

Spatial Analysis and Prediction of Soil Erosion in a Complex Watershed of Cameron Highlands, Malaysia

31

Taofeeq Sholagberu Abdulkadir, Raza Ul Mustafa Muhammad, Olayinka Gafar Okeola, Wan Yusof Khamaruzaman, Bashir Adelodun, and Saheed Adeniyi Aremu

Abstract

Land degradation in the form of erosion is a serious geo-hazard threatening land and water resources sustainability. Its socioeconomic and ecological impacts necessitated its geospatial prediction via susceptibility analysis. Accuracy of susceptibility mapping depends largely on modeling techniques and causative factors (CFs) considered. The study implements scarcely used CFs and support vector machine (SVM) technique for geospatial prediction of soil erosion for low and peak rainfall cycles in a complex watershed of Cameron Highlands Malaysia. The CFs considered are

non-redundant static (drainage density, length slope, lineament density, and erodibility) along with some dynamic (surface temperature, soil moisture index, rainfall erosivity, and vegetation index) CFs. Four kernel functions of SVM with optimized kernel parameters were used for spatial prediction of erosion for both rainfall cycles. The model performances were validated using the commonest evaluation criteria. The results indicated that polynomial SVM outperformed other models and was used in developing soil erosion susceptibility maps. The distribution pattern of erosion showed that majority of the area are “low and moderately” susceptible. The analysis shows that most of the highly susceptible locations are within the urban and agriculture land-use. The study provides information on erosional land forms that could be advanced to gullies. This medium-scale susceptibility map produced could be useful for sustainable watershed management for mitigating erosion.

T. S. Abdulkadir (✉)

Department of Water Resources and Environmental Engineering,
University of Ilorin, Ilorin, Kwara State, Nigeria

Department of Civil and Environmental Engineering,
UniversitiTeknologi PETRONAS, Bandar Seri Iskandar, Perak,
Malaysia

R. U. M. Muhammad

Department of Civil and Environmental Engineering,
UniversitiTeknologi PETRONAS, Bandar Seri Iskandar, Perak,
Malaysia

Centre for Urban Resource Sustainability, Institute of Self-Sustainable
Building, UniversitiTeknologi PETRONAS, Seri Iskandar, Perak,
Malaysia

O. G. Okeola

Department of Water Resources and Environmental Engineering,
University of Ilorin, Ilorin, Kwara State, Nigeria

W. Y. Khamaruzaman

Department of Civil and Environmental Engineering,
UniversitiTeknologi PETRONAS, Bandar Seri Iskandar, Perak,
Malaysia

B. Adelodun

Department of Agricultural and Biosystems Engineering, University of
Ilorin, Ilorin, Kwara State, Nigeria

S. A. Aremu

Department of Water Resources and Environmental Engineering,
University of Ilorin, Ilorin, Kwara State, Nigeria

Lower Niger River Basin Development Authority, Ilorin, Kwara State,
Nigeria

Keywords

SVM · Rainfall · Dynamic factors · Gully erosion ·
Susceptibility

31.1 Introduction

The key basic human needs are intricately linked to water-food nexus, and un-abating world population increase is mounting pressure on land exploitation with resultant impacts on soil and water. Soil erosion, one of the major causes of land degradation (Agnesi et al. 2011; George and Anu 2018; Nampak et al. 2018; Rizzei et al. 2016), is a key factor that endangers water and soil (Arabameri et al. 2018). Decisions made relating to water and food resources are often politically contested (Allouche et al. 2014; Williams et al. 2014). The consequential role of piping in landslides and gully erosion cannot be overemphasized. Piping, the

movement of fine particles out of the soil layer by seepage forces, is a key suspect in the commencement of quite process leading to gully erosion, landslides, and other direct land degradation processes (George and Anu 2018). Asia is ranked second globally with average erosion rate of 16.6 Mg/ha/year followed by South America with a rate of 22.1 Mg/ha/year (Walling and Webb 1983). It results in economic loss, changes of rural livelihood pattern (Kerr 1998), natural resource degradation, and increasing sediment deposition (Kelley and Nater 2000) and has impact on sustainable development and ecosystem services (Pimentel and Burgess 2013; Tinker 1997). Due to these, several studies are ongoing particularly on the predictability of this phenomenon before unleashing havoc on lives and economy. Gully erosion develops in watercourses or other locations where runoff concentrates. It started from rill erosion which may later advances to gullies. It dissects lands, reducing its productivity and values via destruction of farm, harboring vermin, sediment generation, etc. Thus, gully erosion is arguably a huge threat in the attainment of five UN sustainable development goals (SDGs) for this region: SDSs 1 (No poverty), 2 (Zero hunger), 3 (Good health and well-being), 6 (Clean water and sanitation), and 11 (sustainable cities and communities). The SDGs are follow-up of the MDGs established as a response to world poverty, inequality, and insecurity, which have developed into the drivers of resource management. It is imperative to intensify research efforts for the prediction of erosion susceptible locations and mitigate its impacts.

Several studies have been carried out in the last decades on developing models for the evaluation of erosion rates and its susceptibility analysis and mapping at different scales and objectives (Agnesi et al. 2011; Dewitte et al. 2015; Gómez-Gutiérrez et al. 2015; Khatun 2017; Lucà et al. 2011; Mararakanye and Le Roux 2012; Pournader et al. 2018; Sujatha and Sridhar 2018; Wang et al. 2016). Most methods quantify eroded soil volumes by empirically or physically based approach linking soil loss rates to the values of a set of environmental variables or mechanical properties of terrains. The common method is Universal Soil Loss Equation (USLE), and its revised version, RUSLE model (Pham et al. 2018), is widely used in the developing countries because of their simplicity and flexibility and it needs less data than most of the other erosion models (Duarte et al. 2016; Thlakma et al. 2018). Other models that find applicability in soil erosion susceptibility assessment for their compatible integration with GIS are Soil and Water Assessment Tool (SWAT), European Soil Erosion Model (EUROSEM), and Annualized Agricultural Non-Point Source (AnnAGNPS) (Kouli et al. 2009; Thlakma et al. 2018). Other models for assessing soil erosion and gully rate determination quantitatively and qualitatively include Erosion Potential Method; Modified Pacific Southwest Interagency

Committee Model (MPSIAC); Chemicals, Runoff, and Erosion from Agricultural Management Systems (CREAMS); Water Erosion Prediction Project (WEPP); and Ephemeral Gully Erosion Model (EGEM) (Althuwaynee et al. 2014; Barber and Mahler 2010; Pournader et al. 2018).

In measuring, mapping, and monitoring gully erosion, several approaches coupled with remote sensing and GIS have been developed and implemented to obtain spatiotemporal continuous gully information (Wang et al. 2016). For examples, high spatial resolution of classical aerial photography and terrestrial laser scanner (TLS) find applicability in the ability to measure gullies. However, advance use of unmanned aerial vehicles (UAVs) with integrated autopilot technology image acquisition are finding relevancy in the study of gully erosion (d'Oleire-Oltmanns et al. 2012; Hugenholtz et al. 2013; Stöcker et al. 2015). Also, automatic 3D photo-reconstruction techniques enhance image processing for those images are derived from uncalibrated and nonmetric cameras. The 3D photo-reconstruction techniques from the images obtained from UAV to generate high-resolution topographical data and ortho-images are being used for monitoring gully erosion (Stöcker et al. 2015; Wang et al. 2016).

Recently, statistical techniques have been employed for the evaluation of erosion susceptibility by defining the geostatistical relationships between the geographical variability of selected physical attributes and the spatial distribution of the evidence for the water-induced erosion processes. Thus, allowing the generation of maps defining the susceptibility levels and expressing the relative probability of future occurrence of erosion landforms (Conoscenti et al. 2008). Subsequently, the susceptibility analyses advanced with various statistical and machine learning methods such as logistic regression, analytical hierarchy processes, bivariate statistics, boosted regression trees, artificial neural network, random forest, weights-of-evidence, index of entropy, information value, maximum entropy, and frequency ratio for spatial prediction of susceptibility to various erosion types (Akgün and Türk 2011; Conforti et al. 2011; Conoscenti et al. 2014; Lucà et al. 2011; Pourghasemi et al. 2017; Rahmati et al. 2016, 2017). However, some machine learning techniques such as SVM are scarcely used in erosion susceptibility mapping despite its predictive capability. Pradhan (2013) emphasized that both the choice and quality of CFs and the effectiveness of modeling techniques could impact on the susceptibility accuracy. Analysis of the existing literatures indicated lack of specific guidelines in selecting CFs for susceptibility studies (Lee and Pradhan 2007; Tehrany et al. 2015). Therefore, various researchers have applied different CFs (with varying number and types) and modeling techniques for developing susceptibility maps. Extensive review by Prosdocimi et al. (2016) identified CFs as one of the present knowledge gaps in soil erosion

modeling. Researchers believed that the accuracy of susceptibility mapping increases with increase in the number of CFs (Donati and Turrini 2002; Tehrany et al. 2015). Conversely, Remondo et al. (2003) opined that increasing number of CFs does not necessarily increase the susceptibility accuracy if the CFs involved are redundant. Survey of previous literature revealed dominate consideration of static CFs out of which many are redundant that could cause overweighing of the model results (Magliulo 2012). These factors do not respond to change in rainfall cycle and often remain unchanged for a relatively long period of time. However, previous studies are devoid of some crucial dynamic CFs associated with rainfall cycles despite roles in water-induced soil erosion. Therefore, this study aimed to develop geospatial prediction model for erosion susceptibility analysis using SVM technique for low and peak rainfall cycles in a complex watershed of Cameron Highlands, Malaysia. The region, located at 4°19'–4°37'N and 101°21'–101°30'E, is a mountainous landscape with an average elevation of approximately 1180 m, highest at Gunung Brinchang, of about 2032 m above mean sea level (IEA 2006). About 25% of the watershed have elevation less than 1000 m, while two-thirds fall within the range of 1100–1600 m (Gasim et al. 2009) and many of the other peaks are above 1524 m. Cameron Highlands is classified as a complex landscape which is predominantly steep with about 60% of land areas steeper than 20° (Fortuin 2006). This, coupled with high rain, makes it highly susceptible to gully erosion and landslides (Pradhan 2010). Eighty-one percent of the area has high erosion risk due to increasing human activities and nature of the watershed (World Wildlife Fund Malaysia 2002). The adverse impacts currently experienced include deterioration of water quality (Jaafar et al. 2010; Jamil et al. 2014; Khalik et al. 2013), landslide on steep slope terrains (Basith 2011; Conoscenti et al. 2008), and siltation of rivers and built water bodies, i.e., Ringlet Dam, which sometimes triggers flood reducing hydropower generation (IEA 2006; Jansen et al. 2012). Sustainable land-use management in a complex and fragile landscape like Cameron Highlands remains challenging to all the stakeholders (Aminuddin et al. 2005). In mitigating this

menace, susceptibility assessment is required to geospatially predict active/potential erosion distribution and identify critical locations.

31.2 Materials and Methods

31.2.1 Causative Factors for Gully Erosion Process

Gully erosion process is majorly caused by surface runoff which is channeled across an unprotected land and washes away the available soil particles along the drainage lines. This process is often influenced by different watershed characteristics such as land-use/cover, soil type, rainfall characteristics, etc. Thus, several CFs have been implemented by different researchers for susceptibility analyses and mapping. This study considered topographic, land-use/cover, geology, hydrology, and climatic and soil factors which are broadly grouped into static and dynamic causative factors (CFs). These were obtained from different sources as presented in Table 31.1. The CFs were selected due to their impacts on soil erosion processes and previous implementation in erosion susceptibility analysis. The non-redundant static CFs considered are drainage density (Kavzoglu et al. 2015; Pal 2016; Vijith et al. 2012), length slope (Conoscenti et al. 2013; Lucà et al. 2011), lineament density (Pradeep et al. 2015; Vijith et al. 2012), and soil type (erodibility) (Conoscenti et al. 2013; Dube et al. 2014), and their detailed descriptions could be found in the aforementioned literatures. Similarly, dynamic CFs such as *R*-factor (Magliulo 2012), LST (Xue et al. 2011), and SMI (Jamali 2004; Vahabi and Nikkami 2008) have been highlighted by researchers to influence erosion. However, these dynamic factors are rarely considered in the susceptibility analysis despite their roles in erosion processes. Hence, they were introduced to achieve a valid and accurate erosion susceptibility mapping. The preprocessing and processing procedures for the extraction of both static and dynamic CFs are described in the following sections.

