Journal of Geomatics Vol. 20, No. 2, October 2026

Assessment and Mapping of Soil Erosion Hotspots in the Himalayan Region Using RUSLE and Machine Learning Technique

Justin George Kalambukattu[1*], Rokade Rachana Rajendra[1], Suresh Kumar[2], Anu D. Raj[1]
[1]Agriculture and Soils Department, Indian Institute of Remote Sensing (IIRS), ISRO, Dehradun, India
[2]Agriculture, Forestry and Ecology Group, Indian Institute of Remote Sensing (IIRS), ISRO, Dehradun, India
*Corresponding Author email: justin@iirs.gov.in; justingeorgek@gmail.com

(Received on 27 May 2026; In final form on 12 September 2026)

DOI: *https://doi.org/10.58825/jog.2026.20.2.333*

Abstract: Soil erosion is a major driver of land degradation, particularly in fragile and ecologically sensitive mountainous regions such as the Himalayas. This study assessed soil erosion risk and mapped erosion hotspots in the Bhagirathi watershed, Uttarkashi District, Uttarakhand, by integrating the Revised Universal Soil Loss Equation (RUSLE), remote sensing, GIS, field observations, and Random Forest (RF) machine learning. The assessment utilized 80 spatially distributed surface soil samples for estimating soil erodibility, ALOS PALSAR DEM (12.5 m) for terrain characterization and derivation of topographic factors, and Sentinel-2A multispectral data (10 m) for land use/land cover and spectral information. Rainfall, soil, topographic, vegetation, and management-related factors were integrated within the RUSLE framework to estimate soil loss, while terrain and spectral variables were subsequently used to develop the RF regression model. Areas experiencing estimated soil loss exceeding 40 t ha^{-1} yr^{-1} were classified as erosion hotspots. The RUSLE-based erosion risk assessment showed good agreement with field observations, achieving an overall validation accuracy of 85%. The RF regression model demonstrated good predictive performance, with an R^2 of 0.72, RMSE of 0.31, and Nash–Sutcliffe Efficiency of 0.67. Slope, aspect and NDVI were found to be the most important predictors governing the model performance as well as spatial distribution of erosion hotspots. Erosion hotspots occupied approximately 9.34% of the watershed and were predominantly associated with steep slopes and areas with sparse or absent vegetation. These locations were also clearly identifiable in high-resolution satellite imagery. The findings demonstrate that integrating empirical erosion modelling with field observations, geospatial data, and machine learning can effectively assess erosion risk and identify priority hotspots in complex and inaccessible Himalayan terrain. The approach can support targeted soil and water conservation planning and sustainable land management in fragile mountain ecosystems.

Keywords: Soil erosion hotspots, RUSLE model, Terrain analysis, Remote sensing, Machine learning, Random forest, Himalayas

1. Introduction

Land degradation is an important environmental problem, with soil erosion being one of its major driving processes (Quinton and Fiener, 2024). In India, approximately 5.3 billion tonnes of soil are estimated to be lost annually through erosion, posing significant threats to soil fertility, agricultural productivity, and environmental sustainability (Bhattacharyya et al., 2015). In the Himalayan region, land degradation and soil erosion are strongly influenced by steep and complex topography, collapsing of manmade terraces, landsliding, land-use changes, vegetation cover, and anthropogenic disturbances (David Raj et al., 2024). Himalayan region are reported to exhibit extraordinarily significant soil erosion rates, ranging from 20 to 25 t ha^{-1} yr^{-1} to as high as 92 t ha^{-1} yr^{-1} (Sharda and Ojasvi, 2016). However, monitoring soil erosion in this region remains a tough task to researchers and policymakers, due to the predominantly inaccessible nature of terrain and a lack of comprehensive knowledge on soil erosion variables in the hilly and mountainous landscapes (Kalambukattu et al. 2021). Accurately identifying soil erosion-prone areas in detail has become essential before planning and implementing effective soil and water conservation measures at the watershed level, as these interventions are vital for preventing soil loss (Phinzi et al. 2021).

Various erosion models including the empirical models such as USLE (Universal Soil Loss Equation), RUSLE (Revised Universal Soil Loss Equation) and MUSLE (Modified Universal Soil Loss Equation); semi-empirical models like AnnAGNPS (Annualized Agricultural Non-Point Source), SWAT (Soil and Water Assessment Tool) and process based models like WEPP (Watershed Erosion Prediction Project) have been utilised in a variety of ways to predict soil erosion rates and erosion risk at various scales, even though they differ in terms of the processes studied, complexity, as well as the data requirements (Karydas et al., 2014). Among various models, RUSLE a simple empirically based model that predicts soil loss from rain-induced erosion based on factors namely rainfall erosivity, soil erodibility, slope length and slope steepness, cover management and support practice has widely been used for erosion risk assessment spatially at various climatic regions and geographical locations globally, owing to their moderate data requirement, robust structure, ease of understanding and enhanced integration with Remote Sensing (RS) and Geographic Information Systems (GIS). (Shrestha 2009 ; Gupta 2015; Kalambukattu et al. 2021; Ghosh et al. 2023). Although, delineation of soil erosion risk areas can be done using a variety of remote sensing and other measurement techniques, combining RUSLE with remotely sensed data

© Indian Society of Geomatics

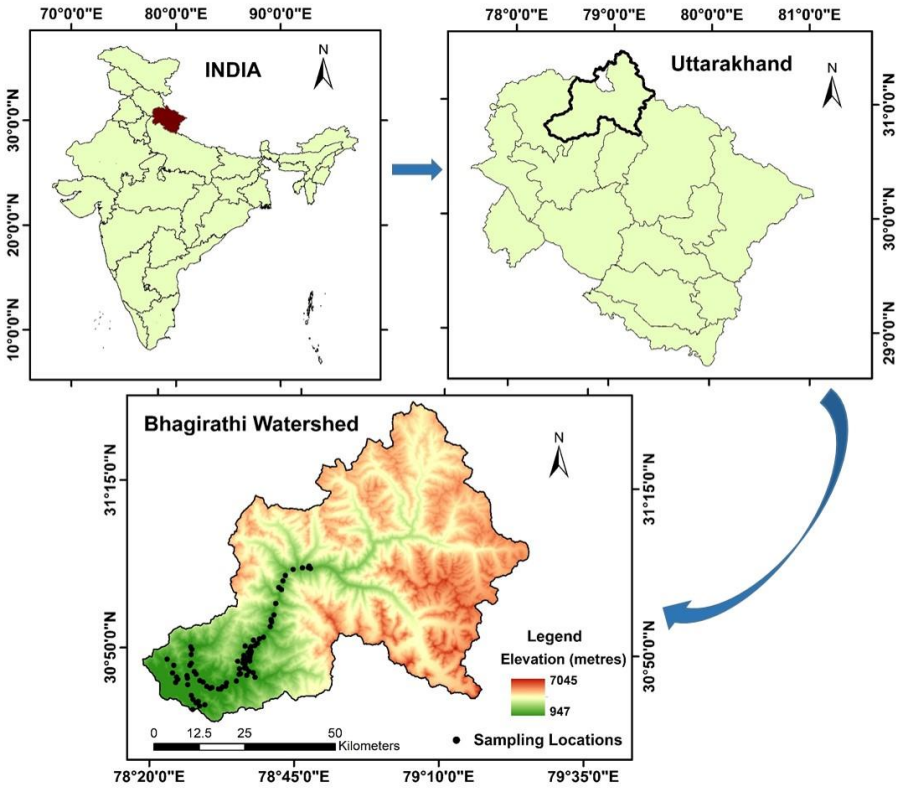


Figure 1: Location of the study area. The sampling locations are overlaid on the elevation values

and GIS can produce time and cost-effective erosion evaluations especially in case of large watersheds as well as areas located in hilly mountainous regions, were accessibility is seriously restricted.

Increased availability of temporal remote sensing images, cloud computing platforms as well as Digital elevaltion models (DEMs) enabled better characterisation of landscape features, land use land cover, topographical features, which are crucial for accurate erosion risk assessment in difficult terrains of hilly regions. Recently there have been increased attempts globally, to apply various machine learning (ML) techniques for erosion assessment and risk mapping due to their ability to handle numerous multisource variables, complex relationships as well as diverse landscapes (Goldblatt et al. 2016). Arabameri et al (2020) did comparative evaluation and reported the applicability of different ML techniques for mapping the susceptibility to gully erosion, using various remote sensing derived environmental variables at watershed scale in hilly mountainous region of Northern Iran. Similarly applicability of RF based modelling for identifying erosion controlling factors and gully erosion suceptible areas by integrating integrating field data with high resolution RS derived spatial inputs has been reported from different watersheds across contrasting agro-ecological environments in the Upper Blue Nile basin of Ethiopia (Setargie et al., 2023). Similar use of RF technique for mapping the spatial distribution of erosion risk classes using multisource data inputs was reported from the Western Zhejiang Province located in the mountainous region of China (Cheng et al., 2018). Applicability of differnet ML techniques and their effective performance for identifying soil erosion susceptible zones utilising various RS and GIS products reperesting erosion controlling factors and the was also reported from a topical sub-humid watershed in India (Bag et al., 2022)

However, attempts combining ML techniques with empirical models like RUSLE, to overcome their shortcomings such as non-availability of soil erosion data, is still in the nascent stage eventhough limited studies have been reported from different regions. Successful delineation of the spatial patterns of soil erosion combining RUSLE parameters and RF algorithm was reported from a catchment scale study conducted in the Eastern Cape province of South Africa (Phinzi et al., 2021). Avand et al (2023) also reported the integration of empirical erosion model and different ML models using various spatial datasets for computation of erosion risk as well as mapping their spatial distribution in hilly mountainous region of Iran. However, the attempts to spatially map erosion hotspots in the rugged challenging Himalayan region, using ML techniques alone or in combination with empirical models is rarely reported. With this background, this study was conducted with the objective of identifying soil erosion risk zones and mapping erosion hotspots by integrating RUSLE model and RF model based ML technique in the Bhagirathi Watershed located in the Uttarkahsi district of Uttarakhand, India. This innovative approach integrating components of empirical and machine learning models is expected to advance the adoption of geospatial techniques for soil erosion risk mapping in the rugged hilly terrains and aid us in making informed decisions towards planning and implementation of sustainable environment management practices.

