GEODESY AND CARTOGRAPHY
Vol. 70, no. 1, article no. e04, 2021

DOI: https://doi.org/10.24425/gac.2021.136679

Original paper

High-resolution soil erodibility K-factor estimation using machine learning generated soil dataset and soil pH levels

Nurlan Mammadli[1]*, Magsad Gojamanov[2]

[1]Azerbaijan National Academy Sciences, Baku, Azerbaijan
e-mail: nurlan.m.h@gmail.com; ORCID: http://orcid.org/0000-0002-8594-5702

[2]Baku State University, Baku, Azerbaijan
e-mail: mgodja@yandex.ru; ORCID: http://orcid.org/0000-0002-5653-5675

*Corresponding author: Nurlan Mammadli, e-mail: nurlan.m.h@gmail.com

Received: 2021-01-23 / Accepted: 2021-05-19

Abstract: Soil Erodibility Factor (K-factor) is a crucial component of a widely used equation for soil erosion assessment known as the USLE (Universal Soil Loss Equation) or its revised version – RUSLE. It reflects the potential of the soil of being detached due to rainfalls or runoffs. So far, an extensive number of researches provide different approaches and techniques in the evaluation of K-factor. This study applies soil erodibility estimation in the soils of the South Caucasian region using soil data prepared by the International Soil Reference and Information Centre (ISRIC) with 250 m resolution, whereas the recent K-factor estimation implemented in the EU scale was with 500 m resolution. Soil erodibility was assessed using an equation involving soil pH levels. The study utilises Trapesoidal equation of soil data processing and preparation, as suggested by ISRIC, for various layers of surface soil data with up to 0-30 cm depth. Both usage of SoilGrids data and its processing as well as estimation of K-factor applying soil pH levels have demonstrated sufficient capacity and accuracy in soil erodibility assessment. The final output result has revealed the K-factor values varying from 0.037 and more than 0.060 t ha h/MJ mm within the study area.

Keywords: soil erodibility, RUSLE, SoilGrids, K factor, soil pH

---

1. Introduction

1.1. Soil erodibility factor

Soil erodibility factor – K, is a crucial component of Revised/Universal Soil Loss Equation (RUSLE) used for estimation of the potential of slope to be susceptible to erosion due to the influence of different soil properties (Renard et al., 1997).

K-factor calculation excludes soil conservation practices and calculates only mean annual soil loss per unit of rainfall erosivity at normal soil condition (Morgan, 2005). Hence, K-factor values are interpreted as high values showing areas more susceptible to erosion and lower values representing less potential for erosion (Adornado et al., 2009).

The general philosophy behind the erodibility factor is that the soils containing high clay content usually have lower K-factor values because they are more resistant to detachment leading to lower erodibility potential. On the other hand, soils with high silt content have more potential to erode due to poor ability to keep integrity, additionally, performing high rates of run-offs are resulting in high K-factor as well (Kirchmeir and Berger, 2019).

Significant number of conditions available which can affect soil's resistance to separation and transportation, such as topography, climate, vegetation cover, soil and water conservation, as well as comprehensive assessment of soil erodibility K-factor play an essential role in estimation of soil loss via RUSLE equation. According to many researches implemented in field of K-factor performance evaluation under different environments, most of them emphasised that precise and accurate determination of erodibility factor significantly affect RUSLE outcomes (Lin et al., 2019).

Considering the fact that, the soil has extremely diverse properties, their accurate estimation is important in order to identify and map locations with potential for erosion.

1.2. Calculations of soil erodibility factor

Soil erodibility factor has been initially defined by Wischmeier and Smith (1978) and further developed by Renard et al. (1997). Most studies refer to the USLE equation by Wischmeier, which consider variables for soil textural information, organic matter, soil structure and profile permeability class. The latter two input information are soil structure codes, which are assumed in four different classes: 1 – very fine granular; 2 – fine granular, 3 – medium or coarse granular; and 4 – blocky, platy.

Furthermore, soil permeability class also plays an important role in the equation being also assumed with 5 or 6 classes as: 1 – rapid, 2 – moderate to rapid; 3 – moderate; 4 – slow to moderate; 5 – slow; 6 – very slow.

