

Research Article

Quantifying Topographic Ruggedness Using Principal Component Analysis

Maan Habib 

Faculty of Civil Engineering, Damascus University, Damascus, Syria

Correspondence should be addressed to Maan Habib; maan.habib@gmail.com

Received 22 June 2021; Revised 16 September 2021; Accepted 7 October 2021; Published 2 November 2021

Academic Editor: Weiping Wen

Copyright © 2021 Maan Habib. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The development of geospatial technologies has opened a new era in terms of data collection techniques and analysis procedures. Digital elevation models as 3D visualization of the Earth's surface have many mapping and spatial analysis applications. The primary terrain factors derived from the raster dataset are usually less critical than secondary ones, e.g., ruggedness index, which plays a vital role in engineering, hydrological information derivation, and geomorphological processes. Surface ruggedness is a significant predictor of topographic heterogeneity by calculating the absolute value of elevation differences within a specified neighborhood surrounding a central pixel. The current study investigates the impacts of various topographic metrics obtained from a digital elevation model on characterizing terrain ruggedness utilizing stepwise principal component analysis. This popular multivariate statistical technique is applied to conduct a comprehensive assessment and treat the information redundancy of terrain parameters. Simultaneously, the standard deviation of elevation is also proposed as an alternative approach to quantifying topographic ruggedness. Besides, quantitative and qualitative method is espoused to validate the algorithms and compare their capabilities to the previously introduced models in the literature. The findings have shown that principal component analysis provides superior performance against other models. Furthermore, they indicated that the standard deviation of elevation could be used instead of the available ones.

1. Introduction

Geomorphometry, which is defined as the science of quantitative Earth's terrain analysis [1], has mainly concentrated on studying terrestrial landscapes and derivation of a large spectrum of environmental variables [2, 3]. The essential data source for performing such spatial analysis is digital elevation models [4]. The digital elevation model is a three-dimensional raster representation of the land surface to identify geographical features [5, 6] by interpolating the landform using modern geospatial techniques [7–9]. It can be established through conventional ground topographic surveys or remote sensing methods [10–13]. In general, DEM is high-quality spatial data that is applied to extract terrain factors such as slopes, curvatures, aspects, and drainage network [14–18].

Generally, irregularity in elevation makes cultivation difficult and costly to traverse [19, 20], but useful for keeping

newborn reindeer calves against weather and windchill [21]. Relief ruggedness is a valuable attribute of wildlife habitat models and one of the prime components affecting vegetation diversity, snowmelt patterns, and water drainage that are key elements of nutrient obtainability [22–25] and the book [26]. Many previous studies have addressed quantifying surface ruggedness, in which some of these investigations were adopted on the statistical dispersion of elevations, slopes, aspects, and vectors orthogonal to planar facets on the landscape. Beasom et al. [27] described an approach for characterizing land surface ruggedness as a function of the total length of contour lines in a study area. More recently, Riley et al. [28] proposed a methodology to determine topographic heterogeneity by calculating the terrain ruggedness index value for each grid cell from the sum elevation changes within a specified neighborhood surrounding a central pixel. Sappington et al. [29] developed and tested a vector ruggedness measure algorithm to

estimate landscape ruggedness as the variation in the 3D orientation of grid cells in a moving window by decomposing slope and aspect into 3D vectors. In contrast, Popit and Verbovšek [30] identified the ruggedness index by computing the differences between the lowest and highest elevation for each cell location on the output raster.

In a multivariate statistical analysis, the dataset's dimension plays a critical role in determining the problem's complexity. One of the most common techniques for reducing the data dimensions to increase its interpretability is the principal component analysis that was introduced by Pearson [31]. Presently, PCA is considered a powerful statistical approach in image processing due to its ability to decrease the number of correlated image bands resulting in a smaller dataset capable of representing most of the original variabilities [32]. Over the last few decades, PCA was used in geosciences and remote sensing [33–37]. For instance, Petrișor et al. [38] highlight the potential of integrating geographic information systems and PCA as a decision-support tool for underdeveloped countries. Huang et al. [39] concluded that PCA could provide an accurate representation of terrain complexity distribution features compared to k-means clustering. However, this method has not yet been introduced to the process of modeling terrain ruggedness.

As stated before, no optimal measure would provide a powerful tool for defining habitat availability [40, 41]. Hence, selecting an appropriate method to evaluate this indicator is an essential task. In contrast, GIS software such as ArcGIS provides a variety of geoprocessing tools, but still has some gaps that must be addressed. The article's main objective is to investigate the variability of quantifying topographic ruggedness characteristics, relying on PCA, a standard deviation of elevation, and traditional indices. Additionally, the spatial tools using ArcGIS Model Builder and Python script will be developed for automating the processes. As the basis for terrain ruggedness analysis, the DEM is generated using synthetic data acquired by a scanned topographic map with varying coefficients to emulate Earth's physical land-surface features. The comprehensive assessment of several terrain metrics extracted from DEM is examined to describe topographic ruggedness using PCA. After that, the surface ruggedness obtained from the PCA and STD methods was compared against the ones from two different literature models by adopting some statistical approaches and visual depictions.

2. Materials and Methods

Topography relates to the shape and characteristics of the Earth's land surface and is a central controlling element in a wide range of natural processes [42, 43]. It must be quantitatively evaluated [44] by building accurate and reliable geospatial data [45]. The terrain analysis is concerned with extracting topographic attributes from DEM raster in GIS as a convenient environment for handling data [46]. In general, DEM quality and resolution have a critical influence on the derived factors [47–49]. The analysis scale and chosen approach for determining

different terrain variables will also play a role in the final results. Surface ruggedness is a significant geomorphological parameter applied in geoscience and environmental studies for indicating the relative change in elevation between adjacent grid cells [50]. Many approaches have been applied to verify which algorithm is the most appropriate for depicting topographic ruggedness. The principal component analysis was utilized in this study to evaluate the correlations between terrain factors and reduce the dimension of the original parameters to characterize surface ruggedness. At the same time, the standard deviation of elevation was also tested to be another ruggedness model. Figure 1 describes a methodology flowchart adopted in the current study for quantifying relief ruggedness.

2.1. Case Study Description. The chosen case study is the Wadi Musa area (Valley of Moses) as one of Ma'an governorate towns in Jordan. It is characterized by gorgeous natural scenery and a heterogeneous landscape with mountains in the eastern part (Figure 2). The geographic position of Wadi Musa has situated about 200 km to the south of Amman, between $30^{\circ} 15' 00''$ N and $30^{\circ} 30' 00''$ N and $35^{\circ} 10' 00''$ E and $35^{\circ} 30' 00''$ E, as shown in Figure 3. The elevation of the terrain surface in this region varies from 49 m to 1725 m above the mean sea level (AMSL). This area has several types of features to assess the topographic ruggedness map under different landform conditions. Indeed, characteristics of terrain relief, such as flat, hilly, and mountainous, have a significant influence on the reliability of DEM [51] and extracted terrain factors. The digital elevation model of the site was generated using a topographic map with a scale of 1/25000. The map was scanned and digitized to capture contour lines with an interval of 10 m. After that, the ANUDEM method, which is entailed in ArcGIS 10.2 under the name of Topo to Raster, is applied to derive the DEM model with 10 m accuracy from CLs. This algorithm was developed by Hutchinson [52] to fit the shape and drainage structure of the land surface. The dataset's quality control was performed by superimposing and inspecting the CLs extracted from the built DEM with the original ones. Figure 4 illustrates the flowchart of data acquisition.

2.2. Surface Ruggedness. Geomorphologic analysis as a measure of landforms geometry was broadly applied to understand various natural processes that impact humans and the environment [53, 54]. Generally, the terrain characteristics have tremendous leverage over the habitat and socioeconomic activities. The emergence of modern geospatial techniques has brought new ways to represent and quantify the bare-earth topography using grid-based DEMs and develop novel algorithms for extracting geomorphometric attributes [55, 56]. Neighborhood tools are common to carry out DEM analysis workflows by computing a specified statistic for a focal cell within a defined moving window that is designated in relation to the central point [57]. In this study, a 3×3 sampling window is effectively shifted across the DEM surface to identify terrain



FIGURE 1: Schematic representation of the research methodology.



FIGURE 2: Aerial view of Wadi Musa town.

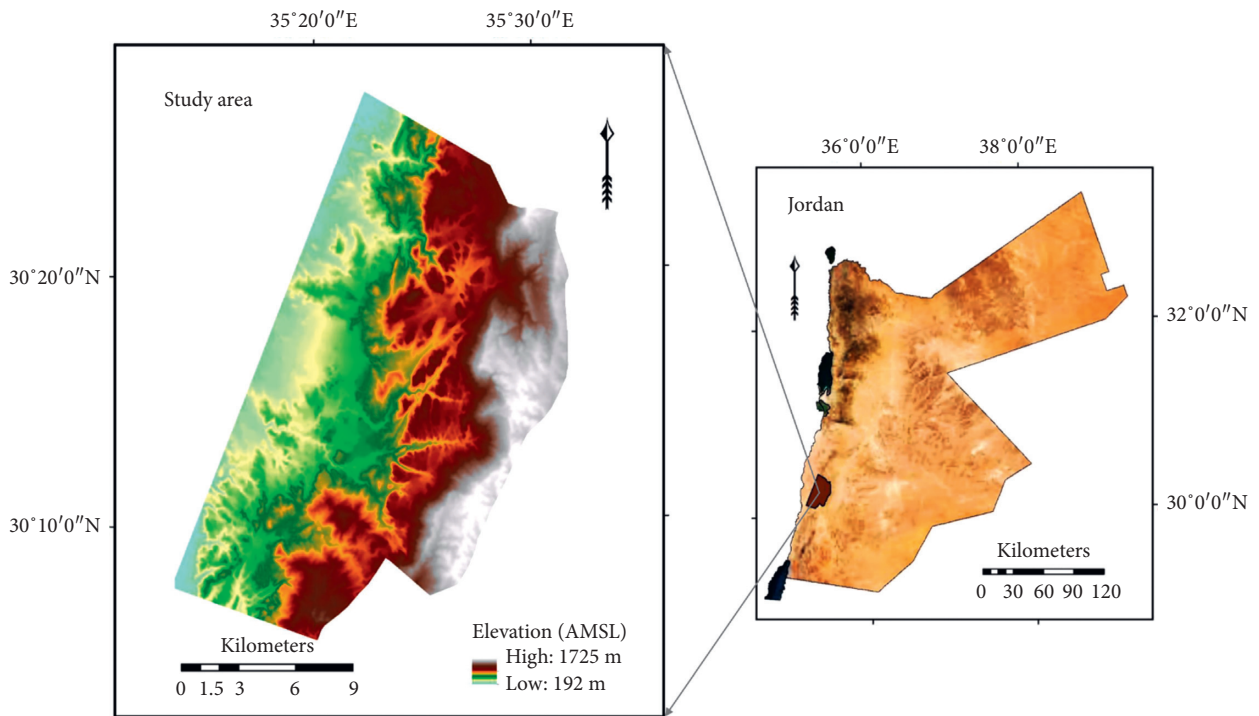


FIGURE 3: Location of Wadi Musa (geographical coordinates; WGS 1984).

information, as indicated in Figure 5. The main reason behind this selection is to limit the effects of the cell size on the results of the investigated models since finer windows require highly accurate DEMs [58].

