Integrating Terrain and Vegetation Indices for Identifying Potential Soil Erosion Risk Area

Arabinda Sharma
Civil Engineering Department, MM University, Mullana-113207, Haryana, India

Abstract The present paper offers an innovative method to monitor the change in soil erosion potential by integrating terrain and vegetation indices derived from remote sensing data. Three terrain indices namely, topographic wetness index (TWI), stream power index (SPI) and slope length factor (LS), were derived from the digital elevation model. Normalized vegetation index (NDVI) was derived for the year 1988 and 2004 using remote sensing images. K-mean clustering was performed on staked indices to categorize the study area into four soil erosion potential classes. The validation of derived erosion potential map using USLE model showed a good agreement. Results indicated that there was a significant change in the erosion potential of the watershed and a gradual shifting of lower erosion potential class to next higher erosion potential class over the study period.

Keywords remote sensing; GIS; NDVI; soil erosion; terrain indices

Introduction

Soil erosion by water is a serious geo-environmental problem causing land degradation all over the world. It causes damage to agricultural lands, sedimentation of reservoir and water quality problems in nearby surface water bodies. The processes involved in soil erosion and its driving forces are largely known. However, the momentum of global and regional socioeconomic changes and the resulting ongoing pressure on land resources goes far beyond current planning opportunities to mitigate soil erosion.[1]

The requirement for identifying and monitoring soil erosion from watershed has led to the development of several empirical and process based models within a geographic information system (GIS). The developed soil erosion models use either an index-based approach such as Sediment Yield Index (SYI),[2] Universal Soil Loss Equation (USLE),[3] Revised Universal Soil Loss Equation (RUSLE)[4] or a dynamic approach like the Water Erosion Prediction Project.[5] In terms of temporal resolution, soil erosion models may be either event-based models (e.g., WEPP) to predict size and timing of sediment discharges or annual-based models like USLE and RUSLE to estimate long-term values of mean annual soil loss. Soil erosion is a continuous process affecting natural resources at landscape level. Its monitoring and mitigation measure require long term evaluation of different driving forces at watershed scale. Factors used to describe erosion that return in most models are land use, slope, precipitation amount and intensity, runoff and peak runoff rates, runoff shear stress, soil cohesion and surface roughness. The problem for the development of models based on
these factors is that many of these factors are often difficult to assess, far from constant in space and time, and interact with each other.\cite{6}

The recent advancement in remote sensing technology can assist in model formulation by providing accurate and timely information on various watershed characteristics, i.e., terrain, vegetation, soil distribution, drainage, etc. These techniques promote the development of empirical or regression models. But unfortunately creating and validating such comprehensive models is a formidable task and require enormous information on important hydrological and biophysical setting of the watershed including rainfall, temperature, evapotranspiration, runoff, sediment yield, representation of topography, soils and land use/land cover.\cite{7} This often hinders the assessment and monitoring of soil erosion hazard, especially in data poor regions or regions under intensive change and crustal deformation. Digital terrain analyses can be used to make interpretations relevant to land resource management and decision-making. Few studies have examined the correspondence between terrain attributes and soil survey data at scales typical of public sources used for resource management planning in large watersheds. A recent study\cite{8} reported a correspondence between terrain and soil survey attributes. The terrain analyses may contribute, as an alternative to traditional basis for soil conservation planning, in identifying target conservation area through qualitative analysis. Numerous researchers have tried using terrain analysis and vegetation information for identifying the erosion potential area\cite{1, 9-10}

The present state of affairs suggest for additional work towards developing a simple method to map potential area of erosion risk even in the absence of essential data required to set and validate erosion model. In this context, this paper is an attempt to develop a generic method for the spatial evaluation of potential soil degradation risks. Given the serious current soil degradation problems in the Maithon catchment owing to deforestation, mining activity and poor agricultural practices, the main objective of the paper was to assess and monitor change in soil erosion risks during 1988-2004 by integrating terrain and vegetation analysis. Thus, terrain analysis may contribute as an alternative to traditional basis for soil conservation planning by helping in the identification of target conservation area through qualitative analysis.