Table 31.1 Spatial data source and their purpose

S/ no.	Spatial data	Format	Source	Purpose used for
1	DEM (5 m resolution)	Digital	Global Pixel Solutions, Malaysia	For extraction of static factors such as length slope (LS) factor, drainage density
2	Soil map	Digital	Malaysian Department of Agriculture, US Food and Agriculture Organization (FAO)	For extraction of soil erodibility factor
3	Multispectral Landsat-8 data (30 m resolution)	Digital	United States Geological Survey (USGS) database	For extraction of factors: land surface temperature (LST), soil moisture index (SMI), lineament, and NDVI
4	Rainfall data (daily data)	Excel format	Malaysian Department of Irrigation and Drainage	For computation and development of rainfall erosivity (<i>R</i> -factor)

31.2.1.1 Extraction of Lineament Density

Due to the roles of lineament density in soil erosion, it was considered in erosion susceptibility modeling. Lineaments were derived from atmospherically corrected Landsat-8 data using ArcMap[®] 10, ENVI 5.3, and Geomatica 2016 software. The following are the procedures for the extraction:

1. The principal component image (PCI) of Landsat-8-pan-sharpened reflected bands was used, as this carries most information suitable for lineament extraction. The pan-sharpened layer of Landsat-8 OLI bands was processed with Gram-Schmidt sharpened surface reflectance (GSSR). This serves as the input file into ENVI 5.3 to obtain PC1 (8-bit grayscale 15 m spatial resolution). The GSSR layer was loaded into ENVI 5.3 Classic in this sequence: *transform > principal_component > forward_PC_ rotation > compute_new_statistics_and_rotation* to obtain PCA. The output PCI was saved as an image file with a resolution of “8-bit grayscale” in TIFF/GeoTIFF format.
2. For the “automatic lineament” extraction, “LINE Module Control Panel” of Geomatica 2016 software was used to further process PC1 to extract the lines. PC1 was launched into Geomatica platform in this sequence: *Tool > Algorithm_Liberian > Line: Lineament Extraction > LINE_Module_Control_Panel*. LINE module control panel settings were adjusted appropriately before running the program. The extracted lineaments were saved as ArcView Shapefile (.shp) as “lineaments” for subsequent processing in the ArcMap[®] software. The PCI-LINE module extracts lineaments from the single 8-bit image through three basic steps: edge detection, thresholding, and curve extraction.
3. The output from Step 2 was processed in ArcMap[®] with split-line model to split the compounded lines into simple lines and edit the lineament attributes. The respective results from Steps 1 and 2 which are PC1 and extracted lineaments were both loaded onto ArcMap[®] platform. The lines were checked for the presence of compounded lines using editor command in ArcMap[®]. However, compounded lines are inevitable in such a large study area. In the context of lineament, compounded lines are the ones that composed of segmented lines which need to be splitted at the respective vertices. This was done by using “Split Line At Vertices” from data management on python modeler of ArcMap[®] to produce simple straight lines with two endpoints. Then, the file was renamed as “lineaments_split”. Attribute Table of the “lineaments_split” was edited to add coordinates (x_1 , y_1) and (x_2 , y_2) to the endpoints of each line. Subsequently, lineament density was estimated around each grid cells within a specific search area using “Line Density” command of Spatial Analyst Tool to display the output raster file (line_density) using “lineaments_split” as input

polyline features. Area unit and output cell size were set to square kilometers and 30, respectively, before running the program for line_density computation. The output (line_density) map was re-sampled using cubic convolution method during display.

31.2.1.2 Extraction of Length Slope Factor

Length slope (LS) is a topographic factor comprised of slope gradient and slope length factors derivable from digital elevation model (DEM). A relatively high-resolution DEM of 5 m was used for its derivation following the procedures presented below:

1. Prior to application of raw DEM, it was first assigned a Universal Transverse Mercator (UTM) projection, Northern zone 47 datum WGS 1984 coordinate system applicable to the study area.
2. The projected DEM was hydrologically corrected by filling the sinks. Then, the flow direction and flow accumulation were determined using flow direction tool in “ArcHydro tool” extension of ArcMap[®] 10.
3. The resulting map obtained was then used for the evaluation of slope map. Finally, LS-factor map was developed using Moore and Burch’s Equation defined in Eq. (31.1) (Parveen and Kumar 2012).

$$LS = \left(\text{FlowAcc} \times \frac{\text{resolution}}{22.1} \right)^{0.6} \times \left(\sin(\text{SlopeDem}) \times \frac{0.0174}{0.09} \right)^{1.3} \quad (31.1)$$

where flowAcc is the flow accumulation and SlopeDem is the slope derived from DEM.

31.2.1.3 Extraction of Drainage Density

Drainage density map was extracted from hydrologically corrected DEM following the procedures described as follows:

1. The “logical option” of “math tool” of ArcMap was used to select the predefined stream lengths to be captured in the stream extraction. All stream features greater than 1000 m were selected using flow accumulation as the input file.
2. In the “hydrology tool,” flow lengths were estimated using “stream lengths” greater than 1000 m as input file.
3. Using the “flow lengths” and “flow direction” as the input files, “stream links” and “stream order” were evaluated.
4. “Stream order” was then converted to “stream feature” for the development of drainage density map for the watershed.

31.2.1.4 Development of Rainfall Erosivity Factor

Daily rainfall data for the weather stations within and around the Cameron Highlands were obtained from Malaysia Department of Irrigation and Drainage (DID) for a period of 11 years (2006–2016). The missing rainfall data were fixed using nearest neighborhood interpolation technique. A widely applied mathematical model (in Eq. 31.2) proposed by Wischmeier and Smith (1958) for the estimation of R -factor from monthly rainfall records was adopted. The model was applied for each of the stations within and around the study area to evaluate the corresponding erosivity values. The estimated R -factor values for the stations with known weather stations' coordinates were imported into ArcMap[®]. Interpolation was carried out for other locations within the watershed using simple kriging interpolation technique to develop R -factor map.

$$R = \sum_{i=1}^{12} 1.735 \times 10^{1.5 \log_{10} \left[\frac{P_i}{P} - 0.08188 \right]} \quad (31.2)$$

where R is the rainfall erosivity factor (MJ mm/ha/year), P_i is the monthly average rainfall for i th month (mm), and P is the annual average rainfall (mm) (Khosrokhani and Pradhan 2014).

31.2.1.5 Development of Soil Erodibility Factor

Soil erodibility is one of the key parameters to be considered in soil erosion susceptibility analysis (Wang et al. 2013). It measures the proneness of a particular soil type to erosion by actions of rainfall and runoff (Ali and Hagos 2016). Digital soil map for Cameron Highlands was extracted from the map obtained from the Malaysia Department of Agriculture and digital soil map of the world (DSMW) for K -factor estimation and development of its map. William's approach for the computation of K -factor (Eq. 31.3) was adopted in this study. The detailed equations and descriptions for estimating component factors are present in Eqs. (31.4)–(31.8) (Wawer et al. 2005). The required corresponding soil properties for applying William's method were extracted from the DSMW database. The estimated K -factor was then exported to ArcMap environment for the development of erodibility map for the entire Cameron Highlands watershed.

$$K_{\text{factor}} = f_{\text{coarse sand}} \times f_{\text{clay silt}} \times f_{\text{org}} \times f_{\text{hi. sand}} \quad (31.3)$$

where $f_{\text{coarse sand}}$ is the factor for high coarse-sand content, $f_{\text{clay silt}}$ is the factor for high clay-to-silt ratios, f_{org} is the factor for high organic carbon content, and $f_{\text{hi. sand}}$ is for soils with extremely high sand content.

$$f_{\text{coarse sand}} = \left[0.2 + 0.3 \times \exp \left(-0.256 \times m_s \left(\frac{100 - m_{\text{silt}}}{100} \right) \right) \right] \quad (31.4)$$

$$f_{\text{clay silt}} = \left[\frac{m_{\text{silt}}}{(m_c + m_{\text{silt}})} \right]^{0.3} \quad (31.5)$$

$$f_{\text{org}} = \left[1 - \frac{0.25 \times \text{OC}}{\text{OC} + \exp(3.72 - 2.95 \times \text{OC})} \right] \quad (31.6)$$

$$f_{\text{hi. sand}} = \left[1 - \frac{0.7 \times A}{A + \exp(-5.51 + 22.9A)} \right] \quad (31.7)$$

$$A = \left[\frac{(100 - m_s)}{100} \right] \quad (31.8)$$

where m_s is the percentage of sand fraction defined to be in the range 0.05–2.00 mm in diameter, m_c is the percentage of clay fraction defined to be <0.002 mm in diameter, m_{silt} is the percentage of silt fraction defined to be in the range 0.002–0.05 mm in diameter, and OC is the percentage of organic carbon present in the soil sample.

31.2.1.6 Extraction of Normalized Difference Vegetation Index

Normalized difference vegetation index (NDVI) map was extracted from Bands 4 (RED) and 5 (near infrared) of Landsat-8 image. In Band 4, vegetation has very high reflectance within the spectral band, while built-up areas have a very high reflectance in Band 5. Prior to application of these Bands, they were treated for atmospheric and sun angle corrections by converting band digital number (DN) to “top of atmospheric (ToA) reflectance” using Eq. (31.9). This was subsequently corrected for sun angle using Eq. (31.10) and the NDVI map evaluated using Eq. (31.11).

$$L_{\lambda} = M_L Q_{\text{cal}} + A_L \quad (31.9)$$

$$\text{DN}_{\text{corr.}} = \frac{\text{DN}_{\text{atm}}}{\sin(\phi)} \quad (31.10)$$

$$\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}} + 1 \quad (31.11)$$

where L_{λ} is ToA spectral radiance with SI unit watt/(m² s rad μm^{−1}), M_L is the radiance multiplicative band number, A_L is the radiance additive band number, Q_{cal} is the quantized and calibrated digital number (DN), and ϕ is the sun angle.

31.2.1.7 Development of Land Surface Temperature

A relatively clear Landsat-8 data was acquired from USGS archive for the study which is located on Row/Path: 57/127. The thermal Band 10 was used for retrieving LST for the Cameron Highlands. The brightness temperatures are measured from the ToA radiance (L_λ) of satellite's thermal infrared (TIR) sensors using Planck's law (Friedl and Davis 1994). TIR band digital number was converted to ToA spectral radiance using the algorithms in Eq. (31.9). The equations required for the computation of land surface emissivity (e), brightness temperature (BT), proportion vegetation (Pv), and land surface temperature (LST) were given, respectively, in Eqs. (31.12)–(31.15). The detailed procedures can be found in Abdulkadir et al. (2017).

$$BT = K_2 / \ln \left(\frac{K_1}{L_\lambda} + 1 \right) \quad (31.12)$$

$$e = 0.004Pv + 0.986 \quad (31.13)$$

$$Pv = \left[\frac{(NDVI - NDVI_{\min})}{(NDVI_{\max} - NDVI_{\min})} \right]^2 \quad (31.14)$$

$$LST = \frac{BT}{\left(1 + w \times \frac{BT}{p} \times \ln(e) \right)} \quad (31.15)$$

where K_1 and K_2 are the band specific thermal conversion constants, p is the Planck's constant, $NDVI_{\max}$ and $NDVI_{\min}$ are maximum and minimum NDVI values obtainable from NDVI map, M_L is the radiance multiplicative band number (rescaling factor), A_L is the radiance additive band number (rescaling factor), and Q_{cal} is the quantized and calibrated digital number (DN).