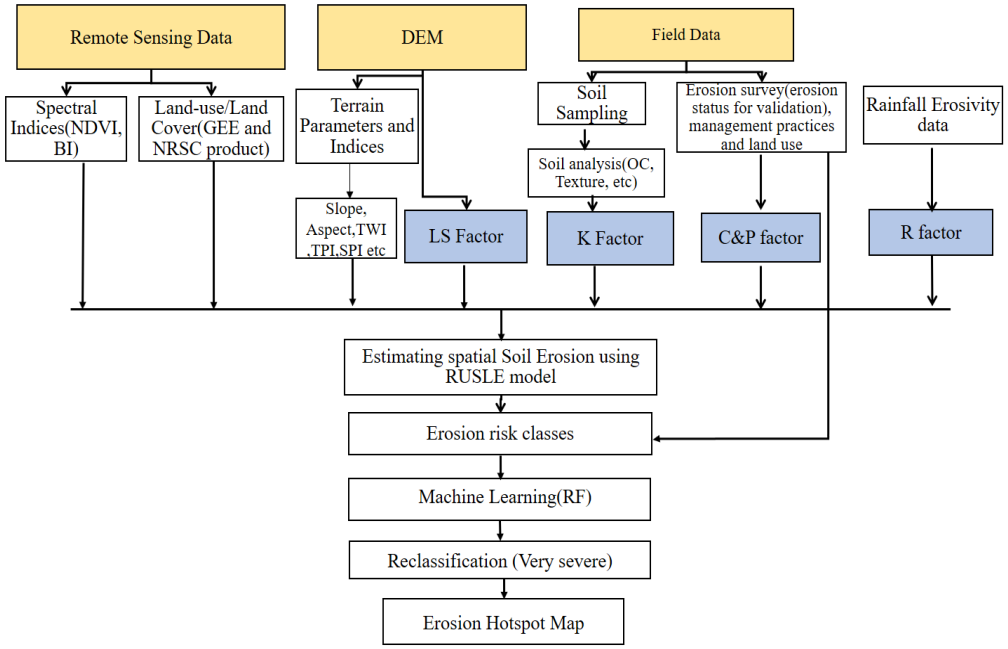


Figure 2: Overall methodology adopted

2. Materials and Methods

2.1 Study area

The Bhagirathi watershed, primarily located in Uttarkashi district, Uttarakhand, with a small portion extending into Himachal Pradesh, was selected as the study area. The watershed encompasses the Bhagirathi River basin up to Maneri Bhali II and covers approximately 4,693 km². and lies between 30°65' to 31°45'N latitudes and 78° 51' to 79° 07' E longitudes. It lies between 30°65′–31°45′ N latitude and 78°51′–79°07′ E longitude, with elevations ranging from 937 to 7,045 m above mean sea level. The region receives a mean annual rainfall of approximately 1,241.8 mm and experiences pronounced seasonal climatic variations, resulting in substantial fluctuations in streamflow. The presence of glaciers, large elevational gradients, and variations in valley aspect contribute to considerable microclimatic heterogeneity. The glacial and paraglacial zones are characterized by active sediment mobilization, fluvial incision, heavy precipitation, and localized slope instability. The landscape comprises snow-clad mountain ranges, broad valleys, deep gorges, and steep spurs. Major land use/land cover types include agriculture, forests, urban areas, barren and scrublands, snow cover, and water bodies. Agriculture is predominantly practiced on terraced slopes, with rice, wheat, potato, pea, rajma, madua, pulses, and seasonal vegetables as major crops. Apple plantations and fallow agricultural lands are also present. The major forest types include deodar, pine, mixed forests, and open and dense scrublands. The dominant soil textures range from loam and sandy loam to silty loam, while soils under forest vegetation generally exhibit relatively high organic carbon content. Figure 1 shows the location of the study area and sampling locations.

2.2 Data and Software Used

Different remote sensing data, derived products and indices were used in the study along with lab generated soil property database for study area characterisation, soil erosion modelling as well as erosion hotspot mapping using ML based technique. The ALOS PALSAR Digital Elevation Model (DEM) of 12.5m spatial resolution obtained from https://asf.alaska.edu/datasets/daac/alos-palsar/ was used for topographical characterisation of study area and generation of various terrain parameters using digital terrain analysis. Sentinel-2 RS data of 10 m resolution from Google Earth Engine (GEE) was used in conjunction with 1:50K Land Use Land Cover (LULC) map from ISRO-BHUVAN portal (https://bhuvan.nrsc.gov.in) for LULC characterisation of the study area. The LULC map and terrain indices were used for the generation of soil-physiographic unit map. In addition, temporal (05 years: 2018-2022) Sentinel-2 images were accessed and processed using GEE platform for generation of long term spectral indices i.e., Normalized Difference Vegetation Index (NDVI) and Brightness Index (BI), respectively. Long term mean values of NDVI and BI represent the peak vegetation cover and fallow periods, respectively and are known to provide qualitative assessment of erosion status (El Jazouli et al. 2017). The global rainfall erosivity product of ~1 km resolution, hosted by European Soil Data Centre (ESDAC) was obtained from http://esdac.jrc.ec.europa.eu for computation of erosivity values within the study area. The erosivity product was generated using rainfall erosivity estimates from 3625 stations across different climatic zones spread across 63 nations globally and contained erosivity estimates from 247 India Meteorological Department (IMD) affiliated stations spread across the country (Panagos et al., 2017). Soil property database generated by laboratory analysis of field collected samples were used for the generation of soil erodibility (K) factor,

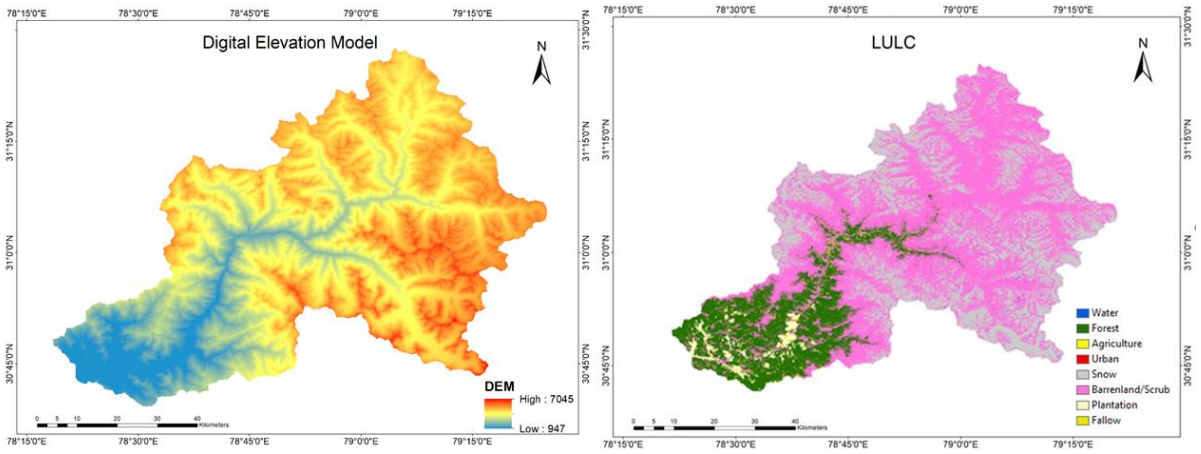


Figure 3: Elevation and LULC map of the area

a vital inputs of RUSLE model. ArcGIS 10.0 and SAGA GIS software were used for digital terrain analysis as well as generation of various output images generation and R software (R 4.2.0) was used for the implementation of ML model.