This is the most common way of applying K-factor calculation by researchers applying USLE technique in various parts of the world like Turkey (Ozsoy et al., 2012), the European Union (Panagos et al., 2014).

However, despite the popularity of the USLE equation it has its limitations depending on the geography of the study area and the set objectives of the study or data availability. In addition to Renard et al. (1997) there are other researchers who modified this equation either to suit their study, or to improve accuracy, or to simplify the large amount of variable required for it. For example, Williams and Renard (1983) and later Chen et al. (2011) managed to develop an equation without using soil structure and profile permeability class.

In this study, David's (1988) K-factor equation will be used considering its simplification of the version used by Wischmeier and Mannering (1969), taking into account only 5 variables: sand (%), silt (%), clay (%), organic matter (%) and pH applied in USA scale. Additionally, Hernandez et al. (2012) have also successfully applied this equation in their study on the Philippines. Considering that this research is focusing on application of RUSLE – the revised versions of K-factor is also used, which usually exclude soil structure and profile permeability inputs from its equation (Benavidez et al., 2018).

The other main objective of using David's (1988) equation is that it is assumption based on the data of local soil, unlike the USLE equation by Wischmeier and Smith (1978) which is based on USA Midwestern soil with specific silt fraction (Renard et al., 1997). According to the study made to test empirical models with actual erodibility performance by Wawer et al. (2005), the main inaccuracy was revealed in the level of silt in soil, i.e. if the silt content in the soil exceeds 70%, which is the main limitation of K-factor equations being sensitive to geographical specificity of study areas (Benavidez et al., 2018).

On the other hand, focusing on South Caucasian scale there has been studies done by Kirchmeir and Michael (2016) in Georgia using local soil maps with 1:200 k scale and Bayramov et al. (2018) in Azerbaijan with 1:100 k scale as a base for K-factor extraction. However, the usage of soil maps that are not up to date, as well as being in large scale provide rough spatial information and high variability due to omission of micro-relief (Kirchmeir and Berger, 2019).

Therefore, this study has exceptionally focused on obtained data from SoilGrids with 250 m high resolution of soil texture raster data, which provides K-factor output with high precision and accuracy. In addition, considering that the calculated silt level (see below) is maximum 50% giving confidence in adopting a revised methods of erosion calculation.

The main purpose of this research is to visualise the extent of the soil erodibility (K-factor) by mapping the impacted area, which can be also used as part of further soil erosion assessments.

2. Materials and methods

2.1. Study area

The study area is located in one of the districts of Azerbaijan – Shamakhi (40°57′N – 40°56′N and 48°61′E – 48°67′E) (Fig. 1). The territory with a total area of 2320 km^2 has a very diverse terrain comprising of highlands with up to 2618 m high mountains and lowlands reaching –25 below sea level. In addition, the territory has a unique landscape with Pirgulu State Nature Reserve on one hand and a very intensive agricultural activities on the other hand.

Almost up to half of Shamakhi's land is used for agricultural purposes 1100 km^2 (48%), however in much higher elevations in the north, it is predominantly covered by vegetation and forest, 388 km^2 (20%). There is also a significant pasture capacity, comprising of 800 km^2 (35%) (Buchhorn et al., 2019).

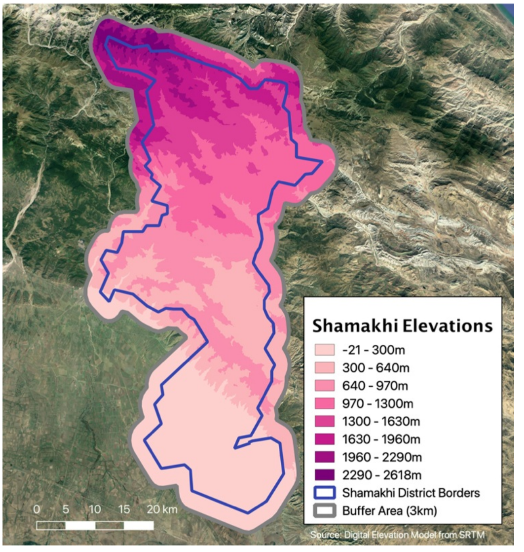

Fig. 1. Location and elevation map of the study area

2.2. Data source

The data used for soil erodibility assessment in this study was obtained from SoildGrids System. SoilGrids is a project developed by ISRIC (International Soil Reference and Information Centre), on open source principle to serve researchers and other users to get accurate soil information for detailed assessment and studies.

