiveness of sustainable land management practices to control runoff, soil erosion, and nutrient loss and the role of rainfall intensity in Mediterranean rainfed agroecosystems. Catena 187, 104352.
Mathieu, R., Cervelle, B., Rémy, D., Pouget, M., 2007. Field-based and spectral indicators for soil erosion mapping in semi-arid mediterranean environments (Coastal Cordillera of central Chile). Earth Surf. Process. Landf. 32 (1), 13–31.
Öttl, L.K., Wilken, F., Auerswald, K., Sommer, M., Wehrhan, M., Fiener, P., 2021. Tillage erosion as an important driver of in-field biomass patterns in an intensively used hummocky landscape. Land Degrad. Dev. 32 (10), 3077–3091.
Poesen, J., 2018. Soil erosion in the Anthropocene: Research needs. Earth Surf. Process. Landf. 43 (1), 64–84.
Pradhan, B., Chaudhari, A., Adinarayana, J., Buchroithner, M.F., 2012. Soil erosion assessment and its correlation with landslide events using remote sensing data and GIS: a case study at Penang Island. Malaysia. Environ. Monit. Assess. 184 (2), 715–727.
Prasuhn, V., Liniger, H., Gisler, S., Herweg, K., Candinas, A., Clément, J.-P., 2013. A high-resolution soil erosion risk map of Switzerland as strategic policy support system. Land Use Policy. 32, 281–291.
Qi, L., Shi, P., Dvorakova, K., Van Oost, K., Sun, Q., Yu, H., van Wesemael, B., 2023. Detection of Soil Erosion Hotspots in the Croplands of a Typical Black Soil Region in Northeast China: Insights from Sentinel-2 Multispectral Remote Sensing. Remote Sens. 15 (5), 1402.
R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Rizzo, R., Medeiros, L.G., Mello, D.C.d., Marques, K.P.P., Mendes, W.D.S., Quiñonez Silvero, N.E., Dotto, A.C., Bonfatti, B.R., Demattê, J.A.M., 2020. Multi-temporal bare surface image associated with transfer functions to support soil classification and mapping in southeastern Brazil. Geoderma 361, 114018.
Roberts, D., Wilford, J., Ghattas, O., 2019. Exposed soil and mineral map of the Australian continent revealing the land at its barest. Nat. Commun. 10 (1), 5297.
Šarapatka, B., Bednář, M., 2021. Agricultural Production on Erosion-Affected Land from the Perspective of Remote Sensing. Agronomy 11 (11), 2216.
Schmid, T., Rodriguez-Rastrero, M., Escribano, P., Palacios-Orueta, A., Ben-Dor, E., Plaza, A., Milewski, R., Huesca, M., Bracken, A., Cicueudez, V., Pelayo, M., Chabrillat, S., 2016. Characterization of Soil Erosion Indicators Using Hyperspectral Data From a Mediterranean Rainfed Cultivated Region. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (2), 845–860.
Sherrod, L.A., Dunn, G., Peterson, G.A., Kolberg, R.L., 2002. Inorganic carbon analysis by modified pressure-calcimeter method. Soil Sci. Soc. Am. J. 66 (1), 299–305.
Shi, P., Six, J., Sila, A., Vanlauwe, B., Van Oost, K., 2022. Towards spatially continuous mapping of soil organic carbon in croplands using multitemporal Sentinel-2 remote sensing. ISPRS J. Photogramm. Remote Sens. 193, 187–199.
Sims, N.C., Newnham, G.J., England, J.R., Guerschman, J., Cox, S.J.D., Roxburgh, S.H., Viscarra Rossel, R.A., Fritz, S., Wheeler, I., 2021. Good Practice Guidance. SDG Indicator 15.3.1, Proportion of Land That Is Degraded Over Total Land Area. Version 2.0. United Nations Convention to Combat Desertification, Bonn, Germany.
Thaler, E.A., Larsen, I.J., Yu, Q., 2021. The extent of soil loss across the US Corn Belt. Proc. Natl. Acad. Sci. 118 (8) e1922375118.
Tziolas, N., Tsakiridis, N., Chabrillat, S., Demattê, J.A.M., Ben-Dor, E., Gholizadeh, A., Zalidis, G., van Wesemael, B., 2021. Earth Observation Data-Driven Cropland Soil Monitoring: A Review. Remote Sens. 13 (21), 4439.
