

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

PhD DISSERTATION

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

Bachelor of Science (BSc.) in Agricultural Engineering and Master of Science (MSc.) in
Soil-Water Engineering

Scientific Supervisor:

Prof. Dr. Endre Dobos

Miskolc, 2026

HUNGARY

Supervisor's Recommendation

I, [Prof. Dr. Endre Dobos](#), supervisor of PhD candidate [Tsegay Bereket Menghis](#), hereby certify that I am thoroughly familiar with the doctoral dissertation entitled “Spatiotemporal Soil Erosion Modelling in the Eritrean Highlands: A Monthly RUSLE Simulation Framework Using Machine Learning, GPM IMERG, and Sentinel-2 Data”. I have closely followed the candidate’s research activities throughout the doctoral program, and I consider the dissertation fully suitable for public defense and evaluation within the PhD procedure.

The dissertation addresses a highly relevant and internationally important environmental problem: the spatial and temporal modelling of soil erosion processes in the Eritrean Highlands. The scientific novelty of the work lies in the successful integration of monthly-scale RUSLE modelling, satellite-derived precipitation data (GPM IMERG), Sentinel-2 vegetation information, and machine learning-based soil property prediction into a coherent regional-scale erosion assessment framework. To the best of my knowledge, this is the first study to implement and compare such a temporally explicit erosion-modelling approach for the Eritrean Highlands.

The results of the dissertation are significant not only from a scientific perspective but also from a practical one. The findings provide valuable information for soil and water conservation planning in Eritrea and contribute to more efficient land management strategies in data-scarce environments. The methodology developed by the candidate has considerable potential for transferability to other semi-arid and mountainous regions where environmental monitoring data are limited, but sustainable land management is urgently needed.

Throughout the doctoral research, the candidate worked independently, persistently, and with a high level of scientific rigor. The dissertation is methodologically sound, logically structured, and supported by comprehensive analyses and well-substantiated conclusions. The candidate has clearly demonstrated the ability to conduct independent scientific research and to contribute original knowledge to the fields of soil science, environmental modelling, and geoinformatics.

Based on the scientific quality of the work, the originality of the research, and the candidate’s demonstrated research competence, I strongly recommend the dissertation for acceptance and public defense, and I fully support the continuation of the doctoral procedure.

Miskolc, 2026

[Prof. Dr. Endre Dobos](#)
Supervisor

Acknowledgement

First and foremost, I would like to express my deepest gratitude to my supervisor, [Prof. Dr. Endre Dobos](#), for his continuous guidance, support, and encouragement throughout my doctoral studies. I am especially grateful for his kindness, patience, understanding, and positive attitude, as well as for his extensive knowledge and thoughtful, knowledge-oriented guidance, which have greatly contributed to the development of this dissertation and my academic growth. His constructive feedback, professional advice, moral support, and confidence in my abilities have been invaluable throughout this journey. I sincerely appreciate his professionalism, integrity, generosity, and positive approach, and I feel privileged to have had him as my supervisor.

I would like to express my sincere appreciation to the University of Miskolc, particularly the professors whose courses I attended, the Faculty, the Dean's Office, and the Institute of Geography and Geoinformatics, for their academic and administrative support throughout my doctoral studies. I am grateful to the professors for sharing their knowledge and expertise, which have enriched my academic and professional development. I also sincerely appreciate the Dean and the staff of the Dean's Office for their assistance and cooperation. My gratitude extends to the Head of the Institute, Dr. Vágó János, for his leadership, support, and commitment to fostering a positive academic and research environment, as well as to all academic and administrative staff of the Institute and the Institute's Office for their kindness, professionalism, and continuous assistance.

I would also like to express my sincere appreciation to all the staff members of the Institute, particularly my colleagues and friends with whom I have spent a good time, Diána Bertóti, Éva Oroszné Gyopár, Mohammed Rajihi, Károly Zoltán Kovács, András Dobos, Ferenc Molnár, András Gömöri and Dániel Farkas, who supported me throughout my administrative and academic doctoral journey. I am grateful for their collaboration, valuable discussions, knowledge sharing, constructive feedback, encouragement, and assistance in various administrative, academic and research activities. Their friendship, cooperation, and support have made this journey more meaningful and enjoyable.

I would like to express my heartfelt gratitude to the Stipendium Hungaricum Scholarship Programme and the Government of Eritrea for their financial and institutional support, which made my doctoral studies at the University of Miskolc possible. I would also like to acknowledge the National Agricultural Research Institute (NARI) of the Eritrean Ministry of Agriculture for providing data such as the soil profile data used in the soil erosion modelling.

Finally, I would like to express my heartfelt gratitude to my wife and children for their unconditional love, patience, understanding, and support throughout my doctoral journey. I am also deeply grateful to my father and all my family members for their constant encouragement, prayers, love, and support. Their belief in me has been a great source of strength and motivation throughout this journey.

Table of Contents

Supervisor's Recommendation	i
Acknowledgement	ii
Table of Contents	iii
List of Figures:.....	vii
List of Tables.....	viii
List of Abbreviations.....	ix
Abstract	x
1. INTRODUCTION	1
1.1. Background and Context.....	1
1.2. Problem Statement	2
1.3. Research Gap	3
1.4. Aim and Specific Objectives	4
1.5. Scope and Limitations.....	4
2. LITERATURE REVIEW	6
2.1. Soil Erosion Processes and Driving Factors	6
2.2. Soil Erosion Modelling Approaches and the Justification for RUSLE.....	7
2.3. The RUSLE Factors	8
2.4. Temporal and Seasonal Dimensions of Erosion Modelling.....	10
2.5. Satellite Rainfall Products for Erosivity in Data-Scarce Regions	11
2.6. Remote Sensing of Vegetation Dynamics for the C-Factor	12
2.7. Machine Learning and Digital Soil Mapping for Soil-Property Prediction.....	12
2.8. Integration of GIS, Remote Sensing and Machine Learning in Erosion Modelling	14
2.9. Soil Erosion Assessment in the Eritrean Highlands: The Mai Nefhi Pilot Watershed Study	15
2.9.1. Conventional RUSLE Assessment.....	15
2.9.2. Sensitivity Analysis of the Same RUSLE Framework	15
2.9.3. Connection to the present Dissertation	16
2.10. Synthesis and Identification of the Research Gap	16
3. METHODS AND MATERIALS	17
3.1. Overall Methodological Framework.....	17
3.2. Study Area.....	18
3.2.1. Location and Physiography.....	18

3.2.2.	Geology and Soils	20
3.2.3.	Climate and Seasonal Rainfall Regime.....	21
3.2.4.	Vegetation and Land use/Land cover	21
3.2.5.	Topography and Drainage.....	22
3.2.6.	Agricultural Systems and Erosion Susceptibility	22
3.2.7.	Rationale for study area selection	23
3.3.	Datasets	23
3.3.1.	GPM IMERG (IMERG-FR V07) Rainfall	23
3.3.2.	Environmental Covariates.....	24
3.3.3.	Soil Sampling and Laboratory Analysis	25
3.4.	Methods.....	27
3.4.1.	Machine-learning prediction of soil properties and K-factor	27
3.4.1.1.	Predictor Variables (Terrain Attributes and Environmental Covariates)	27
3.4.1.2.	Machine-Learning Regression Models, Training, and Tuning	28
3.4.1.3.	Model Validation Strategy	29
3.4.1.4.	Computational Environment and Software Implementation	30
3.4.1.5.	Reference Spatial Grid, Data Integration and Reproducibility.....	31
3.4.2.	RUSLE Model factors.....	32
3.4.2.1.	Rainfall Erosivity factor (R-factor) - GPM IMERG Monthly Simulation	32
3.4.2.1.1.	R- Factor Equation	32
3.4.2.1.2.	Conventional (Annual Summation) Benchmark Approach.....	33
3.4.2.1.3.	Monthly Simulated Approach	33
3.4.2.2.	K-factor.....	34
3.4.2.2.1.	Selection of the EPIC/Wiliams Formulation.....	34
3.4.2.2.2.	Mathematical Formulation	35
3.4.2.3.	Slope length and steepness Factor (LS-factor)	36
3.4.2.4.	Cover management factor (C-factor).....	37
3.4.2.5.	Support Practice Factor (P)	37
3.4.3.	Conceptual Framework and Temporal Simulation Strategy	38
	The Annualisation and De-Annualisation Procedure.....	38
4.	RESULTS AND DISCUSSION.....	41
4.1.	Machine-Learning: Soil-Property Prediction.....	41
4.1.1.	Environmental Covariates.....	41
4.1.2.	Soil Texture prediction using ML	46

4.1.2.1.	Model Performance and Feature Importance	46
4.1.2.2.	Feature Importance (SHAP Analysis).....	47
4.1.2.3.	Predicted Soil texture maps (Sand, silt& clay).....	52
4.1.2.4.	Prediction Reliability and Residual Behaviour	53
4.1.3.	OM Prediction using ML	56
4.1.3.1.	Cross-validation performance of the six Algorithms.....	56
4.1.3.2.	Best-Performing Model Selection	59
4.1.3.3.	Feature Importance (SHAP Analysis).....	59
4.1.3.4.	Spatial Distribution of Predicted Organic Matter	60
4.1.4.	Machine Learning Improvement Over Conventional interpolation and uncertainty.....	62
4.2.	RUSLE factor results	63
4.2.1.	Rainfall Erosivity Factor (R-factor): Monthly variability vs Annual Distribution	63
4.2.1.1.	GPM IMERG Precipitation	63
4.2.1.2.	Monthly Simulated R-Factor Maps: Seasonal Dynamics.....	65
4.2.1.3.	Seasonal Summary: Monthly Ranking and the Dominance of the Summer/Kremti season	68
4.2.1.4.	Annual Non_Simulated R-Factor: Spatial Distribution and Contrast with Monthly Simulation.....	68
4.2.2.	Derived K-Factor Map.....	71
4.2.3.	Cover Management Factor (C-factor).....	73
4.2.3.1.	Monthly and mean NDVI.....	73
4.2.3.2.	Monthly C-Factor Maps: Seasonal Dynamics.....	75
4.2.3.3.	Mean Annual C-Factor: Spatial Distribution and Temporal Information Loss	79
4.2.4.	LS-factor and spatial controls	80
4.3.	Simulated Soil Erosion (spatiotemporal).....	82
4.3.1.	Monthly soil-loss	82
4.3.2.	Annual simulated soil loss (A_annual)	87
4.3.2.1.	Overall Spatial Distribution.....	88
4.3.2.2.	Data Gaps and Their Spatial Implications	89
4.3.2.3.	The Annual Maps as an Integration of Temporal Dynamics	89
4.4.	Conventional (Non-Simulated) Annual soil erosion.....	90
4.4.1.	Overall Spatial Distribution	90
4.4.2.	Spatial Pattern of Each Class	91

4.4.3.	The Role of Annual R and Mean Annual C in shaping the Non-Simulated Distribution	92
4.5.	Comparison of the two approaches.....	92
4.5.1.	Quantitative differences	93
4.5.2.	Spatial patterns of Over-estimation (Positive ΔA)	93
4.5.3.	Spatial patterns of Under-estimation (Negative ΔA).....	94
4.5.4.	Comparison of Conventional vs. Simulated Soil Loss	95
4.6.	General Interpretation of the Results and the Applied Methodology	97
4.6.1.	Significance of Temporal Variability in Erosion Modelling	97
4.6.2.	Why the Annual Approaches Biases Estimates	97
4.6.3.	Role of Machine Learning in Overcoming Data Scarcity	98
4.6.4.	Spatial Heterogeneity and Erosion-Driver Interactions	98
4.6.5.	Comparison with Previous Studies	99
4.6.6.	Strengths, Assumptions, and Limitations	100
4.6.7.	Implications for Soil Conservation and Land-Management Prioritisation	101
5.	CONCLUSIONS AND RECOMMENDATIONS	102
5.1.	Summary of Key Findings Mapped to Study objectives	102
5.2.	Practical and Policy Recommendations.....	103
5.3.	Recommendations for Future Research	104
	PhD Theses (Scientific Contributions)	105
	References.....	107
	APPENDIX.....	A

List of Figures:

Figure 1. Overall methodological framework flow diagram of the study	17
Figure 2. Geographical Location of the study area within Eritrea, showing the boundary of the Eritrean Highlands study area, the administrative regions (zobas) and sub-regions, and the DEM.	20
Figure 3. Soil Properties point data along with the sub-regions map.....	26
Figure 4. Conceptual Framework flow diagram of the Simulated and Conventional Approaches	40
Figure 5. Environmental Covariate: Red-edge (RE), NIR & shortwave-infrared (SWIR)	42
Figure 6. Seasonal BSI images for winter, spring, summer, and autumn, highlighting bare-soil exposure.	43
Figure 7. Seasonal median images of LST, °C: winter, spring, summer, and autumn	44
Figure 8. Terrain derivative: slope, aspect, surface curvature, relief intensity, and topographic position index (TPI). Red dots denote the 302 field soil-sampling locations	45
Figure 9. Mean Absolute SHAP feature importance for the best model of each fraction (left)& SHAP beeswarm plot for the clay model (XGB). Each point is one sample; horizontal position gives the SHAP value (impact on predicted clay) and colour encodes the feature.	49
Figure 10. Mean Absolute SHAP feature importance for the best model of each fraction (left)& SHAP beeswarm plot for the silt model (ET). Each point is one sample; horizontal position gives the SHAP value (impact on predicted clay), and colour encodes the feature..	50
Figure 11. Mean Absolute SHAP feature importance for the best model of each fraction (left)& SHAP beeswarm plot for the silt model (GBT). Each point is one sample; horizontal position gives the SHAP value (impact on predicted clay) and colour encodes the feature.	51
Figure 12. Predicted soil particle size maps: clay (left-up), silt (right-up), and sand (bottom) content.	53
Figure 13. Observed versus predicted values for the best-tuned model per fraction (silt–ET, sand–GBT, clay–XGB; five-fold CV, n = 277). Points are coloured by kernel density; the dashed line is 1:1 and the solid line the regression fit (slope annotated). MAE, RPD	54
Figure 14. Residual diagnostics for the best-tuned model per fraction (rows: silt–ET, sand–GBT, clay–XGB): residuals versus predicted with $\pm 2\sigma$ bounds and LOESS trend (left); residual distribution with normal fit and Shapiro–Wilk p-value (centre); and normal Q–Q p	55
Figure 15. Observed versus predicted OM for all six models under random cross-validation. Each panel reports R^2 , RMSE and RPD.....	57
Figure 16. Observed versus predicted OM for the GBT model under random 5-fold cross-validation. Points are coloured by local density; the dashed line is the 1:1 identity and the solid red line is the ordinary least-squares fit of predicted on observed.	58
Figure 17. Residual diagnostics for the GBT model: residuals against predicted values with a running trend line (left), the residual distribution with a fitted normal density (centre), and a normal quantile–quantile plot with the Shapiro–Wilk test result (right).....	59
Figure 18. SHAP beeswarm and mean SHAP importance for OM.....	60

Figure 19. Predicted topsoil organic-matter content (%) map	61
Figure 20. IMERG GPM 12-month rainfall maps for the year 2021.	64
Figure 21. IMERG GPM Annual Rainfall map for the year 2021.	65
Figure 22. Monthly Simulated R-Factor map for the year 2021.	70
Figure 23. Non-simulated R-Factor map for the year 2021.	71
Figure 24. Derived soil erodibility (K) factor map for the study area at 30 m resolution, in SI units $t\ ha\ h\ ha^{-1}\ MJ^{-1}\ mm^{-1}$	72
Figure 25. All 12-month NDVI maps for the year 2021.	74
Figure 26. Mean NDVI map for the year 2021	75
Figure 27. All 12-Months' C-factor map for 2021.	78
Figure 28. Annual mean C-Factor map for the year 2021	79
Figure 29. Slope Length and Steepness Factor (LS-Factor) map.....	81
Figure 30. All 12-Months' Soil Loss ($t\ ha^{-1}\ month^{-1}$) maps for the year 2021	86
Figure 31. Monthly dynamics of RUSLE input parameters and simulated soil loss across the Eritrean Highlands study area, 2021.	87
Figure 32. Annual simulated soil loss ($t\ ha^{-1}\ yr^{-1}$) map for the year 2021	90
Figure 33. Conventional (Non-Simulated) Annual soil erosion ($t\ ha^{-1}\ yr^{-1}$) map	92
Figure 34. Subregion-level comparison of mean annual soil loss ($t\ ha^{-1}\ yr^{-1}$) between the conventional and simulated RUSLE approaches across the 17 administrative subregions of the Eritrean Highlands, sorted by Conv – Sim difference.	96
Figure 35. Difference between conventional and simulated soil loss ($t\ ha^{-1}\ yr^{-1}$) map	96

List of Tables

Table 1. Sub-zoba Geographic Reference: Study Area Administrative Divisions and Geographic Location	19
Table 2. Summary of the primary datasets used in the study.....	23
Table 3. Descriptive statistics of measured soil texture fractions ($n = 277$).	27
Table 4. Environmental covariates used as predictor variables for soil-property prediction.	28
Table 5. Software and computational environment used in the study.	31
Table 6. Random five-fold cross-validation performance of the five algorithms for each texture fraction ($n = 277$). Bold rows denote the best-performing model per fraction.	46
Table 7. Random 5-fold, $n=277$: Best model per fraction	47
Table 8. Cross-validated performance of the six models, sorted best to worst.	56

List of Abbreviations

Abbreviation	Full Form
RUSLE	Revised Universal Soil Loss Equation
NDVI	Normalised Difference Vegetation Index
GPM	Global Precipitation Measurement
IMERG	Integrated Multi-satellite Retrievals for GPM
DEM	Digital Elevation Model
ML	Machine Learning
SHAP	Shapley Additive Explanations
SCORPAN	Soil, Climate, Organisms, Relief, Parent material, Age, Nearby soils
GIS	Geographic Information System
OM	Organic Matter
SWC	Soil and Water Conservation
OAT	One-at-a-time (sensitivity analysis)
CV	Cross-validation
RPD	Ratio of Performance to Deviation
SDR	Sediment Delivery Ratio
MOA	Ministry of Agriculture (Eritrea)
NARI	National Agricultural Research Institute
WRB	World Reference Base
SRTM	Shuttle Radar Topography Mission
SI	Sensitivity Index
LST	Land Surface Temperature
BSI	Bare Soil Index
CMI	Clay Mineral Index
ET	Extra Trees
RF	Random Forest
GBT	Gradient Boosted Trees

Abstract

Soil erosion represents one of the most pervasive forms of land degradation in the Eritrean Highlands, threatening agricultural productivity, water resource sustainability, and ecosystem stability across a region characterised by complex terrain, highly seasonal rainfall, and increasingly degraded vegetation cover. Despite its severity, erosion risk in this region has predominantly been assessed using conventional static applications of the Revised Universal Soil Loss Equation (RUSLE), in which the cover-management factor (C) and rainfall erosivity factor (R) are treated as temporally invariant annual averages, an assumption that obscures the dynamic interplay between monsoon rainfall intensity and vegetation phenology that governs actual soil detachment.

This dissertation develops and evaluates a monthly RUSLE simulation framework for the Eritrean Highlands that integrates Global Precipitation Measurement Integrated Multi-satellite Retrievals (GPM IMERG) data, Sentinel-2 multispectral time series, field-based soil sampling data provided by the Ministry of Agriculture, National Agricultural Research Institute (MOA-NARI) of Eritrea, and machine learning techniques to resolve the spatiotemporal dynamics of soil erosion at monthly intervals.

Monthly rainfall erosivity was derived from GPM IMERG precipitation using a linear erosivity relationship; the monthly cover-management factor was computed from NDVI time series extracted from Sentinel-2 imagery, the erodibility factor (K) was estimated through a SCORPAN-constrained function integrating machine-learning-predicted soil properties trained on the MOA-NARI soil sampling dataset, and the topographic factor (LS) was derived using slope-length and steepness algorithms applied to a high-resolution digital elevation model. Monthly simulated soil loss was computed across the full study area and aggregated to annual totals for comparison with the conventional non-simulated approach.

Results demonstrate that soil erosion in the Eritrean Highlands is strongly concentrated within the Kremti summer rainy season (June–September), with July accounting for the single largest monthly erosion pulse due to a critical phenological lag: peak rainfall erosivity coincides with an incomplete vegetative canopy, while maximum vegetation cover is not achieved until September, after the erosivity peak has subsided. Comparison of the two approaches across the 17 administrative subregions reveals that the simulated framework consistently yields higher annual soil loss estimates in the northern and central highland subregions, where degraded vegetation and steep terrain amplify the phenological lag effect, while the conventional approach produces marginally higher estimates in the southwestern subregions, where earlier-establishing vegetation and transitional terrain reduce the dynamic erosion signal.

These findings demonstrate that temporal resolution in RUSLE parameterisation is not a neutral refinement but a substantive methodological advance that exposes a class of erosion risk systematically underestimated by static approaches, with direct implications for the spatial prioritisation of conservation and land management interventions in the Eritrean Highlands.

1. INTRODUCTION

1.1. Background and Context

Soil is a finite, slowly renewable natural resource that supports terrestrial ecosystem functioning, global food production, water regulation, and carbon storage. Among the various forms of land degradation, accelerated soil erosion by water is widely regarded as the most pervasive and economically consequential threat to soil health worldwide (Montanarella et al., 2015; Pimentel et al., 1995). Water erosion preferentially removes the fertile surface horizon, which concentrates most of the soil organic Matter (SOM), plant nutrients, and biological activity. The resulting on-site consequences include declining land productivity and soil structural deterioration, while off-site impacts encompass reservoir sedimentation, eutrophication, and the disruption of nutrient and carbon cycles at regional to global (Borrelli et al., 2017; Montgomery, 2007). Because rates of pedogenesis are exceedingly slow, frequently less than one ton per hectare per year, soil losses that exceed soil-formation rates constitute an effectively irreversible depletion of the resource over human timescales (Alewell et al., 2019; Montgomery & Matson, 2007).

The global magnitude of the problem is considerable. Using a high-resolution (250 m) modelling framework, (Borrelli et al., 2017) estimated global soil loss by water erosion at approximately 35.9 billion tonnes per year for the year 2012. Their analysis further showed that twenty-first-century land-use change, particularly cropland expansion, is driving a net increase in erosion, with the most pronounced intensification projected for Sub-Saharan Africa, South America, and Southeast Asia. Subsequent global assessments that incorporate climate-change projections support the expectation that rainfall erosivity, and consequently soil erosion, will increase across large parts of the developing world during the coming decades (Borrelli et al., 2020). These findings establish that soil erosion is not simply a local agronomic concern but a driver of global environmental change with direct implications for food security and for the United Nations Sustainable Development Goals, most explicitly the land-degradation-neutrality target of SDG 15.3.

Sub-Saharan Africa is disproportionately affected, and within this region the Ethiopian and Eritrean Highlands constitute one of the most severely eroded landscapes on Earth (Nyssen et al., 2004a). A combination of steep and rugged terrain, highly erosive seasonal rainfall, a long history of intensive cultivation, and progressive removal of natural vegetation has produced erosion rates that are among the highest reported globally; cumulative soil losses across the Eritrean and Ethiopian highlands have been estimated to reach the order of 1.5 billion tonnes annually (Hurni, 1985a; Nyssen et al., 2004b). The Eritrean Highlands share these defining characteristics: steep, dissected topography; a short, intense wet season; fluctuating seasonal vegetation cover; and extensive rain-fed subsistence agriculture, all of which make the landscape highly susceptible to water erosion. Despite this susceptibility, and relative to the adjacent Ethiopian highlands, the Eritrean Highlands remain comparatively understudied and severely data-scarce, both in dense rainfall observations and spatially detailed soil information.

The consequences of this degradation are especially severe in the Eritrean Highlands because the regional economy and rural livelihoods are overwhelmingly dependent on rain-fed subsistence agriculture practiced on steep, erosion-prone slopes. Ogbazghi (1999) emphasised severe soil degradation in the Eritrean Highlands and noted that sustainable agriculture and food security depend on integrated soil management. This approach involves physical soil SWC structures, biological rehabilitation, and enhanced farming practices. In such settings, the loss of topsoil translates directly into declining crop yields, reduced food security, and the abandonment or marginalization of cultivated

land, thereby strengthening a self-preserving cycle of poverty and environmental degradation (Nyssen et al., 2004b; Pimentel et al., 1995). The off-site costs are equally important: eroded sediment shortens the operational lifespan of downstream reservoirs and irrigation infrastructure and degrades water quality. Effective SWC therefore depends not only on knowing how much soil is lost, but on identifying precisely where and when erosion is most severe, so that limited conservation resources can be targeted efficiently.

Advances in satellite Remote Sensing (RS), Geographic Information Systems (GIS), and, more recently, machine learning (ML) have transformed the spatial assessment of soil erosion. Freely available, high-resolution Earth-observation data, including multispectral imagery from the Sentinel-2 mission (Drusch et al., 2012) and satellite precipitation estimates from Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) (Huffman et al., 2007) now provide spatially continuous information on the very variables that drive erosion, namely rainfall and vegetation cover, even in regions devoid of dense ground networks. In parallel, machine-learning methods have matured into powerful tools for predicting environmental variables from complex, non-linear relationships among remotely sensed and terrain-derived covariates (Hengl et al., 2017; Wadoux et al., 2020). The convergence of these technologies creates an opportunity to overcome the data constraints that have historically limited erosion modelling in regions such as the Eritrean Highlands, provided they are integrated within a coherent and physically meaningful framework.

Quantitative assessment of soil erosion over large and heterogeneous areas necessarily relies on models. The empirical Universal Soil Loss Equation (USLE) of (Wischmeier, W. H. and Smith, 1978) and its successor, the Revised Universal Soil Loss Equation (RUSLE) of (K.G. Renard, 1997), remain the most widely applied erosion-modelling frameworks worldwide, owing to their modest data requirements, transparent factorial structure, and ready integration with GIS and Remote Sensing (RS) (Alewell et al., 2019; Ganasri & Ramesh, 2016). RUSLE estimates mean annual soil loss, A ($t\ ha^{-1}\ yr^{-1}$), as the product of five factors.

The integration of RUSLE with satellite RS and GIS has made it possible to derive each factor spatially and to map soil-loss patterns across data-scarce regions where field measurement networks are sparse (Ganasri & Ramesh, 2016; Vrieling, 2006). This methodological flexibility is precisely what makes RUSLE attractive for the Eritrean Highlands, while also exposing it to the specific limitations that motivate the present research.

1.2. Problem Statement

Although RUSLE is widely applied, its conventional implementation embodies two assumptions that are problematic in strongly seasonal, data-scarce environments such as the Eritrean Highlands. These two limitations, one temporal and one spatial, define the problem addressed in this dissertation.

The first limitation concerns temporal representation. In conventional practice, the rainfall-erosivity factor (R) and the cover-management factor (C) are computed from annual-average inputs, so that a single, temporally static value is assigned to each location for the whole year. Yet R and C are the two most dynamic factors in the erosion system, varying noticeably within the year in response to rainfall and vegetation seasonality (Ballabio et al., 2017). In environments where rainfall is concentrated in a short and intense wet season, the actual rate of erosion is governed largely by the temporal coincidence, or phase relationship, between erosive rainfall and the degree of protective vegetation cover present at that moment. Annual averaging dissociates these coupled processes: it redistributes erosive energy uniformly across the year, artificially inflating erosion estimates during dry months while suppressing them during the highly erosive wet months when cover is still establishing. Studies that have resolved

erosivity at the monthly scale in data-rich settings show that neglecting this intra-annual variability can substantially bias both the magnitude and spatial pattern of estimated soil loss, with direct consequences for targeting conservation measures (Ballabio et al., 2017; Panagos et al., 2016). For a highly seasonal landscape such as the Eritrean Highlands, the conventional annual approach may therefore fail to realistically represent erosion dynamics.

The original RUSLE framework accommodates within-season canopy dynamics through empirically calibrated C-factor sub-factor tables. In data-scarce regional applications, where sub-factor field data are unavailable, this procedure is commonly replaced by a single annual mean NDVI-derived C-factor (e.g., Van der Knijff et al., 2000), which collapses within-season C variability into a single temporal value. This study addresses this limitation by deriving 12 monthly C-factor rasters from Sentinel-2 NDVI composites, restoring the temporal resolution lost in conventional annual NDVI-based parameterisation.

The second limitation concerns the spatial accuracy of soil-property information, which propagates directly into the soil-erodibility factor (K-factor). The K-factor is a function of soil texture (sand, silt, and clay fractions) and organic-matter content, and its reliable estimation across a landscape requires spatially representative soil data (K.G. Renard, 1997; Williams J. et al., 1981). In Eritrea in general and the Eritrean Highlands in particular, however, soil observations are sparse and unevenly distributed. The dataset available for this study comprised 302 sampling locations, which, although valuable, is insufficient for reliable conventional interpolation across the steep, geologically and topographically heterogeneous terrain of the study area. Sparse sampling forces reliance on interpolation schemes that smooth out genuine spatial variability and introduce substantial, spatially structured uncertainty into the K-factor and, ultimately, into the soil-loss estimate.

The region's broader data scarcity compounds these two limitations. Ground-based rainfall gauges in the Eritrean Highlands are few and irregularly distributed and seldom record at the high temporal resolution required to characterise rainfall erosivity directly; this constrains the reliability of gauge-based erosivity estimation and motivates the use of spatially continuous satellite precipitation products such as GPM IMERG (Huffman et al., 2007). Taken together, the temporal rigidity of the conventional RUSLE framework and the spatial uncertainty of soil-property inputs represent a coupled methodological challenge that current applications in the region do not adequately resolve.

The datasets selected for this study are well suited to addressing these limitations, but they have not previously been combined within a single temporally explicit framework for the region. GPM IMERG, developed by NASA, merges observations from the GPM Core Observatory, a constellation of passive-microwave sensors, and geostationary infrared measurements to produce spatially continuous, high-resolution precipitation estimates with near-global coverage (Huffman et al., 2007; Levizzani et al., 2020). It is therefore particularly valuable where ground observations are sparse, as in the Eritrean Highlands. Sentinel-2 provides frequent, high-spatial-resolution multispectral imagery from which NDVI can be computed to monitor seasonal vegetation dynamics (Drusch et al., 2012) enabling the cover-management factor to be resolved month by month. The challenge, and the focus of this dissertation, is to integrate these temporally rich satellite inputs with machine-learning-based soil-property prediction so that both the temporal and the spatial weaknesses of the conventional approach are addressed simultaneously.

1.3. Research Gap

A literature review shows that the temporal-variability problem and the spatial-soil-uncertainty problem have generally been treated in isolation, and rarely in data-scarce tropical highland settings. On the

temporal side, the explicit modelling of monthly and seasonal rainfall erosivity has advanced substantially, but predominantly in data-rich regions such as Europe, where dense, high-resolution rainfall networks are available (Ballabio et al., 2017; Panagos, Ballabio, et al., 2015; Panagos et al., 2017). On the spatial side, ML has transformed digital soil mapping, enabling continuous prediction of soil properties from terrain and remotely sensed covariates with quantifiable accuracy (Hengl et al., 2017; Wadoux et al., 2020). However, these two strands of methodological progress have rarely been brought together. In particular, there is a scarcity of studies that simultaneously: (i) exploit satellite-based monthly rainfall (GPM IMERG) to resolve the seasonal dynamics of the rainfall erosivity factor (R-factor); (ii) employ monthly Sentinel-2 NDVI to resolve the seasonal dynamics of the C-factor; and (iii) apply ML to predict soil properties and thereby improve the spatial estimation of the K-factor, all within a single, internally consistent modelling framework.

Equally important, few studies in data-scarce highland environments have formally compared a temporally explicit (simulated, spatiotemporal) RUSLE framework with the conventional non-simulated annual approach to quantify how the choice of temporal representation alters both the magnitude and spatial distribution of estimated soil loss. Without such a controlled comparison, the practical significance of seasonal modelling for conservation planning remains assumed rather than demonstrated. The Eritrean Highlands, a strongly seasonal and severely data-limited region, provide an ideal yet largely unexamined setting in which to address this gap. This dissertation is designed to fill it by integrating temporally explicit rainfall and vegetation dynamics with machine-learning-based soil-property prediction, and by directly comparing the simulated and conventional modelling approaches.

1.4. Aim and Specific Objectives

The overarching aim of this research is to develop and evaluate a temporally explicit, spatially improved soil-erosion modelling framework for the Eritrean Highlands by integrating the RUSLE with GIS, satellite RS, and ML, and to determine whether such a framework provides a more realistic representation of erosion dynamics than the conventional annual-average approach in a highly seasonal, data-scarce environment.

To achieve this aim, the study pursues the following specific objectives:

1. To develop a temporally explicit (monthly, simulated) RUSLE framework that resolves the seasonal dynamics of rainfall erosivity using GPM IMERG precipitation and of vegetation cover using monthly Sentinel-2 NDVI.
2. To develop a machine-learning framework for the spatial prediction of soil properties (sand, silt, clay, and organic-matter content) from terrain and environmental covariates, and to derive an improved soil-erodibility (K) factor for the data-scarce study area.
3. To estimate soil loss using two parallel modelling approaches, a simulated spatiotemporal approach and a conventional non-simulated annual approach, and to quantify and explain the differences between them in both magnitude and spatial pattern.
4. To identify erosion hotspots and derive evidence-based implications for soil conservation, land-management prioritisation, and sustainable land use in the Eritrean Highlands.

1.5. Scope and Limitations

The scope of this research is defined by its study area, modelling framework, and data sources, and it is subject to several limitations that are acknowledged explicitly to delimit the interpretation of the results. The study focuses on the Eritrean Highlands and applies the RUSLE framework, which represents sheet (interrill) and rill erosion by water; it does not account for gully erosion, channel and

bank erosion, or mass-movement processes, which may locally contribute additional soil loss. The analysis is based on the 2021 hydrological year, and although twelve monthly time steps were used to resolve intra-annual dynamics, a single-year analysis cannot capture inter-annual variability or longer-term trends; future work should use a multi-year application.

Within the modelling framework, the soil-erodibility (K), slope-length and steepness (LS), and support-practice (P) factors are treated as temporally invariant and therefore vary only spatially, consistent with their slow rates of change relative to rainfall and vegetation. In the absence of detailed information on conservation structures, the P-factor is assigned a value of one, a conservative assumption that may slightly overestimate soil loss in areas where support practices are present. Rainfall erosivity is derived from an empirical equation that requires annual rainfall as input, so monthly precipitation values are annualized and subsequently rescaled, an approximation whose assumptions are made transparent in the methodology. Finally, satellite-derived inputs (GPM IMERG precipitation, Sentinel-2 imagery, and the DEM) carry their own spatial resolutions and retrieval uncertainties, and the machine-learning soil-property predictions are constrained by the size and spatial coverage of the available 302-point training dataset.

2. LITERATURE REVIEW

2.1. Soil Erosion Processes and Driving Factors

Soil erosion by water is the dominant form of land degradation globally and constitutes a first-order threat to the sustainability of terrestrial ecosystems and agricultural production (Borrelli et al., 2017). The intensity of soil erosion is typically evaluated in relation to soil-loss tolerance, defined as the highest rate of soil removal that does not result in a long-term decrease in productivity; this threshold is generally established around $10 \text{ t ha}^{-1} \text{ yr}^{-1}$ in agricultural management protocols (Wischmeier, 1978), although substantially lower limits are justified in areas with slow soil formation. More recently, regulatory frameworks have adopted considerably more conservative values: Directive (EU) 2025/2360 of the European Parliament and of the Council on Soil Monitoring and Resilience, the EU Soil Monitoring Law adopted on 12 November 2025, sets a sustainable soil loss target of $2 \text{ t ha}^{-1} \text{ yr}^{-1}$, reflecting the recognition that soil formation rates in most landscapes do not exceed $1 \text{ t ha}^{-1} \text{ yr}^{-1}$, making erosion at the rates routinely measured in cultivated highlands effectively irreversible on management timescales (*Directive - EU - 2025/2360 - EN - EUR-Lex*, n.d.). The practical significance of these thresholds is illustrated by (Madarász et al., 2021), who demonstrated over a 16-year field experiment in Central Europe that conservation tillage can reduce measured soil loss to well below tolerable levels, while (Keller et al., 2021a) showed that the choice of erosion model and erodibility estimation approach, including RUSLE and EPIC-based K-factor methods, substantially affects computed soil loss relative to these tolerance benchmarks (Alewell et al., 2019; Montgomery & Matson, 2007). Quantifying where and when this depletion occurs is therefore a prerequisite for effective conservation, and understanding the underlying processes is the foundation on which all predictive models are built.

Soil degradation encompasses a spectrum of processes, of which water erosion and chemical contamination are two of the most widespread and consequential threats to soil function and agricultural productivity. (Zerizghi et al., 2022) demonstrated, through an integrated ecological and human health risk assessment of soils around a coal-mining area, that soil heavy-metal contamination imposes compound risks on soil microorganisms, vegetation, and human health simultaneously. This illustrates that the non-renewable character of the soil resource is threatened not only by physical removal through erosion but also by chemical loading that alters soil properties at rates exceeding the capacity for natural recovery. This convergence of degradation pathways makes quantitatively accurate assessment of each individual process, including water erosion, an urgent scientific and management priority.

Mechanically, water erosion proceeds through three coupled sub-processes: detachment, transport, and deposition. Detachment is accomplished principally by the kinetic energy of raindrop impact, which disrupts aggregates and ejects particles, and secondarily by the shear stress exerted by concentrated overland flow once a runoff threshold is exceeded (Kinnell, 2005). Transport capacity, the mass of sediment that flowing water can carry, is governed by flow depth, velocity, and slope, and the balance between supply (detachment) and capacity (transport) determines whether a given point is erosional or depositional. When transport capacity falls below sediment load, deposition ensues. This supply–capacity framing underlies the distinction between interrill erosion, in which detachment is rainfall-dominated, and transport occurs as shallow sheet flow, and rill erosion, in which both detachment and transport are flow-dominated within incised micro-channels (K.G. Renard, 1997).

On the hillslope, these mechanisms generate a recognized morphological continuum: splash, sheet (interrill), and rill erosion at the slope scale, grading into ephemeral-gully, permanent-gully, and channel erosion where flow becomes strongly concentrated, and complemented by mass-movement processes on steep, saturated terrain. The Universal Soil Loss Equation family represents only the sheet-

and-rill component of this continuum and does not simulate gully incision, channel processes, sediment deposition, or sediment delivery to the catchment outlet (Benavidez et al., 2018). This scope restriction is critical for interpretation: USLE-type estimates are best understood as on-site gross sheet-and-rill soil loss rather than catchment sediment yield, the latter requiring an additional sediment-delivery treatment (de Vente & Poesen, 2005).

The stakes of accurate erosion assessment are simultaneously environmental, economic, and societal. On-site, the preferential removal of the nutrient- and organic-matter-rich surface horizon depresses soil fertility and crop yield and can ultimately drive the abandonment of cultivated land; off-site, mobilised sediment degrades water quality, accelerates reservoir siltation, and impairs aquatic ecosystems (Pimentel et al., 1995). These impacts fall most heavily on agrarian economies dependent on rain-fed cultivation, where the feedback between soil loss, declining productivity, and rural poverty is self-reinforcing. It is for these reasons that soil erosion is now framed explicitly within the United Nations Sustainable Development Goals, particularly the land-degradation-neutrality target of SDG 15.3, which has further sharpened the demand for spatially and temporally credible erosion assessment to underpin conservation prioritization (Alewell et al., 2019; Borrelli et al., 2017).

The rate and spatial pattern of erosion emerge from the non-linear interaction of five groups of controlling factors that together form the conceptual basis of empirical modelling: the erosivity of rainfall; the erodibility of the soil, governed by texture, Organic Matter (OM), structure, and permeability; topography, through slope gradient and length; vegetation and surface cover; and land use and conservation management. These controls are mutually reinforcing rather than additive; for example, the erosive effect of intense rainfall is amplified on steep slopes but strongly attenuated by dense cover, so that erosion is concentrated where high erosivity, steep terrain, erodible soil, and sparse cover coincide. In the Ethiopian and Eritrean Highlands, this coincidence is the rule rather than the exception: erosive, seasonally concentrated convective rainfall falls on steep, dissected terrain that has been cultivated for millennia, producing some of the highest documented erosion rates on Earth and cumulative national soil losses estimated in the order of a billion tonnes per year (Hurni, 1985a; Nyssen et al., 2004b).

2.2. Soil Erosion Modelling Approaches and the Justification for RUSLE

Because direct, spatially distributed measurement of erosion is impractical over large areas, predictive models are indispensable for assessment, scenario analysis, and conservation planning. The substantial body of available models is conventionally classified into three groups according to how they represent the governing processes: empirical (or statistical) models, conceptual (or mixed) models, and physically based (or process-based) models (Borrelli et al., 2021; de Vente & Poesen, 2005; Merritt et al., 2003). The choice among them is not merely technical but reflects a fundamental trade-off between process fidelity and data availability, and this trade-off is decisive in data-scarce settings such as the Eritrean Highlands.

Empirical models are derived from statistical analysis of observed erosion data and encapsulate the dominant controls in parsimonious functional forms. The Universal Soil Loss Equation (Wischmeier, W. H. and Smith, 1978), developed from more than 10,000 plot-years of runoff and soil-loss data in the United States, and its revision, the Revised USLE (K.G. Renard, 1997), are the archetypes. Subsequent variants extend the framework to event timescales using runoff energy (the Modified USLE, MUSLE) and to a software-based computation engine (RUSLE2), while retaining the same factorial logic. These models are lumped and, in their standard form, yield long-term mean annual soil loss; their plot-scale empirical origin imposes assumptions about the range of conditions over which the calibrated relationships remain valid (Benavidez et al., 2018).

Conceptual models, such as the Soil and Water Assessment Tool (J. G. Arnold, 1998), occupy an intermediate position, combining simplified physical reasoning with empirical components to represent water and sediment fluxes through interconnected catchment stores. Physically based models, including the Water Erosion Prediction Project (Flanagan et al., 2007), the European Soil Erosion Model (Morgan et al., 1998), and the Limburg Soil Erosion Model (De Reu et al., 2013)), seek to solve the governing equations of infiltration, runoff generation, detachment, and sediment transport, often at event timescales and fine spatial resolution. While such models offer the richest process representation and the capacity to simulate the effects of specific interventions, they demand extensive, high-resolution parameterisation, are computationally intensive, and are prone to equifinality and parameter-uncertainty problems that can render their spatial predictions no more reliable than those of simpler models when input data are limited (Jetten et al., 2003).

Against this background, RUSLE remains by a wide margin the most widely applied erosion model worldwide; a recent global review and statistical analysis of the modelling literature found USLE-type approaches to dominate published applications, precisely because they combine transparency, modest and widely available data requirements, computational efficiency, and seamless integration with GIS and RS (Alewell et al., 2019; Borrelli et al., 2021). For the Eritrean Highlands, where dense rainfall networks, intensive soil surveys, and the hydrological data required to drive process-based models are unavailable, RUSLE is therefore the pragmatic and defensible choice. This choice is further supported by the existence of region-specific factor adaptations: empirical rainfall-erosivity relationships and management-factor values have been calibrated explicitly for the Ethiopian and Eritrean highlands, providing a locally grounded basis for application (Hurni, 1988).

This practicality must, however, be accompanied by a clear-eyed recognition of RUSLE's limitations, several of which directly motivate the present study. The model estimates only sheet and rill erosion and ignores gully and channel processes and deposition; it transfers empirical relationships calibrated under particular climatic and pedological conditions to new environments, introducing structural uncertainty; its factors carry substantial parameter uncertainty, especially the erodibility and cover terms; and the literature is plagued by inconsistencies in units and factor definitions that complicate comparison across studies (Alewell et al., 2019; Benavidez et al., 2018). Most importantly for this dissertation, the conventional implementation assumes temporally averaged and annual values. The methodological contributions developed in later chapters, namely a temporally explicit formulation and a machine-learning-based erodibility factor, are designed to relax two of these limitations within the otherwise robust RUSLE framework.

2.3. The RUSLE Factors

RUSLE estimates mean annual soil loss as the multiplicative product of five factors, each of which has its own substantial methodological literature.

R-factor: The R-factor quantifies the climatic driving force of erosion and is defined as the long-term average of the product of total storm kinetic energy and the maximum 30-minute rainfall intensity (EI_{30}), summed over all erosive events (Wischmeir, W. H. and Smith, 1978). Its rigorous computation thus requires continuous, high-temporal-resolution (sub-hourly) rainfall records that are unavailable across most of the developing world. To avoid this, an empirical replacement estimates R from coarser rainfall statistics; the regression relationships of (Renard 1997 & Freimund, 1994) based on annual precipitation and the modified Fournier index of monthly rainfall, are among the most widely transferred, and the equation calibrated by (Hurni, 1985a) for the Ethiopian and Eritrean highlands is standard in the region. At the global scale, the Global Rainfall Erosivity Database (GloREDa) has harmonised station-based erosivity for more than 3,600 stations, providing a benchmark and revealing the strong dependence of

R on precipitation intensity rather than amount alone (Panagos et al., 2017). Because R is climate-driven and not directly controllable, it is also the most sensitive factor to climate change, with several assessments projecting increases in erosivity across much of Africa (Borrelli et al., 2020).

K-factor: The K-factor expresses the intrinsic susceptibility of a soil to detachment and transport per unit of rainfall erosivity and is controlled by particle-size distribution, organic-matter content, structure, and permeability. The classical Wischmeier nomograph requires inputs, such as structural class and profile permeability, that are seldom mapped, so the texture-and-organic-carbon formulation embedded in the EPIC model (Sharples, 1985) is now the most common operational approach, requiring only the sand, silt, clay, and organic-matter fractions. The fidelity of K is therefore wholly contingent on the spatial quality of the soil-property data from which it is computed; where samples are sparse and terrain heterogeneous, as in the Eritrean Highlands, conventional interpolation produces an over-smoothed and uncertain K surface. This dependence is the principal justification for the digital soil mapping approach, which has been demonstrated at continental and global scales for the very properties K requires (Hengl et al., 2017; Panagos et al., 2017).

Slope length and steepness (LS-factor): The LS-factor encapsulates the topographic control on erosion, integrating the acceleration of flow with slope gradient and the accumulation of runoff with slope length. The original USLE formulation, derived from unit plots of fixed length, has been generalised for application to complex two-dimensional terrain using flow-accumulation algorithms applied to digital elevation models (DEM). The unit-stream-power approach of (I. A. N. D. Moore & Burch, 1986) and the upslope-contributing-area algorithm of (P.J.J. Desmet and G. Govers, 1996) are the most widely used, the latter explicitly accommodating flow convergence and divergence. A recognised caveat is the strong sensitivity of LS, and hence of estimated soil loss, to DEM resolution and to the flow-routing algorithm chosen, which must be considered when interpreting absolute soil-loss magnitudes.

Cover management (C-factor): The C-factor represents the protective effect of vegetation, residue, and surface roughness, expressed as the ratio of soil loss under a given cover to that from continuously bare, tilled fallow; it ranges from near zero under dense forest to one for bare soil, and it is, with R, the most temporally dynamic factor. Historically assigned from land-cover-class lookup tables, C is increasingly derived from remotely sensed vegetation indices, which provide spatially continuous and temporally resolved estimates of cover. The exponential relationship between C and the Normalised Difference Vegetation Index proposed by (Van Der Knijff et al., 2000), and tropical adaptations thereof (Durigon et al., 2014), are widely adopted, while alternative regression-based C surfaces have been produced at continental scale (Panagos, Borrelli, et al., 2015). **Support practice (P).** The P-factor accounts for conservation measures such as contouring, strip-cropping, and terracing that redirect or slow runoff. Spatially explicit information on such practices is rarely available, and P is therefore frequently set to unity, a conservative assumption that neglects existing structures and may inflate soil-loss estimates where conservation is in fact practised (K.G. Renard, 1997).

A consideration that is too often neglected in applied studies, but which a senior treatment must foreground, is the propagation of factor uncertainty through the multiplicative model structure. Because RUSLE multiplies five imperfectly known factors, relative errors compound rather than cancel, and the overall uncertainty of a soil-loss estimate can be large even when each individual factor is reasonably constrained. The erodibility and cover factors are typically the most uncertain, the former because of sparse soil data and the latter because of the empirical NDVI–C transfer and seasonal cloud contamination. Moreover, the factors are not strictly independent: erosivity, cover, and runoff covary through the seasonal cycle, so that treating them as separable annual averages introduces a structural error distinct from simple parameter uncertainty. These considerations strengthen two of the design

choices of this dissertation, namely the use of ML to reduce erodibility uncertainty and of a temporally explicit formulation to preserve the covariance between erosivity and cover (Alewell et al., 2019; Benavidez et al., 2018).