The database is build based on a large collection of georeferenced soil profile data generated and managed by World Soil Information Services. There are up to 6 million soil records with georeferenced point data gathered worldwide in order to provide accurate data for SoilGrids. This also includes multiple reference points gathered for South Caucasus as well (Fig. 2).

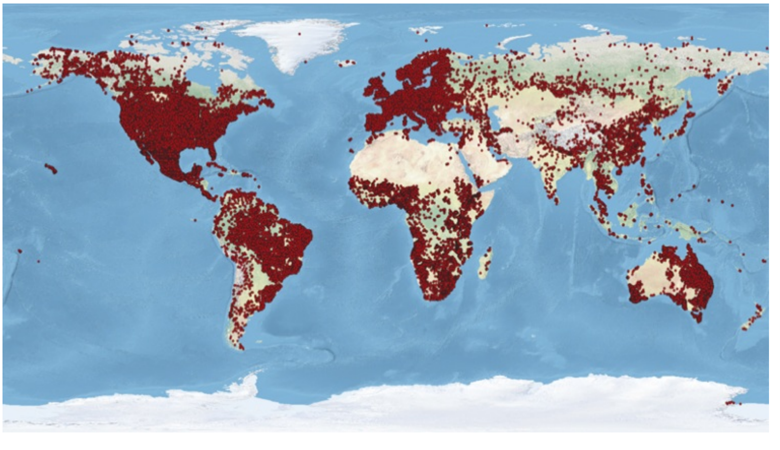

Fig. 2. Worldwide location of soil profiles

The dataset was further processes using a machine learning method to map the spatial distribution of soil properties around the world. SoilGrids system is available with 250 m resolution and latest updated in June 2016. The dataset includes such profile properties as: organic carbon, bulk density, pH, soil texture fractions, chemical compositions and other valuable data sources. The datasets are available in seven depths (0, 5, 15, 30, 60, 100 and 200 cm). The remote sensing data, SoilGrids is based on, is obtained from MODIS, SRTM DEM satellite derivatives, climatic images and other landform maps (Hengl et al., 2014).

2.3. Data collection processing

In order to prepare the data for further K-factor calculation, initially several layers of soil data are required. The type of soil layers to be obtained are identified according to the variables in the soil erodibility equation to be applied.

Nevertheless, SoilGrids provide data for seven layers at certain depth, however for this study an overall depth about 30 cm was considered. In order to generate depths for different soil properties, vertical disintegration calculation was required (Arrouays et al., 2014).

Four layers of data for various soil properties with depth in 0 mm, 5 mm, 15 mm, 30 mm were acquired for each below layer:

- clay layer (Clay Content 0–2 um mass fraction in % at 0, 5, 15, 30 mm);
- sand layer (Sand Content 50–2000 um mass fraction in % at 0, 5, 15, 30 mm);
- silt layer (Silt Content 2–50 um mass fraction in % at 0, 5, 15, 30 mm);
- soil Organic Matter layer (fine earth fraction in g/kg);
- soil pH level (at 0, 5, 15, 30 mm).

The Figure 3 shows the diagram with consequent steps of soil erodibility computation, including data collection, data processing and preparation and K-factor quantifications with final output map.

All the layers are available in percentage format, except for the soil organic matter that was subsequently converted to the percentage format from g/km.

In order to obtain certain depth values, depth at 30 cm is generally used for further processing, which also might depend on soil depth data availability. The processing is based on ISRIC (International Soil Reference Information Centre) – World Soil prediction methodology using Trapezoidal rule described by Hengl et al. (2017).