Terrain ruggedness index is assigned for each raster cell using the algorithm described by Riley et al. [28] by

calculating the square root of the sum of the squared elevation differences between the central pixel and the adjacent ones within a specified neighborhood, as presented in equation (1), in which z_o is the elevation of the central cell, z is the elevation of each neighbor grid cell to the focal one, and N is the width of the roving window (number of pixels).

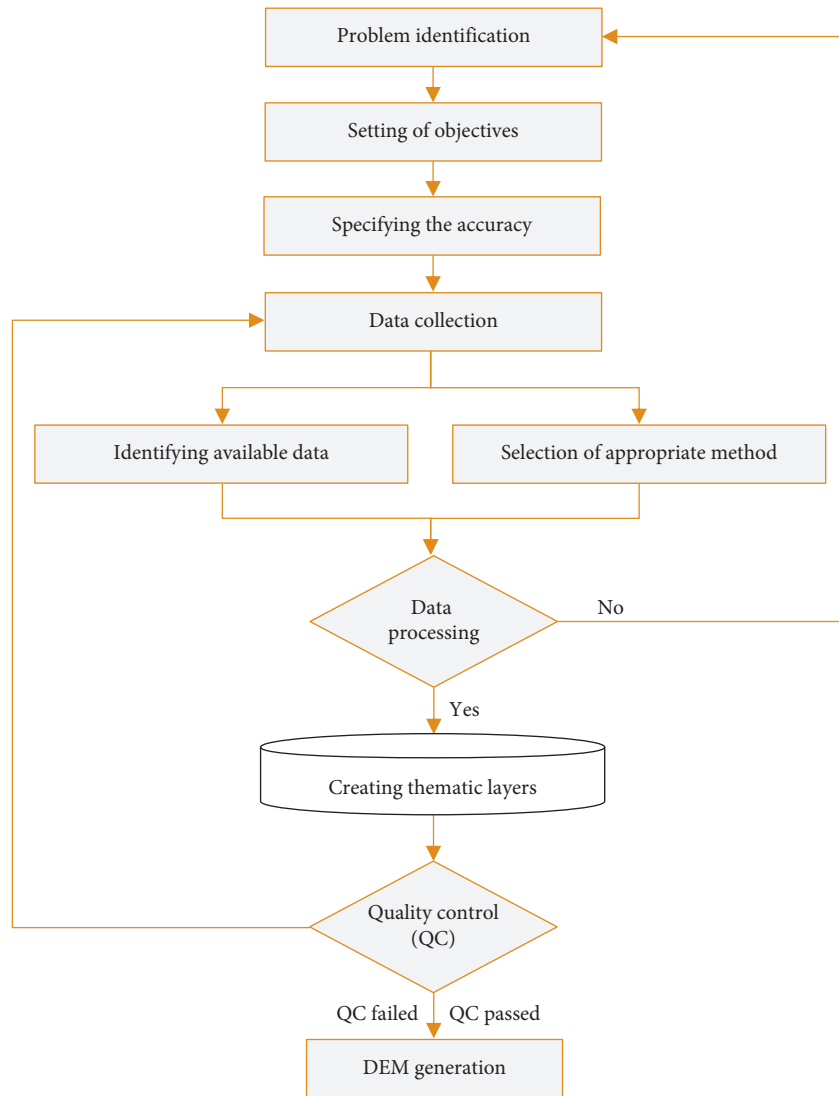


FIGURE 4: Flowchart of data acquisition.

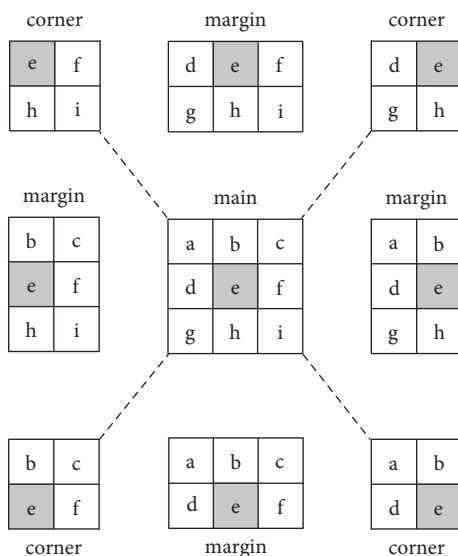


FIGURE 5: Neighborhood operations of a 3 × 3 input cell block.

Hence, according to Figure 5, equation (2) is written. The source code was developed in Python programming in the context of ArcPy to calculate the TRI algorithm, as indicated in Figure 6.

$$TRI = \left[\frac{1}{N^2 - 1} \sum_{k=1}^{N^2} (z_o - z_k)^2 \right]^{1/2}, \quad (1)$$

$$\sum_{k=1}^{N^2} (z_o - z_k)^2 = (e - a)^2 + (e - b)^2 + \dots + (e - i)^2. \quad (2)$$

Vector ruggedness measure provides another procedure to quantify terrain ruggedness as the variation in cells' 3D orientation within a neighborhood, as presented in Figure 7. It was initially developed by Hobson [59] and later espoused by Sappington et al. [29]. The vector ruggedness measure can be calculated from the slope and aspect using equation (3). Figure 8 illustrates a complete workflow ArcGIS Model Builder scheme to generate a VRM grid from the DEM dataset.


```

# Import system moduls
import arcpy
import math
import numpy as np
import time, sys, traceback
import numpy.ma as ma

#Specify DEM dataset
inRas = r'C:\Maan Data\Wadi Musa\Wadi_Musa.mdb\DEM_Surface'

#Determine an output directory
outRas = r'C:\Maan Data\Wadi Musa\Wadi_Musa.mdb'

try:
    # Convert Raster to numpy array
    arr = arcpy.RasterToNumPyArray(inRas,nodata_to_value=0)
    rows, cols = np.shape(arr)

    # Create a new array to hold TRI values
    arrTRI = np.zeros_like(arr)

    # Make a mask for the split array
    arrMask = [[1,1,1], [1,0,1], [1,1,1]]
    arrMask = np.ma.make_mask(arrMask)

    i = 0
    while i < rows-2:
        j = 0
        while j < cols-2:
            # Slice out a 3x3 bit of the array
            arrSplit = arr[i:i+3, j:j+3]

            # Get the value of central pixel
            centerCell = arrSplit[1, 1]

            # Apply the array mask to include only surrounding cells
            m_arrSplit = arrSplit[arrMask]
            elevDiff = pow((centerCell - m_arrSplit), 2)
            mean = np.mean(elevDiff)
            triResult = math.sqrt(mean)
            arrTRI[i + 1, j + 1] = triResult
            j += 1

        i += 1

    # Get geoprocessing result object (DEM properties)
    bottomResult = arcpy.GetRasterProperties_management(inRas, "BOTTOM")
    bottom = float(bottomResult.getOutput(0))
    leftResult = arcpy.GetRasterProperties_management(inRas, "LEFT")
    left = float(leftResult.getOutput(0))
    cellSizeXResult = arcpy.GetRasterProperties_management(inRas, "CELLSIZEX")
    cellSizeX = float(cellSizeXResult.getOutput(0))
    cellSizeYResult = arcpy.GetRasterProperties_management(inRas, "CELLSIZEY")
    cellSizeY = float(cellSizeYResult.getOutput(0))

    # Convert final array to TRI raster
    newRaster = arcpy.NumPyArrayToRaster(arrTRI)
    newRaster.save(outRas + '/temp0')
    arcpy.Shift_management(outRas + '/temp0', outRas + '/temp1', left, bottom)
    arcpy.Rescale_management(outRas + '/temp1', outRas + '/DEM_TRI', cellSizeX, cellSizeY)

    # Define TRI raster projection
    arcpy.DefineProjection_management(outRas + '/DEM_TRI', inRas)

    #Clean up temporary layers
    arcpy.Delete_management(outRas + '/temp0')
    arcpy.Delete_management(outRas + '/temp1')

    # Print message when tool runs successfully
    print(arcpy.GetMessages(0))

except arcpy.ExecuteError:
    print(arcpy.GetMessages(2))

except Exception as ex:
    print(ex.args[0])

```

FIGURE 6: TRI calculator code in Python script.

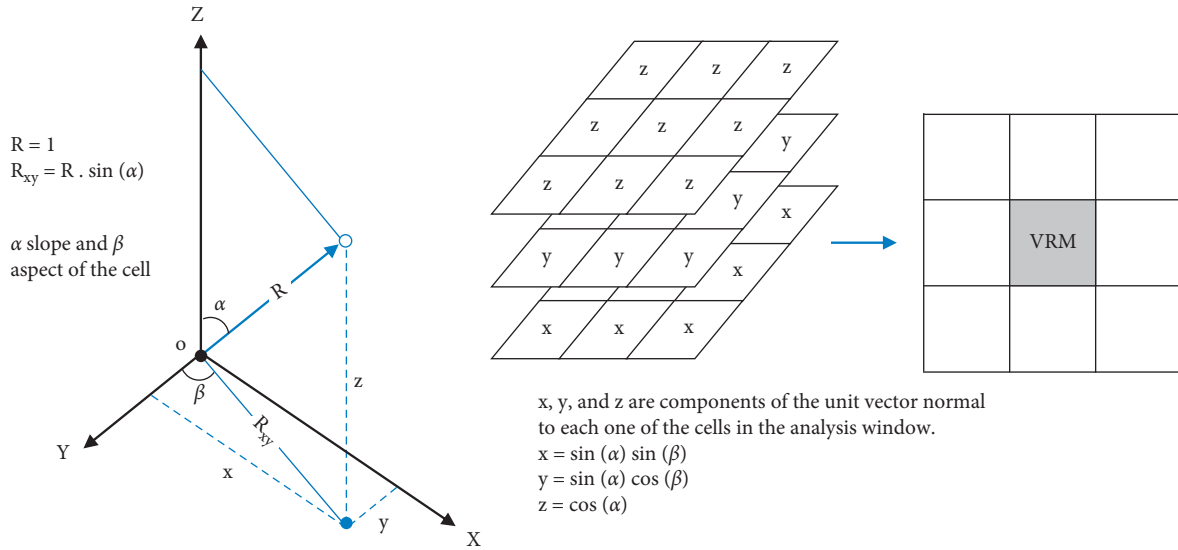


FIGURE 7: Computation of VRM utilizing vector analysis and a gridded DEM.