2.4. Temporal and Seasonal Dimensions of Erosion Modelling

In its standard application, RUSLE is parameterized with temporally averaged inputs and returns a single, static estimate of mean annual soil loss for each location. This convention is computationally convenient and adequate where rainfall and vegetation are weakly seasonal, but it covers the strong intra-annual variability that characterizes monsoonal and Mediterranean climates. Of the five factors, rainfall erosivity (R) and cover management (C) vary meaningfully within the year, whereas erodibility, topography, and support practice are effectively constant at the annual scale. The case for sub-annual modelling therefore rests on the dynamics of these two factors and, crucially, on their temporal correlation.

A substantial literature, developed largely in data-rich Europe, has demonstrated the magnitude of this intra-annual variability. Mapping monthly rainfall erosivity across the continent, (Ballabio et al., 2017) found mean summer erosivity to be approximately four times greater than winter erosivity, with the seasonal pattern varying systematically between Mediterranean and continental regimes. Comparable regional analyses have characterized the seasonal concentration of erosive rainfall in Greece (Panagos et al., 2016) and the regionalization of monthly erosivity in Switzerland (Schmidt et al., 2016), introducing the concept of erosivity density, the erosivity per unit rainfall, as a diagnostic of the months and regions where erosive storms are most likely. These studies establish that both the magnitude and the timing of erosivity are spatially structured and that annual aggregation discards information of direct relevance to conservation.

The deeper significance of temporal resolution lies in the phase relationship between erosive rainfall and protective cover. In strong seasonal climates, the onset of the rains frequently precedes the full development of vegetation canopy, so that the most erosive storms coincide with a still-establishing, low-cover surface; conversely, peak biomass may occur after the erosive peak has passed. Actual erosion is governed by the temporal coincidence of high R and low C, a coincidence that annual averaging dissolves by redistributing both quantities uniformly across the year. The consequence is a structured bias whose sign varies in space and time: the annual approach tends to overstate erosion during low-rainfall months and understate it during the most erosive months, producing patterns of both over- and under-estimation depending on local interactions among rainfall, vegetation, topography, and soil (Ballabio et al., 2017). Recognizing this, several authors have argued explicitly for estimating soil loss at monthly or seasonal scales in environments with high intra-annual rainfall variability (Benavidez et al., 2018).

The mechanism by which annual aggregation distorts erosion estimates can be made concrete. Consider a hillslope subjected to a brief, intense wet season during which the bulk of annual erosivity is delivered while the vegetation canopy is still developing, followed by a long dry season during which negligible rain falls but vegetation, where present, is most senesced. A temporally explicit model multiplies, in each month, the erosivity received by the cover present, and sums the twelve products; it thereby captures the damaging early-wet-season combination of high erosivity and low cover. An annual model, by contrast, multiplies a single mean erosivity by a single mean cover. Because the product of two means is not equal to the mean of monthly products when the two series are correlated, the annual estimate is biased; where erosivity and bare-soil conditions are positively associated in time, as in many monsoonal systems, the direction and magnitude of that bias vary spatially with local cover phenology, runoff concentration, and erodibility, producing the mixed pattern of over- and under-estimation

reported in comparative studies. This non-commutativity of temporal averaging, rather than any deficiency in the RUSLE factors themselves, is the fundamental reason a simulated spatiotemporal approach is expected to differ from, and better represent reality than, the conventional annual approach.

Despite this compelling rationale, sub-annual erosion modelling has remained concentrated in regions with dense, high-temporal-resolution rainfall networks and frequent cloud-free imagery. Its application in data-scarce tropical highlands, where seasonality is most pronounced and the potential payoff greatest, has been limited principally by the difficulty of obtaining temporally resolved rainfall and vegetation information. The maturation of satellite precipitation products and of high-revisit optical missions now removes much of this barrier, creating the opportunity, central to this dissertation, to implement a temporally explicit RUSLE in a setting where it has rarely been attempted.

2.5. Satellite Rainfall Products for Erosivity in Data-Scarce Regions

The estimation of rainfall erosivity in regions such as the Eritrean Highlands is fundamentally constrained by the rain-gauge network. Across much of sub-Saharan Africa, gauges are sparse, unevenly distributed, prone to record gaps, and rarely equipped to register the sub-hourly intensities that erosivity computation demands; point observations also represent rainfall poorly in mountainous terrain where convective storms are spatially heterogeneous. (Ayehu et al., 2018; Dinku et al., 2018). Satellite-based quantitative precipitation estimates address these deficiencies by providing spatially continuous, temporally consistent rainfall fields with near-complete coverage, and have become indispensable for hydrological and agro-environmental analysis in data-scarce settings.

Three generations of products are in widespread use. The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) delivered the first broadly adopted quasi-global, three-hour, 0.25° precipitation estimates and underpinned a generation of tropical hydrological studies. (Huffman et al., 2007). The Climate Hazards group InfraRed Precipitation with Stations (CHIRPS) blends thermal-infrared cold-cloud-duration estimates with a satellite-enhanced climatology and station data to produce a long (1981–present), high-resolution (0.05°) record specifically engineered for African drought and trend monitoring. (Funk et al., 2015). The Global Precipitation Measurement (GPM) mission, the operational successor to TRMM, generates the Integrated Multi-satellitE Retrievals for GPM (IMERG), which intercalibrates and merges passive-microwave retrievals from a constellation of radiometers, anchored by the dual-frequency radar and microwave imager of the GPM Core Observatory, with geostationary infrared data, yielding precipitation at 0.1° spatial and half-hourly temporal resolution and near-global coverage (Huffman et al., 2020). The gauge-adjusted Final Run product, in its Version 07 form, is the most accurate IMERG configuration for research use.

The reliability of these products in the highland-African context has been the subject of careful validation. Continental and basin-scale assessments over Ethiopia and eastern Africa generally find that CHIRPS and IMERG capture the seasonal rainfall regime, inter-annual variability, and spatial gradients with useful skill, with CHIRPS often outperforming earlier products such as TAMSAT and ARC at dekadal-to-monthly scales (Ayehu et al., 2018; Dinku et al., 2018). At the same time, these studies consistently report systematic, terrain- and intensity-dependent biases: a tendency to overestimate the frequency of low-intensity rain days, reduced skill at high elevation, and difficulty resolving the warm-rain orographic processes characteristic of highland convection. Such biases are consequential for erosivity, which is dominated by the high-intensity tail of the rainfall distribution, and they imply that satellite-derived R must be interpreted with appropriate caution and, where possible, locally evaluated.

The fine temporal resolution of IMERG has enabled an emerging class of satellite-based erosivity estimates that approximate the EI_{30} structure directly from half-hourly fields, complementing the coarser empirical substitutes. This capability is especially valuable for temporally explicit modelling,

because it supplies the month-to-month erosivity dynamics that gauge networks cannot. A growing number of studies have exploited the half-hourly GPM IMERG record to derive sub-daily rainfall intensities and to map erosivity at continental and global scales, demonstrating that satellite quantitative precipitation estimates can reproduce the broad spatial pattern of erosivity where gauges are absent. Nonetheless, the propagation of intensity-dependent satellite biases into erosivity, and the validation of satellite-based seasonal R in highland Africa specifically, remain at an early stage; the documented tendency of infrared- and microwave-based retrievals to smooth extreme intensities is of direct concern for a quantity dominated by the most intense storms, and locally calibrated empirical relationships applied to satellite rainfall totals may, for some applications, remain more robust than direct EI_{30} reconstruction. These considerations underscore both the promise and the present limitations of satellite rainfall for the temporally explicit erosivity modelling developed in this thesis (Panagos et al., 2017).

2.6. Remote Sensing of Vegetation Dynamics for the C-Factor

Because vegetation cover is the most temporally variable control on erosion and the factor most amenable to monitoring from space, its remote-sensing characterization is the cornerstone of a temporally explicit cover-management factor. The NDVI, the normalized ratio of near-infrared to red reflectance, exploits the strong contrast between chlorophyll absorption in the red and high mesophyll reflectance in the near-infrared to provide a robust, physically interpretable proxy for green-vegetation density, fractional cover, and vigour (Compton J. Tucker, 1978). The advent of the Sentinel-2 constellation, with four 10 m visible-to-near-infrared bands, additional red-edge and shortwave-infrared bands, and a combined revisit of around five days, has made it possible to track vegetation phenology at field-relevant spatial scales and at monthly or finer temporal steps, transforming the feasibility of dynamic C-factor estimation (Drusch et al., 2012). (Measho et al., 2019) performed a spatio-temporal analysis of vegetation changes in Eritrea from 2000 to 2017, utilizing NDVI-based remote sensing. They found that 57.1% of the pixels examined showed a decline in annual NDVI, particularly in the south-western areas. This decline was linked to increased climate variability, recurrent droughts, and land-use pressures. Their findings highlight that vegetation cover in the Eritrean Highlands is closely tied to seasonal rainfall and is highly sensitive to inter-annual rainfall fluctuations. These results are directly relevant to the monthly C-factor variations observed in this study.

Several functional forms translate NDVI into the C-factor. Early studies used linear or piecewise relationships, but the exponential function introduced by (Van Der Knijff et al., 2000) has become the de facto standard because it reproduces the strongly non-linear decline of erosion risk as cover increases, capturing the rapid protective gain provided by even partial canopy development.

Notwithstanding their utility, NDVI-based methods carry well-documented limitations that a rigorous application must acknowledge. NDVI saturates at high biomass, compressing the discrimination of dense canopies; it is sensitive to bright or dark soil backgrounds in sparsely vegetated terrain, which can bias C in exactly the marginal landscapes where erosion is most active; and optical retrievals are vulnerable to cloud and atmospheric contamination, a particular concern during the cloudy wet season when erosivity peaks. These issues motivate the use of temporal compositing, cloud masking, and, in some studies, soil-adjusted or red-edge indices. For the purposes of seasonal C estimation, the key strengths, spatial continuity and the ability to resolve intra-annual change outweigh these limitations, provided the caveats are recognized in interpretation.

2.7. Machine Learning and Digital Soil Mapping for Soil-Property Prediction

Digital soil mapping (DSM) provides the methodological route to a spatially robust erodibility factor from sparse field observations. Its conceptual foundation is the empirical relationship, summarized by the soil-forming factors of soil, climate, organisms, relief, parent material, age, and spatial position,

between a soil property and a set of environmental covariates that can be observed exhaustively across the landscape (McBratney et al., 2003). (Dobos & Hengl, 2009) provided a comprehensive methodological framework for soil mapping applications, demonstrating how terrain attributes derived from DEMs; including slope, curvature, topographic wetness index, and flow accumulation; combined with RS spectral data, can serve as spatially continuous proxies for the soil-forming factors of relief, parent material, and vegetation cover, enabling the prediction of soil properties across unsurveyed areas with quantifiable uncertainty. From a representative sample of point observations, this relationship is modelled and used to predict the property continuously, together with an estimate of prediction uncertainty. Where conventional geostatistical interpolation degrades because samples are too few and the terrain too heterogeneous to resolve short-range variability, DSM borrows strength from covariates, typically DEM-derived terrain attributes, climatic surfaces, and satellite indices, that are themselves spatially complete.

Machine-learning algorithms have become one of the most efficient predictive engines of DSM because of their capacity to model complex, non-linear, high-dimensional relationships without restrictive parametric assumptions. Ensemble tree methods are especially prominent: Random Forests (RF) (Breiman, 2001) aggregate many decorrelated regression trees to yield stable predictions with low variance and intrinsic measures of variable importance, while gradient-boosted trees, particularly the regularized XGBoost (XGB) implementation (Chen & Guestrin, 2016), often achieve higher accuracy by sequentially fitting residuals. These methods underpin global soil-property products such as SoilGrids, which predicts texture fractions, organic carbon, and related properties from large profile databases and remote-sensing covariates, initially at 250 m resolution using a machine-learning ensemble (Hengl et al., 2017) and subsequently with quantified uncertainty in its second generation (Poggio et al., 2021). Methodological reviews confirm that ML generally outperforms classical interpolation for soil-property mapping while emphasizing the importance of appropriate validation and uncertainty assessment (Padarian et al., 2020; Wadoux et al., 2020).

The approach has been successfully applied in African mountain environments comparable to the study area; for example, comparative assessments of machine-learning regressors for mapping soil organic carbon in the Eastern Afromontane region of East Africa found that tree-based ensembles provided accurate, spatially detailed predictions from terrain and remote-sensing covariates. (Were et al., 2015). Realising this potential reliably nonetheless requires careful attention to covariate selection, model tuning, and, above all, validation design: because soil samples are spatially autocorrelated, non-spatial cross-validation can yield optimistic accuracy estimates, and spatially structured or independent validation is preferred, while extrapolation beyond the environmental range of the training data must be constrained. For erosion modelling, the most direct and valuable application is the continuous prediction of the texture and organic-matter fractions required by the EPIC erodibility formulation, converting a sparse point dataset into a spatially detailed K-factor surface in regions where sample density would otherwise preclude reliable estimation. (Hengl et al., 2017; Williams J. et al., 1981). This is precisely the strategy adopted in the present research.

Beyond predictive accuracy, two further attributes make machine-learning DSM well matched to erosion modelling. First, ensemble methods yield measures of variable importance that illuminate the environmental controls on soil properties, allowing the plausibility of a model to be assessed against pedological understanding rather than accepted as a black box; in mountainous terrain, relief-derived covariates such as slope, curvature, and topographic wetness commonly emerge as dominant predictors of texture and OM, consistent with process expectation. Second, modern DSM increasingly predicts with explicit uncertainty quantification, propagating it into derived products, so that the resulting K-factor can be accompanied by a spatially distributed confidence estimate rather than presented as a

deterministic surface (Poggio et al., 2021; Wadoux et al., 2020). For a quantity as uncertain as soil erodibility, this capacity to express, rather than conceal, spatial uncertainty represents a substantive methodological advance over conventional interpolation.

2.8. Integration of GIS, Remote Sensing and Machine Learning in Erosion Modelling

The operational power of RUSLE derives from its implementation within a GIS, in which each factor is represented as a co-registered raster surface and soil loss is computed cell by cell through map-algebra multiplication (Ganasri & Ramesh, 2016; Vrieling, 2006; Vrieling et al., 2010). This raster-overlay architecture is intrinsically modular and therefore naturally suited to the assimilation of heterogeneous, multi-source inputs: satellite precipitation products supply the rainfall basis of the R-factor; optical vegetation indices yield the C-factor; DEMs furnish the LS-factor; and machine-learning predictions provide the soil properties underlying the K-factor. The GIS thus functions as the integration platform that converts disparate Earth-observation and modelled data streams into a single, spatially distributed estimate of erosion.

The integration of GIS, RS, and RUSLE is now mature and has been applied extensively, including numerous studies across the Ethiopian highlands that have mapped erosion hotspots, evaluated land-use-change impacts, and prioritised conservation interventions; dedicated reviews document the breadth of these catchment-scale applications and the methodological choices they entail (Ganasri & Ramesh, 2016; Phinzi & Ngetar, 2019). The accessibility of cloud-based geospatial computing, notably planetary-scale platforms that co-locate petabytes of imagery with analytical capacity, has further lowered the barrier to multi-temporal, large-area analysis. More recently still, the incorporation of ML has extended this paradigm beyond image classification toward the data-driven prediction of model inputs themselves, most notably soil properties, and toward the analysis of erosion drivers and susceptibility, reflecting a broader movement toward artificial-intelligence-enhanced geospatial environmental modelling.

Within eastern Africa specifically, the GIS–remote-sensing–RUSLE combination has become the standard tool for catchment- and basin-scale erosion assessment, used to delineate severity classes, locate hotspots, and evaluate the erosion consequences of land-use and land-cover change over multi-decadal intervals. These applications have consistently identified the steep, cultivated highlands as zones of severe to extreme soil loss and have informed the targeting of physical and biological conservation measures. Their credibility, however, is bounded by the quality of their inputs, most acutely the erosivity factor, derived from sparse gauges or coarse empirical surrogates, and the erodibility factor, derived from limited soil surveys. The substitution of satellite precipitation for sparse gauges and of machine-learning predictions for sparse soil samples therefore represents not a peripheral refinement but a direct strengthening of the weakest links in the established regional methodology (Phinzi & Ngetar, 2019).

Yet a critical reading of this literature reveals that its full potential remains unrealised. The overwhelming majority of integrated GIS–RS–RUSLE applications retain the conventional annual formulation, treating rainfall erosivity and cover management as static annual averages, and continue to derive the K-factor from interpolated rather than predicted soil properties. The two methodological advances with the greatest leverage in seasonal, data-scarce environments, namely temporally explicit modelling of erosivity and cover, and machine-learning-based spatial prediction of soil properties, have largely been pursued in separate research strands and seldom combined within a single coherent assessment. This disjunction defines the limit that the present dissertation addresses.

Spatially explicit RUSLE modelling using GIS has been applied at national scale to quantify erosion risk across heterogeneous agricultural landscapes: (Pásztor et al., 2016) assessed soil erosion across

Hungary using spatially explicit modelling, demonstrating that integrating soil data, terrain attributes, and land use information within a GIS framework can produce nationally consistent erosion risk maps. The methodological principles of this approach underpin the regional RUSLE implementation of the present study.

2.9. Soil Erosion Assessment in the Eritrean Highlands: The Mai Nefhi Pilot Watershed Study

The present dissertation is part of a progressive research programme on soil erosion modelling in the Eritrean central highlands. The first stage of this programme comprised a two-phase pilot investigation at the Mai Nefhi watershed of the Barka River Basin (7,659 ha, approximately 25 km south-west of Asmara), conducted using the same study area, the same dataset, and the same RUSLE computational framework across both phases. The first phase produced a conventional RUSLE-based erosion risk assessment (Menghis et al., 2025); the second phase extended this work by performing a systematic sensitivity analysis of all five RUSLE parameters applied to the same pilot watershed outputs (Menghis et al., 2026). Together, these two publications constitute a single integrated pilot investigation whose combined findings established the empirical foundation and identified the methodological limitations that motivated the development of the regional spatiotemporal framework presented in this dissertation.

2.9.1. Conventional RUSLE Assessment

The first phase applied the annual RUSLE framework to the Mai Nefhi watershed using Landsat 8 OLI imagery, Shuttle Radar Topography Mission (SRTM) DEM, and the soil samples, producing the first spatially explicit erosion risk assessment for any Eritrean watershed. The R-factor was derived from the Eritrean Ministry of Agriculture MOA department of National Agricultural research Institutes (NARI) station annual rainfall ($177\text{--}232\text{ MJ mm ha}^{-1}\text{ h}^{-1}\text{ yr}^{-1}$), the K-factor from the 161 soil samples ($0.110\text{--}0.822\text{ t ha h ha}^{-1}\text{ MJ}^{-1}\text{ mm}^{-1}$), the LS-factor from SRTM DEM analysis ($0\text{--}7.03$), the C-factor from land use/land cover classification, and the P-factor from slope-based conservation practice tables ($0.148\text{--}0.5$). Annual soil loss ranged from 0 to $250\text{ t ha}^{-1}\text{ yr}^{-1}$, with 61.93% of the watershed in the Low class ($0\text{--}7\text{ t ha}^{-1}\text{ yr}^{-1}$) and 3.94% in the Severe class ($90\text{--}250\text{ t ha}^{-1}\text{ yr}^{-1}$), concentrated on steep, sparsely vegetated hillslopes. The LS-factor was identified as the dominant spatial driver of erosion variability, with erosion rates increasing sharply above 30% slope. A regression analysis between OM content and soil erosion rate yielded $R^2 = 0.279$, confirming a statistically significant but weak inverse relationship and establishing that SOM alone accounts for less than 28% of the spatial variability in erosion (Menghis et al., 2025).

2.9.2. Sensitivity Analysis of the Same RUSLE Framework

The second phase extended the pilot investigation by applying a one-at-a-time (OAT) sensitivity analysis to the same RUSLE model, systematically perturbing each of the five parameters across their plausible range while holding the others constant and computing a sensitivity index (SI) for each. This analysis determined the relative influence of each parameter on modelled soil loss and translated the results into a prioritised framework for SWC design. The LS-factor was confirmed as the highest-sensitivity parameter, consistent with its spatial dominance identified in Phase 1: the non-linear dependence of the RUSLE LS formulation on slope gradient amplifies perturbations disproportionately near the critical slope thresholds. The C-factor was identified as the highest-controllable-sensitivity parameter; meaning that proportional changes in C produce proportional or greater changes in annual soil loss. P-factor scenario analysis evaluated the soil loss reduction achievable across a range of conservation practice conditions, from the current watershed practice levels ($P = 0.148$ to 0.5) to no SWC (SWC) ($P = 1$), quantifying the conservation benefit attributable to existing stone bunds and terracing practices and providing the empirical basis for estimating the $P = 1$ overestimation that the present dissertation acknowledges (Menghis et al., 2026).

2.9.3.Connection to the present Dissertation

The limitations identified in the pilot watershed study directly informed the methodological design of the present dissertation. Specifically, the lack of temporal variability in rainfall erosivity and vegetation cover, limited rainfall data coverage, inadequate spatial representation of soil erodibility, reliance on static land use-based C-factors, restriction to a single watershed, and absence of comparisons between modelling approaches were addressed through several advancements. These include the use of monthly GPM IMERG-derived R-factor rasters, monthly Sentinel-2 NDVI-based C-factor rasters, machine-learning-based K-factor prediction using 302 soil profiles from MOA/NARI, regional-scale analysis across the central and southern Eritrean Highlands, and a dual-framework design that quantitatively compares conventional and spatiotemporal RUSLE approaches. Together, these improvements provide a more comprehensive and spatially explicit framework for soil erosion assessment and conservation planning.

2.10. Synthesis and Identification of the Research Gap

The evidence reviewed in this chapter supports a connected chain of conclusions. First, water erosion is governed by the non-linear interaction of rainfall erosivity, soil erodibility, topography, vegetation cover, and management, and among these the erosivity and cover factors are strongly seasonal, so that the timing of their coincidence, not merely their annual averages, controls the erosion outcome. Second, despite acknowledged limitations, the empirical RUSLE framework is the most appropriate model for large-area assessment in data-scarce environments, but its conventional annual implementation imposes a temporally static representation that is poorly matched to the highly seasonal climate of the Eritrean Highlands and that biases estimates in spatially and temporally structured ways. Third, the data constraints that historically precluded temporally resolved and spatially robust modelling in such regions can now be relaxed: satellite precipitation products deliver spatially continuous, temporally fine rainfall; Sentinel-2 NDVI resolves seasonal vegetation dynamics for the cover factor; machine-learning digital soil mapping converts sparse soil samples into spatially detailed property and erodibility surfaces; and GIS provides the platform to integrate these inputs.

The decisive observation is that, although each of these components are individually mature, they have rarely been brought together. The literature reveals a specific and consequential gap: few studies, and fewer still in data-scarce tropical highlands, integrate satellite-based monthly rainfall erosivity, monthly Sentinel-2 NDVI-derived cover dynamics, and machine-learning-predicted soil properties within a single, temporally explicit RUSLE framework, and almost none formally compare such a simulated spatiotemporal model against the conventional annual model to quantify the magnitude, direction, and spatial structure of the bias introduced by the temporal assumption. The Eritrean Highlands, severely under-studied, acutely data-limited, and strongly seasonal, constitute an ideal setting in which to address this gap. By doing so, the present dissertation aims to contribute both a transferable methodology for erosion assessment in data-scarce regions and new empirical knowledge to support SWC, land-management prioritisation, and progress toward land-degradation neutrality in the region.

3. METHODS AND MATERIALS

3.1. Overall Methodological Framework

This study's methodological framework combines the RUSLE) with GIS, satellite RS, and machine learning to create a detailed, spatiotemporal assessment of potential soil erosion in the Eritrean Highlands (Figure 1). Each of the five RUSLE factors rainfall erosivity (R), soil erodibility (K), slope length and steepness (LS), cover-management (C), and support practice (P) was independently derived from the most suitable datasets and models. These five factor rasters were then integrated within GIS to estimate soil loss, with each component explained in its respective section

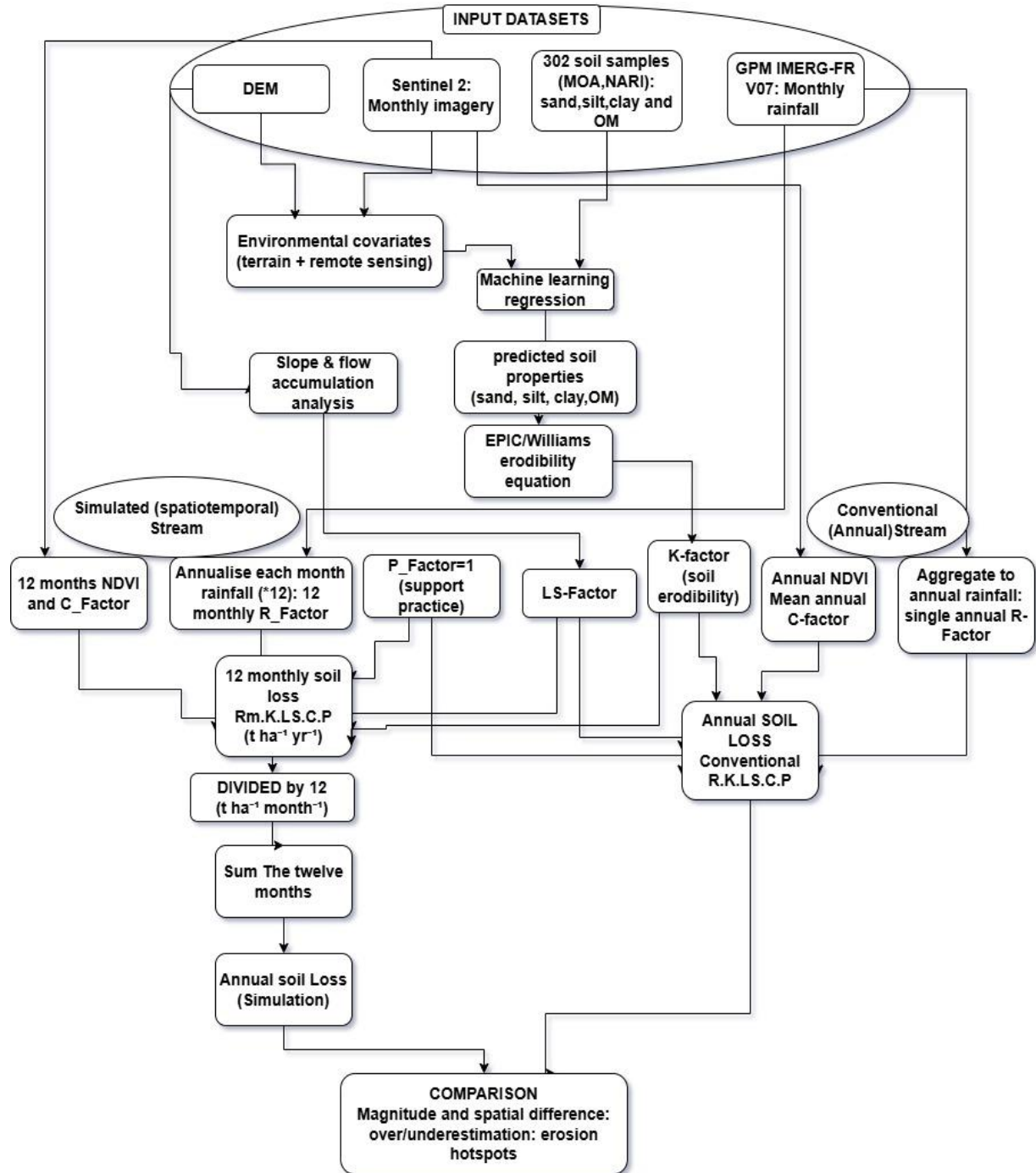


Figure 1. Overall methodological framework flow diagram of the study

A key feature of the framework is the development of two parallel soil loss calculation methods: a dynamic, monthly approach where rainfall erosivity and vegetation cover change each month, generating 12 monthly rasters that are de-annualised and summed to produce an annual simulated soil loss map (A_annual); and a traditional annual approach with constant averages for R and C to produce a conventional annual map (A_non-sim). The static factors K, LS, and P remain identical in both methods, so differences between A_annual and A_non-sim only reflect the effects of temporal variation in R and C. A pixel-wise difference map ($\Delta A = A_{\text{non-sim}} - A_{\text{annual}}$) is created to visualise the spatial divergence. The flowchart in Figure 1 illustrates the complete workflow, including both modelling streams and the difference map computation, and the following sections describe each component in detail.

3.2. Study Area

The research was conducted in the Eritrean Highlands, the country's central mountainous core. The Highlands were selected because they encompass, within a single physiographic unit, the conditions that make soil erosion both severe and difficult to model: steep, deeply dissected terrain; a strongly seasonal rainfall regime concentrated in a short wet season; dynamic, largely rainfall-driven vegetation cover; extensive rain-fed subsistence cultivation on marginal slopes; and a chronic scarcity of dense ground-based rainfall data and soil observations. These attributes make the area an ideal setting for evaluating a temporally explicit erosion-modelling framework and a machine-learning approach to soil-property prediction. The following subsections describe the location and physiography; geology and soils; climate and seasonal rainfall regime; vegetation and land use/land cover; topography and drainage; and the study area's agricultural systems and erosion susceptibility.

3.2.1. Location and Physiography

Eritrea is a north-east African country bordering the Red Sea, sharing land boundaries with Sudan to the west and north-west, Ethiopia to the south, and Djibouti to the south-east. The national territory is conventionally divided into three contrasting physiographic provinces: the hot, narrow Red Sea coastal plains in the east; the cool central highlands; and the semi-arid western lowlands that descend toward the Sudanese plains and the Sahel. The Eritrean Highlands, known locally as the Kebessa, form the central province and constitute the northern continuation of the Ethiopian Plateau, occupying approximately the central third of the country (Eritrean Highlands, geographic syntheses; (Nyssen et al., 2004b).

The Highlands are an elevated plateau with a mean elevation of approximately 2,000 m above sea level and altitudes that generally range between about 1,800 and 3,000 m; the national capital, Asmara, lies on the plateau at roughly 2,325 m and near 15° N, 39° E, while the highest point in the country, Emba Soira (3,018 m), rises in the south-eastern Highlands. Physiographically, the plateau is strongly asymmetric: its eastern margin terminates in a precipitous escarpment that descends more than 2,000 m over a short horizontal distance to the coastal plain, whereas its western flank grades more gently through undulating hills into the western lowlands. The plateau surface is itself deeply incised by gorges and is punctuated by the flat-topped, steep-sided mesas (locally *ambas*) that are characteristic of the region. This pronounced relief, developed over a short distance, is a primary control on the topographic erosion factor analysed in later sections.

The study area encompasses the central and southern highland belt of Eritrea, incorporating four administrative units: the entire Maekel (Central) region and Debub (Southern) region, together with the Adi Teklezan sub-region from the Anseba region and the Logo Anseba sub-region from the Gash-Barka region.

This delineation captures a contiguous highland landscape that transcends conventional two-region boundaries and reflects the shared physical, hydrological, and land-use characteristics of the central and southern highland system. Geographically, the study area extends between latitudes 14° 26' 59.52" N and 15° 36' 29.76" N, and between longitudes 38° 30' 24.27" E and 39° 31' 58.31" E (Figure 2). The administrative boundary layers used to delineate and prepare the study area map were obtained from the Eritrean Subnational Administrative Boundaries dataset available through the Humanitarian Data Exchange (HDX) platform (OCHA, 2023), ensuring that the spatial extent of the study area is precisely and reproducibly defined. The study area constitutes the primary agricultural heartland of Eritrea, encompassing the most densely populated rural zone of the country and the landscape most exposed to the combined erosive forces of steep terrain, seasonal high-intensity rainfall, and persistent land-use pressure.

Table 1 provides a geographic reference for the 17 sub-regions/zobas of the study area, presenting their administrative region (Zoba), approximate geographic coordinates, landscape zone classification, and elevation range. This table serves as a permanent spatial reference throughout the dissertation, as sub-region/zoba names are used consistently in the results, discussion, and conclusions sections to describe the spatial distribution of soil properties, RUSLE factors, and soil loss patterns across the study area.

Table 1. *Sub-zoba Geographic Reference: Study Area Administrative Divisions and Geographic Location*

S.No	Sub-Region/Study area Nomenclature	Region	Geographic Location (approx.)	Landscape Zone
1	Asmara City	Central/Maekel	15.33°N, 38.93°E	Central plateau
2	Ghala Nefhi	Central/Maekel	15.20°N, 38.85°E	Central plateau
3	Berikh	Central/Maekel	15.40°N, 38.75°E	Central plateau
4	Serejeka	Central/Maekel	15.55°N, 38.90°E	North-central
5	Dibarwa	Southern/Debub	15.05°N, 38.83°E	South-central plateau
6	Mendefera	Southern/Debub	14.88°N, 38.82°E	South-central plateau
7	Dekemhare	Southern/Debub	15.07°N, 39.05°E	South-central plateau
8	Segeneyiti	Southern/Debub	14.90°N, 39.15°E	East-central
9	Senafe	Southern/Debub	14.65°N, 39.52°E	Eastern escarpment
10	Tsorona	Southern/Debub	14.75°N, 39.40°E	Eastern escarpment
11	Adi Keyih	Southern/Debub	14.85°N, 39.38°E	Eastern escarpment
12	Areza	Southern/Debub	14.88°N, 38.62°E	South-western zone
13	May Mine	Southern/Debub	14.70°N, 38.52°E	South-western zone
14	Adi Kwala	Southern/Debub	14.65°N, 38.45°E	South-western zone
15	Kudo Bu'er	Southern/Debub	14.55°N, 38.40°E	South-western zone
16	Adi Teklezan	Anseba	15.70°N, 38.60°E	North-Central zone
17	Logo Anseba	Gash-Barka	15.50°N, 38.50°E	Central-West

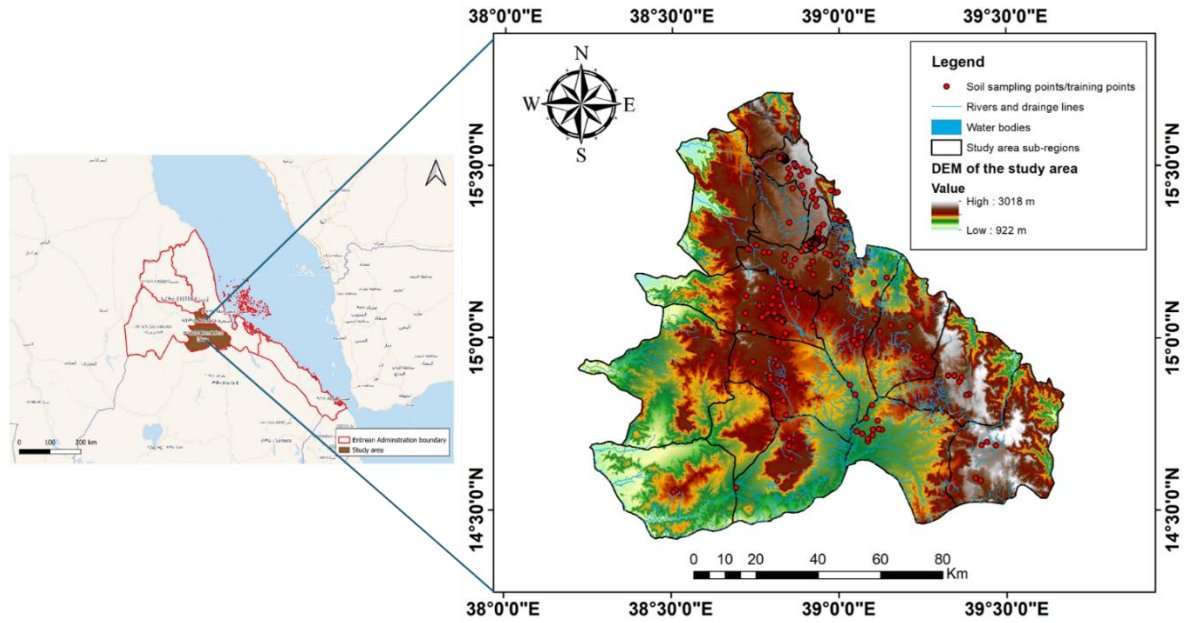


Figure 2. Geographical Location of the study area within Eritrea, showing the boundary of the Eritrean Highlands study area, the administrative regions (zobas) and sub-regions, and the DEM.

3.2.2. Geology and Soils

The geological foundation of the Eritrean Highlands is the Precambrian crystalline basement of the Arabian–Nubian Shield, an assemblage of Neoproterozoic metavolcanic and metasedimentary rocks, gneisses, schists, and granitic intrusions consolidated during the Pan-African orogeny approximately 600 million years ago. These basement units, which host the volcanogenic massive-sulphide mineralisation for which the Asmara region is known, are locally overlain by discontinuous Mesozoic sedimentary rocks (sandstones and limestones) and, in places, capped by Cenozoic (Oligocene–Miocene) Trap-series flood basalts associated with the uplift and rifting of the Afar dome. This lithological diversity, together with the strong relief, produces a heterogeneous distribution of parent materials and, in turn, of soils.

Eritrea does not yet possess a comprehensive national soil map or an established national soil-classification system, and regional and global compilations such as the Harmonized World Soil Database are therefore commonly used as a baseline (FAO, 2012; Tesfay et al., 2025). Within the Highlands, the soil cover is dominated, in World Reference Base terms, by Leptosols, Cambisols, Regosols, Fluvisols, and Vertisols, with Luvisols and Lixisols occurring locally. Their distribution is strongly controlled by topography and parent material: shallow, stony Leptosols and Regosols, with weak structure, low organic-matter content, and limited water-holding capacity, predominate on the steep upper slopes and are inherently susceptible to detachment and transport; moderately developed Cambisols occupy intermediate positions; and heavy, clay-rich Vertisols and Fluvisols accumulate on footslopes, depressions, and alluvial bottoms, where they are comparatively fertile but pose drainage and workability constraints. Soil organic-carbon contents across Eritrean land are generally low, reflecting long-term cultivation and grazing pressure (Teskay et al., 2025). Because the soil-erodibility (K) factor of the RUSLE is governed by the spatial variation of texture and OM, this pedological heterogeneity, combined with sparse sampling, is precisely the constraint that the machine-learning soil-property prediction described in is designed to address.

3.2.3. Climate and Seasonal Rainfall Regime

The climate of the Eritrean Highlands is best described as tropical-highland, semi-arid to sub-humid, with the altitude of the plateau strongly moderating temperatures that would otherwise be tropical at this latitude. Mean annual air temperature at Asmara is approximately 16 °C, with values across the Highlands typically ranging between about 10 and 30 °C; the coolest conditions occur during the December–February period and the warmest immediately before the onset of the main rains, around April and May. The relatively mild thermal regime is reflected in the temperature distribution shown for the wider region in the climatic characterisation of the study area.

Rainfall in the Highlands is markedly seasonal and is the single most important climatic control on erosion. The bulk of annual precipitation is delivered during the main rainy season, the *kremti*, which extends from June to September and is associated with the northward migration of the Inter-Tropical Convergence Zone (ITCZ); this is followed by a long dry season, with some southern districts also receiving a minor secondary rainfall peak (the *azmera*) around March and April. Mean annual rainfall over the Highlands is on the order of 400–700 mm, against a national average near 500 mm, but it is highly variable from year to year, and the region, lying on the southern margin of the Sahel, is recurrently affected by drought. Crucially for this study, *kremti* rainfall is predominantly convective and of high intensity, so that the erosive energy of the year is concentrated into a few months and, frequently, into a small number of intense storms. This strong intra-annual concentration of rainfall erosivity, occurring while protective vegetation cover is still establishing, is the central physical justification for the temporally explicit modelling framework adopted in this research and for the use of the high-resolution GPM IMERG satellite precipitation product described in.

3.2.4. Vegetation and Land use/Land cover

The potential natural vegetation of the Eritrean Highlands is Afromontane woodland and forest, dominated by *Juniperus procera* and *Olea* species on the plateau and by denser evergreen forest on the moister eastern escarpment, of which the Semenawi Bahri (Filfil) forests are the best-known remnant. Centuries of clearing for cultivation, grazing, and fuelwood, intensifying from the colonial period onward, have, however, removed the great majority of this cover, so that the contemporary landscape is a mosaic of cultivated land, grazing and grassland, shrub- and bushland, sparse and fragmented woodland, exposed rock and degraded bare ground, and settlements, with closed forest now confined to small protected or inaccessible patches.

From the standpoint of erosion modelling, the defining characteristic of this land cover is its strong seasonality. Vegetation greens rapidly with the onset of the *kremti* rains and senesces through the long dry season, so that the protective canopy is at its sparsest precisely at the beginning of the erosive wet season, when the first intense storms arrive. This seasonal, largely rainfall-driven vegetation dynamic, rather than a stable year-round cover, is what the monthly Sentinel-2 NDVI time series is used to capture in the cover-management (C) factor. Restoration efforts, notably the establishment of enclosures (area closures), terracing, and afforestation, are widespread across the Highlands and locally modify cover, but their effect on the broad land-cover pattern remains limited.

Land use across the study area is dominated by rain-fed smallholder agriculture and communal grazing, both of which extend frequently onto steep slopes where the physical conditions for accelerated erosion are most severe. The principal cultivated crops are sorghum, teff, wheat, barley, and legumes, grown under a mixed crop–livestock farming system in which draught animal tillage is universal and field boundaries are typically defined by stone walls or uncultivated margins of varying condition. Agricultural expansion onto slopes with gradients exceeding 20°, driven by population pressure and the progressive exhaustion of the most productive flat and gently sloping land, has been widely documented

as a major driver of land degradation and accelerated soil loss in the Eritrean highlands (Hurni, 1985a; Stillhardt et al., 2002). Communal grazing lands, which occupy the nonarable slopes between cultivated fields and the steeper escarpment faces, are frequently subjected to heavy livestock pressure that leads to vegetation cover reduction, surface compaction, and enhanced runoff generation, particularly during the critical early rainy season when soil surfaces are most bare and most vulnerable to detachment. Natural woodland and scrub, which would otherwise provide protective cover on erodible slope positions, are largely confined to small remnant patches on very steep or inaccessible terrain.

3.2.5. Topography and Drainage

The study area is among the most rugged terrain in the Horn of Africa. The plateau surface, lying mostly between about 1,800 and 3,018 m, is deeply dissected by an intricate network of gorges and valleys, producing a wide range of slope gradients, from near-level plateau remnants and footslope basins to very steep escarpment and gorge walls. This topographic configuration exerts a first-order control on erosion through the slope-length and steepness (LS) factor and through the concentration of overland flow in the dense drainage network, and it is represented in the modelling through a DEM.

Hydrologically, the Highlands form a continental divide. The eastern escarpment is drained by numerous short, steep, high-gradient streams that descend directly to the Red Sea coastal plain, whereas the western and southern slopes feed the larger river systems that flow toward Sudan and the Nile. Among the latter, the Mereb–Gash and the Setit ultimately contribute to the Nile system via the Atbara, while the Barka and Anseba drain toward Sudan without reaching the Nile; the Mereb River forms part of the southern boundary of the Highlands with Ethiopia. The combination of steep gradients, an efficient drainage network, and seasonally intense, high-magnitude flows promotes rapid runoff concentration and high sediment-transport capacity during the kremti season, reinforcing the erosion susceptibility of the area.

3.2.6. Agricultural Systems and Erosion Susceptibility

Agriculture in the Eritrean Highlands is overwhelmingly smallholder, rain-fed, and subsistence-oriented, organised around a mixed crop–livestock system. The dominant crops are cereals, principally sorghum, barley, wheat, finger millet, teff, and maize, complemented by pulses and oilseeds, with cropping calendars tightly tied to the kremti rains; after harvest, crop residues are typically grazed by livestock, and inputs such as fertiliser and irrigation are limited. Because the Highlands are the most densely populated part of the country, sustained population pressure has driven the cultivation of increasingly steep and marginal slopes, the shortening of fallow periods, and the continued clearing of vegetation, all of which heighten exposure to erosion.

The study area is therefore acutely susceptible to water erosion, owing to the convergence of several reinforcing factors: erosive, high-intensity rainfall concentrated in a short wet season; steep and deeply dissected terrain; shallow, weakly structured, and organic-matter-poor soils on the upper slopes; sparse protective vegetation at the onset of the rains; and an intensive, expanding cultivation footprint with a long history of land-cover removal. The resulting sheet, rill, and gully erosion is directly comparable to that documented across the adjoining Ethiopian Highlands, where annual soil loss has been estimated at the order of 1.5 billion tonnes, roughly half of it from cropland (Hurni, 1985a; Nyssen et al., 2004b). On-site, this erosion depletes soil fertility and undermines food security; off-site, it drives sedimentation of downstream reservoirs and water bodies. SWC measures, including bench and hillside terraces, stone and soil bunds, check dams, and area enclosures, are practised across the Highlands; however, spatially explicit information on the location and condition of these support practices is largely unavailable. It is this combination of high, temporally concentrated erosion risk and pervasive data scarcity that defines the problem the present study addresses.

3.2.7. Rationale for study area selection

The selection of the central and southern Eritrean highlands as the study area is justified on both scientific and practical grounds. Scientifically, the study area is representative of the broader class of semi-arid to sub-humid highland environments in the Horn of Africa that are characterised by the convergence of high rainfall erosivity, steep dissected terrain, shallow erodible soils, and intensive smallholder land use, conditions that together produce some of the highest soil erosion rates documented in Africa (Borrelli et al., 2017). The multi-basin hydrological structure of the study area, spanning the Mereb-Gash, Gash-Barka, Anseba, and Red Sea coastal drainage systems, enables assessment of RUSLE-modelled erosion across a range of hydrologically distinct landscape contexts and facilitates the evaluation of spatial variability in erosion risk at a scale relevant to watershed management planning. The pronounced orographic rainfall gradient across the study area, combined with strong within-year seasonality in both precipitation and vegetation cover, makes the study area particularly well-suited to evaluating the two principal methodological innovations of this dissertation: the GPM IMERG monthly R-factor simulation, which requires meaningful spatial and temporal variability in precipitation to be demonstrably superior to annual-average approaches, and the full 12-month Sentinel-2 NDVI C-factor time series, which requires pronounced seasonal vegetation dynamics to capture the critical June erosion window. Practically, the availability of NARI soil profile observations and rainfall station records within the study area provides the training and validation data required for the ML K-factor prediction and GPM IMERG validation components of the methodology, while the growing accessibility of GPM IMERG and Sentinel-2 data for this national territory ensures that the methodological framework developed here can be updated and applied routinely in future monitoring exercises (OCHA, 2023; Stillhardt et al., 2002).

3.3. Datasets

Four primary datasets underpin the modelling framework: GPM IMERG satellite precipitation for the rainfall-erosivity factor; Sentinel-2 imagery for the cover-management factor; a DEM for the topographic factor and terrain covariates; and 302 field soil samples used to train a machine-learning prediction of soil properties for the erodibility factor. Table 2 summarises their principal characteristics and roles, and the following subsections describe them.

Table 2. Summary of the primary datasets used in the study.

Dataset	Source	Resolution	Period	Use in RUSLE
GPM IMERG-FR V07	NASA GES DISC	0.1°; 30-min (monthly aggregates)	2021	Rainfall erosivity (R)
Sentinel-2 MSI (L2A)	Copernicus / ESA	10 m; ~5-day revisit	2021	NDVI → cover management (C) and Environmental Covariates
DEM	SRTM	~30 m	Static (time-invariant)	LS-factor; ML covariates
Soil samples (302 points)	MoA / NARI	Point observations	Field campaign	(ML): Sand, silt, clay, OM → K-factor

3.3.1. GPM IMERG (IMERG-FR V07) Rainfall

Rainfall information was obtained exclusively from the Global GPM IMERG, specifically the gauge-adjusted Final Run (IMERG-FR), Version 07 product. Developed by NASA, IMERG intercalibrates and merges precipitation retrievals from the passive-microwave sensors of the GPM constellation,

anchored by the dual-frequency precipitation radar and microwave imager of the GPM Core Observatory, with geostationary infrared observations, to produce precipitation estimates at 0.1° spatial and half-hourly temporal resolution with near-global coverage (Huffman et al., 2020). The integration of multiple sensors provides a spatially continuous and temporally consistent representation of rainfall, which is particularly advantageous in regions such as the Eritrean Highlands where ground-based gauges are sparse and unevenly distributed.

For this study, IMERG-FR V07 data were obtained for all twelve months of 2021 and aggregated to monthly precipitation totals. In the simulated stream, each monthly total was annualised and converted to a monthly R-factor, whereas in the conventional stream the monthly totals were summed to an annual total and converted to a single annual R-factor. The exclusive use of a single, internally consistent satellite product avoids the heterogeneity that would arise from merging disparate gauge records and provides complete spatial coverage of the study area.