The averages of each layer of overall depth intervals for 0 to 30 cm are derived by taking a weighted average of the predictions within the depth interval using numerical integration through Trapezoidal equation:

$$\frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x \approx \frac{1}{(b-a)}\frac{1}{2}\sum_{k=1}^{N-1}(x_{k+1}-x_k)(f(x_k)+f(x_{k+1}))$$

where:

N – number of depth,

x_k – k-th depth,

$f(x_k)$ – value of target variable at k'th depth.

This equation aggregates different available layers over standard depth interval of 0–30 cm depth by taking a weighted average using trapezoidal rule.

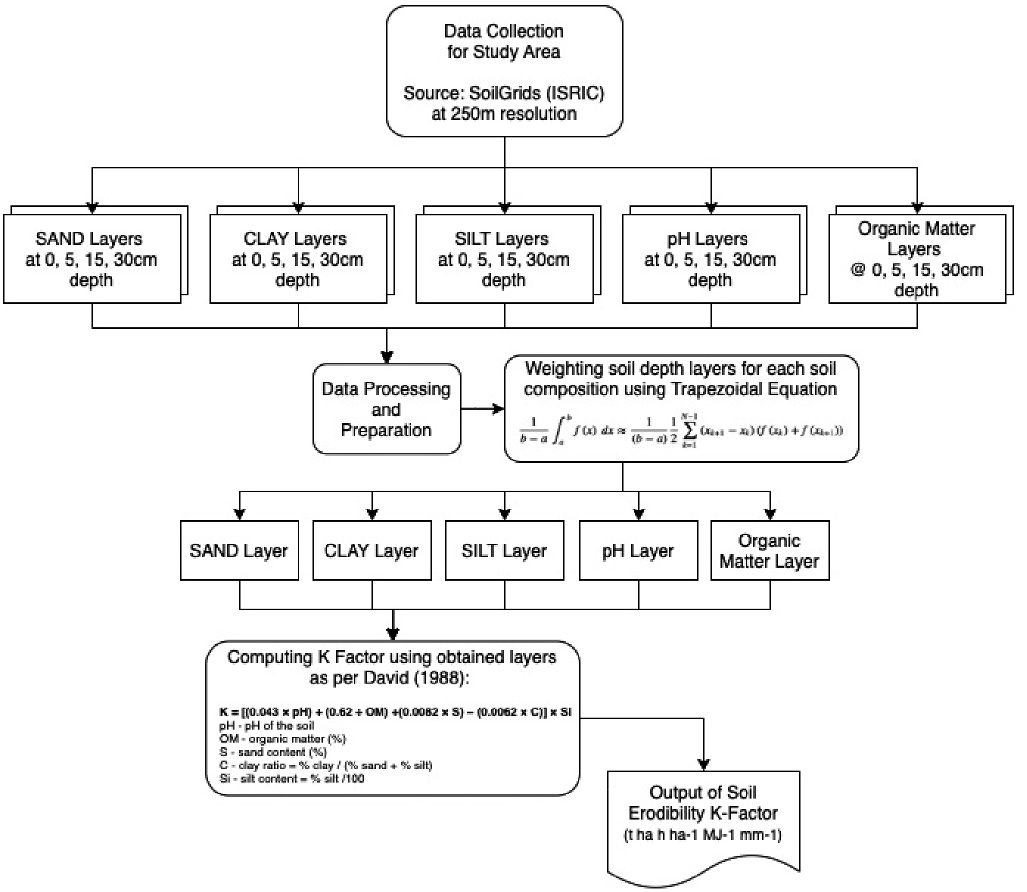

Fig. 3. The methodological diagram for soil erodibility K-factor computation

The above equation has been translated and processed with GIS in the following way:

$$(5-0)\times(0+5)+(15-5)\times(5+15)+(30-15)\times(15+30)\div 30.2$$

This equation shows the application of Trapezoidal rule in an example of weighted average of the Organic matter of soil content of the study area in 30m depth which was calculated based on the 4 depth layers: 0, 5, 15, 30 cm (Fig. 4).

This way was adopted to the all other remaining layers: clay, sand and silt. The following table summarises calculated values for each soil property layers (Table 1), and Figure 5, show five raster output maps of evaluated layers.

