

UNIVERSITY OF MISKOLC

Faculty of Earth and Environmental Sciences and Engineering

Institute of Geography and Geoinformatics



Mikoviny Sámuel Doctoral School of Earth Sciences

Head of the Doctoral School: Prof. Dr. Péter Szűcs

THESES BOOKLET

Spatiotemporal Soil Erosion Modelling in the Eritrean Highlands:

A Monthly RUSLE Simulation Framework Using Machine
Learning, GPM IMERG, and Sentinel-2 Data

Prepared by:

TSEGAY BEREKET MENGHIS

Scientific Supervisor: Prof. Dr. Endre Dobos

Miskolc, 2026

Part I: Summary of Research Aims

1.1. Background and problem statement

Soil erosion is one of the most critical land degradation processes in the Eritrean Highlands, threatening agricultural productivity, water resources, and long-term food security across a predominantly rain-fed agricultural landscape of approximately 10,000 km² (Nyssen et al., 2004; Stillhardt et al., 2002). The Kremti monsoon (June–September) delivers 70–80% of annual rainfall in two to four intense months, creating a strongly seasonal erosion regime in which the timing and spatial distribution of erosion are as important as the annual total. Despite this, assessments in the region have historically relied on the conventional annual RUSLE approach (Wischmeier & Smith, 1978; Renard et al., 1997), which applies a single annual R-factor and annual mean C-factor, collapsing the seasonal erosion cycle into one temporal snapshot and losing the within-year dynamics that drive actual erosion risk.

A further critical gap is the absence of a spatially continuous soil erodibility (K-factor) map for the Eritrean Highlands. Researchers have previously estimated erodibility from sparse point observations without machine-learning-based spatial prediction (Menghis et al., 2025), limiting the resolution and transferability of regional erosion assessments. No study had yet integrated monthly GPM IMERG rainfall erosivity (Huffman et al., 2020), monthly Sentinel-2 NDVI-derived vegetation cover, and machine-learning-predicted soil erodibility within a SCORPAN framework (McBratney et al.,

2003) into a spatially explicit, temporally resolved erosion assessment at regional scale for the Eritrean Highlands.

1.2. Research Objectives

Objective 1: To predict soil texture (sand, silt, clay) and organic matter spatially across the Eritrean Highlands using ensemble machine learning within the SCORPAN framework (McBratney et al., 2003), and to derive a spatially continuous K-factor map at 30 m resolution using the EPIC/Williams equation.

Objective 2: To develop a monthly RUSLE simulation framework integrating 12 monthly GPM IMERG rainfall erosivity rasters (Huffman et al., 2020) and 12 monthly Sentinel-2 NDVI-derived C-factor rasters (Van der Knijff et al., 2000) resolving the seasonal dynamics of soil erosion in the Eritrean Highlands for 2021.

Objective 3: To compare the monthly simulated annual soil loss (A_{annual}) with the conventional non-simulated approach ($A_{\text{non-sim}}$) through a pixel-by-pixel difference map ($\Delta A = A_{\text{non-sim}} - A_{\text{annual}}$), quantifying the spatial distribution, direction, and magnitude of divergence between the two frameworks.

Objective 4: To identify the critical erosion window, the seasonal migration of erosion hotspots, and the conservation planning implications of the simulation results across the 17 sub-zobas of the study area. Ultimately, to provide scientific information to the decision-makers.

1.3. Study Area

The study area encompasses the central and southern zobas of the Eritrean Highlands, covering approximately 10,000 km² across 17 administrative sub-zobas. Elevation ranges from 922 m to 3,018 m above mean sea level. Rain-fed agriculture, rangeland, and natural vegetation dominate the landscape across four topographic zones: the central plateau, south-western zone, eastern escarpment, and north-western zone. Three distinct rainfall systems govern the study area: the Kremti monsoon (June–September) as the primary erosive system delivering 70–80% of annual rainfall (350–650 mm); the Hamle minor rainy season (March) in the north-western zone; and Indian Ocean post-monsoon moisture (October–December) sustaining year-round erosion in the eastern escarpment (130–300 mm/yr). The study area has a long history of rain-fed cultivation, overgrazing, and deforestation, making it one of the most erosion-vulnerable landscapes in the Horn of Africa (Herweg & Ludi, 1999; FAO, 1994).