2.3 Field survey, sampling and soil analysis

Physiogrpahic/soil –landscape unit based stratified random sampling was mainly adopted for collection of soil samples in the area, to ensure proper representation of the existing variability. Considering the rugged terrain and inaccessible nature of many area, transects were also used for covering different locations. Surface soil samples (0-15cm) from 80 different georeferenced locations representing different LULC types and soil-landscape units were collected following standard sampling procedures during the field survey conducted in October – 2022. During sampling, areas with limited access such as areas under snow cover and no road connectivity, could not be sampled extensively. Also caution was exercised to avoid sampling from areas adjacent to rivers, roads, settlements as well as recently manured or fertilized fields, to ensure proper sample collection. In addition to soil samples, details related to presence and characteristics of erosion features as well as management practices via erosion survey were also collected from each sampling location to determine the past erosion status/class of the area along and for estimation of soil erosion risk rates. The samples were transported, air-dried, pre-processed and analysed in the central analytical laboratory of IIRS Dehradun, for estimation of different soil parameters namely, pH, electrical conductivity (EC), organic carbon(OC) and textural composition (%s of sand, silt and clay) following standard analytical procedures. pH and EC were analysed using pH and EC meters respectively (Burt, 1992). Organic carbon (OC) was estimated by wet-oxidation method (Walkley and Black, 1934), while textural composition was estimated by using hydrometer method (Bouyoucos, 1962).

The overall methodology adopted in the study is depicted in Figure 2.

2.4 Study area Characterization

2.4.1 Land Use land cover

The soil erosion rates and risk are highly determined to the nature of land cover and vegetation types existing on the surface. Sentinel-2 RS data of 10 m resolution was accessed and processed using Google earth engine (GEE) cloud computing platform employing RF classification algorithm to generate the current broad level land use land cover (LULC) map of the area. Six LULC classes namely forest, agriculture, urban, barrenland /scrub, snow and waterbodies were classified and mapped using Sentinel 2 data and training samples. The 1:50K LULC map from ISRO-BHUVAN portal (https://bhuvan.nrsc.gov.in) in the WMS layer was digitized for identification of areas under plantation and fallow lands. The outputs from both sources were combined to generate the final LULC map of the watershed.

2.4.2 Terrain Characterisation and soil-landscape map generation

Terrain plays a crucial role in determining the impact of rainfall and runoff on soil hydrological processes and erosion risk. Elevation, slope, aspect, curvature, Stream Power lndex (SPl), Terrain Wetness lndex (TWl) and Terrain Position lndex (TPl) were generated using the ALOS DEM employing ArcGIS software. Slope is the first derivative of DEM and is known to affect the erosion rates. Aspect indicates the direction in which the slope is facing, with south facing slopes found to have drier less developed shallow soils which support less vegetation cover compared to northern slopes. Curvature indicates the concave or convex nature of the terrain. It is known to influence the convergence/divergence and acceleration/deceleration patterns of runoff, which plays a vital role in erosion and deposition of sediments across the landscape. The possibility of flow erosion at specified location on the topographic surface can be described using the SPI. The Equation 1 was used for the computation of SPI. The volume of the water supplied by upslope areas and rate of water flow increase with catchment area and slope gradient, leading to an increase in SPI value and erosion risk. TWI is used to measure topographic controls on hydrological processes and was generated using equation 2. It depicts the region or specific site where soil

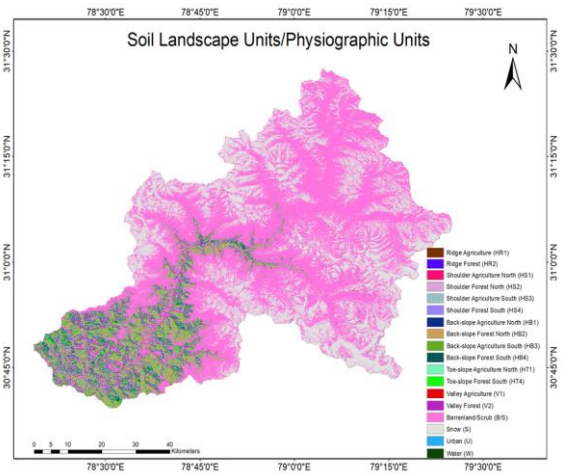


Figure 4: Soil-landscape map of the area

water is distributed and the surface is saturated. Catchment areas with the same TWI values are believed to have the same hydrological progression to rainfall when other environmental conditions are the same or can be viewed as being the same. TPI, a measure of relative positon of a pixel with respect to its neighbours and is calculated by comparing each pixel's elevation to that of its surrounding pixels using neighbourhood analysis. Local high values indicating peaks are represented as positive values, whereas local lows or depressions are represented as negative values. TPI could be used for identification of different landforms or hillslope elements such as flat plain, hilltop, valley bottom, exposed ridge, higher slope, lower slope etc by employing Eq 3. In the study, inner and outer radii values were specified as 2 and 22 m, respectively for landform delineation.

$$\boldsymbol{SPI = ln\,(As * tan\beta)} \qquad \text{Eq: 1}$$

$$\boldsymbol{TWI = ln\left(\frac{As}{tan\beta}\right)} \qquad \text{Eq: 2}$$

$$\boldsymbol{TPI = int((DE - Focal\ Mean(Dem, Annulus, i\ rad, o\ rad) + 0.5))} \qquad \text{Eq: 3}$$

Where, As: Upstream contributing area in m2; β = Local slope gradient; DEM: input elevation data; Annulus – neighbourhood; i rad—interior radius of the annulus; and o rad—exterior radius of the annulus.

In the study area, which has a very steep slope, the annulus neighbourhood proved highly helpful in classifying the different types of landforms. The ERDAS imagine model builder was used to merge TPI and aspect maps, an aspect-based landform category map was created. This was further combined with LULC map of the area to generate the final physiographic/soil Landscape Unit map of the area. Each unit of the final map represents a unique combination of the landform, slope aspect, as well as LULC representing the major pedogenic factors responsible for soil variation.

2.4.3 Spectral Indices

Sentinel-2 images of October and May months were accessed and processed using GEE platform for generation of long term monthly mean values of spectral indices i.e., Normalized Difference Vegetation Index (NDVI) and Brightness Index (BI), respectively. Long term mean values of NDVI and BI represent the peak vegetation cover and fallow periods, respectively and are known to provide qualitative assessment of erosion status (El Jazouli et al. 2017). These two indices are extensively used for soil erosion risk studies and modelling due to their ability to capture variations in vegetation health and cover density.

2.5 Soil erosion modelling using Revised Universal Soil Loss Equation (RUSLE)

RUSLE is the most widely employed empirical model for calculating soil erosion. Rainfall, soil characteristics, topography and slope, land cover, and support practices are five parameters used in the RUSLE model, which have been widely utilized for estimating erosion rates (Mohammed et al. 2020). The RSULE model is expressed as the equation given below:

$$\boldsymbol{A = R * K * LS * C * P} \qquad \text{Eq: 4}$$

where, A = average annual soil loss ($t\ ha^{-1}\ yr^{-1}$), R = rainfall erosivity factor (MJ mm^{-1} ha^{-1} h^{-1} yr), K = soil erodibility factor (t ha h ha^{-1} MJ^{-1} mm^{-1}), LS= slope length and slope steepness factor (dimensionless), C = cover management factor (dimensionless), P = soil conservation or support practices factor (dimensionless).

2.5.1 R Factor (Rainfall Erosivity)

Rainfall Erosivity represents the ability of rainfall to erode soil off hill slopes. It makes an estimate of how well a downpour can separate aggregates and transport soil particles. The intensity, duration and volume of the precipitation all have a significant impact on the kinetic energy of the rain, which in turn determines how erosive it is. The intensity and kinetic energy of rainfall are measured by the R-factor, a multi-annual average indicator, to ascertain how it influences sheet and rill erosion. In the current study, global rainfall erosivity product of European Soil Data Centre (ESDAC), as discussed in *Data and software section* was used for computation of erosivity values. Erosivity values corresponding to sampling locations were retrieved using a 3*3 window, and used further for modelling erosion risk rates.

2.5.2 K Factor (Soil Erodibility)

The inherent ability of a soil to resist or permit erosion is quantified by the K factor. It evaluates how easily soil particles will separate and move during precipitation and runoff. Different soil characteristics such as primary particle size, porosity, permeability, infiltration rate, structural stability, soil organic matter, etc., influence how quickly or slowly different soils erode. K factor value at all the sampling locations were computed using the following equation (Eq 5) given by Sharpley et al. (1990), using the soil parameters estimated during laboratory analysis. The equation was successfully used for soil erosion risk modelling and mapping in different studies reported from Himalayan region such as Tibet (Yang et al., 2018) as well as entire Uttarakhand (Kalambukattu et al., 2021).