After obtaining values for all the layers improved K-factor equation has been used based on the most recent research on Tibet by Yuanyuan et al. (2018).

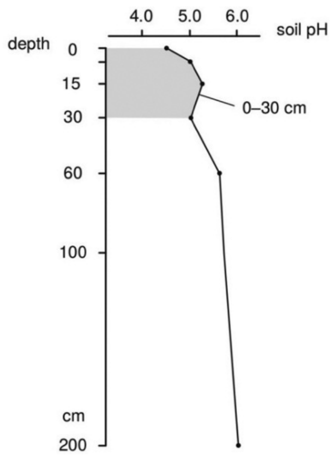

Fig. 4. Standard soil depth based on Global Soil Map specifications and the result of the numerical integration according to the trapezoidal rule (Hengl et al., 2017)

2.4. K-factor computation

In the introductory section provides details on the applied K-factor equation was provided. The equation has been proposed by David (1988) for K-factor calculation can be expressed as follows:

$$K = \big[(0.043 \times \mathrm{pH}) + (0.62 \div \mathrm{OM}) + (0.0082 \times \mathrm{S}) - (0.0062 \times \mathrm{C})\big] \times \mathrm{Si}$$

where:
- pH – pH of the soil,
- OM – organic matter (%),
- S – sand content (%),
- C – clay ratio = % clay / (% sand + % silt),
- Si – silt content = % silt /100.

K-factor equation has been formulated based on imperial units providing output as in ton acre per hundreds of acres per foot per tons per inch. In order to convert to SI units from imperial units the final result was multiplied by 0.1317, providing the outputs as metric tons hectare hour per hectare per mega-joule per millimetre (Renard et al., 1997).

3. Results and discussion

3.1. Soil properties

The processed soil property layers which were further used to quantify the soil erodibility factor are presented in the Figure 5, as well as highest and lowest values for each soil property layer is presented in the Table 1.

Created each layer of the soil property map clearly shows diverse distribution of soil content within the area depending on elevation and vegetation cover. Hence, silt percentage (Fig. 5b) is significantly higher in relatively low elevated zones with low vegetation, whereas in the northern parts of the study area silt content is significantly low.

Table 1. Values of soil property layers generated applying Trapezoidal equation

Soil Property	High Value	Low Value	Avg. Value
Sand %	11.41	59.41	35.44
Silt %	29.66	50.08	39.90
Clay %	8.75	46.25	27.49
Organic Matter %	0.71	24.67	12.69
pH	5.52	8.29	6.90