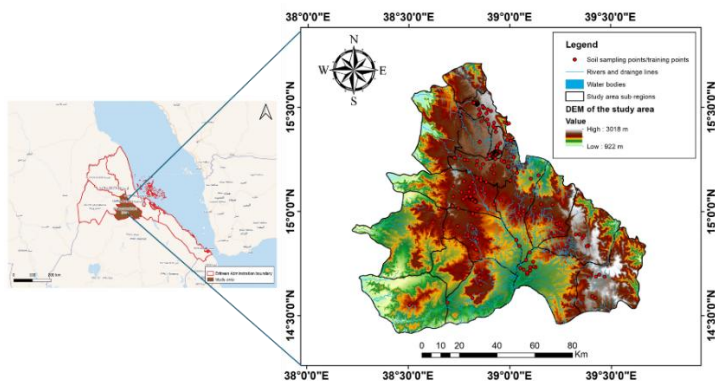


Figure 1. Location of the study area within Eritrea

PART II: Description of Investigation, Methods, and Data Sources.

2.1. Data Sources

Soil profiles: 302 MOA/NARI soil profiles (277 used in cross-validation) providing laboratory-measured sand (%), silt (%), clay (%), and organic matter (%) across the study area, collected by the Eritrean Ministry of Agriculture (MOA) and National Agricultural Research Institute (NARI).

GPM IMERG precipitation: GPM IMERG Final Run Version 07 monthly precipitation data for 2021 at approximately 10 km resolution, providing 12 monthly precipitation rasters for R-factor computation (Huffman et al., 2020).

Sentinel-2 NDVI: Sentinel-2 L2A surface reflectance imagery for 2021, cloud-composited into 12 monthly NDVI rasters at 30 m resolution for C-factor derivation (Rouse et al., 1974; Van der Knijff et al., 2000).

SRTM DEM: SRTM digital elevation model at 30 m resolution, hydrologically conditioned, used for LS-factor derivation (Moore & Burch, 1986; Desmet & Govers, 1996) and terrain covariates.

Environmental covariates: Sentinel-2 spectral indices (NDVI, BSI, CMI, NIR, SWIR1, SWIR2), Landsat 8/9 land surface temperature (LST; Jiménez-Muñoz et al., 2014) for four seasonal composites (Diek et al., 2017), and SRTM terrain derivatives (slope, aspect, TPI, curvature, flow accumulation, TRI; Conrad et al., 2015), all at 30 m resolution.

2.2. Machine Learning K-Factor Prediction (SCORPAN Framework)

Soil texture and organic matter were predicted spatially using five ensemble machine learning algorithms: Extra Trees, Random Forest, Gradient Boosted Trees (GBT), XGBoost, and LightGBM, trained on 277 soil profiles within the SCORPAN framework (McBratney et al., 2003) with 5-fold spatial cross-validation. Environmental covariates proxy the five primary soil-forming factors (climate, organisms, relief, parent material, time). SHAP analysis (Lundberg & Lee, 2017) identified process-consistent predictors. The best-performing model per property was selected on R^2 , RMSE, MAE, bias, and RPD. Predicted soil property surfaces were used as inputs to the EPIC/Williams equation (Williams, 1995) to produce the spatially continuous K-factor raster ($t\ ha\ h\ ha^{-1}\ MJ^{-1}\ mm^{-1}$). The resulting texture and erodibility surfaces represent regional gradients rather than exact point predictors, within the bounds defined by spatial cross-validation and explicit range compression (Wadoux et al., 2020).