1 Material and methods

1.1 Study area and data used

The study site for the present research work is the Maithon reservoir catchment (85.41°–86.90° E longitude and 23.75°–24.56° N latitude), situated in Jharkhand state of India. The watershed covers an area of about 5553 km². The elevation ranges from 120 m to 1360 m above mean sea level. It is predominantly an agricultural watershed with scattered areas of active mines and sparsely distributed forest patches.

The main spatial data used in the present study, shuttle radar topographic mission (SRTM), digital elevation model (DEM) and satellite imageries of Landsat-5 TM and IRS P6 LISS III sensors were procured for the year 1988 and 2004, respectively. The ancillary data used in the present study includes survey of India (SOI) topographic maps and substantial amount of field data collected during ground truthing to support image classification and validation.

1.2 Methodology

The occurrence of soil erosion is largely dictated by topographic features and vegetation of the landscape. Thus, the spatial distribution of areas with topographic potential for erosion was modeled using the approach based on the terrain and vegetation indices. For the purpose three compound terrain indices such as Length-Slope factor, Stream Power Index and Topographic Wetness Index along with normalized vegetation indices were selected. These indices represent the underlying physics of natural processes that have important hydrological and geomorphological consequences in many landscapes. In this study, terrain indices were derived with ArcView scripts based on the D8 algorithm while NDVI was calculated using the spatial modeler of Erdas Imagine software.

1.2.1 RUSLE length-slope factor (LS factor)

Length-Slope factor\cite{11} implemented in the revised universal soil loss equation (RUSLE) was used in this study. It gives a value for the water erosion potential
relative to a slope of 22.13 m length and a slope angle of 5 degree. It can be represented as the following equation:

$$LS = (m + 1) \left( \frac{A}{22.13} \right)^n \left( \sin \beta / 0.0896 \right)^\alpha$$

(1)

Where \(m = 0.4\) and \(n = 1.3\) for a slope length <100 m and a slope angle <14°. Slope length and slope gradient (\(\sin \beta\)) are most often considered topographic factors influencing soil loss. Total soil loss per unit area in erosion events tends to increase with increasing slope length with a well-fitted power law relationship. RUSLE considers erosion only along the flow line without considering the influence of flow convergence/divergence (depositional areas) and hence the equations can be properly applied only to areas experiencing net erosion.

1.2.2 Stream power index (SPI)

Stream power is the time rate of energy expenditure and has been used extensively in studies of erosion and sediment transport as a measure of the erosive power of flowing water. This index calculates a spatially distributed sediment transport capacity and may be more suited to landscape assessments of erosion than other approaches because it accounts for flow convergence and divergence. It computes the spatial distribution of soil loss potential by assuming uniform rainfall excess runoff and that the erosion rate is transport rather than detachment limited.[12] SPI can be calculated as follows:

$$SPI = A_s \tan \beta$$

(2)

Where \(A_s\) and \(\tan \beta\) are specific catchment area and local slope, respectively. In SPI, the measure of erosive power of flowing water is based on the assumption that discharge (\(q\)) is proportional to specific catchment area (\(A_s\)). It predicts net erosion in areas of profile convexity and tangential concavity (flow acceleration and convergence zones) and net deposition in areas of profile concavity (zones of decreasing flow velocity).

1.2.3 Topographic wetness index (TWI)

Topographic wetness index (TWI) has been used extensively to describe the effects of topography on the location and size of saturated source areas of runoff generation. The concept was first used in TOPMODEL and further developed in the 1990s.[13] It can be expressed as follows:

$$W = \ln \left( \frac{A_s}{\tan \beta} \right)$$

(3)

Where \(A_s\) is calculated from specific catchment area (\(m^2m^{-1}\)) of a point and the \(\tan \beta\) is local slope gradient in degrees. This particular equation assumes steady-state conditions and uniform soil properties (i.e., transmissivity is constant throughout the catchment and equal to unity). The TWI empirically represents the spatial distribution of soil moisture, surface saturation and hence surface runoff which is an important factor to stimulate soil erosion processes.