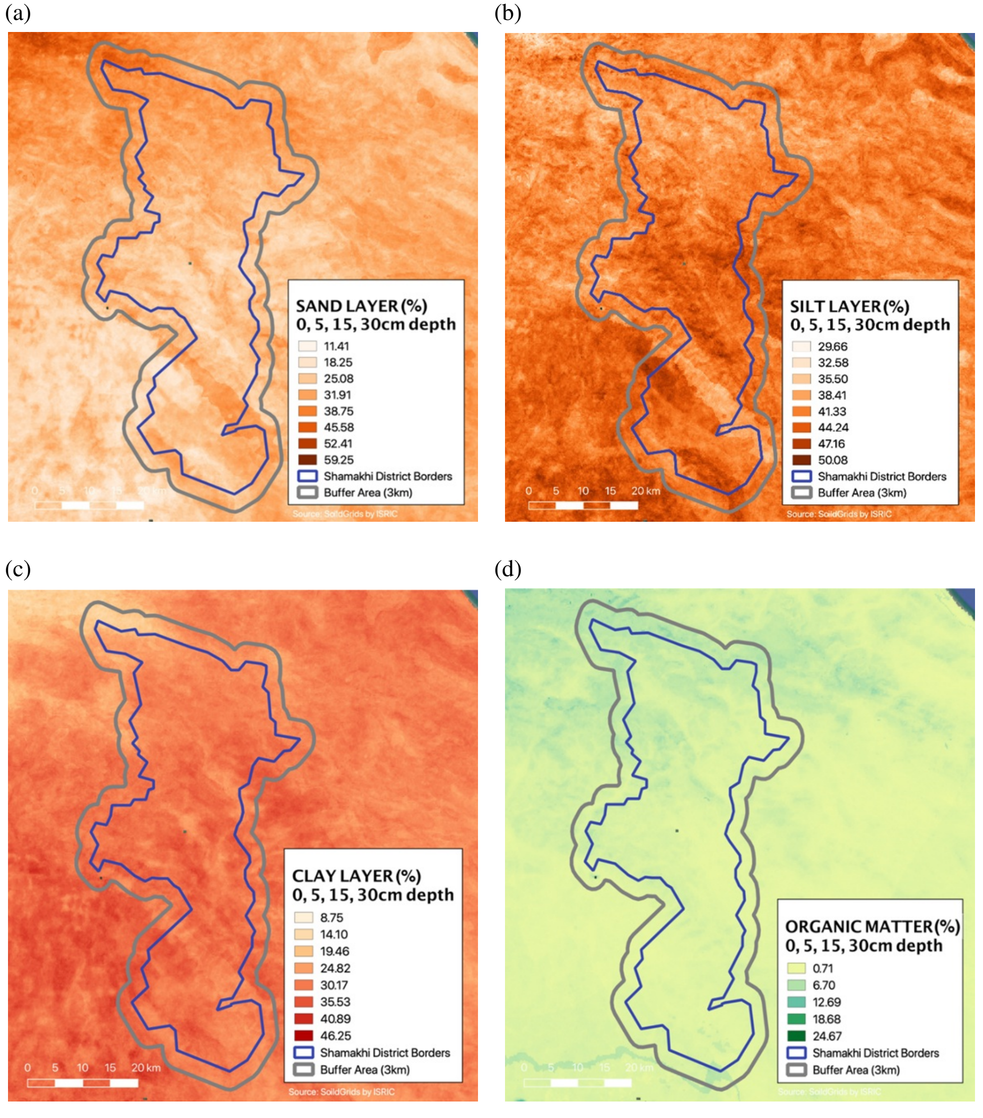

Fig. 5. Raster output maps for five soil property layers: (a) sand, (b) silt, (c) clay and (d) organic matter generated as a result of Trapezoidal rule

However, the percentage of organic matter (Fig. 5d) in the northern parts is higher, correlating with high density of vegetation in this area.

The percentage of sand (Fig. 5a) did not show high content overall with insignificantly high levels in some areas only.

On the other hand, the clay content (Fig. 5c) similarly to the silt, in overall area, is significantly higher reaching up to 46% and 50% respectively, intensifying in more central parts especially.

Despite the fact that, sand shows highest percentage (60%) overall, its dissipation around the study area averages reaching only approximately 35%.

The pH level (Fig. 6) is particularly higher in the lowlands with low and poor vegetation cover reaching up to 8 pH and in dense forest areas as low as 5.5 pH.

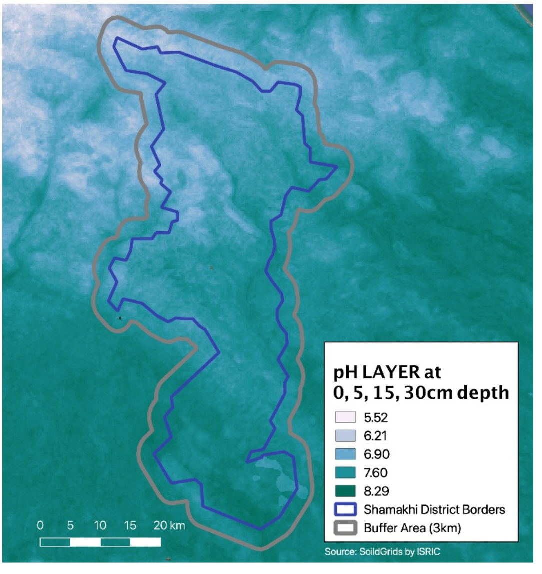

Fig. 6. Raster output maps for pH level of soil in study area generated as a result of Trapesoidal rule

3.2. Soil erodibility factor

The Soil erodibility factor was calculated using generated soil property percentages and soil pH levels. As seen from the final map shown in Figure 7, some areas produced high erodibility levels, reaching 0.060 t ha h/MJ mm and the lowest amount with less than 0.037 t ha h/MJ mm averaging in 0.039. The distribution of erosion prone regions is observed not to be even, but rather concentrated in lower elevated zones, primarily in south-west and mid-western parts of the study area.