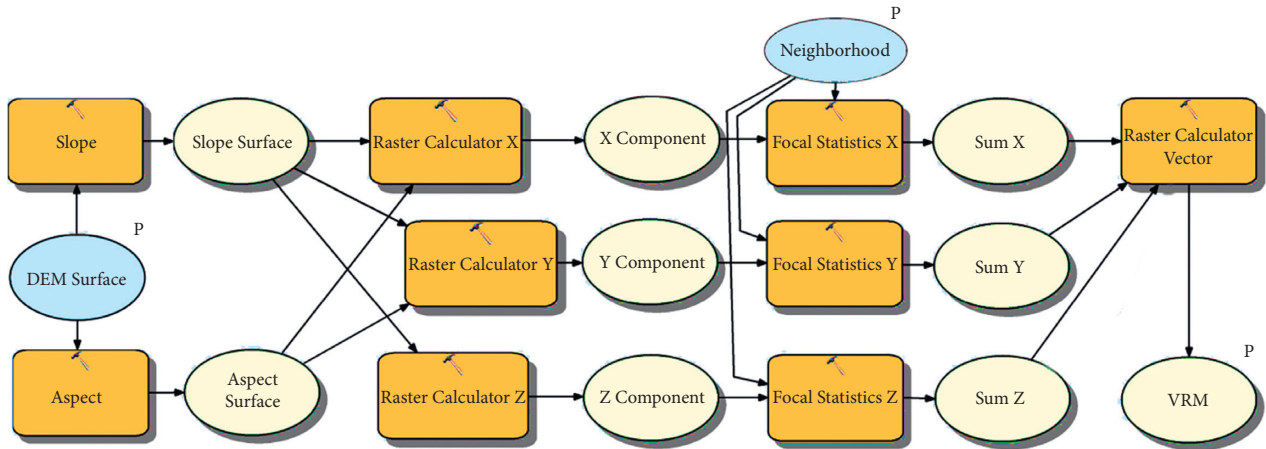


FIGURE 8: Model builder diagram to produce VRM map from DEM raster.

$$r = \left[\left(\sum_{k=1}^n x_k \right)^2 + \left(\sum_{k=1}^n y_k \right)^2 + \left(\sum_{k=1}^n z_k \right)^2 \right]^{1/2}, \quad (3)$$

$$VRM = 1 - \frac{|r|}{n}, \quad (4)$$

where n is the sample size (number of pixels in the moving window) and r is the resultant vector of central cells in a given neighborhood.

2.3. Ruggedness Conditioning Parameters. The terrain metrics shown in Table 1 are tested to evaluate and analyze ruggedness mapping to determine the most influential predictive factors. These parameters include six topographic features: slope, aspect, total curvature, surface roughness, topographic relief, and topographic position index. The relevant hydrological factors were the stream power index, sediment transport index, and topographic wetness index. Figure 9 demonstrates an

automated workflow in ArcGIS Model Builder that uses standard tools to extract morphometric variables from DEM.

2.4. Standard Deviation of Elevation. The standard deviation of elevation is determined by analyzing a DEM at a pixel level via sliding window scanning using equation (5). The standard deviation of elevation is examined in this study as an alternative ruggedness measure. The key point about choosing this method is to ease calculating its parameters using focal statistics in ArcGIS with a reasonable processing time for a large spatial dataset.

$$STD = \left[\frac{1}{N^2 - 1} \sum_{k=1}^{N^2} (z_k - \bar{z})^2 \right]^{1/2}. \quad (5)$$

where N is the width of the moving window, \bar{z} is the mean elevation in the window, and z is cell elevation. Hence, according to Figure 5, it can be written as follows:

TABLE 1: Definition and formula of terrain parameters.

Factor	Definition	Formula
Slope (°)	Angle between a tangent plane and a horizontal one at a certain point [56]	$\beta = \tan^{-1}(\sqrt{p^2 + q^2})$
Aspect (°)	Clockwise angle from north to the horizontal projection of an external normal vector at a specific point [56]	$\alpha = \tan^{-1}(p/q)$
Ground curvature (m ⁻¹)	Degree of bending variation at a given point [60, 61]	$C = q^2 r - 2pq\beta + p^2 t / (q^2 + p^2) \sqrt{1 + q^2 + p^2}$ <i>p and q are the rate of elevation change in x and y direction, respectively, and r and t are curvature radii of point</i>
Topographic relief (m)	Elevation difference between the highest (z _{max}) and lowest (z _{min}) point over an area	$TR = z_{max} - z_{min}$
Surface roughness	Describes surface relief and soil erosion degree [62]	$SR = 1/\cos \beta$
Topographic position index (m)	Elevation difference between central pixel (z _c) and the average elevation (\bar{z}) of neighboring cells [63]	$TPI = z_c - \bar{z}$
Topographic wetness index	Defines the cell tendency to accumulate water [64]	$TWI = \ln(A_s/\tan \beta)$ <i>A_s catchment area</i>
Stream power index	Measures potential flow erosion at the specified point [65]	$SPI = A_s \tan \beta$
Sediment transport index	Characterizes the process of erosion and deposition [65]	$STI = (A_s/22.13)^m (\sin \beta/0.0896)^n$ <i>m = 0.6, n = 1.3</i>

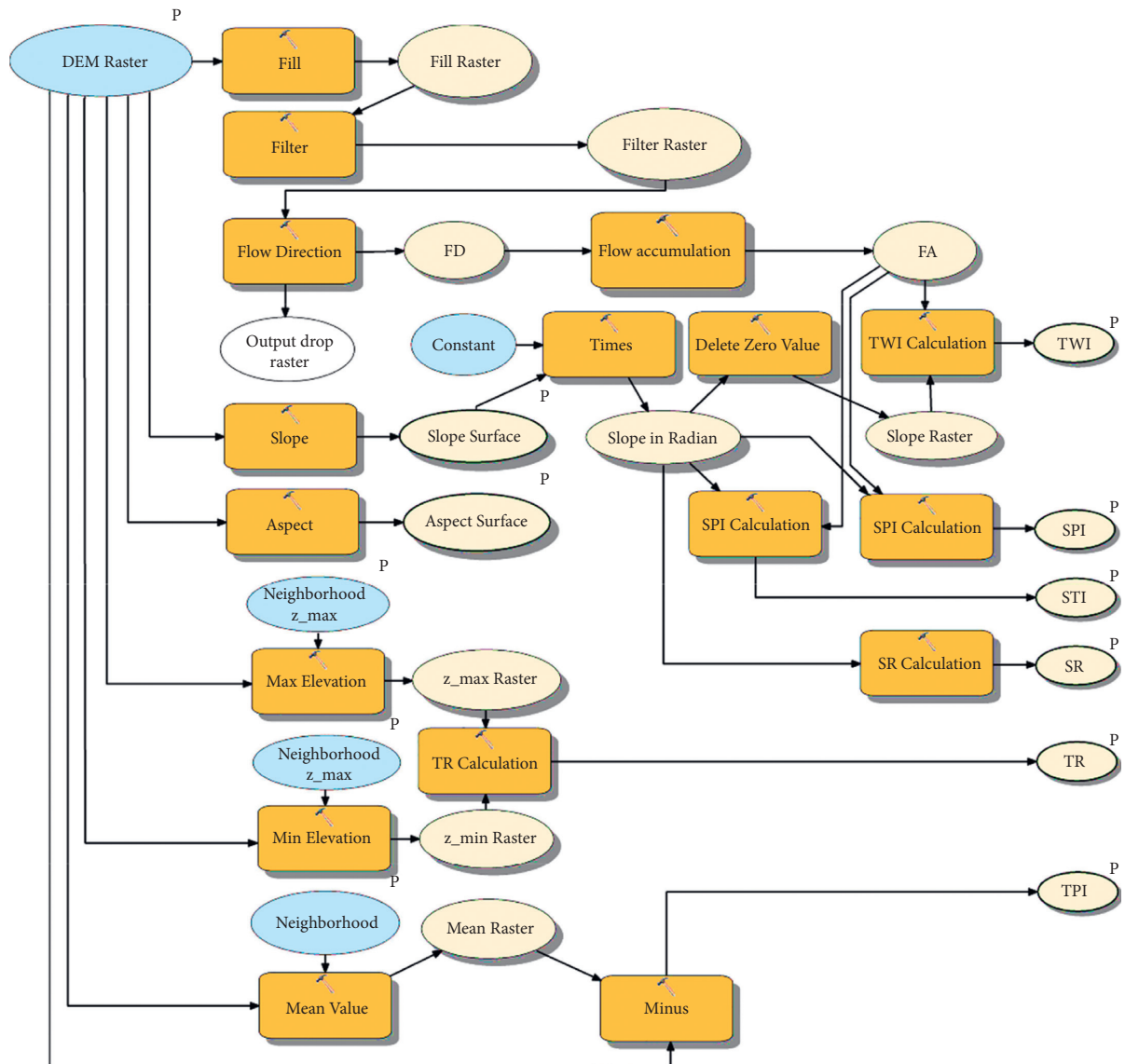


FIGURE 9: Geoprocessing tool in ArcGIS to calculate topographic metrics.

$$\bar{z} = \frac{a + b + \dots + i}{N^2}, \quad (6)$$

$$\sum_{k=1}^{N^2} (z_k - \bar{z})^2 = (a - \bar{z})^2 + (b - \bar{z})^2 + \dots + (i - \bar{z})^2.$$

2.5. Principal Component Analysis. The simplified algorithm for conducted PCA is seen in Figure 10. In general, four main steps are used to perform this method based on terrain factors. The correlation and covariance matrices are computed, and then the principal components are defined, and finally the dataset is shifted and rotated to the new axis. Usually, to improve the performance, avoid the numerical conflicting, and training stability of the ruggedness values, the following equation is used to normalize the obtained results [66].

$$\text{normalized values} = \frac{\text{Raster}_{\text{value}} - \text{Raster}_{\text{min}}}{\text{Raster}_{\text{max}} - \text{Raster}_{\text{min}}}. \quad (7)$$

2.6. Model's Performance Assessment. As mentioned previously, the obtained ruggedness grids from each method were classified into five different classes. These classes were defined as level (1), slight (2), moderate (3), high (4), and very high (5). In line with this study's aim, the identified classes from each model were compared by adopting the principle of confusion matrix that is always used for evaluating the performance of classification algorithms in machine learning. Each matrix's row represents the instances in a true class, whereas each column accounts for the predicted class cases. For a specific class in a multiple classes matrix such as the one shown in Figure 11, it can be divided into four main categories to indicate the number of times that the model: (1) true positive (TP) has rightly predicted the positive class, (2) true negative (TN) has precisely estimated the negative one, (3) false positive (FP) has incorrectly predicted the positive class, and (4) false negative (FN) has incorrectly estimated the negative class. Thus, using this matrix provides an easy and direct way for investigating the matching between two different approaches of the selected ruggedness modeling methods.

In general, matching any two models in this study is evaluated by computing the confusion matrix's overall accuracy. The accuracy is defined as the number of corresponding points in all cases divided by the total number of points selected for the investigation. Furthermore, the models' performance at the class level is assessed using two different approaches: precision and recall, as demonstrated in Figure 12.

3. Results and Discussion

Digital elevation models are fundamental to many applications in Earth sciences such as geomorphology, geology, ecology, and engineering [67–69]. Hence, the accuracy of DEM-derived elements must be deemed to reduce the inherent errors. The terrain ruggedness is extracted from DEM to measure local elevation variability within a defined window. This study's main objective is to carry out a comparative analysis to assess the quality of surface ruggedness generated by the PCA and STD algorithm. This analysis includes representing topographic ruggedness indices and verifying their performance. The standard deviation of elevation was determined using focal statistics over a 3×3 cell neighborhood of the DEM dataset. The chosen conditioning parameters (Figure 13) were normalized, and then the PCA approach was applied to describe surface ruggedness. On the other hand, the geoprocessing workflow in ArcGIS 10.2 software was automated by Python Script and Model Builder tool to calculate TRI and VRM indices, respectively.