1.2.4 Normalized vegetation indices (NDVI)

Vegetation cover is possibly the most crucial element in the process of soil erosion study and management. It is the most dynamic factor in a watershed which can be readily altered to control the loss of water and soil. We estimated the vegetation fraction in the study area using satellite images derived NDVI, first used in [14]. It is a numerical indicator, which uses the visible and near-infrared bands of the remote sensing data for identifying green vegetation in the study site. The NDVI algorithm subtracts the red reflectance values from the near-infrared and divides it by the sum of near-infrared and red bands.

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

(4)

Theoretically, NDVI values are represented as a ratio ranging in value from \(-1\) to 1 but in practice extreme negative values represent water, values around zero represent bare soil and values over 0.5 represent dense green vegetation.

1.2.5 Soil erosion potential mapping

The ultimate goal of this study was to integrate terrain indices along with NDVI in GIS to identify the area of topographic potential of soil erosion. Obtained TWI, SPI, LS, and inversely signed NDVI (because high NDVI corresponds to high vegetation cover and hence less soil erosion) map were rescaled from 0-255. In the normalized indices the zero and 255
values correspond to lowest and highest soil erosion in a relative scale. In the subsequent step, the normalized indices were combined (stacking) to form a multi band grid. The multi band grids obtained for year 1988 and 2004 were put to multivariate analysis in ArcGIS 9.1 environment to identify the soil erosion potential categories. For this purpose, ISODATA clustering was performed to determine the characteristics of the natural groupings of cells in multidimension attribute space and the results are stored in an output ASCII signature file. The signature file was then used in the maximum likelihood classifier to obtain four clusters corresponding to four erosion potential categories, i.e., very low, low, medium and high.

1.3 Validation of erosion potential map

For validation purposes, a reference soil erosion potential map was prepared using universal soil loss equation (USLE) model of the same area for the year 2004. USLE is the most frequently applied worldwide for estimating the annual soil loss from rainfall erosivity, topography and land-use data. The model can be represented as follows:

$$A = R \times K \times L \times S \times C \times P$$  

(5)

Where $A$ is the soil erosion calculated in tons/ha/year; $R =$ rainfall erosivity factor (MJ mm ha$^{-1}$ h$^{-1}$ yr$^{-1}$); $K =$ soil erodibility factor (t ha h$^{-1}$ MJ$^{-1}$ mm$^{-1}$); $L =$ slope length factor; $S =$ slope steepness factor; $C =$ cover and management factor; and $P =$ erosion control practice factor. For calculating $R$-factor and $K$-factor and monthly average rainfall data for the year 2004 and a soil map with related attributes were collected from soil conservation office of Damodar Valley Corporation (DVC) situated at Hazaribag, India.

2 Results and discussion

The results of terrain and vegetation analysis, and change in spatial distribution of topographic potential of soil erosion in Maithon catchment are presented below.

2.1 Terrain indices

SRTM DEM of 90 m resolution was used in the present study to calculate the various compound terrain indices. Fig.1(a) shows the presence of hilly terrain with steep slope in the southern and north western part of the study site. The low altitude, relatively flat areas are generally found in the south eastern region while areas with moderate slope are found in the central part of the watershed. During terrain analysis, we encountered unrealistic high values (for LS factor and SPI in steep slopes and for TWI in local valley bottoms) with large contrast to neighboring values which was also reported by earlier researchers. Thus, all the three indices were obtained by applying mean filter to the original one that smoothens the large contrasts of values (removal of outliers or noise).