31.2.1.8 Development of Soil Moisture Index

Having developed NDVI and LST maps as previously described, pixel-based values for each of the parameters were extracted and plotted against each other. The soil moisture index (SMI) was evaluated by defining parameters $LST_{i, \max}$ and $LST_{i, \min}$ for a given NDVI plot. From LST-NDVI space diagram, linear regression equations for the dry and wet edges were obtained. These are typically represented as in Eqs. (31.16) and (31.17). The LST_i empirical parameters for dry and wet edges, when modeled as a linear fit to the observation data (i.e., constants a , b , a' , and b'), were determined from a two-dimensional scatter plot of LST against NDVI. Hence, SMI on pixel basis was evaluated with Eq. (31.18).

$$LST_{i, \min} = a + b \times NDVI_i \quad (31.16)$$

$$LST_{i, \max} = a' + b' \times NDVI_i \quad (31.17)$$

$$SMI = \frac{LST_{i, \max} - LST_i}{LST_{i, \max} - LST_{i, \min}} \quad (31.18)$$

where $LST_{i, \max}$ and $LST_{i, \min}$ are maximum and minimum surface temperatures for a given NDVI. LST_i is the observed instantaneous surface temperature at the specific pixel. The constants a , b , a' , and b' are empirical parameters evaluated from the dry and wet edges of linear fits (Ridd 1995).

31.2.2 Support Vector Machine Technique for Susceptibility Modeling

Support vector machine (SVM) is a statistical learning technique proposed by Sujay and Pares (2014) for solving classification and regression problems by constructing n -dimensional hyperplanes that optimally separate data values into categories (Hearst et al. 1998). It is one of the most sophisticated supervised techniques in artificial intelligence derived from statistical learning theory and structural risk minimization principle (Tehrany et al. 2015; Vapnik 1998) whose accuracy exceeds many other data mining, bivariate and multivariate techniques (Sujay and Pares (2014)). It uses kernel tricks on training dataset such as linear, sigmoid, polynomial, or radial basis function (RBF) with their parameters (tolerance, γ ; polynomial degree, d ; and cost function, C) (Bui et al. 2012; Tehrany et al. 2015) to reshape nonlinear problems into linear and process-able classes by generating separating hyperplanes (Jebur et al. 2014). The flowchart for the implementation of SVM for soil erosion susceptibility modeling is depicted in Fig. 31.1, and detailed procedures are defined in the following sections:

31.2.2.1 Data Preparation and Partitioning for SVM Application

The first stage in performing soil erosion susceptibility mapping is to develop erosion inventory map followed by development of geodatabase of erosion-related CFs. Some of the surveyed gully and rill erosion locations in the study area are shown in Fig. 31.2a–d. These were documented during multiple field survey for the development of inventory map for the area under consideration. In a watershed, there is possibility of occurrence (denoted with 1) or non-occurrence (denoted with 0) of soil erosion at different locations. This makes the response variable to be binary or dichotomous in nature. It is recommended to have

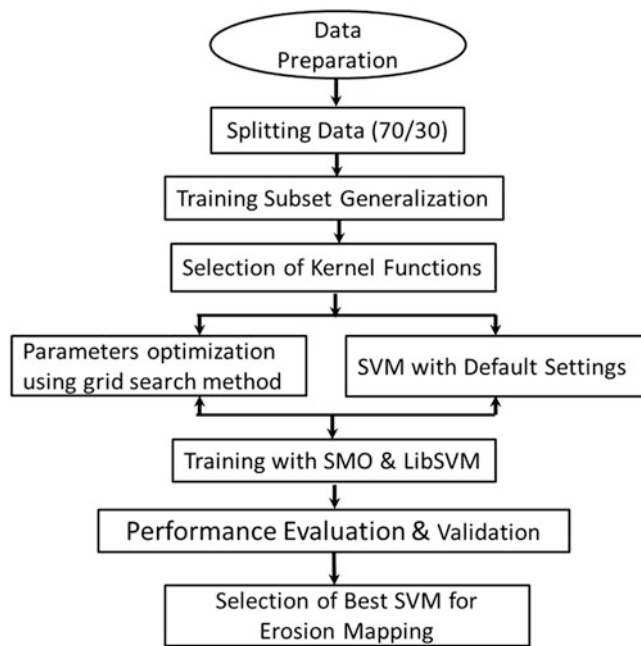


Fig. 31.1 Flowchart for application of SVM

approximately equal number of erosion-present points (1) and erosion-absent (0) points (Ayalew and Yamagishi 2005). It is mandatory that all the CFs have the same projection for easier acquisition of data. Moreover, Conoscenti et al. (2014) recommended that all CFs should have uniform scale/extent (number of columns and rows) and range as the dependent variable. From the developed erosion inventory map, equal number of erosion-absent points were randomly created using slope map as a guide. All the CFs were projected to UTM 47 North, WGS 84, and have similar grid sizes of 30 by 30 m and column by row of 1025×1112 . Based on the response variable, the corresponding erosion CFs' (predictor variables) data values were generated for SVM modeling.

In preprocessing stage, the generated dataset was partitioned into training and testing datasets using unsupervised filtering algorithms to avoid duplication. Each of these datasets is composed of both response and predictor variables. Trend in the previous studies on application of machine learning showed that there is no specific rule of thumb for the proportion of training and testing datasets. However, Guyon (1996) recommended that the proportion for validation (testing) set should be inversely proportional to the square root of the number of free adjustable parameters. The testing set (v) to training set (t) size ratio, v/t , scales yielded 33.33% proportion of the dataset for testing, due to the consensus of the researchers to have more training dataset than testing. Thus, the dataset was partitioned into 70% training and 30% testing datasets. This conforms with the proportion of dataset considered in similar studies (Jebur et al. 2014; Tehrany et al. 2014, 2015).

31.2.2.2 Selection of SVM Kernel Functions, Kernel Parameters, and Its Optimization

The nature of environmental problems like soil erosion is often nonlinear due to complexity of the watershed characteristics. Thus, it is imperative to apply different kernel functions such as linear, polynomial, radial basis function (RBF), and sigmoid. Selection of kernel type is very vital in SVM modeling due to their impacts on the successful training and classification precision (Tehrany et al. 2015). After the preparation of dataset as described in Sect. 31.2.2.1, the kernel functions were selected and implemented one after the other. For each of the kernel function, the corresponding kernel parameters described in Table 31.2 were defined. The C controls the trade-off between training errors and margin that often prevent over-fitting of the model (Marjanović et al. 2011), while γ controls the degree of nonlinearity of the SVM model (Bui et al. 2012). It is recommended that these parameters be predefined for specific studies to obtain accurate modeling results (Pourghasemi et al. 2013). In most studies, it's often difficult to select kernel parameters that will produce optimum model results particularly when modeling nonlinear problems. Hence, parameter optimization is indispensable to select optimum parameters. In this study, grid-search method was adopted for selecting optimal cost parameter, C ; tolerance, γ ; and degree of polynomial, d .

31.2.2.3 SVM Susceptibility Modeling and Performance Evaluation

The training dataset (x_i, y_i) are used in mapping the input dataset into high dimensional feature space. The pair of training dataset (x_i, y_i) for $x_i \in R^n$, $y_i \in (1, 0)$ for $i = 1, 2, \dots, m$. The elements x are the array of erosion CFs consisting of LS-factor, R -factor, K -factor, lineament density, drainage density, NDVI, LST, and SMI. The separating hyperplane is generated in the original space of n coordinates (x_i parameters in vector x) between the points of two distinct classes (Marjanović et al. 2011). The two classes are erosion-present and erosion-absent pixels represented by 1 and 0, respectively, for the training dataset. If the point is located over the hyperplane, it will be classified as 1, otherwise, 0. Hence, SVM is implemented to establish optimal separating hyperplanes that would classify the training dataset into the two pixel classes. The training topology for this study is depicted in Fig. 31.3.

The most popular WEKA packages, LibSVM and SMOreg algorithms, were used for SVM classification or regression. The packages can handle both linear and nonlinear SVM analyses. Sequential minimal optimization (SMO) is one of the optimization algorithms for the training of a given dataset. It uses heuristics to partition the training dataset into smaller units that can be solved analytically. The training of dataset was applied to find Lagrange multipliers (α_i) by maximizing a dual quadratic program

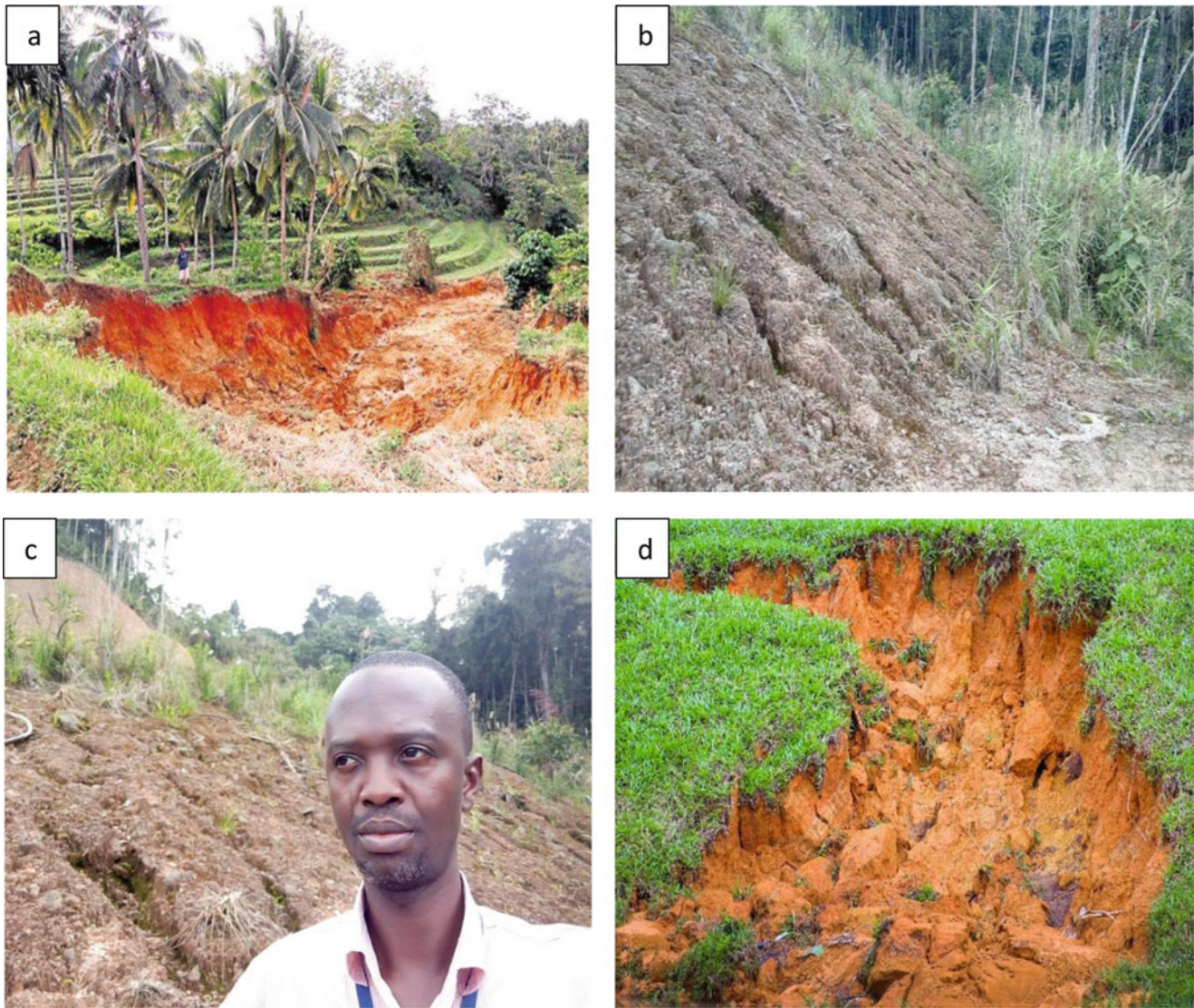


Fig. 31.2 Some erosional landforms in the study area

Table 31.2 Different kernel functions (Bui et al. 2012; Tehrany et al. 2015)