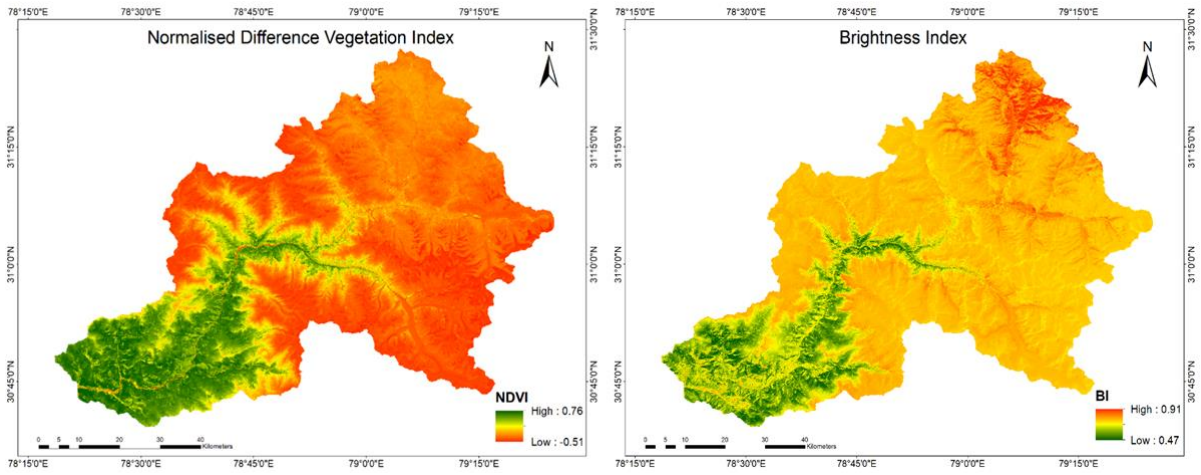


Figure 5: Long term spectral indices (NDVI and BI) of the area

$$K = 0.1317 * (0.2 + 0.3 * e^{\left(-0.0256*SAN\left(1-\frac{SIL}{100}\right)\right)} * \left(\frac{SIL}{CLA+SIL}\right)^{0.3} * \left[1-\frac{0.25*TOC}{TOC+e^{(3.72-2.95*TOC)}}\right] * \left[1-\frac{0.7*SN1}{SN1+e^{(22.9*SN1-5.51)}}\right] \quad \text{Eq: 5}$$

Where, K = soil erodibility factor (t ha h ha^{-1} MJ^{-1} mm^{-1}), SAN = content of sand (%), SIL is content of silt (%), CLA = content of clay (%), TOC = total content of Soil Organic Carbon (%), SN1=1 – (SAN (%) / 100).

2.5.3 LS Factor (Slope Length and Steepness Factor)

In hilly and mountainous areas, topography plays the most prominent role in determining erosion rates. The LS Factor indicates how much erosion is caused by the uninterrupted runoff flow length (slope's length) and slope steepness. Unit-contributing area based concept (Desmet and Govers, 1996) was used for the generation of L factor (12.5m grid size) using SAGA GIS, while equations by McCool et al. (1987) was employed for computation of S factor values (Eq 7 & 8).

$$L_{ij} = \frac{(A_{i,j-in}+D^2)^{m+1} - A_{i,j-in}^{m+1}}{D^{m+2} * x_{i,j}^{m} * 22.13^{m}} \quad \text{Eq: 6}$$

Where, $A_{i,j-in}$ = contributing area at the inlet of grid cell (i,j) measured in m^2. D = grid cell size in meters, $x_{i,j}^{m} = \sin a_{i,j} + \cos a_{i,j}$; $a_{i,j}$ = aspect direction of the grid cell (i,j) and m is related to the ratio of rill to inter rill erosion.

S factor equations:

- For milder slopes (< 9%) :

  $S = 10.8 \cdot \sin\theta + 0.03$ Eq: 7

- For steeper slopes (≥9%):

  $S = 16.8 \cdot \sin\theta - 0.50$ Eq: 8

L ad S factors were integrated to produce LS factor map of the area. LS factor values corresponding to sampling locations were retrieved using a 3*3 window, and used further for modelling erosion risk rates.

2.5.4 C Factor (Cover Management Factor)

The C factor encompasses the combined impact of type of crops, rotation of crops etc, as well as the impact of different land covers like forest canopy, grass cover in non-agricultural areas, on decreasing the rates of soil erosion. The values for different land cover types were identified from previous reported literatures (Panagos et al. 2015, Benavidez et al. 2018, Kebede et al. 2021). Field collected information related to major kharif crops cultivated in the area was considered while fixing the values for agricultural areas, whereas the density of forest and scrub types were considered for assigning C factor values for the respective land cover types. The vegetation canopy and ground coverage can lessen soil erosion depending on the season and crop production strategy.

2.5.5 P Factor (Support practice Factor)

The P factor encompasses all management techniques that decrease the rate and amount of runoff and subsequently soil erosion rates, such as grass bunding, stone patching, bunding, terracing etc (Ganasri and Ramesh 2016). The field collected information related to different management practices adopted under various LULC types and cropping systems within area was used for assigning literature identified P values, for different sampling locations (Ebabu et al. 2022).

Finally, the soil erosion risk at different locations (80 nos) were computed employing RUSLE model by combining the location specific values of different model parameters, whose generation has been discussed in the previous sections. The model computed erosion rates were then validated against the erosion classes (low to very severe erosion) assigned for each sampling location during field survey, based on the observed past erosion features, land cover types, land management etc. (Morgan, 2005).

For validation, the RUSLE computed erosion rates were converted to different erosion classes namely low (e1), moderate (e2), severe (e3) and very severe (e4) corresponding to erosion rates of <10 tha^{-1}yr^{-1}, 10-20 tha^{-1}yr^{-1}, 20 - 40 tha^{-1}yr^{-1} and >40 tha^{-1}yr^{-1}, respectively (Morgan, 2005). Areas of extremely high erosion risk (erosion rates of >40 tha^{-1}yr^{-1}) were considered as erosion hotspots.

2.6 Machine Learning Technique

Random forest (RF) regression based ML technique was used for hotspot identification in the study. It is a non-parametric supervised learning approach for regression that employs ensemble learning (Mohammady and Davudirad, 2024). This model is known for its robustness, resistance to overfitting data handling abilities etc. and hence is widely used. The overall prediction or model output is an ensemble combination of outputs from individual decision trees fit during the process. It also gives an account of the relative importance of different input variables in determining the final results. The various spectral and terrain parameters, know to govern erosion risk in hilly terrains were used as input variables for the model (Aspect, Slope, Elevation, Curvature, TPI, TWI, SPI, BI and NDV). The input variables were tested for their multicollinearity using Variance Inflation Factor (VIF) analysis before model development. The dataset was randomly split in the ratio of 70:30 and used as calibration and validation of the model, respectively (Arabameri et al. 2020). The RF model was implemented using *randomforest* package available in R software (R 4.2.0).

2.6.1 Model Evaluation

Performance evaluation of the model was done using the observed and model predicted values employing different parameters namely, coefficient of determination (R^2), Root Mean Square Error (RMSE) and Nash Sutcliffe model Efficiency coefficient (NSE). R^2 ranges from 0 to 1 and describes the proportion of model explained variance, whereas RMSE indicates the average magnitude of residuals. NSE indicates the how effectively the observed and predicted values align with respect to 1:1 line.

3. Results and Discussion

3.1 Terrain and Land cover characterization

Elevation within the study area ranged from 947m to 7045m above the mean sea level (Figure 3). Slope values within the area were classified into six classes namely gentle (0-5%), undulating (5-10%), moderate (10 - 25%), moderately steep (25 – 33%), steep (33 – 50%) and very steep slope (>50%) (modified from FAO, 1990 and FAO, 2006). Most of the area in the watershed was as having steep slope (34.17%) followed by Moderate slope (27.49%), Moderately steep slope (22.35%), Very steep slope (7.61%), Undulating slope (5.41%) and Gentle slope (2.93%). On analysis of slope aspect, it was observed that a greater proportion of agricultural and cultivation related activities are located in the south aspect because of the higher solar radiation and better temperature conditions. Whereas most of the north facing slopes were primarily covered with dense forests as well as snow cover. Curvature, the second derivative of slope indicating the rate of change, ranged from -91.52 to 90.88 indicating the convex and concave nature of slope, thus influencing the runoff patterns. SPI values ranged from -13.81 to 16.78, with negative values indicating promising areas for deposition and positive values indicating places with high erosion potential. Highest SPI values were correlated with regions having steep slope gradients. TPI values ranged from -220.637 to 252.437 and were classified into different hillslope elements such as valley, toe slope, back slope, shoulder, and ridge. The positive TPI scores show areas that are greater above the neighbourhood average for their surrounds (ridges). Locations with negative TPI levels are those that are inferior to their surroundings (valleys). TPI values close to 0 represent either flat regions (where the slope is close to 0) or regions with a constant slope. TWI values ranged from 0.45 to 26.63 in the watershed. The positive values of TWI indicate the saturated zones or depositional areas, whereas the negative values are known to indicate drier locations in the watershed. The TPI based landform elements were combined with prominent aspect classes of north and south to generate seven different classes namely ridge, shoulder (North), shoulder (South), back slope (North), back slope (South), toe slope, and valley (Kumar et al., 2023). Among the different landforms toeslope covered the 29.58% of the area followed by Backslope N (24.36%), Shoulder N (11.88%), Backslope S (11.66%), valley (9.81%), Shoulder S (6.43%), and Ridge (6.27%). LULC map revealed that nearly 44.28% of the watershed is covered by barrenland/scrub, followed by snow cover (27 %) and forest (16.9%). The rest 12 % of the study area is mainly occupied by agricultural land comprising of both croplands, fallow lands as well as plantation areas. The combined landform map was integrated with the LULC map to produce the final physiography/soil-landscape unit map (Figure 4). Physiographic map represents unique combinations of landform, topography and land cover, which are known to influence the variability of soil properties. The area comprised of 18 soil-landscape units of which barrenland/scrub Landscape Units occupied the highest area. The percentage distribution of area under different soil landscape units could be seen in Table 1.