The terrain variables in Table 1 were all used to develop an initial PCA model and choose the most suitable ones through its correlation matrix and eigenvalues. Indeed, the curvature, TWI, TPI, and aspect have indicated very low correlations than other metrics. Pearson's correlation coefficient of each pair for a selected dataset (sample size 64635 points) and descriptive statistics were computed to investigate the impact of various terrain parameters on the ruggedness indices, as shown in Tables 2 and 3. Hence, the influencing parameters, which highly correlate with the ruggedness index, were selected for building the PCA model. The Pearson correlation coefficient varies between -1 and 1 , with values close to the limits reflect high significance. In general, the blue color refers to have a positive correlation in which the ruggedness index is directly proportional to the terrain feature, while the orange color represents having a negative correlation means that the ruggedness index is inversely proportional to the terrain feature. Accordingly, the slope, TR, and SR have revealed the highest correlations; besides, SPI and STI gave a good correlation with different ruggedness methods. Hence, it is vital to use them to generate an accurate PCA map. Although TWI shows a significant relationship with all approaches, it minimized the model's capabilities when utilized in the PCA due to its negative interaction. Additional analysis was conducted to identify the bivariate correlation of factors with the PCA model. Table 3 indicates that the trends in PCA results follow those of the other approaches. In contrast, the correlation coefficients between the examined parameters and VRM are significantly lower than the other models. For instance, the correlation coefficient between the slope and the ruggedness index reduced from about 0.99 to 0.291 when using the VRM

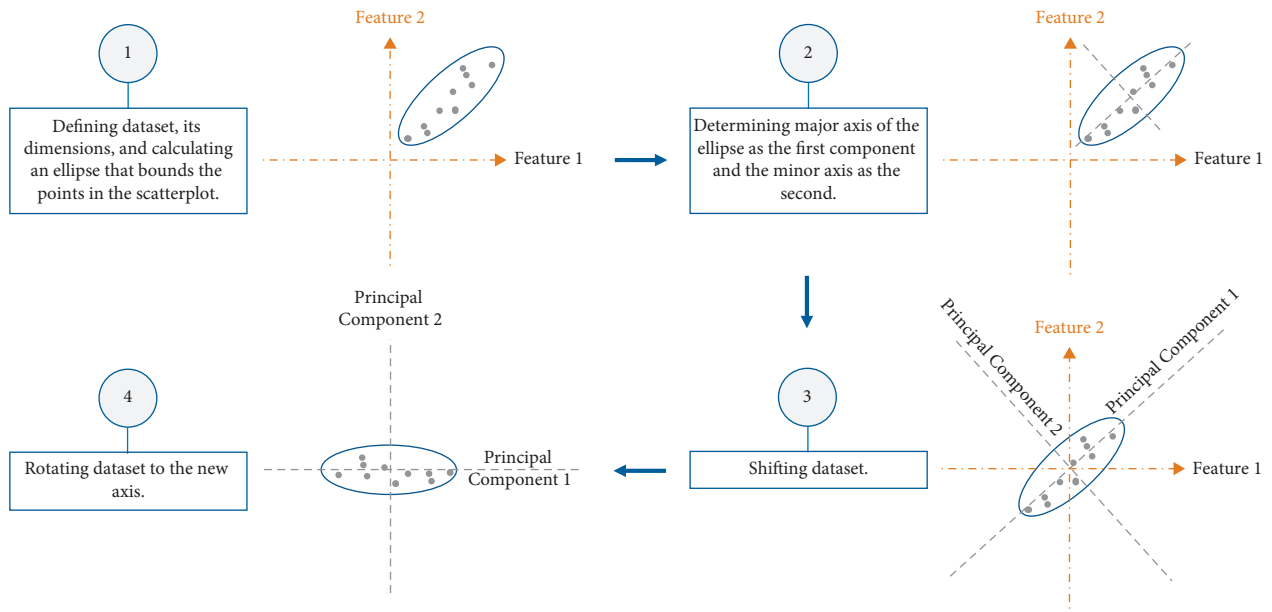


FIGURE 10: Flowchart of basic operations related to PCA.

True classes	$C_0 \dots C_{k-1}$	TN	FP	TN
	C_k	FN	TP	FN
	$C_{k+1} \dots C_n$	TN	FP	TN
		$C_0 \dots C_{k-1}$	C_k	$C_{k+1} \dots C_n$
		Predicted classes		

FIGURE 11: Simple illustration of a confusion matrix with multiple classes.

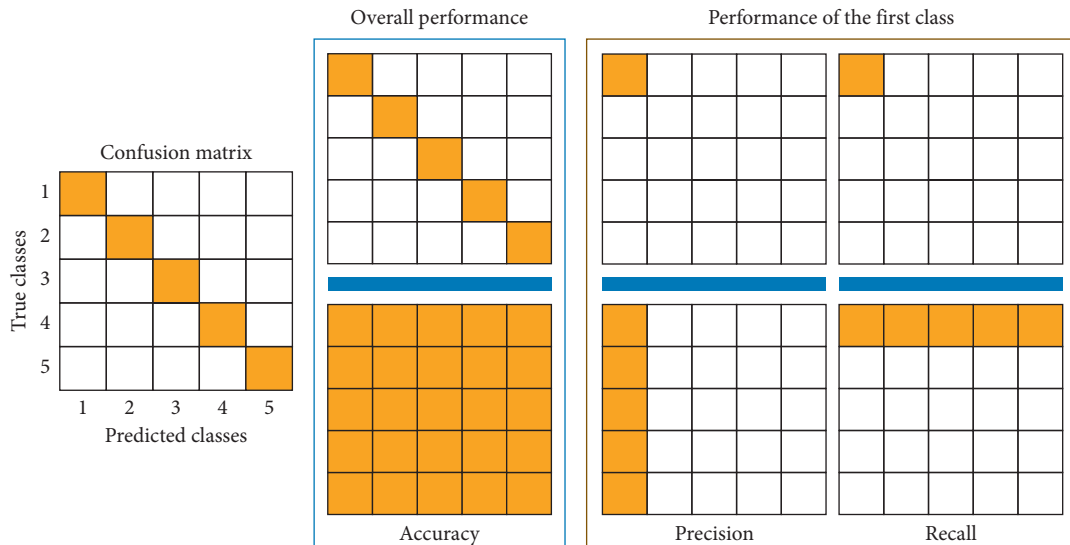


FIGURE 12: Examples of the procedure used for calculating the accuracy and precision.

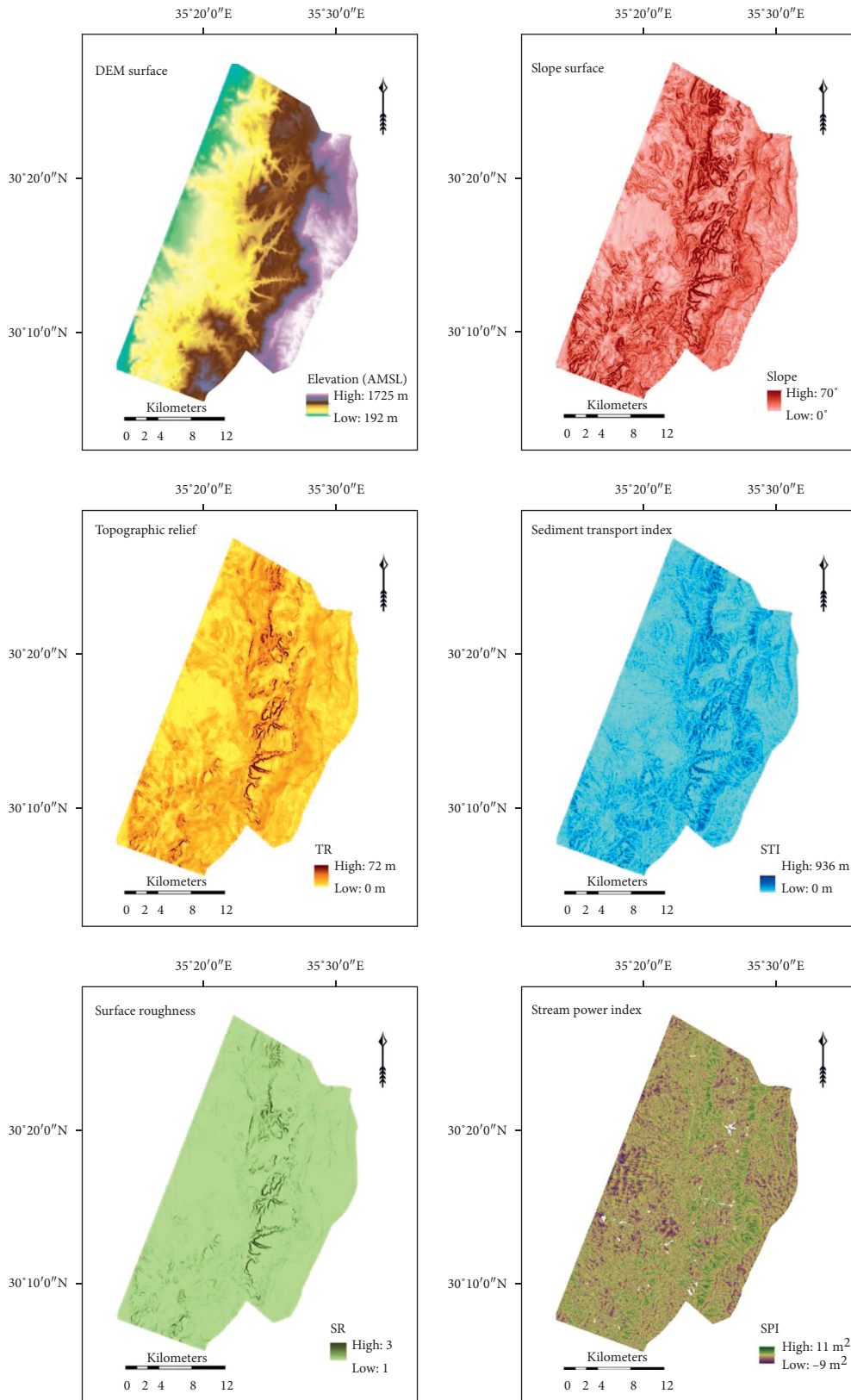


FIGURE 13: DEM dataset and topographic metrics applied for characterizing terrain ruggedness.

TABLE 2: Descriptive statistics of the selected dataset.

Variable	Mean	Standard deviation	Maximum	Minimum	Q1	Median	Q3
PCA	0.212	0.105	0.885	0.0137	0.132	0.192	0.267
TRI	0.083	0.071	0.802	0.0001	0.032	0.067	0.114
STD	0.083	0.071	0.799	0.0001	0.032	0.066	0.113
VRM	0.000	0.001	0.064	0.0000	0.000	0.000	0.000
SPI	0.498	0.082	0.985	0.0757	0.452	0.502	0.547
STI	0.004	0.009	0.754	0.0000	0.001	0.003	0.004
Slope	0.179	0.136	0.937	0.0001	0.072	0.150	0.251
TR	0.083	0.072	0.812	0.0001	0.031	0.065	0.113
SR	0.348	0.026	0.822	0.3336	0.335	0.339	0.350
Curvature	0.544	0.019	0.951	0.2593	0.538	0.544	0.550
TWI	0.268	0.090	0.917	0.0131	0.209	0.257	0.314
TPI	0.545	0.019	0.956	0.2697	0.538	0.545	0.551
Aspect	0.583	0.287	1.000	0.0000	0.346	0.671	0.813

TABLE 3: Bivariate correlation of terrain metrics and ruggedness algorithms.