Kernel function $K(x_i, x_j)$	Kernel representation	Kernel parameter
Linear, LNR	x_i^T, x_j	–
Polynomial, Poly	$(-y \cdot x_i^T \cdot x + C)^d$, C is often given as 1	y, d
Sigmoid, Sigmd	$\tanh((-y \cdot x_i^T \cdot x + C)^d)$	y
Radial basis function, RBF	$\exp(-y x_i - x_j^2)$, for $y > 0$	y

(QP). The SMO algorithms search through the feasible region of the dual quadratic problems and maximize Eq. (31.19) by decomposing it into fixed size QP subproblems and solving the smallest possible optimization problem at each step (Hearst et al. 1998; Platt 1998).

$$\begin{aligned} \max L_D(\alpha) = & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \end{aligned} \quad (31.19)$$

Subject to $0 \leq \alpha_i \leq C$

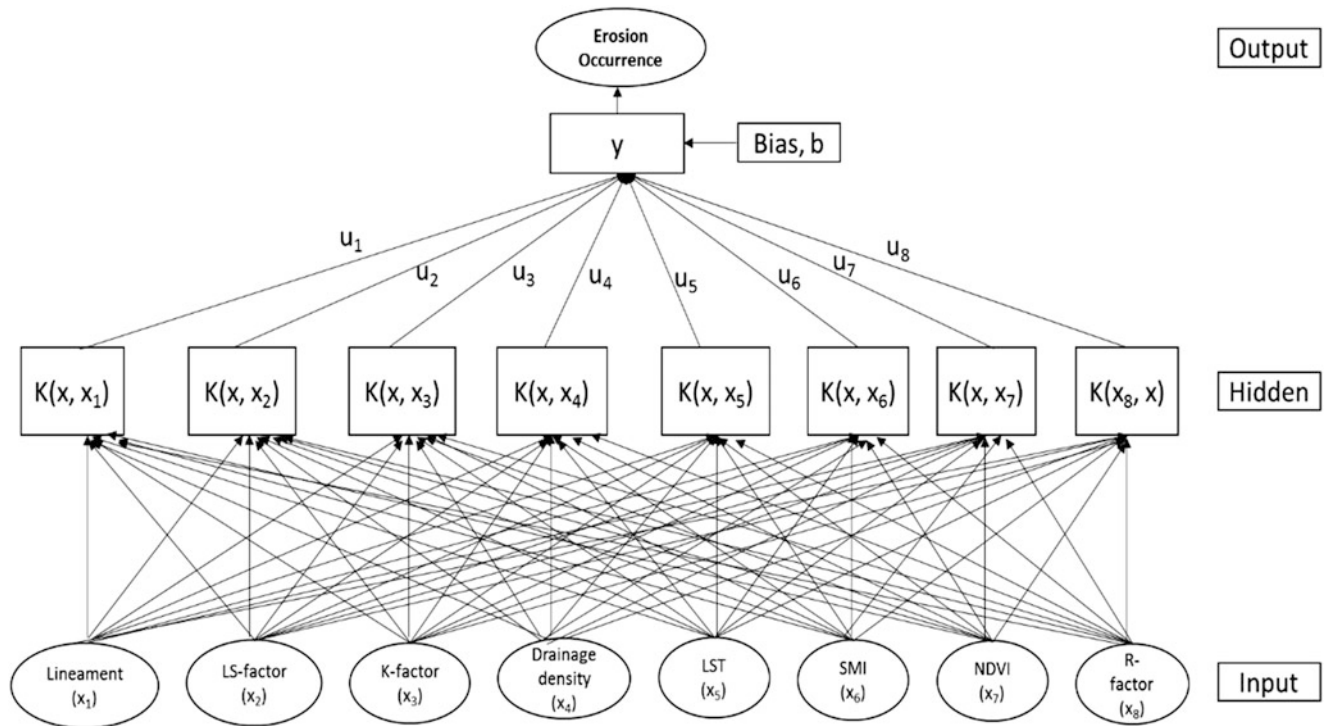


Fig. 31.3 Training architecture for erosion susceptibility

$$\sum_{i=1}^n \alpha_i y_i = 0, \forall_i$$

where C is the SVM hyperparameter and $K(x_i, x_j)$ is kernel functions.

The optimized kernel parameters were used for the respective kernel functions for the classification purposes. SVM in WEKA requires that one of the classification training approaches must be selected for the training of the model. Hence, model performance in terms of success and prediction rates can be evaluated for the training and testing datasets, respectively. There are four available training techniques in WEKA machine learning workbench which include training dataset, supplied test data, percentage split, and cross validation techniques. In cross validation technique, the dataset was partitioned into n -folds ($n = 10$ for this study). The training will be conducted on all the partitions except one that is held out as the test set. The repetition of this process will create n -number of the models and avail each fold the chance of being held out as test set. The cross validation was selected in this study for the training of the model because of its accuracy over the others. Then, the success rate of the model was evaluated. In the same manner, test dataset was introduced through “supplied test data” option of the work bench to evaluate the prediction rate of the model with the test dataset.

Finally, performance of the models were evaluated with the indexes: area under the curve (AUC), kappa index (k),

prediction rate (PR), success rate (SR), and classification errors (Althuwaynee et al. 2012; Bui et al. 2012; Jebur et al. 2014; Shit et al. 2015; Tehrany et al. 2014, 2015). Comparative analysis among all the SVM kernel functions was conducted to evaluate their relative performances. In the end, SVM classification results were used for the development of erosion susceptibility mapping for Cameron Highlands.

31.2.3 Erosion Susceptibility Modeling

Susceptibility technique is used to assess the relative probability of occurrence of erosion at a particular location compared to other locations under the influence of triggering factors known as CFs. Eight CFs were considered as predictor variables and a response variable (which is a binary class composed of eroded and non-eroded pixels). The CFs were grouped into static and dynamic factors. In order to evaluate the seasonality impact of dynamic CFs on the erosion susceptibility, the low and peak rainfall cycles were considered since the set of CFs respond to change in rainfall cycles. Hence, Landsat images for February 26, 2016, and November 8, 2016, were downloaded representing the respective low and peak rainfall cycles. NDVI, LST, R -factor, and SMI were extracted for these two periods for multitemporal analysis of erosion susceptibility mapping. The erosion base map in Fig. 31.4a that consists of

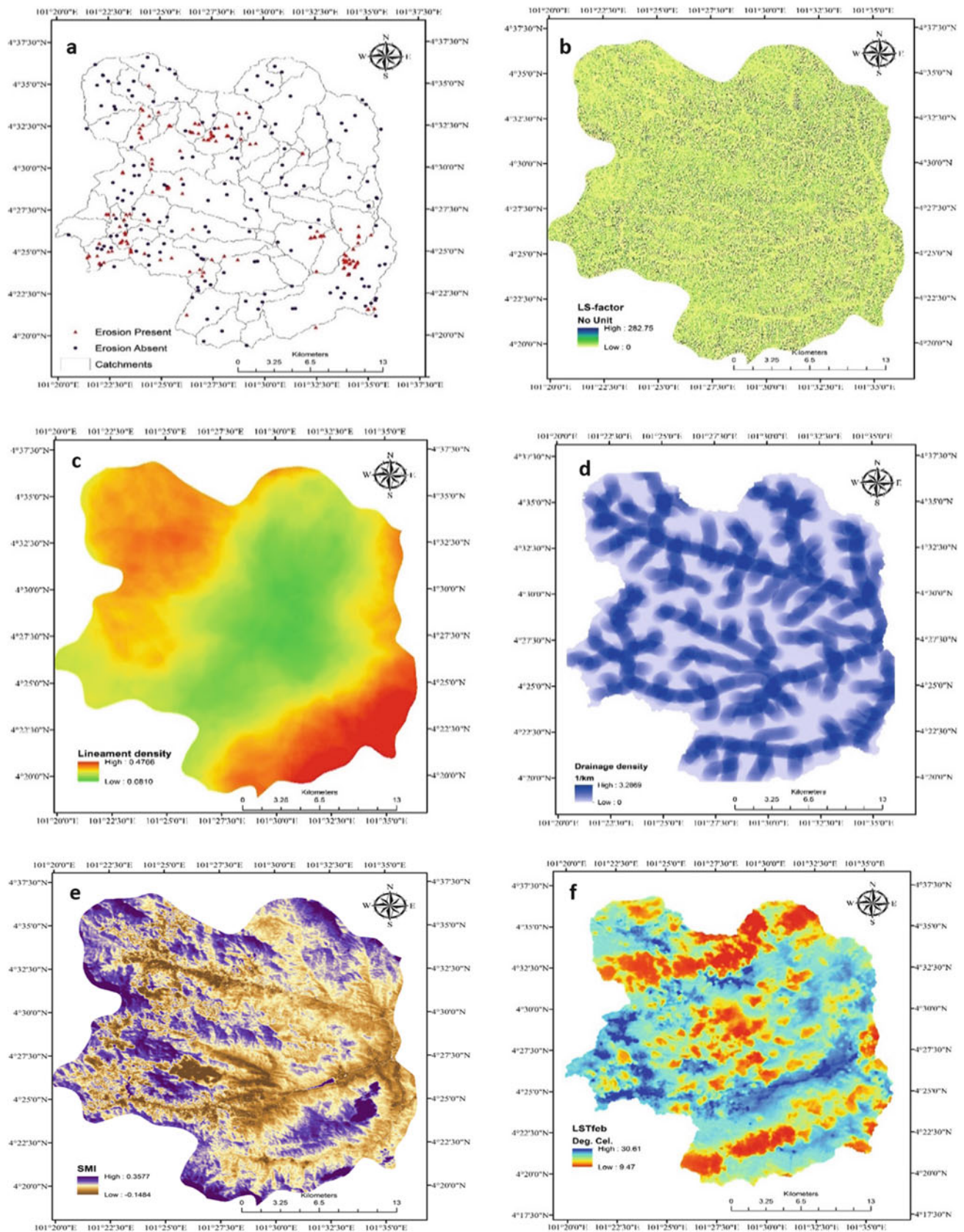


Fig. 31.4 (a) Inventory map. (b) Length slope factor. (c) Lineament density. (d) Drainage density. (e) Soil moisture index. (f) Land surface temperature. (g) NDVI. (h) Rainfall erosivity. (i) Soil erodibility factor