2.3. Conceptual Framework and Temporal Simulation Approaches

Two parallel RUSLE computation frameworks were implemented. In the simulated framework, monthly GPM IMERG precipitation (P_m) is annualised ($P^*_m = P_m \times 12$) before entry into the Hurni (1985) erosivity equation, producing 12 monthly R-factor rasters ($R_sim_m = -8.12 + 0.562 \times P^*_m$). The RUSLE equation $A_sim_m = R_sim_m \times K \times LS \times C_m \times P$ is applied 12 times; each result is de-annualised ($A_actual_m = A_sim_m \div 12$); and the 12

contributions summed ($A_{\text{annual}} = \sum A_{\text{actual_m}}$). Mathematical validation confirms no erosivity bias. In the conventional framework, a single annual R-factor ($R_{\text{annual}} = -8.12 + 0.562 \times \sum P_m$) and annual mean C-factor (from mean annual NDVI) are applied in a single RUSLE computation, producing $A_{\text{non-sim}}$. K, LS, and P are identical in both frameworks.

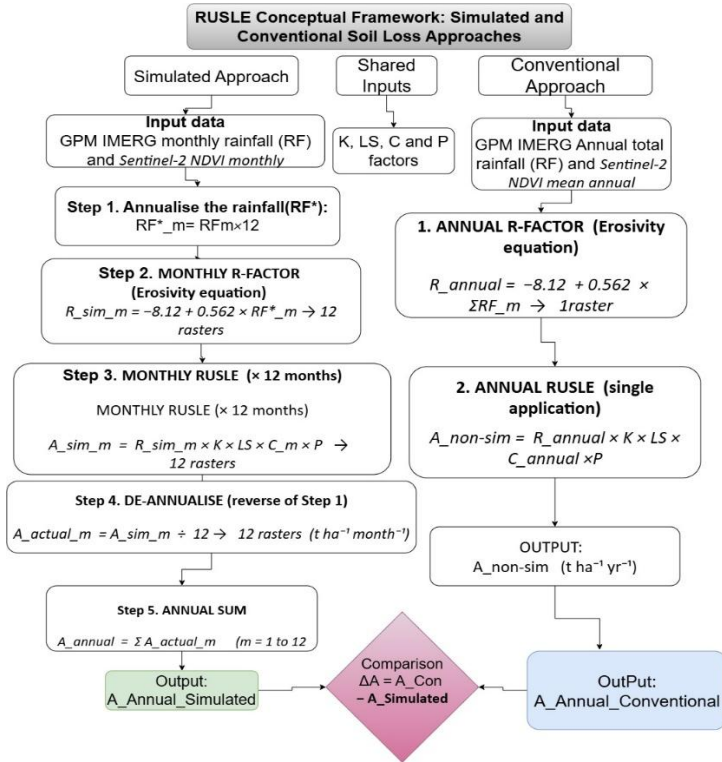


Figure 2. Conceptual Framework of the Simulated and Conventional Approaches

2.4. RUSLE Factors Derivation

R-factor: Hurni (1985) empirical equation calibrated for the East African Highlands: $R = -8.12 + 0.562 \times P_{\text{ann}}$. Applied monthly (simulated) and annually (conventional).

K-factor: EPIC/Williams equation (Williams, 1995) applied to ML-predicted sand, silt, clay, and OM surfaces. K-factor range:

LS-factor: Moore and Burch (1986) and Desmet and Govers (1996) algorithms applied to a hydrologically conditioned SRTM DEM.

C-factor: Van der Knijff et al. (2000) exponential equation: $C_m = \exp[-2 \times \text{NDVI}_m / (1 - \text{NDVI}_m)]$, $\alpha = 2$, $\beta = 1$. Applied monthly.

P-factor: $P = 1$ throughout the study area, reflecting the absence of spatially documented SWC practice data at regional scale (Herweg & Ludi, 1999; Menghis et al., 2025). All results represent potential soil loss.