Parameter	PCA	TRI	STD	VRM
SPI	0.537	0.415	0.415	0.112
STI	0.328	0.273	0.272	0.123
Slope	0.992	0.989	0.989	0.291
TR	0.977	0.989	0.989	0.288
SR	0.868	0.929	0.929	0.247
Curvature	0.002	0.037	0.038	-0.098
TWI	-0.473	-0.533	-0.533	-0.177
TPI	0.003	0.038	0.039	-0.098
Aspect	0.032	0.021	0.021	-0.012

model. A similar drop was observed in the VRM case, whether the correlation is positive or negative. On the other side, Figure 14 illustrates the correlation plots to measure the degree of agreement between the models and chosen topographic metrics. The x -axis of all diagrams represents the ruggedness index for each of the investigated methods (i.e., PCA, TRI, STD, and VRM) stated on the top, and the y -axis is the topographic parameters. In general, the TRI and STD algorithms display nearly the same, and TRI is highly correlated with the slope and the SR. In contrast, the PCA is related to several variables while there is no correlation between the VRM and comparative factors.

Figure 15 represents topographic ruggedness maps generated by TRI, VRM, PCA, and STD algorithms examined in the research area. These datasets are split into five homogeneous classes (level, slight, moderate, high, and very high) using the Jenks natural breaks optimization [70, 71]. Indeed, the smoothest areas represented by shades of a gray tone are entirely related to plain regions. Moreover, the more

rugged zones, clarified by cyan color, are fully seen around the mountains' topmost, as illustrated in Figure 16. The graphical representations of the ruggedness index indicate that the texture and colors/tones of the STD model closely match the corresponding TRI one. At the same time, these two surfaces show a slight difference from that of PCA, especially in places with very high ruggedness. Also, the VRM index reveals a considerable variation in the tone color patches from other algorithms. As a result, visual inspection has highlighted the PCA approach as practically the most superior model among the tested ones. Nevertheless, further analysis will be performed to evaluate tested models using statistical methods.

The PCA model was developed using selected topographic parameters in the case study, and covariance and correlation matrices are shown in Tables 4 and 5. Furthermore, the eigenvectors and eigenvalues are provided in Tables 6 and 7, respectively. Typically, the first component of the updated PCA cumulative contribution reached about

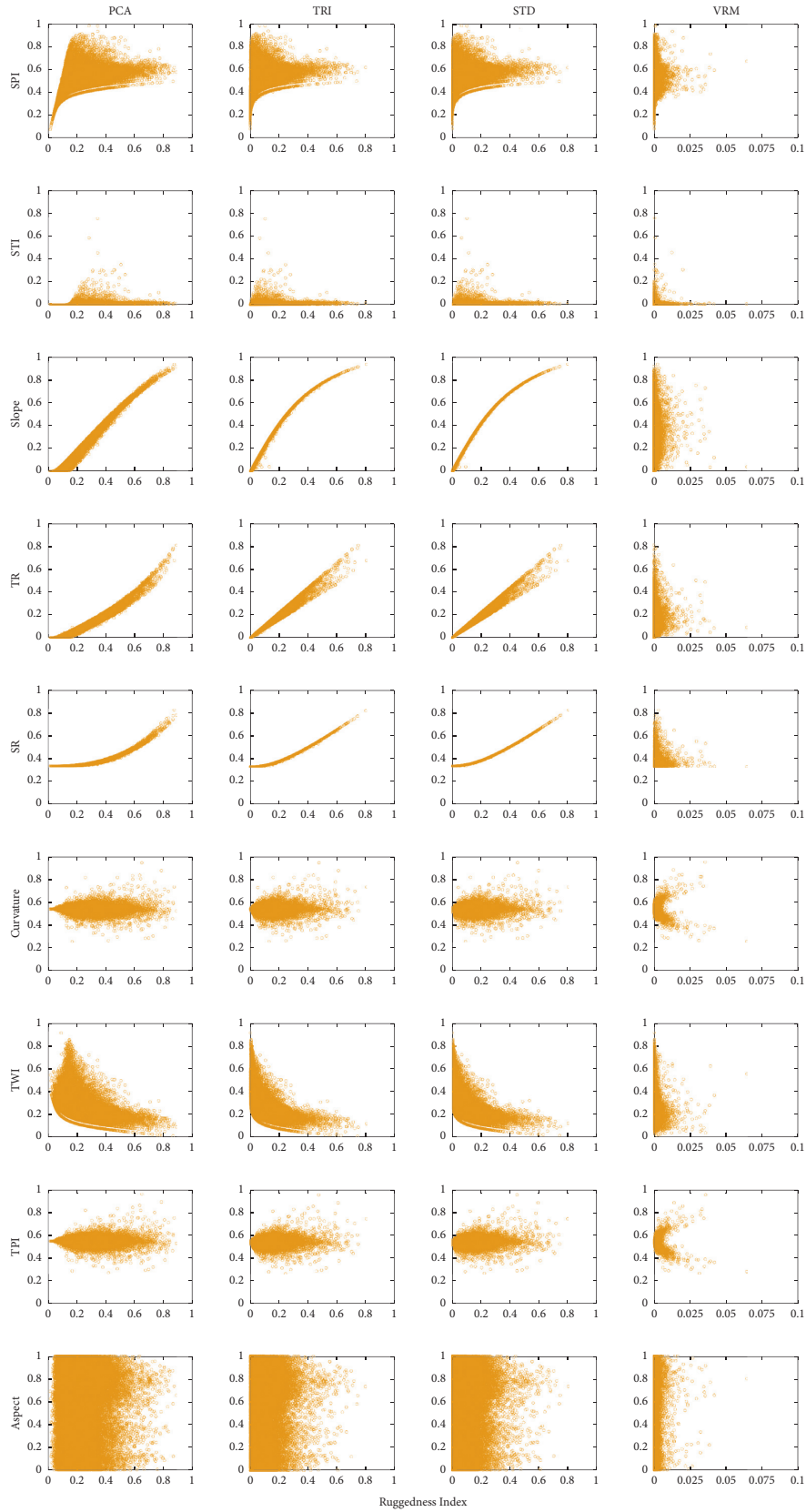


FIGURE 14: Plots of the correlation between the models and tested parameters.

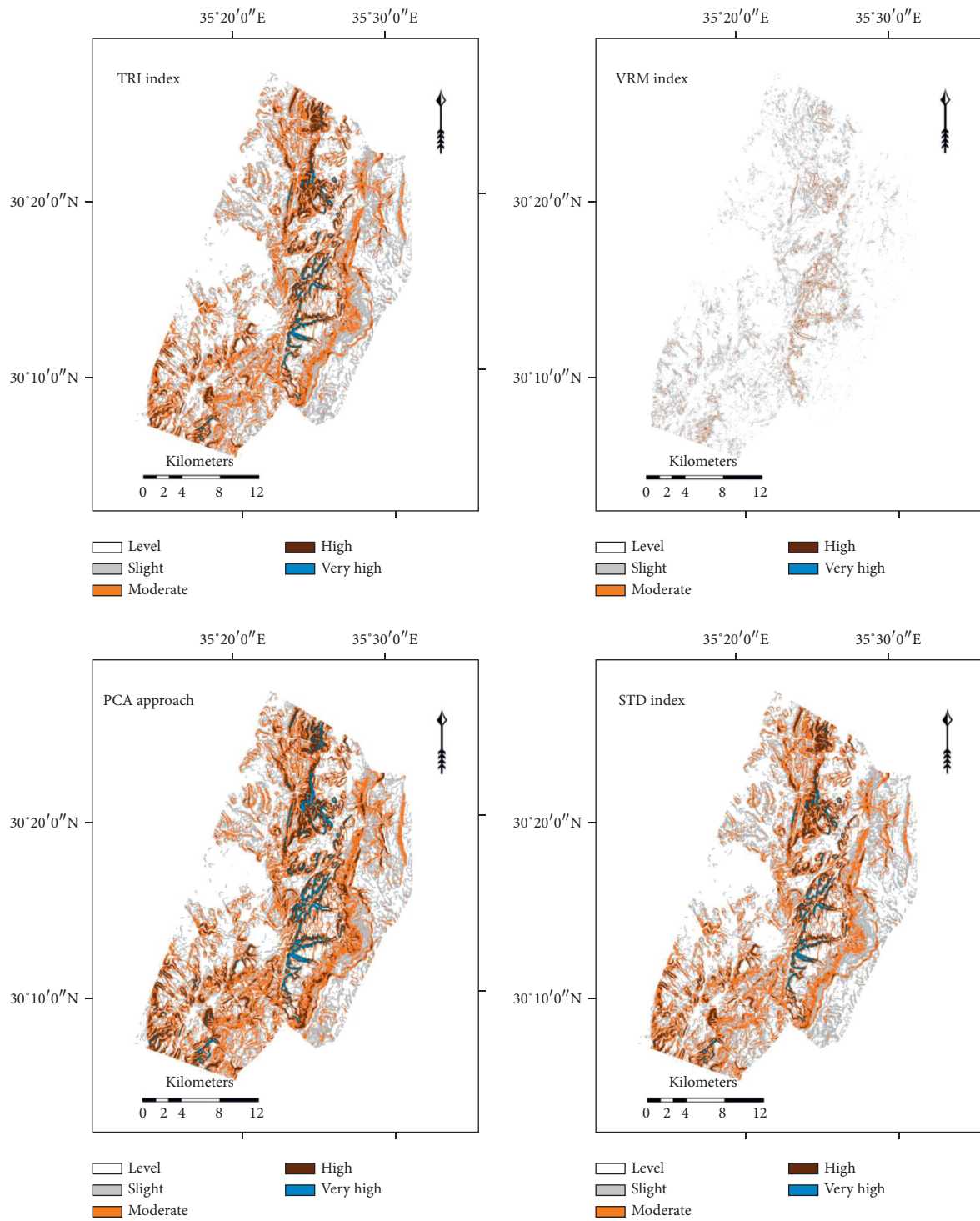


FIGURE 15: Topographic ruggedness maps of investigated models.

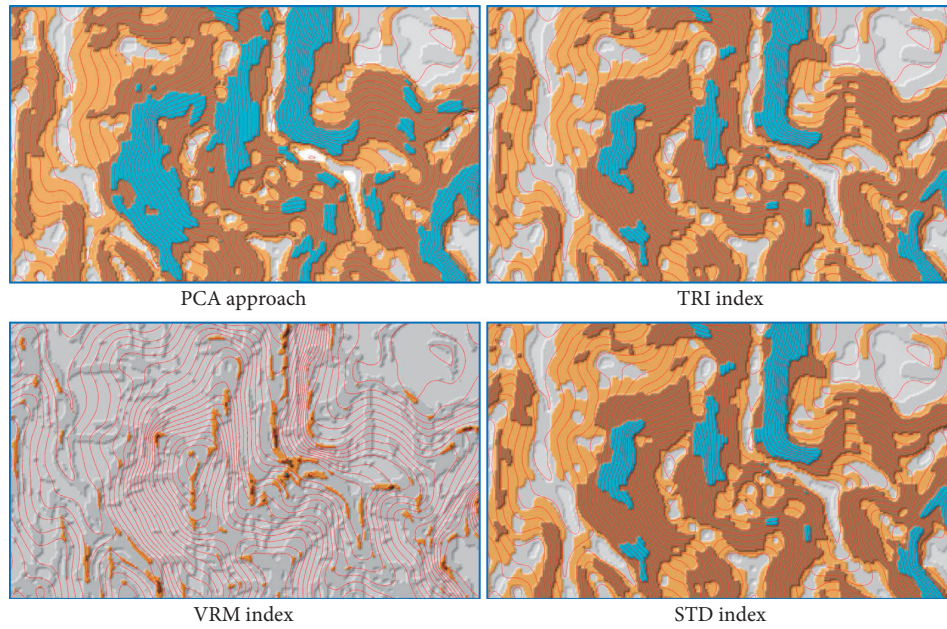


FIGURE 16: Terrain ruggedness and CLs of the sample area.