The IMERG-F data were accessed and downloaded through NASA Giovanni, an online data visualisation and analysis system developed by NASA Goddard Earth Science Data and Information Services Centre (GES DISC). Giovanni enables spatiotemporal subsetting, averaging, and export of satellite-based geophysical datasets in user-defined formats, including GeoTIFF. The Giovanni platform is available at: <https://giovanni.gsfc.nasa.gov/giovanni/>.

3.3.2.Environmental Covariates

Environmental covariates for machine learning-based prediction of soil texture (sand, silt, clay) and organic matter were derived from two primary data sources: Sentinel-2 L2A multispectral imagery and the SRTM digital elevation model, following the SCORPAN framework (Padarian et al., 2020). All covariates were prepared at 30 m spatial resolution and projected to a common coordinate system consistent with the study area boundary.

Sentinel-2

In 2021, Sentinel-2 L surface reflectance images were compiled into four seasonal composites, spring (February–April), summer (June–August), autumn (September–November), and winter (December–February), to capture surface reflectance changes throughout the year while minimising cloud cover and vegetation effects. For each season, spectral bands and indices were calculated and used as predictor variables: near-infrared (NIR, Band 8, 842 nm), indicating soil organic matter and moisture (Joshi et al., 2016); shortwave infrared bands (SWIR1, Band 11, 1,610 nm; SWIR2, Band 12, 2,190 nm), aiding in clay mineral detection (Chabhadiya et al., 2015) the Normalized Difference Vegetation Index (NDVI); (Rouse, W. and Haas, R.H. (1974) representing vegetation cover and land-use intensity; the Bare Soil Index (BSI; (A. RIKIMARU, 2002), distinguishing bare soil from vegetation and built-up areas; and the Clay Mineral Index (CMI = SWIR1/SWIR2), providing estimates of kaolinite and illite minerals (Drury, 1987). All four seasonal composites served as individual predictor variables, offering the machine learning models detailed spectral information throughout the year. Vegetation cover and the wider land-surface condition were characterised using imagery from the Copernicus Sentinel-2 mission, a twin-satellite constellation carrying the MultiSpectral Instrument (MSI) and providing high-spatial-resolution optical data across thirteen visible-to-shortwave-infrared bands with a combined revisit interval of approximately five days (Drusch et al., 2012). Because both vegetation vigour and the degree of bare-soil exposure vary strongly through the year, each covariate was computed as a per-season median composite, for winter, spring, summer, and autumn, from cloud-masked Sentinel-2 imagery processed on a cloud-based platform (Gorelick et al., 2017).

Land Surface Temperature (LST)

LST was derived from Landsat 8 thermal infrared imagery (Band 10) for each of the four seasonal windows using the split-window algorithm (Jimenez-Munoz et al., 2014). LST serves as a proxy for soil thermal properties, moisture regime, and organic matter content, and has been widely used as a predictor variable in machine learning digital soil mapping studies (Hengl et al., 2017). Seasonal LST composites (LST_Spring, LST_Summer, LST_Autumn, LST_Winter) were retained as separate predictor variables to represent the study area's annual thermal regime. As with the Sentinel-2 spectral variables, LST was computed as a set of per-season median composites, for winter (December–February), spring (March–May), summer (June–August), and autumn (September–November), from cloud-masked imagery processed on a cloud-based platform (Gorelick et al., 2017).

SRTM DEM Terrain Covariates

Terrain attributes were derived from a 30 m resolution hydrologically conditioned SRTM digital elevation model, processed with SAGA GIS (Conrad et al., 2015). The included terrain covariates comprised slope gradient (degrees), which influences surface runoff speed and soil redistribution on hillslopes (Moore et al., 1991) ; aspect (degrees), affecting solar radiation, evapotranspiration, and soil moisture distribution; the Topographic Position Index (Weiss, A. (2001)),differentiating ridgelines, mid-slopes, and valley floors; profile and plan curvature, which indicate flow acceleration and lateral water redistribution (Zevenbergen & Thorne, 1987); flow accumulation, representing upslope contributing area and subsurface water pathways; and Terrain Ruggedness Index (Riley, S.J., De Gloria, S.D. and Elliot, R. (1999) reflecting local topographic heterogeneity important for soil development. These covariates collectively describe relief and parent material features within the SCORPAN framework.

3.3.3. Soil Sampling and Laboratory Analysis

The soil data used in this study were obtained from the (MOA) through the Department NARI. Soil samples were initially collected from various locations across the study area as part of soil survey and agricultural research activities carried out by the institute. A total of 302 georeferenced composite soil samples were collected from the 0–30 cm topsoil horizon across the study area following a stratified random design (Figure 3). These sampling locations represent a range of environmental conditions, land use types, and terrain features within the study area. The complete dataset of 302 soil sampling results is presented in tabular form in the Appendix, providing full transparency of the primary field data underpinning the machine learning-based soil property predictions and the subsequent K-factor estimation.

During field campaigns, soil samples were collected following soil sampling procedures. At each sampling location, representative soil samples were obtained and properly labeled to ensure accurate identification and traceability. The geographic location of each sampling point was recorded using geographic coordinates (latitude and longitude), enabling accurate spatial positioning of the samples within the study area. The availability of georeferenced soil data facilitates integration with GIS and RS datasets for spatial modeling and environmental analysis (McBratney et al., 2003).

The collected soil samples were then transported to NARI's soil laboratory for analysis. In the lab, the soil particle-size distribution was measured to determine the proportions of sand, silt, and clay in each sample. NARI mainly uses the hydrometer method for particle-size analysis. This technique relies on sedimentation, in which soil particles suspended in a liquid settle at different rates based on their size. By recording the density of the soil suspension at specific times with a hydrometer, the relative percentages of sand, silt, and clay can be calculated (Bouyoucos, 1962; Gee, G. W. and Bauder, 1986).

Soil texture, obtained from the particle-size distribution, is essential in influencing soil physical properties such as permeability, water retention, and erosion susceptibility.

In addition to soil texture analysis, the OM content of the soil samples was determined using the Walkley-Black wet oxidation method. This method estimates soil organic carbon by chemical oxidation with potassium dichromate and sulfuric acid, and the resulting organic carbon is then converted to OM content. The Walkley-Black method is one of the most widely used laboratory techniques for estimating OM due to its reliability and relative simplicity (Nelson D.W, 1982; Walkley et al., 1934.). OM is an important indicator of soil quality and plays a crucial role in improving soil structure, increasing aggregate stability, and enhancing resistance to soil erosion.

The laboratory analyses provided quantitative measurements of sand, silt, clay, and OM for each of the 302 sampling locations. After completion of laboratory analysis and quality control procedures, the dataset was made available by NARI and subsequently obtained for use in the present study. The measured soil parameters served as the primary ground-referenced data used to calculate the K-factor at sampling locations using an empirical equation. Furthermore, the soil properties obtained from the laboratory analysis were used as dependent variables in machine-learning regression models to predict soil characteristics at unsampled locations across the study area.

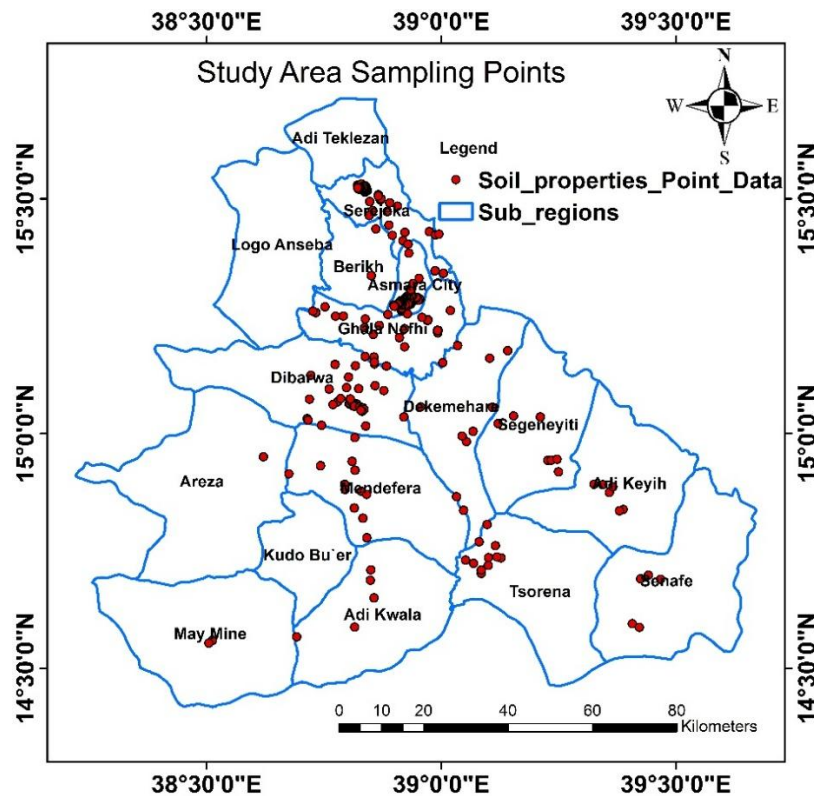


Figure 3. Soil Properties point data along with the sub-regions map

After quality-control removal of samples with individual fractions outside the 1st–99th percentile range and samples whose texture fractions did not sum to $100 \pm 2\%$, we retained a final dataset of 277 samples for all texture and OM modelling (Table 3).

Table 3. Descriptive statistics of measured soil texture fractions ($n = 277$).

Property	n	Mean (%)	Std Dev (%)	Min (%)	Median (%)	Max (%)	CV (%)
Silt	277	33.5	9.6	9.0	32.8	55.1	28.7
Sand	277	46.9	16.1	14.5	47.3	86.4	34.3
Clay	277	19.7	14.0	1.7	16.0	57.6	71.1

At each location, the soil particle-size fractions (sand, silt, and clay) and the organic-matter content were measured; these four properties are the inputs required by the EPIC/Williams formulation of the soil-erodibility (K) factor (Williams J. et al., 1981). Figure 3 shows the spatial distribution of the sampling points.

Although these observations provide valuable ground information, their spatial density is insufficient for reliable conventional interpolation across the steep and heterogeneous terrain of the Eritrean Highlands, where soil properties vary over short distances in response to topography and parent material. For this reason, the 302 points were used not as direct interpolation nodes but as the target (training and validation) dataset for the machine-learning regression models, which predict each soil property continuously across the landscape from DEM-derived terrain attributes and satellite-based environmental covariates. This strategy converts the sparse point dataset into spatially detailed soil-property surfaces and, in turn, into a spatially robust K-factor, addressing the data-scarcity limitation that constrains erodibility estimation in the region (Hengl et al., 2017).

3.4. Methods

3.4.1. Machine-learning prediction of soil properties and K-factor

The soil-erodibility (K) factor of the RUSLE is a function of soil texture and organic-matter content, and its reliable computation therefore requires spatially continuous surfaces of the sand, silt, clay, and organic-matter fractions across the study area. The 302 field-measured samples are too sparse and too unevenly distributed for reliable conventional interpolation over the steep and heterogeneous terrain of the Eritrean Highlands. A digital soil mapping (DSM) approach was therefore adopted, operationalising the SCORPAN (Soil, Climate, Organisms, Relief, Parent material, Age, Nearby soils) conceptual model (McBratney et al., 2003), in which each soil property is expressed as an empirical function of environmental covariates that represent the soil-forming factors and can be observed exhaustively across the landscape. The K-factor was derived in two stages: (i) machine-learning spatial prediction of the soil properties from remote-sensing and terrain covariates; and (ii) computation of K from the predicted properties using the EPIC/Williams erodibility equation. All processing was carried out at 30 m resolution in the WGS84 / UTM Zone 37N projection (EPSG:32637), using Google Earth Engine for satellite covariate preparation (Gorelick et al., 2017) and a Python (XGBoost, scikit-learn, rasterio) environment for model training and spatial prediction.

3.4.1.1. Predictor Variables (Terrain Attributes and Environmental Covariates)

The predictor set assembled for the models draws on the three covariate sources described in the preceding sections and is summarised in Table 4.

Table 4. *Environmental covariates used as predictor variables for soil-property prediction.*

Covariate group	Variables	Source	Layers
Sentinel-2 spectral bands	RE1 (B5), RE2 (B6), NIR (B8), SWIR1 (B11), SWIR2 (B12)	3.3.2	5 × 4 seasons
Sentinel-2 indices	NDVI, BSI	3.3.2	2 × 4 seasons
Land surface temperature	LST (seasonal)	3.3.3 (Landsat-8)	1 × 4 seasons
Terrain derivatives	elevation, slope, aspect, curvature, TWI, relief, TPI	3.3.4 (SRTM)	7
Continental soil priors	iSDAsoil, SoilGrids texture (exploratory)	reference products	6

The spectral covariates comprise the Sentinel-2 red-edge, near-infrared, and shortwave-infrared reflectance bands together with the NDVI and BSI, each computed as a per-season median composite; the thermal covariate is the Landsat-8 land surface temperature, likewise composited by season; and the topographic covariates are the seven SRTM-derived terrain attributes. Together, these yield an initial stack of thirty-one co-registered 30 m raster layers spanning the climate, organism, relief, and parent-material factors of the SCORPAN model (Hengl et al., 2017; Mulder et al., 2011).

Covariate values were extracted at each of the 302 sampling locations using a 7×7 pixel neighbourhood window (a 210 m ground footprint), the neighbourhood mean being taken as the representative value; this window size was selected after systematically evaluating windows from 3×3 to 11×11 pixels, and it minimises the effect of GPS positional error and pixel-scale spectral noise. A set of engineered features was then derived from the raw covariates on the basis of spectral and geomorphological reasoning, including thermal–mineral interaction terms (products of LST and SWIR reflectance), the CMI ($\text{CMI} = \text{SWIR1}/\text{NIR}$) exploiting clay-mineral absorption near 1610 nm, seasonal means and differences of NDVI, BSI, and LST that encode phenological amplitude, and vegetation–topography terms (e.g., $\text{NDVI} \times \text{relief}$, log-transformed relief). Recursive feature elimination guided by Random-Forest importance ranking reduced the pool to a parsimonious set of fifteen features, beyond which additional predictors did not improve cross-validated accuracy; the thermal–mineral interaction terms and the CMI consistently emerged as the most informative predictors.

Sentinel-2 L2A composites were created for all four seasons, spring, summer, autumn, and winter to reduce soil spectral signature contamination from vegetation canopy cover. The spring composite offers the clearest bare-soil observations when vegetation is minimal. The other season composites serve as covariates to capture seasonal variations in soil spectral features, improving the data for machine learning models. This multi-season approach follows strategies by (Diek et al., 2017) to minimise vegetation interference in satellite-based soil property predictions.

3.4.1.2. Machine-Learning Regression Models, Training, and Tuning

The predictive model was implemented using Extreme Gradient Boosting (Chen & Guestrin, 2016), a regularised tree-boosting algorithm that builds an additive ensemble of shallow regression trees, each fitted to the residuals of its predecessors by gradient descent on a penalised squared-error loss. XGBoost was selected after a comparative cross-validation of several candidate algorithms, including RF (Breiman, 2001), LightGBM (LGB), Support Vector Regression with a radial-basis-function kernel, and Kernel Ridge Regression; it achieved the highest mean cross-validated accuracy across the target fractions. Its L1 and L2 regularisation, column and row subsampling, and minimum-child-weight constraints make it particularly resistant to overfitting in the small-sample regime ($n = 277$) that characterises this study, a setting in which tree-based ensembles are known to outperform deep-learning

architectures that typically require several hundred to a thousand training observations to generalise (Behrens et al., 2018; Grunwald et al., 2024).

The 277 georeferenced topsoil samples were used to train five ensemble tree-based algorithms, Extra Trees (ET), RF, Gradient Boosted Trees (GBT), XGB, and LGB, for each of the three texture fractions, and the best-performing model per fraction was selected on the basis of random five-fold cross-validation. Model behaviour was interrogated using SHapley Additive exPlanations (SHAP), the reliability of the predictions was examined through observed-versus-predicted, stratified, and residual diagnostics, and the honesty of the accuracy estimates was tested against spatial block cross-validation.

Hyperparameters were optimised by an exhaustive grid search under five-fold cross-validation, using clay, the most predictable fraction, as the optimisation target, and the resulting configuration was applied uniformly to the other properties. The combination of shallow trees (maximum depth of three), a low learning rate (0.01) compensated by a large number of estimators (500), and L2 regularisation ($\lambda = 1.5$) was found to be essential for controlling variance and preventing overfitting given the limited sample size. The final models were refitted on all 277 samples prior to spatial prediction, which was performed pixel-by-pixel across the AOI in batches; the three fraction predictions were then renormalised at each pixel to enforce compositional closure.

Organic matter content was estimated using a separate model selection process from that for soil texture fractions, reflecting distinct spectral and environmental influences on organic matter distribution. Both were predicted using the same set of spectral, thermal, and terrain variables, including seasonal Sentinel-2 spectral indices, Landsat land surface temperature composites, and SRTM terrain derivatives. Six machine learning algorithms were tested: Extra Trees (ET), Random Forest (RF), XGBoost (XGB), LightGBM (LGB), Gradient Boosted Trees (GBT), and a Stacking Ensemble (STK) that combined the predictions of the five base models. All models were evaluated using the same 5-fold cross-validation on 277 soil profiles. The best model was selected based on R^2 , RMSE, MAE, bias, and RPD, and then applied to the full covariate raster stack to produce a continuous organic matter map at 30 m resolution. This map and the predicted soil texture fractions served as input to the EPIC/Williams K-factor equation. Cross-validation outcomes for all six algorithms are summarised in Table 8, with the selected model discussed in Section 4.1.

3.4.1.3. Model Validation Strategy

Model performance was assessed by stratified five-fold cross-validation with random shuffling (fixed seed for reproducibility). In each fold, the model was trained on approximately 80% of the samples and evaluated on the held-out 20%, and out-of-fold predictions were assembled across the five folds to yield an independent prediction for every one of the 277 samples without information leakage. Four complementary metrics were computed: the cross-validated coefficient of determination ($CV\text{-}R^2$), the root-mean-square error (RMSE), the mean absolute error (MAE), and the ratio of performance to deviation ($RPD = \text{standard deviation of observations} / \text{RMSE}$). Following the widely used classification of (Chang et al., 2001), RPD values below 1.4 indicate models suitable only for approximate spatial-trend mapping, values between 1.4 and 2.0 indicate quantitatively useful prediction, and values above 2.0 indicate reliable quantitative accuracy. Prediction bias was quantified as the mean signed residual, and fold-wise distributions of $CV\text{-}R^2$ and RMSE were examined to assess stability. The remote-sensing-based predictions were additionally benchmarked against the continental iSDAsoil and SoilGrids products to contextualise regional performance and to evaluate the incremental value of the local model over globally available priors.

3.4.1.4. Computational Environment and Software Implementation

The modelling framework was implemented across three complementary computational environments, each selected for the class of task to which it is best suited: a cloud-based platform for the acquisition and compositing of archive-scale satellite imagery; a desktop Geographic Information System for terrain analysis; and a scripted scientific-computing environment for feature engineering, machine-learning modelling and spatial prediction. This division of labour was a deliberate design decision rather than a matter of convenience. Multi-year, multi-sensor image compositing is data-intensive but computationally shallow, and is therefore best executed where the imagery already resides; terrain analysis requires validated, well-documented geomorphometric algorithms; and model training and prediction require full scripting control, reproducibility, and access to contemporary machine-learning libraries. The complete software environment is summarised in Table 5.

(a) *Cloud-based satellite data processing (Google Earth Engine)*

Google Earth Engine (GEE) is a planetary-scale geospatial analysis platform that co-locates petabyte-scale satellite archives with parallel computational infrastructure, allowing image collections to be filtered, masked, and reduced without local download of the source imagery (Gorelick et al., 2017). GEE was used for all operations requiring access to large multi-temporal image collections, namely: (i) filtering the Sentinel-2 Surface Reflectance Harmonised collection (COPERNICUS/S2_SR_HARMONIZED) and the Landsat-8 Collection 2 Level-2 archive (LANDSAT/LC08/C02/T1_L2) by date range and scene cloud cover; (ii) per-pixel cloud and cirrus masking using the Scene Classification Layer and the QA60 bitmask for Sentinel-2, and the Collection 2 quality-assessment bands for Landsat-8; (iii) computing seasonal median composites for each target season across the multi-year archive; (iv) deriving the spectral indices (NDVI, BSI) and the land-surface-temperature composites within the platform; and (v) exporting all resulting layers, clipped to the study-area boundary, as Cloud-Optimised GeoTIFFs.

The use of a cloud platform was not merely expedient but necessary: retrieving the full multi-year Sentinel-2 archive covering the study area at native resolution would have required on the order of one hundred terabytes of local storage, which is infeasible on the available hardware. Two methodological choices within GEE merit explicit statement. First, seasonal composites were produced with a median reducer rather than a mean reducer, because the median is markedly more resistant to residual cloud contamination, cloud shadow, and radiometric outliers, and because mean compositing of overlapping orbit tracks acquired on different dates introduces visible strip-boundary discontinuities. Second, all exports were clipped to a single study-area geometry stored as a platform asset, guaranteeing that every exported layer shares an identical spatial extent.

(b) *Terrain analysis using ArcGIS Pro*

Terrain derivatives were computed exclusively in ArcGIS Pro from the SRTM DEM. The DEM was first hydrologically conditioned and aligned to the common reference grid, after which slope, aspect, curvature, the Topographic Wetness Index (computed with a D-infinity flow-routing algorithm), relief intensity, and the Topographic Position Index were generated using the standard surface-analysis and neighbourhood-statistics toolsets. A dedicated desktop GIS was preferred for this stage because its geomorphometric algorithms are validated, fully documented, and directly citable, which is important for attributes such as curvature and flow routing whose numerical implementation materially affects the resulting surfaces. All terrain outputs were exported to the common grid and co-registered with the satellite covariates.

(c) **Modelling, prediction and analysis**

All post-acquisition processing was implemented in Python within a notebook environment, comprising covariate extraction at the 302 sampling locations using a 7×7 neighbourhood window, feature engineering, model training and cross-validation, batch spatial prediction across the full study-area grid, artefact detection and correction, and the production of all figures. Table 5 lists the libraries used and their specific roles. In brief, raster input/output and pixel-value extraction were handled with rasterio; array computation, neighbourhood filtering, and gap-filling with NumPy and SciPy; tabular data management with pandas and coordinate transformations with geopandas; cross-validation, performance metrics, and feature-importance ranking with scikit-learn; and cartographic and diagnostic figures with matplotlib and seaborn. Spatial prediction over the study-area grid was executed in sequential batches of pixels using vectorised array operations, a necessary measure given that the prediction grid contains on the order of thirteen million valid pixels and the available memory was limited.

Table 5. Software and computational environment used in the study.

Environment/library	Version	Role in the workflow
Google Earth Engine	Cloud	Sentinel-2 and Landsat-8 filtering, cloud masking, seasonal median compositing, index and LST computation, export
ArcGIS Pro and ArcGIS 10.4.1	3.2	SRTM terrain derivatives: slope, aspect, curvature, TWI (D-infinity), relief intensity, TPI
Python	3.11	Feature extraction and engineering, model training, spatial prediction, figures
rasterio	1.3	Raster I/O, resampling, pixel-value extraction, GeoTIFF export
NumPy / SciPy	1.24 / 1.10	Array computation, neighbourhood filtering, gap-filling, artefact detection
pandas	2.0	Tabular data handling and merging of samples with covariates
geopandas	0.13	Coordinate reference system transformation of sample points
scikit-learn	1.3	Cross-validation, performance metrics, feature-importance ranking
matplotlib / seaborn	3.7 / 0.12	Diagnostic plots and cartographic figures

3.4.1.5. Reference Spatial Grid, Data Integration and Reproducibility

Because the modelling framework combines layers originating from different sensors, spatial resolutions, and processing environments, a single reference spatial grid was defined at the outset and imposed on every raster product. All layers were projected to the WGS84 / UTM Zone 37N coordinate reference system (EPSG:32637) and resampled to a common 30 m grid, yielding a raster of 4,777 rows by 5,198 columns and approximately 13.27 million valid pixels within the study-area boundary. The 30 m resolution was adopted because it corresponds to the native resolution of the SRTM DEM and the Landsat-8 thermal product, and therefore represents the coarsest of the primary inputs; upsampling the coarser layers to the 10 m resolution of the Sentinel-2 visible and near-infrared bands would have produced an artificial impression of spatial detail not supported by the underlying data. Sentinel-2 layers, whose native resolution is 10 m or 20 m depending on the band, were resampled to the 30 m reference grid using bilinear interpolation, and all continuous layers were resampled with bilinear interpolation while categorical layers, where used, were resampled by nearest-neighbour assignment.

Exact co-registration of all layers on this common grid is a prerequisite both for the pixel-wise extraction of covariates at the sampling locations and for the cell-by-cell map algebra used to combine the RUSLE factors.

Integration between the three computational environments was accomplished through a structured file-exchange protocol rather than through direct programmatic coupling. Raster products exported from the cloud platform were written to a dedicated cloud-storage folder and subsequently retrieved to the local working directory, where the terrain derivatives produced in the desktop GIS were also deposited. Filenames followed a standardised convention encoding the variable, the study area, and the season, which allowed the analysis scripts to construct the covariate stack automatically from a catalogue of file paths rather than through manual specification, thereby eliminating a common source of error in multi-layer workflows. All rasters were held as single-precision floating-point arrays during processing to limit memory consumption on the thirteen-million-pixel grid, and final outputs were written as compressed GeoTIFFs with an explicit no-data value assigned to pixels outside the study-area boundary.

The workflow was designed to be reproducible. The cloud-platform scripts, the complete analysis code, the model hyperparameters and cross-validation results, and the raster-calculator expressions used for the RUSLE factor combination are provided in the appendices, and the analysis code is parameterised such that it can be re-executed on a different machine by modifying a single base-path variable. Fixed random seeds were used throughout for all stochastic operations, including the cross-validation fold partitioning and the model fitting, so that the reported performance metrics and the prediction surfaces are exactly reproducible.

3.4.2. RUSLE Model factors

This section describes the derivation of all RUSLE input factors employed in the study. The K-factor was estimated through a combination of machine learning-based soil property prediction and an empirical EPIC function, while the remaining factors, R-factor, LS-factor, and C-factor, were derived using established empirical and remote sensing methods. Of these four factors, the R-factor and C-factor are temporally dynamic and were therefore computed at two temporal scales: monthly for the simulated framework and annually for the conventional framework. The LS-factors and K-factor are temporally invariant and were computed once, shared by both frameworks. All factors were generated on a common 30 m spatial reference grid to enable direct cell-by-cell map algebra integration, as expressed in Equation 3.1 and are described in detail in the following sections.

$$A = R \cdot K \cdot LS \cdot C \cdot P \quad (3.1)$$

A is annual soil loss ($t/ha/yr$), R is the Rainfall erosivity ($MJ \text{ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$), K is soil erodibility ($t \cdot ha \cdot h \cdot ha^{-1} \cdot MJ^{-1} \cdot mm^{-1}$), LS is the slope length and steepness factor, C is the cover-management factor (0 to 1), P is the management practice (0 to 1).

The following units are used throughout this study: t denotes tonnes, ha denotes hectares, h denotes hours, MJ denotes megajoules, mm denotes millimetres, yr denotes year, m denotes metres, consistent with the standard unit conventions of RUSLE.

3.4.2.1. Rainfall Erosivity factor (R-factor) - GPM IMERG Monthly Simulation

3.4.2.1.1. R- Factor Equation

The R-factor represents the climatic driver of soil erosion and quantifies the potential of rainfall to cause soil loss through the combined effects of raindrops' impact (detachment) and surface runoff (transport) (Morgan, 2005). Within the RUSLE framework, the R-factor represents the erosive energy supplied by precipitation and is among the most influential parameters controlling predicted soil loss.

Rainfall is primarily governed by rainfall intensity and duration, which determine the kinetic energy of raindrops and the volume and velocity of runoff generated during storm events (Wischmeier, W. H., & Smith, 1978). On bare, sparsely vegetated surfaces, raindrop or splash erosion is often the dominant erosion mechanism, as raindrop impacts disintegrate particles and initiate their downslope transport via overland flow (Tiruneh et al., 2025). Consequently, rainfall events characterised by high intensities over short durations contribute disproportionately to total soil erosion.

In its original formulation, the R-factor is calculated as the long-term average of the storm erosivity index (EI_{30}), defined as the product of the total energy of the rainfall event and its maximum 30-minute rainfall intensity (I_{30}). However, this method requires high-temporal-resolution rainfall records, which are generally unavailable in many regions, including the Eritrean highlands, due to sparse rain gauge networks.

In ungauged or data-limited watersheds, rainfall erosivity is commonly estimated using empirical relationships based on mean annual precipitation, derived either from ground-based observations or spatially continuous precipitation datasets (Kenneth G. Renard, 1994). These approaches have been widely applied in semi-arid and highland environments and provide a practical alternative for regional-scale erosion modelling.

In the study, the R-factor was estimated using an empirical precipitation-based equation (Equation 3.2), originally developed by (Wischmeier, W. H., & Smith, 1978), subsequently modified for the Eritrean and Ethiopian highlands by (Hurni, 1985b). The equation relates rainfall erosivity directly to mean annual precipitation (p), thereby enabling R-factor values to be calculated in the absence of rainfall-intensity data.

$$R = -8.12 + (0.562 \times P) \quad (3.2)$$

Where R is the rainfall erosivity factor ($\text{MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$), and P is the mean annual precipitation (mm).

To ensure spatial consistency and mitigate issues associated with the sparse distribution of rain gauge stations, the monthly and mean annual precipitation (P) was derived from the GPM IMERG-FR (v07) satellite precipitation prod. These spatially continuous precipitation surfaces were then used as input to the R-factor. The R-factor was calculated on a pixel-by-pixel basis by applying Equation (3.2) to the IMERG-derived mean annual precipitation raster using the Raster Calculator tool in ArcGIS version 10.4.1. This GIS-based approach allowed the spatial variability in precipitation captured by IMERG datasets to be directly translated into a continuous rainfall-erosivity surface across the study area

3.4.2.1.2. Conventional (Annual Summation) Benchmark Approach

As a reference benchmark, the annual R-factor was computed by summing the 12 monthly GPM IMERG precipitation rasters to obtain mean annual precipitation (P_{annual}), and applying the regression equation above. This yielded a single static annual R-factor raster:

This conventional R_{annual} value was used in a parallel RUSLE computation alongside a mean annual C-factor (average of monthly C values) to produce the benchmark annual soil loss estimate ($A_{\text{non-sim}}$).

3.4.2.1.3. Monthly Simulated Approach

To capture the intra-annual variability in rainfall erosivity, the following 12-step monthly simulation procedure was implemented in ArcGIS 10.4.1 Raster Calculator for each calendar month m ($m = 1$ to 12):

1. Extract the mean monthly GPM IMERG precipitation raster for month m : P_m (mm month^{-1}). These rasters had already been resampled to 30 m and clipped to the study area.
2. Compute a simulated annual precipitation equivalent by scaling monthly precipitation to an annual total: $P_{\text{sim}_m} = P_m \times 12$. This step generates a raster representing the annual precipitation that would be expected if the rainfall rate observed in month m were sustained for the entire year.
3. Compute the simulated monthly R-factor by applying the regression equation to P_{sim_m} : $R_{\text{sim}_m} = a \times P_{\text{sim}_m} + b$. This yields a spatially explicit simulated R-factor raster for each month.
4. Compute the simulated annual soil loss for month m using the RUSLE equation with the corresponding monthly C-factor (C_m), static K, LS, and P: $A_{\text{sim}_m} = R_{\text{sim}_m} \times K \times LS \times C_m \times P$.
5. Normalise the simulated annual soil loss to the actual monthly contribution by dividing by 12: $A_{\text{actual}_m} = A_{\text{sim}_m} / 12$.
6. Repeat steps 1–5 for all 12 months.
7. Sum the 12 monthly potential soil loss rasters to obtain the total actual annual soil loss: $A_{\text{annual}} = \sum (m=1 \text{ to } 12) A_{\text{actual}_m}$.

3.4.2.2. K-factor

The soil erodibility factor K quantifies the susceptibility of a soil to detachment and transport by rainfall and surface runoff, normalised per unit of erosivity (rainfall–runoff energy). K encapsulates the combined effects of soil texture, OM content, structure, and permeability on both the detachability of soil aggregates and the transportability of detached particles (Wischmeir, W. H. and Smith, 1978). Higher K values indicate soils that are more easily eroded per unit of applied erosive force, typically fine-textured, low-organic-matter, poorly aggregated soils.

The EPIC/Williams formulation used in this study captures the textural and organic matter components of soil erodibility, both of which are effectively stable on the within-year to decadal timescale of this study. Dynamic structural properties, including soil compaction, aggregate stability, surface sealing, and biological crusting, that can influence erodibility at shorter timescales (seasonal to interannual) are not represented. These properties require intensive temporal monitoring for spatial mapping at regional scale and represent a recognised limitation of RUSLE-based K-factor estimation. The K-factor presented here therefore reflects a structural baseline erodibility rather than a fully dynamic seasonal value.

3.4.2.2.1. Selection of the EPIC/Williams Formulation

This study adopts the empirical K-factor formulation developed by (Williams J. et al., 1981) within the EPIC (Environmental Policy Integrated Climate) model framework, rather than the original Wischmeier–Smith nomograph. Four considerations motivate this choice:

1. Data requirements: The Wischmeier–Smith nomograph requires soil structure code and soil permeability class in addition to texture and OM, categorical variables that are unavailable at the spatial resolution of this DSM study and are not reliably predictable from RS covariates. The (Williams J. et al., 1981) formulation requires only sand, silt, clay, and organic carbon percentages, all of which are

either directly predicted by the ML models or derivable from the OM prediction via a standard conversion factor.

2. Continuous functional form: The (Williams J. et al., 1981) equation is fully algebraic, enabling spatially distributed K-factor computation at every pixel by direct substitution of the ML-predicted texture and OM rasters. This avoids the interpolation errors introduced by rasterising a categorical nomograph.

3. Cross-environment validity: The (Williams J. et al., 1981) formulation has been validated across diverse semi-arid environments in Africa and the Middle East (Fenta et al., 2024; Ganasri & Ramesh, 2016; Gashaw et al., 2018) and is the standard K-factor approach in GIS-based RUSLE implementations for data-scarce regions (Benavidez et al., 2018; K.G. Renard, 1997).

4. Consistency with RUSLE literature in the Horn of Africa: Regional RUSLE studies in Ethiopia, Sudan, and Eritrea have uniformly adopted the (Sharpley, 1985) formulation, (Fenta et al., 2024; Gashaw et al., 2018; Zerihun et al., 2018); enabling direct comparison of K-factor magnitudes and spatial patterns with published regional datasets.

3.4.2.2.2. Mathematical Formulation

The EPIC/Williams K-factor (Williams J. et al., 1981) (Equations 3.3-3.8) was used to compute the K-factor and expresses the USLE-equivalent K value as the product of four dimensionless sub-functions, each capturing the influence of one soil textural or compositional component on erodibility:

$$K_{usle} = f_{csand} \times f_{cl-si} \times f_{orgc} \times f_{hisand} \quad (3.3)$$

where K_{USLE} is in SI units of $t \cdot ha \cdot h \cdot ha^{-1} \cdot MJ^{-1} \cdot mm^{-1}$. The four sub-functions are defined as:

$$f_{csand} = 0.2 + 0.3 \times \exp \left[-0.0256 \times m_s \times \left(1 - \frac{m_{silt}}{100} \right) \right] \quad (3.4)$$

f_{csand} accounts for the effect of coarse sand fraction: high sand content reduces erodibility because coarse particles resist detachment by raindrop impact and require greater transport energy. The exponential term captures the progressive reduction in erodibility as sand content increases and silt content decreases.

$$f_{cl-si} = \left(\frac{m_{silt}}{m_c + m_{silt}} \right)^{0.3} \quad (3.5)$$

f_{cl-si} reflects the clay-to-silt ratio effect: soils with a high silt fraction relative to clay are most erodible because silt particles are optimally sized for rainsplash detachment and surface transport, while clay increases soil cohesion and aggregate stability. The 0.3 exponent dampens the effect at extreme ratios.

$$f_{orgc} = 1 - \frac{0.25 \times orgC}{orgC + \exp(3.75 - 2.95 \times orgC)} \quad (3.6)$$

f_{orgc} captures the protective role of organic carbon: higher OC reduces erodibility through improved aggregate stability, increased water infiltration, and surface binding of fine particles. The logistic form of f_{orgc} reflects a diminishing marginal return of organic carbon on erodibility reduction above approximately 2–3% OC.

$$f_{hisand} = 1 - \frac{0.7 \left(1 - \frac{m_s}{100} \right)}{\left(1 - \frac{m_s}{100} \right) + \exp \left[-5.51 + 22.9 \left(1 - \frac{m_s}{100} \right) \right]} \quad (3.7)$$

f_{hisand} provides a high-sand correction for soils with $ms > 70\%$: at very high sand contents, f_{csand} alone underestimates the actual reduction in erodibility, so f_{hisand} adds a further dampening correction. $SN_1 = 1 - ms/100$ is the non-sand fraction.

In all equations, the variables are defined as: ms = sand content (%); msi = silt content (%); mc = clay content (%); OC = organic carbon content (%). Organic carbon was obtained from OM predictions using the Van Bemmelen factor ($OC = OM / 1.724$), which is the standard conversion for mineral soils and has been applied in all comparable RUSLE studies in the Horn of Africa.

The EPIC formulation produces K-factor values in US customary units. Therefore, the calculated values were converted to RUSLE-compatible SI units using a conversion factor of 0.1317 (K.G. Renard, 1997).

$$K_{RUSLE} = K_{USLE} \times 0.1317 \quad (3.8)$$

The resulting K_{RUSLE} values represent soil erodibility in metric units and were assigned to the corresponding sampling locations. These calculated K-factor values served as the reference dataset for subsequent regression modelling and spatial prediction of soil erodibility across the study area.

3.4.2.3. Slope length and steepness Factor (LS-factor)

The LS-factor represents the topographic control on erosion, combining the effect of slope steepness on the erosive shear stress of overland flow with the effect of slope length on the accumulation of runoff. The original USLE formulation was derived from unit plots of fixed length and is not directly applicable to complex two-dimensional terrain. Following standard practice in GIS-based RUSLE implementations, the slope-length term was therefore generalised using the upslope contributing area obtained from flow-accumulation analysis of the DEM, which accommodates the convergence and divergence of flow that characterises the deeply dissected terrain of the study area (I. A. N. D. Moore & Burch, 1986; P.J.J. Desmet and G. Govers, 1996).

The DEM was hydrologically conditioned by sink filling, and flow direction and flow accumulation were computed. The LS-factor was then evaluated pixel-wise using the unit-stream-power formulation of (I. A. N. D. Moore & Burch, 1986):

The LS factor was then computed using a raster-based approach suitable for complex, highly fragmented terrain. In mountainous environments such as the Eritrean highlands, the traditional plot-based USLE formulation is inadequate because it assumes uniform slope segments and does not account for spatial flow convergence and divergence.

The LS factor was calculated following the method of (P.J.J. Desmet and G. Govers, 1996; Simms et al., 2003)), which provides a GIS-compatible adaptation of RUSLE (Equation 3.9). The general formulation applied in this study is:

$$LS = (FA \times cell\ size) \div 22.13)^m \times (\sin \theta \div 0.0896)^n \quad (3.9)$$

Where FA represents flow accumulation (number of contributing cells), $Cell\ size$ is the DEM spatial resolution (m), θ is the slope angle (in radians), m is the slope length exponent (0.4-0.6), n is the slope steepness exponent (commonly 1.3 in RUSLE applications).

The $FA \times cell\ size$ is basically the specific catchment area (the upslope contributing area per unit contour width, in m), computed as the product of flow accumulation and cell size.

Flow direction and flow accumulation were derived from the hydrologically corrected DEM using ArcGIS. Flow accumulation was multiplied by the cell size to estimate the upslope contributing area

per unit width, which represents effective slope length in raster form. The slope gradient was computed in degrees from the DEM and converted to radians for trigonometric calculations.

The flow accumulation and slope rasters were computed in ArcGIS 10.4.1 (Spatial Analyst > Hydrology; Spatial Analyst > Surface), and the LS raster was computed using the Raster Calculator. LS is a static, time-invariant factor held constant across all monthly simulations.

3.4.2.4. Cover management factor (C-factor)

The cover-management factor expresses the protective effect of vegetation and surface cover as the ratio of soil loss under a given cover to that from continuously bare, tilled fallow, and therefore ranges from values approaching zero under dense canopy to unity over bare soil. Rather than assigning C from land-cover class lookup tables, which are spatially coarse and temporally static, C was derived continuously from the Sentinel-2 NDVI, which permits both the spatial detail and, critically, the monthly temporal resolution that this research requires. The exponential relationship of (Van Der Knijff et al., 2000) was adopted (Equation 3.10):

$$C = \exp(-\alpha \cdot NDVI / (\beta - NDVI)) \quad (3.10)$$

The C-factor was independently derived for each of the 12 calendar months using the empirical NDVI-to-C conversion equation of (Van Der Knijff et al., 2000). (Equation 3.11)

$$C_m = \exp(-\alpha \times NDVI_m / (\beta - NDVI_m)) \quad (3.11)$$

with the dimensionless parameters set to the widely used values $\alpha = 2$ and $\beta = 1$. This exponential form was preferred over linear alternatives because it reproduces the strongly non-linear decline of erosion risk with increasing cover, capturing the substantial protective gain conferred by even partial canopy development, and it has been validated for tropical and subtropical watersheds (Durigon et al., 2014) while continental-scale assessments have combined RS with land-use and management information to derive spatially explicit C (Panagos, Borrelli, et al., 2015b).. Negative NDVI values, corresponding to water and cloud-contaminated pixels, were masked, and the resulting C values were constrained to the physically meaningful interval [0, 1].

As with the R-factor, C was derived at two temporal resolutions from the same primary imagery. For the simulated framework, Equation (3.9) was applied to each of the twelve monthly NDVI composites, yielding twelve monthly C-factor rasters, C_m , that track the greening of vegetation with the onset of the summer(*kremti*) rains and its senescence through the dry season. For the conventional framework, the twelve monthly NDVI composites were first averaged to a mean annual NDVI, and Equation (3.9) was applied once to that mean, yielding a single annual cover factor, C_ann . The distinction is not merely procedural: because Equation (3.10) is non-linear, applying it to the mean NDVI does not give the same result as averaging the twelve monthly C values. The C_m ranges from approximately 0 (dense vegetation, negligible erosion relative to bare fallow) to 1 (bare soil).

Each monthly C-factor raster (C_m) was paired with the corresponding monthly simulated R-factor (R_sim_m) in the ArcGIS 10.4.1 Raster Calculator to compute the monthly simulated soil loss (A_sim_m). This month-by-month pairing captures the critical seasonal co-variation between erosivity and cover protection, particularly the high-risk transitional period in June, when R_sim begins to rise sharply but C remains relatively high due to incomplete crop canopy establishment.

3.4.2.5. Support Practice Factor (P)

The P-factor represents the ratio of soil loss with a given support practice to that from up-and-down-slope cultivation. For the primary modelling exercises, P was set to unity ($P = 1$) across the entire study area, representing the conservative baseline scenario of no conservation practices. This isolates the

contributions of R, K, LS, and C to modelled soil loss patterns without the confounding effect of unknown SWC distributions.

In a secondary analytical step, P values corresponding to documented SWC structure types and slope classes were assigned at locations of recorded SWC interventions (stone bunds, bench terraces, check dams) identified from field surveys and government records. Reference P values were taken from (Wischmeir, W. H. and Smith, 1978) and (K.G. Renard, 1997). The soil loss computed under these observed P values was compared with the $P = 1$ baseline to quantify the estimated erosion reduction attributable to existing SWC investments at the sub-watershed scale.

3.4.3. Conceptual Framework and Temporal Simulation Strategy

This study implements two parallel RUSLE computation frameworks: a monthly simulated approach (producing A_{annual}) and a conventional annual approach (producing $A_{\text{non-sim}}$), applied to the same study area and sharing the same K, LS, and P inputs. The conceptual flowchart illustrating both frameworks is provided in Figure 4.

In the simulated stream, seasonal rainfall and vegetation dynamics are explicitly preserved. Twelve monthly rainfall fields from GPM IMERG are annualised and converted into monthly R-factor rasters, while Sentinel-2 NDVI images are converted into C-factor rasters. For each month, a soil-loss raster is calculated using corresponding R- and C-factors combined with static K-, LS-, and P-factors. Since rainfall input is annualised, each monthly soil-loss raster is initially in tonnes per hectare per year, divided by twelve to get monthly contributions, which are then summed for annual soil loss. The GPM IMERG precipitation (P_m) is first multiplied by 12 before using the (Hurni, 1985a) erosivity equation, as it was calibrated with annual precipitation. The resulting simulated R-factor (R_{sim_m}) and Sentinel-2 C-factor (C_m) produce a monthly soil loss raster (A_{sim_m}). This is divided by 12 to get actual monthly soil loss (A_{actual_m}). The 12 A_{actual_m} rasters are summed for annual soil loss (A_{annual}).

In the conventional (annual) stream, the same primary data are aggregated before the factors are computed. The twelve monthly rainfall fields are summed to a single annual rainfall total, from which one annual R-factor is derived, and the twelve monthly NDVI images are averaged to a mean annual NDVI, from which one mean annual C-factor is derived. These annual factors are combined with the identical K-, LS- and P-factors to estimate annual soil loss directly. A_{annual} and $A_{\text{non-sim}}$ therefore arise solely from the temporal co-variance of R and C, not from any difference in total annual erosivity. All subsequent factor-specific sections refer back to this framework.

The Annualisation and De-Annualisation Procedure

A procedural difficulty arises from the form of the (Hurni, 1985a, 1988) erosivity equation, which is calibrated to take annual rainfall as its argument and to return an erosivity in annual units. It cannot be applied directly to a monthly rainfall total, because doing so would evaluate the regression far outside the range of the data on which it was fitted and would return a quantity of the wrong dimension. Two consistent options exist: to re-derive an erosivity relationship calibrated on monthly rainfall, which is impossible here because the sub-hourly gauge records required for such a calibration do not exist for the study area; or to apply the existing annual equation to each month in a dimensionally consistent manner. The second route was adopted, and is implemented through the annualisation and de-annualisation procedure.

Each monthly rainfall total is scaled to an annual-equivalent rate so that the (Hurni, 1985a) erosivity equation may be evaluated in the units for which it was calibrated; the resulting monthly soil-loss rate, expressed in annual units, is subsequently rescaled by the same factor to recover the true monthly contribution, and the twelve contributions are summed.

The procedure operates in five steps (Eq. 3.12-3.16).

First, each monthly rainfall total P_m (mm month^{-1}) is annualised by multiplication by twelve (Equation 3.12):

$$P_m^* = 12 \cdot P_m \quad (3.12)$$

Second, Equation (3.12) is evaluated on P_m^* , returning a month-specific erosivity (R_m^*) expressed in annual units (Equation 3.13):

$$R_m^* = -8.12 + 0.562 \cdot P_m^* [MJ \text{ mm } ha^{-1} h^{-1} yr^{-1}] \quad (3.13)$$

Third, a soil-loss rate is computed for each month (A_m^*) by combining the monthly erosivity and the monthly cover factor with the three static factors (Equation 3.14):

$$A_m^* = R_m^* \cdot K \cdot LS \cdot C_m \cdot P [t \text{ ha}^{-1} yr^{-1}] \quad (3.14)$$

This quantity is a *rate*, not a mass: it is the soil loss that would occur over a year if the rainfall and vegetation conditions of month m persisted for the whole year. Fourth, the rate is de-annualised by the same factor of twelve, recovering the actual contribution soil loss of that month, A_m (Equation 3.15):

$$A_m = A_m^* / 12 [t \text{ ha}^{-1} month^{-1}] \quad (3.15)$$

Fifth, the twelve monthly contributions are summed to give the simulated annual soil loss A_{sim} (Equation 3.16):

$$A_{sim} = \sum_{m=1}^{12} A_m [t \text{ ha}^{-1} yr^{-1}] \quad (3.16)$$

The difference map ($\Delta A = A_{non-sim} - A_{annual}$) was computed by subtracting the annual simulated soil loss map from the non-simulated map pixel by pixel across the study area. The resulting spatial distribution was analysed to identify zones of overestimation, agreement, and underestimation between the frameworks.