The distribution pattern of LS factor is shown in Fig.1(b). The LS factor indicates spots on the field that have both a steep slope and relatively large length of a slope. Thus, it can be assumed as an indicator of higher amount of incoming water with high kinetic energy. The value of LS factor in the study site ranges from 0 to 11. The zero value is mainly due to the presence of flat area or area with negligible slope. The pattern shown in Fig.1(b) confirms the presence of high LS factor corresponds to hilly region in the southern part of the watershed and LS factor of intermediate value are concentrated in the central part of the watershed. However, majority of the watershed has relatively small value of LS factor. LS factor is directly proportional to erosion potential of the terrain, i.e., a high value of LS factor has high soil erosion risk and vice versa. The LS factor predicts high erosion potential at the hillslope regions (lower concave parts) but these are the areas, where deposition is usually observed. Since the influence of flow convergence is not included, the LS factor predicts relatively low erosion in areas with convergent water flow and higher erosion in some areas with dispersal water flow (Fig.1(b)). This limitation of LS factor is largely taken care of by stream power index.

The process-based stream power index (SPI) predicts similar flow lines like the TWI. However, the grid value representing the flow acceleration that water experiences downhill increases with increasing slope steepness (Fig.1(c)). The value of SPI ranges from 0 to 2.9894. Since SPI is directly proportional to erosion potential, areas with high magnitude of SPI will have high erosion potential while SPI value zero
indicates area of deposition potential, which is generally located at the local extremes.

This is in agreement with the description of the variation in erosion processes along complex slopes as presented in [16]. The high magnitude of SPI almost corresponds to LS factor, found in the southern hilly region of the watershed. However, high LS value is also found in the central moderately sloped area of the watershed. This dramatic increase in stream power index in valleys (Fig.1(c)) indicates that the stream formation (areas of water flow convergence) is even more striking when LS factor is small (Fig.2(b)). The area where deposition occurs (area of reduced velocity of water) can be derived by calculating the slope of the SPI or areas where the stream power changes, i.e., along the slopes. In Fig.1(c) the drainage pattern or flow lines of the watershed could be clearly observed, thus showing areas where soil conservation measure is to be set up to mitigate the flow acceleration and the erosive force.

Often in soil erosion studies, it is important to characterize the presence of saturation excess overland flow area that favors ephemeral gully formation.[16] TWI reflects the multiple influences of the terrain on saturation excess runoff processes and closely related to sorting of suspended sediment material. A homogeneous distribution of soil conditions was assumed for TWI calculation and the spatial distribution of TWI in the watershed is shown in Fig.1(d). The value of TWI ranges from 5.2 to 11.0 in the study site. Fig.1(d) illustrates that high values of the TWI are found mostly in the northern part of the watershed (converging flat terrain) while low values are typical to the southern part of the watershed (steep, diverging areas). The TWI maps indicate zones of high potential soil moisture (high values) and zones that dry up first (low values). High TWI is of significance because gullies are often formed where saturated surface soil loses its strength and then slump when seepage occurs. The slumped soil is carried away by overland flow (rainfall excess plus saturation excess) that further scour the bed and wall of the channel causing gullies to deepen and widen. Hence, the areas with high TWI will have more risk of gully erosion than areas with low TWI value and vice versa.

2.2 Vegetation indices

In this study, we consider NDVI as a measure of vegetation extent and density. NDVI images for 1988 and 2004 were obtained from the Landsat5 TM and IRS (Indian remote sensing satellite) P6 LISS III (linear imaging self-scanning system) data.

Keeping in view different sensor characteristics, due care was taken to reduce the uncertainty accounted
for sensor calibration and viewing angle. Two separate linear regression models, one each for Red band (Band 2 of IRS LISS III vs. Band 3 Landsat TM) and NIR band (Band 3 of IRS LISS III vs. Band 4 Landsat TM) were developed. The values of coefficient of determination ($R^2$) were found to be 0.91 and 0.95 for Red band and NIR band, respectively. Since both scenes used were of the January month, it can be assumed that there were no uncertainties in NDVI difference due to seasonal variation and vegetation phenology. In this study, we did not have shadow effect and a very negligible portion of 2004 imagery was under cloud cover so that was compensated by using NDVI value from 1988 imagery. However, the effect of rainfall variation on NDVI value were not considered due to lack of rainfall data for the study period and a constant uniform annual rainfall was considered for the present study. The spatial temporal distributions of NDVI value during the study period are shown in Fig.2.