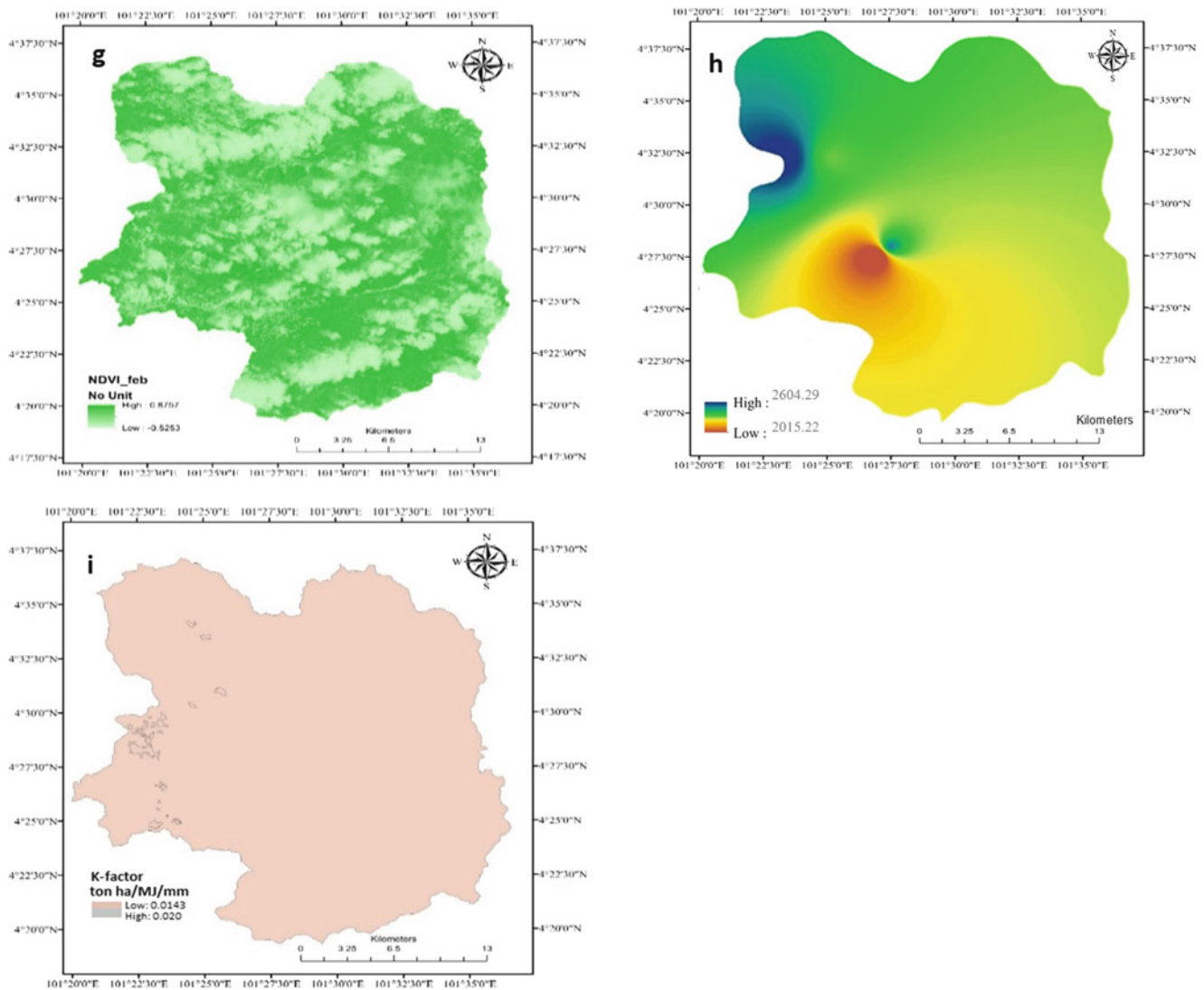


Fig. 31.4 (continued)

159 erosion-present and equal numbers of erosion-absent locations was randomly created across the watershed. The erosion-present locations are made up of both gullies and rill erosional landforms. The weights for the CFs of the model having the best performance were applied for development of erosion susceptibility maps for the study area.

31.3 Results and Discussion

The analysis involved nine attributes (i.e., 8 CFs as the predictor and the response variables, binary classes of absence or presence) with total of three hundred and eighteen cases. The cases comprised of erosion-present and erosion-absent cases with their corresponding values of the erosion CFs. Figure 31.4b–i shows the geospatial distribution of all

the CFs within the study area. The dataset is composed of static and dynamic CFs for the month of low rainfall cycle. In order to achieve accurate modeling results, kernel parameters for each of the kernel functions were optimized. The cost function (C) and tolerance (γ) were optimized for the SVM-RBF and sigmoid SVM, C , γ , and the polynomial degree (d) for polynomial SVM and only C for linear SVM using grid search technique. Table 31.3 shows the optimized kernel parameter obtained which were used for SVM modeling. The non-zero positive C values obtained indicate that the polynomial kernel is inhomogeneous in nature. This is a free parameter trading off the influence of higher-order versus lower-order terms in the polynomial.

These parameters were applied on low rainfall cycle dataset for erosion susceptibility mapping and the results are presented in Table 31.4. The results therein comprised

Table 31.3 Optimized kernel parameters

Kernel	Cost parameter, C	Tolerance, γ	Polynomial degree, d
Linear SVM	10	–	–
Polynomial SVM	14	5.2	4
RBF SVM	12	1.6	–
Sigmoid SVM	16	2.0	–

Table 31.4 Kernel functions for low rainfall cycle

Kernel	SR	PR	Kappa index		AUC		MAE		RMSE	
			Training	Testing	Training	Testing	Training	Testing	Training	Testing
Linear SVM	0.9258	0.8778	0.8477	0.7552	0.828	0.838	0.1732	0.1214	0.2014	0.3091
Polynomial SVM	0.9287	0.8844	0.8512	0.7911	0.896	0.885	0.1001	0.1012	0.3029	0.5511
RBF SVM	0.8022	0.7104	0.6872	0.6324	0.804	0.791	0.3003	0.3321	0.4130	0.4014
Sigmoid SVM	0.6286	0.5527	0.1601	0.1840	0.510	0.501	0.4148	0.4011	0.6181	0.7047

of SR, PR, k , AUC, mean absolute error (MAE), and root mean squared error (RMSE) obtained for the training and testing for model performance evaluation. However, AUC and kappa index are the commonest validation criteria in susceptibility studies. The analysis shows an SR of 0.9258 representing the training cases that were correctly classified with respective MAE and RMSE of 0.1732 and 0.2014 for training dataset. However, the SR only is insufficient to valid the efficacy of the model because it gives model performance on the training dataset. Hence, PR was evaluated on the test dataset that were not included in the training exercise to ascertain their performance. Implementing the results on test dataset indicated a prediction rate of 0.8778. The respective classification errors MAE and RMSE were 0.1214 and 0.3091.

Using kappa index and AUC as validation criteria showed that polynomial SVM outperformed other models during testing followed by linear SVM. The performance of polynomial SVM suggests the nonlinearity and complexity of watershed characteristics with respect to soil erosion. This is slightly similar to the results obtained by Tehrani et al. (2015) who evaluated the capability of SVM kernel function in flood susceptibility mapping. Furthermore, sigmoid SVM had the least performance and the highest classification errors. The closeness of its AUC value to 0.5 and the k value to 2.0 underscores its weakness in erosion susceptibility analysis and mapping (Althuwaynee et al. 2012; Tehrani et al. 2014).

Table 31.5 presents the weights obtained for each of the CFs for different kernel functions. These underscore the strength of each CFs which appear slightly similar. However,

their magnitudes differ from one another. The results showed that all the CFs had positive relationships with occurrence of erosion except NDVI. From the three kernel functions, topographic landscape appeared to be most significant of all the CFs. Specifically, in polynomial SVM, LS-factor's weight of 2.632 was the highest followed by lineament density. Being the best model, the corresponding weights of polynomial SVM were considered for the development of soil erosion susceptibility map (Fig. 31.4).

From the developed erosion susceptibility map (Fig. 31.5), the classes are very low, low, moderate, high, and very high susceptible zones constituting 3.6%, 68.8%, 25.1%, 1.8%, and 0.7% of the total area of Cameron Highlands watershed, respectively. The output map showed most of the areas are in the categories of “low and moderate” susceptibilities. It was observed that the areas within “high and very high” susceptibilities fell in the urban and agriculture areas. This underscores the results of human disturbance in the watershed thereby causing land degradation.

Similar analyses were performed for the peak rainfall cycle dataset. The previous dynamic CFs for the low rainfall cycle were replaced with that of the month of November (i.e., peak rainfall cycle) for another round of SVM susceptibility modeling for multitemporal analysis. However, static CFs remain unchanged because they do not respond to rainfall cycles. With optimized kernel parameters, SVM classification was carried out for all the kernel functions. The results of the analyses are presented in Table 31.6. The results also showed that polynomial SVM had the best performance among all the SVM models.

Table 31.5 Variable of importance of each CFs for low rainfall cycle

Weights	K -factor	LS-factor	Lineament density	Drainage density	LST	SMI	NDVI	R -factor
Linear SVM	0.027	2.572	1.520	0.214	1.041	1.423	–3.133	0.175
Polynomial SVM	0.021	2.632	1.641	0.201	1.442	1.376	–3.820	0.153
RBF SVM	0.043	2.632	1.492	0.234	1.984	1.403	–2.811	0.148

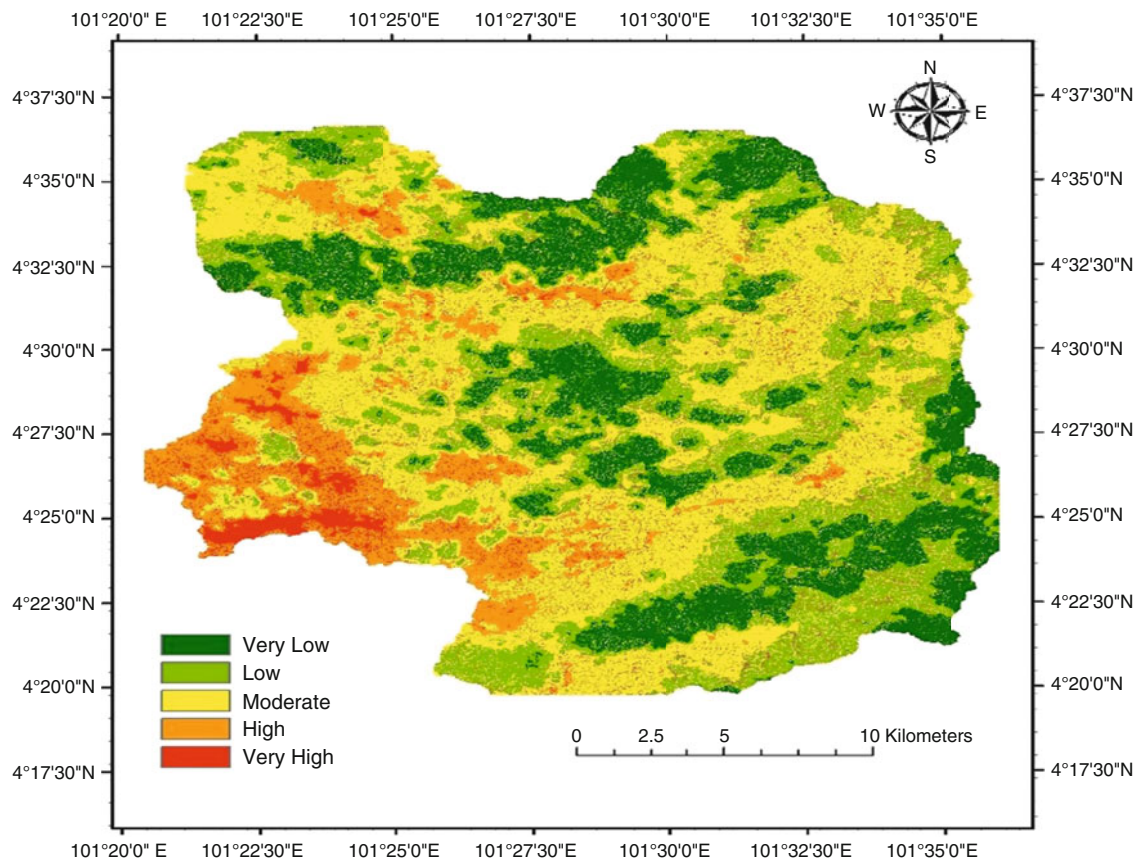


Fig. 31.5 Soil erosion susceptibility map for polynomial SVM-Poly for low rainfall cycle