Table 1: Soil-physiographic units and their percentage area coverage

Soil Landscape Units	Area (%)
Ridge Agriculture (HR1)	0.02
Ridge Forest (HR2)	0.84
Shoulder Agriculture North (HS1)	0.29
Shoulder Forest North (HS2)	2.73
Shoulder Agriculture South (HS3)	0.98
Shoulder Forest South (HS4)	1.12
Back-slope Agriculture North (HB1)	1.34
Back-slope Forest North (HB2)	4.2
Back-slope Agriculture South (HB3)	1.81
Back-slope Forest South (HB4)	1.51
Toe-slope Agriculture North (HT1)	2.06
Toe-slope Forest South (HT2)	4.41
Valley Agriculture (V1)	4.01
Valley Forest (V2)	2.09
Barrenland/Scrub (R/S)	44.28
Snow(S)	27.05
Urban(U)	0.07
Water(W)	0.16

3.2 Spectral Indices

Long term NDVI of 5 years (2018-2022) for the month of October ranged its values from -0.688 to 0.86 with a mean value of 0.19, within the area (Figure 5). The positive

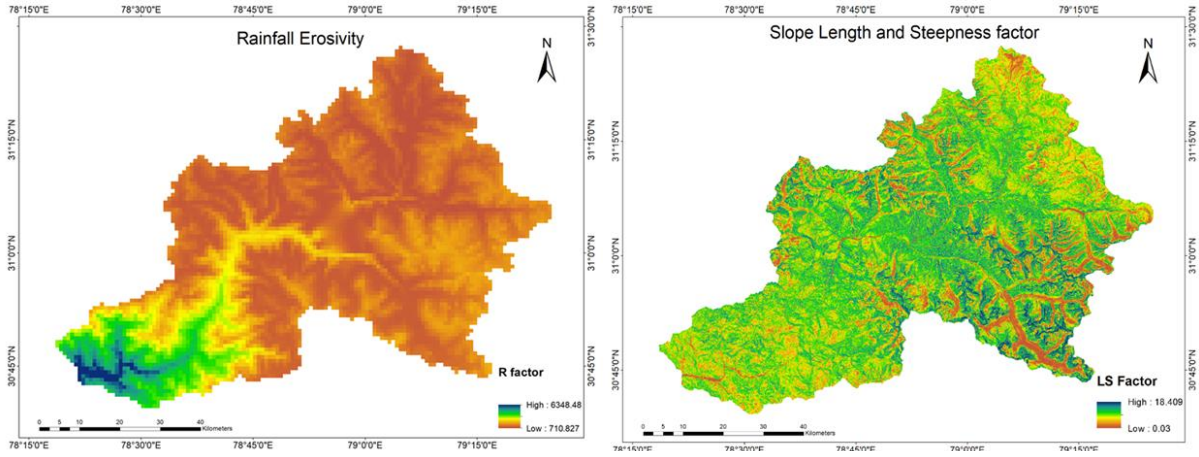


Figure 6: Rainfall erosivity and slope length –steepness factors in the area

values of the NDVI indicate the presence of healthy and green vegetation. Whereas, the negative values indicate that barren land or presence of water. Long term BI of 5 years (2018-2022) for the month of May generated varied were observed to be in the range 0.47 – 0.91, with a mean value of 0.72 (Figure 5). The larger values indicated higher soil brightness may be due to reduced vegetation cover, fallow period as well as exposure of subsurface layers or parent materials due top past erosion processes.

3.3 RUSLE model parameters

The rainfall erosivity values within the area ranged from 710.83 to 6348.48 MJ mm^{-1} ha^{-1} h^{-1} yr with a mean value of 1520.62 MJ mm^{-1} ha^{-1} h^{-1} yr (Figure 6). Lower erosivity values were found towards the higher elevation regions which are prominently under snow cover, towards the northern portion of the area. Whereas, higher values were observed towards the southern portion of the study area around Maneri Bhali II, as well as in the valley portions. The K factor values of different sampling locations, generated using the laboratory estimated soil properties' varied from 0.004 t ha h ha^{-1} MJ^{-1} mm^{-1} to 0.046 t ha h ha^{-1} MJ^{-1} mm^{-1}, with a mean value of 0.020 t ha h ha^{-1} MJ^{-1} mm^{-1}. Similar K-factor values from hilly mountainous terrains of Himalayan region have been previously reported (Kalambukattu et al., 2021). High values of K factor were found in agriculture (mean value of 0.023), followed by forest (0.018) and scurb (0.016). The soil erodibility analysis revealed that organic carbon content plays a vital role in determining the erodibility of the soil.

Table 2: C factor values of different LULC types

	LULC Type	C factor values
Agriculture		
1	Rajma, Potato, Soyabean, Pea, Madua,	0.10
2	Rice, Wheat	0.15
3	Apple	0.20
Forest		
1	Dense Deodar forest	0.04
2	Moderate Deodar forest	0.06
3	Dense Pine forest	0.08
4	Open Pine forest	0.01
5	Very Dense Mixed forest	0.04
Scrub		
1	Open Scrub	0.3
2	Dense Scrub	0.1

LS factor in the area had a minimum value of 0.03 and maximum value of 18.34 (Figure 6). Higher LS factor values were observed to coincide with areas having steeper slopes. The LULC values extracted for different sampling locations, revealed that agricultural areas exhibit lower LS factor values (mean value of 5.21), compared to regions under forest (7.96) as well as scrubs (6.98). This could be inferred in light of the restrictive nature of slope steepness for agricultural activities in the area.

C factor values were assigned considering the cover status as well as information regarding the cultivated crops at different sampling locations, identified in the field during field survey (Table 2). For agriculture areas, the values for each crop in the Kharif season were assigned based on the value reported by Panagos et al., 2015. Major crops grown and observed in the area were Rajma, Potato, Soyabean, Pea, Madua, Rice, Wheat and Apple. The C factor values for forest were given according to the type and density of the forest cover. The major forest tyeps observed in the area were dense deodar forest, moderate deodar forest, dense pine forest, open pine forest and very dense mixed forest. Similarly, the values for scrub types, i.e. dense scrub and open scrub were also assigned. Similarly, literature based (Ebabu et al., 2022) P factor values were also assigned based on the adopted management practices observed at different locations, during field survey (Table 3). Croplands in the area are terraced and farmers maintain stone bunds, stone patching and grass bunds. Some areas do not have any conservation practices and erosion features indicating higher erosion risk were observed in those areas. At many locations within forest areas also, management practices in the form of stone bunds were observed.

Table 3: P factor values used in the study

Sr. No.	LULC/ Support Practice	P factor values
1	Terracing	0.3
2	Stone Bunding	0.61
3	Fields with Bunds	0.2
4	Grass Bunding	0.36
5	No Conservation Practice	1
6	Dense Forest	0.8
7	Open forest and Scrub	1

3.4 Estimation of point scale erosion rates and validation

Soil erosion rates at different sampling locations were computed using the different model parameters employing RUSLE model. Further computed erosion rates were further classified into different erosion severity classes (e1, e2, e3 and e4), as discussed earlier. The computed classes were compared with the field estimated erosion class of each location to assess the overall validity of our point-scale model results. Out of the total 80 locations, field identified erosion classes of 68 locations matched with the model estimated soil erosion class, thus yielding an 85% agreement between both. The erosion values were observed to be highest in the points belonging to scrub areas due to lack of vegetation cover, absence of proper management practices and their occurrence in comparatively steeper slopes. Lower erosion values were observed in areas with good vegetation cover and presence of support practices such as bunds, terraces etc. Higher erosion was observed in areas of steeper slopes with higher LS factor values and higher rainfall erosivity values. The C and P factor values had considerable influence on erosion, especially in cases of areas under agriculture.

3.5 Modelling for erosion hotspot identification and mapping

Multicollinearity analysis carried out using nine inputs parameters before training the model for hotspot identification, helped to reduce the number of predictor