![Spatial distribution of NDVI in the study site](image)

During the year 1988, the values of NDVI range between −1.0 and 0.67 while in the year 2004 they range between −0.66 and 0.65. Fig.2 shows that most of the high NDVI values are located in the southern part of the watershed and correspond to forest area while the lowest values of NDVI are obtained over reservoir and other water pixels in the watershed. Pixel based analysis of NDVI images of both years confirm that a value of NDVI above 0.45 is vegetation (forest) while a value below −0.40 correspond to water in the studied landscape. Comparison of NDVI images of both years reveals some interesting changes in spatial distribution of NDVI values. The amount of pixels above 0.45 (blue color) was decreased during the studied period. Most of the loss was observed just above and below the main river in the central part of the watershed. It was also observed that there was a decrease in intensity of blue color (high NDVI value). This is an indication of a decrease in greenness of the watershed in terms of vegetation extent and density as well. Thus, the landscape of the watershed has been deemed more susceptible to soil erosion due to loss of protective vegetation cover. Using NDVI as a factor for soil erosion estimation is more specifically related to affect land management on soil loss. Recently, several researchers used NDVI value for direct derivation of $C$-factor of the USLE model through an empirical model.[17]

### 2.3 Assessment of erosion potential and temporal change

All the indices rescaled between 0-255 were integrated and put into multivariate analysis to identify the potential erosion area. Given that the reservoir area of the watershed resulted in unrealistic clustering patterns, it was kept out from the erosion potential analysis. Fig.3 shows the geographic pattern of erosion potential (erosion is likely to occur) in the study area predicted with a simple terrain based model. Fig.3 demonstrates that the areas of high potential of soil erosion are generally found in the southern part of the watershed which corresponds to high value of LS factor and stream power index and NDVI. Most of the vegetative areas (high NDVI value) of the watershed fall either in high class or in medium erosion intensity class so by virtue of their location on steep slopes.

The very low class of erosion potential is found to be distributed all over the watershed and more or less follow a dendrite pattern close to the natural drainage of the watershed. This suggests that the proposed method is suited to identify the area of rill and sheet erosion. In order to depict the area of gully erosion, proper weight
should be assigned to individual indices but that needs the help of some local expert followed by validation using field check or using high resolution imagery from other sources, i.e., Google earth images.

2.4 Validation of erosion potential map

The erosion potential map generated by integrating the terrain and vegetation indices was also validated using soil erosion category map obtained through universal soil loss equation (USLE) model for the year 2004. USLE model requires data on rainfall, soil characteristics and conservation practice in addition to terrain and vegetation data. The rainfall erosivity (R-factor) for three gauging station was calculated from average monthly rainfall data available for the 2004 using \(^{[18]}\) method. The value for three gauging stations were found to be 320, 345 to 385 MJ mm ha\(^{-1}\)h\(^{-1}\)yr\(^{-1}\) and for unguided points it was estimated using Thiessen polygon method. Soil erodibility (K-factor) was calculated for different soil categories using the formula in [19]. Slope length and steepness factor (LS-factor) was directly calculated from SRTM DEM in ArcView using the formula suggested in [20]. The values of K-factor range from 0.18 to 0.26 in the study area. The magnitude of LS-factor varies considerably between 0.008 (in flat area) to 4.64 (in steep slope). The cover factor C was obtained by assigning value to individual LULC class water body (1.000), settlement (0.002), cropland (0.320), forest (0.004) and wasteland (0.100). The values inside the bracket indicate the C value used for a particular LULC based on field data and literature survey. The practice factor P was invariably considered to be 1 (no conservation structure). The resultant soil erosion maps from both the methods were classified into four soil erosion intensity classes (Table 1) as suggested in [21].

| Sl. No. | Soil erosion (ton/ha/year) | Erosion intensity class |
|--------|---------------------------|------------------------|
| 1      | 0-5                        | Very Low               |
| 2      | 5-10                       | Low                    |
| 3      | 10-20                      | Medium                 |
| 4      | >20                        | High                   |