TABLE 4: The covariance matrix of terrain factors.

Feature	SPI	STI	Slope	TR	SR
SPI	0.00379	0.00022	0.00273	0.00131	0.00037
STI	0.00022	0.00005	0.00020	0.00010	0.00003
Slope	0.00273	0.00020	0.01087	0.00558	0.00180
TR	0.00131	0.00010	0.00558	0.00300	0.00100
SR	0.00037	0.00003	0.00180	0.00100	0.00040

TABLE 5: The correlation coefficient of terrain factors.

Feature	SPI	STI	Slope	TR	SR
SPI	1	0.51582	0.42526	0.38804	0.30153
STI	0.51582	1	0.26438	0.25196	0.22384
Slope	0.42526	0.26438	1	0.9777	0.86108
TR	0.38804	0.25196	0.9777	1	0.91594
SR	0.30153	0.22384	0.86108	0.91594	1

TABLE 6: Eigenvectors of all principal components.

Feature	SPI	STI	Slope	TR	SR
SPI	0.26467	0.9617	0.04072	-0.05455	0.02105
STI	0.0183	0.05405	0.03683	0.92304	-0.37865
Slope	0.84721	-0.2136	-0.43802	0.09144	0.19076
TR	0.43779	-0.14837	0.607	-0.26591	-0.5892
SR	0.1421	-0.06757	0.66081	0.25679	0.68748

TABLE 7: Percent and accumulative eigenvalues of the five topographic metrics.

Feature	Eigen value	Percent of eigen values (%)	Accumulative of eigen values (%)
SPI	0.01491	82.3431	82.3431
STI	0.00297	16.4057	98.7488
Slope	0.00016	0.8627	99.6115
TR	0.00004	0.2033	99.8148
SR	0.00003	0.1852	100

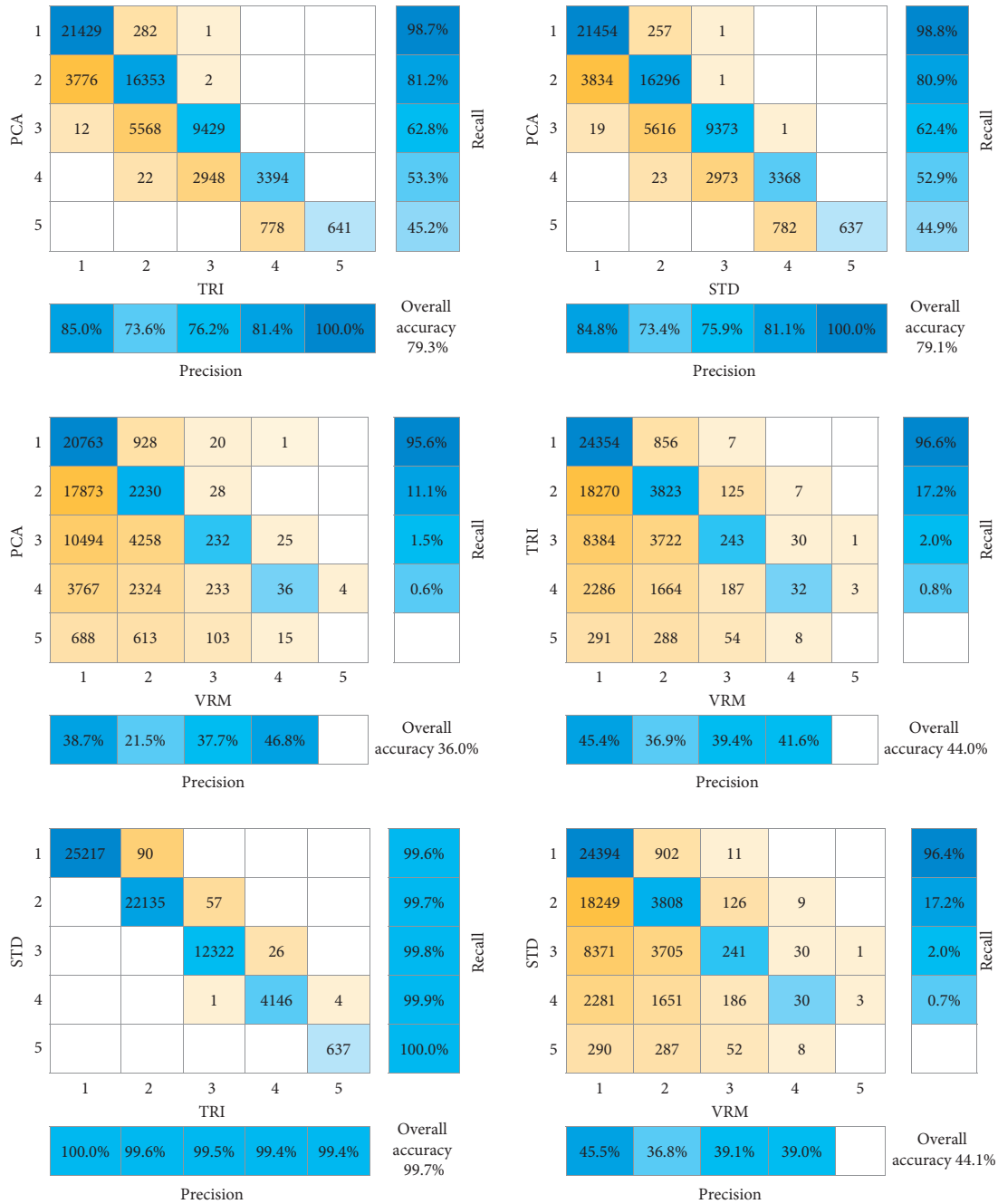


FIGURE 17: Confusion matrix for each pair of the investigated models.

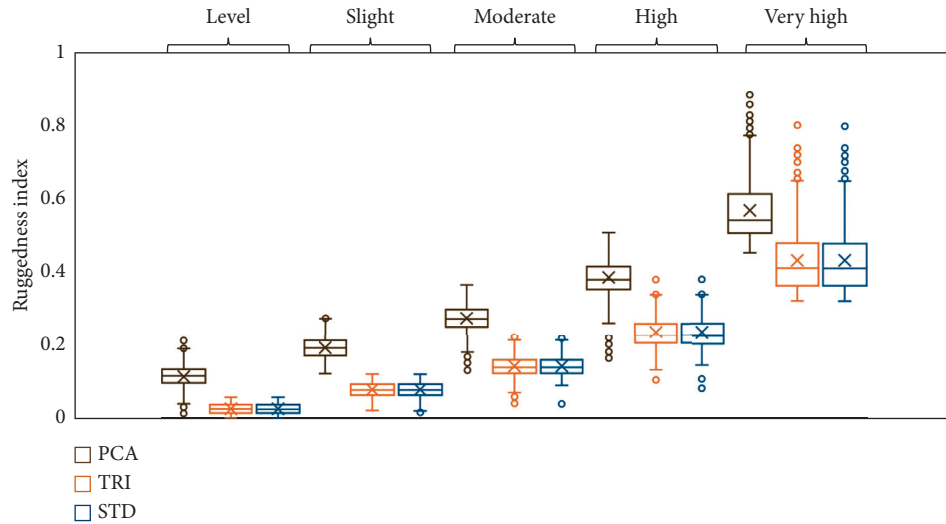


FIGURE 18: Box plot comparison between the investigated models.

82.3%, while it was possible to exceed 95% using the first and second components only. However, the surface ruggedness was created using all the components to reach the best accuracy.

The confusion matrix was defined for each pair of methods to evaluate their capabilities and variations, as clarified in Figure 17. It is observed that both TRI and STD are almost identical in which they provided the same zonal classification with an overall accuracy (matching) of 99.7%. This means that the STD model can replace the TRI without any negative influence. The difference between PCA and STD and TRI can be represented by an overall variation of about 20%. The PCA classified more points in the very high category compared to the TRI one that was capable of correctly identifying about 45% of the points only. This means that the PCA model is more sensitive for rugged areas in comparison with other methods. Pairwise comparison between the PCA and VRM models showed an overall difference of about 70%; besides, the ratio of points classified in the first class by VRM was almost 97% of the testing population. This observation is in line with the previous findings in Figure 14 and means that the VRM method is incapable of detecting moderate, high, and very high roughness areas. Finally, both STD and TRI results indicated about 44% correlation with the ones extracted from the VRM due to having a majority of the points classified in the level class of the VRM model.

The following comparisons and postanalysis of the results will omit the VRM model and consider PCA, TRI, and STD ones as most VRM points concentrated in the level

category. Figure 18 provides a class-based disparity between the PCA, TRI, and STD results through box plots. In general, it can be seen that the points' distribution in each PCA class takes a larger range as compared to the TRI and STD models that shrink the classes into minimal extents of propagation. Furthermore, it can be noticed that the limits of the level class in the PCA model are equal to that of the level, slight, and moderate of the TRI and STD ones combined. This means that the PCA method increases the ruggedness index values of the entire area as compared to the TRI and STD results even though their maps look very similar. On the other hand, the datasets' normal distribution in each of the investigated models is shown in Figure 19. It can be observed from the plots that the only significant difference between the PCA and the other algorithm is defined in the first class (level), where the PCA follows the normal distribution perfectly. Also, the spread of the other classes looks pretty similar in all of the indices. In general, Figure 20 shows a strong relationship between the results of PCA and those of TRI and STD. Thus, the regression equation (8) can be used to transfer TRI and STD models to the PCA one in Wadi Musa were proposed in this paper. The correlation coefficient of these equations is 0.9925 and 0.9926 in TRI and STD, respectively. These models can find application in many instances to overcome the need for high computational efforts.

$$\begin{aligned} \text{PCA} &= -1.116 \text{TRI}^2 + 1.805 \text{TRI} + 0.0745 \\ \text{PCA} &= -1.118 \text{STD}^2 + 1.807 \text{STD} + 0.0748. \end{aligned} \quad (8)$$