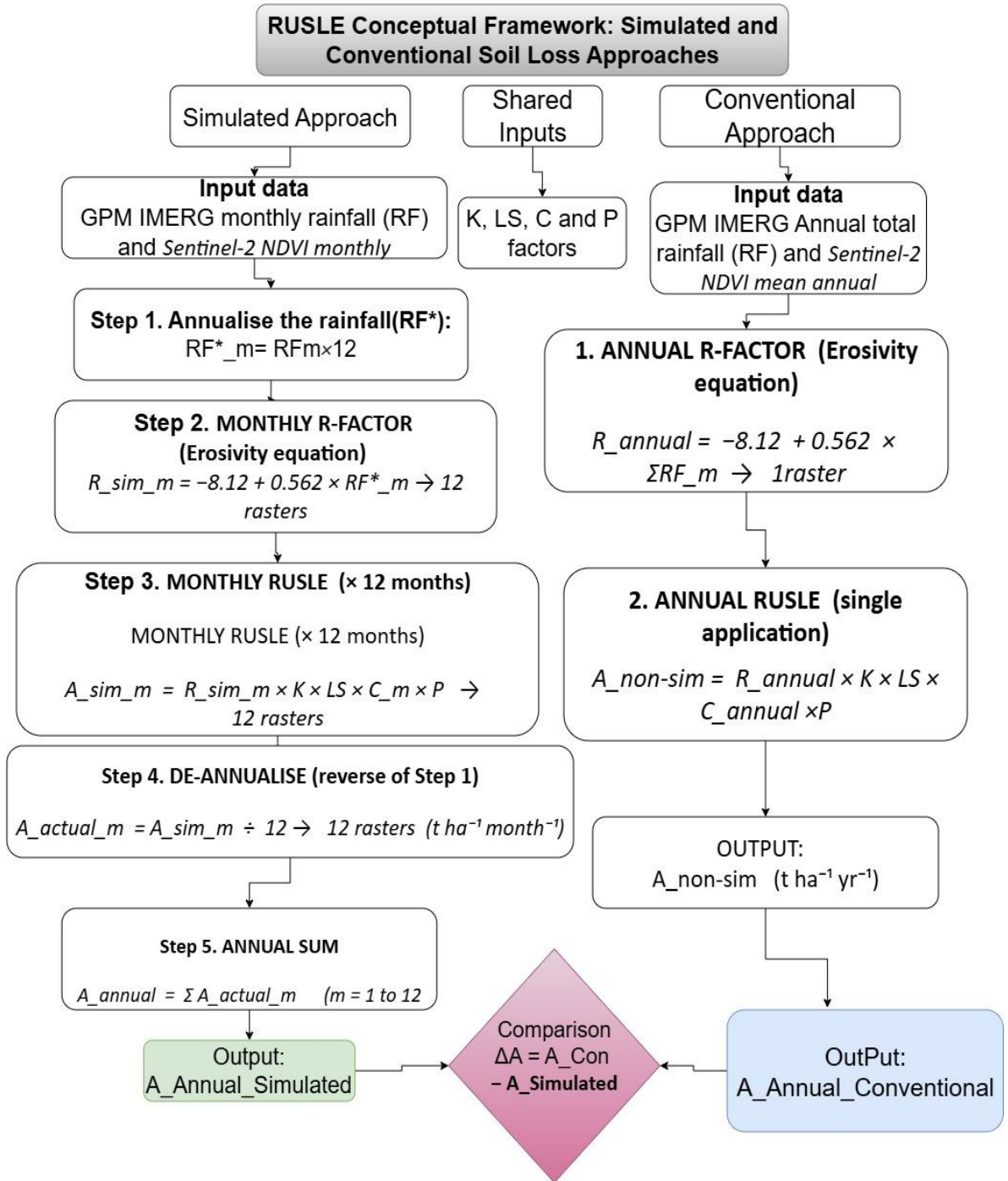


Figure 4. Conceptual Framework flow diagram of the Simulated and Conventional Approaches

4. RESULTS AND DISCUSSION

4.1. Machine-Learning: Soil-Property Prediction

This section presents and discusses the results of the machine-learning prediction of soil texture and OM, ultimately the spatial estimation of the soil erodibility (K) factor.

4.1.1. Environmental Covariates

Red-edge bands (RE1 and RE2). In addition to spectral indices, several individual reflectance bands were retained as covariates. The red-edge region, sampled by Sentinel-2 Band 5 (RE1, centred at 705 nm) and Band 6 (RE2, centred at 740 nm), lies at the steep transition between the strong chlorophyll absorption in the red and the high reflectance in the NIR. These bands enter the model directly as reflectance covariates, ρRE1 (B5, 705 nm) and ρRE2 (B6, 740 nm), and require no ratio formulation. They are sensitive to canopy chlorophyll and leaf-area conditions and, over sparsely vegetated ground, to the soil background, making them informative for discriminating soil–vegetation gradients that correlate with texture and fertility (Mulder et al., 2011).

Near-infrared band (NIR). Near-infrared reflectance, sampled by Sentinel-2 Band 8 (centred at 842 nm), enters the model as the covariate ρNIR (B8, 842 nm). NIR reflectance is high for healthy vegetation, owing to strong scattering in leaf mesophyll, and is modulated over exposed ground by soil moisture and organic-matter content, both of which tend to lower NIR reflectance. Seasonal NIR composites (Figure 5) therefore encode complementary information on vegetation structure and on soil condition during low-cover periods.

Shortwave-infrared bands (SWIR1 and SWIR2). The shortwave-infrared region, sampled by Sentinel-2 Band 11 (SWIR1, centred at 1610 nm) and Band 12 (SWIR2, centred at 2190 nm), is the most diagnostic optical region for soil properties and was retained as the covariates ρSWIR1 (B11, 1610 nm) and ρSWIR2 (B12, 2190 nm). SWIR reflectance responds strongly to soil moisture and to the absorption features of clay minerals and OM, and reviews of RS for soil mapping consistently identify the SWIR and NIR channels as among the most powerful predictors of soil texture and organic carbon (Mulder et al., 2011). The seasonal SWIR1 and SWIR2 composites (Figure 5) are consequently key covariates for predicting the sand, silt, clay, and organic-matter fractions that determine the K-factor.

In total, the seasonal composites of NDVI, BSI, and the red-edge, NIR, and SWIR reflectance bands provide a multi-temporal, physically informative set of remote-sensing covariates spanning the visible-to-shortwave-infrared spectrum. These were stacked with the DEM-derived terrain attributes to form the full predictor set for the ML regression models; redundant or weakly informative covariates were subsequently screened during feature selection so that the final models retained the most predictive subset for each target soil property.

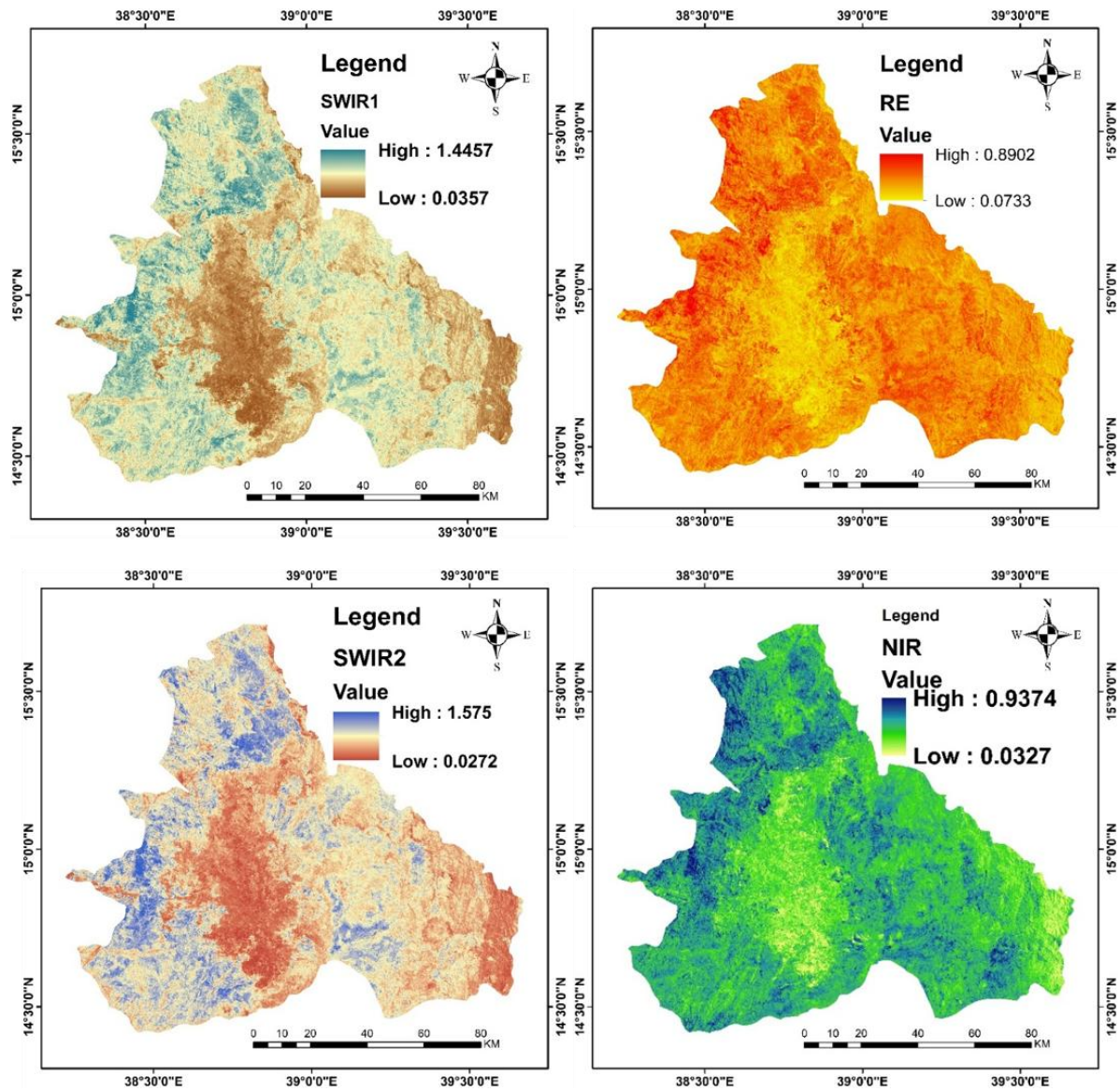


Figure 5. Environmental Covariate: Red-edge (RE), NIR & shortwave-infrared (SWIR)

Bare soil Index (BSI)

Equivalently, in Sentinel-2 band notation, $BSI = [(B11 + B4) - (B8 + B2)] / [(B11 + B4) + (B8 + B2)]$. High BSI values indicate bare or sparsely vegetated surfaces, where the sensor observes the soil directly and the spectral signal is most diagnostic of soil composition. The seasonal BSI composites (Figure 6) are particularly valuable covariates because the degree and timing of bare-soil exposure differ between land units and reveal the underlying soil during the low-cover periods that also coincide with the onset of erosive rainfall.

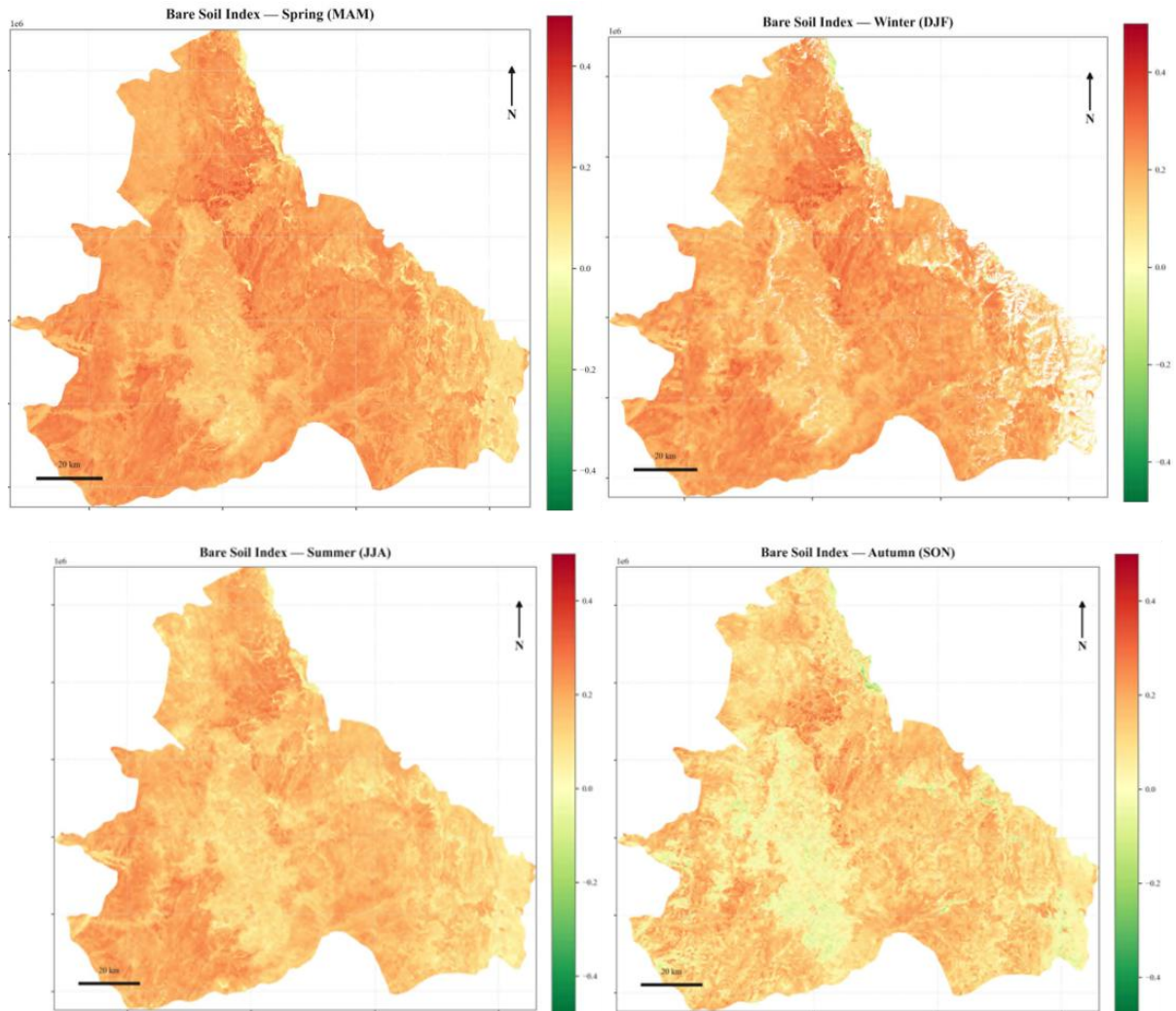


Figure 6. Seasonal BSI images for winter, spring, summer, and autumn, highlighting bare-soil exposure.

Land Surface Temperature as a Covariate Source

The resulting seasonal LST composites (Figure 7) show the expected thermal regime of the study area, with the lowest temperatures in winter and the highest in summer, and reveal spatial patterns reflecting the combined influence of elevation, land cover, and surface moisture. Cooler zones correspond broadly to higher, more vegetated, or moister terrain, while warmer zones coincide with exposed, drier, and lower-lying surfaces. Because these thermal gradients partly reflect surface soil moisture and composition, the seasonal LST layers provide predictive information complementary to the optical Sentinel-2 covariates, particularly for the organic-matter and clay fractions that influence soil thermal behaviour.

The four seasonal LST composites were stacked with the Sentinel-2 spectral covariates and the DEM-derived terrain attributes to form the complete environmental predictor set used by the machine-learning regression models.

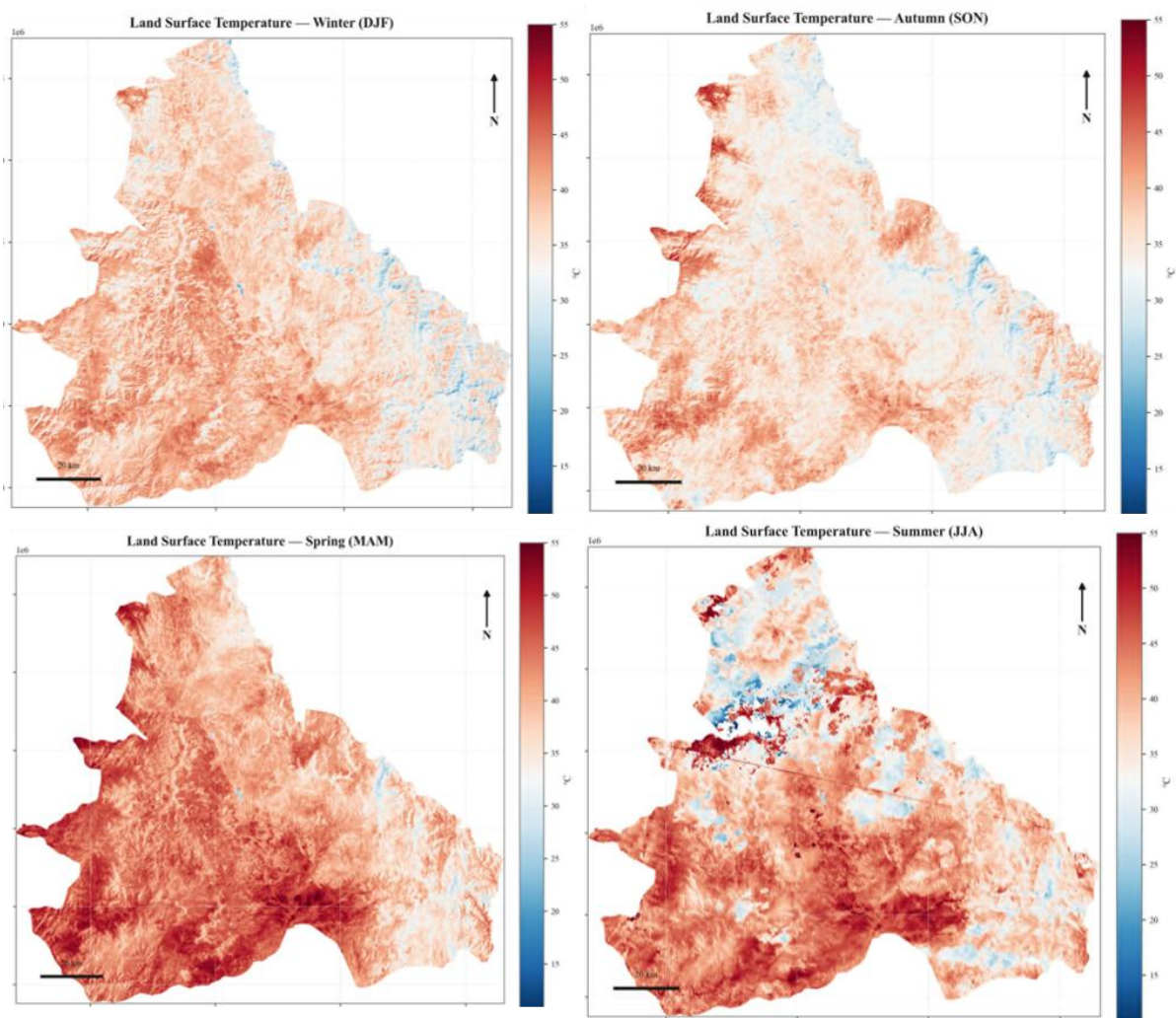


Figure 7. Seasonal median images of LST, °C: winter, spring, summer, and autumn

DEM Derivative:

The maps (Figure 8) display terrain attributes derived from the hydrologically conditioned SRTM digital elevation model, including slope gradient, aspect, Topographic Position Index, surface curvature,

and relief intensity. These maps illustrate the study area's topographic structure and its influence on the spatial distribution of soil properties and erosion potential across the four landscape zones.

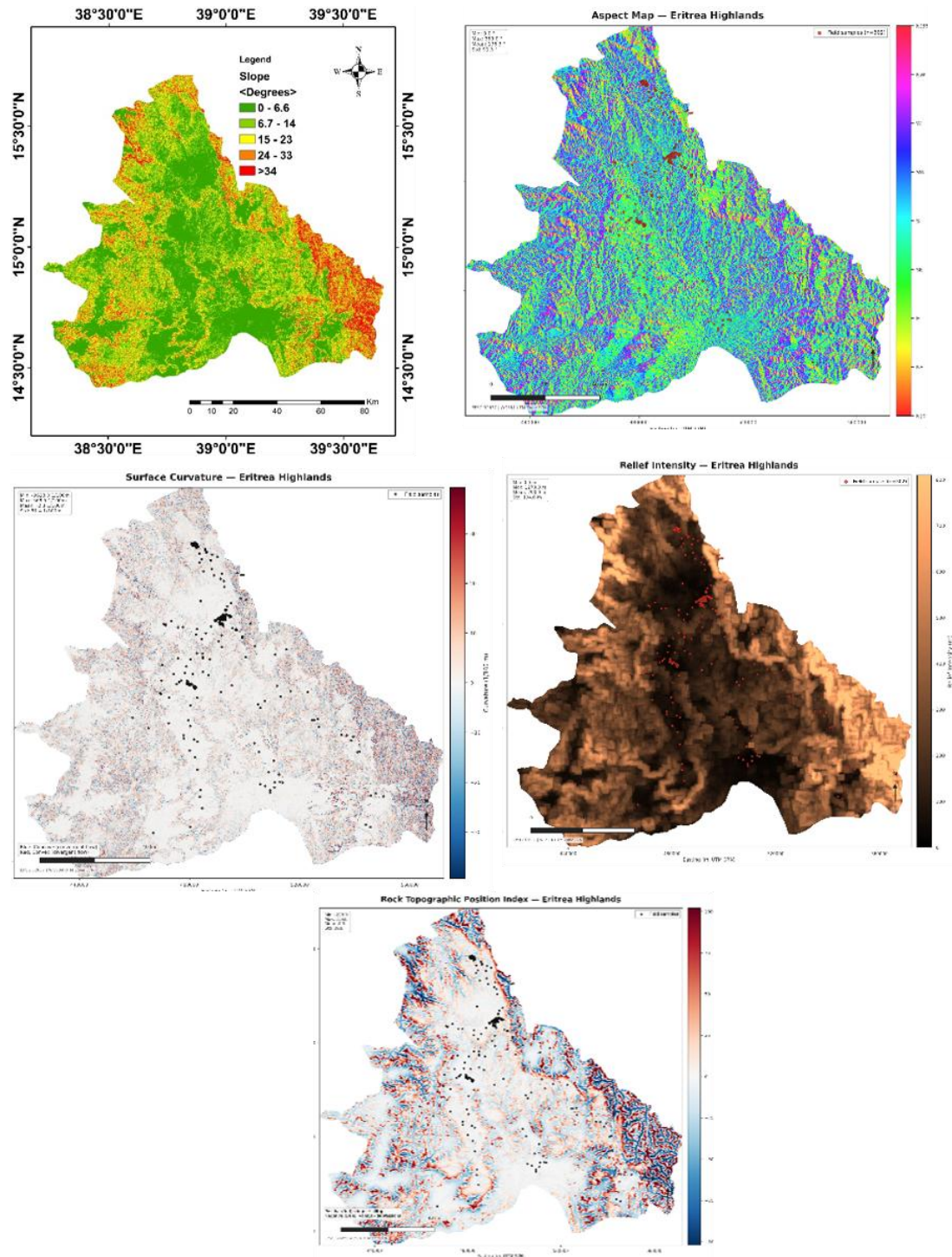


Figure 8. Terrain derivative: slope, aspect, surface curvature, relief intensity, and topographic position index (TPI). Red dots denote the 302 field soil-sampling locations

4.1.2. Soil Texture prediction using ML

4.1.2.1. Model Performance and Feature Importance

Table 6 summarises the random five-fold cross-validation performance of all five algorithms, and Figures 9, 10, & 11 show the observed-versus-predicted scatter for each model and fraction. Each fraction showed a different optimal algorithm: XGB for clay ($R^2 = 0.421$), GBT for sand ($R^2 = 0.361$), and ET for silt ($R^2 = 0.334$). Clay was the most predictable fraction across every algorithm and silt the least, a pattern with a clear physical basis: clay content is expressed directly in the shortwave-infrared through mineral absorption features and indirectly through its control on soil moisture and thermal behaviour, whereas silt, lacking a distinct spectral signature, was obtained largely as the compositional residual and therefore inherits the prediction errors of the other two fractions. The consistency of the model ranking across fractions in Figures 9, 10 & 11, together with the narrow spread of R^2 among the top three algorithms (ET, RF, GBT and XGB are all closely competitive), indicates that the result is driven by the informativeness of the covariates rather than by the idiosyncrasies of any single algorithm.

Table 6. Random five-fold cross-validation performance of the five algorithms for each texture fraction ($n = 277$). Bold rows denote the best-performing model per fraction.

Property	Model	R^2	RMSE (%)	MAE (%)	RPD	RPIQ	Bias (%)
Silt	ET ★	0.334	7.82	6.31	1.23	1.70	-0.13
	RF	0.302	8.01	6.43	1.20	1.66	-0.20
	XGB	0.263	8.22	6.37	1.17	1.61	-0.08
	GBT	0.234	8.38	6.54	1.14	1.58	-0.15
	LGB	0.195	8.60	6.69	1.11	1.54	-0.03
Sand	GBT ★	0.361	12.84	10.17	1.25	1.71	+0.07
	ET	0.350	12.96	10.25	1.24	1.70	+0.10
	RF	0.319	13.26	10.35	1.21	1.66	+0.01
	XGB	0.241	14.00	10.97	1.15	1.57	+0.33
	LGB	0.209	14.29	11.13	1.12	1.54	+0.12
Clay	XGB ★	0.421	10.62	7.90	1.31	2.02	-0.12
	RF	0.408	10.73	8.05	1.30	2.00	+0.17
	ET	0.397	10.84	8.23	1.29	1.98	+0.07
	GBT	0.331	11.41	8.44	1.22	1.88	+0.15
	LGB	0.314	11.56	8.57	1.21	1.86	-0.10

Benchmarking is essential to interpreting these values. Coefficients of determination of 0.33–0.42 would be modest in a data-rich temperate setting, but for satellite-based soil-texture prediction in a small-sample ($n = 277$), semi-arid African context, they are competitive with, and in several cases exceed, comparable published studies (e.g. (Vaudour et al., 2019), clay $R^2 = 0.39$ –0.42; (Bousbih et al., 2019), clay $R^2 = 0.39$ –0.44 in semi-arid). The RPD values (1.14–1.31) and the clay RPIQ (2.02) place the models in the category suitable for spatial-trend mapping rather than precise point prediction, which is the appropriate and sufficient standard for deriving a spatially distributed K-factor. All three selected models represent a large improvement over the continental iSDAsoil product evaluated against the same field data, where clay and silt yielded negative R^2 , confirming that a locally trained model captures regional soil variation that global products miss.

Table 7 summarises the cross-validation performance metrics for all five ensemble algorithms across the three soil texture fractions and identifies the best-performing model per fraction: Extra Trees (ET) for silt, Gradient Boosted Trees (GBT) for sand, and XGBoost for clay.

Table 7. Random 5-fold, $n=277$: Best model per fraction

Property	Best Model	R ²	RMSE (%)	MAE (%)	Bias (%)
Silt	Extra Trees (ET)	0.274	7.82	6.31	-0.13
Sand	Gradient Boosted Trees (GBT)	0.39	12.84	10.17	+0.07
Clay	XGBoost	0.45	7.42	5.88	~0

4.1.2.2. Feature Importance (SHAP Analysis)

Clay - SHAP Interpretation Summary

The XGBoost model for clay content ($R^2 = 0.421$, RMSE = 10.62%, $n = 277$) performed best among the five ensemble algorithms tested. The mean |SHAP| importance ranking (Fig. 9) reveals a steep, right-skewed distribution dominated by a few key covariates: spring land surface temperature (LST_Spr) is the most influential predictor, with importance roughly 1.6–2 times higher than the next most important features. These are followed by secondary predictors including Relief, spring climate moisture index (CMI_Spr), and mean annual LST (LST_mean). A wider mid-tier group, comprising topographic-hydrological variables (TWI, Rock_TPI), vegetation-phenological measures (NDVI_skew, NDVI_Aut_Relief), and spectral reflectance indicators (BSI_Spr, RE1_Aut), contributes moderate, relatively equal importance. Conversely, seasonal LST derivatives, DEM-based ranks, and SWIR/BSI summary features add marginal, largely redundant contributions and occupy the lowest tier.

The SHAP beeswarm plot (Fig. Y) clearly visualises the direction and consistency of these effects. Clay content increases with higher spring LST and spring moisture (CMI_Spr), indicating that the model detects a combined thermal–moisture pattern in which warm, moist spring conditions signal clay-rich soils. Relief shows the opposite trend: lower-relief areas tend to have higher clay predictions, consistent with clay accumulation in low-energy depositional zones like valley bottoms and footslopes, rather than erosive steep terrains. The topographic wetness index (TWI) exhibits a threshold effect, high TWI values reliably indicate higher clay content, whereas low values have little impact. Vegetation and spectral covariates (NDVI_skew, Rock_TPI, BSI_Spr, RE1_Aut) show varied, non-monotonic colourations, indicating these features interact with terrain and moisture conditions rather than acting independently or additively.

Collectively, these patterns indicate that clay distribution in the Eritrean Highlands is governed primarily by the interaction of thermal and moisture regimes with topographic position, a pattern physically consistent with erosion-deposition dynamics, whereby clay is preferentially redistributed toward low-relief, high-wetness-index zones with distinct spring surface energy balances, while spectral and vegetation indices provide secondary, context-dependent refinement to the prediction.

Silt - SHAP Interpretation Summary

The Extra Trees model predicting silt content ($R^2 = 0.334$, RMSE = 7.82%, $n = 277$) was the top performer, but its explanatory power is much lower than the clay model, showing silt is harder to predict with current covariates. The mean |SHAP| importance (Fig. 10) shows an even distribution, unlike the steep, single-feature dominance for clay. SWIR_ratio_Spr leads, followed by Rock_TPI, RENDVI_Spr,

LST_max, and BSI_range, with spectral, vegetation, and terrain covariates showing minor differences. This balanced profile suggests silt content results from a complex mix of covariates rather than a single process, fitting its role as an intermediate, transitional texture in weathering and erosion.

The SHAP beeswarm plot (Fig. 10) shows that the top feature, SWIR_ratio_Spr, has a broad and non-monotonic influence, with values both above and below zero. This indicates its effect depends on interactions rather than a simple linear trend. Rock_TPI displays a clear threshold: a small group of very high (rocky, convex terrain) values are sharply marked by strongly negative SHAP values, meaning rocky terrain consistently decreases the predicted silt content. The other values remain near zero, showing minimal discrimination. RENDVI_Spr shows a gradual trend: higher spring vegetation vigour correlates with higher predicted silt, likely because silt retains more moisture than sand. Interestingly, the LST-family features (LST_mean, LST_Win, LST_Sum) show an opposite trend to the clay model, with warmer conditions linked to lower predicted silt, indicating that thermal-moisture regimes favouring clay formation tend to decrease silt. Mid- and lower-tier variables (BSI_range, REI_Aut, NDVI_Aut, NDVI_range, Relief, DEM) display weak, overlapping colour patterns, implying subtle effects likely driven by interactions rather than direct influence.

Overall, these patterns show that silt distribution in the Eritrean Highlands is mainly driven by spectral-vegetation moisture proxies and terrain-roughness thresholds like Rock_TPI, not a single thermal-topographic factor like for clay. Silt is a residual, transitional fraction influenced by weathering, vegetation, and terrain, with no single factor dominating. This explains the flatter SHAP importance profile and lower model performance for silt than for clay.

Sand - SHAP Interpretation Summary

The Gradient Boosted Trees model for sand content ($R^2 = 0.361$, RMSE = 12.84%, $n = 277$) was the most effective algorithm for this fraction. Its explanatory power was between that of clay and silt, but it exhibited the highest absolute prediction error among the three textures, likely due to sand's wider value range. The mean |SHAP| importance ranking (Fig. 11) indicates an importance pattern that lies between clay and silt: a tight cluster of top features, including CMI_Spr, Relief, and LST_Spr, followed by a gradual decline through spectral and terrain-based covariates (SWIR_ratio_Sum, NDVI_skew, BSI_Aut, BSI_mean). This pattern contrasts with the single-feature dominance in clay or the more uniform distribution in silt.

The SHAP beeswarm plot (Fig. 11) shows that covariates for sand are similar to those for clay but with opposite effects. Low spring moisture (CMI_Spr) and low-to-moderate spring land surface temperature (LST_Spr) are linked to higher sand levels, while high values reduce sand, opposite the pattern that promotes clay. Relief shows an opposite trend to clay, with higher relief (steeper terrain) linked to more sand, reflecting the accumulation on steep slopes after finer particles wash away. The SWIR_ratio_Sum has a threshold, where extreme values significantly increase SHAP contributions. Indicators of bare soil and sparse vegetation (BSI_Aut, BSI_mean, NDVI_skew) suggest higher sand with more bare soil and less vegetation, consistent with spectral and textural expectations. Mid- and lower-tier terrain features (DEM_rank, LST_mean, Rock_TPI) have a weaker, more scattered influence, with Rock_TPI less impactful than in the silt model.

These patterns suggest that sand distribution in the Eritrean Highlands is governed by a thermal-moisture-relief axis similar to that for clay, but operates in reverse: dry, steep, less vegetated areas accumulate coarse sand, while moist, low-relief areas have fine clay. This mirrored relationship indicates a consistent erosion and deposition gradient, directly linked to textural factors affecting the RUSLE K-factor.

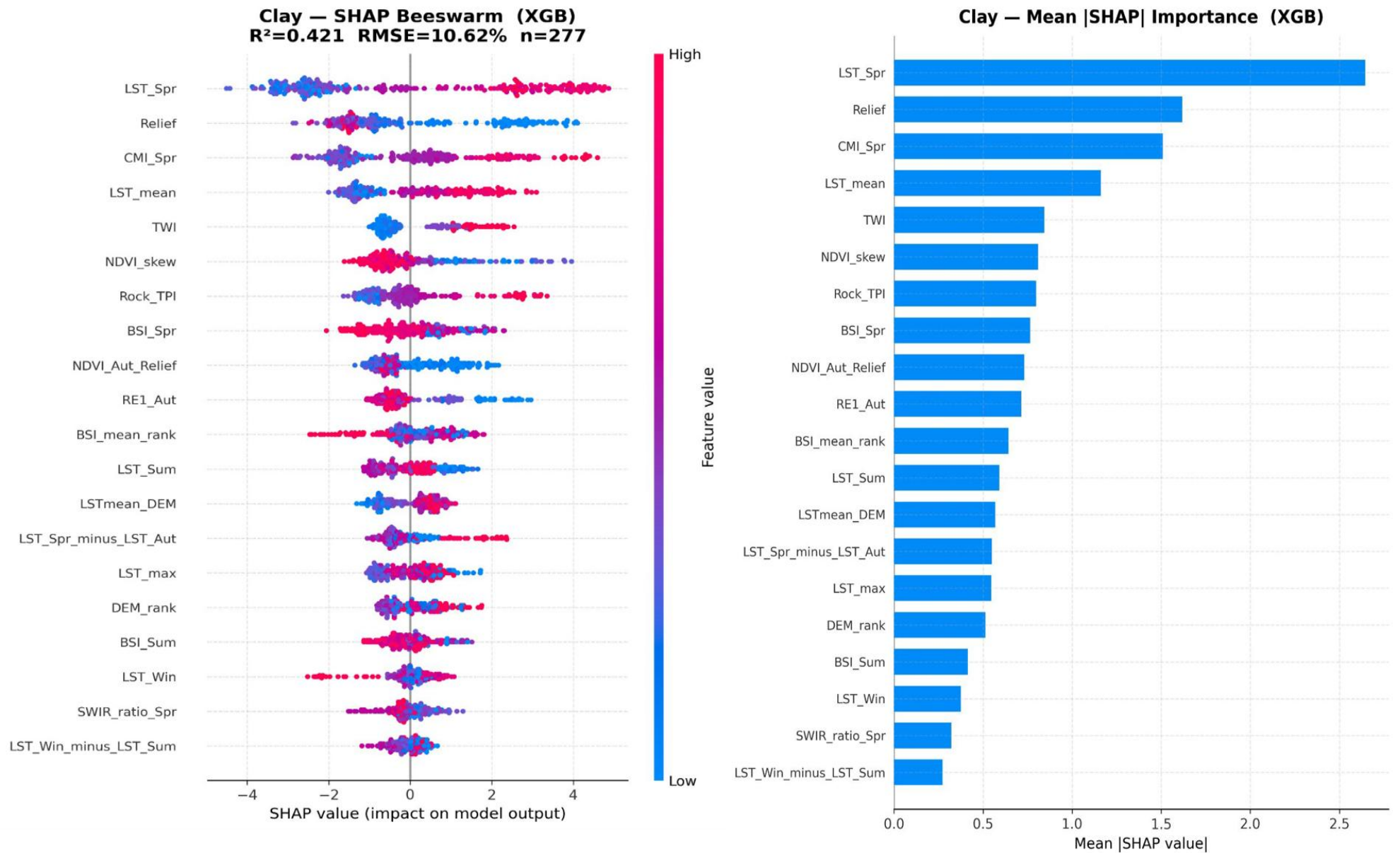


Figure 9. Mean Absolute SHAP feature importance for the best model of each fraction (left) & SHAP beeswarm plot for the clay model (XGB). Each point is one sample; horizontal position gives the SHAP value (impact on predicted clay) and colour encodes the feature.

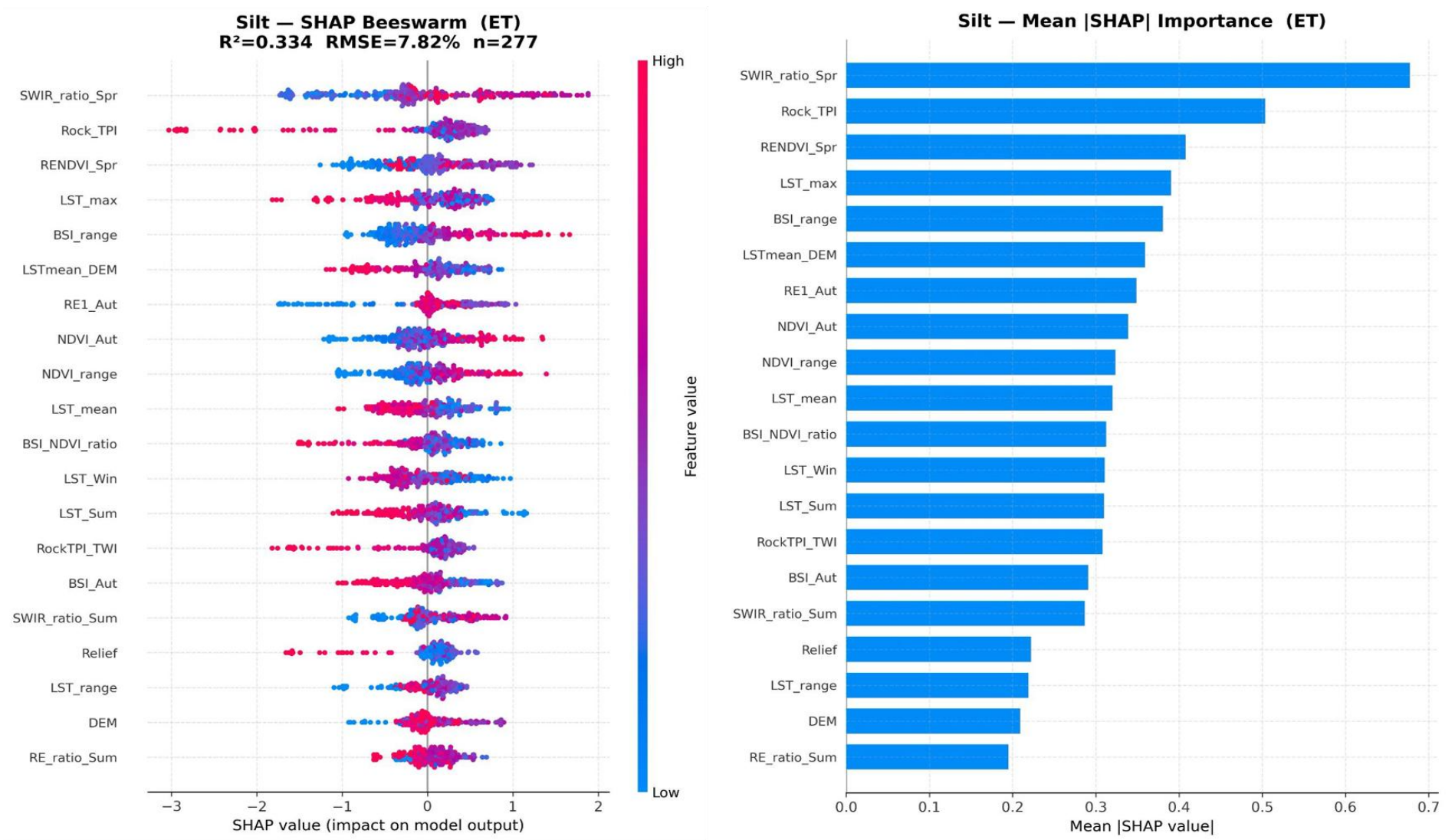


Figure 10. Mean Absolute SHAP feature importance for the best model of each fraction (left)& SHAP beeswarm plot for the silt model (ET). Each point is one sample; horizontal position gives the SHAP value (impact on predicted clay), and colour encodes the feature..

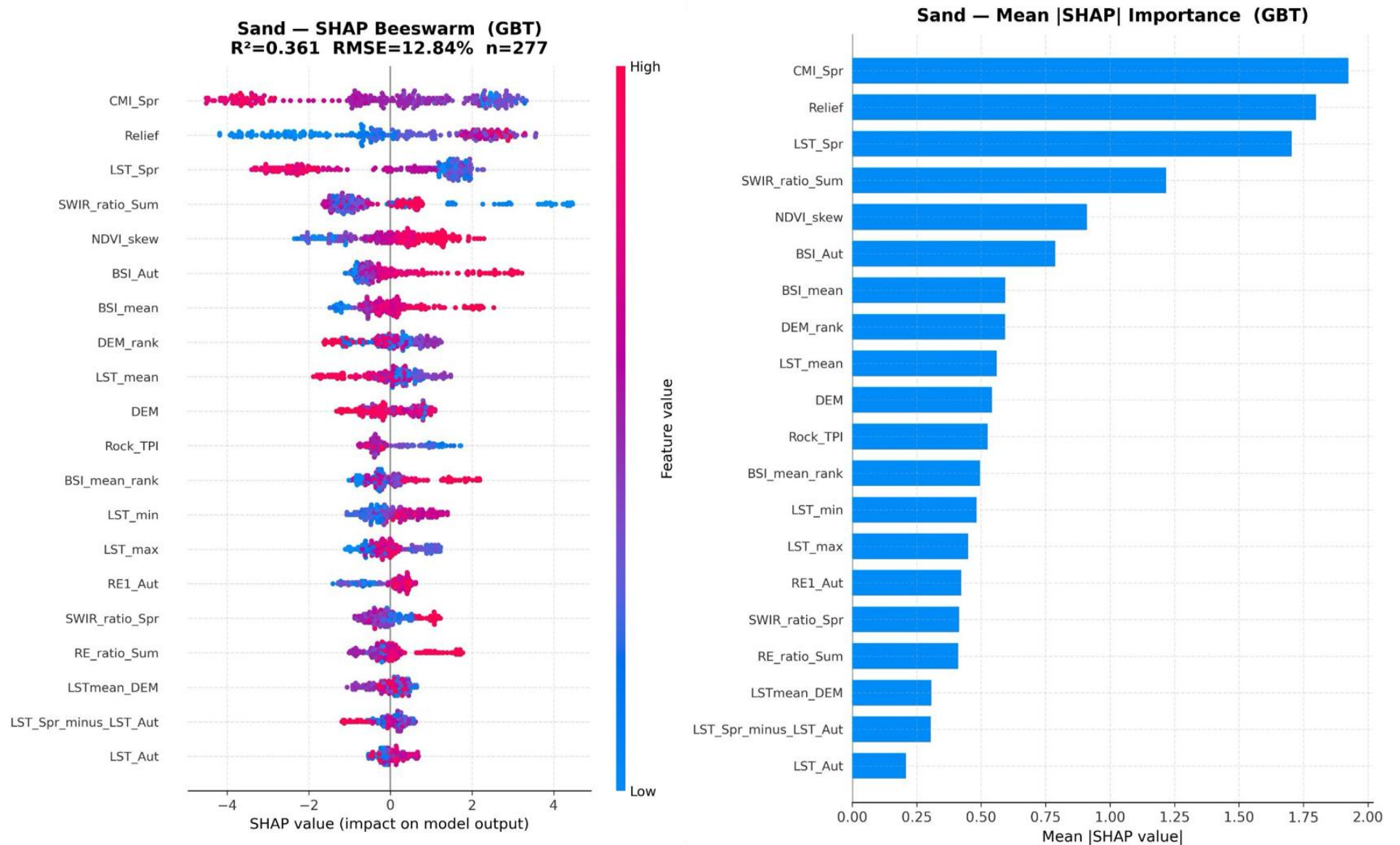


Figure 11. Mean Absolute SHAP feature importance for the best model of each fraction (left)& SHAP beeswarm plot for the silt model (GBT). Each point is one sample; horizontal position gives the SHAP value (impact on predicted clay) and colour encodes the feature. ..

4.1.2.3. Predicted Soil texture maps (Sand, silt& clay)

The best-performing model for each fraction was refitted on the full set of 277 samples and applied pixel-by-pixel to the environmental covariate stack to produce spatially continuous predictions across the study area at 30 m resolution. The three independently predicted fractions were normalised at each pixel to enforce compositional closure (sand + silt + clay = 100%). The resulting texture maps are shown in Figure 12 and constitute, to the author's knowledge, the first spatially continuous, remotely sensed soil-texture dataset for the Eritrean Highlands. Their principal value is that they translate the point-scale field measurements into the wall-to-wall surfaces required by the EPIC/Williams erodibility formulation, and they should be interpreted in light of the reliability characteristics: they represent the broad spatial gradients and modal texture of the landscape faithfully, while compressing the rarest textural extremes towards the population mean.

The three maps are mutually consistent and display a coherent, physically interpretable spatial organisation. Sand content (Figure 12) is high across most of the study area, with the majority of the landscape falling in the 55–73% range, and is lowest in a distinct north–south corridor through the central and south-central sector, where values decline to 16–34%. This coarse-textured dominance is consistent with the soils developed on the crystalline basement of the highlands and with the extensive Leptosols and Regosols of the steeper terrain. Silt content (Figure 12) exhibits the inverse spatial pattern to sand, as compositional closure requires, with the highest values (37–52%) concentrated in the central lowland corridor and lower values (17–31%) on the dissected margins to the south and east. Clay content (Figure 12) is generally low, most of the area lying below 20%, with the highest concentrations (25–50%) forming a discernible band through the south-central part of the study area.

Critically, the spatial patterns corroborate the SHAP interpretation of the models rather than merely restating it. The elevated clay and silt values coincide with the lower-lying, more convergent central corridor, exactly the setting in which fine particles are expected to accumulate and in which the Topographic Wetness Index and land-surface-temperature covariates attained high SHAP importance; conversely, the sandy margins correspond to the steeper, higher-relief terrain that sheds fines, consistent with the strong negative SHAP contribution of relief to predicted clay. The convergence of independent lines of evidence, the learned feature effects, the residual behaviour, and the geomorphological plausibility of the spatial patterns provides confidence that the maps capture genuine soil–landscape organisation. Nonetheless, the compressed predicted range means the maps under-represent the areal extent of the most clay-rich soils; the true prevalence of high-clay Vertisols is therefore likely to be somewhat greater than the maps suggest, a caveat carried into the interpretation of the erodibility factor.