2.5. Difference Map and Spatial Comparison

The difference map ($\Delta A = A_{\text{non-sim}} - A_{\text{annual}}$) was computed by pixel-by-pixel subtraction across the entire study area. The resulting spatial distribution was analysed to identify three outcome zones: the zone of agreement, the zone of conventional overestimation (positive ΔA), and the zone of simulation-higher underestimation (negative ΔA). Both biases are predictable from the seasonal co-variance of R and C (Vrieling et al., 2008; Nearing et al., 2005). All analyses were conducted in ArcGIS 10.4.1.

PART III. Summary of Scientific Results (Thesis)

The following new scientific findings/contributions (Theses) contribute to spatiotemporal soil erosion modelling in the Eritrean Highlands through RUSLE simulations, machine-learning predictions, and comparisons with conventional methods. They address temporal, spatial, and methodological aspects of erosion assessment in a data-scarce, seasonal environment and support soil and water conservation planning.

Thesis 1: The development of a Machine Learning-based methodology for soil particle size estimation of the Eritrean highlands.

I have developed a soil texture database for the Eritrean Highlands using ML. The soil texture is predictable based on soil-forming factors derived from spectral and thermal signatures and terrain features. Three models were trained on 277 samples. SHAP analysis indicates the model uses relevant covariates, and the resulting terrain-organised soil texture maps align with geomorphological understanding, reflecting underlying processes. Spatial cross-validation and range compression ensure the maps represent regional gradients accurately, providing a reliable basis for soil erodibility assessment rather than precise point predictions.

Thesis 2: The development of a Machine Learning-based methodology for OM estimation of the Eritrean highlands.

I developed an organic matter database for the Eritrean Highlands using machine learning trained on 277 field-sampled soil profiles within the SCORPAN framework, demonstrating that organic matter is a predictable expression of soil-forming

factors encodable in spectral, thermal, and terrain covariates derived from Sentinel-2 imagery and the SRTM digital elevation model. SHAP attribution confirms that the model relies on process-consistent predictors, and the predicted organic matter surface exhibits terrain-organised structure concordant with independent geomorphological reasoning, establishing the recovered spatial pattern as mechanistic rather than a statistical artefact. The predicted organic matter field provides a faithful representation of regional gradients and a defensible spatial input to the EPIC/Williams K-factor equation, within the bounds defined by spatial cross-validation and explicit range compression, rather than an exact point predictor.

Thesis 3: The integration of the GPM IMERG dataset into the RUSLE system as a Physically Consistent Erosivity Dataset:

GPM IMERG Final Run Version 07 offers a consistent monthly rainfall erosivity dataset suitable for regional RUSLE simulations in data-scarce highlands. In the Eritrean highlands, where station coverage is limited, GPM IMERG provides the only continuous spatial precipitation data needed for erosion modelling. Its R-factor range matches literature values for Ethiopian and Blue Nile highlands, confirming regional applicability. This study is the first to apply GPM IMERG Version 07 to monthly RUSLE R-factor calculation in Eritrea, with methodology transferable to other highland regions with sparse ground data.

Thesis 4: I have found that despite the high topographic variability of the Eritrean highlands, the most important factors determining the annual soil loss are the R and C.

The highest annual soil loss in the Eritrean highlands is driven more by the duration of the monsoon erosion season than by topography or peak. Identifying this requires monthly simulation of cumulative erosion. The primary erosion factors- steep terrain (LS), erodible soils (K), and peak rainfall (R)- do not always align spatially. The area with the highest LS has low K and R; the highest K area has moderate LS and rainfall; and the highest R area has neither extreme LS nor K. The simulated soil loss map shows that the highest soil loss occurs in south-western zones during May-September, due to sustained erosion over five months from pre-monsoon and monsoon rainfall. This demonstrates that maximum erosion results from combined monthly factors, not single RUSLE factors, unlike simplified priority maps that misjudge complex landscapes. Conservation efforts should focus on multi-season strategies, especially in south-western zones, rather than peak monsoon periods alone.