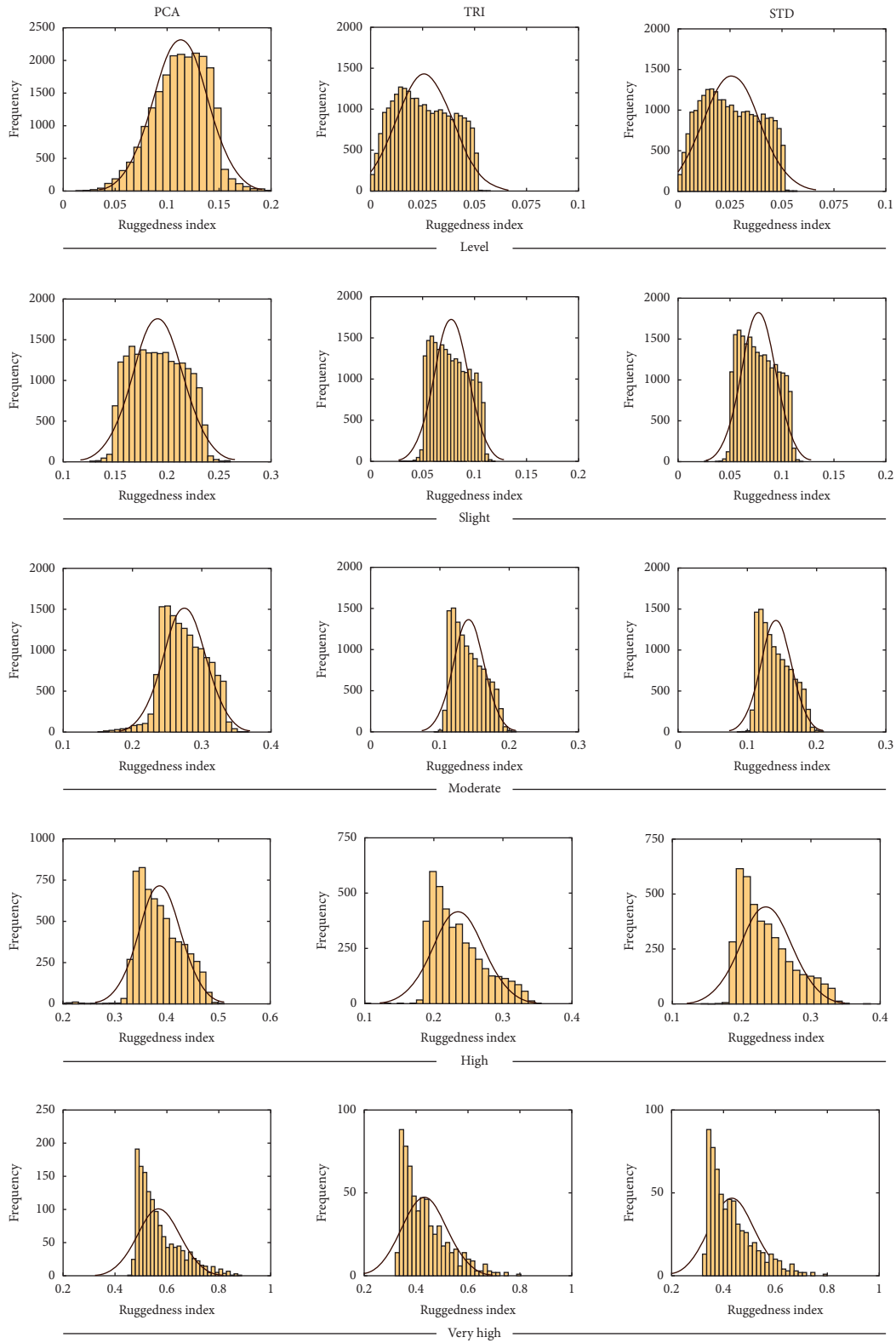


FIGURE 19: Distribution of dataset in each class of the investigated models.

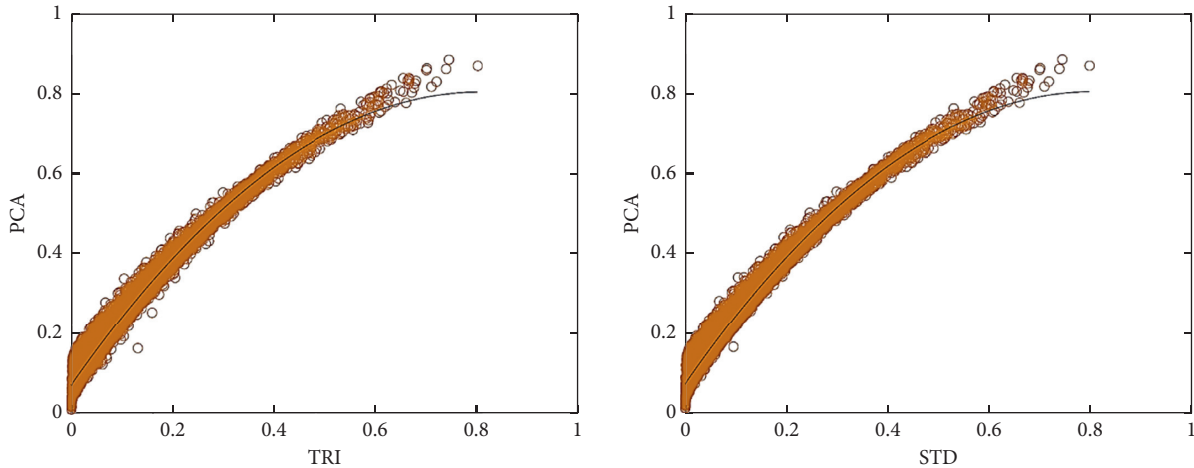


FIGURE 20: Regression models to transfer TRI and STD to PCA method.

4. Conclusion

Topographic ruggedness is an essential factor of terrain analysis for many applications in engineering and geomorphology. Several algorithms were introduced in the literature to show the spatial distribution of surface ruggedness but did not necessarily provide an optimal solution. This study presents a substitutional approach to generate terrain ruggedness using PCA and STD algorithms. Firstly, the effects of various DEM-derived topographic metrics on the ruggedness model were investigated using PCA. Secondly, the proposed methods are compared with the most common ones, namely, TRI and VRM. Indeed, the performance analysis of the ruggedness indices was conducted relying on visual inspection and statistical methods. The findings indicated the feasibility of applying PCA to generate an accurate ruggedness map consistent with the actual topographic features. STD model is simple to implement, available in ArcGIS software, and produces nearly identical results to TRI, the more complex method, while VRM showed the lowest performance. According to the statistical analysis, the ruggedness index's most influential variables are slope, TR, and SR. Finally, additional research is required to understand PCA's capabilities for describing terrain characteristics across various landscapes [72] and at different scales.

Abbreviation

TRI:	Terrain ruggedness index
DEM:	Digital elevation model
VRM:	Vector ruggedness measure
PCA:	Principal component analysis
GIS:	Geographic information systems
STD:	Standard deviation of elevation
CLs:	Contour lines
AMSL:	Above mean sea level
α :	Aspect
β :	Slope
C:	Curvature

SR:	Surface roughness
TR:	Topographic relief
TPI:	Topographic position index
TWI:	Topographic wetness index
SPI:	Stream power index
STI:	Sediment transport index.

Data Availability

Some or all data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] R. J. Pike, "Geomorphometry—diversity in quantitative surface analysis," *Progress in Physical Geography*, vol. 24, no. 1, pp. 1–20, 2000.
- [2] V. Lecours, M. F. J. Dolan, A. Micallef, and V. L. Lucieer, "A review of marine geomorphometry, the quantitative study of the seafloor," *Hydrology and Earth System Sciences*, vol. 20, no. 8, pp. 3207–3244, 2016.
- [3] G. Amatulli, D. McNerney, T. Sethi, P. Strobl, and S. Domisch, "Geomorpho 90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers," *Scientific Data*, vol. 7, no. 1, pp. 162–218, 2020.
- [4] G. P. B. Garcia and C. H. Grohmann, "DEM-based geomorphological mapping and landforms characterization of a tropical karst environment in southeastern Brazil," *Journal of South American Earth Sciences*, vol. 93, pp. 14–22, 2019.
- [5] T. X. Yue, Z. P. Du, D. J. Song, and Y. Gong, "A new method of surface modeling and its application to DEM construction," *Geomorphology*, vol. 91, no. 1–2, pp. 161–172, 2007.
- [6] L. Drăguț and C. Eisank, "Object representations at multiple scales from digital elevation models," *Geomorphology*, vol. 129, no. 3–4, pp. 183–189, 2011.