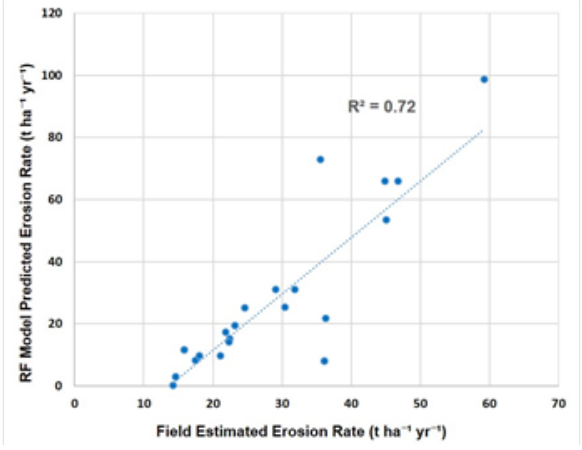


Figure 8: Scatterplot showing field estimated Vs. model predicted erosion rates

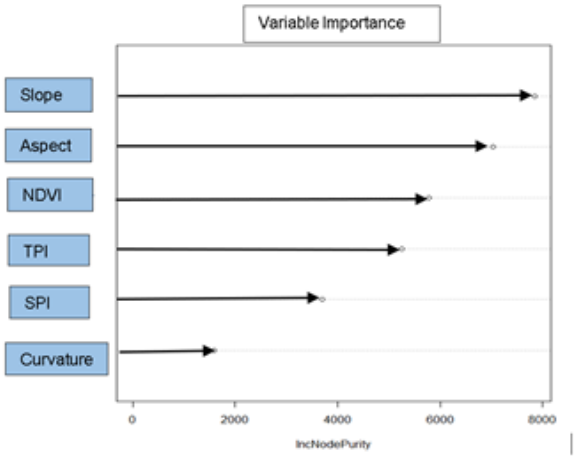


Figure 7: Plot showing relative importance of different predictor variables

variables by estimating their VIF values. Variables with VIF values more than 10 were excluded from further analysis, thus reducing the total number of predictors to six. Aspect (VIF: 3.952), Curvature (1.581), Slope (8.517), SPI (1.538), TPI (1.273) and NDVI (9.588) were selected for model building using RF regression. The validated model could predict the erosion values with considerable accuracy as indicted by $R^2$and RMSE values of 0.72 and 0.31, respectively (Figure 7). NSE value of 0.67 also revealed a satisfactory model performance for hotspot identification. The relative importance of different input variables in predicting the erosion values are given in Figure 8. The results revealed that slope and aspect played the most prominent role in predicting the erosion values in the hilly mountainous terrain. The validated model was used for spatial prediction and mapping of erosion risk in the entire area (Figure 9). Adoption of ML techniques as well as their integration with empirical erosion models yielding comparable results for assessing erosion risk and susceptibility to gully erosion is increasingly reported from various agro-ecological regions globally (Arabameri et al., 2020; Bag et al., 2022; Avand et al., 2023; Setargie et al., 2023, Phinsi et al., 2026). Among the various ML techniques, RF has frequently shown strong predictive performance, demonstrating its suitability for handling the complex and non-linear relationships between erosion and environmental factors.

The integration of ML with RS and GIS provides an effective means of analysing multiple erosion-controlling factors, including topography, soil properties, land use/land cover, vegetation, and climatic variables. For example, Avand et al. (2023) integrated RUSLE-derived erosion information with ML techniques to identify erosion hotspots in the Talar watershed, Iran, and reported superior performance of RF compared with other tested algorithms. Similarly, Phinzi et al. (2021) demonstrated that combining RUSLE and RF could effectively capture the spatial variability of soil erosion in the Umzintlava catchment. These findings support the complementary use of empirical erosion models and data-driven ML approaches.

The effectiveness of RF has also been demonstrated in mountainous and heterogeneous landscapes. Cheng et al.

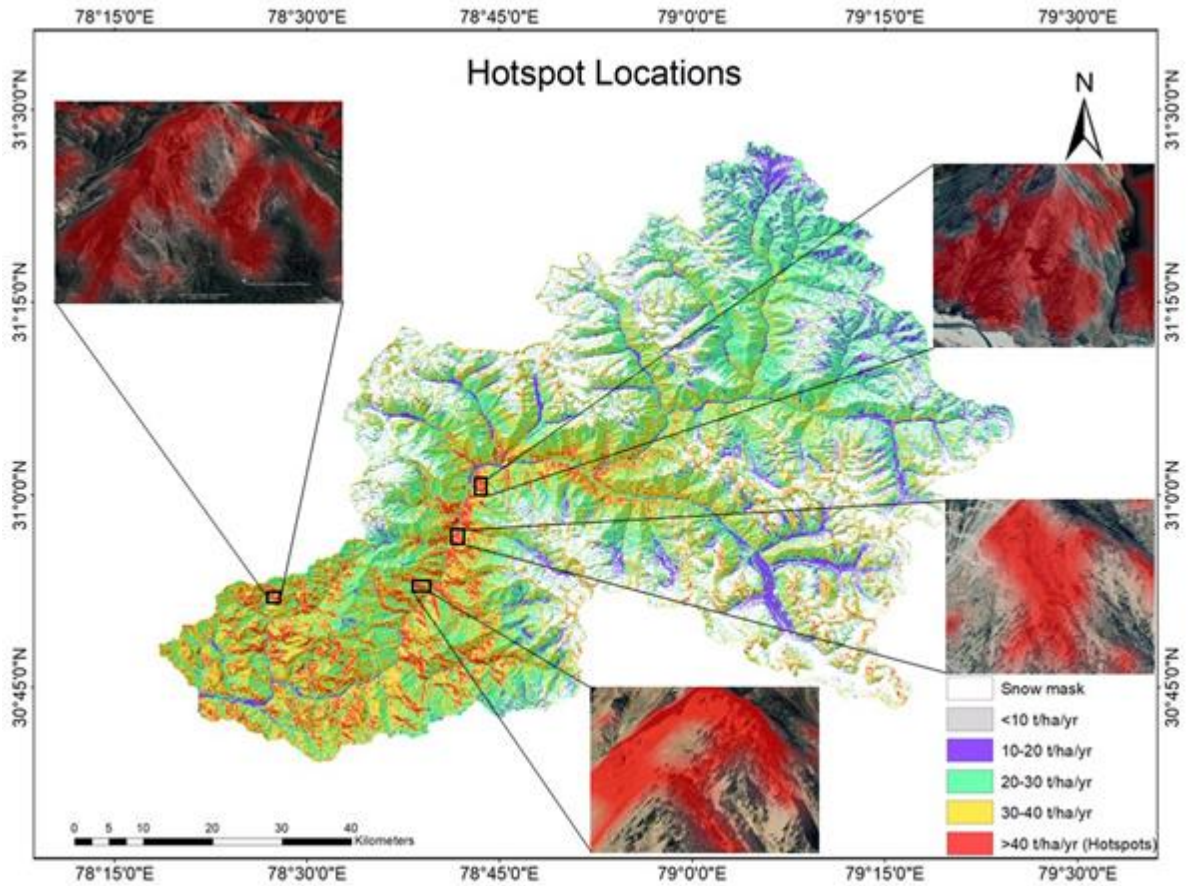


Figure 9: Soil Erosion Hotspot Map, along with erosion features seen on high resolution satellite imagery at few hotspot locations

(2018) successfully applied RF to identify erosion-controlling factors in hickory plantations in western Zhejiang, China, while Setargie et al. (2023) used RF and high-resolution remote sensing and field data to assess gully erosion susceptibility across different agroecological settings in the Upper Blue Nile basin. Their results highlighted the ability of RF to identify the relative importance of environmental factors controlling erosion and to provide reliable spatial predictions.

The findings of the present study are consistent with these previous studies, demonstrating that ML can effectively complement RUSLE for erosion assessment in complex Himalayan terrain. While RUSLE provides an established framework for estimating soil loss from major erosion-controlling factors, RF offers an additional advantage by capturing complex relationships among terrain and spectral variables. The integration of these approaches, supported by field observations and high-resolution geospatial datasets, therefore provides a useful framework for identifying and prioritizing soil erosion hotspots in mountainous and relatively inaccessible regions.

The spatial soil erosion risk values were divided into five classes representing Very low (<10 t/ha/yr), low (10 to 20 t/ha/yr), moderate (20 - 30 t/ha/yr), severe (30 – 40 t/ha/yr) and Very Severe (>40 t/ha/yr) erosion risk (Bhattacharyya et al. 2015). Areas belonging to very Severe class was identified as soil erosion hotspots in the area considering the high mean soil loss rate of 40.11 t/ha/yr estimated for very steep sloped in Uttarakhand state (Kalambukattu et al., 2021). The soil erosion hotspot map as per the above mentioned classes are given in Figure 9. Snow mask indicating areas under permanent snow cover in the watershed, were generated using GEE by comparing between the seasonal and permanent snow images. Very low erosion risk class covered only 0.04% of the area. Low, moderate, severe and very severe erosion risk classes covered 20.20%, 43.17%, 27.14% and 9.34% of the area respectively (Figure 10). The erosion / hotspot map was further overlaid on high resolution satellite imagery of the area employing Google Earth Pro platform and few hotspot locations were validated against high resolution imagery (given inside Figure 9). The overlay clearly revealed that erosion hotspots are prominently occurring in steep sloping slopes with sparse vegetation/ devoid of vegetation, especially near ridges or shoulder slopes. These locations were also characterised by comparatively higher rainfall erosivity values, thus highlighting the combined effect of rainfall, topography as well as land cover in accelerating the erosion processes and thus generation of erosion hotspot areas.

The erosion risk zones were further analysed with respect to slope, aspect as well as LULC types in the area. The area was broadly divided into four slope classes namely <10%, 10 - 25%, 25 - 50% and >50%. Nearly 56.54% of the area was under 25-50% slope category and contributed most towards erosion risk. Nearly 4.96% and 4.06% of the area under 25-50% and >50% slope categories were found to be falling under the very severe erosion risk class representing hotspots. Most of the areas with >50% slope

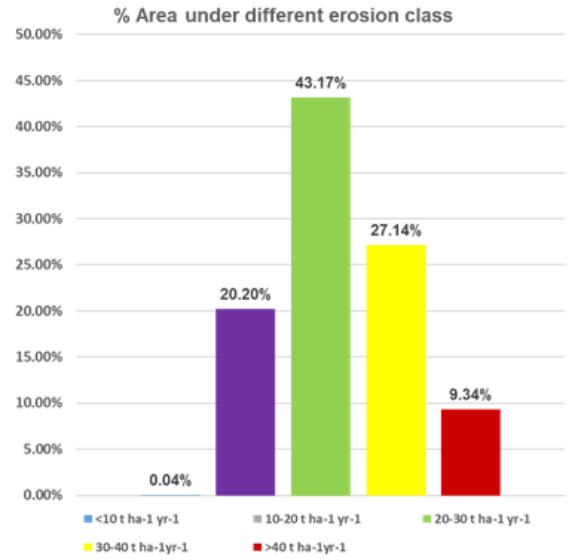


Figure 10: Percentage distribution of different erosion risk classes

occurred towards the higher elevation regions with snow cover, receiving lower rainfall and thus low values of rainfall erosivity thus resulting in lower rates of runoff based erosion processes.