During the validation process the erosion potential map of year 2004 obtained by integrating terrain-vegetation indices and erosion potential map from USLE model were checked for one-to-one correspondence of erosion intensity category for each pixel using Kappa statistics. The Kappa value of 0.74 indicates a good agreement of the erosion potential map, prepared using integration of biophysical indices, with one obtained from USLE model. It supported the general perception (hypothesis) that the erosion hazard is related particularly to the pattern of the vegetation cover and terrain characteristics for a given soil and rainfall characteristics. Simultaneously, a Kappa coefficient value of 0.74 also signified that the factors such as rainfall and soil were intricately related to soil erosion process and hence, a very high Kappa coefficient could not be obtained during validation. However, the validation process indicated that integration of terrain indices such as TWI, LS and SPI with NDVI can be used for rapid identification of the areas that are prone to soil erosion in the watershed.

2.5 Change in erosion potential

Subsequently, the erosion potential maps of both the years were compared. It was observed that there is a gradual shifting from lower erosion potential class to
next higher erosion potential class. Fig. 4 shows the percentage distribution of each erosion potential class during the study period. There is an increase in areas of high and medium erosion potential classes while areas falling under low and very low erosion potential class were decreased over a span of 16 years.

![Percentage distribution of erosion potential classes](image)

The statistical significance of the change in the erosion potential of the watershed was carried out using Kappa statistics. The different class agreement values in confusion matrix, obtained through cross tabulation of the erosion potential maps, were used to calculate the Kappa coefficient. A Kappa coefficient of 0.61 suggests that there was a significant change in the erosion potential of the watershed over the study period.

The main limitations of the above methods are the relatively coarse resolution data and little scope to validate the model. However, the presented method still hold’s good in view of its simplicity, reduced cost and time as compared to conventional RUSLE method. The latter method is time and labor intensive and needs rainfall and soil texture data. Also, the consistency of data is hard to maintain. The efficiency of the described model can be increased with the use of high resolution and error free DEM, which is easily available nowadays with laser altimetry, radargrammetry and satellite photogrammetry techniques. The effectiveness of the method could further be increased by reducing the possible source of uncertainty in measuring the true NDVI value.

### 3 Conclusion

The goal of this study was to utilize remote sensing as the basis for identifying the potential area of soil erosion in a watershed. This assessment process identifies watershed area prone to soil erosion by considering both erosion and sedimentation processes. This study also focused on mapping the geographic distribution of change in soil erosion potential across the Maithon reservoir catchment during 1988 to 2004 using terrain and vegetation indices. Topographic variables derived from remote sensing data were related to the geographic distribution of erosion and were found to be important indicators of erosion. The erosion potential map generated by integrating different terrain indices and NDVI showed a good agreement with the erosion potential map generated using USLE model. A gradual shifting from lower erosion potential class to next higher erosion potential was observed. The areas of high and medium erosion potential class were increased while those of low and very low erosion potential class were decreased during the same period. A Kappa coefficient of 0.61 suggested that there was a significant change in the erosion potential of the watershed over the study period. This approach might not have produced predicted maps with highest accuracy, but it provided a quick, realistic and simple method for combining topographic variables for mapping and monitoring the potential location of soil erosion change. The method can also be transferred to other data poor region with similar environmental conditions for rapid assessment and monitoring of change in erosion potential. The final erosion potential map produced can be used in decision making regarding the prioritization of study areas or for further modeling for selection of conservative measures to reduce soil loss. Thereby, the presented method provided a tool to help with the implementation of soil conservation measurement and sustainable management of other natural resources.

### References

1. Bou Kheir R, Abdallah C, Khawlie M (2008) Assessing soil erosion in Mediterranean karst landscapes of Lebanon using remote sensing and GIS [J]. *Engineering Geology*, 99(3-4): 239-54
2. Bali Y P, Karale R L (1977) Sediment yield index as a criterion for choosing priority basins [J]. *IAHS-AISH Publication*, 122: 180-8
3. Wischmeier W H, Smith D D (1978) Predicting rainfall...
Arabinda Sharma / Integrating Terrain and Vegetation Indices ...