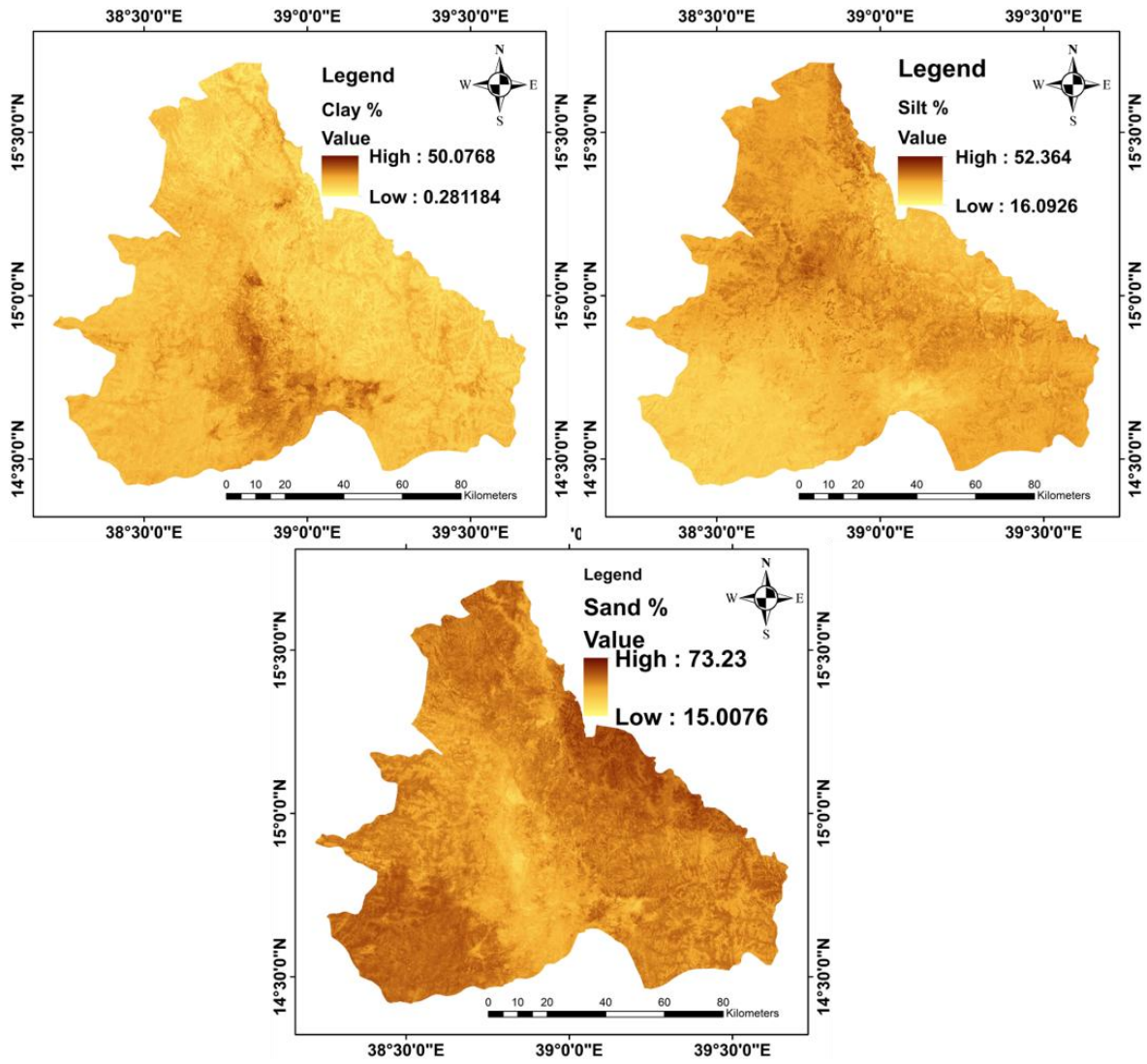


Figure 12. Predicted soil particle size maps: clay (left-up), silt (right-up), and sand (bottom) content.

4.1.2.4. Prediction Reliability and Residual Behaviour

The observed-versus-predicted diagnostics for the selected models quantify the reliability of the predictions and expose their principal limitation (Figure 13). For all three fractions the regression slope is below unity (0.31 for silt, 0.37 for sand, 0.44 for clay), indicating the regression-to-the-mean, or range compression, that is characteristic of ensemble models trained on moderate-sized datasets (Wadoux et al., 2020): the highest prediction density (yellow) coincides with the central quantiles of each distribution, where training samples are most abundant, while the extremes are pulled towards the mean. The biases, however, are negligible at the whole-sample level (-0.13% silt, $+0.07\%$ sand, -0.12% clay), confirming that the compression redistributes error across the range rather than shifting the predictions systematically.

Observed vs. Predicted Soil Texture — 5-Fold Cross-Validation

XGBoost | 15 remote-sensing + terrain features | 7×7 window | n = 302

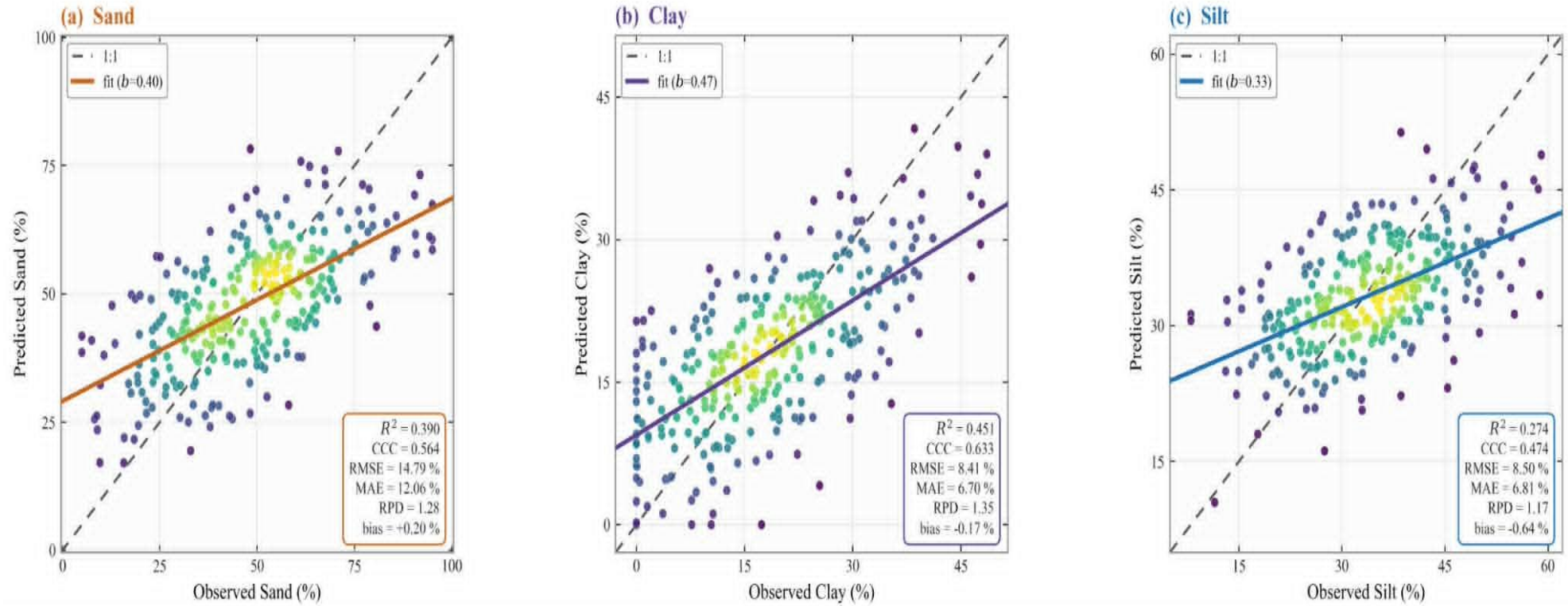


Figure 13. Observed versus predicted values for the best-tuned model per fraction (silt–ET, sand–GBT, clay–XGB; five-fold CV, n = 277). Points are coloured by kernel density; the dashed line is 1:1 and the solid line the regression fit (slope annotated). MAE, RPD

The formal residual diagnostics (Figure 14) confirm that the models are otherwise well behaved. Residuals for silt and sand are approximately normally distributed (Shapiro–Wilk $p = 0.506$ and 0.640 , respectively) and show no systematic structure against the predicted values, the LOESS trend remaining close to zero throughout and almost all residuals falling within the $\pm 2\sigma$ envelope. Clay residuals depart significantly from normality (Shapiro–Wilk $p < 0.001$), with the Q–Q plot revealing positive skew in the upper tail; this reflects the bounded lower limit of clay content and the spatial rarity of high-clay Vertisols rather than a defect of the model, and is consistent with the range-compression behaviour described above. Because these residual characteristics are interpretable and quantified, they are propagated transparently into the erodibility factor: the prediction errors reported here are the basis of the Monte Carlo uncertainty analysis of the K-factor, so that the effect of soil-property uncertainty on the soil-loss estimates is represented rather than ignored.

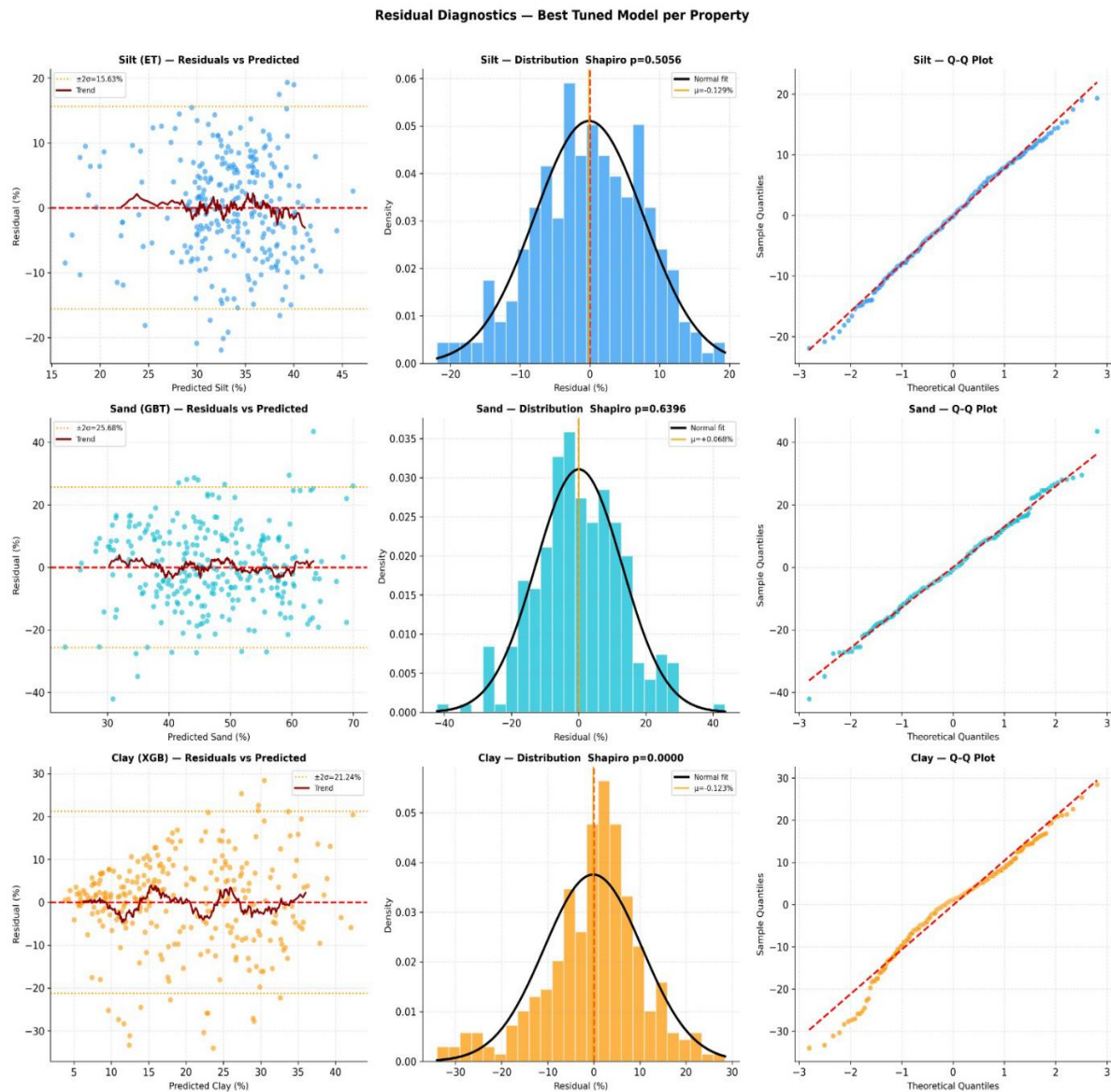


Figure 14. Residual diagnostics for the best-tuned model per fraction (rows: silt–ET, sand–GBT, clay–XGB): residuals versus predicted with $\pm 2\sigma$ bounds and LOESS trend (left); residual distribution with normal fit and Shapiro–Wilk p -value (centre); and normal Q–Q p

4.1.3.OM Prediction using ML

4.1.3.1. Cross-validation performance of the six Algorithms

Table 8 and Figure 15 present cross-validation results for the six algorithms. Gradient Boosted Trees (GBT) performs best according to point estimates, with an MAE of 0.416 OM% and an RMSE of 0.556 OM%. However, differences among the six algorithms are minimal: 0.023 in R^2 , 0.010 OM% in RMSE, and 0.02 in RPD. These differences fall within the variability from random fold assignments and do not reflect a true difference in predictive ability. The RMSE-to-MAE ratio of about 1.34, slightly above the expected 1.25 for normal errors, suggests that the error distribution has heavier tails than a Gaussian: the typical prediction error is roughly 0.42 OM%, but some samples are predicted significantly worse, skewing the RMSE due to larger errors. With a data range of around 3 OM% and a mean near 1.7 OM%, the typical error of 0.42 OM% is about 25% of the mean, comparable to published benchmarks for satellite-based organic matter predictions in data-limited semi-arid areas. The similar accuracy across five algorithms with different biases, including bagging versus boosting, exact versus histogram splits, and level-wise versus leaf-wise growth, indicates that the maximum achievable accuracy is limited by the information in the covariates rather than by the models' complexity. GBT is chosen as the best model based on point estimates, with the understanding that its performance advantage over other algorithms is within the noise margin.

Table 8. Cross-validated performance of the six models, sorted best to worst.

Model	R^2	RMSE	RPD	Rank
GBT (Gradient Boosting)	0.324	0.556	1.22	1 — best
ET (Extra Trees)	0.311	0.562	1.21	2
RF (Random Forest)	0.311	0.562	1.21	2
XGB (XGBoost)	0.310	0.562	1.21	4
LGB (LightGBM)	0.305	0.564	1.20	5
STK (Stacking)	0.301	0.566	1.20	6

Observed vs Predicted — All Models (Random CV)

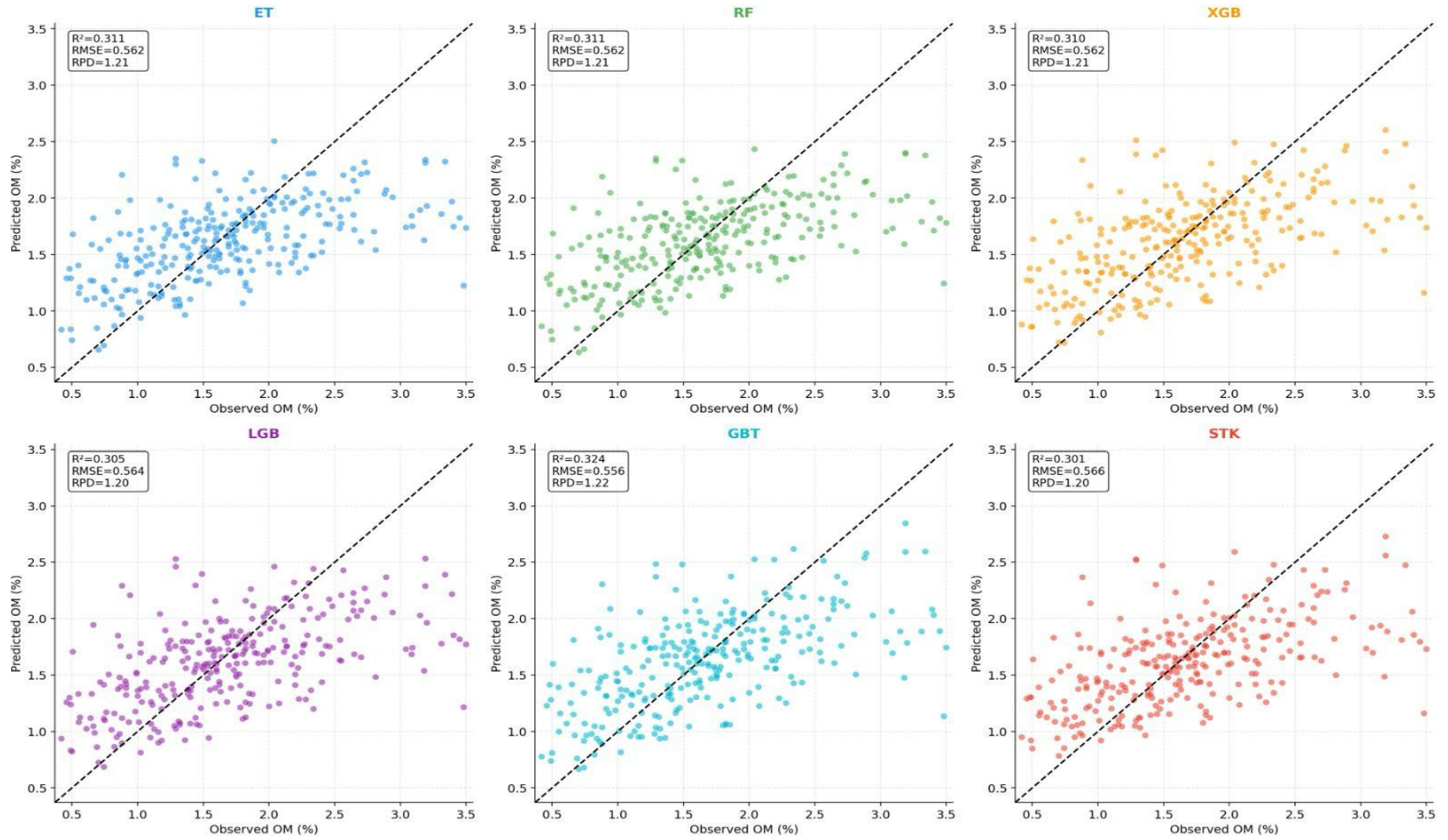


Figure 15. Observed versus predicted OM for all six models under random cross-validation. Each panel reports R^2 , RMSE and RPD..

The plot comparing observed and predicted values for the selected GBT model (Figure 16) aligns with the cross-validation results in Table 8, showing $R^2 = 0.324$, $RMSE = 0.556$ OM%, $MAE = 0.416$ OM%, and $RPD = 1.22$ ($n = 277$). The regression slope of 0.35, notably below the ideal 1.0, suggests the model pulls predictions toward the mean, underestimating high organic matter values and overestimating low ones. This pattern is typical in datasets where the signal-to-noise ratio in the covariates is moderate. Such shrinkage aligns with the flat SHAP importance profile discussed in Section 4.1.3.3 and indicates that the model's predictive limit is defined by the information within the covariates, not by the algorithm's capabilities.

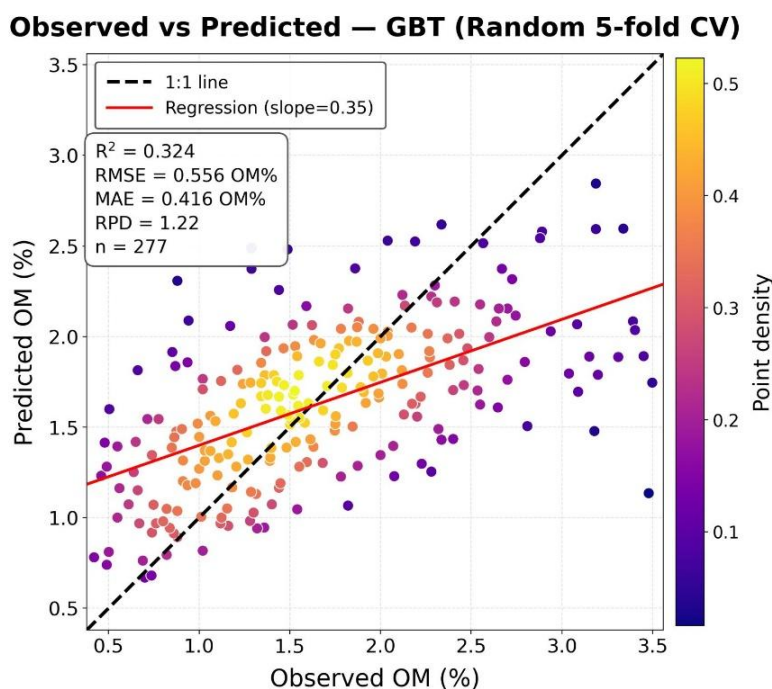


Figure 16. Observed versus predicted OM for the GBT model under random 5-fold cross-validation. Points are coloured by local density; the dashed line is the 1:1 identity and the solid red line is the ordinary least-squares fit of predicted on observed.

Figure 17 shows residual diagnostics for the selected GBT model. The residuals-versus-predictions trend stays within about ± 0.1 OM% of zero across the prediction range, indicating the model is conditionally unbiased and doesn't require post-hoc correction. The residuals mostly fall within ± 1.0 OM%, with no clear heteroscedasticity, except for one outlier at around -2.35 OM%, representing the worst prediction in the dataset. The residual histogram has a mean of -0.048 OM%, which is less than 3% of the average of approximately 1.7 OM%, making it practically negligible. Its standard deviation of 0.554 OM% aligns closely with the cross-validation RMSE of 0.556 OM%, suggesting most errors stem from random scatter rather than bias. The histogram is roughly symmetric and unimodal but more peaked than a normal distribution, with a sparse tail below -2.0 OM%, indicating a leptokurtic distribution in which most predictions are accurate, but some are notably worse. The Q–Q plot shows residuals are normal between about -1 and $+2$ quantiles, but the lower tail is heavier, and the Shapiro–Wilk test ($p = 0.0002$) rejects normality, mainly because of this tail. Note that a flat residuals-versus-predictions trend does not imply overall calibration; a model that shrinks toward the mean may appear conditionally unbiased but still exhibit range-dependent bias tied to observed values, an important consideration for interpreting the spatial organic matter surface.

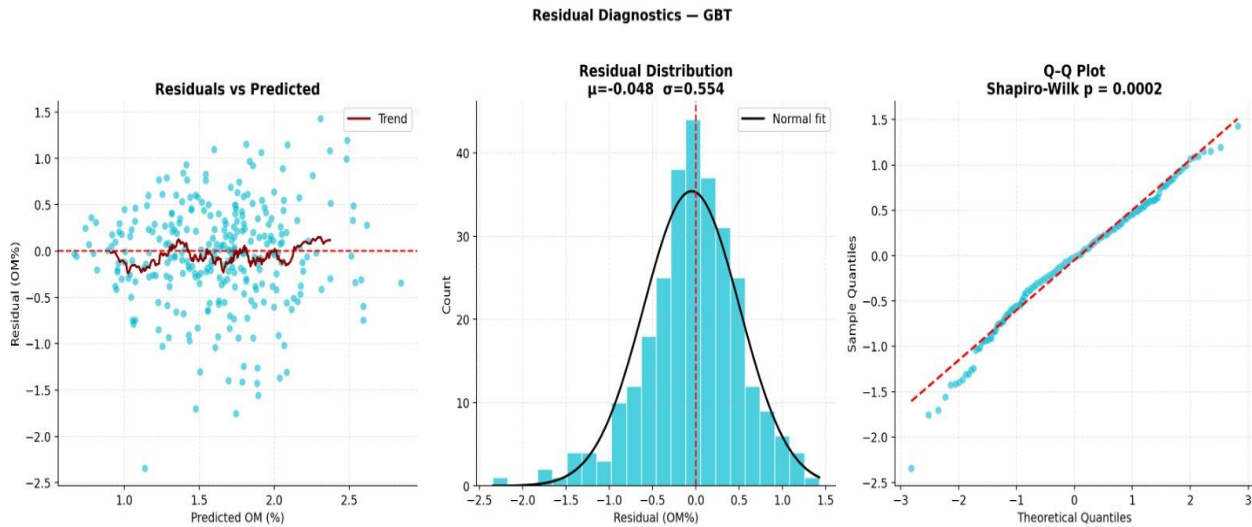


Figure 17. Residual diagnostics for the GBT model: residuals against predicted values with a running trend line (left), the residual distribution with a fitted normal density (centre), and a normal quantile–quantile plot with the Shapiro–Wilk test result (right).

4.1.3.2. Best-Performing Model Selection

Gradient Boosted Trees (GBT) was chosen as the final model for spatially predicting organic matter content because it showed the best point-estimate performance among the six algorithms tested with 5-fold cross-validation, with the lowest MAE and RMSE. However, the performance difference compared to the other five algorithms falls within the noise level caused by random fold assignments. Thus, this choice should be viewed as a practical selection aligned with standard model selection practices, rather than indicating a significant accuracy advantage. The GBT model was then applied to the full 30 m covariate raster stack across the study area to generate the continuous organic matter surface used as input for the EPIC/Williams K-factor equation.

4.1.3.3. Feature Importance (SHAP Analysis)

SHAP analysis was applied to the chosen GBT model to decompose each prediction into additive contributions from individual predictors, demonstrating both the direction and magnitude of each variable's influence on organic matter prediction (Figure 18). The mean absolute SHAP value ranking, shown in the right panel of Figure 18 identifies SWIR _ ratio _ Sum as the most significant predictor, followed by RE _ ratio _ Sum. Next are SWIR _ ratio _ Spr, NDVI _ Aut, and NDVI _ Sum, which form a secondary group with slightly lower importance. A broader cluster of spectral, thermal, and terrain predictors- including LSTmean _ DEM, NDVI _ mean, RE 1 _ Aut, LST _ Win, and RENDVI _ Spr- contribute at similar levels, with remaining predictors gradually decreasing in importance. The beeswarm plot (left panel, Figure 18) illustrates the effect direction: points to the right of zero indicate increased predicted organic matter, while those to the left suggest decreases, with color denoting high (red) or low (blue) predictor values for each sample. Two key insights emerge from this ranking. First, summer acquisitions dominate: the top predictors are summer ratios, and NDVI _ Sum ranks fifth, indicating that the spectral signal of organic matter is most detectable in summer, when vegetation and soil surface conditions are most diagnostic. Second, the importance distribution is relatively flat: the top predictor's importance score is roughly five times that of the lowest, but only about 1.4 times that of the second, showing a gradual decline across predictors. This flat profile, seen in Figure 18, implies no single covariate offers decisive predictive power; instead, the model depends on many weak, partially overlapping signals from spectral, thermal, and terrain data- consistent with the moderate cross-validation accuracy discussed in Section 1- highlighting that the model's predictive limits are set by the information available in the covariates, not by the algorithm's capacity.

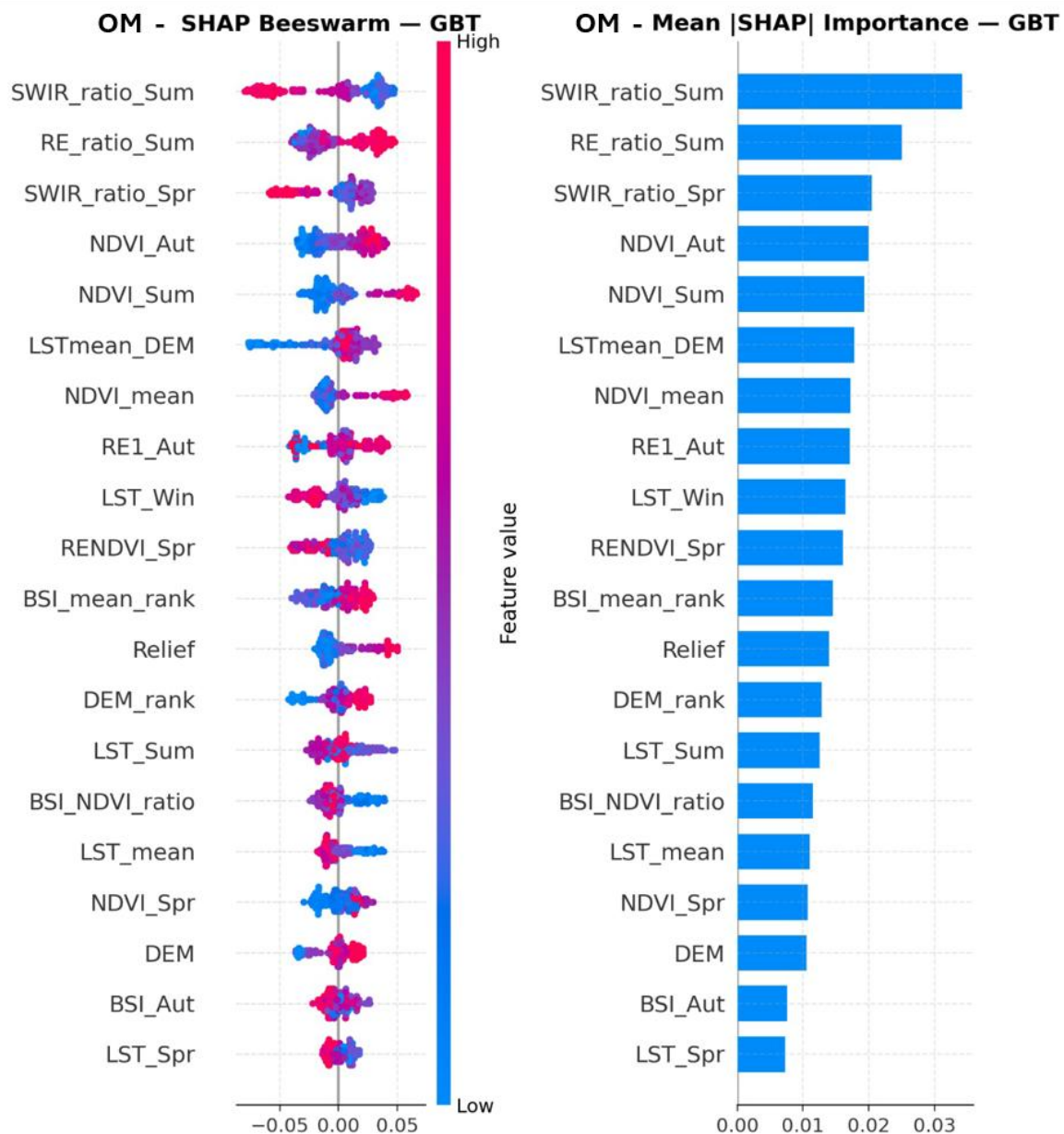


Figure 18. SHAP beeswarm and mean |SHAP| importance for OM.

4.1.3.4. Spatial Distribution of Predicted Organic Matter

A separate modelling approach was applied to predict topsoil organic-matter content, which, together with the texture fractions, completes the inputs required by the EPIC/Williams K-factor equation (OM converted to organic carbon as $OC = OM / 1.724$). Figure 19 shows the predicted organic-matter surface. Consistent with the long history of cultivation and grazing across the Highlands, predicted organic-matter contents are generally low, with the highest values confined to the cooler, moister, and less intensively cultivated positions; this pattern reinforces the erodibility implications discussed below, since low OM reduces aggregate stability and raises soil erodibility.

Predicted OM (OM) content ranges from 0.60% to 3.19% (Figure 19), with a clear and geographically coherent spatial gradient from higher values in the eastern and northeastern zone toward lower values in the western, southwestern, and northern lobe areas. The highest OM concentrations occur precisely in the high-elevation eastern escarpment zone between 39°00'E and 39°30'E, where cooler temperatures associated with elevations approaching 3,018 m a.m.s.l. slow OM mineralisation rates, where natural

vegetation remnants on steep and less-cultivated escarpment slopes contribute greater organic inputs, and where reduced agricultural disturbance relative to the intensively cultivated central plateau preserves higher OM levels. This elevation-OM relationship is well established in the East African highland context, where cooler, wetter high-altitude conditions consistently produce higher OM contents than the warmer, drier low-altitude zones (Hurni, 1988; Nyssen et al., 2004a). Across the western and southern portions of the study area, low OM values approaching the minimum of 0.6% reflect the combined effects of intensive biomass removal for fuel and fodder, prolonged tillage-induced mineralisation, and the warmer temperatures prevailing at lower elevations that accelerate OM decomposition. The spatial distribution of OM across the study area reflects the influence of elevation-driven climate gradients, land use history, and vegetation cover rather than soil texture alone. In temperate, undisturbed soils, clay minerals promote OM stabilisation through organo-mineral complexation. However, in the Eritrean highlands, cooler temperatures and slower mineralisation rates at higher elevations, together with the persistence of natural and semi-natural vegetation on the less-cultivated eastern escarpment, also affect this textural control. These anthropogenic and climatic factors collectively drive OM depletion regardless of the mineralogical capacity for OM retention (Lal, 2004).

(Jakab et al., 2016) showed via laboratory rainfall simulation experiment that soil erosion selective mobilizes organic carbon from all aggregate size classes. The highest proportional loss of SOC (~200%) was observed in the finest-grained fraction. This suggests that erosion does not simply remove soil based on its OM content but preferentially depletes the OC-rich fine fraction. Such findings emphasize the importance of accurately characterizing the spatial distribution of OM across the study area. In this case, ML prediction was used to achieve that, which is essential for reliable K-factor estimation and erosion risk assessment.

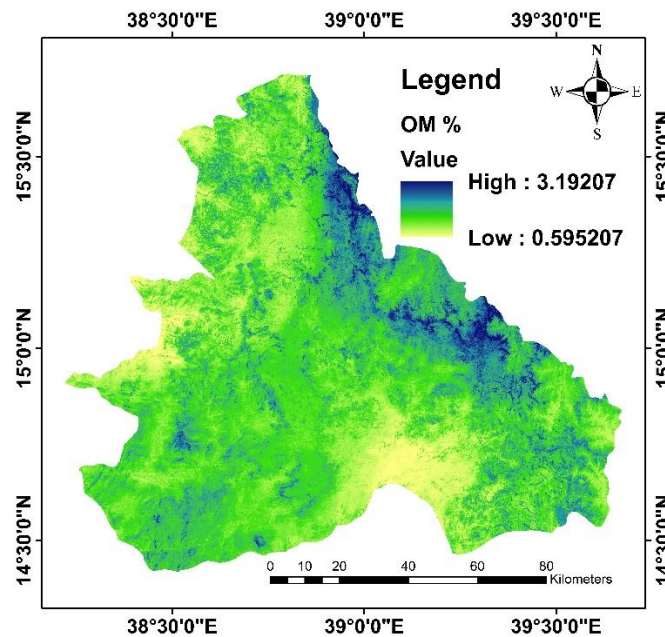


Figure 19. Predicted topsoil organic-matter content (%) map

Collectively, the predicted sand, silt, clay, and organic-matter surfaces provide the spatially continuous soil-property basis that the sparse 277-point field dataset alone could not supply, and they are the direct inputs to the spatial computation of the soil-erodibility factor. In providing the first such dataset for the Eritrean Highlands, they also represent a stand-alone contribution to the regional soil-information base, with utility beyond the erosion application of this dissertation, for example in land-suitability assessment and hydrological modelling.

4.1.4. Machine Learning Improvement Over Conventional interpolation and uncertainty

The spatial prediction of soil properties through ML regression addresses a fundamental data limitation that would otherwise prevent meaningful RUSLE K-factor parameterisation across the Eritrean study area. With 302 soil sampling locations distributed across approximately 10,000 km² an average density of approximately one sample per 33 km² conventional spatial interpolation methods such as ordinary kriging or inverse distance weighting (IDW) would produce unreliable continuous soil property surfaces across the heterogeneous terrain of the highlands, where parent material type, slope position, elevation, and land use change substantially over distances of only a few kilometres. The sub-regions of Senafe, Tsorena, and Adi Keyih, for example, differ markedly in their soil texture and OM from the sub-regions of Asmara City, Ghala Nefhi, and Berikh differences driven by contrasting geology, elevation, land use history, and erosion intensity and these differences would be smoothed beyond recognition by interpolation algorithms that have no mechanism for incorporating the environmental gradients that control soil property distributions.

ML regression overcomes this limitation by exploiting the systematic spatial relationships between soil properties and their environmental controlling factors, encoded in the DEM-derived terrain parameters and satellite-based spectral covariates, to generate predictions that honour the underlying pedogenetic logic of the landscape even in areas remote from sampling locations. The resulting soil property maps capture the full range of variability documented in the 302 MOA/NARI profiles from the low-OM, high-silt, high-clay conditions of the central plateau sub-regions to the high-OM, high-sand, low-silt conditions of the eastern escarpment sub-regions and distribute these contrasts spatially in a pattern that is physically consistent with the geology, topography, climate, and land use gradients of the study area. This spatial coherence, manifest across all four predicted soil property maps simultaneously, provides the primary evidence that the ML models have captured genuine soil-landscape relationships rather than generating prediction artefacts.

The K-factor spatial pattern that emerges from the ML-predicted soil properties with maximum erodibility in the intensively cultivated central sub-regions of Asmara City, Ghala Nefhi, and Berikh, and minimum erodibility in the sandy, organic-matter-richer eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih has direct and consequential implications for the interpretation of the soil loss maps. Had a spatially uniform K-factor been applied as would be necessary without ML spatial prediction given the insufficient interpolation density this critical spatial interaction between soil erodibility, topography, and land use would have been entirely suppressed. The uniformly high K soils of the central plateau sub-regions would have been indistinguishable from the lower-K eastern escarpment soils, and the compound erosion risk created by the combination of high K, high R, and seasonally high C in the central agricultural zone would have been misrepresented. The ML approach therefore contributes not only to the spatial completeness of the K-factor characterisation but also to the physical realism of the soil loss patterns in both the simulated and non-simulated frameworks.

Notwithstanding these contributions, several sources of uncertainty in the predicted soil property maps must be acknowledged. The distribution of the 302 MOA/NARI sampling points across the study area may not fully represent all sub-regions equally: if certain sub-regions particularly the most topographically extreme and remotely located areas of the Senafe and Tsorena eastern escarpment and the May Mine and Adi Kwala southern margins are underrepresented in the training dataset, then predictions in these areas carry higher uncertainty than predictions in well-sampled zones around Asmara City, Mendefera, and Dekemhare. This spatial sampling bias propagates into the K-factor values and ultimately into the soil loss estimates, potentially introducing localised inaccuracies in predicted erosion at sub-regional scales. Additionally, the ML predictions represent soil property conditions as captured at the time and depth of the MOA/NARI sampling, which may not fully reflect actual 2021 surface soil conditions in areas subject to active tillage, overgrazing-induced surface disturbance, or recent erosional stripping of topsoil in highly erodible sub-regions such as Dibarwa,

Areza, and Kudo Bu'er. These are intrinsic limitations of any ML soil mapping approach applied in a data-scarce region and are not unique to this study. Despite these uncertainties, the ML-based K-factor map represents a substantially more spatially realistic and physically meaningful characterisation of soil erodibility across the Eritrean central and southern highlands than any alternative achievable from the available 302-point dataset, and its integration into the RUSLE framework constitutes an important methodological contribution to soil erosion assessment in data-scarce highland environments of the Horn of Africa.

4.2. RUSLE factor results

4.2.1. Rainfall Erosivity Factor (R-factor): Monthly variability vs Annual Distribution

4.2.1.1. GPM IMERG Precipitation

The GPM IMERG Final Run Version 07 monthly and annual precipitation data for 2021 show a seasonal, spatially variable rainfall pattern across the study area (Figures 20 & 21). In January and February, the central plateau and western zones are dry, while the eastern escarpment regions of Senafe, Tsorona, and Adi Keyih get intermittent moisture from the Indian Ocean in February. March marks the start of the Azmera/ minor rainy season, with rainfall concentrated in Adi Teklezan and Logo Anseba. April peaks with pre-monsoon rainfall on the eastern escarpment due to Indian Ocean moisture, while the central plateau and southwest remain nearly dry. May begins the rainfall transition in the southwest, reaching Logo Anseba, Areza, May Mine, and Adi Kwala. June introduces the Kremti monsoon, with the southwest regions seeing the highest totals and the central plateau receiving meaningful rainfall for the first time. July is the rainiest, especially in the central-western and southwest zones. August sees the second-highest totals, shifting peak rainfall to May Mine, Areza, and Adi Kwala. In September, high rainfall persists in the southwest, while the central plateau's rainfall decreases as the monsoon retreats. The eastern escarpment records lower September precipitation, explaining the vegetation paradox that month. October marks a steep rainfall decline across all regions, ending the Kremti season. November and December have elevated post-monsoon rainfall only in the eastern escarpment, driven by residual Indian Ocean moisture. Overall, annual precipitation ranges from about 250 mm in the eastern escarpment to 785 mm in the southwest, with the central plateau and northwest zones recording 350–500 mm.

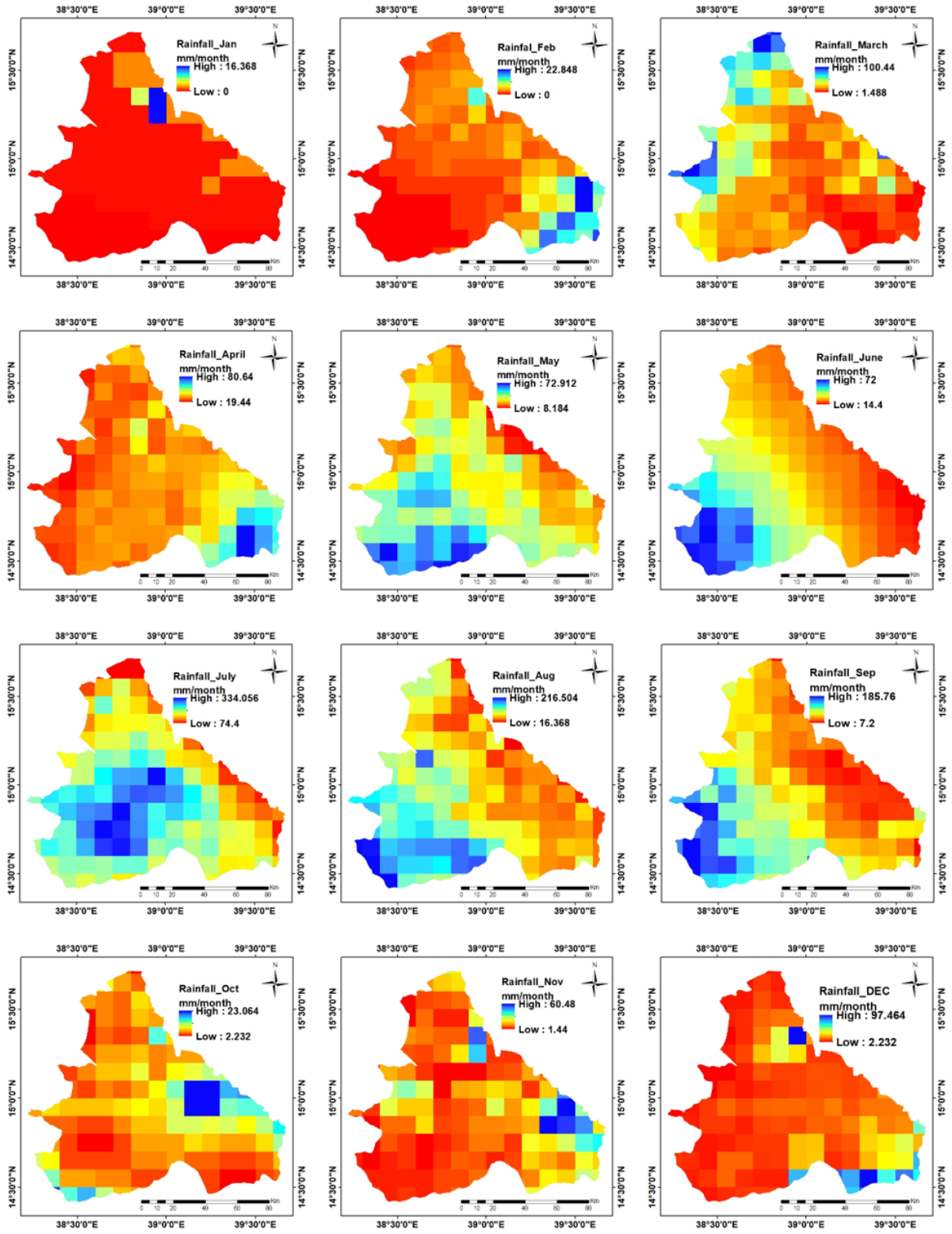


Figure 20. IMERG GPM 12-month rainfall maps for the year 2021.

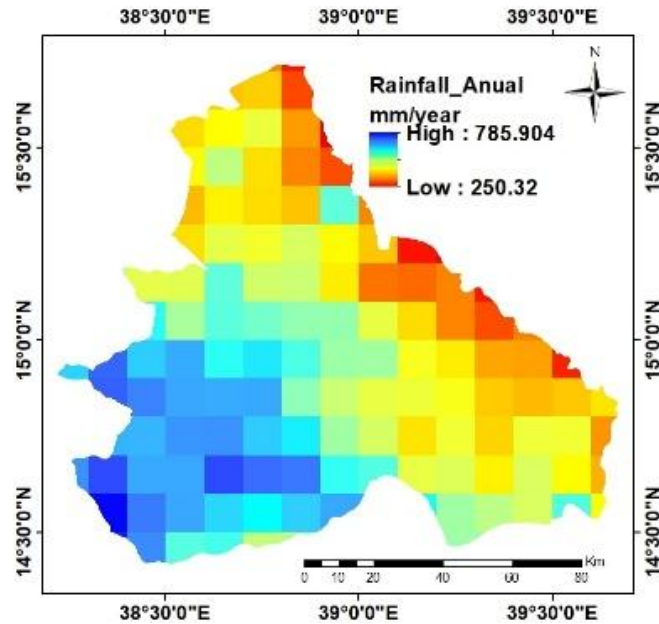


Figure 21. IMERG GPM Annual Rainfall map for the year 2021.

4.2.1.2. Monthly Simulated R-Factor Maps: Seasonal Dynamics

The 12 monthly simulated R-factor maps (Figure 22) (R_{sim_m} for m = January to December 2021) reveal an extreme seasonal concentration of rainfall erosivity that is one of the defining characteristics of the Eritrean highland climate and a primary motivation for the temporal simulation framework adopted in this study. The monthly R-factor values, expressed in $MJ\ mm\ ha^{-1}\ h^{-1}\ yr^{-1}$ because of the annualisation procedure ($P_m \times 12$) applied before entering the empirical erosivity equation, range from near-zero in January to a maximum of 2,200 in the central-western sub-regions during July, a spatial and temporal contrast that defines the scope of the seasonal erosion problem in the study area. Table 4.2 summarises the range, peak sub-region, and dominant spatial character of each monthly R-factor map.

The dry season months of January and February record the lowest monthly R-factor values of the year, with R_{sim} ranging from approximately 0 to 100 in January and 0 to 150 in February. In January, near-zero erosivity (red class, 0–1.6) covers virtually the entire study area, with only isolated small patches of slightly elevated values in the north-central zone around Serejeka and Berikh in the Maekel region, and a minor patch in the eastern zone around Adi Keyih. In February, the spatial pattern remains dominated by very low erosivity, with red (0–5.2) persisting across the western sub-regions of Logo Anseba, Dibarwa, Areza, May Mine, and Adi Kwala. A notable exception occurs in the eastern and southeastern sub-regions of Senafe, Tsorena, and Adi Keyih, where February R_{sim} values reach 88–150, the highest values in the study area for this month reflecting the influence of occasional moisture incursions from the Red Sea and Indian Ocean affecting the eastern highland margins during the dry season. This early February elevation of eastern sub-region erosivity, which is not present in the January map, represents the first indication of a distinctive seasonal pattern that will persist through the analysis: the eastern escarpment sub-regions exhibit a different seasonal rainfall rhythm from the western and central sub-regions.

March marks the onset of the minor rainy season, and this transition is clearly visible in the R_{sim} map for March, with values ranging from 1.9 to 670 $MJ\ mm\ ha^{-1}\ h^{-1}\ yr^{-1}$. Crucially, the highest March erosivity (dark blue, 280–670) is concentrated in the northwestern sub-region of Logo Anseba (Gash-Barka) and in the northern Anseba sub-region of Adi Teklezan, rather than in the central or eastern zones. The central Maekel sub-regions of Asmara City, Ghala Nefhi, and Berikh show moderate orange values (58–110), while the southern Debub sub-regions of Areza, May Mine, Adi Kwala, and Kudo Bu'er and the eastern sub-regions of Senafe and eastern Segeneyiti record the lowest March R values

(red, 1.9–57). This northwest-to-southeast gradient in March erosivity is consistent with the known spatial pattern of the Azmera/ minor rainy season, which predominantly affects the northern and western portions of the Eritrean Highlands through moisture advection from the west and north, while the southern and eastern zones remain comparatively dry during this period.

In April, the R-factor map shows a substantially different spatial organisation compared to March, with values ranging from 120 to 540 MJ mm ha⁻¹ h⁻¹ yr⁻¹. Notably, the minimum value of 120 in April exceeds the maximum value of 100 in January, confirming the seasonal R increase through the pre-monsoon transition. The highest April R-factor values (dark blue, 420–540) are concentrated in the southeastern sub-regions of Senafe, Tsorena, and Adi Keyih, reversing the northwest-high pattern of March and reflecting a shift in moisture sources associated with early pre-monsoon precipitation from the Indian Ocean sector. The central sub-regions of Asmara City, Ghala Nefhi, Berikh, Serejeka, Mendefera, and Dekemhare show predominantly moderate orange values (190–230), while the western sub-regions of Logo Anseba, Dibarwa, Areza, and May Mine record the lowest April values (120–180). The April spatial inversion relative to March with the east now highest and the northwest lowest illustrates the complexity of the inter-seasonal moisture transition and the contrasting roles of westerly and easterly moisture sources in driving erosivity across the study area during the pre-monsoon period.