Among the two broad aspect classes, southern aspect covered 66.29% of the area and contributes more towards the erosion. Nearly 3.92% and 5.41% area of the north and southern aspects were found to be falling under the very severe erosion risk class representing hotspots. This emphasis the increased erosion vulnerability of southern aspects, which have more area under scrub vegetation compared to northern slopes, mainly under forest cover.

On analysis of erosion risk zones with respect to different LULC types of forest, agriculture and scrubland, nearly 5.86% area belonging to scrublands were found to be erosion hotspots, followed by forest (5.30%). However, the hotspot area under agriculture areas was found to be very low (0.33%). The better slope management as well as adoption of conservation practices could be the major reasons behind reduced occurrence of hotspot area under agricultural lands.

3.6 Sources of Uncertainty

The soil erosion estimates as well as hotspot mapping carried out in the study still suffer from issues related to uncertainty. The uncertainties in the RUSLE estimates may arise from RUSLE assumptions, input datasets, field observations, and machine learning predictions. The use of the approximately 1 km resolution global rainfall erosivity product introduces uncertainty in mountainous terrain, where rainfall intensity can vary substantially over short distances because of strong elevation gradients and orographic effects. Consequently, localized variations in rainfall erosivity may not be fully captured, potentially affecting the accuracy of site-specific soil erosion estimates, although the dataset remains suitable for assessing broad, long-term watershed-scale erosion patterns. Similarly, K-factor estimates derived from the spatial interpolation of point-scale K-factor based on limited number of field samples may not fully represent soil variability, particularly across the remote, inaccessible, and unsampled portions of the watershed. Additional uncertainty may arise from the assumptions used to derive the LS, C, and P factors. The LS factor may not fully capture localized runoff concentration, abrupt slope changes, and gully formation, while seasonal variations in vegetation, crop growth, and management may not be completely represented in the C factor. Similarly, the P factor is generally assigned based on assumed conservation practices and land-use conditions, while the actual implementation and effectiveness of such measures may vary spatially across the watershed. These assumptions can therefore contribute to uncertainty in the estimated soil-loss values, particularly at local scales. Thus, in the complex Himalayan terrain, localized variations in rainfall, soil characteristics, land cover, and management practices may not be fully represented by the available spatial datasets.

Uncertainty in the RF model was also assessed through independent validation, which yielded an R^2 of 0.72, RMSE of 0.31, and NSE of 0.67, indicating good predictive performance with some residual uncertainty. The RUSLE-derived erosion classes were also validated against field observations, achieving an overall accuracy of 85%. Considering these limitations, the results should be interpreted primarily as estimates of long-term relative erosion risk rather than precise measurements of soil loss at every location. The uncertainty in the final predictions also reflects errors and resolution differences in the input datasets. Although high-resolution terrain and Sentinel-2 data were used, the spatial resolution and temporal characteristics of individual input layers differ. Furthermore, the limited number of field observations and the inherent empirical assumptions of RUSLE may influence the accuracy of the estimated erosion rates.

Despite these limitations, the combination of field validation, model-performance statistics, and multi-source geospatial data provides confidence in the overall spatial patterns and identification of major erosion hotspots. Future studies could further quantify uncertainty through approaches such as sensitivity analysis, uncertainty propagation, and ensemble modelling using additional high-resolution rainfall and denser field erosion observations.

4. Conclusion

The study demonstrated the potential of integrating RUSLE, remote sensing, GIS, field observations, and Random Forest (RF) modelling for assessing soil erosion and mapping the spatial distribution of erosion hotspots in the complex Himalayan terrain. Different satellite based RS datasets including ALOS DEM, Sentinel 2 data as well as various geospatial data products such as ISRO LULC, global rainfall erosivity, IMD gridded data and soil analysis data were used along with information generated during field survey for detailed study area characterisation and generation of various model inputs. The RF model was able to predict the erosion rates with an R^2 value of 0.72. The results revealed that slope and aspect played the most prominent role in predicting the erosion values in the hilly

mountainous terrain. The combined approach provided a reliable spatial representation of erosion risk and identified critical areas requiring priority conservation attention. The results indicate that erosion hotspots are mainly associated with steep terrain and sparse or absent vegetation, highlighting the importance of topography and land-cover conditions in controlling erosion across the watershed.

A major contribution of the study is the complementary use of an established empirical erosion model and a data-driven ML approach. While RUSLE provides a practical framework for estimating long-term soil loss, RF enhances the spatial prediction by capturing complex relationships between erosion and terrain and spectral characteristics. This integrated framework is particularly useful in mountainous regions where extensive field-based monitoring is constrained by difficult accessibility and heterogeneous environmental conditions.

The identified erosion hotspots should be prioritized for site-specific soil and water conservation measures, including vegetation restoration, slope stabilization, improved agricultural management, and appropriate runoff-control practices. Future research should focus on incorporating denser field observations, higher-resolution and locally calibrated rainfall erosivity data, multi-temporal remote sensing, and additional erosion indicators. Uncertainty propagation and ensemble machine learning approaches could also be explored to improved hotspot prediction reliability under diverse terrains and ecosystems. Such improvements would strengthen the applicability of the framework for long-term erosion monitoring, watershed prioritization, and sustainable land and water management in the Himalayan region.

Acknowledgments
The authors sincerely thank the Indian Space Research Organization (ISRO) for providing financial support under Earth Observation Applications Mission (EOAM) Project (ISRO/DOS) on Mountain Ecosystem Processes and Services - Phase - II -'Soil Erosion Estimation based on Radio Tracer Technique and Soil Quality Assessment in Mountainous Landscape of North-West Himalaya', which made this research possible. We express our gratitude to the Director, Indian Institute of Remote Sensing (IIRS) for providing all the necessary facilities and support for this research.

Conflict of 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 Statement
The geospatial and climate datasets used in the study could be obtained from the respective sources cited in the text. The soil database generated and used during the research work are currently not publicly available, as Indian Institute of Remote Sensing (IIRS), Indian Space Research Organization (ISRO), Department of Space, Government of India sponsored the project.

Author's Contribution
Conceptualization: JGK, RRR and SK; Methodology: RRR, JGK, SK and ADR; Formal analysis, Data collection, analysis and investigation: RRR, JGK and ADR; Writing original draft preparation: JGK and RRR; Visualization - RRR, Writing-review and editing JGK, RRR and ADR; Supervision: JGK and SK. All authors read and approved the final manuscript.

References

Arabameri, A., W. Chen, M. Loche, X. Zhao, Y. Li, L. Lombardo, A. Cerda, B. Pradhan, and D. T. Bui. 2020. "Comparison of Machine Learning Models for Gully Erosion Susceptibility Mapping." *Geoscience Frontiers* 11(5):1609–1620. https://doi.org/10.1016/j.gsf.2019.11.009

Avand, M., M. Mohammadi, F. Mirchooli, A. Kavian, and J. P. Tiefenbacher. 2023. "A New Approach for Smart Soil Erosion Modeling: Integration of Empirical and Machine-Learning Models." *Environmental Modeling & Assessment* 28(1):145–160. https://doi.org/10.1007/s10666-022-09858-x

Bag, R., I. Mondal, M. Dehbozorgi, S. P. Bank, D. N. Das, J. Bandyopadhyay, Q. B. Pham, A. M. Fadhil Al-Quraishi, and X. C. Nguyen. 2022. "Modelling and Mapping of Soil Erosion Susceptibility Using Machine Learning in a Tropical Hot Sub-Humid Environment." *Journal of Cleaner Production* 364:132428. https://doi.org/10.1016/j.jclepro.2022.132428

Benavidez, R., B. Jackson, D. Maxwell, and K. Norton. 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." *Hydrology and Earth System Sciences* 22(11):6059–6086. https://doi.org/10.5194/hess-22-6059-2018

Bhattacharyya, R., B. Ghosh, P. Mishra, B. Mandal, C. Rao, D. Sarkar, K. Das, K. Anil, M. Lalitha, K. Hati, and A. Franzluebbers. 2015. "Soil Degradation in India: Challenges and Potential Solutions." *Sustainability* 7(4):3528–3570. https://doi.org/10.3390/su7043528

Bouyoucos, G. J. 1962. "Hydrometer Method Improved for Making Particle Size Analyses of Soils." *Agronomy Journal* 54(5):464–465.

Burt, R. 1992. *Soil Survey Laboratory Methods Manual*. USDA.