- [7] M. J. Smith and C. D. Clark, "Methods for the visualization of digital elevation models for landform mapping," *Earth Surface Processes and Landforms*, vol. 30, no. 7, pp. 885–900, 2005.
- [8] M. Uysal, A. S. Toprak, and N. Polat, "DEM generation with UAV Photogrammetry and accuracy analysis in Sahitler hill," *Measurement*, vol. 73, pp. 539–543, 2015.
- [9] S. Ghuffar, "DEM generation from multi satellite PlanetScope imagery," *Remote Sensing*, vol. 10, no. 9, p. 1462, 2018.
- [10] H. Taud, J.-F. Parrot, and R. Alvarez, "DEM generation by contour line dilation," *Computers & Geosciences*, vol. 25, no. 7, pp. 775–783, 1999.
- [11] J. Shan and S. Aparajithan, "Urban DEM generation from raw LiDAR data," *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 2, pp. 217–226, 2005.
- [12] D. Bolkas, G. Fotopoulos, A. Braun, and I. N. Tziavos, "Assessing digital elevation model uncertainty using GPS survey data," *Journal of Surveying Engineering*, vol. 142, no. 3, Article ID 04016001, 2016.
- [13] M. Habib, Y. Alzubi, A. Malkawi, and M. Awwad, "Impact of interpolation techniques on the accuracy of large-scale digital elevation model," *Open Geosciences*, vol. 12, no. 1, 2020.
- [14] W. Zhang and D. R. Montgomery, "Digital elevation model grid size, landscape representation, and hydrologic simulations," *Water Resources Research*, vol. 30, no. 4, pp. 1019–1028, 1994.
- [15] R. Turcotte, J. P. Fortin, A. N. Rousseau, S. Massicotte, and J. P. Villeneuve, "Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network," *Journal of Hydrology*, vol. 240, no. 3-4, pp. 225–242, 2001.
- [16] Q. Zhou and X. Liu, "Analysis of errors of derived slope and aspect related to DEM data properties," *Computers & Geosciences*, vol. 30, no. 4, pp. 369–378, 2004.
- [17] H. Saadat, R. Bonnell, F. Sharifi, G. Mehuis, M. Namdar, and S. Ale-Ebrahim, "Landform classification from a digital elevation model and satellite imagery," *Geomorphology*, vol. 100, no. 3-4, pp. 453–464, 2008.
- [18] A. K. Al-Fugara, M. Ahmadlou, A. R. Al-Shabeeb et al., "Spatial mapping of groundwater springs potentiality using grid search-based and genetic algorithm-based support vector regression," *Geocarto International*, vol. 79, pp. 1–20, 2020.
- [19] N. Nunn and D. Puga, "Ruggedness: the blessing of bad geography in Africa," *The Review of Economics and Statistics*, vol. 94, no. 1, pp. 20–36, 2012.
- [20] M. Khawlie, M. Awad, A. Shaban, R. B. Kheir, and C. Abdallah, "Remote sensing for environmental protection of the eastern Mediterranean rugged mountainous areas, Lebanon," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 57, no. 1-2, pp. 13–23, 2002.
- [21] T. Skogland, "Life History Characteristics of Wild Reindeer (*Rangifer Tarandus* L.) in Relation to Their Food Resources; Ecological Effects and Behavioral Adaptations," *Direktoratet for vilt og ferskvannsfisk*, vol. 5, 1985.
- [22] U. Matthes-Sears, W. C. Matthes-Sears, S. J. Hastings, and W. C. Oechel, "The effects of topography and nutrient status on the biomass, vegetative characteristics, and gas exchange of two deciduous shrubs on an arctic tundra slope," *Arctic and Alpine Research*, vol. 20, no. 3, pp. 342–351, 1988.
- [23] B. K. Pokhrel, P. Chevallier, V. Andréassian et al., "Comparison of two snowmelt modelling approaches in the Dudh Koshi basin (eastern Himalayas, Nepal)," *Hydrological Sciences Journal*, vol. 59, no. 8, pp. 1507–1518, 2014.
- [24] H. Han, W. Chung, J. Song, A. Seol, and J. Chung, "A terrain-based method for selecting potential mountain ridge protection areas in South Korea," *Landscape Research*, vol. 41, no. 8, pp. 906–921, 2016.
- [25] D. Asfaw and G. Workineh, "Quantitative analysis of morphometry on Ribb and Gumara watersheds: implications for soil and water conservation," *International Soil and Water Conservation Research*, vol. 7, no. 2, pp. 150–157, 2019.
- [26] C. Nellemann and G. Fry, "Quantitative analysis of terrain ruggedness in reindeer winter grounds," *Arctic*, vol. 48, pp. 172–176, 1995.
- [27] S. L. Beasom, E. P. Wiggers, and J. R. Giardino, "A technique for assessing land surface ruggedness," *Journal of Wildlife Management*, vol. 47, no. 4, pp. 1163–1166, 1983.
- [28] S. J. Riley, S. D. DeGloria, and R. Elliot, "A terrain ruggedness index that quantifies topographic heterogeneity," *Intermountain Journal of Sciences*, vol. 5, no. 1–4, pp. 23–27, 1999.
- [29] J. M. Sappington, K. M. Longshore, and D. B. Thompson, "Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert," *Journal of Wildlife Management*, vol. 71, no. 5, pp. 1419–1426, 2007.
- [30] T. Popit and T. Verbovšek, "Analysis of surface roughness in the Sveta Magdalena paleo-landslide in the Rebrnice area," *RMZ Materials and Geoenvironment*, vol. 60, pp. 197–204, 2013.
- [31] K. Pearson, "LIII. On lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901.
- [32] Q. Cheng, G. Bonham-Carter, W. Wang, S. Zhang, W. Li, and X. Qinglin, "A spatially weighted principal component analysis for multi-element geochemical data for mapping locations of felsic intrusions in the Gejiu mineral district of Yunnan, China," *Computers & Geosciences*, vol. 37, no. 5, pp. 662–669, 2011.
- [33] H. Guo, K. J. Marfurt, and J. Liu, "Principal component spectral analysis," *Geophysics*, vol. 74, no. 4, pp. P35–P43, 2009.
- [34] H. Faraji Sabokbar, M. Shadman Roodposhti, and E. Tazik, "Landslide susceptibility mapping using geographically-weighted principal component analysis," *Geomorphology*, vol. 226, pp. 15–24, 2014.
- [35] Y. Farhan, A. Anbar, N. Al-Shaikh, and R. Mousa, "Prioritization of semi-arid agricultural watershed using morphometric and principal component analysis, remote sensing, and GIS techniques, the Zerqa River Watershed, Northern Jordan," *Agricultural Sciences*, vol. 8, no. 1, pp. 113–148, 2016.
- [36] J. P. Praene, B. Malet-Damour, M. H. Radanielina, L. Fontaine, and G. Rivière, "GIS-based approach to identify climatic zoning: a hierarchical clustering on principal component analysis," *Building and Environment*, vol. 164, Article ID 106330, 2019.
- [37] T. Basu, A. Das, and S. Pal, "Application of geographically weighted principal component analysis and fuzzy approach for unsupervised landslide susceptibility mapping on Gish River Basin, India," *Geocarto International*, vol. 1-24, 2020.
- [38] A. I. Petrişor, I. Ianoş, D. Iurea, and M. N. Văidianu, "Applications of principal component analysis integrated with GIS," *Procedia Environmental Sciences*, vol. 14, pp. 247–256, 2012.
- [39] F. Huang, J. Yang, B. Zhang, Y. Li, J. Huang, and N. Chen, "Regional terrain complexity assessment based on principal component analysis and geographic information system: a

- case of Jiangxi province, China,” *ISPRS International Journal of Geo-Information*, vol. 9, no. 9, p. 539, 2020.
- [40] T. McKinney, S. R. Boe, and J. C. deVos Jr, “GIS-based evaluation of escape terrain and desert bighorn sheep populations in Arizona,” *Wildlife Society Bulletin*, vol. 31, pp. 1229–1236, 2003.
- [41] N. Bragin, S. Amgalanbaatar, G. Wingard, and R. P. Reading, “Creating a model of habitat suitability using vegetation and ruggedness for *Ovis ammon* and *Capra sibirica* (Artiodactyla: bovidae) in Mongolia,” *Journal of Asia-Pacific Business*, vol. 10, no. 3, pp. 390–395, 2017.
- [42] D. R. Montgomery and M. T. Brandon, “Topographic controls on erosion rates in tectonically active mountain ranges,” *Earth and Planetary Science Letters*, vol. 201, no. 3–4, pp. 481–489, 2002.
- [43] S. Das, P. P. Patel, and S. Sengupta, “Evaluation of different digital elevation models for analyzing drainage morphometric parameters in a mountainous terrain: a case study of the Supin-Upper Tons Basin, Indian Himalayas,” *SpringerPlus*, vol. 5, no. 1, p. 1544, 2016.
- [44] P. Tarolli, “High-resolution topography for understanding Earth surface processes: opportunities and challenges,” *Geomorphology*, vol. 216, pp. 295–312, 2014.
- [45] M. Habib and A. K. M. S. Al-Fugara, “Evaluation of alternative conformal mapping for geospatial data in Jordan,” *Journal of Applied Geodesy*, vol. 13, no. 4, pp. 335–344, 2019.
- [46] J. P. Wilson and J. C. Gallant, Eds., *Terrain Analysis: Principles and Applications*, John Wiley & Sons, Hoboken, NJ, USA, 2000.
- [47] J. Gao, “Resolution and accuracy of terrain representation by grid DEMs at a micro-scale,” *International Journal of Geographical Information Science*, vol. 11, no. 2, pp. 199–212, 1997.
- [48] J. A. Thompson, J. C. Bell, and C. A. Butler, “Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling,” *Geoderma*, vol. 100, no. 1–2, pp. 67–89, 2001.
- [49] J. Vaze, J. Teng, and G. Spencer, “Impact of DEM accuracy and resolution on topographic indices,” *Environmental Modelling & Software*, vol. 25, no. 10, 2010.
- [50] Y. Kubota, H. Murata, and K. Kikuzawa, “Effects of topographic heterogeneity on tree species richness and stand dynamics in a subtropical forest in Okinawa Island, southern Japan,” *Journal of Ecology*, vol. 92, no. 2, pp. 230–240, 2004.
- [51] V. Chaplot, F. Darboux, H. Bourennane, S. Legu  dois, N. Silvera, and K. Phachomphon, “Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landform types and data density,” *Geomorphology*, vol. 77, no. 1–2, pp. 126–141, 2006.
- [52] M. F. Hutchinson, “Calculation of hydrologically sound digital elevation models,” in *Proceedings of the 3rd International Symposium on Spatial Data Handling*, vol. 133, pp. 117–133, Sydney, Australia, August 19 1988.
- [53] D. M. Mark, “Geomorphometric parameters: a review and evaluation,” *Geografiska Annaler-Series A: Physical Geography*, vol. 57, no. 3–4, pp. 165–177, 1975.
- [54] A. Bolongaro-Crevenna, V. Torres-Rodr  guez, V. Sorani, D. Frame, and M. A. Ortiz, “Geomorphometric analysis for characterizing landforms in Morelos State, Mexico,” *Geomorphology*, vol. 67, no. 3–4, pp. 407–422, 2005.
- [55] W. Tobler, “The development of analytical cartography: a personal note,” *Cartography and Geographic Information Science*, vol. 27, no. 3, pp. 189–194, 2000.
- [56] P. A. Burrough, R. McDonnell, R. A. McDonnell, and C. D. Lloyd, *Principles of Geographical Information Systems*, Oxford University Press, Oxford, UK, 2015.
- [57] A. Redecker, J. Burian, N. Moos, and K. Mack  , “Spatial analysis in geomatics,” in *Spatonomy*, pp. 65–92, Springer, Cham, Berlin, Germany, 2020.
- [58] C. H. Grohmann, M. J. Smith, and C. Riccomini, “Multiscale analysis of topographic surface roughness in the Midland Valley, Scotland,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 4, pp. 1200–1213, 2010.
- [59] R. D. Hobson, *Surface Roughness in Topography: Quantitative approach. Spatial Analysis in Geomorphology*, pp. 221–245, Methuen, London, UK, 1972.
- [60] L. W. Zevenbergen and C. R. Thorne, “Quantitative analysis of land surface topography,” *Earth Surface Processes and Landforms*, vol. 12, no. 1, pp. 47–56, 1987.
- [61] I. D. Moore, R. B. Grayson, and A. R. Ladson, “Digital terrain modelling: a review of hydrological, geomorphological, and biological applications,” *Hydrological Processes*, vol. 5, no. 1, pp. 3–30, 1991.
- [62] I. K. Robinson, “Crystal truncation rods and surface roughness,” *Physical Review B*, vol. 33, no. 6, pp. 3830–3836, 1986.
- [63] K. J. Beven and M. J. Kirkby, “A physically based, variable contributing area model of basin hydrology/Un mod  le    base physique de zone d’appel variable de l’hydrologie du bassin versant,” *Hydrological Sciences Bulletin*, vol. 24, no. 1, pp. 43–69, 1979.
- [64] P. F. Quinn, K. J. Beven, and R. Lamb, “The in(a/tan β) index: how to calculate it and how to use it within the topmodel framework,” *Hydrological Processes*, vol. 9, no. 2, pp. 161–182, 1995.
- [65] I. D. Moore, G. J. Burch, and D. H. Mackenzie, “Topographic effects on the distribution of surface soil water and the location of ephemeral gullies,” *Transactions of the ASAE*, vol. 31, no. 4, pp. 1098–1107, 1988.
- [66] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*, Elsevier, Amsterdam, Netherlands, 2011.
- [67] G. Jord  n, B. M. L. Meijninger, D. J. J. v. Hinsbergen, J. E. Meulenkamp, and P. M. v. Dijk, “Extraction of morphotectonic features from DEMs: development and applications for study areas in Hungary and NW Greece,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 7, no. 3, pp. 163–182, 2005.
- [68] M. Wiczeorek and P. Migo  n, “Automatic relief classification versus expert and field based landform classification for the medium-altitude mountain range, the Sudetes, SW Poland,” *Geomorphology*, vol. 206, pp. 133–146, 2014.
- [69] V. Moudr  y, V. Lecours, K. Gdulov   et al., “On the use of global DEMs in ecological modelling and the accuracy of new bare-earth DEMs,” *Ecological Modelling*, vol. 383, pp. 3–9, 2018.
- [70] G. F. Jenks, “The data model concept in statistical mapping,” *International Yearbook of Cartography*, vol. 7, pp. 186–190, 1967.
- [71] J. Chen, S. T. Yang, H. W. Li, B. Zhang, and J. R. Lv, “Research on geographical environment unit division based on the method of natural breaks (Jenks),” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 47–50, 2013.
- [72] L. U. Huaxing, “Modelling terrain complexity,” in *Advances in Digital Terrain Analysis*, pp. 159–176, Springer, Berlin, Germany, 2008.