Table 31.6 Kernel functions peak rainfall cycle

Kernel	SR	PR	Kappa index		AUC		MAE		RMSE	
			Training	Testing	Training	Testing	Training	Testing	Training	Testing
Linear SVM	0.8901	0.8687	0.7828	0.7568	0.7924	0.8001	0.1661	0.1003	0.1988	0.2145
Polynomial SVM	0.9226	0.8830	0.8642	0.8012	0.8226	0.8605	0.0913	0.1001	0.1093	0.2111
RBF SVM	0.8837	0.8855	0.7011	0.7112	0.7320	0.7522	0.2002	0.2046	0.2210	0.3441
Sigmoid SVM	0.5843	0.5088	0.1982	0.2002	0.5013	0.5116	0.2133	0.2822	0.4008	0.5145

Table 31.7 shows the CFs' weights obtained for each of the SVM models highlighting the strength and types of the relationships existing between CFs and the occurrence of soil erosion. Comparing these results with that of low rainfall cycle showed relatively similar behaviors in terms of the strengths and nature of the relationships of the CFs with soil erosion. Since polynomial SVM outperformed others,

thus, their weights were used in developing erosion susceptibility map (Fig. 31.6) for Cameron Highlands. The only difference observed when compared with the previous analysis was in the percentage area susceptible to soil erosion. The percentage area susceptible to erosion during the peak rainfall cycle was observed to be 1.8%, 66.5%, 28.1%, 2.5%, and 1.1% for the very low, low, moderate, high, and very high

Table 31.7 Weights for CFs for each kernel functions for peak rainfall cycle

Weights	K-factor	LS-factor	Lineament density	Drainage density	LST	SMI	NDVI	R-factor
SVM-LNR	0.021	2.554	1.600	0.301	1.116	1.394	−3.164	0.224
SVM-Poly	0.027	2.661	1.688	0.202	1.443	1.365	−3.843	0.302
SVM-RBF	0.050	2.645	1.498	0.244	1.984	1.355	−2.755	0.146

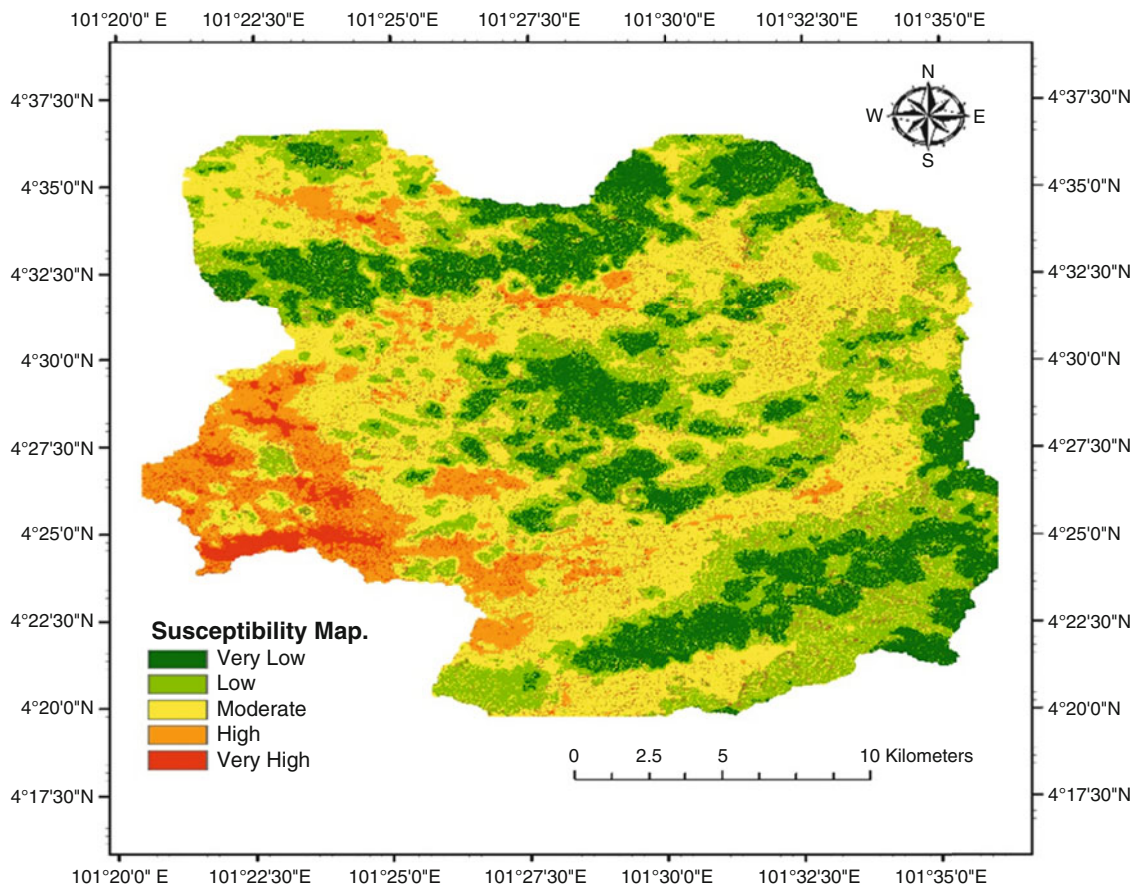


Fig. 31.6 Soil erosion susceptibility map for SVM-Poly for peak rainfall cycle

susceptible zones. This shows a slight increase in the area “high and very high” susceptible regions compare to low rainfall cycle period.

The distribution pattern of erosion susceptibility showed that majority of the areas are under “very low, low, and moderate” category of susceptibility. The analysis shows that most of the highly eroded locations are within the urban and agriculture land-use. This is because erosion rates are often enhanced at locations that are denuded and reshaped for urban and agricultural development. These susceptibility maps would provide early warning to the watershed users and other stakeholders on the proneness of some certain locations to soil erosion. This would greatly assist in sustainable management of the watershed that will alleviate soil erosion and its related hazards.

31.4 Conclusion

The study implements SVM for geospatial prediction of soil erosion under both static and dynamic CFs for low and peak rainfall cycles. This machine learning technique was applied to the complex watershed of Cameron Highlands to model its

erosion susceptibility. A total of eight CFs non-redundant CFs were selected including the unpopular dynamic CFs that are often neglected in many related studies. Four SVM kernel functions such as linear, polynomial, sigmoid, and RBF were implemented with optimized kernel parameters. The performances of SVM models were evaluated using AUC, SR, PR, and kappa index. The results indicated that polynomial SVM had the highest performance with sigmoid SVM having the lowest accuracy with relatively large classification errors. The performance of polynomial SVM on the other hand proved its suitability and superiority in modeling the complex and nonlinear soil erosion processes. Due to its outstanding predictive capability of this model, its corresponding weights for all CFs were used for the development of soil erosion susceptibility map for the study area. The analysis of the CFs considered in this study indicated that all had positive relationships with the occurrence of soil erosion except NDVI. The distribution patterns of erosion showed that majority of the area are “low and moderately susceptible.” The analysis shows that most of the highly susceptible locations are within the urban and agriculture land-use. This is because erosion rates are often prevalent where land are denuded and reshaped for anthropogenic activities.

Susceptibility analyses of both low and peak rainfall cycles yielded similar results except that there was increase in susceptible areas during the peak rainfall cycle. This study could provide information on the erosional landforms that could advance to gullies especially during the peak rainfall cycle. This medium-scale erosion susceptibility map produced would be of great benefits to the watershed managers in mitigating the environmental and socioeconomic-related challenges of soil erosion and thus contribute to achieving six SDGs earlier mentioned.

Acknowledgment This research was funded by Universiti Teknologi PETRONAS via 2016 URIF grant 0153AA-G04.

References

- Abdulkadir TS, Muhammad MR, Khamaruzaman WY, Ahmad HM (2017) Assessing the influence of terrain characteristics on spatial distribution of satellite derived land surface parameters in mountainous areas. In: 37th IAHR World Congress, Kuala Lumpur Malaysia, 13-18th August 2017. pp 2897-2906
- Agnesi V, Angileri S, Cappadonia C, Conoscenti C, Rotigliano E (2011) Multi-parametric GIS analysis to assess gully erosion susceptibility: a test in southern Sicily, Italy *Landform Analysis* 17:15-20
- Akgün A, Türk N (2011) Mapping erosion susceptibility by a multivariate statistical method: a case study from the Ayvalık region, NW Turkey *Computers & Geosciences* 37:1515-1524
- Ali SA, Hagos H (2016) Estimation of soil erosion using USLE and GIS in Awassa catchment, Rift valley, Central Ethiopia *Geoderma Regional* 7:159-166
- Allouche J, Middleton C, Gyawali D (2014) Nexus Nirvana or Nexus Nullity? A dynamic approach to security and sustainability in the water-energy-food nexus. In: STEPS Working Paper no. 63, STEPS Centre, Institute of Development Studies. University of Sussex, Brighton,
- Althuwaynee OF, Pradhan B, Lee S (2012) Application of an evidential belief function model in landslide susceptibility mapping *Computers & Geosciences* 44:120-135
- Althuwaynee OF, Pradhan B, Park H-J, Lee JH (2014) A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping *Catena* 114:21-36
- Aminuddin B, Ghulam M, Abdullah WY, Zulkefli M, Salama R (2005) Sustainability of current agricultural practices in the Cameron Highlands, Malaysia *Water, Air, & Soil Pollution: Focus* 5:89-101
- Arabameri A, Pradhan, B., Pourghasemi, H.R., Rezaei, K. and Kerle, N. (2018) Spatial Modelling of Gully Erosion Using GIS and R Programing: A Comparison among Three Data Mining Algorithms. *Applied Sciences* 8:1369
- Ayalew L, Yamagishi H (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan *Geomorphology* 65:15-31
- Barber M, Mahler R (2010) Ephemeral gully erosion from agricultural regions in the Pacific Northwest, USA. *Ann. Wars. Univ. Life Sci Land Reclam* 42:23-29
- Basith A (2011) Landslide susceptibility modelling under environmental changes: A case study of Cameron Highlands, Malaysia., Universiti Teknologi Petronas
- Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB (2012) Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS *Computers & Geosciences* 45:199-211
- Conforti M, Aucelli PP, Robustelli G, Scarciglia F (2011) Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (Northern Calabria, Italy) *Natural hazards* 56:881-898
- Conoscenti C, Agnesi V, Angileri S, Cappadonia C, Rotigliano E, Marker M (2013) A GIS-based approach for gully erosion susceptibility modeling: A test in Sicily, Italy *Environmental Earth Sciences* 70:1179-1195
- Conoscenti C, Angileri S, Cappadonia C, Rotigliano E, Agnesi V, Märker M (2014) Gully erosion susceptibility assessment by means of GIS-based logistic regression: a case of Sicily (Italy) *Geomorphology* 204:399-411
- Conoscenti C, Di Maggio C, Rotigliano E (2008) Soil erosion susceptibility assessment and validation using a geostatistical multivariate approach: a test in Southern Sicily *Natural Hazards* 46:287-305
- d'Oleire-Oltmanns S, Marzolf I, Peter KD, Ries JB (2012) Unmanned aerial vehicle (UAV) for monitoring soil erosion in Morocco *Remote Sensing* 4:3390-3416
- Dewitte O, Daoudi M, Bosco C, Eeckhaut MVD (2015) Predicting the susceptibility to gully initiation in data-poor regions *Geomorphology* 228:101-115
- Donati L, Turrini M (2002) An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: application to an area of the Apennines (Valnerina; Perugia, Italy) *Engineering Geology* 63:277-289
- Duarte L, Teodoro A, Gonçalves J, Soares D, Cunha M (2016) Assessing soil erosion risk using RUSLE through a GIS open source desktop and web application *Environ Monit Assess* 188:1-16
- Dube F, Nhapi I, Murwira A, Gumindoga W, Goldin J, Mashauri D (2014) Potential of weight of evidence modelling for gully erosion hazard assessment in Mbire District-Zimbabwe *Physics and Chemistry of the Earth, Parts A/B/C* 67:145-152
- Fortuin R (2006) Soil erosion in Cameron Highlands: An erosion rate study in a highland area Saxion University, Deventer, the Netherlands, Regional Environmental Awareness Cameron Highlands:1-83
- Friedl M, Davis F (1994) Sources of variation in radiometric surface temperature over a tallgrass prairie *Remote Sensing of Environment* 48:1-17
- Gasim MB, Surif S, Toriman ME, Rahim SA, Elfithri R, Lun PI (2009) Land-use change and climate-change patterns of the Cameron Highlands, Pahang, Malaysia *The Arab World Geographer* 12:51-61
- George CM, Anu VV (2018) Predicting piping erosion susceptibility by statistical and artificial intelligence approaches-a review *International Research Journal of Engineering and Technology* 5:239-243
- Gómez-Gutiérrez Á, Conoscenti C, Angileri SE, Rotigliano E, Schnabel S (2015) Using topographical attributes to evaluate gully erosion proneness (susceptibility) in two Mediterranean basins: advantages and limitations *Natural Hazards* 79:291-314
- Guyon I (1996) A scaling law for the validation-set training-set ratio. Unpublished Technical Report, AT&T Bell Laboratories
- Hearst MA, Dumais ST, Osman E, Platt J, Scholkopf B (1998) Support vector machines. *Intelligent Systems and their Applications IEEE Geoscience and remote sensing letters* 13:18-28
- Hugenholtz CH et al. (2013) Geomorphological mapping with a small unmanned aircraft system (sUAS): Feature detection and accuracy assessment of a photogrammetrically-derived digital terrain model *Geomorphology* 194:16-24
- IEA (2006) Hydropower Good Practices: Environmental mitigation measures and benefits case study 04-03: Reservoir sedimentation-cameron highlands hydroelectric scheme, Malaysia. Hydropower Implementing Agreement Annex VIII, International Energy Agency
- Jaafar O et al. (2010) Modeling the impacts of ringlet reservoir on downstream hydraulic capacity of bertain river using XPSWMM in