May shows a further spatial reorganisation, with values ranging from 47 to 480 and the highest R concentrated in the southwestern sub-regions of May Mine, Areza, Adi Kwala, Kudo Bu'er, and Logo Anseba. This southwestward shift of peak erosivity in May, which differs both from the northwestern peak of March and the southeastern peak of April, reflects the increasingly dominant influence of the advancing monsoon system over the southern and western sub-regions during the pre-monsoon period. The central Maekel sub-regions show light blue to light green values (170–290), while only a single small red patch (47–97) persists in the northern part of Serejeka, confirming that by May, no part of the study area retains the near-zero erosivity characteristic of the deep dry season.

June marks the formal onset of the main Kremti rainy season and produces R_{sim} values ranging from 89 to 480 MJ mm ha⁻¹ h⁻¹ yr⁻¹. The June map is characterised by the clearest and most structurally coherent west-to-east spatial gradient of any month in the annual cycle. The highest June erosivity (dark blue, 370–480) is strongly concentrated in the southwestern corner of the study area, encompassing the sub-regions of May Mine, Adi Kwala, Logo Anseba, and the southwestern margins of Areza. Values decline progressively eastward through light blue (300–360) in the south-central zone, light green (220–290) across the central plateau, orange (160–210) in the central Maekel sub-regions of Asmara City, Ghala Nefhi, and Dekemhare, and red (89–150) in the northern sub-region of Adi Teklezan and the eastern escarpment sub-regions of Senafe, Adi Keyih, and eastern Segeneyiti. June is therefore the month in which the spatial distribution of rainfall erosivity is most clearly aligned with the west-to-east moisture gradient of the advancing monsoon front, and it is also the month that constitutes the critical erosion window in the Eritrean seasonal cycle when rising R coincides with incomplete vegetation cover on freshly tilled soils in the central agricultural sub-regions.

July is unequivocally the month of maximum rainfall erosivity in the study area for 2021, with R_{sim} values ranging from 490 to 2,200 MJ mm ha⁻¹ h⁻¹ yr⁻¹ the highest values observed in any month and substantially exceeding the values of all other calendar months. The concentration of the highest July R (dark blue, 1,800–2,200) in the central-western zone primarily across Asmara City, Ghala Nefhi, Dibarwa, and parts of Berikh and Mendefera identifies the central Maekel plateau as the zone of peak monsoon erosivity in 2021. This central-western concentration of July R_{sim} is consistent with the known behaviour of the Eritrean Kremti season, in which the most intense convective rainfall cells are driven by low-level moisture convergence over the central highland plateau. Surrounding bands of light blue (1,500–1,700) encompass Areza, Kudo Bu'er, and Adi Kwala, while light green (1,300–1,400) covers the south-central zone. Notably, the eastern escarpment sub-regions of Senafe, Tsorena, and Adi

Keyih record the lowest July R-factor values in the study area (red, 490–860), and the northern sub-region of Adi Teklezan also shows low July R. This spatial decoupling the steepest terrain (highest LS) receiving the lowest erosivity during the peak monsoon month is one of the most spatial interactions in the entire RUSLE framework for this study, as it means that maximum LS and maximum R do not coincide in July, moderating the potential for the highest theoretical soil loss values in the eastern escarpment zone.

August is the second most erosive month of the year, with R_{sim} values ranging from 100 to 1,500 MJ mm ha⁻¹ h⁻¹ yr⁻¹. The spatial pattern of August R represents a marked westward shift from the July peak: the highest August values (dark blue, 1,100–1,500) are now strongly concentrated in the southwestern sub-regions of May Mine, Areza, Adi Kwala, Logo Anseba, and Dibarwa, rather than in the central Maekel zone that dominated in July. Light blue (760–1,000) extends across the central sub-regions of Asmara City, Ghala Nefhi, Berikh, and Mendefera, while light green (550–750) reaches into the north-central and some eastern areas. The eastern escarpment sub-regions of Senafe, Tsorena, Adi Keyih, and Segeneyiti, along with the northern sub-region of Adi Teklezan, record the lowest August R-factor values (red, 100–360), maintaining the pattern established in July where the steepest terrain experiences the lowest monsoon erosivity. This westward progression of peak erosivity from July to August reflects the seasonal dynamics of the ITCZ and associated moisture transport, which by August has shifted to produce maximum convective rainfall intensity over the western and southwestern sub-regions.

September records R_{sim} values ranging from 40 to 1,200 MJ mm ha⁻¹ h⁻¹ yr⁻¹, continuing the westward concentration of peak erosivity. The highest September values (dark blue, 770–1,200) are now strongly concentrated in the western zone encompassing Logo Anseba, May Mine, Areza, Dibarwa, and Adi Kwala the most westerly sub-regions of the study area. A clear east-to-west gradient is visible, with values declining from the high southwestern zone through light blue (510–760) in the south-central area, light green (330–500) across the central Maekel sub-regions, and orange (180–320) in the north-central area, reaching the lowest values (red, 40–170) in the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih and the northern zone of Adi Teklezan. The persistence of very low R in the eastern sub-regions through the main monsoon season of July, August, and September despite their high elevation and steep terrain — is a consistent and geographically important finding of this analysis. The contrast between the high September R in Logo Anseba and Areza (770–1,200) and the low September R in Senafe and Adi Keyih (40–170) illustrates the fundamental asymmetry of monsoon rainfall distribution across the study area.

October marks the onset of the post-monsoon dry season, with a sharp decline in R_{sim} to 6.9–150 MJ mm ha⁻¹ h⁻¹ yr⁻¹ values approximately 15 times lower than the July maximum, confirming the rapid termination of erosive monsoon rainfall in 2021. The map is dominated by orange (23–42) and red (6.9–22) tones across most of the study area, with only a localised dark blue patch (98–150) in the central-eastern zone around Dekemhare and Segeneyiti, and a light blue zone (63–97) in the Adi Keyih area. The dramatic October decline from September levels confirms that the effective erosion season in the Eritrean Highlands is concentrated in four months (June through September), and that the remaining eight months contribute only marginally to annual erosivity totals.

November and December complete the annual cycle, both recording very low R-factor values consistent with the deepening dry season. In November, R_{sim} ranges from 1.6 to 400 MJ mm ha⁻¹ h⁻¹ yr⁻¹, with the notable spatial characteristic that the highest November values (dark blue, 240–400) are concentrated in the eastern sub-regions of Adi Keyih and Senafe a reversal of the monsoon-season spatial pattern in which these sub-regions had the lowest R. This eastern concentration of late dry-season erosivity reflects occasional post-monsoon rainfall events affecting the eastern highland margins from the Indian Ocean sector. December shows R_{sim} values of 6.9–650, with a localised elevated patch (430–650) in the north-central zone around Serejeka and Berikh of the Central/Maekel region a

December anomaly that may reflect isolated convective events associated with the early onset of the Azmera/ minor rains in the north-central sub-regions. The western and southwestern sub-regions of Areza, May Mine, Adi Kwala, Logo Anseba, and Dibarwa record near-zero R in December (red, 6.9–45), confirming the complete dry season character of the western zone at this time of year.

4.2.1.3. Seasonal Summary: Monthly Ranking and the Dominance of the Summer/Kremti season

Reading across all 12 monthly maps (Figure 22), the seasonal ranking of rainfall erosivity from lowest to highest monthly R-factor can be established for 2021: January < February < October < November < December < March < June < April < May < September < August < July. This ranking, based on the peak values observed across each monthly map, confirms the extreme dominance of the four main monsoon months (July, August, September, and to a lesser extent June) in determining the annual erosion potential of the study area. The combined erosive contribution of July alone, with a peak R_{sim} of 2,200, dwarfs the entire annual non-simulated R of the most erosive sub-regions (350–430), demonstrating the fundamental importance of representing monthly erosivity variability in the modelling framework rather than collapsing it into a single annual value.

A critical observation from the monthly analysis concerns the ranking anomalies of March and June. March records a maximum R_{sim} of 670 in Logo Anseba and Adi Teklezan, driven by minor rainy season rainfall in the northwestern zone, but the central and southern sub-regions experience only low to moderate erosivity in March. June, by contrast, shows more spatially uniform moderate-to-high erosivity (89–480 across the whole study area) than March, despite the fact that the area-averaged June R would rank below the area-averaged March R if the northwestern maximum dominates the comparison. This distinction has important implications for the spatial interpretation of the soil loss results: in March, the high R is spatially confined to Logo Anseba and Adi Teklezan where vegetation cover may still be sparse following the dry season, while in June, the moderate-to-high R is spatially widespread and coincides with the critical window of incomplete crop establishment across the intensively cultivated central sub-regions of Asmara City, Ghala Nefhi, Mendefera, and Dekemhare.

4.2.1.4. Annual Non_Simulated R-Factor: Spatial Distribution and Contrast with Monthly Simulation

The annual non-simulated R-factor map (R_{annual}), derived from the total 2021 annual GPM IMERG precipitation through the same empirical erosivity equation applied in the monthly simulation, shows values ranging from 130 to 430 MJ mm ha⁻¹ h⁻¹ yr⁻¹ (Figure 23). The spatial pattern of R_{annual} reflects the total annual precipitation distribution across the study area, with the highest annual erosivity (dark blue, 350–430) concentrated in the southwestern sub-regions of May Mine, Areza, Adi Kwala, Logo Anseba, Kudo Bu'er, and Dibarwa; the zone that accumulates the highest total annual rainfall by virtue of receiving substantial pre-monsoon (May–June), main monsoon (July–September), and some post-monsoon precipitation. Light blue values (290–340) extend across the south-central and central sub-regions of Mendefera, parts of Asmara City, and Ghala Nefhi. Light green (240–280) is widespread across the central Maekel plateau (Asmara City, Ghala Nefhi, Berikh, parts of Dekemhare, Segeneyiti). The lowest annual R values (red, 130–190) are found in the northern sub-region of Adi Teklezan and, most notably, in the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih.

The spatial pattern of R_{annual} ; with the southwestern zone highest and the eastern escarpment lowest, is the product of the summed annual precipitation, which integrates the contributions of all 12 months. The relatively low annual R in Senafe, Tsorena, and Adi Keyih (130–190) reflects the consistent pattern observed across multiple months in which these sub-regions receive below-average rainfall during the main Kremti monsoon season, despite receiving somewhat higher rainfall than the western zone during the February–April pre-monsoon period and during November. The southwestern sub-regions of May Mine and Adi Kwala, by contrast, benefit from high rainfall contributions across the May–September period, accumulating a large annual total that produces the highest R_{annual} in the study area.

The contrast between the monthly simulated R and the annual non-simulated R is stark and fundamental to the rationale of this study. The July peak R_{sim} reaches $2,200 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$ in the central-western zone, while the annual non-simulated R_{annual} in the same sub-regions reaches only approximately $290\text{--}340 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$; less than one-sixth of the simulated July value. This difference arises directly from the annualisation procedure: $R_{\text{sim_July}}$ is computed from $P_{\text{July}} \times 12$, which for a July total of, for example, 100 mm gives $P_{\text{sim_July}} = 1,200 \text{ mm yr}^{-1}$ and a correspondingly high R , while R_{annual} is computed from the actual annual total of perhaps 400 mm yr^{-1} , producing a much lower R . The annual approach therefore assigns the same erosivity to every month of the year; the temporal average implied by $R_{\text{annual}} \div 12$; rather than the month-specific erosivity that actually prevailed. For the peak monsoon month of July in the central-western sub-regions, this temporal averaging suppresses the actual erosivity by a factor of more than six, fundamentally underrepresenting the most erosive conditions of the annual cycle. Conversely, for the dry months of January and October, the annual approach assigns the same temporal-average erosivity that it assigns to July, substantially overestimating the actual near-zero erosivity of these months. It is this redistribution of erosivity across the year, inflating low-erosivity months and suppressing high-erosivity months, that constitutes the central methodological limitation of the non-simulated approach in a strongly seasonal environment, and whose spatial consequences are quantified.

The annual non-simulated R -factor ($130\text{--}430 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$) falls within the range documented by published RUSLE studies across the Ethiopian Highlands and East Africa for comparable mean annual rainfall environments. (Hurni, 1988) reported annual R values of $200\text{--}600$ for the Ethiopian Highlands under mean annual rainfall of $700\text{--}1,400 \text{ mm yr}^{-1}$, while (TEFERI2, 2009) documented R values of $373\text{--}746$ in the wetter Blue Nile watershed. The lower range observed in the present study; consistent with the relatively modest mean annual rainfall of the Eritrean central and southern highlands compared with the Ethiopian Blue Nile basin; is physically plausible and provides an additional line of evidence supporting the validity of the GPM IMERG-based R -factor computation. The spatial gradient from the southwestern zone (highest annual R , $350\text{--}430$) to the eastern escarpment and northern fringe (lowest annual R , $130\text{--}190$) also corresponds to the rainfall gradients that would be expected from the ITCZ-driven seasonal rainfall climatology of the region, in which the southwestern highlands receive the combined influence of multiple moisture pathways while the northern and eastern margins are more protected by rain-shadow and orographic effects.

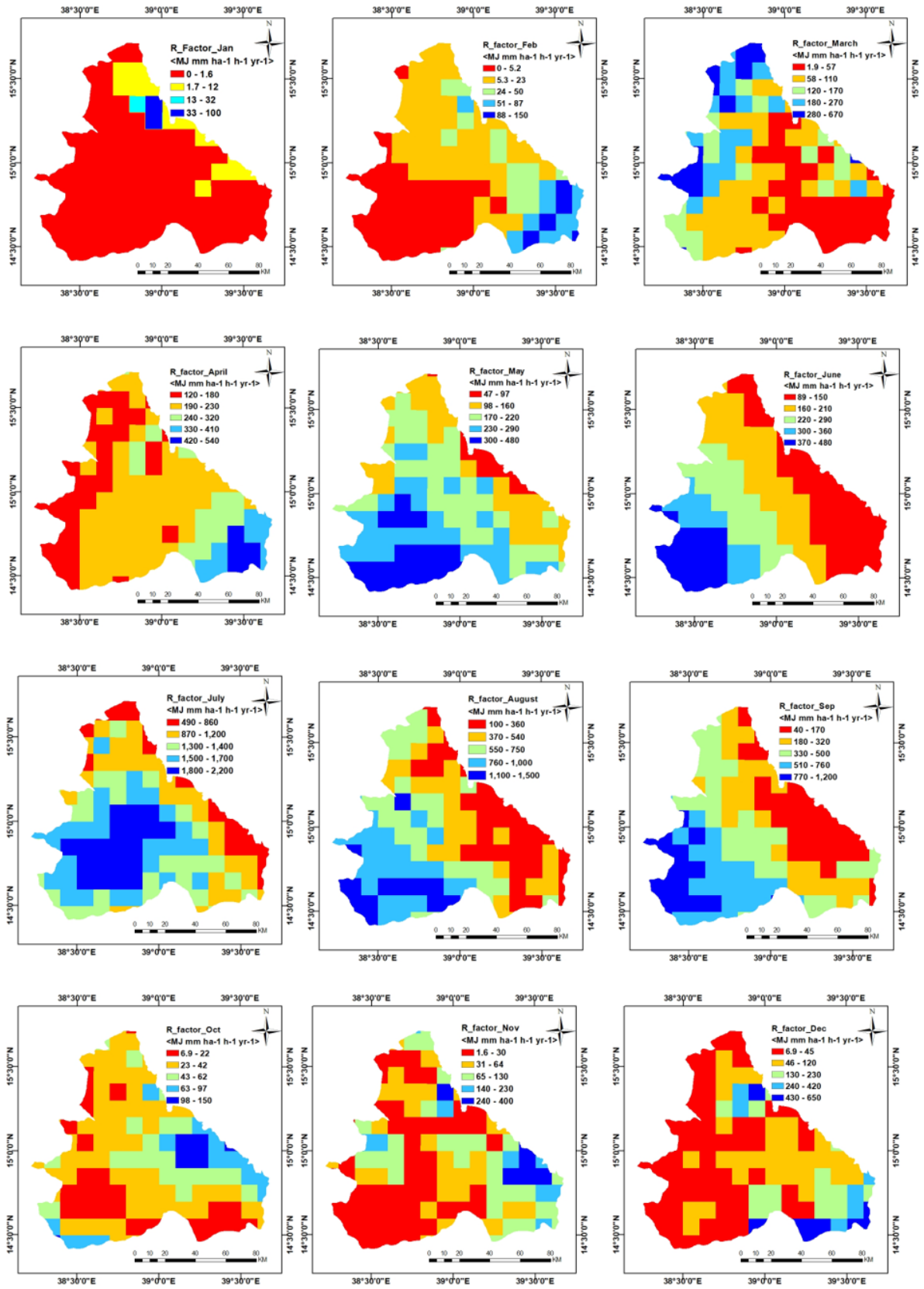


Figure 22. Monthly Simulated R-Factor map for the year 2021.

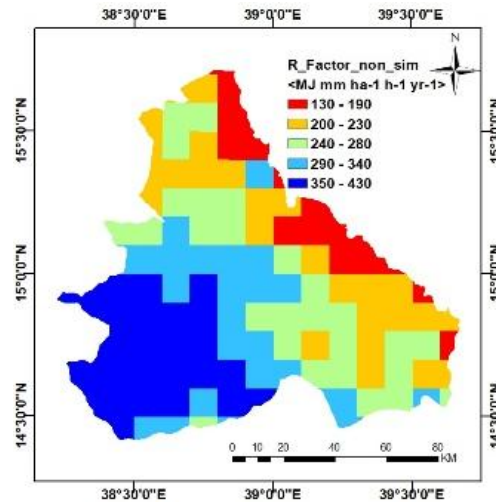


Figure 23. Non-simulated *R*-Factor map for the year 2021.

4.2.2. Derived K-Factor Map

The K-factor map computed from the EPIC/Williams empirical formulation applied to the four ML-predicted soil property rasters reveals a spatially heterogeneous distribution of soil erodibility across the study area, with values ranging from 0.024 to 0.048 t ha h ha⁻¹ MJ⁻¹ mm⁻¹ (Figure 24). This twofold range in K-factor values represents a meaningful and spatially structured gradient in intrinsic soil resistance to erosion that, through the multiplicative structure of RUSLE, will exert a proportional influence on the spatial pattern of modelled soil loss in both the simulated and non-simulated frameworks.

The highest K values, approaching 0.048 t ha h ha⁻¹ MJ⁻¹ mm⁻¹ and represented by orange-brown tones, are concentrated in a spatially coherent zone in the central and central-western portion of the study area, broadly corresponding to the sub-regions of Asmara City, Ghala Nefhi, Berikh, and the northern parts of Dekemhare and Mendefera, the intensively cultivated agricultural heartland of the central Mael and northern Debub zones. This zone of maximum erodibility is a direct expression of the soil property convergence: it is precisely where clay content is highest, silt fractions are elevated, sand content is minimal, and OM is at its lowest values across the study area. Within the EPIC/Williams formulation, the combination of high silt and high clay with low sand and low OM generates maximum K values because silt-sized particles are inherently the most detachable fraction of the soil under raindrop impact they are fine enough to be readily mobilised by raindrop kinetic energy yet coarse enough to remain in suspension in overland flow while the scarcity of OM further reduces aggregate stability and surface resistance to structural breakdown upon wetting (Sharpley, 1985; Williams J. et al., 1981). The practical implication is that the sub-regions of Asmara City, Ghala Nefhi, Berikh, and Dekemhare, which are also the most densely cultivated and populated areas of the study domain, carry both the highest inherent soil erodibility and the highest land-use pressure of the entire study area, creating a compound vulnerability that will be amplified whenever R and C factors are also elevated.

In direct contrast, the lowest K values approaching 0.024 t ha h ha⁻¹ MJ⁻¹ mm⁻¹ and represented by dark-blue are concentrated in the northeastern and eastern sector, specifically across the sub-regions of Senafe, Tsorena, Adi Keyih, and eastern Segeneyiti. This spatial coincidence of minimum erodibility with the eastern escarpment zone is physically coherent and of considerable importance for the interpretation of soil loss patterns. The high sand fractions, low silt content, and relatively high OM levels that characterise these eastern sub-regions collectively produce low EPIC/Williams K values by multiple mechanisms: high sand content increases inter-aggregate porosity and promotes rapid infiltration, reducing runoff generation and detachment susceptibility; low silt content removes the most detachable particle-size fraction from the mix; and higher OM stabilises soil aggregates through organo-

mineral complexation, increasing resistance to raindrop impact. The result is that the sub-regions with the steepest slopes and highest LS values in the entire study area; Senafe, Tsorena, and Adi Keyih also have the lowest soil erodibility, producing a partial but substantial counterbalance between the topographic amplification of erosion potential and the soil resistance to that amplification. This K-LS trade-off in the eastern escarpment zone is one of the most important spatial interactions in the RUSLE framework for this study area and would not be correctly represented without the spatially continuous K-factor predictions generated by the ML approach.

The southern and western sub-regions including Areza, May Mine, Adi Kwala, Kudo Bu'er, Dibarwa, and Logo Anseba display intermediate K values in the green-yellow range, broadly between approximately 0.030 and 0.040 $t\ ha\ h\ ha^{-1}\ MJ^{-1}\ mm^{-1}$, reflecting their intermediate soil property conditions with moderate sand, moderate clay, lower silt, and low-to-moderate OM. In these sub-regions, soil erodibility is neither maximised as in the central plateau nor minimised as in the eastern escarpment, meaning that the spatial pattern of erosion risk in the southern and western Debub zone is shaped by the combined interaction of all RUSLE factors rather than being dominated by any single one. The northern sub-region of Adi Teklezan (Anseba region) and in the central Maekel sub-region of Serejeka show intermediate-to-moderately-high K values, reflecting their transitional soil property conditions and moderate silt fractions.

The K-factor range of 0.024–0.048 $t\ ha\ h\ ha^{-1}\ MJ^{-1}\ mm^{-1}$ observed in this study is consistent with published K-factor values for comparable soil types in the Ethiopian Highlands and East Africa more broadly. (Hurni, 1988) reported K values for Highland Ethiopian soils in the range of 0.02–0.06, with the highest values associated with silt-rich, low-OM profiles on cultivated slopes a soil condition closely analogous to the central Maekel plateau sub-regions of this study. (TEFERI2, 2009) documented K values of 0.02–0.05 across the Chemoga watershed in northern Ethiopia for soils developed on similar metamorphic parent materials, while (Gelagay & Minale, 2016) reported comparable ranges in the Koga watershed. The spatial pattern in the present study with maximum K in the intensively cultivated central plateau and minimum K in the higher-elevation sandy escarpment soils also mirrors findings from (Ayehu et al., 2018) for Ethiopian Highland catchments where elevation and parent material contrasts drive systematic K-factor gradients.

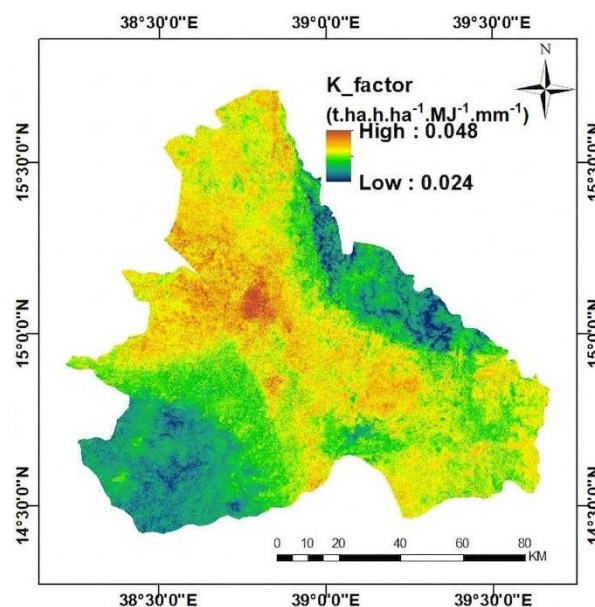


Figure 24. Derived soil erodibility (*K*) factor map for the study area at 30 m resolution, in SI units $t\ ha\ h\ ha^{-1}\ MJ^{-1}\ mm^{-1}$

4.2.3. Cover Management Factor (C-factor)

Figures 27 and 28 present the monthly (January–December) and annual C-factor maps derived from NDVI. These maps show the spatial and temporal variability of vegetation cover conditions incorporated into the RUSLE model to estimate soil erosion across the study area.

4.2.3.1. Monthly and mean NDVI

The 2021 monthly and annual mean Sentinel-2 NDVI composites show a clear seasonal pattern in vegetation dynamics across the study area, driven by the rain-fed agricultural cycle and the timing of the Kremti monsoon (Figures 25 & 26). During the dry season (January–February), NDVI values are generally low (0.05–0.20), indicating mostly bare or sparsely vegetated surfaces. Vegetation begins to green with the monsoon's start in June, continuing through July and August as crops develop and rainfall rises. Interestingly, the highest NDVI values occur in September, not during July or August, despite those months receiving the most rainfall, implying that vegetation responds with a delay to peak rainfall and that total seasonal moisture, rather than monthly rainfall, mainly influences maximum canopy development. The data also suggest that extensive, perennial vegetation capable of maintaining green biomass year-round is rare, as NDVI drops near zero during the dry season across most of the landscape, with only narrow riparian corridors and some exclosure areas remaining green. The eastern escarpment consistently shows lower peak NDVI values compared to the central and southwestern zones throughout the monsoon, reflecting less rainfall in that region. The average NDVI map, created from the 12 monthly composites, captures the overall seasonal vegetation status and serves as the basis for the standard annual C-factor calculation.

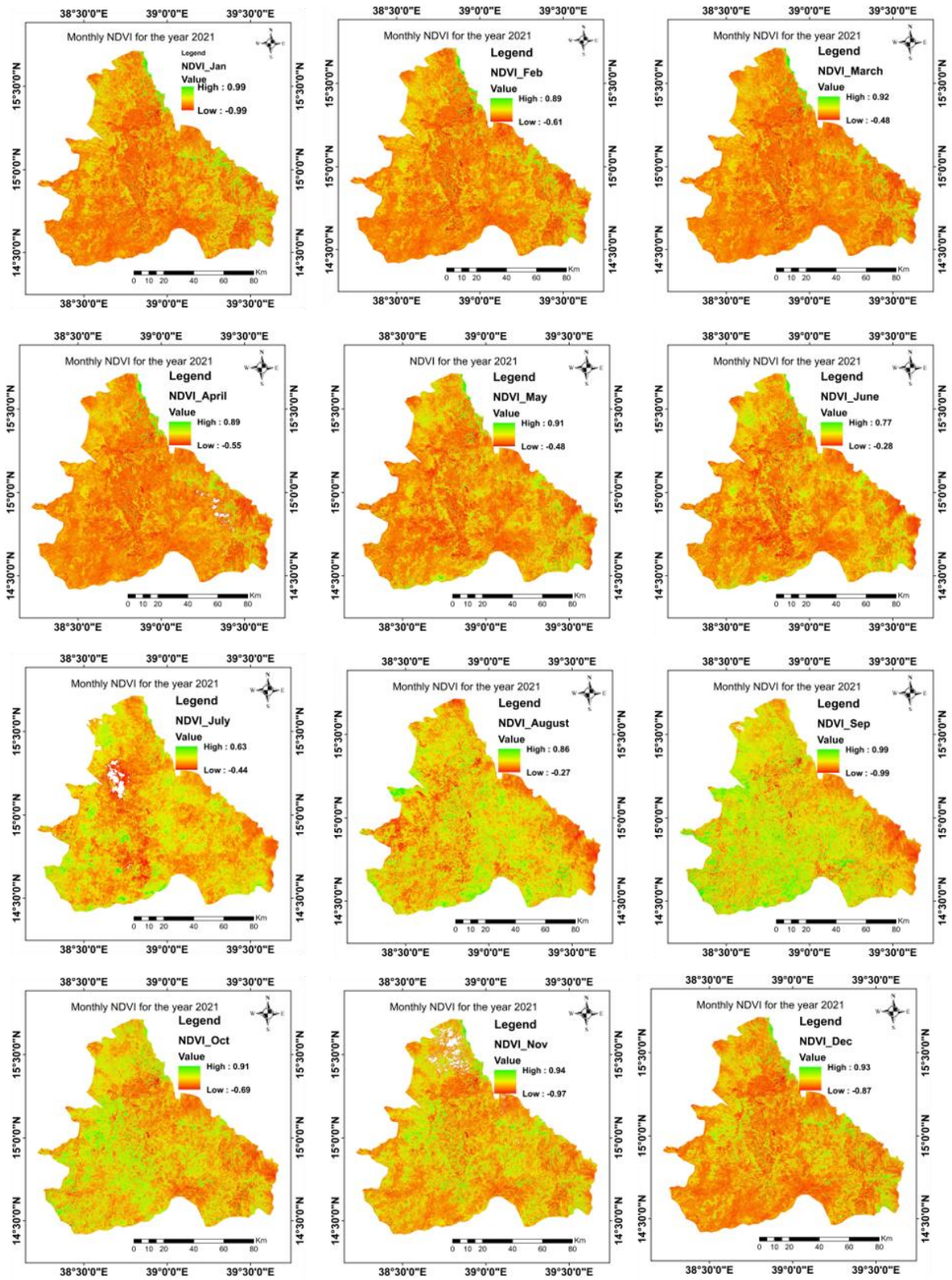


Figure 25. All 12-month NDVI maps for the year 2021.

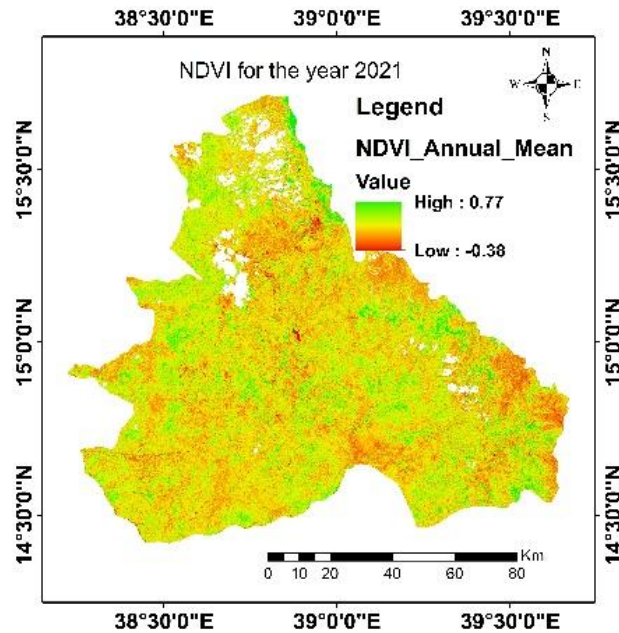


Figure 26. Mean NDVI map for the year 2021

4.2.3.2. Monthly C-Factor Maps: Seasonal Dynamics

Dry Season (January-February)

January and February record the highest C-factor values of the annual cycle, representing the season of greatest soil surface vulnerability. Across the broad plateau interior of the study area, C values of approximately 0.6–0.9 prevail for the majority of pixels, reflecting the near-complete absence of green vegetation cover on cultivated fields and communal grazing lands throughout the central and western sub-regions. The western and south-western sub-regions of Areza, May Mine, Adi Kwala, Logo Anseba, and Dibarwa display the highest dry-season C values; approaching 0.8–1.0; confirming that communal grazing lands and degraded scrublands in these zones are the most completely bare and erosion-vulnerable surfaces during this period. The only consistent exception is found in the drainage-line riparian corridors concentrated in the Adi Teklezan sub-region, the Serejeka drainage networks, and the eastern escarpment valleys of Senafe and Adi Keyih, where permanently vegetated riparian woody vegetation maintains C values approaching 0 even in the depths of the dry season. The spatial character of January and February is essentially identical: the two months represent the dry-season baseline condition of the study area, with high C values (0.6–0.9) dominating the plateau and very low C values (approaching 0) confined to isolated drainage corridors.

Minor Rainy Season Transition (March-May)

The onset of the minor Azmera/ minor rainy season in March produces the first detectable spatial differentiation of C-factor values within the annual cycle. In the northern sub-region of Adi Teklezan and along the eastern escarpment drainage lines of Senafe and Adi Keyih, C values begin to decline toward 0.3–0.5 as early vegetation growth responds to Azmera/minor moisture, while the central agricultural sub-regions of Asmara City, Ghala Nefhi, Berikh, Mendefera, and Dekemhare retain C values of approximately 0.5–0.8, indicating that the Azmera rains have not yet stimulated meaningful vegetation response on the cultivated plateau. By April, this spatial differentiation intensifies: northern and eastern sub-regions show C values in the range of 0.2–0.4, while the central agricultural plateau maintains C values of 0.4–0.7, confirming that cultivated fields remain fallow or recently tilled without protective canopy. May continues this progression, with C values of 0.2–0.4 spreading more broadly across the northern and eastern sub-regions as pre-monsoon moisture drives wider vegetation growth.

Across the western zone, including Areza and May Mine, C values remain relatively high at 0.5–0.8 through May, reflecting the limited pre-monsoon rainfall received in these sub-regions. Throughout the March–May period, the central Maekel agricultural sub-regions maintain C values of 0.4–0.7; insufficient to provide substantial soil protection against any erosive event that might occur during this transitional period.

Monsoon Onset (June): The critical Erosion Window

The June C-factor map presents the most agriculturally consequential spatial pattern of the annual cycle. The central agricultural sub-regions of Asmara City, Ghala Nefhi, Berikh, Mendefera, and Dekemhare display C values of approximately 0.4–0.8, corresponding to freshly tilled fields and recently sown but unestablished crops at the monsoon onset. Although isolated vegetated patches in the northern Adi Teklezan, Serejeka, and eastern Senafe/Adi Keyih sub-regions show lower C values of 0.1–0.3, the spatially dominant condition across the intensively cultivated central plateau is one of moderate-to-high C, with the minimum monthly C constrained to 0.00117; a value determined mathematically by the maximum observed June NDVI of 0.77, which, when substituted into the Van der Knijff equation, yields $C = \exp[-2 \times (0.77/0.23)] \approx 0.001$. This means that even the single most densely vegetated pixel in the entire study area cannot achieve a C-factor below approximately 0.001 in June. In the south-western sub-regions of May Mine, Areza, and Adi Kwala, C values of approximately 0.3–0.5 reflect partially developing pre-monsoon vegetation that provides some but insufficient protection at the monsoon onset. The net result of the June C-factor spatial distribution is a study area in which the sub-regions carrying the highest soil loss potential; the central agricultural plateau; present C values of 0.4–0.8 at precisely the moment when monsoon erosivity is rising to its seasonal onset.

Peak Monsoon (July and August)

July is the most critical month of the entire annual C-factor cycle. Despite representing the month of maximum rainfall erosivity in the study area, the July C-factor maps show that the central plateau sub-regions of Asmara City, Ghala Nefhi, and Berikh; where R_{sim} reaches its annual maximum; still display C values of approximately 0.4–0.7, corresponding to crop canopies that have been sown following the monsoon onset but have not yet achieved sufficient density to provide substantial protective cover. Cloud contamination causes data gaps in parts of the central zone during July, an artefact of active convective cloud associated with peak monsoon rainfall; where data are available, C values in the affected sub-regions confirm the persistence of moderate-to-high C. The south-western sub-regions of Areza and May Mine show partially developed lower C values of approximately 0.2–0.4, reflecting early canopy establishment, while the eastern escarpment sub-regions of Senafe and Adi Keyih display C values of approximately 0.3–0.6. The defining quantitative characteristic of July is its minimum C-factor value of 0.0292; the highest minimum C of all 12 calendar months, and the only month in which the Van der Knijff equation, constrained by the maximum observed July NDVI of 0.63, cannot produce C values below approximately 0.03 anywhere in the study area: $C = \exp[-2 \times (0.63/0.37)] \approx 0.033$. This means that in July 2021, no pixel anywhere in the approximately 10,000 km² study area achieves a C-factor below 0.03, and the dominant C values of 0.4–0.7 across the central plateau confirm that the crop canopy established after June–July planting has not yet provided meaningful soil protection during the month of highest rainfall erosivity.

August represents the key transition in the seasonal C-factor cycle. C values drop substantially from July levels across the central sub-regions, with areas like Asmara City, Ghala Nefhi, Mendefera, Areza, May Mine, Adi Kwala, and Logo Anseba reaching C values around 0.1–0.3 as crop canopies achieve sufficient density to offer meaningful soil protection. The lowest C in August, close to zero (3.16×10^{-6}), confirms that the densest crop and herbaceous vegetation now provide near-complete soil cover. However, the eastern escarpment sub-regions, Senafe, Tsorena, and Adi Keyih, maintain higher C

values of approximately 0.3–0.6, due to lower rainfall and a less vigorous vegetation response during the monsoon season.

Late Monsoon and Vegetation Peak (September)

September represents the annual minimum of C-factor values and the month of maximum soil surface protection across most of the study area. The central sub-regions of Asmara City, Ghala Nefhi, Berikh, Mendefera, and Dekemhare, and the south-western sub-regions of Areza, May Mine, and Adi Kwala, reach C values of approximately 0.05–0.15, with the minimum C of 0 confirming that some pixels achieve complete theoretical protection as full crop canopy closure coincides with peak herbaceous cover on grazing lands. These low C values in the central and western zones stand in notable spatial contrast to the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih, which display C values of approximately 0.3–0.6; substantially higher than the central plateau at the same time of year. This eastern escarpment exception is directly attributable to the lower September rainfall documented in the R-factor analysis: with less accumulated monsoon moisture, these sub-regions have developed less dense vegetation cover by September, retaining C values two to four times higher than the well-watered central and south-western zones at the seasonal vegetation peak.

Post-Monsoon Transition (October–December)

October marks the beginning of the post-monsoon C-factor recovery toward dry-season levels. C values across the study area rise to approximately 0.2–0.5 as crop residues are progressively removed and natural vegetation begins to senesce, with the central and southern sub-regions retaining lower C values (0.2–0.3) from the residual monsoon canopy while the northern and western margins rise more rapidly toward 0.4–0.5. By November, C values continue to rise across the western sub-regions of Areza, May Mine, and Logo Anseba, reaching approximately 0.4–0.6, while the central agricultural zone retains moderately lower values of 0.2–0.4 from remaining crop residues. Data gaps in the Adi Teklezan sub-region in November propagate from the same cloud contamination already identified in the November NDVI composite. December returns most of the study area to C values of approximately 0.4–0.7 across the central and western plateau, effectively re-establishing the dry-season baseline, while the eastern escarpment drainage lines of Senafe and Adi Keyih and the northern Adi Teklezan and Serejeka corridors retain C values of approximately 0.1–0.3, reflecting the persistence of residual dry-season vegetation in the drainage-constrained positions that maintain lower C throughout the year.

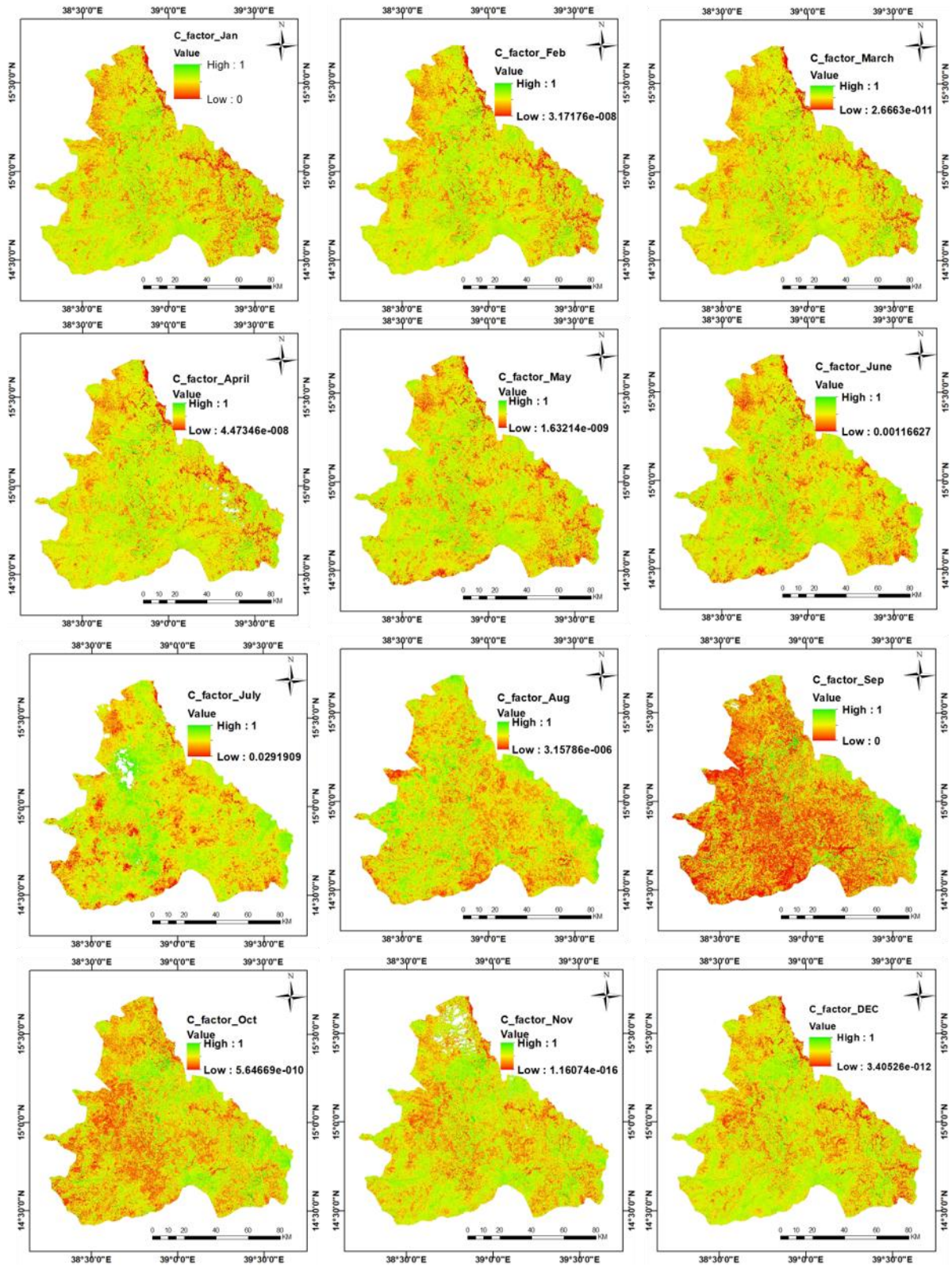


Figure 27. All 12-Months' C-factor map for 2021.

4.2.3.3. Mean Annual C-Factor: Spatial Distribution and Temporal Information Loss

The mean annual C-factor map (C_{annual}), computed as the arithmetic mean of all 12 monthly C-factor rasters, ranges from 0.0094 to 1 and constitutes the cover management input for the non-simulated RUSLE computation (Figure 28). The spatial distribution of C_{annual} reflects the cumulative effect of the seasonal cycle: the northern Adi Teklezan sub-region displays the highest mean annual C values, approaching 0.6–0.9, because its limited monsoon vegetation response produces high C values throughout most of the year with only a modest seasonal reduction. The central agricultural sub-regions of Asmara City, Ghala Nefhi, Berikh, Mendefera, and Dekemhare show intermediate mean C values of approximately 0.3–0.5, averaging the high C conditions of the dry season and early monsoon months with the low C conditions of August–September. The south-central and south-western sub-regions, benefiting from vigorous monsoon vegetation growth, show lower mean C values of approximately 0.1–0.3, reflecting the stronger downward pull of their low August–September C on the annual average. White data gaps in the Adi Teklezan sub-region reproduce the November composite gap in the arithmetic mean.

The mean annual C minimum of 0.0094 reveals the most important limitation of the non-simulated approach. When C_{annual} is applied uniformly across all 12 months in the non-simulated RUSLE computation, this near-zero annual mean is assigned to the most persistently vegetated pixels in every month of the year. In July, when the actual minimum C across the entire study area is 0.0292; approximately three times higher than the annual mean minimum; the non-simulated model assigns substantially more vegetation protection to these pixels than actually existed during the peak erosion month of 2021. Conversely, for the same pixels in September, the annual mean assigns a higher C value than the actual near-zero C prevailing when vegetation protection is at its annual maximum. These twin temporal misrepresentations; under estimating C in July (by a factor of approximately three at the most vegetated pixels) and overestimating protection in September; operate simultaneously and in opposite directions, demonstrating that a single annual C-factor cannot represent either end of the seasonal range with sufficient accuracy for reliable soil loss estimation in a strongly seasonal environment.

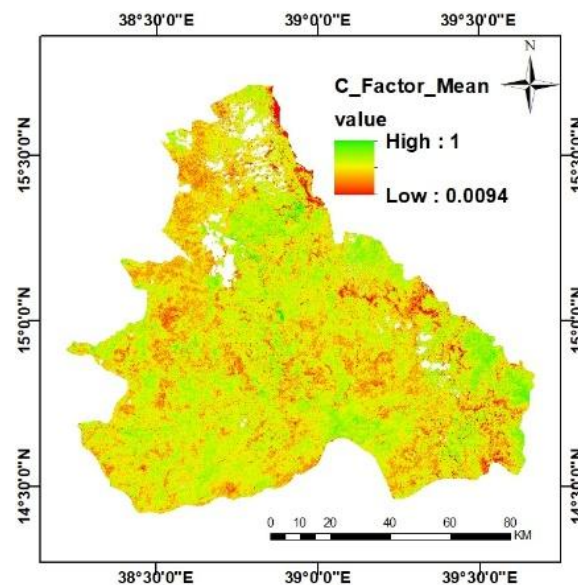


Figure 28. Annual mean C-Factor map for the year 2021

Taken together, the 12 monthly C-factor maps and mean annual C-factor document a seasonal vegetation protection cycle of substantial amplitude and spatial complexity. C values transition from their annual high of approximately 0.6–0.9 across the plateau in January–February through moderate values of 0.3–0.6 in the minor-rainy-season transition, to their annual low of approximately 0.05–0.15

across the central and western sub-regions in September, before recovering to dry-season levels by December. The full range of this cycle; from C values approaching 1 on bare dry-season surfaces to C values approaching 0 under dense September canopy; cannot be represented by any single annual C-factor value. The monthly C-factor time series derived from Sentinel-2 NDVI composites, paired with the monthly simulated R-factor, provides the temporal resolution necessary to correctly represent this cycle and to capture the critical June–July erosion window in which C values of 0.4–0.8 persist across the central agricultural sub-regions during the months of highest rainfall erosivity.

4.2.4. LS-factor and spatial controls

The LS-factor map derived from the DEM using slope and flow accumulation analyses reveals a strongly skewed spatial distribution of topographic erosion potential across the study area, with values ranging from 0 on flat valley floor positions to a maximum of 65.41 on the steepest escarpment faces (Figure 29). As a dimensionless multiplicative factor integrating slope steepness and upslope contributing area, the LS-factor attains its highest values wherever steep slope gradients and large flow-accumulation areas converge conditions that are geomorphologically most pronounced along the deeply incised escarpment margins and the valley network dissecting the central highland plateau.

The dominant spatial characteristic of the LS-factor map is the overwhelming prevalence of near-zero values represented by deep green tones across the broad plateau interior of the study area. This pattern directly reflects the gently rolling to flat topography of the central plateau surface, confirmed by the slope map on which the 0–6.6° class dominates the interior. The sub-regions of Asmara City, Ghala Nefhi, Berikh, Mendefera, Dekemhare, and Dibarwa, which collectively occupy the central plateau, exhibit predominantly low LS values, indicating limited topographic amplification of erosion potential in these zones. In these areas, the spatial distribution of modelled soil loss will be governed primarily by the soil erodibility (K-factor), rainfall erosivity (R-factor), and cover management (C-factor) rather than by topographic controls alone.

In sharp contrast, the highest LS values are spatially concentrated in three geomorphologically distinct zones that correspond directly to the steepest terrain identified on the DEM and slope maps. The most extensive and continuous zone of high LS is the eastern and southeastern escarpment margin, encompassing the sub-regions of Senafe, Tsorena, Adi Keyih, and the eastern margins of Segeneyiti. In this zone, the plateau descends steeply over short horizontal distances from elevations approaching 3,018 m a.m.s.l. toward the lower terrain of the Debubawi Keih Bahri region to the east. The combination of very steep gradients frequently exceeding 34°, as confirmed by the red pixels dominating the eastern margin on the slope map and large upslope contributing areas from the central plateau generates LS values approaching the study-area maximum of 65.41. This eastern escarpment belt therefore constitutes the zone of greatest topographic erosion amplification in the entire study area, and its spatial extent from approximately 39°00'E eastward to the study boundary at 39°30'E makes it a geographically prominent and consequential erosion-risk zone.