Thesis 5: September Vegetation/erosion/rainfall paradox was determined in the Eastern Escarpment of the Eritrean highlands.

In September, the eastern escarpment sub-regions (Senafe, Tsorona, Adi Keyih) retain moderate-to-high C-factor values and sustained soil loss, while the central plateau achieves near-zero C-factor and minimal erosion, a spatial paradox in which the zone of highest topographic erosion potential simultaneously receives the least monsoon moisture, limiting canopy development and preventing the vegetation protection

that occurs naturally elsewhere. This pattern cannot be detected from annual soil-loss maps and requires year-round, vegetation-based SWC interventions in the eastern escarpment rather than strategies focused on the peak monsoon period alone.

Thesis 6: I found that the annual RUSLE equation consistently underestimates the actual erosion where topographic conditions are favourable for erosion

I found that the yearly average is lower when using the conventional RUSLE, with annual values estimating much lower values than the simulated ones in regions where topographic conditions favour erosion. This shows that annual dynamics of vegetation cover and rainfall strongly affect actual erosion and must be considered in the Eritrean highlands. This provides crucial information for soil conservation activities.

Part IV. Utilization and Possibilities of utilization

i. Direct Practical Application

The monthly simulated soil loss maps and A_{annual} provide the first spatially explicit, temporally resolved erosion baseline for the Eritrean Highlands at 30 m resolution. These outputs directly support: (1) national SWC planning by the Eritrean Ministry of Agriculture at sub-zoba and monthly resolution, enabling month-specific prioritisation of interventions; (2) identification of June–July as the critical erosion window, the primary target period for agronomic conservation measures including early-cover cropping and reduced tillage; (3) identification of the eastern escarpment as requiring year-round

SWC strategies rather than monsoon-season-only approaches; and (4) benchmarking against the EU Soil Monitoring Law Directive 2025/2360 sustainable erosion threshold of $2 \text{ t ha}^{-1} \text{ yr}^{-1}$ (Ogbazghi, 1999; Tesfay et al., 2018).

ii. Indirect Practical Application

The monthly RUSLE simulation framework is directly extensible to: (1) climate scenario analysis using downscaled CMIP6 precipitation projections (IPCC, 2021) as replacements for GPM IMERG inputs; (2) multi-year erosion monitoring using the full GPM IMERG record available from 2000; and (3) actual soil loss estimation by replacing $P = 1$ with a spatially variable P-factor map derived from a planned regional SWC field survey (Herweg & Ludi, 1999; Menghis et al., 2025), the single highest-priority methodological improvement identified.

iii. Proportion of the discipline

This dissertation contributes to the scientific discipline by: (1) establishing the first ML-based spatially continuous K-factor map for the Eritrean Highlands, providing a replicable workflow for data-scarce East African Highland environments; (2) providing the mathematical proof that the annualisation/de-annualisation procedure introduces no erosivity bias, offering a methodological foundation for monthly RUSLE simulation in any strongly seasonal environment; (3) producing the first pixel-level quantitative comparison of monthly simulated and conventional annual RUSLE at regional scale in East Africa; and (4) documenting the September vegetation paradox as a counter-intuitive coupled topography–vegetation–rainfall phenomenon unique to the eastern escarpment of the Eritrean Highlands.