- cameron highlands, Malaysia Research Journal of Applied Sciences 5:47-53
- Jamali HMB (2004) Study of soil moisture in relation to soil erosion in the proposed Tamcitaro Geopark, Central Mexico: A case of the Zacandaro sub-watershed., International Institute for Geo-information Science and Earth Observation
- Jamil NR, Ruslan MS, Toriman ME, Idris M, Razad AA (2014) Impact of Landuse on Seasonal Water Quality at Highland Lake: A Case Study of Ringlet Lake, Cameron Highlands, Pahang. In: From Sources to Solution. Springer, pp 409-413
- Jansen L, Lariyah MS, Mohamed NMD, Pierre YJ (2012) Challenge in running hydropower as source of clean energy: Ringlet reservoir, Cameron Highlands case study. Paper presented at the Proceedings National Graduate Conference, Universiti Tenaga Nasional, Malaysia, November 8-10
- Jebur MN, Pradhan B, Tehrany MS (2014) Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale Remote Sensing of Environment 152:150-165
- Kavzoglu T, Sahin EK, Colkesen I (2015) Selecting optimal conditioning factors in shallow translational landslide susceptibility mapping using genetic algorithm Engineering Geology 192:101-112
- Kelley DW, Nater EA (2000) Historical sediment flux from three watersheds into Lake Pepin, Minnesota, USA J Environ Qual 29:561-568
- Kerr J (1998) The economics of soil degradation: from national policy to farmers' fields Soil Erosion at Multiple Scales: Principals and Methods for Assessing Causes and Impacts International Board for Soil Research and Management (IBSRAM), CABI Publishers, Wallingford, UK:21-38
- Khalik WMAWM, Abdullah MP, Amerudin NA, Padli N (2013) Physicochemical analysis on water quality status of Bertam River in Cameron Highlands, Malaysia Journal of Materials and Environmental Science 4 488-495
- Khatun S (2017) Detection of Soil Erosion Potential Zones and Estimation of Soil Loss in Kushkarani River Basin of Eastern India Journal of Geography, Environment and Earth Science International 12:1-17
- Khosrokhani M, Pradhan B (2014) Spatio-temporal assessment of soil erosion at Kuala Lumpur metropolitan city using remote sensing data and GIS Geomatics, Natural Hazards and Risk 5:252-270
- Kouli M, Soupios P, Vallianatos F (2009) Soil erosion prediction using the revised universal soil loss equation (RUSLE) in a GIS framework, Chania, Northwestern Crete, Greece Environmental Geology 57:483-497
- Lee S, Pradhan B (2007) Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models Landslides 4:33-41
- Lucà F, Conforti M, Robustelli G (2011) Comparison of GIS-based gully susceptibility mapping using bivariate and multivariate statistics: Northern Calabria, South Italy Geomorphology 134:297-308
- Magliulo P (2012) Assessing the susceptibility to water-induced soil erosion using a geomorphological, bivariate statistics-based approach Environmental Earth Sciences 67:1801-1820
- Mararakanye N, Le Roux J (2012) Gully location mapping at a national scale for South Africa South African Geographical Journal 94:208-218
- Marjanović M, Kovačević M, Bajat B, Voženilek V (2011) Landslide susceptibility assessment using SVM machine learning algorithm Engineering Geology 123:225-234
- Nampak H, Pradhan B, Mojaddadi Rizeei H, Park H-J (2018) Assessment of Land Cover and Land Use Change Impact on Soil Loss in a Tropical Catchment by Using Multi-Temporal SPOT-5 Satellite Images and RUSLE model Land Degrad Dev 29:3440-3455
- Pal S (2016) Identification of soil erosion vulnerable areas in Chandrabhaga river basin: a multi-criteria decision approach Modeling Earth Systems and Environment 2:5
- Parveen R, Kumar U (2012) Integrated approach of universal soil loss equation (USLE) and geographical information system (GIS) for soil loss risk assessment in Upper South Koel Basin, Jharkhand Journal of Geographic Information System 4:588
- Pham TG, Degener J, Kappas M (2018) Integrated universal soil loss equation (USLE) and Geographical Information System (GIS) for soil erosion estimation in A Sap basin: Central Vietnam Int Soil Water Conserv Res 6:99-110
- Pimentel D, Burgess M (2013) Soil erosion threatens food production Agriculture 3:443-463
- Platt J (1998) Sequential minimal optimization: A fast algorithm for training support vector machines
- Pourghasemi HR, Jirandeh AG, Pradhan B, Xu C, Gokceoglu C (2013) Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran J Earth Syst Sci 122:349-369
- Pourghasemi HR, Yousefi S, Kornejady A, Cerdà A (2017) Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling Science of Total Environment 609:764-775
- Pournader M, Ahmadi H, Feiznia S, Karimi H, Peirovan HR (2018) Spatial prediction of soil erosion susceptibility: an evaluation of the maximum entropy model Earth Science Informatics 11:389-401 doi: <https://doi.org/10.1007/s12145-018-0338-6>
- Pradeep GS, Krishnan MN, Vijith H (2015) Identification of critical soil erosion prone areas and annual average soil loss in an upland agricultural watershed of Western Ghats, using analytical hierarchy process (AHP) and RUSLE techniques Arabian Journal of Geosciences 8:3697-3711
- Pradhan B (2010) Application of an advanced fuzzy logic model for landslide susceptibility analysis International Journal of Computational Intelligence Systems 3:370-381
- Pradhan B (2013) A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS Computers & Geosciences, 51:350-365
- Prosdoci M, Cerdà A, Tarolli P (2016) Soil water erosion on Mediterranean vineyards: A review Catena 141:1-21
- Rahmati O, Pourghasemi HR, Zeinivand H (2016) Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golestan Province, Iran Geocarto International 31:42-70
- Rahmati O, Tahmasebipour N, Haghizadeh A, Pourghasemi HR, Feizizadeh B (2017) Evaluating the influence of geo-environmental factors on gully erosion in a semi-arid region of Iran: An integrated framework. Sci Total Environ 579:913-927
- Remondo J, González A, De Terán JRD, Cendrero A, Fabbri A, Chung C-JF (2003) Validation of landslide susceptibility maps; examples and applications from a case study in Northern Spain Natural Hazards 30:437-449
- Ridd MK (1995) Exploring a VIS (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities International journal of remote sensing 16:2165-2185
- Rizeei HM, Maryam AS, Pradhan B, Ahmad N (2016) Soil erosion prediction based on land cover dynamics at the Semenyih watershed in Malaysia using LTM and USLE models Geocarto international 31:1158-1177
- Shit PK, Paira R, Bhunia G, Maiti R (2015) Modeling of potential gully erosion hazard using geo-spatial technology at Garbheta block, West Bengal in India Modeling Earth Systems and Environment 1:1-16
- Stöcker C, Eltner A, Karrasch P (2015) Measuring gullies by synergetic application of UAV and close range photogrammetry - A case study from Andalusia, Spain Catena 132:1-11
- Sujatha ER, Sridhar V (2018) Spatial Prediction of Erosion Risk of a Small Mountainous Watershed Using RUSLE: A Case-Study of the Palar Sub-Watershed in Kodaikanal South India Water 10:1608
- Sujay RN, Paresh CD (2014) Support vector machine applications in the field of hydrology: a review Applied Soft Computing 19:372-386
- Tehrany MS, Pradhan B, Jebur MN (2014) Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS Journal of hydrology 512:332-343