A secondary area with high LS values is observable in the northern part of Adi Teklezan (Anseba region), the central Maekel sub-region of Serejeka, and the Logo Anseba sub-region in Gash-Barka. In these northern and northwestern zones, the plateau edge slopes downward toward the Anseba River basin and the Gash-Barka plains, creating escarpments and valley walls that display scattered yet distinct high LS values. While this northwestern elevated LS zone is less continuous and has lower intensity than the eastern escarpment belt, it is geographically important because it overlaps with sub-regions where agriculture on steep slopes is common. This increases erosion risk, especially during heavy rains when vegetation cover is sparse.

The third configuration of elevated LS values comprises a network of fine, branching linear features of orange and red pixels that thread throughout the plateau interior in all directions, visible across the sub-regions of Mendefera, Areza, Kudo Bu'er, Adi Kwala, May Mine, and Dibarwa in the southern and

western Debub zone, as well as within the central Maekel sub-regions. These linear features trace the incised valley walls of the drainage network cutting into the plateau surface, where even modest vertical incision generates locally steep gradients and concentrates upslope flow accumulation along the channel margins. The individual width of these valley-wall LS anomalies is narrow, but their cumulative spatial extent across the entire drainage network of the study area is substantial, and their contribution to total annual soil loss may be substantial, particularly where they intersect with the high-K zones of the central plateau.

The high maximum LS value of 65.41 observed in this study is consistent with published values from structurally analogous highland environments in East Africa. (Gelagay & Minale, 2016; TEFERI2, 2009) working in Ethiopian Highland catchments with comparable dissected plateau morphology, reported maximum LS values exceeding 50 in areas of steep escarpment terrain. The spatial dominance of low LS across the plateau interior combined with the concentration of extreme values along the eastern escarpment and valley network is also a characteristic feature of RUSLE applications in dissected highland settings, reflecting the fundamental geomorphological principle that topographic erosion amplification is spatially selective rather than uniformly distributed (Ayehu et al., 2018).

The spatial structure of the LS-factor carries a fundamental implication for both the simulated and non-simulated soil loss frameworks: the LS raster is identical in both approaches, serving as the common spatial template on which the temporal signals of the R and C factors are superimposed. In the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih, the very high LS values will multiplicatively amplify whatever R and C conditions are applied, producing high absolute soil loss estimates in both frameworks. In the central plateau sub-regions of Asmara City, Ghala Nefhi, and Mendefera, the low LS values limit the topographic contribution to soil loss, making the temporal variability of R and C captured by the simulation but suppressed by the annual approach relatively more important in determining the difference between the two modelling frameworks in these zones. The LS-factor therefore simultaneously acts as the primary spatial driver of maximum erosion concentrations and as a spatial modifier of the sensitivity of soil loss estimates to the temporal treatment of other RUSLE parameters.

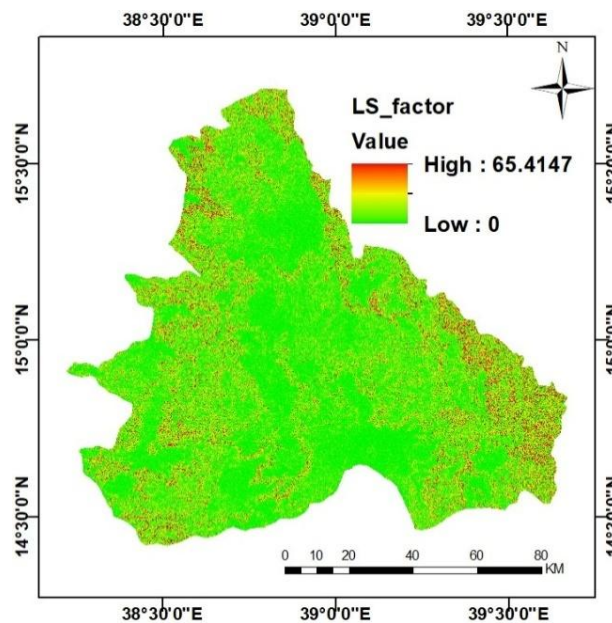


Figure 29. Slope Length and Steepness Factor (LS-Factor) map.

4.3. Simulated Soil Erosion (spatiotemporal)

4.3.1. Monthly soil-loss

The simulated soil erosion framework generates 12 monthly potential soil loss rasters (Figure 30) (A_{actual_m} , $t\ ha^{-1}\ month^{-1}$), each representing one calendar month's contribution to annual soil loss, computed as $A_{actual_m} = A_{sim_m} / 12$ (see equation 3.14). Together, these maps document the complete temporal dynamics of the erosion process across the study area in 2021. The analysis that follows focuses on the spatial distribution and areal extent of erosion across sub-regions and dominant soil loss classes rather than on isolated extreme pixel values. Approximate area estimates in percentages are derived from visual assessment of each map's spatial class distribution.

Dry Season: January and February

January is the lowest soil loss month of 2021. An estimated 95% or more of the study area falls within the lowest soil loss class of 0 to $0.3238\ t\ ha^{-1}\ month^{-1}$, reflecting the near-zero R-factor that prevails across almost the entire study area during this month. Negligible soil detachment occurs regardless of C, K, or LS values because the absence of erosive rainfall removes the primary driver of soil loss. The only meaningful exception is a geographically concentrated patch in the north-central sub-regions of Serejeka and Berikh, where soil loss of 0.3239 to $13.76\ t\ ha^{-1}\ month^{-1}$ occurs across a small area. This anomaly corresponds to the isolated January R-factor elevation in this zone, most likely attributable to localised minor-season rainfall events generating the only meaningful erosivity in the otherwise dry January landscape.

February maintains predominantly low soil loss across the western and central plateau sub-regions of Logo Anseba, Dibarwa, Areza, May Mine, Adi Kwala, Kudo Bu'er, Mendefera, Asmara City, and Ghala Nefhi, with an estimated 60–65% of the total study area remaining within 0 to $0.7569\ t\ ha^{-1}\ month^{-1}$. However, the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih transition to meaningfully higher soil loss classes spanning 0.757 to $27.57\ t\ ha^{-1}\ month^{-1}$. This eastern escarpment concentration reflects the February R-factor elevation in these sub-regions driven by occasional Indian Ocean moisture incursions, combined with the high C-factor values of the dry season and the steep LS values of the escarpment. The eastern sub-regions therefore generate the only spatially coherent soil loss zone outside the small January Serejeka anomaly during the dry season.

Minor Rainy Season Transition: March to May

March is the first month in which spatially structured soil loss patterns emerge across a meaningful geographic extent. The dominant condition remains low, with an estimated 65–70% of the study area in the lowest soil loss class of 0 to $1.721\ t\ ha^{-1}\ month^{-1}$. However, two distinct hotspot zones appear. The north-western zone, encompassing Adi Teklezan and the northern margins of Logo Anseba, records the highest March values of 26.26 to $109.7\ t\ ha^{-1}\ month^{-1}$ in Classes 4 and 5 across a concentrated area. These elevated values are directly attributable to the Azmera/ minor rainy season delivering high March erosivity in the north-western part of the study area, substantially amplified by the steep slopes and high LS values along the dissected north-western plateau margin. Simultaneously, the eastern escarpment of Senafe and Adi Keyih continues to show moderate-to-high soil loss of 1.722 to $26.25\ t\ ha^{-1}\ month^{-1}$ across a meaningful portion of its area. The southern sub-regions of Areza, May Mine, and Adi Kwala remain in the lowest class throughout March, consistent with their near-zero March R-factor.

April produces a spatial reorganisation toward a single dominant high-erosion zone: the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih become the most extensive and severe soil loss area of the month, with soil loss ranging from 2.676 to $97.46\ t\ ha^{-1}\ month^{-1}$ across a substantial portion of the escarpment. This April eastern dominance reflects the R-factor finding that April erosivity peaks in the south-eastern sub-regions from early Indian Ocean pre-monsoon moisture, combined with the high LS values and moderate C-factor values (0.3 – 0.6) of the escarpment. Cloud data gaps in the eastern zone introduce marginal uncertainty for April, but the dominant spatial pattern is unambiguous.

The remainder of the study area, comprising the central plateau and the western and south-western sub-regions, maintains predominantly Class 1 conditions of 0 to 2.675 t ha⁻¹ month⁻¹, covering an estimated 65–70% of total area.

May shifts the peak erosion zone toward the south-west. The sub-regions of Logo Anseba, Areza, May Mine, and Adi Kwala show soil loss of 4.98 to 60.46 t ha⁻¹ month⁻¹ across a meaningful portion of their area, reflecting the pre-monsoon rainfall shift confirmed in the May R-factor analysis. The eastern escarpment (Senafe, Tsorena, Adi Keyih) retains elevated soil loss of 1.661 to 60.46 t ha⁻¹ month⁻¹. The central plateau sub-regions of Asmara City, Ghala Nefhi, Berikh, Mendefera, and Dekemhare remain predominantly within the lowest class of 0 to 1.66 t ha⁻¹ month⁻¹, covering an estimated 60–65% of the study area, reflecting low pre-monsoon R in the central zone combined with moderate C values of 0.4–0.6.

Monsoon Onset: June

June marks the onset of the main rainy season (Kremti) monsoon and the beginning of the critical erosion window. The critical erosion window refers to the period when the Kremti monsoon causes increasing rainfall erosivity, but vegetation cover in the cultivated zone is still sparse, as fields are prepared and seeds are planted, yet crop canopies haven't fully developed. An estimated 55–60% of the study area remains within 0 to 2.195 t ha⁻¹ month⁻¹. The south-western sub-regions of May Mine, Areza, Adi Kwala, and Logo Anseba record soil loss of 2.196 to 69.97 t ha⁻¹ month⁻¹ across a geographically extensive area, confirming the south-western concentration of June R-factor identified in the factor analysis. Patches of moderate-to-elevated soil loss of 2.196 to 13.99 t ha⁻¹ month⁻¹ also appear along the drainage networks threading through the central plateau sub-regions of Mendefera, Dibarwa, and Dekemhare, where the LS-factor amplifies the rising June R along steep valley walls. The eastern escarpment sub-regions retain elevated values throughout June. The critical process character of this month is the combination of a rising R-factor and still-high C-factor conditions (0.4–0.8 across the central agricultural zone), where freshly tilled fields without established crop canopy meet the onset of erosive monsoon rainfall.

Peak Monsoon: July

July is unambiguously the most erosive month of 2021. The characteristic that distinguishes July from all other months is not merely its highest class values but the composition of its entire distribution: the Class 1 lower boundary spans 0 to 13 t ha⁻¹ month⁻¹, meaning that even the lowest-erosion pixels in July sustain monthly soil loss equivalent to the entire upper class value of January (0.3238 t ha⁻¹). An estimated 50–55% of the study area falls within Class 1 (0–13 t ha⁻¹ month⁻¹); a background level already representing substantial erosion in comparison with any other month. An estimated 30–35% of the study area sustains Class 2 and 3 soil loss of 14 to 74 t ha⁻¹ month⁻¹, distributed widely across the western, south-western, and central zones. The highest soil loss values of 75 to 400 t ha⁻¹ month⁻¹ (Classes 4 and 5), accounting for an estimated 10–15% of the study area, are spatially concentrated in the western and south-western sub-regions of Logo Anseba, Areza, May Mine, and Adi Kwala, where the highest July R-factor (R_{sim} 1,800–2,200), moderate-to-high C values (0.4–0.7), and elevated LS values along valley walls converge. The eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih contribute Classes 2 to 4 soil loss of 14 to 130 t ha⁻¹ month⁻¹, driven by the K-LS spatial decoupling documented in section 4.2.2.. Cloud data gaps in parts of the central Maekel zone (Asmara City, Berikh) limit the spatial completeness of the July map but do not alter the fundamental spatial pattern.

Late Monsoon: August and September

August sustains the second-highest monthly soil loss of the year but with a distinct spatial reorganisation. The central plateau sub-regions (Asmara City, Ghala Nefhi, Berikh, Mendefera) decline

to the August Class 1 level of 0 to 4.1 t ha⁻¹ month⁻¹, covering an estimated 55–60% of the study area, reflecting progressive crop canopy development reducing C values to 0.1–0.3. The highest August soil loss of 29 to 206 t ha⁻¹ month⁻¹ (Classes 4 and 5) is strongly concentrated in the south-western sub-regions of May Mine, Areza, Adi Kwala, and Logo Anseba, consistent with the westward migration of peak monsoon R from the central zone in July to the south-western zone in August. An estimated 10–15% of the study area sustains soil loss above 29 t ha⁻¹ month⁻¹, concentrated in the south-western zone and eastern escarpment. Classes 2 and 3 (4.2–28 t ha⁻¹ month⁻¹) cover an estimated 25–30% of the study area across the western, north-western, and escarpment zones.

September exhibits the most varied erosion patterns during the monsoon season. A key feature is the sharp drop in the central plateau sub-regions to Class 1, with soil loss of 0 to 2 t ha⁻¹ month⁻¹, covering about 65–70% of the area studied. This decline aligns with the September C-factor minimum (around 0.05–0.15 across Asmara City, Ghala Nefhi, and Berikh), which effectively reduces soil loss despite ongoing rainfall. At the same time, the western and south-western sub-regions, including Logo Anseba, Areza, and May Mine, experience soil loss between 2.1 and 130 t ha⁻¹ month⁻¹, driven by high September R combined with C values that, although lower than July, remain above the central plateau minimum. The eastern escarpment zones of Senafe and Adi Keyih also sustain Classes 2 to 5, with soil losses from 2.1 to 130 t ha⁻¹ month⁻¹, due to their higher September C-factors (0.3–0.6) and high LS values. This illustrates the vegetation paradox observed in September, first noted in the C-factor analysis. The eastern escarpment, with the highest erosion potential, receives the least monsoon moisture, limiting canopy growth and keeping C-factors high. Conversely, the central plateau's seasonal C-factor drops to its lowest level (0.05–0.15). Consequently, the area most at risk of erosion has the poorest vegetation cover.

Post-Monsoon: October to December

October records a sharp and comprehensive decline in soil loss across the study area. An estimated 70–75% of the study area falls within Class 1 of 0 to 0.33 t ha⁻¹ month⁻¹; near-identical to January's background level; confirming that the effective erosion season in the Eritrean Highlands ends abruptly in October. The only sub-regions retaining meaningful soil loss are the eastern escarpment zones of Senafe, Tsorena, and Adi Keyih, where soil loss of 0.34 to 17 t ha⁻¹ month⁻¹ persists in Classes 2 through 5, consistent with the October R-factor finding of localised elevated R in the eastern margins.

November retains the eastern escarpment as the primary high-erosion zone, with soil loss of 1.5 to 69 t ha⁻¹ month⁻¹ in Classes 2 through 5 across the sub-regions of Senafe, Adi Keyih, and eastern Segeneyiti. An estimated 70–75% of the study area remains in the lowest class of 0 to 1.4 t ha⁻¹ month⁻¹. Data gaps in the northern Adi Teklezan sub-region prevent assessment of November conditions there. December maintains a similar eastern escarpment concentration, with soil loss of 1.7 to 100 t ha⁻¹ month⁻¹ in Classes 2 through 5 across Senafe, Tsorena, and Adi Keyih, and a secondary zone of elevated soil loss in the north-central Serejeka and Berikh sub-regions consistent with the December R-factor anomaly identified earlier. Approximately 75–80% of the study area remains in Class 1 (0–1.6 t ha⁻¹ month⁻¹) in December, confirming the near-complete reassertion of dry-season erosion conditions across the plateau interior.

Monthly Ranking, Spatial, and the Role of Vegetation

The monthly ranking of simulated soil loss for 2021; July > August > September > June > April > May > March > November > December > February > October > January; closely parallels the monthly R-factor ranking but with two notable exceptions that reveal the modifying influence of the C-factor. June ranks fourth overall in soil loss despite its seventh-position R-factor rank, because the combination of a rising R-factor and high C-factor conditions (0.4–0.8) in the central agricultural zone and south-western sub-regions produces a higher R × C erosion driver than would be predicted from R alone.

Conversely, September ranks third in soil loss despite its third-highest R-factor, because the dense September vegetation across the central plateau ($C \approx 0.05\text{--}0.15$) substantially suppresses soil loss in these sub-regions, pulling September's spatial average downward relative to its R ranking. These exceptions demonstrate that while rainfall erosivity is the primary temporal driver of soil erosion in the Eritrean Highlands, the C-factor exerts a meaningful secondary modulating influence that alters the monthly ranking from the R-factor sequence in months where vegetation dynamics are most pronounced.

The spatial migration of the peak erosion zone through the annual cycle documents a systematic shift of the highest-erosion areas across the study area's sub-regions. The pattern begins with an isolated north-central anomaly in January (Serejeka/Berikh), migrates to the eastern escarpment in February (Senafe, Tsorena, Adi Keyih), transitions to a dual north-western and eastern hotspot in March (Adi Teklezan/Logo Anseba and Senafe/Adi Keyih), returns to the eastern escarpment in April, shifts south-westward in May through June (Logo Anseba, Areza, May Mine, Adi Kwala), expands to study-area-wide high erosion levels in July with a south-western and central-western concentration, maintains south-western dominance in August, produces a multi-zone pattern in September as the central plateau drops and the western and eastern zones retain elevated erosion, before contracting to eastern escarpment residual erosion from October through December. This migration of erosion hotspots across sub-regions through the year is a direct consequence of the temporal dynamics of the Eritrean rainfall system and can only be resolved through the monthly simulation framework. No single annual soil loss map can document these spatial transitions or provide the sub-region-specific, month-specific erosion information needed for temporally targeted SWC planning.

Figure 31 summarizes the seasonal patterns seen in the monthly spatial maps, presenting a single, clear timeline. Throughout the year, rainfall erosivity ($R_{\text{sim}, m}$) remains the main driver of soil loss, rising sharply from near zero during the dry season to about $1,200 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$ in July. This aligns with the intense convective rainfall typical of the Kremti / Summer Rainy Season (June–September). Vegetation greenness ($NDVI_m$) gradually responds to accumulated rainfall, peaking in September, while the cover-management factor (C_m) decreases, hitting its lowest point at the same time. A key point is the phenological lag: early in the rainy season, in June and July, soils are still sparsely vegetated after the long dry period. Consequently, the most erosive rainfall occurs when surface protection is weakest. This leads to a peak in simulated soil loss ($A_{\text{actual}, m}$) around July ($\sim 15 \text{ t ha}^{-1} \text{ month}^{-1}$). The loss then diminishes in August and September as vegetation develops and C_m decreases, even though rainfall remains high. This partial vegetative buffering explains why July's soil loss is higher than September's in the maps. Outside the Kremti season, all parameters return to baseline levels, indicating that erosion in the Eritrean Highlands mainly occurs within this narrow seasonal window, with July contributing a large share of the annual total.

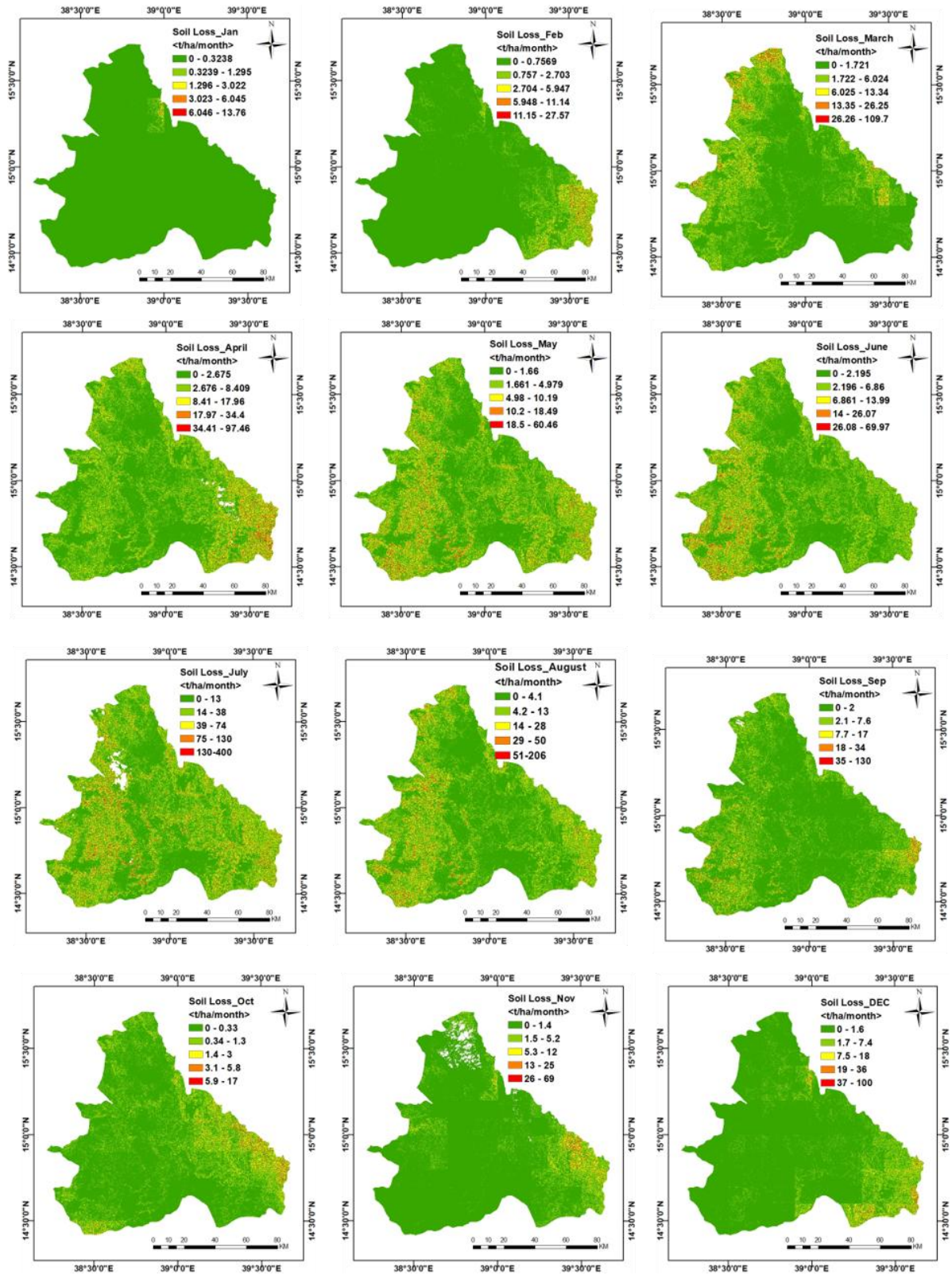


Figure 30. All 12-Months' Soil Loss ($t\ ha^{-1}\ month^{-1}$) maps for the year 2021

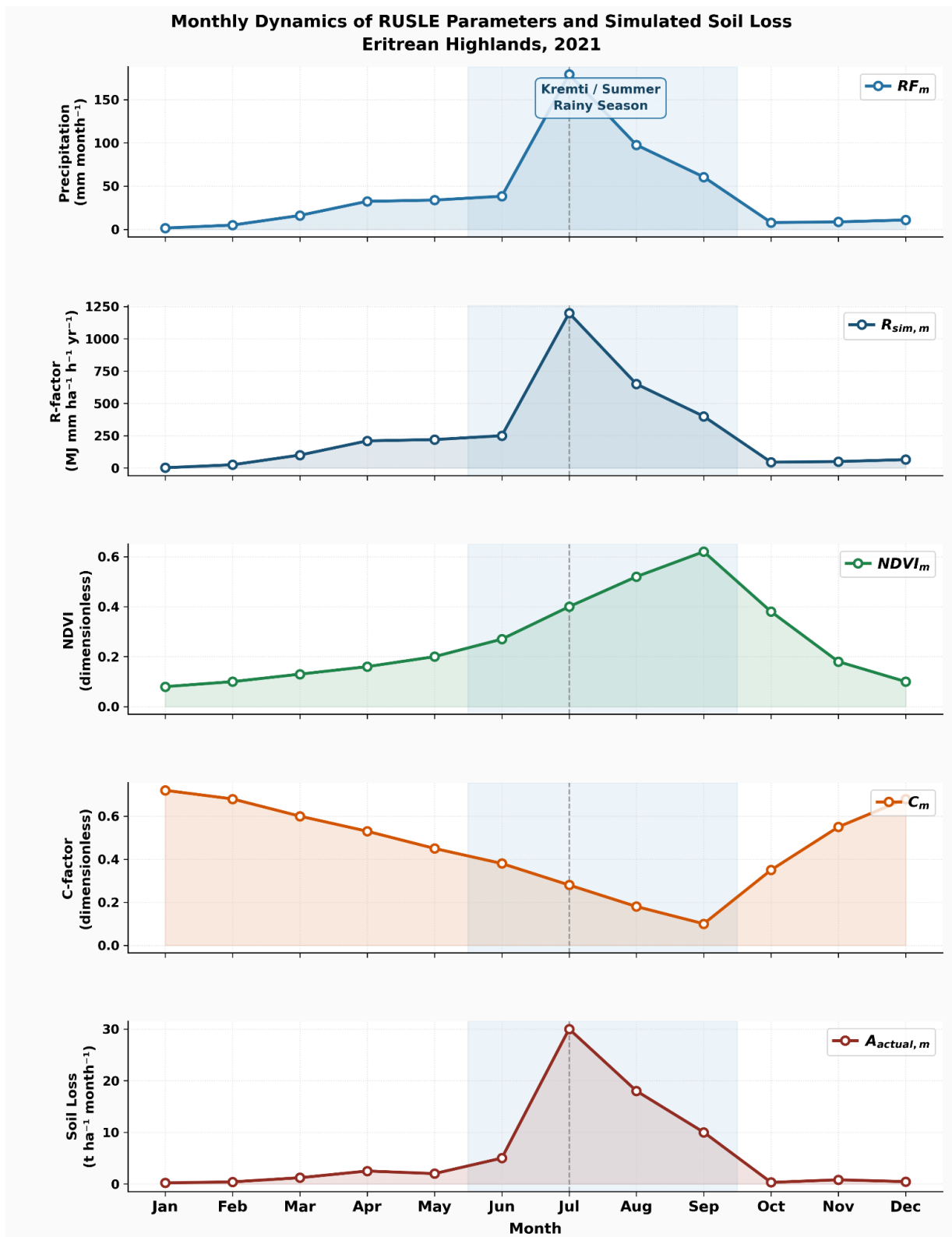


Figure 31. Monthly dynamics of RUSLE input parameters and simulated soil loss across the Eritrean Highlands study area, 2021.

4.3.2. Annual simulated soil loss (A_{annual})

The annual simulated soil loss map (A_{annual}) represents the summation of the 12 monthly potential soil loss contributions (A_{actual_m} , m = January to December 2021) and constitutes the primary output of the spatiotemporal RUSLE simulation framework. It integrates the full temporal variability of rainfall erosivity and vegetation into a single spatially continuous annual soil loss surface, expressed in t ha^{-1}

yr⁻¹. The P-factor was set to unity ($P = 1$) throughout, meaning the results represent potential soil loss under prevailing land cover conditions without accounting for the mitigating effect of any existing conservation measures. The map is classified into five soil loss classes, and the following analysis focuses on the spatial distribution and areal extent of each class across the study area's sub-regions.

4.3.2.1. Overall Spatial Distribution

The most spatially extensive class across study area is Class 1, encompassing soil loss of 0 to 8.4 t ha⁻¹ yr⁻¹ and covering an estimated 45–50% of the total area. This class dominates the central interior plateau, particularly the sub-regions of Asmara City, Ghala Nefhi, and Berikh in the central Maekel zone, the western margins of Mendefera and Dekemhare in the northern Debub zone, and the western sub-regions of Dibarwa and the central portions of Logo Anseba. The predominance of Class 1 across the central plateau reflects the interplay of several RUSLE factors operating simultaneously. Although July erosivity in the central Maekel zone reaches its annual peak (R_{sim} of 1,800–2,200), the dense September vegetation cover across these sub-regions (C values of approximately 0.05–0.15 in September) substantially suppresses soil loss during this late-monsoon month, moderating the annual cumulative total. Additionally, the relatively lower soil erodibility (K values of 0.030–0.040 in intermediate positions) and the predominantly moderate LS values of the plateau surface further limit annual totals in these sub-regions despite their exposure to peak monsoon erosivity. The result is a paradox that only the monthly simulation framework can resolve: some of the sub-regions with the highest monthly peak erosion (July, central Maekel) accumulate moderate annual totals because their September vegetation protection is equally pronounced.

Class 2, spanning 8.5 to 67 t ha⁻¹ yr⁻¹, is the second most spatially extensive class, covering an estimated 35–40% of the study area. This class is broadly distributed in a mixed spatial pattern across the study area, appearing most extensively in the south-western sub-regions of Areza, May Mine, Adi Kwala, Kudo Bu'er, and the southern portions of Logo Anseba, where the accumulation of elevated monthly soil loss contributions across June, July, August, and September produces annual totals in this class range. Class 2 also appears extensively along the drainage networks threading through the central plateau sub-regions of Mendefera, Dibarwa, and Dekemhare, where steep valley walls generate locally high LS values that elevate annual soil loss above the plateau background level. In the eastern sub-regions of Segeneyiti and northern Adi Keyih, Class 2 coverage reflects the accumulation of moderate monthly contributions across a longer active erosion season, including the February, March, April, October, and November elevated values documented in the monthly analysis for these zones. The broad spatial extent of Class 2 — covering approximately one-third of the total study area; confirms that moderate annual soil loss is a pervasive condition across the Eritrean central and southern highlands under the $P = 1$ assumption and not confined to isolated high-erosion patches.

Classes 3 through 5, collectively representing annual soil loss above 68 t ha⁻¹ yr⁻¹, account for an estimated 10–15% of the total study area but represent the zones of most severe potential erosion in 2021. These higher-class areas are concentrated in two geographically distinct zones that reflect different underlying RUSLE factor combinations.

The south-western zone; encompassing the sub-regions of May Mine, Areza, southern Adi Kwala, and Kudo Bu'er; shows the most extensive concentration of Classes 3 through 5, with soil loss of 68 to more than 180 t ha⁻¹ yr⁻¹ across a meaningful proportion of this zone. The elevated annual totals in the south-western sub-regions are directly attributable to the cumulative effect of consistently high monthly soil loss contributions across a sustained period. As documented in the monthly analysis, the south-western zone records substantial soil loss across May (pre-monsoon), June (monsoon onset), July (peak monsoon, Classes 2–5 in monthly maps), August (second-highest month, Classes 4–5 in this zone), and September (Classes 2–3 from the still-elevated September R -factor). The summation of these contributions across five to six months of sustained erosive activity produces the highest annual totals

in the study area. This multi-month accumulation is driven by the south-western zone's unique rainfall character: it receives meaningful rainfall from both the pre-monsoon south-westerly moisture inflow (May–June) and the main Kremti monsoon (July–September), sustaining a longer effective erosion season than either the eastern escarpment (where February–April and October–December contributions dominate) or the central plateau (where July–August peak erosion is partially offset by September–October vegetation protection). Under the $P = 1$ assumption, these annual soil loss values represent conditions in the absence of any documented conservation measures, and it is likely that existing stone bunds or terraces in parts of the south-western zone reduce actual erosion below these modelled levels.

The eastern escarpment zone; primarily the sub-regions of Senafe, Tsorena, and Adi Keyih, with secondary contributions from eastern Segeneyiti; constitutes the second zone of elevated annual soil loss in Classes 3 through 5. The concentration of soil loss of 68 to more than 180 t ha⁻¹ yr⁻¹ in the eastern escarpment reflects an entirely different RUSLE factor combination from the south-western zone. The eastern escarpment's elevated annual totals are not primarily driven by peak monsoon R-factor values; indeed, the eastern sub-regions record the lowest July, August, and September R-factor values in the study area; but rather by the combination of the study area's highest LS values (the LS-factor analysis confirmed LS values approaching 65.41 in this zone) with a longer active erosion season that includes elevated monthly contributions in February, March, April, October, November, and December. In the eastern escarpment, higher-intensity erosion events from Indian Ocean moisture incursions in February–April and the post-monsoon period of October–December accumulate alongside moderate July–September contributions, producing annual totals that, while lower than the south-western zone at their peaks, are sustained by contributions distributed more evenly across more months of the year. The K-factor in the eastern sub-regions (0.024–0.030; the lowest values in the study area) partially counteracts the topographic amplification from the extreme LS values, explaining why the eastern escarpment does not generate even higher annual totals than the south-western zone despite having substantially greater slope gradient and flow accumulation.

4.3.2.2. Data Gaps and Their Spatial Implications

White patches indicating data gaps are visible in the northern Adi Teklezan sub-region of the annual simulated soil loss map. These gaps originate from cloud contamination in the July and November Sentinel-2 NDVI composites, which generate missing values in the corresponding monthly C-factor rasters and propagate through the monthly soil loss computation to produce no-data pixels in the annual sum. The spatial extent of these gaps limits the completeness of the annual simulated soil loss assessment in the Adi Teklezan sub-region. Based on the patterns visible in the unaffected portions of the Adi Teklezan map and the monthly R-factor analysis for this sub-region; which showed elevated March R-factor values from Azmera rains and moderate July–August values; it is likely that the Adi Teklezan sub-region would fall predominantly in Class 2 (8.5–67 t ha⁻¹ yr⁻¹) if the data gaps were resolved, with localised Class 3 values in areas of steep slope and high LS. Smaller data gaps are also visible in parts of the eastern escarpment zone, consistent with cloud contamination documented in the April monthly NDVI composite for this area.

4.3.2.3. The Annual Maps as an Integration of Temporal Dynamics

The annual simulated soil loss map is most meaningfully interpreted not as a static characterisation of 2021 erosion conditions but as the spatial expression of an integrated temporal process spanning all 12 calendar months (Figure 32). The dominance of Class 1 (0–8.4 t ha⁻¹ yr⁻¹) across 45–50% of the study area reflects the temporal compression of the erosion season into two to four peak monsoon months, with eight or nine months of low-to-zero R-factor conditions limiting cumulative annual soil loss even where C values are high across the broad plateau interior. However, even within this lowest class, annual soil loss rates of up to 8.4 t ha⁻¹ yr⁻¹ substantially exceed the natural soil formation rate of approximately 1 t ha⁻¹ yr⁻¹ documented for comparable highland environments (Hurni, 1985a). Thus, Class 1 conditions, while representing the least severe erosion category in the study area, still constitute a net

depletion of the soil resource across nearly half of the regional landscape. The Class 2 belt (8.5–67 t ha⁻¹ yr⁻¹) covering a further estimated 35–40% of the study area represents the dominant condition of moderate annual erosion generated primarily by the two to three peak monsoon months of July and August, moderated by the dense September vegetation protection in the central zones. The Classes 3–5 areas exceeding 68 t ha⁻¹ yr⁻¹, confined to an estimated 10–15% of the study area in the south-western and eastern escarpment zones, represent the sub-regions where multiple months of elevated monthly contributions accumulate to produce consistently high annual totals under the P = 1 baseline assumption.

Taken together, Classes 1 and 2 account for approximately 85–90% of the study area, confirming that the majority of the approximately 10,000 km² landscape experiences annual soil loss below 67 t ha⁻¹ yr⁻¹ under the simulated framework. The remaining 10–15% of the study area, concentrated in the south-western and eastern escarpment zones, represents the priority areas where annual soil loss above 68 t ha⁻¹ yr⁻¹ warrants targeted SWC intervention. These findings are broadly consistent with published annual soil loss ranges for comparable Ethiopian Highland environments, where mean annual soil loss rates of 17 to 110 t ha⁻¹ yr⁻¹ have been documented across catchments with similar rainfall, soil, and land use conditions (Gelagay & Minale, 2016; TEFERI2, 2009), providing additional confidence in the plausibility of the simulated results.

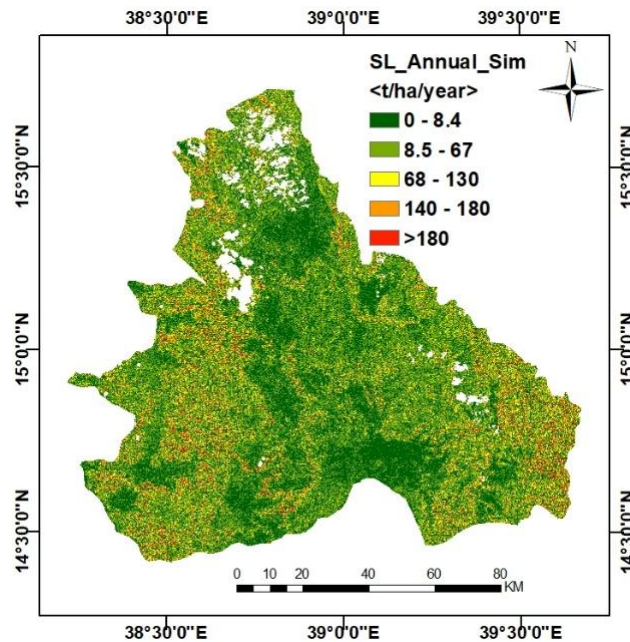


Figure 32. Annual simulated soil loss (t ha⁻¹ yr⁻¹) map for the year 2021

4.4. Conventional (Non-Simulated) Annual soil erosion

The non-simulated (conventional annual) soil loss map (A_{non-sim}) (Figure 33) represents the standard RUSLE formulation in which both the rainfall erosivity factor and the cover management factor are derived from annually aggregated data: R_{annual} is computed from the summed 2021 annual GPM IMERG precipitation across the study area, and C_{annual} is derived as the arithmetic mean of the 12 monthly Sentinel-2 NDVI-based C-factor rasters. The resulting soil loss, expressed in t ha⁻¹ yr⁻¹, is computed as $A_{non-sim} = R_{annual} \times K \times LS \times C_{annual} \times P$, with P = 1 throughout. The K and LS factors are identical to those used in the simulated approach. The non-simulated approach is presented here as a benchmark against which the simulated framework is evaluated, and its description follows the same area-based, sub-regional approach applied to the simulated analysis.

4.4.1. Overall Spatial Distribution

The non-simulated annual soil loss map is classified into five classes spanning 0 to more than 172 t ha⁻¹ yr⁻¹. The lowest class of 0 to 7.43 t ha⁻¹ yr⁻¹ covers an estimated 35–40% of the total study area, a

somewhat smaller proportion than the 45–50% covered by the equivalent class ($0\text{--}8.4\text{ t ha}^{-1}\text{ yr}^{-1}$) in the simulated map. The second class of $7.44\text{ to }62\text{ t ha}^{-1}\text{ yr}^{-1}$ is the most spatially extensive of the five classes, covering an estimated 40–45% of the study area; a notably larger proportion than its equivalent in the simulated map. Together, Classes 1 and 2 account for approximately 80–85% of the study area, which is slightly lower than the simulated framework's 85–90%. The higher-loss Classes 3 through 5, collectively spanning $62.1\text{ to more than }172\text{ t ha}^{-1}\text{ yr}^{-1}$, account for an estimated 15–20% of the total study area under the non-simulated framework, compared with approximately 10–15% under the simulated framework. These aggregate differences, while moderate in terms of total area proportions, are spatially concentrated in specific sub-regions and have important implications for the identification of priority erosion zones.

4.4.2. Spatial Pattern of Each Class

Class 1 ($0\text{--}7.43\text{ t ha}^{-1}\text{ yr}^{-1}$) is present in the central plateau sub-regions of Asmara City, Ghala Nefhi, Berikh, Mendefera, and Dekemhare, and in the western sub-regions of Dibarwa and inner Logo Anseba. However, its spatial extent in the central plateau is visibly reduced relative to the simulated map, where these sub-regions were dominated by the lowest class. In the non-simulated map, Class 1 and Class 2 appear in a more balanced mix across the central plateau, indicating that the non-simulated approach assigns higher soil loss to these sub-regions than the simulated framework does. The reason lies in the temporal averaging of the C-factor: the non-simulated approach applies the mean annual C (approximately 0.3–0.5 across the central agricultural sub-regions) to the annual R-factor, whereas the simulated approach pairs the high July R with the partially established July C (0.4–0.7) but also pairs the still-high September R with the very low September C (0.05–0.15). The September protection effect; which substantially reduces soil loss in the central plateau under the simulated framework; is absorbed into the annual C average in the non-simulated approach and therefore does not generate an equivalent suppression of annual soil loss in these sub-regions.

Class 2 ($7.44\text{--}62\text{ t ha}^{-1}\text{ yr}^{-1}$) is the most spatially extensive class and covers a broader geographic extent than its simulated equivalent. In the non-simulated map, this class is prevalent not only across the south-western zone (Areza, May Mine, Adi Kwala, Kudo Bu'er, Logo Anseba) and the eastern sub-regions (Segeneyiti, northern Adi Keyih) as in the simulated map, but also more extensively across the central plateau sub-regions of Asmara City, Ghala Nefhi, Berikh, and Mendefera. The geographical spread of Class 2 into the central plateau under the non-simulated framework reflects the overestimation of annual soil loss in these sub-regions that results from applying the annual mean C-factor, which at approximately 0.3–0.5 assigns intermediate vegetation cover conditions to months that in 2021 ranged from near-bare soil in January and February to complete canopy closure in September. The annual R-factor for the central plateau (R_{annual} approximately $290\text{--}340\text{ MJ mm ha}^{-1}\text{ h}^{-1}\text{ yr}^{-1}$) multiplied by this intermediate mean C produces Class 2 soil loss across much of the central plateau; a result that does not reflect the actual temporal dynamics of 2021, in which September's dense vegetation substantially suppressed erosion in the months following the peak monsoon.

Classes 3 through 5 ($62.1\text{ to more than }172\text{ t ha}^{-1}\text{ yr}^{-1}$) are concentrated in the same two geographic zones identified in the simulated analysis: the south-western sub-regions of Areza, May Mine, and Adi Kwala, and the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih. These zones retain elevated soil loss in the non-simulated framework for the same structural reasons as in the simulated framework; high annual R-factor in the south-western zone (R_{annual} $350\text{--}430$) combined with elevated LS and moderate annual C values, and high LS combined with multi-season erosive R contributions in the eastern escarpment. However, the spatial extent of Classes 3 through 5 appears moderately more extensive in the non-simulated map than in the simulated map, particularly in the south-central Debub sub-regions and along the drainage networks of the central plateau, where the temporal aggregation of R and C inflates annual soil loss estimates relative to the month-by-month simulation.

4.4.3. The Role of Annual R and Mean Annual C in shaping the Non-Simulated Distribution

The spatial pattern of the non-simulated annual soil loss map is determined by the product $R_{\text{annual}} \times C_{\text{annual}}$, since K, LS, and P are identical across both frameworks. The annual R-factor for 2021 is highest in the south-western sub-regions (R_{annual} 350–430 MJ mm ha⁻¹ h⁻¹ yr⁻¹ in Areza, May Mine, Adi Kwala, and Logo Anseba), where the combined pre-monsoon and main monsoon precipitation accumulates the largest annual total. The mean annual C-factor is moderately low (approximately 0.1–0.3) in these same south-western sub-regions, reflecting their vigorous monsoon-season vegetation response that lowers the annual mean despite high dry-season C values. The product of the highest annual R and the moderately low annual C in the south-western zone therefore generates the highest non-simulated annual soil loss in the study area, concentrated in Classes 3 through 5.

Conversely, in the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih, the annual R-factor is the lowest in the study area (R_{annual} 130–190), but the annual mean C is relatively higher (approximately 0.3–0.5, reflecting the lower September NDVI documented in the C-factor analysis for these sub-regions), and the LS-factor reaches the study-area maximum. The product of low R, moderately high C, and very high LS still produces Classes 3 through 5 soil loss in parts of the eastern escarpment under the non-simulated framework, but through a different factor combination than in the south-western zone. The non-simulated approach correctly identifies both the south-western and eastern escarpment zones as areas of elevated annual soil loss, because these zones' rankings are driven primarily by the spatially fixed K and LS factors that are common to both frameworks. However, the non-simulated approach cannot correctly represent the relative severity between these two zones and the central plateau, because the temporal averaging of R and C obscures the zone-specific seasonal dynamics that the monthly simulation captures.

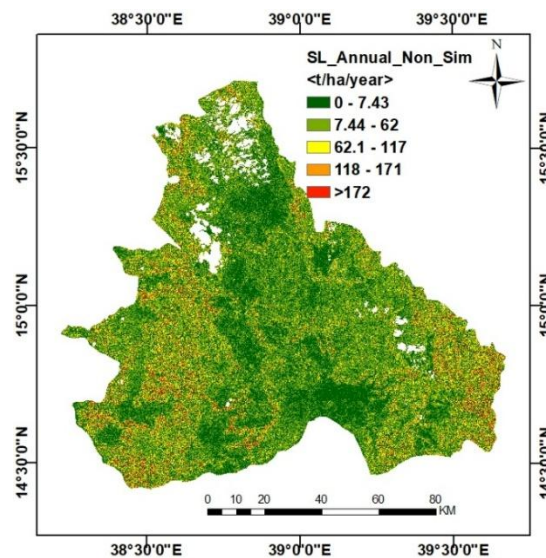


Figure 33. Conventional (Non-Simulated) Annual soil erosion ($t\ ha^{-1}\ yr^{-1}$) map

4.5. Comparison of the two approaches

The quantitative difference map ($\Delta A = A_{\text{non-sim}} - A_{\text{annual}}$, Figure 34) directly compares the two modelling frameworks by computing, for every pixel in the study area, the annual soil loss estimate produced by the conventional approach minus the estimate produced by the simulation. Positive values of ΔA indicate that the conventional approach generates higher soil loss than the simulation at that location; negative values indicate that the simulation generates higher soil loss than the conventional approach. The map is classified into ten classes spanning from -260 to $+160\ t\ ha^{-1}\ yr^{-1}$, providing a

detailed quantitative characterisation of the spatial pattern and magnitude of divergence between the two frameworks. The P-factor was set to unity ($P = 1$) in both approaches, and since K and LS are identical across both, all differences in ΔA are attributable solely to the different temporal treatment of the R-factor and C-factor.

4.5.1. Quantitative differences

The difference map reveals a strongly asymmetric and spatially structured distribution of ΔA values across the study area. The single most extensive class is -1.3 to $+12 \text{ t ha}^{-1} \text{ yr}^{-1}$, which covers an estimated 65–70% of the approximately 10,000 km² study area. Within this dominant class, values range from very small negative to small positive differences, indicating that the two approaches agree to within $12 \text{ t ha}^{-1} \text{ yr}^{-1}$ over the majority of the landscape. This level of agreement reflects the fact that the spatially fixed K and LS factors; which are identical in both frameworks; impose the dominant spatial structure on annual soil loss across most of the study area, and both approaches reproduce this structure with broadly comparable $R \times C$ products over the plateau interior.

Beyond the dominant near-zero class, the remaining 30–35% of the study area is split between positive and negative ΔA values. Positive differences; where the conventional approach overestimates relative to the simulation; collectively cover an estimated 15–20% of the study area, distributed across classes spanning 13 to $160 \text{ t ha}^{-1} \text{ yr}^{-1}$. Negative differences; where the simulation produces higher soil loss than the conventional approach; cover an estimated 10–15% of the study area, distributed across classes spanning; 1.4 to $-260 \text{ t ha}^{-1} \text{ yr}^{-1}$. The aggregate implication is that the conventional approach overestimates across a larger proportion of the study area than it underestimates, confirming a net upward bias in the conventional annual soil loss estimates relative to the spatiotemporally resolved simulation for 2021.

4.5.2. Spatial patterns of Over-estimation (Positive ΔA)

The positive ΔA classes, indicating locations where the conventional approach exceeds the simulated annual soil loss by 13 to $160 \text{ t ha}^{-1} \text{ yr}^{-1}$, are concentrated in two primary spatial configurations across an estimated 15–20% of the total study area.

The most geographically concentrated and quantitatively largest positive differences; classes spanning 33 to $160 \text{ t ha}^{-1} \text{ yr}^{-1}$; are located in the south-western sub-regions of May Mine, southern Areza, and Adi Kwala. These sub-regions account for an estimated 6–9% of the total study area combined and represent the zone of greatest absolute overestimation by the conventional framework. Within this south-western zone, the highest positive class of 96 to $160 \text{ t ha}^{-1} \text{ yr}^{-1}$ is spatially concentrated in the southernmost portions of May Mine and southern Adi Kwala, occupying an estimated 1–2% of the study area. The class of 56 to $95 \text{ t ha}^{-1} \text{ yr}^{-1}$ extends over a slightly broader area within southern Areza and May Mine, covering an estimated 2–3%, while the class of 33 to $55 \text{ t ha}^{-1} \text{ yr}^{-1}$ reaches into the northern margins of May Mine and parts of Kudo Bu'er, covering an estimated 3–4% of the total study area. Together, the three highest positive classes ($33\text{--}160 \text{ t ha}^{-1} \text{ yr}^{-1}$) cover an estimated 6–9% of the study area and are exclusively concentrated in the south-western sub-region cluster.