The current study demonstrated the quantification of soil erodibility areas showing most vulnerable ground surfaces for potential erosion in Shamakhi region, in particular with significant levels in mid-eastern and southern parts.

Even though there are multiple ways of soil erodibility quantification, this study focuses on the application of the equation additionally focusing on soil pH levels. Con-

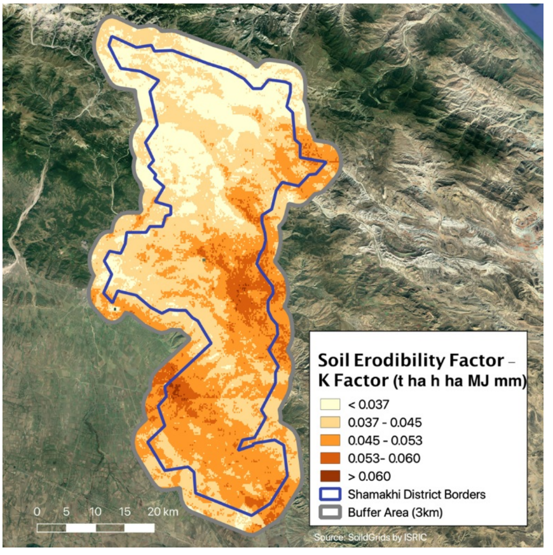

Fig. 7. Soil Erodibility Factor of study area

sidering limitations of the USLE equation and availability of pH data provides a more efficient way of obtaining soil erodibility results with high resolution outputs.

In addition, the high-resolution end results due to open source SoilGrids data with 250 m sized soil data used for this study provides confidence in accuracy of the results. Most of the studies mentioned previously, undertaken on the regional level, rely on the scaled maps produced with very low resolution, due to unavailability of satellite data on soil properties for certain regions. This can compromise the accuracy of further erosion assessment results by giving a very wide range of erosion prone areas with potential miscalculation of soil loss quantification, which is eventually leading to high costs for erosion control and management measures.

4. Conclusions

The achievement of accuracy in soil erodibility factor estimation plays crucial role to reach precision in overall soil water erosion assessment. The presented research aims to show a possible way of developing a high-resolution map for soil erodibility factor.

Therefore, the soil dataset that has been processed using machine learning technique by SoilGrids has been used. Despite the fact that digitally achieved soil data, based on physically sampled soil profiles, is a relatively recently developing technique, its high-resolution helps to achieve accurate mapping of erodibility areas.

Hence, the study at hand provides practical insights into how to achieve accuracy in calculation of RUSLE factors for further soil loss quantification. Moreover, it also provides usage of soil property layers in combination with the pH level to obtain the soil erodibility factor particularly for soils in the Southern Caucasus region.

Author contributions

Conceptualization: M.N.; Methodology development, Data analysis, Writing – original and editing: M.N..; Review and editing: G.M.

Data availability statement

The datasets used in this research study are available at SoilGrids under the Open Database License (ODbl) v1.0 and can be downloaded from www.soilgrids.org and/or ftp.soilgrids.org without restrictions. SoilGrids250m data has already been released in July 2016.

Acknowledgements

This research study has not received any external funding.

References

Adornado, H.A., Yoshida, M. and Apolinares, H. (2009). Erosion Vulnerability Assessment in REINA, Quezon Province, Philippines with Raster-based Tool Built within GIS Environment. *J. Agric. Res.*

Arrouays, D., Grundy M.G., Hartemink A.E. et al. (2014). Chapter Three – GlobalSoilMap: Toward a Fine-Resolution Global Grid of Soil Properties. In: Sparks D.L. (Eds.) *Soil carbon*. Advances in Agronomy, vol. 125. United States: Academic Press.