- Tehrany MS, Pradhan B, Mansor S, Ahmad N (2015) Flood susceptibility assessment using GIS-based support vector machine model with different kernel types *Catena* 125:91-101
- Thlakma SR, Iguisi E, Odunze AC, Jeb DN (2018) Estimation of Soil Erosion Risk in Mubi South Watershed, Adamawa State, Nigeria. *J Remote Sensing & GIS* 7:226
- Tinker PB (1997) The environmental implications of intensified land use in developing countries *Philos Trans R Soc Lond B* 352:1023-1033
- Vahabi J, Nikkani D (2008) Assessing dominant factors affecting soil erosion using a portable rainfall simulator *International Journal of Sediment Research* 23:376-386
- Vapnik V (1998) *Statistical Learning Theory*. Wiley, New York, NY
- Vijith H, Suma M, Rekha V, Shiju C, Rejith P (2012) An assessment of soil erosion probability and erosion rate in a tropical mountainous watershed using remote sensing and GIS *Arabian Journal of Geosciences* 5:797-805
- Walling DE, Webb BW (1983) Patterns of sediment yield. Gregory KJ (Ed) *Background to Paleohydrology*. Wiley, Chichester
- Wang B, Zheng F, Römkens MJ, Darboux F (2013) Soil erodibility for water erosion: A perspective and Chinese experiences *Geomorphology* 187:1-10
- Wang R et al. (2016) Gully Erosion Mapping and Monitoring at Multiple Scales Based on Multi-Source Remote Sensing Data of the Sancha River Catchment, Northeast China *ISPRS Int J Geo-Inf* 5:200
- Wawer R, Nowocien E, Podolski B (2005) Eal and calculated Kusle erodibility factor for selected Polish soils *Polish Journal of Environmental Studies* 14:655-658
- Williams J, Bouzarovski S, Swyngedouw E (2014) Politicising the Nexus: Nexus Technologies, Urban Circulation, and the Coproduction of Water-energy. *Nexus Network Think Piece Series*, Paper 001,
- Wischmeier WH, Smith DD (1958) Rainfall energy and its relationship to soil loss *Eos, Transactions American Geophysical Union* 39:285-291
- World Wildlife Fund Malaysia (2002) Community and non-governmental organisation (NGO) partnership in highland catchment management in Malaysia.
- Xue X, Luo Y, Zhou X, Sherry R, Jia X (2011) Climate warming increases soil erosion, carbon and nitrogen loss with biofuel feedstock harvest in tallgrass prairie *Gcb Bioenergy* 3:198-207
- Taofeeq Sholagberu Abdulkadir** is a Lecturer in the Department of Water Resources and Environmental Engineering (Formerly in Civil Engineering Department), University of Ilorin, Nigeria. He received his Bachelor and Master of Civil Engineering degrees from University of Ilorin, Nigeria and his PhD in Civil and Environmental Engineering from Universiti Teknologi Petronas, Malaysia. His research interests include water resources modeling, watershed management and conservation, artificial neural network, machine learning, GIS and remote sensing. He has published few research articles in reputable national and international journals and conference proceedings.
- Raza Ul Mustafa Muhammad** is a Senior Lecturer in the Department of Civil and Environmental Engineering, Universiti Teknologi Petronas, Malaysia. He completed his Bachelor of Agricultural Engineering from the University of Agriculture Faisalabad, Pakistan; master's degree in Water Resources Engineering, University of Engineering and Technology, Pakistan; and PhD degree from Universiti Teknologi Petronas, Malaysia. His research interest includes artificial neural network, machine learning applications to water resources, and sediment modeling, among others.
- Olayinka Gafar Okeola** is a Civil Engineer in the Department of Civil Engineering, University of Ilorin. He received his PhD degree from University of Ilorin in 2010. He has been involved in teaching, mentoring, and research. His key research interest includes decision support systems, watershed hydrology, policy, urban water supply, hydro-economics, and GIS. He has publications in reputable journals and conference proceedings. He received the Pillars of Nation Building Award in 2012 as a Distinguished Academic ICON from the Strategic Institute for Natural Resources and Human Development of Nigeria (SINRHD).
- Wan Yusof Khamaruzaman** is an Associate Professor, Department of Civil and Environmental Engineering, Universiti Teknologi Petronas, (UTP) Malaysia. He obtained his Bachelor of Science in Civil Engineering from Loughborough University of Technology, United Kingdom, Master of Science in Irrigation Engineering from University of Southampton, Southampton, and PhD in Water Resources and GIS from Coventry University, Coventry, United Kingdom. He has over 30 years of teaching and research experience during which he has graduated several research students. He is currently the cluster leader of Water and Environmental Engineering, UTP. He has won several grants and gold medals in academic exhibitions. His research interests include storm water management, flood modeling, hydrological modeling, coastal erosion, watershed management, computational fluid dynamics, and GIS modeling.
- Bashir Adelodun** is a Lecturer in the Department of Agricultural and Biosystems Engineering, University of Ilorin, Nigeria. He received both his Bachelor of Engineering and Master of Engineering degrees from the University of Ilorin. He is currently doing his PhD research program in land and water engineering in the Department of Agricultural Civil Engineering, Kyungpook National University, South Korea. He has over 8 years of experience in both teaching and research. His research includes water resources management, water quality and wastewater treatment, soil remediation and environmental pollution, and hydrologic and water resources modeling and simulation.
- Saheed Adeniyi Aremu** is an Associate Professor in the Department of Water Resources and Environmental Engineering, University of Ilorin, Nigeria. Presently, he is the Managing Director, Lower Niger River Basin, Ilorin, Nigeria. He bagged his Master of Civil Engineering from University of Ibadan and Doctor of Philosophy from University of Ilorin, Nigeria. He has been teaching and researching for over two decades and has published several research papers in reputable national and international Journals including conference proceedings books and chapters. His research interests include water resources management, environmental modeling, pollution, waste management, GIS, and remote sensing.

Index

- A**
Afforestation, 13, 18, 160, 236, 241
Aggradation and degradation theory, 6
Akarsa watershed, 187–203
Analytical hierarchy process (AHP), 93–106, 405, 416, 428, 462
Arc-GIS, 51, 118, 188, 189, 199, 203
- B**
Badlands, ix, 46, 48, 65, 69, 105, 109–112, 114–116, 118, 122–125, 137, 155, 208, 212, 214, 216, 222, 231
Bamboo based technology, 307–318
Bamboo plantation, 13, 308, 310–312, 315–317
Bed width, 11, 12, 254, 386, 387
Below ground biomass, 295
Bio-engineering, 13, 14
Bivariate analyses, 370, 371, 373, 376, 377, 379, 405
Buffer strips, 222, 224, 226, 228–230
Bulk density, 22, 25, 120, 222, 295–297, 308, 327, 360, 386, 387
- C**
Chambal valley, 3, 6, 12, 307
Check dams, 14, 15, 17, 158, 160, 238, 248, 292, 310, 313, 322, 337–343, 456
Chotanagpur plateau, 36, 47, 148, 152
Cluster analysis, 49, 385, 386, 388–390
Community resilience, 449–458
Concentrated flow, 2, 6, 65, 87, 93, 169, 170, 221–230, 235, 295–297, 300, 302, 304, 382
Cross-section area, 256–258
- D**
Darcy-Weisbach friction factor, 223, 224, 227, 229, 230
Dendrogram, 387
Denudational processes, 70, 79, 134, 198
Detritus layer, 21, 22, 24, 26, 27, 29
Disintegration index, 22
Drainage density, 37, 41, 94, 96, 98, 99, 101, 104, 106, 109, 153, 163, 170, 188, 189, 191–193, 197, 199, 236, 240, 255, 256, 259, 260, 340, 398, 408, 411, 416, 418, 422, 432, 433, 436, 443, 463, 464, 467, 470, 472, 473
Drainage network, 69, 70, 189, 240, 253, 259, 341, 346, 378, 434
- E**
Ephemeral channel, 163–183
Erosion pin, 110, 111, 115, 118–120, 123, 124
- F**
Factor analysis, 101, 386, 388–390
Fish resources, 350
Fourfold plot, 423
Froude Number (Fr), 223, 227–230, 249
- G**
Gangani, 22–24, 27, 28, 69, 208, 211, 212, 214, 218, 346, 348–350
Garbha badland, 70
Geo-environmental factors, 134, 137, 144, 187, 188, 203, 415, 435, 443
Geomorphic threshold, ix, 45–66
Gorganrood watershed, 427–444
Groundwater fluctuation, 268–271, 273
Gully beds erosion, 221–230
Gully classification, 10, 11, 238
Gully collapse, ix, 21–32
Gully head, 6, 7, 10, 12–14, 21, 22, 27, 28, 46, 47, 49, 50, 53, 56, 57, 59, 63, 66, 115, 136, 208, 211, 214, 222, 238, 258, 280, 284, 290, 307–309, 317, 327, 358, 360, 362, 369–379, 381–383, 385–390, 394
Gully headcut, 14, 51, 63, 66, 142, 208, 372, 386, 394, 399
Gully morphology, 85, 87, 134, 148, 150, 207–218
Gully reclamability, 10, 11
Gully stabilize, 14, 231, 317
Gully susceptibility, 99, 101, 142–143, 428, 432, 440
- H**
Hydraulic share stress, 223
HydroCAD, 331–334
Hydrogeomorphic, 46, 53, 341
- I**
Index of diversity, 347
Initiation of gully, 53, 56, 251–261, 279, 427
In-situ scouring experiment, 309
- K**
Kangsabati basin, 164–166, 172, 346, 348–350, 354
Kappa index, 39, 416, 420, 423, 469, 472–474
Kernel functions, 419, 467–469, 471–474
Khoai badland, 147–160
- L**
Landslides, 40, 57, 133, 134, 238–240, 249, 321, 323, 324, 378, 421, 428, 434, 436, 457, 461, 463
Lateritic region, 148
Lateritic tract, 22, 208
Lateritic uplands, ix, 208

Livelihood improvement, 318
 Logistic regression (LR), 397, 405, 410, 416, 428, 462

M

Machine learning, ix, 39–41, 398, 406, 415–423, 428, 429, 431, 434, 443, 462, 467, 469, 474
 Maharlou watershed, 406–411
 Mahi ravine, 8, 12
 Manning roughness coefficient, 223, 224, 230
 Maximum entropy model (ME), 398, 405, 427–444, 462
 Micro-watershed, 309, 337–343
 Multicollinearity, 432
 Multi-influencing factor (MIF), 187–203
 Multivariate adaptive regression splines (MARS) model, ix, 405–411, 428

N

Normalized difference vegetation index (NDVI), 189, 192, 408, 411, 463, 465–467, 469, 470, 472–474

O

Oceanic upwelling theory, 7
 Orthophoto, 360, 369, 374

P

Palaeogenesis, 47
 Particle size distribution (PSD), 22, 24, 26, 27, 31, 32, 370, 372
 Pedogeomorphology, 46
 Photogrammetry, 93, 207–218
 Pielou's evenness index, 347, 348
 Potential land degradation zone (PLDZ), 187–203
 Profilometer, 110, 111, 114–118, 123

Q

QGIS, 254

R

Rainfall simulator, 110, 111, 119, 121–122, 325, 332
 Random forest (RF) model, ix, 35–43
 Rangamati badland, 222, 231
 Ravine area restoration, 11, 12
 R environment, 416, 420
 Resource conservation, 307–318
 Reynolds number (Re), 223, 226–230
 River water quality, 345–354
 Run-off, 2, 39, 46, 69, 98, 118, 140, 147, 163, 199, 222, 235, 252, 280, 296, 307, 322, 337, 345, 358, 369, 382, 393, 408, 416, 433, 462

S

Sediment yield (SY), ix, 54, 104, 109–111, 118–120, 122, 125, 147, 151, 157, 163–183, 222, 236, 240–242, 248, 279–292, 302, 321, 415
 Shannon diversity index, 347, 348

Siddheswari river basin, 279–292
 Silabati river basin, 346, 349, 350, 353, 354
 Simpson's index, 347
 Slope aspect, 37, 41, 141, 198, 398, 406, 416, 418, 433, 443
 Slope instability, 322
 Soil and Water Assessment Tool (SWAT) model, 238, 240, 279–292, 462
 Soil anti-scourability, 295, 296, 300, 303–305
 Soil detachment rate, 296, 300–303
 Soil erodibility, 55, 112, 166, 190, 191, 194, 197, 199, 281, 283, 295, 296, 300, 301, 354, 370, 450, 463, 465, 470
 Soil erosion, 2, 21, 35, 45, 69, 93, 109, 133, 148, 163, 188, 211, 222, 236, 251, 279, 295, 307, 321, 337, 345, 357, 378, 381, 393, 405, 415, 427, 450, 461
 Soil moisture index (SMI), 463, 466, 467, 469, 470, 472, 473
 Soil organic matter (SOM), 6, 7, 22, 24–27, 29, 31, 32
 Spatial analysis, 370–372, 461–475
 Stream length-gradient index, 148, 150
 Stream power index (SPI), 134, 137, 141, 144, 381, 432, 436
 Support vector machine (SVM), 406, 415–423, 428, 462, 463, 466–469, 471–474
 Surface runoff, ix, 14, 46, 53, 55, 105, 147, 151, 163, 168, 222, 228, 280, 281, 283, 284, 287–289, 296, 312, 337, 354, 358, 360, 365, 382, 388, 389, 416, 418, 423, 433, 434, 463
 Susceptibility map, 35, 36, 41, 133, 145, 405–410, 427–430, 435, 439–441, 443, 462, 463, 466, 469, 471, 472

T

Tension cracks, 8, 427
 Terai region, 251–261
 Topographical factors, 175, 416, 418
 Top width, 10, 12, 366, 386, 387
 Tropical plateau, 163–181, 202
 True skill statistics (TSS), 416, 419, 420, 423
 Tunneling, 7–9, 358

U

Univariate, 370, 372, 373, 375, 377, 379
 Upliftment theory, 6
 Upstream drainage system, 256

V

Vegetative cover, 109, 231, 312, 318, 322, 323, 455
 Vetiver root, 325–328, 331–334
 Vulnerability, 26, 57, 59, 66, 94, 198, 316, 451, 452, 455–458

W

Weighted linear combination (WLC), 96, 99, 101, 105
 Weight of evidence model (WOE), 133, 145, 405
 Wetness index (WI), 37, 134, 137, 140–142, 144, 381, 398, 416, 418, 422, 433, 436, 443
 Width-depth ratio, 74, 76, 77, 256, 258, 387

Y

Yamuna valley, 3, 6, 12, 17, 309, 311, 312, 315