PART V: List of the Publications: The most relevant and significant publications related to the dissertation

1. Menghis, T. B., Rajhi, M., Zdruli, P., & Dobos, E. (2026). Sensitivity Analysis of RUSLE Parameters: A GIS-Based Approach for Soil Erosion Control at Mai-Nefhi Watershed, Eritrea. *Spectrum Science Journal*, 3(1), 8–22. <https://doi.org/10.21608/SASJ.2026.460668.1012>
2. Menghis, T. B., Zdruli, P., & Dobos, E. (2025). A GIS-Based Approach to Soil Erosion Risk Assessment Using RUSLE: The Case of the Mai Nefhi Watershed, Barka River Basin, Eritrea. *Earth (Switzerland)*, 6(2), 1–14. <https://doi.org/10.3390/earth6020058>
3. Menghis, T.B. and Dobos, E. (2023). Characterisation of Land Use And Land Cover Changes Using Historical Satellite Datasets and GIS: A Case Study Of Mai-Nefhi Watershed, Eritrea. XXVI. Spring Wind Conference 2023 - Study Volume II.
4. Menghis, T.B., & Dobos, E. Effect of soil erosion on organic matter content: A case study of Xiashiziwan watershed, China. 14th Geospatial Information Conference and Exhibition, Debrecen, Book excerpt (Abstract / Excerpt) | ISBN: 9789636150846, 2023
5. Menghis, T.B., & Dobos, E. Soil Erosion and Severity class analysis based on the RUSLE model. A case study of Mai/Nefhi watershed, Eritrea. XV Geospatial Conference & Exhibition, University of Debrecen, Book excerpt (Abstract / Excerpt) |ISBN: 9789634906193, 2024

6. DOBOS, E., SZABÓ, E., CSENKI, S., **MENGHIS, T. B.**, MOLNÁR, F., DOBAI, A., & RAJHI, M. (2023). DIGITAL SOIL MAPPING APPROACH TO ESTIMATE SOIL PLASTICITY USING GEOREFERENCED TECHNICAL DATA OF TRACTORS. *Geosciences and Engineering*, 11(1), 56–67. <https://doi.org/10.33030/GEOSCIENCES.2023.01.005>
7. Mallu, T. S., Okbagabir, S. G., Negassi, B. T., Tesfay, T., Negash, T. T., **Menghis, T. B.**, Medhn, T. A., Teklay, S. G., Gebremeskal, Y. H., Tewolde, R. H., Gebremariam, G. K., Tecleab, A. A., Tewelde, K. G., Brhane, D., Asghede, K. M., Adem, A. H., & Solomon, M. T. (2025). The Agronomic, Nutritional, Environmental, and Health Benefits of Teff (*Eragrostis tef*) as an Underutilized Gluten-Free Superfood and Feed Source. *International Journal of Agronomy*, 2025(1), 6128837. <https://doi.org/10.1155/IOA/6128837>
8. Mallu, T. S., Okbagabir, S. G., Negassi, B. T., Tesfay, T., Negash, T. T., **Menghis, T. B.**, Medhn, T. A., Teklay, S. G., Gebremeskal, Y. H., Tewolde, R. H., Gebremariam, G. K., Tecleab, A. A., Tewelde, K. G., Brhane, D., Asghede, K. M., Adem, A. H., & Solomon, M. T. (2026). A comprehensive review on finger millet (*Eleusine coracana* (L.) Gaertn): A gluten-free superfood and feed. *Journal of Agriculture and Food Research*, 25, 102528. <https://doi.org/10.1016/J.JAFR.2025.102528>

Reference

- Conrad, O., et al. (2015). *System for Automated Geoscientific Analyses (SAGA) v.2.1.4*. Geoscientific Model Development, 8(7), 1991–2007
- Desmet, P.J.J. & Govers, G. (1996). *A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units*. Journal of Soil and Water Conservation, 51(5), 427–433
- Diek, S., et al. (2017). *Barest pixel composite for agricultural areas using Landsat time series*. Remote Sensing, 9(12), 1245
- FAO (1994). *Land and Water Resources Survey, Eritrea*. Food and Agriculture Organization of the United Nations, Rome
- Hengl, T., et al. (2017). *SoilGrids250m: Global gridded soil information based on machine learning*. PLOS ONE, 12(2), e0169748
- Hengsdijk, H., Meijerink, G.W. & Mosugu, M.E. (2005). *Modelling the effect of three soil and water conservation practices in Tigray, Ethiopia*. Agriculture, Ecosystems & Environment, 105(1–2), 29–40
- Herweg, K. & Ludi, E. (1999). *The performance of selected soil and water conservation measures, case studies from Ethiopia and Eritrea*. Catena, 36(1–2), 99–114
- Huffman, G.J., et al. (2020). *Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG)*. In: Levizzani, V. et al. (Eds), Satellite Precipitation Measurement. Springer. DOI: 10.1007/978-3-030-24568-9_19