Bayramov, E., Schlager, P., Kada, M. et al. (2019). Quantitative assessment of climate change impacts onto predicted erosion risks and their spatial distribution within the landcover classes of the Southern Caucasus using GIS and remote sensing. *Mod. Earth Sys. and Env.*, V5

Benavidez, R., Bethanna, J., Deborah, M. et al. (2018). A review of the (Revised) Universal Soil Loss Equation ((R)USLE). *Hydrol. Earth Syst. Sci.*, 22.

Buchhorn, M., Smets, B., Bertels, L. et al. (2019). Copernicus Global Land Service: Land Cover 100m, epoch “year”. *Globe* (V2.0.2).

Chen, L., Qian, X. and Shi, Y. (2011). Critical Area Identification of Potential Soil Loss in a Typical Watershed of the Three Gorges Reservoir Region. *Water. Resour. Manage*. 25, 3445. DOI: 10.1007/s11269-011-9864-4.

David, W. P. (1988). Soil and Water Conservation Planning: Policy Issues and Recommendations. *J. Philipp. Dev.*, 15, 47–84.

Hengl, T, de Jesus, J.M., MacMillan, R.A. et al. (2014). SoilGrids1km – global soil information based on automated mapping. *PLoS One*, 9(8):e105992.

Hengl, T., de Jesus, M.J., Heuvelink, G.B.M. et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, 12(2):e0169748.

Hernandez, E.C., Henderson, A. and Oliver, D.P. (2012). Effects of changing land use in the Pagsanjan–Lumban catchment on suspended sediment loads to Laguna de Bay, Philippines, *Agric. Water Manag.*, 106, 8–16.

Kirchmeir, H. and Michael, H. (2016). *Remote Sensing Concepts on Erosion Control and Pasture Management Report. Integrated Erosion control measures in Ismayilli, Azerbaijan*. IBIS, GIZ, 11.

Kirchmeir, H. and Berger, V. (2019). *Development of land cover and erosion risk map based on remote sensing for Tusheti protected areas*. E.C.O. Institute of Ecology, GIZ.

Lin, B.S., Chen, C.K, Thomas, K. et al. (2019). Improvement of the K-Factor of USLE and Soil Erosion Estimation in Shihmen Reservoir Watershed. *Sustainability*, 11(2), 355. DOI: 10.3390/su11020355.

Morgan, R.P.C. (2005). Soil Erosion and Conservation, National Soil Resources Institute. Cranfield University, ch22

Ozsoy, G., Aksoy, E., Dirim, M.S. et al. (2012). Determination of soil erosion risk in the Mustafakemalpasa river basin, Turkey, using the revised universal soil loss equation, geographic information system, and remote sensing. *Environ. Manage.*, 50.

Panagos, P., Meusburger, K., Ballabio, C. et al. (2014). Soil erodibility in Europe: A high-resolution dataset based on LUCAS. *Sci. Total Environ.*, 5, 461–487.

Renard, K., Foster, G., Weesies, G. et al. (1997). *Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE)*. United States: Agricultural Research Services.

Yuanyuan, Y., Ruiying, Zh., Zhou, Sh. et al. (2018). Integrating multi-source data to improve water erosion mapping in Tibet, China. *CATENA*, 169, 31–45.

Wawer, R., Nowocien, E. and Podolski, B. (2005). Real and calculated K-USLE erodibility factor for selected Polish soils. *Polish J. Environ. Studies*, 14(5), 655–658.

Williams, R.J. and Renard, K.G. (1983). EPIC – a new method for assessing erosions effect on soil productivity. *J. Soil Water Conserv.*, 38, 381–383.

Wischmeier, W.H. and Mannering, J.V. (1969). Relation of Soil Properties to its Erodibility. Soil and Water Management and Conservation. *Soil Sci. Soc. Am. J.*, 15, 131–137. DOI: 10.2136/sssaj1969.03615995003300010035x.

Wischmeier, W.H. and Smith, D.D. (1978). *Predicting rainfall erosion losses*. United States: Agricultural Research Services.