Vagen, T.G., Winowiecki, L.A., 2019. Predicting the spatial distribution and severity of soil erosion in the global tropics using satellite remote sensing. Remote Sens. 11 (15), 1800.
Vågen, T.-G., Winowiecki, L.A., 2019. Predicting the Spatial Distribution and Severity of Soil Erosion in the Global Tropics using Satellite Remote Sensing. Remote Sens. 11 (15), 1800.
Van Oost, K., Govers, G., De Alba, S., Quine, T.A., 2006. Tillage erosion: a review of controlling factors and implications for soil quality. Progress in Physical Geography: Earth and Environment 30 (4), 443–466.
Van Oost, K., Quine, T.A., Govers, G., De Gryze, S., Six, J., Harden, J.W., Ritchie, J.C., McCarty, G.W., Heckrath, G., Kosmas, C., Giraldez, J.V., da Silva, J.R., Merckx, R., 2007. The impact of agricultural soil erosion on the global carbon cycle. Science 318, 626–629.
Vaudour, E., Gholizadeh, A., Castaldi, F., Saberioon, M., Borůvka, L., Urbina-Salazar, D., Fouad, Y., Arrouays, D., Richer-de-Forges, A.C., Biney, J., Wetterlind, J., van Wesemael, B., 2022. Satellite Imagery to Map Topsoil Organic Carbon Content over Cultivated Areas: An Overview. Remote Sens. 14 (12), 2917.
Wang, H., Yang, S., Wang, Y., Gu, Z., Xiong, S., Huang, X., Sun, M., Zhang, S., Guo, L., Cui, J., Tang, Z., Ding, Z., 2022. Rates and causes of black soil erosion in Northeast China. CATENA 214, 106250.
Wang, Z., Zhang, Y., Govers, G., Tang, G., Quine, T.A., Qiu, J., Navas, A., Fang, H., Tan, Q., Van Oost, K., 2023. Temperature effect on erosion-induced disturbances to soil organic carbon cycling. Nat. Clim. Change 13 (2), 174–181.
Wuepper, D., Borrelli, P., Finger, R., 2019. Countries and the global rate of soil erosion. Nat. Sustain. 3 (1), 51–55.
Yan, B., Tang, J., 2004. Study on reference cesium-137 inventory of black soil in Northeast China. J. Soil Water Conserv. 04, 33–36.
Ye, Y., Fang, X., Ren, Y., Zhang, X., Chen, L., 2009. Cropland cover change in Northeast China during the past 300 years. Science in China Series d: Earth Sciences 52 (8), 1172–1182.
Zepp, S., Heiden, U., Bachmann, M., Wiesmeier, M., Steininger, M., van Wesemael, B., 2021. Estimation of soil organic carbon contents in croplands of cavaria from SCMaP soil reflectance composites. Remote Sens. 13 (16), 3141.
Zhang, X.C., Polyakov, V.O., Liu, B.Y., Nearing, M.A., 2019. Quantifying geostatistical properties of 137Cs and 210Pbex at small scales for improving sampling design and soil erosion estimation. Geoderma 334, 155–164.
Zhao, P., Li, S., Wang, E., Chen, X., Deng, J., Zhao, Y., 2018. Tillage erosion and its effect on spatial variations of soil organic carbon in the black soil region of China. Soil Tillage Res. 178, 72–81.
Zhu, Y., Wang, D., Wang, X., Li, W., Shi, P., 2021. Aggregate-associated soil organic carbon dynamics as affected by erosion and deposition along contrasting hillslopes in the Chinese Corn Belt. CATENA 199, 105106.
Žížala, D., Zádorová, T., Kapička, J., 2017. Assessment of soil degradation by erosion based on analysis of soil properties using aerial hyperspectral images and ancillary data. Czech Republic. Remote Sens. 9 (1), 28.
Žížala, D., Juřicová, A., Zádorová, T., Zelenková, K., Minařík, R., 2019. Mapping soil degradation using remote sensing data and ancillary data: South-East Moravia. Czech Republic. Eur. J. Remote Sens. 52 (sup1), 108–122.