Hurni, H. (1985). *Erosion-productivity-conservation systems in Ethiopia*. Proceedings of the 4th International Conference on Soil Conservation, Maracay, Venezuela, 3–4

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report*. Cambridge University Press

Jiménez-Muñoz, J.C., et al. (2014). *Land surface temperature retrieval methods from Landsat-8 thermal infrared sensor data*. IEEE Geoscience and Remote Sensing Letters, 11(10), 1840–1843

Lundberg, S.M. & Lee, S.-I. (2017). *A unified approach to interpreting model predictions*. Advances in Neural Information Processing Systems, 30

McBratney, A.B., Mendonça Santos, M.L. & Minasny, B. (2003). *On digital soil mapping*. Geoderma, 117(1–2), 3–52

Menghis, T. B., Zdruli, P., & Dobos, E. (2025). A GIS-Based Approach to Soil Erosion Risk Assessment Using RUSLE: The Case of the Mai Nefhi Watershed, Barka River Basin, Eritrea. *Earth (Switzerland)*, 6(2), 1–14. <https://doi.org/10.3390/earth6020058>

Nyssen, J., et al. (2004). *Human impact on the environment in the Ethiopian and Eritrean highlands, a state of the art*. Earth-Science Reviews, 64(3–4), 273–320

Ogbazghi, W. (1999). *Integrated soil management for sustainable agriculture and food security in Eritrea*. In: Proceedings of the Workshop on Integrated Soil Management for Sustainable Agriculture and Food Security in Southern and Eastern Africa. FAO, Rome, pp. 193–202

Renard, K.G., Foster, G.R., Weesies, G.A., McCool, D.K. & Yoder, D.C. (1997). *Predicting Soil Erosion by Water: A Guide*

to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE). USDA Agriculture Handbook No. 703, Washington D.C.

Rouse, J.W., et al. (1974). *Monitoring vegetation systems in the Great Plains with ERTS*. Proceedings of the 3rd ERTS Symposium, NASA SP-351, 309–317

Stillhardt, B., Herweg, K. & Hurni, H. (2002). *Long-term monitoring of soil erosion and soil conservation in Eritrea, 1994–2000*. Centre for Development and Environment (CDE), University of Bern Working Paper

Tesfay, T., Ogbazghi, W., Singh, B. & Tsegai, T. (2018). *Factors influencing soil and water conservation adoption in Basher, Gheshnashm and Shmangus Laelai, Eritrea*. IRA-International Journal of Applied Sciences, 12(2), 7

Van der Knijff, J.M., Jones, R.J.A. & Montanarella, L. (2000). *Soil erosion risk assessment in Europe*. EUR 19044 EN. European Commission, Joint Research Centre, Ispra

Vrieling, A., et al. (2008). *Spatially detailed rainfall erosivity estimates for Africa*. Hydrology and Earth System Sciences, 14, 2635–2647

Wadoux, A., Minasny, B. & McBratney, A.B. (2020). *Machine learning for digital soil mapping: Applications, challenges and suggested solutions*. Earth-Science Reviews, 210, 103359

Williams, J.R. (1995). *The EPIC model*. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, CO, pp. 909–1000

Wischmeier, W.H. & Smith, D.D. (1978). *Predicting Rainfall Erosion Losses, A Guide to Conservation Planning*. USDA Agriculture Handbook No. 537, Washington, D.C.