Cheng, Z., D. Lu, G. Li, J. Huang, N. Sinha, J. Zhi, and S. Li. 2018. "A Random Forest-Based Approach to Map Soil Erosion Risk Distribution in Hickory Plantations in Western Zhejiang Province, China." *Remote Sensing* 10(12):1899. https://doi.org/10.3390/rs10121899

David Raj, A., S. Kumar, K. R. Sooryamol, J. G. Kalambukattu, and S. Kumara. 2024. "Soil Redistribution Rates along the Forested and Cultivated Steep Hillslope in the Mid-Himalayas Using Fallout-137Cs and RUSLE Model." *Land Degradation & Development* 35(16):4795–4813. https://doi.org/10.1002/ldr.5258

Desmet, P. J. J., and G. Govers. 1996. "A GIS Procedure for Automatically Calculating the USLE LS Factor on Topographically Complex Landscape Units." *Journal of Soil and Water Conservation* 51(5):427–433.

Ebabu, K., A. Tsunekawa, N. Haregeweyn, M. Tsubo, E. Adgo, A. A. Fenta, D. T. Meshesha, M. L. Berihun, D.

Sultan, M. Vanmaercke, P. Panagos, P. Borrelli, E. J. Langendoen, and J. Poesen. 2022. "Global Analysis of Cover Management and Support Practice Factors That Control Soil Erosion and Conservation." *International Soil and Water Conservation Research* 10(2):161–176. https://doi.org/10.1016/j.iswcr.2021.12.002

El Jazouli, A., A. Barakat, A. Ghafiri, S. El Moutaki, A. Ettaqy, and R. Khellouk. 2017. "Soil Erosion Modeled with USLE, GIS, and Remote Sensing: A Case Study of Ikkour Watershed in Middle Atlas (Morocco)." *Geoscience Letters* 4(1):9. https://doi.org/10.1186/s40562-017-0091-6

FAO. 1990. *Guidelines for Soil Description*. 3rd ed. Food and Agriculture Organization of the United Nations.

FAO. 2006. *Explaining Success in Reducing Under Nourishment Numbers in Ghana*. ESA Working Paper No. 03-10. Food and Agriculture Organization of the United Nations.

Ganasri, B. P., and H. Ramesh. 2016. "Assessment of Soil Erosion by RUSLE Model Using Remote Sensing and GIS: A Case Study of Nethravathi Basin." *Geoscience Frontiers* 7(6):953–961. https://doi.org/10.1016/j.gsf.2015.10.007

Ghosh, A., S. Rakshit, S. Tikle, S. Das, U. Chatterjee, C. B. Pande, A. Alataway, A. A. Al-Othman, A. Z. Dewidar, and M. A. Mattar. 2023. "Integration of GIS and Remote Sensing with RUSLE Model for Estimation of Soil Erosion." *Land* 12(1):116. https://doi.org/10.3390/land12010116

Goldblatt, R., W. You, G. Hanson, and A. K. Khandelwal. 2016. "Detecting the Boundaries of Urban Areas in India: A Dataset for Pixel-Based Image Classification in Google Earth Engine." *Remote Sensing* 8(8):634. https://doi.org/10.3390/rs8080634

Gupta, S., and S. Kumar. 2017. "Simulating Climate Change Impact on Soil Carbon Sequestration in Agro-Ecosystem of Mid-Himalayan Landscape Using CENTURY Model." *Environmental Earth Sciences* 76(11):394. https://doi.org/10.1007/s12665-017-6720-8

Kalambukattu, J. G., S. Kumar, and M. R. Hole. 2021. "Geospatial Modelling of Soil Erosion and Risk Assessment in Indian Himalayan Region—A Study of Uttarakhand State." *Environmental Advances* 4:100039. https://doi.org/10.1016/j.envadv.2021.100039

Karydas, C. G., P. Panagos, and I. Z. Gitas. 2014. "A Classification of Water Erosion Models According to Their Geospatial Characteristics." *International Journal of Digital Earth* 7(3):229–250. https://doi.org/10.1080/17538947.2012.671380

Kebede, B., A. Tsunekawa, N. Haregeweyn, E. Adgo, K. Ebabu, D. T. Meshesha, M. Tsubo, T. Masunaga, and A. A. Fenta. 2021. "Determining C- and P-Factors of RUSLE for Different Land Uses and Management Practices across Agro-Ecologies: Case Studies from the Upper Blue Nile Basin, Ethiopia." *Physical Geography* 42(2):160–182. https://doi.org/10.1080/02723646.2020.1762831

Kumar, S., N. Murugan, A. David Raj, and J. G. Kalambukattu. 2023. "Hillslope Elements and Soil–Landscape Analysis in Himalayas for Characterizing Soil Quality Parameters Using Digital Terrain Model and Remote-Sensing Data." *Environmental Earth Sciences* 82(20):469. https://doi.org/10.1007/s12665-023-11166-w

McCool, D. K., L. C. Brown, G. R. Foster, C. K. Mutchler, and L. D. Meyer. 1987. "Revised Slope Steepness Factor for the Universal Soil Loss Equation." *Transactions of the ASAE* 30(4):944–950. https://doi.org/10.13031/2013.30576

Mohammady, M., and A. Davudirad. 2024. "Gully Erosion Susceptibility Assessment Using Different Machine Learning Algorithms: A Case Study of Shazand Watershed in Iran." *Environmental Modeling & Assessment* 29(2):249–261. https://doi.org/10.1007/s10666-023-09910-4

Mohammed, S., K. Alsafadi, S. Talukdar, S. Kiwan, S. Hennawi, O. Alshihabi, M. Sharaf, and E. Harsányi. 2020. "Estimation of Soil Erosion Risk in Southern Part of Syria by Using RUSLE Integrating Geo-Informatics Approach." *Remote Sensing Applications: Society and Environment* 20:100375. https://doi.org/10.1016/j.rsase.2020.100375

Morgan, R. P. C. 2005. *Soil Erosion and Conservation*. 3rd ed. Wiley-Blackwell.

Panagos, P., P. Borrelli, K. Meusburger, C. Alewell, E. Lugato, and L. Montanarella. 2015. "Estimating the Soil Erosion Cover-Management Factor at the European Scale." *Land Use Policy* 48:38–50. https://doi.org/10.1016/j.landusepol.2015.05.021

Panagos, P., P. Borrelli, K. Meusburger, B. Yu, A. Klik, K. Jae Lim, J. E. Yang, J. Ni, C. Miao, N. Chattopadhyay, S. H. Sadeghi, Z. Hazbavi, M. Zabihi, G. A. Larionov, S. F. Krasnov, A. V. Gorobets, Y. Levi, G. Erpul, C. Birkel, … C. Ballabio. 2017. "Global Rainfall Erosivity Assessment Based on High-Temporal Resolution Rainfall Records." *Scientific Reports* 7(1). https://doi.org/10.1038/s41598-017-04282-8

Phinzi, K., N. S. Ngetar, and O. Ebhuoma. 2021. "Soil Erosion Risk Assessment in the Umzintlava Catchment (T32E), Eastern Cape, South Africa, Using RUSLE and Random Forest Algorithm." *South African Geographical Journal* 103(2):139–162. https://doi.org/10.1080/03736245.2020.1716838

Phinzi, K. 2026. "Integrating an Empirical Erosion Model and Machine Learning for Assessing Soil Loss and Sediment Delivery Dynamics in a Sub-Humid Catchment." *Modeling Earth Systems and Environment* 12:125. https://doi.org/10.1007/s40808-026-02780-1

Quinton, J. N., and P. Fiener. 2024. "Soil Erosion on Arable Land: An Unresolved Global Environmental Threat." *Progress in Physical Geography: Earth and Environment* 48(1):136–161. https://doi.org/10.1177/03091333231216595

Setargie, T. A., A. Tsunekawa, N. Haregeweyn, M. Tsubo, A. A. Fenta, M. L. Berihun, D. Sultan, M. Yibeltal, K. Ebabu, B. Nzioki, and T. M. Meshesha. 2023. "Random Forest-Based Gully Erosion Susceptibility Assessment

across Different Agro-Ecologies of the Upper Blue Nile Basin, Ethiopia." *Geomorphology* 431:108671. https://doi.org/10.1016/j.geomorph.2023.108671

Sharda, V. N., and P. R. Ojasvi. 2016. "A Revised Soil Erosion Budget for India: Role of Reservoir Sedimentation and Land-Use Protection Measures." *Earth Surface Processes and Landforms* 41(14):2007–2023. https://doi.org/10.1002/esp.3965

Sharpley, A. N., and J. R. Williams. 1990. *EPIC—Erosion/Productivity Impact Calculator: 1. Model Documentation*. Technical Bulletin No. 1768. U.S. Department of Agriculture, Agricultural Research Service.

Shrestha, D. P. 2009. "Assessment of Soil Erosion in the Nepalese Himalaya: A Case Study in Likhu Khola Valley, Middle Mountain Region." *Land Husbandry* 2(1):59–80.

Walkley, A., and I. A. Black. 1934. "An Examination of the Degtjareff Method for Determining Soil Organic Matter, and a Proposed Modification of the Chromic Acid Titration Method." *Soil Science* 37(1):29–38.

Yang, Y., R. Zhao, Z. Shi, R. A. Viscarra Rossel, D. Wan, and Z. Liang. 2018. "Integrating Multi-Source Data to Improve Water Erosion Mapping in Tibet, China." *Catena* 169:31–45. https://doi.org/10.1016/j.catena.2018.05.021