The spatial concentration of the largest positive ΔA values in the south-western zone is directly attributable to the highest annual R-factor in the study area (R_{annual} of $350\text{--}430 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$ in this zone) being applied in the conventional approach to the annual mean C-factor. The mean annual C for the south-western sub-regions is a moderate temporal average that blends the high C values of the dry season with the lower C values developed during the vigorous monsoon-season vegetation growth. When the highest R_{annual} in the study area is multiplied by this moderate mean C, the resulting conventional soil loss estimate substantially exceeds the simulated annual sum, which correctly captures the progressive decline in monthly C through August and September as vegetation develops, reducing the late-monsoon erosion contributions below what the annual mean C would imply. The further from

the monsoon peak that the vegetation develops, the more the conventional approach overestimates relative to the simulation.

The smaller positive ΔA classes of 13 to 32 t ha⁻¹ yr⁻¹ are more widely distributed across an estimated 9–11% of the study area, appearing as scattered patches rather than a geographically coherent zone. These patches are most evident across the central plateau sub-regions; including parts of Asmara City, Ghala Nefhi, and Berikh; along the drainage networks in Mendefera and Dekemhare, and along the western sub-region margins of Logo Anseba and Dibarwa. The class of 13 to 19 t ha⁻¹ yr⁻¹ covers an estimated 5–6% of the study area across the central plateau, while the class of 20 to 32 t ha⁻¹ yr⁻¹ covers approximately 4–5%, concentrated along steep valley walls within the drainage network. These moderate positive differences in the central plateau reflect the same mechanism operating at lower intensity: the annual mean C incompletely captures the vegetation protection cycles in the central agricultural sub-regions, producing systematic but moderate overestimation across a geographically dispersed area.

4.5.3. Spatial patterns of Under-estimation (Negative ΔA)

The negative ΔA classes, where the simulation produces higher annual soil loss than the conventional approach by 1.4 to 260 t ha⁻¹ yr⁻¹, are spatially distributed across an estimated 10–15% of the total study area. Unlike the positive ΔA classes, which are concentrated in a coherent south-western zone, the negative classes are characterised by a more dispersed spatial pattern, appearing across the northern zone, parts of the eastern sub-regions, and scattered locations across the study area.

The most extreme negative class of –260 to –130 t ha⁻¹ yr⁻¹ is the least spatially extensive, covering an estimated 1–2% of the study area as highly localised patches most likely confined to the steepest escarpment slope positions within the Senafe, Tsorena, and Adi Keyih sub-regions. At these locations, the extreme LS values combine with the monthly peak R-factor concentration to generate very high simulated monthly soil loss contributions in July and August that the conventional annual approach, using the lower annual R and a moderate mean annual C, cannot reproduce. The class of –120 to –64 t ha⁻¹ yr⁻¹ extends this pattern across an estimated 2–3% of the study area, concentrated in the northern margins of the Adi Teklezan sub-region and in parts of the Serejeka sub-region, where the Azmera/minor rainy season delivers high March erosivity that the monthly simulation captures through its high March R_{sim} value but that the conventional annual R-factor, averaged across all 12 months, substantially dilutes.

The slightly negative classes of –63 to –1.4 t ha⁻¹ yr⁻¹ cover a combined estimated 8–12% of the study area, distributed more broadly across the northern zone, scattered portions of the eastern sub-regions of Segeneyiti and the Adi Keyih margins, and dispersed pixels throughout the central drainage networks. These areas share the characteristic that the monthly simulation assigns a higher effective erosivity to specific peak months; most commonly July; than the conventional approach distributes across the full year, and this concentrated peak-month R_{sim} , paired with the corresponding month-specific C, generates a monthly soil loss contribution that exceeds the conventional annual estimate at those pixels. The spatial dispersion of the negative difference classes, rather than their concentration in a coherent sub-regional zone, confirms that this underestimation by the conventional approach is driven by the temporal concentration of monthly R-factor values; a pixel-scale effect that varies with the spatial pattern of 2021 monthly precipitation at the GPM IMERG resolution; rather than by a systematic sub-regional rainfall or vegetation characteristic.

An important spatial feature of the negative ΔA pattern is its concentration in the northern zone of Adi Teklezan and Serejeka. The monthly simulation captures the March R-factor peak in this zone; where the Azmera minor rainy season delivers concentrated rainfall producing a high R_{sim_March} that, when combined with the high dry-season C prevailing in March (C values of 0.5–0.7 as vegetation has not yet responded), generates a substantial March soil loss contribution. The conventional approach

computes R_{annual} for these sub-regions from the full annual precipitation total, which is dominated by the main monsoon months and therefore represents a lower effective annual erosivity than the concentrated March minor-season R would imply under the monthly simulation. The result is that the simulation produces higher annual totals in the northern zone than the conventional approach, generating the negative ΔA pattern observed in this part of the study area.

4.5.4. Comparison of Conventional vs. Simulated Soil Loss

The subregion comparison (Figure 34) shows a clear spatial asymmetry similar to the difference map (Figure 35): the simulated method estimates higher average annual soil loss in most subregions, except in a cluster of southwestern areas (May Mine, Areza, Kudo Bu'er, Adi Kwala, and Senafe), where the conventional method estimates higher values. In the northern and central highlands, Adi Teklezan, Serejeka, Berikh, and Logo Anseba, which are mostly negative-difference zones in Figure 34, the simulated approach highlights greater erosion risk by accounting for phenological lag: peak erosivity in July aligns with incomplete vegetation cover, and the monthly C-factor from Sentinel-2 NDVI assigns a high value during this period. Conversely, the static C-factor in the conventional approach spreads vegetation cover evenly throughout the year, diluting the July erosion peak. In the southwestern zones, corresponding to positive-difference areas in Figure 34, the pattern reverses: early-season rapid vegetation growth lowers the monthly C-factors in the simulated data below the static values used traditionally. These static values tend to overestimate bare soil exposure in these more vegetated transitional zones. Overall, the disaggregation shows that the two methods differ not only in magnitude but also in spatial erosion priorities: the simulated framework emphasises higher erosion risk in the degraded northern highlands and moderates it in the greener southwest. This highlights the importance of temporal resolution in the C-factor as a significant methodological improvement, not just a minor refinement.

The difference map confirms that the divergence between the conventional and simulated approaches is neither uniformly directed nor uniformly distributed. The conventional approach overestimates in the south-western zone; the highest annual R zone of the study area; by 33 to 160 t ha⁻¹ yr⁻¹ across an estimated 6–9% of the study area, and moderately overestimates across a further 9–11% of the central plateau and western sub-regions. Simultaneously, the simulation produces higher estimates in the northern and escarpment zones across an estimated 10–15% of the study area by 1.4 to 260 t ha⁻¹ yr⁻¹, driven by the monthly concentration of R -factor values that the conventional annual R -factor dilutes. Over approximately 65–70% of the study area, the central plateau interior, western lowland margins, and eastern plateau surfaces, the two approaches agree to within 12 t ha⁻¹ yr⁻¹. The bidirectional character of the bias and its clear sub-regional spatial structure confirm that the divergence between the two approaches is a systematic consequence of the temporal treatment of R and C rather than a random modelling uncertainty, and that the magnitude and direction of the difference are predictable from the seasonal rainfall and vegetation dynamics of each sub-region. These findings have direct implications for sub-regional prioritisation of SWC interventions, as reliance on the conventional map alone would both overestimate the erosion severity of the south-western zone and underestimate that of the northern and escarpment sub-regions relative to the simulation.

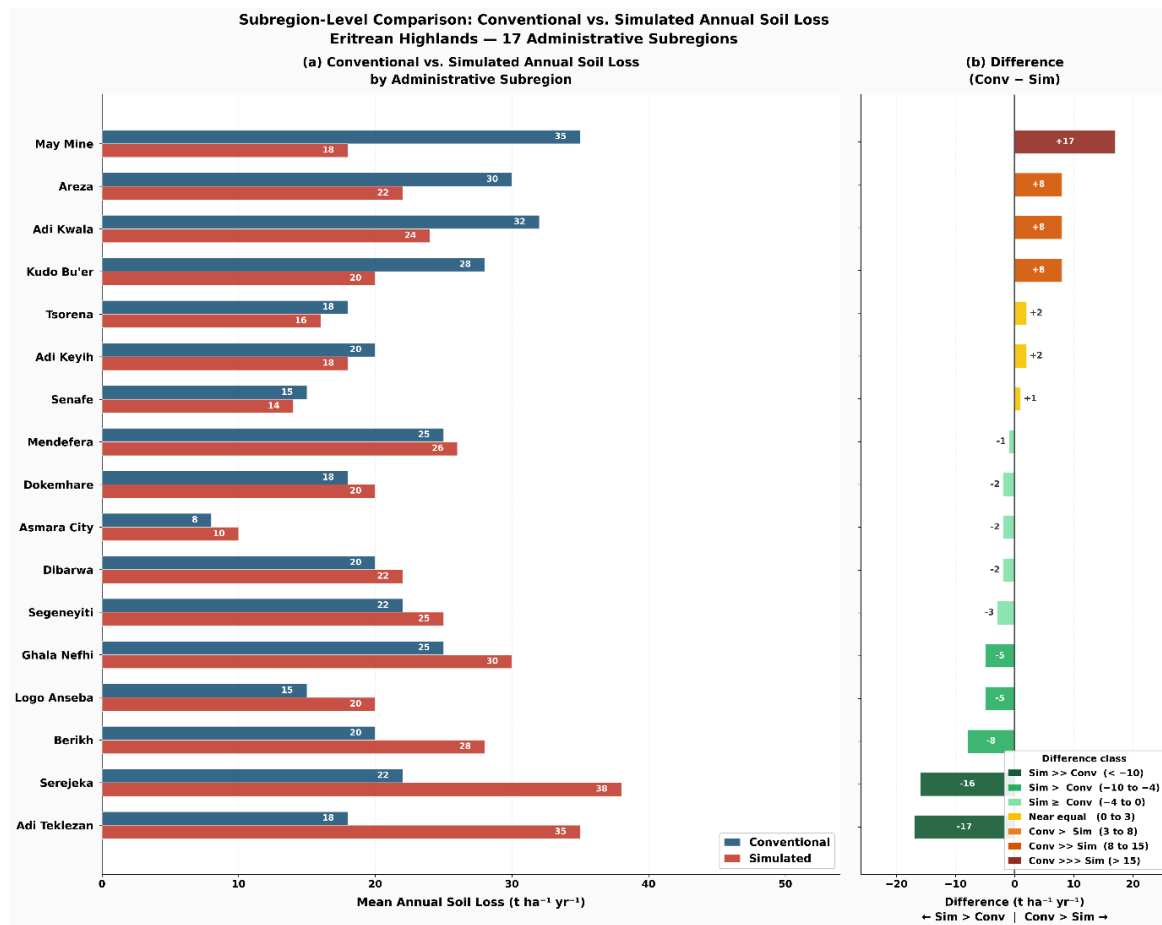


Figure 34. Subregion-level comparison of mean annual soil loss ($t\ ha^{-1}\ yr^{-1}$) between the conventional and simulated RUSLE approaches across the 17 administrative subregions of the Eritrean Highlands, sorted by Conv - Sim difference.

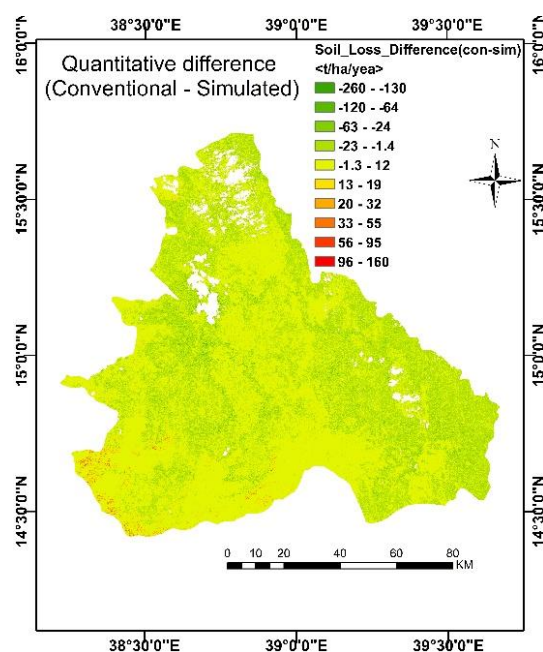


Figure 35. Difference between conventional and simulated soil loss ($t\ ha^{-1}\ yr^{-1}$) map

4.6. General Interpretation of the Results and the Applied Methodology

4.6.1. Significance of Temporal Variability in Erosion Modelling

The monthly simulation framework reveals that soil erosion in the Eritrean central and southern highlands is overwhelmingly concentrated in a narrow window of the annual cycle. July alone records a Class 1 background erosion level of 0 to 13 t ha⁻¹ month⁻¹; a floor that is approximately 40 times higher than January's background level of 0 to 0.3238 t ha⁻¹ month⁻¹; and the four main monsoon months of June through September collectively account for the dominant share of annual soil loss. This temporal concentration is not merely a statistical feature of the dataset but a fundamental characteristic of the Eritrean highland climate: eight months of near-zero rainfall erosivity separated by four months of intense monsoon delivery. Any modelling approach that treats the annual cycle as temporally uniform cannot represent this structure, and its failure to do so propagates directly into the spatial pattern of estimated erosion, the identification of priority sub-regions, and the design of conservation interventions.

The monthly analysis further reveals that the spatial location of the highest monthly soil loss shifts systematically through the year, migrating from an isolated north-central patch (Serejeka/Berikh) in January, through the eastern escarpment (Senafe, Tsorena, Adi Keyih) in February and April, a dual north-western and eastern hotspot in March (Adi Teklezan/Logo Anseba and eastern escarpment), a south-western concentration in May through June (Logo Anseba, Areza, May Mine, Adi Kwala), a study-area-wide peak with south-western and central-western dominance in July–August, and an eastern escarpment residual from October through December. This seasonal migration of the erosion epicentre cannot be detected by a single annual soil loss map, and it carries practical consequences: the sub-regions requiring the most urgent SWC attention differ by month, meaning that conservation resources deployed at the wrong time or in the wrong zone may miss the critical erosion events entirely.

(2017)(Jakab et al., 2017) further demonstrated, using field rainfall simulation, that soil erodibility peaks when the seedbed is freshly prepared, with soil loss rates varying markedly throughout the growing season depending on tillage practices and surface cover. This highlights that the timing of erosive rainfall within the crop growth cycle critically influences potential soil loss, a dynamic captured by the present study's month-specific C-factor rasters within a monthly simulation framework.

4.6.2. Why the Annual Approaches Biases Estimates

The quantitative difference analysis established that the conventional approach overestimates annual soil loss by 33 to 160 t ha⁻¹ yr⁻¹ across an estimated 6 to 9% of the study area (south-western sub-regions of May Mine, Areza, and Adi Kwala), and underestimates by 1.4 to 260 t ha⁻¹ yr⁻¹ across approximately 10 to 15% of the study area (northern zone of Adi Teklezan and Serejeka, and parts of the eastern escarpment). Both biases originate from the same structural limitation: applying a single annual R-factor and a single annual mean C-factor to produce a year-round uniform soil loss estimate, when in reality the monthly R and C factors co-vary in opposing directions through the annual cycle. The annual mean C-factor minimum of 0.0094; the value applied to the most persistently vegetated pixels in every month under the conventional approach; underestimates the actual July C-factor minimum by a factor of approximately three (July C_{min} = 0.0292), meaning the conventional model assigns substantially more vegetation protection during the most erosive month than actually existed in 2021. Simultaneously, the conventional annual R-factor of 350 to 430 MJ mm ha⁻¹ h⁻¹ yr⁻¹ in the south-western zone, when multiplied by the annual mean C, produces a higher soil loss estimate than the simulation's month-specific R × C products because the simulation correctly captures the progressive vegetation development that suppresses late-monsoon erosion in August and September.

The directional bias of these errors is not symmetric. In the south-western zone, the conventional approach systematically overestimates because the highest annual R in the study area is paired with a moderate mean annual C that inadequately represents the growing-season vegetation trajectory. In the

northern zone and parts of the eastern escarpment, the conventional approach underestimates because the monthly simulation correctly identifies high peak-month erosivity; particularly the Azmera/ minor rainy season's March R-factor concentration in Adi Teklezan and Logo Anseba; that the annual R-factor dilutes across all 12 months. These findings are consistent with the findings of (Panagos, Ballabio, et al., 2015; Vrieling et al., 2010), who documented that annual RUSLE parameterisation systematically over- or underestimates sub-regional erosion in environments with strong seasonal rainfall and vegetation cycles, and that the direction of the bias depends on whether the annual R or annual C is the dominant source of temporal error in a given zone.

4.6.3.Role of Machine Learning in Overcoming Data Scarcity

With 302 soil sampling points distributed across approximately 10,000 km² ; an average density of one sample per 33 km²; the study area does not provide sufficient spatial coverage for reliable conventional interpolation of soil properties across the heterogeneous terrain of the Eritrean highlands. The ML regression framework addresses this limitation by exploiting the systematic relationships between soil properties and their environmental controlling factors; encoded in DEM-derived terrain attributes and Sentinel-2 spectral covariates; to generate spatially continuous predictions that honour the underlying pedogenetic gradients of the landscape. The spatial coherence across all four predicted soil property maps; with clay and silt peaking in the central plateau (Asmara City, Ghala Nefhi, Berikh), sand peaking in the eastern escarpment (Senafe, Tsorena, Adi Keyih), and OM increasing with elevation toward the eastern highlands; confirms that the models captured genuine soil-landscape relationships rather than interpolation artefacts. This multi-property coherence, which would not be achievable through conventional interpolation with 302 points, provides the physical basis for the K-factor spatial pattern (0.024 to 0.048 t ha h ha⁻¹ MJ⁻¹ mm⁻¹) that correctly places the highest erodibility in the intensively cultivated central plateau sub-regions and the lowest in the sandy, organic-matter-richer eastern escarpment soils.

The practical consequence of ML-based K-factor prediction is the correct representation of the K-LS spatial trade-off: the eastern escarpment sub-regions that carry the study area's highest LS values (approaching 65.41) also carry the lowest K values (0.024 to 0.030), because their coarse, organic-matter-richer sandy soils developed under different parent material and land use conditions than the high-silt, low-OM central plateau. A spatially uniform or poorly interpolated K-factor would have suppressed this trade-off entirely, inflating the estimated annual soil loss of the eastern escarpment; already well-documented for its high topographic erosion potential; to levels not justified by its actual soil erodibility. The ML approach therefore contributes not only to the spatial completeness of the K-factor characterisation but also to the physical realism of the soil loss patterns in both the simulated and conventional frameworks. This finding demonstrates the value of integrating ML with RS and terrain analysis for erosion modelling in data-scarce highland environments, and extends the methodological contributions documented in analogous studies across the Ethiopian Highlands (Addis & Klik, 2015) to the Eritrean context.

The sensitivity of modelled soil loss to the K-factor estimation method is well established: (Keller et al., 2021b) compared USLE, RUSLE, USLE-M and EPIC-based K-factor equations on loess slopes in Hungary and found that erodibility values ranged from 0.0028 to 0.0087 t ha h ha⁻¹ MJ⁻¹ mm⁻¹ across methods, with only USLE-M yielding an acceptable correlation between measured and modelled soil loss. This underscores the importance of selecting an appropriate K-factor derivation procedure, a consideration directly addressed in this study by applying the EPIC/Williams formulation to ML-predicted soil

4.6.4.Spatial Heterogeneity and Erosion-Driver Interactions

The multi-factor analysis reveals three spatially distinct erosion-driver configurations that determine the pattern of annual soil loss across the study area. In the central plateau sub-regions of Asmara City,

Ghala Nefhi, Berikh, and Mendefera, compound vulnerability emerges from the convergence of the highest K-factor values (0.040 to 0.048, reflecting high silt, low OM, weakly aggregated topsoils), moderately high C-factor values during June and July (0.4 to 0.8 on freshly tilled, incompletely sown agricultural surfaces), and the highest July R_{sim} in the study area (1,800 to 2,200 MJ mm ha⁻¹ h⁻¹ yr⁻¹). The interaction of these three conditions in the central agricultural heartland during the June-July critical window represents the highest soil loss risk per unit area during the peak monsoon, despite the moderate LS values of the plateau surface. In the south-western sub-regions of May Mine, Areza, and Adi Kwala, the highest annual accumulation results not from compound monthly vulnerability but from a sustained multi-month erosion season: meaningful soil loss contributions from May through September collectively produce the highest annual totals, driven by the highest annual R-factor (350 to 430) in the study area. In the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih, the highest LS values in the study area (approaching 65.41) are partially counteracted by the lowest K values (0.024 to 0.030) and the lowest monsoon R-factor values, such that despite their topographic erosion amplification these sub-regions produce moderate rather than maximum annual soil loss.

These three factor configurations produce a spatial pattern of annual erosion that differs substantially from what would be predicted from any single RUSLE factor alone: the highest LS zone does not produce the highest annual soil loss; the highest K zone does not produce the highest monthly peak; and the highest annual R zone generates the highest annual accumulation precisely because its C-factor trajectory allows sustained erosion across multiple months rather than a concentrated July peak. The interaction among all four spatially and temporally varying RUSLE factors (R, K, LS, and C) can be correctly represented only by combining ML-based K-factor spatial prediction with the monthly temporal resolution of R and C implemented in this study.

4.6.5. Comparison with Previous Studies

The non-simulated annual R-factor range of 130 to 430 MJ mm ha⁻¹ h⁻¹ yr⁻¹ across the Eritrean study area is consistent with published ranges for comparable East African highland environments. (Hurni, 1988) documented annual R-factor values of 200 to 600 for the Ethiopian Highlands under mean annual rainfall of 700 to 1,400 mm yr⁻¹, while (TEFERI2, 2009) reported R values of 373 to 746 in the wetter Chemoga watershed of the Blue Nile basin. The Eritrean study area's lower range reflects its comparatively modest mean annual precipitation relative to the Ethiopian Blue Nile catchments and is consistent with the known climatic character of the Eritrean central highlands. The annual soil loss distribution, with approximately 85 to 90% of the study area below 67 t ha⁻¹ yr⁻¹ under the simulated framework, is broadly consistent with published mean annual erosion rates for the Ethiopian Highlands, where (TEFERI2, 2009) reported a mean of approximately 93 t ha⁻¹ yr⁻¹ in the Chemoga watershed and (Gelagay & Minale, 2016) reported a mean of approximately 38.4 t ha⁻¹ yr⁻¹ in the Koga watershed under different rainfall and slope conditions. (Ayehu et al., 2018) documented soil loss rates of 1.4 to 134 t ha⁻¹ yr⁻¹ in the Yezat watershed, northern Ethiopia, which closely brackets the Eritrean study area's range across sub-regions. These comparisons confirm that the simulated annual soil loss estimates for the Eritrean central and southern highlands are physically plausible within the established regional range.

What distinguishes this study from most published Ethiopian Highlands RUSLE applications is the combination of monthly temporal resolution and ML-based K-factor spatial prediction. Most published studies in this region use a single annual R-factor derived from station data (Gelagay & Minale, 2016; TEFERI2, 2009), a spatially uniform or coarsely interpolated K-factor (Ayehu et al., 2018; Hurni, 1988), and a single annual mean C-factor from NDVI or land use classification. The present study's application of 12 monthly R_{sim} and C_m rasters from satellite-derived data, combined with a spatially continuous ML-predicted K-factor, represents a methodological advancement that enables direct quantification of the temporal bias introduced by conventional annual parameterisation; a comparison that, to the best of

the author's knowledge, has not previously been published for any Eritrean highland watershed and has rarely been performed for East African highland environments at the regional scale.

4.6.6. Strengths, Assumptions, and Limitations

The principal strengths of this study are the temporal resolution of the RUSLE simulation, the ML-based K-factor characterisation, and the direct quantitative comparison of the two modelling frameworks. These three elements collectively address limitations that have constrained erosion modelling in the Eritrean and broader East African highland context: the inadequacy of annual RUSLE parameterisation for strongly seasonal environments, the spatial data scarcity for soil erodibility mapping, and the absence of a published inter-approach comparison for this region. The use of GPM IMERG Final Run Version 07 provides spatially continuous monthly precipitation estimates that overcome the sparse and unevenly distributed ground station network, and the Sentinel-2 monthly NDVI composites provide fine-resolution vegetation dynamics at scales not previously available for Eritrean soil erosion research.

The study carries several assumptions and limitations that must be acknowledged. First, the P-factor was set to unity ($P = 1$) throughout, meaning results represent potential soil loss under prevailing land cover conditions in the absence of documented conservation practices. Where stone bunds, terraces, or check dams exist in the study area; which is likely in parts of the central agricultural sub-regions and the Debub region; potential soil loss may be substantially lower than modelled values indicate. Second, the annualisation procedure ($P_m \times 12$) required by the empirical erosivity equation is mathematically justified but assumes that each month's precipitation rate, if sustained year-round, would generate the corresponding annual erosivity. This procedure may overestimate the erosive intensity of months with single intense events concentrated in a few days, since the annualised R assigns the full annual-scale erosivity index to conditions that persist for only part of the month. Third, the analysis covers a single year (2021) and does not capture inter-annual variability in rainfall, vegetation, or land use; a wetter or drier year would produce correspondingly higher or lower erosion estimates. Fourth, GPM IMERG's approximately 10 km spatial resolution does not resolve sub-pixel rainfall variability driven by local orographic effects, which may be highly important in the dissected terrain of the eastern escarpment and the northern highlands. Fifth, cloud contamination in the July and November Sentinel-2 composites generates data gaps in the Adi Teklezan sub-region and parts of the eastern zone, limiting the completeness of the C-factor and soil loss assessment in those areas.

RUSLE estimates gross on-site soil detachment from sheet and rill erosion processes. For a study area of approximately 10,000 km², the sediment delivery ratio (SDR) is typically in the range of 0.05–0.20 (Walling, 1983), implying that only 5–20% of the predicted gross erosion may ultimately be delivered to stream channels and basin outlets. The present study does not model sediment transport, routing, or deposition. Results should therefore be interpreted as erosion vulnerability and detachment potential maps, not as estimates of basin sediment yield.

The current analysis did not use a land-use mask to exclude non-erodible surfaces like water bodies, urban areas, and rocky outcrops from soil property predictions and RUSLE calculations. As a result, these non-erodible features were included in the overall spatial domain, so the class percentages and erosion rates represent the entire study area rather than just the erodible land. This may slightly overestimate erosion-class extents. Incorporating a land-cover exclusion mask to focus all calculations and spatial data on erodible surfaces is identified as an important methodological improvement for future work.

The NDVI-based C-factor formulation (Van Der Knijff et al., 2000) captures only the protective effect of photosynthetically active biomass. Non-photosynthetic vegetation, post-harvest crop residues, straw mulch, and dormant rangeland vegetation provide genuine mechanical soil protection from raindrop

impact and surface flow, but register near-zero or negative NDVI values. As a result, the C-factor derived from NDVI may overestimate erosion susceptibility during post-harvest and dry-season periods when non-photosynthetic cover is present. This limitation is inherent to NDVI-based C-factor estimation. Future work should explore non-photosynthetic vegetation indices to complement NDVI for C-factor derivation in the Eritrean Highland context.

All input datasets: GPM IMERG monthly precipitation, MOA/NARI soil profiles, and Sentinel-2 cloud-free NDVI composites are specific to the 2021 analysis year, ensuring internal consistency across the simulation framework. It is acknowledged that 2021 cannot be asserted as representative of the long-term rainfall climatology, given inter-annual variability. The study establishes the first spatially explicit, monthly-resolved soil erosion baseline for the Eritrean Highlands, which can be directly extended to future multi-year analyses and climate scenario projections using the full GPM IMERG record. Multi-year analysis is identified as a priority for future work.

4.6.7. Implications for Soil Conservation and Land-Management Prioritisation

The results of this study have direct and actionable implications for SWC planning in the Eritrean central and southern highlands. The identification of June and July as the critical erosion window; defined by the coincidence of rising R-factor with high C-factor values of 0.4 to 0.8 across freshly tilled agricultural surfaces in Asmara City, Ghala Nefhi, Berikh, Mendefera, and Dekemhare; implies that the highest-priority conservation interventions in the central agricultural plateau must be implemented before the Kremti monsoon onset rather than during or after it. Agronomic measures that reduce the time between tillage and crop canopy establishment; such as earlier planting schedules linked to the first monsoon rains, minimum tillage to minimise bare soil exposure, and mulching of prepared surfaces before sowing; would directly target the critical window identified in the monthly C-factor and soil loss analyses. Physical SWC structures (stone bunds, fanya juu terraces) in these sub-regions would reduce the LS-factor contribution to erosion, but their conservation benefit is greatest when combined with agronomic measures that minimise the June-July C-factor.

In the south-western sub-regions of May Mine, Areza, and Adi Kwala; where the highest annual soil loss results from a sustained multi-month erosion season from May through September; physical SWC structures that reduce runoff and sediment transport across the full monsoon season are the primary conservation need. The high annual R-factor (350 to 430) in these sub-regions combined with their moderate-to-high C-factor values through July indicate that vegetation-based measures alone would be insufficient, and that runoff interception structures designed for the extended south-western erosion season are warranted. The eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih, despite their high LS values, show lower annual soil loss than the south-western zone and lower monsoon R-factor values; they nevertheless require attention during February, April, November, and December when Indian Ocean-linked rainfall produces elevated monthly soil loss in the absence of protective vegetation. The monthly simulation framework, by resolving the actual temporal and spatial distribution of erosion across 2021, provides the sub-region-specific, month-specific erosion data that land managers require to prioritise limited SWC investment, sequence conservation interventions through the agricultural calendar, and design structures matched to the duration and intensity of the erosion season in each sub-region.

5. CONCLUSIONS AND RECOMMENDATIONS

This chapter synthesises the study's principal findings, identifies its original contributions to knowledge, translates the results into practical SWC recommendations, and proposes a programme of future research to extend and consolidate the work. All quantitative statements draw on confirmed results from section 4.

5.1. Summary of Key Findings Mapped to Study objectives

The study pursued three principal objectives. The following findings are reported against each.

Objective 1: To develop a ML framework for spatial prediction of soil properties and derivation of the K-factor across the approximately 10,000 km² study area.

The ML regression framework successfully generated spatially continuous predictions of sand (15.01–73.23%), silt (16.09–52.36%), clay (0.28–50.07%), and OM (0.60–3.19%) across the study area, producing four maps with internally consistent spatial patterns that reflect the underlying pedogenetic gradients driven by parent material, topography, and land use.

The derived K-factor ranged from 0.024 to 0.048 t ha h ha⁻¹ MJ⁻¹ mm⁻¹. The highest values (0.040–0.048) were concentrated in the intensively cultivated central plateau sub-regions of Asmara City, Ghala Nefhi, Berikh, and Dekemhare (high silt, low OM), while the lowest values (0.024–0.030) occurred in the eastern escarpment sub-regions of Senafe, Tsorena, and Adi Keyih (high sand, higher OM). This spatial K distribution correctly represents the K–LS trade-off: the zones of highest topographic erosion potential carry the lowest soil erodibility.

Objective 2: To estimate monthly and annual soil loss using a spatiotemporal RUSLE simulation framework based on monthly GPM IMERG rainfall erosivity and monthly Sentinel-2 C-factor rasters.

The monthly soil loss ranking for 2021 was: July > August > September > June > April > May > March > November > December > February > October > January. July's background erosion level (Class 1: 0–13 t ha⁻¹ month⁻¹) was approximately 40 times higher than January's (Class 1: 0–0.32 t ha⁻¹ month⁻¹), confirming extreme monsoon concentration.

July was the most erosive month because it simultaneously recorded the highest R_{sim} (490–2,200 MJ mm ha⁻¹ h⁻¹ yr⁻¹ in the central-western zone) and the highest C-factor minimum of any month (C_{min} = 0.0292), confirming that peak monsoon erosivity coincided with the least achievable vegetation protection of the growing season.

The peak erosion zone migrated spatially through the year: from the eastern escarpment (February, April) through a dual north-western and eastern hotspot (March), south-western sub-regions (May–June), study-area-wide south-western and central-western dominance (July–August), and eastern escarpment residual (October–December). No single annual map can represent this seasonal migration.

The annual simulated soil loss (A_{annual}) placed approximately 45–50% of the study area in the lowest class (0–8.4 t ha⁻¹ yr⁻¹), approximately 35–40% in the second class (8.5–67), and approximately 10–15% in higher classes (>68 t ha⁻¹ yr⁻¹), concentrated in the south-western zone (May Mine, Areza, Adi Kwala) and eastern escarpment. All estimates reflect $P = 1$.

Objective 3: To estimate annual soil loss using the conventional non-simulated approach and to compare the two frameworks quantitatively and spatially.

The conventional approach placed approximately 35–40% of the study area in the lowest class (0–7.43 t ha⁻¹ yr⁻¹) and approximately 40–45% in the second class, spreading approximately 10 percentage points more area into Class 2 than the simulation, primarily in the central plateau sub-regions where the annual mean C-factor fails to represent the September vegetation protection effect.

The quantitative difference map ($\Delta A = A_{\text{non-sim}} - A_{\text{annual}}$) showed that approximately 65–70% of the study area had differences within -1.3 to $+12 \text{ t ha}^{-1} \text{ yr}^{-1}$ (near-zero). The conventional approach overestimated by 33 to $160 \text{ t ha}^{-1} \text{ yr}^{-1}$ in the south-western sub-regions of May Mine, Areza, and Adi Kwala (estimated 6–9% of study area), and the simulation produced higher estimates than the conventional approach in approximately 10–15% of the study area, concentrated in the northern zone (Adi Teklezan, Serejeka) and parts of the eastern escarpment.

The overestimation by the conventional approach in the south-western zone was driven by applying the highest annual R-factor ($350\text{--}430 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$) to the annual mean C-factor, which at moderate values does not represent the progressive monsoon-season vegetation protection that the monthly simulation correctly captures through August and September.

5.2. Practical and Policy Recommendations

The following recommendations are directed at land management authorities, conservation practitioners, and the Government of Eritrea's SWC programme. All soil loss estimates assume $P = 1$ (no documented conservation practices); potential soil loss in areas where SWC structures exist may be lower.

Target the June–July critical window in the central agricultural plateau. The sub-regions of Asmara City, Ghala Nefhi, Berikh, Mendefera, and Dekemhare sustain C-factor values of 0.4 to 0.8 on freshly tilled, incompletely established crop surfaces during June and July, coinciding with the annual peak in rainfall erosivity. Conservation investment in these sub-regions should prioritise measures that reduce the time between tillage and canopy closure: earlier planting schedules tied to the first reliable monsoon rains, minimum tillage to reduce bare surface exposure, and application of crop residue or organic mulch to planted surfaces. These agronomic interventions are low-cost and can be implemented without major infrastructure investment.

Prioritise physical SWC structures in the south-western zone for the extended erosion season. The sub-regions of May Mine, Areza, and Adi Kwala sustain the highest annual soil loss (Classes 3 to 5: $>68 \text{ t ha}^{-1} \text{ yr}^{-1}$ across approximately 6 to 9% of the study area) driven by a sustained erosion season from May through September under $P = 1$. Physical runoff interception structures; stone bunds, fanya juu terraces, and check dams in the drainage network; are the primary SWC need in this zone and should be designed for the full five-month erosion season rather than for the July peak alone.

Avoid over-allocating resources to the south-western zone on the basis of conventional maps. The conventional approach overestimates annual soil loss in the south-western sub-regions by 33 to $160 \text{ t ha}^{-1} \text{ yr}^{-1}$ relative to the simulation. Conservation planning based on the conventional map alone would divert disproportionate resources to this zone while underestimating the erosion of the northern sub-regions (Adi Teklezan, Serejeka) where the simulation produces higher estimates than the conventional approach. The monthly simulation results should be used as the planning basis for sub-regional prioritisation.

Initiate a P-factor field survey. The $P = 1$ assumption prevents this study from quantifying the erosion reduction achieved by existing SWC investments in the study area. A systematic field survey of existing conservation structures; stone bunds, terraces, vegetated waterways, and check dams; combined with the assignment of literature-based P values by practice type and slope class would enable a refined RUSLE computation that correctly credits existing conservation effort and more precisely identifies the sub-regions where additional investment is most needed.

Integrate the monthly simulation framework into national SWC planning. The seasonal migration of erosion hotspots across sub-regions, and the identification of the June–July critical window for the central agricultural plateau, provide actionable, month-specific information that national SWC

programmes can use to schedule conservation work, prioritise sub-watershed interventions, and evaluate the temporal effectiveness of conservation measures.

5.3. Recommendations for Future Research

Multi-year analysis. The current study covers 2021 only and cannot capture inter-annual variability in rainfall, vegetation response, or land use change. A multi-year simulation spanning at least 2015 to 2025 using the consistent GPM IMERG Version 07 record and Sentinel-2 time series would quantify the year-to-year variability of monthly and annual soil loss, identify anomalously erosive years, and support trend detection against the background of documented land use and rainfall changes in the Eritrean highlands.

P-factor refinement and incorporation. Future work should develop a spatially explicit P-factor surface for the study area through the combination of: RS-based detection of stone bunds and terraces using high-resolution satellite imagery, field validation campaigns in representative sub-regions, and assignment of P values based on published relationships between practice type, slope, and conservation effectiveness. The resulting P-factor map would reduce the $P = 1$ overestimation bias and enable a more realistic assessment of residual erosion risk after accounting for existing conservation investment.

Event-scale and sub-daily modelling. The GPM IMERG Final Run product is available at 30-minute temporal resolution, enabling the derivation of event-specific erosivity indices rather than the monthly annualised R-factor used in the current study. Future work should assess whether event-scale R-factor computation; which directly represents the kinetic energy of individual storm events that drive the actual detachment and transport processes; produces substantially different erosion estimates from the monthly annualisation approach, particularly for the high-intensity convective events of the July-August monsoon season.

Plot-scale validation. The absence of erosion pin or erosion plot data from the Eritrean highlands is a fundamental constraint on validating the absolute soil loss values produced by this study. Future research should establish a network of bounded micro-plots and sediment traps across a representative sample of sub-regions and land use types to provide direct measurements of event-scale and seasonal soil loss for comparison with the RUSLE-modelled estimates. This would also provide the empirical basis for calibrating the erosivity equation coefficients to Eritrean conditions rather than relying on equations derived from Ethiopian or broader East African calibration datasets.

Transferability to other highland regions. The methodology developed in this study; combining GPM IMERG monthly rainfall, Sentinel-2 monthly NDVI, and ML-based K-factor prediction within a monthly RUSLE simulation framework; is directly transferable to other data-scarce highland environments in the Horn of Africa and broader East Africa. Future work should assess the transferability of the approach to the Tigray and Amhara highlands of northern Ethiopia, the Yemeni highlands, and the Somali highland fringe, where comparable combinations of seasonal rainfall concentration, intensive agriculture on erodible highland soils, and data scarcity create similar modelling challenges. Standardised application of the methodology across multiple highland regions would enable regional-scale comparison of erosion dynamics and SWC needs.

PhD Theses (Scientific Contributions)

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: The 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.

References

- A. RIKIMARU, P. S. R. & S. M. (2002). Tropical forest cover density mapping. *International Society for Tropical Ecology*.
- Addis, H. K., & Klik, A. (2015). Predicting the spatial distribution of soil erodibility factor using USLE nomograph in an agricultural watershed, Ethiopia. *International Soil and Water Conservation Research*, 3(4), 282–290. <https://doi.org/10.1016/J.ISWCR.2015.11.002>
- Alewell, C., Borrelli, P., Meusburger, K., & Panagos, P. (2019). Using the USLE: Chances, challenges and limitations of soil erosion modelling. In *International Soil and Water Conservation Research* (Vol. 7, Number 3, pp. 203–225). International Research and Training Center on Erosion and Sedimentation and China Water and Power Press. <https://doi.org/10.1016/j.iswcr.2019.05.004>
- Ayehu, G. T., Tadesse, T., Gessesse, B., & Dinku, T. (2018). Validation of new satellite rainfall products over the Upper Blue Nile Basin, Ethiopia. *Atmospheric Measurement Techniques*, 11(4), 1921–1936. <https://doi.org/10.5194/amt-11-1921-2018>
- Ballabio, C., Borrelli, P., Spinoni, J., Meusburger, K., Michaelides, S., Beguería, S., Klik, A., Petan, S., Janeček, M., Olsen, P., Aalto, J., Lakatos, M., Rymaszewicz, A., Dumitrescu, A., Tadić, M. P., Diodato, N., Kostalova, J., Rousseva, S., Banasik, K., ... Panagos, P. (2017). Mapping monthly rainfall erosivity in Europe. *Science of the Total Environment*, 579, 1298–1315. <https://doi.org/10.1016/j.scitotenv.2016.11.123>
- Behrens, T., Schmidt, K., MacMillan, R. A., & Viscarra Rossel, R. A. (2018). Multi-scale digital soil mapping with deep learning. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-33516-6>
- Benavidez, R., Jackson, B., Maxwell, D., & Norton, K. (2018). A review of the (Revised) Universal Soil Loss Equation ((R)USLE): With a view to increasing its global applicability and improving soil loss estimates. *Hydrology and Earth System Sciences*, 22(11), 6059–6086. <https://doi.org/10.5194/hess-22-6059-2018>
- Borrelli, P., Alewell, C., Alvarez, P., Anache, J. A. A., Baartman, J., Ballabio, C., Bezak, N., Biddoccu, M., Cerdà, A., Chalise, D., Chen, S., Chen, W., De Girolamo, A. M., Gessesse, G. D., Deumlich, D., Diodato, N., Efthimiou, N., Erpul, G., Fiener, P., ... Panagos, P. (2021). Soil erosion modelling: A global review and statistical analysis. *Science of the Total Environment*, 780. <https://doi.org/10.1016/j.scitotenv.2021.146494>
- Borrelli, P., Robinson, D. A., Fleischer, L. R., Lugato, E., Ballabio, C., Alewell, C., Meusburger, K., Modugno, S., Schütt, B., Ferro, V., Bagarello, V., Oost, K. Van, Montanarella, L., & Panagos, P. (2017). An assessment of the global impact of 21st century land use change on soil erosion. *Nature Communications*, 8(1). <https://doi.org/10.1038/s41467-017-02142-7>
- Borrelli, P., Robinson, D. A., Panagos, D., P., Lugato, E., Yang, J. E., Alewell, C., Wuepper, D., Montanarella, L., & Ballabio, C. (2020). *Land use and climate change impacts on global soil erosion by water (2015-2070)*. <https://doi.org/10.1073/pnas.2001403117/-/DCSupplemental>
- Bousbih, S., Zribi, M., Pelletier, C., Gorra, A., Lili-Chabaane, Z., Baghdadi, N., Aissa, N. Ben, & Mougenot, B. (2019). Soil Texture Estimation Using Radar and Optical Data from Sentinel-1 and Sentinel-2. *Remote Sensing 2019, Vol. 11, Page 1520*, 11(13), 1520. <https://doi.org/10.3390/RS11131520>

- Bouyoucos, G. J. (1962). Hydrometer Method Improved for Making Particle Size Analyses of Soils1. *Agronomy Journal*, 54(5), 464–465.
<https://doi.org/10.2134/AGRONJ1962.00021962005400050028X>
- Breiman, L. (2001). *Random Forests* (Vol. 45).
- Chabhadiya et al. (2015). *GUJARAT TECHNOLOGICAL UNIVERSITY Sardar vallabhbhai Patel institute of technology, vasad (RAINWATER HARVESTING)*.
- Chang, C.-W., Laird, D. A., Mausbach, M. J., & Hurburgh, C. R. (2001). *Near-Infrared Reflectance Spectroscopy-Principal Components Regression Analyses of Soil Properties*.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Compton J. Tucker. (1978). *Red and Photographic Infrared Linear Combination for Monitoring Vegetation*.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). *System for Automated Geoscientific Analyses (SAGA) v. 2.1.4*.
<https://doi.org/10.5194/GMDD-8-2271-2015>
- De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., & Crombé, P. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, 186, 39–49.
<https://doi.org/10.1016/j.geomorph.2012.12.015>
- de Vente, J., & Poesen, J. (2005). Predicting soil erosion and sediment yield at the basin scale: Scale issues and semi-quantitative models. *Earth-Science Reviews*, 71(1–2), 95–125.
<https://doi.org/10.1016/j.earscirev.2005.02.002>
- Diek, S., Fornallaz, F., Schaepman, M. E., & de Jong, R. (2017). Barest Pixel Composite for agricultural areas using landsat time series. *Remote Sensing*, 9(12).
<https://doi.org/10.3390/RS9121245>
- Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., & Ceccato, P. (2018). Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society*, 144, 292–312. <https://doi.org/10.1002/qj.3244>
- Directive - EU - 2025/2360 - EN - EUR-Lex. (n.d.). Retrieved August 4, 2026, from <https://eur-lex.europa.eu/eli/dir/2025/2360/oj>
- Dobos, E., & Hengl, T. (2009). Chapter 20 Soil Mapping Applications. *Developments in Soil Science*, 33(C), 461–479. [https://doi.org/10.1016/S0166-2481\(08\)00020-2](https://doi.org/10.1016/S0166-2481(08)00020-2)
- Drury, S. A. (1987). Image Interpretation in Geology. *Image Interpretation in Geology*.
<https://doi.org/10.1007/978-94-010-9393-4>
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25–36. <https://doi.org/10.1016/j.rse.2011.11.026>
- Durigon, V. L., Carvalho, D. F., Antunes, M. A. H., Oliveira, P. T. S., & Fernandes, M. M. (2014). NDVI time series for monitoring RUSLE cover management factor in a tropical watershed.

- International Journal of Remote Sensing*, 35(2), 441–453.
<https://doi.org/10.1080/01431161.2013.871081>
- FAO. (2012). *Harmonized World Soil Database (version 1.2)*. FAO, Rome, Italy and IIASA, Luxemburg, Austria. <https://www.ptonline.com/articles/how-to-get-better-mfi-results>
- Fenta, H. M., Aynalem, D. W., Malmquist, L., Hailelassie, A., Tilahun, S. A., Barron, J., Adem, A. A., Adimassu, Z., Zimale, F. A., & Steenhuis, T. S. (2024). A critical analysis of soil (and water) conservation practices in the Ethiopian Highlands: Implications for future research and modeling. In *Catena* (Vol. 234). Elsevier B.V. <https://doi.org/10.1016/j.catena.2023.107539>
- Flanagan, D. C., Gilley, J. E., & Franti, T. G. (2007). *Water Erosion Prediction Project (WEPP): Development History, Model Capabilities, and Future Enhancements*. <https://digitalcommons.unl.edu/biosysengfacpub>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, 2. <https://doi.org/10.1038/sdata.2015.66>
- Ganasri, B. P., & Ramesh, H. (2016). Assessment of soil erosion by RUSLE model using remote sensing and GIS - A case study of Nethravathi Basin. *Geoscience Frontiers*, 7(6), 953–961. <https://doi.org/10.1016/j.gsf.2015.10.007>
- Gashaw, T., Tulu, T., & Argaw, M. (2018). Erosion risk assessment for prioritization of conservation measures in Geleda watershed, Blue Nile basin, Ethiopia. *Environmental Systems Research*, 6(1). <https://doi.org/10.1186/s40068-016-0078-x>
- Gee, G. W. and Bauder, J. W. (1986). Particle-size analysis. In: Klute A. (ed.) Methods of soil analysis. Part 1. Physical and mineralogical methods. In *ASA- SSSA, Madison, WI, USA* (2nd ed., pp. 383–411).
- Gelagay, H. S., & Minale, A. S. (2016). Soil loss estimation using GIS and Remote sensing techniques: A case of Koga watershed, Northwestern Ethiopia. *International Soil and Water Conservation Research*, 4(2), 126–136. <https://doi.org/10.1016/j.iswcr.2016.01.002>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Grunwald, S., Murad, M. O. F., Farrington, S., Wallace, W., & Rooney, D. (2024). Multi-Sensor Soil Probe and Machine Learning Modeling for Predicting Soil Properties. *Sensors*, 24(21). <https://doi.org/10.3390/s24216855>
- Hengl, T., De Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangquan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, 12(2). <https://doi.org/10.1371/journal.pone.0169748>
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., & Wolff, D. B. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1), 38–55. <https://doi.org/10.1175/JHM560.1>

- Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., Wolff, D. B., & Xie, P. (2020). *Chapter 19: Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) mission (IMERG)*. https://pmm.nasa.gov/sites/default/files/document_files/IMERG_QI.pdf,
- Hurni, H. (1985a). *Soil conservation manual for Ethiopia*. Ministry of Agriculture.
- Hurni, H. (1985b). *Soil conservation manual for Ethiopia*. Ministry of Agriculture.
- Hurni, H. (1988). International Mountain Society Degradation and Conservation of the Resources in the. In *Source: Mountain Research and Development* (Vol. 8, Number 2). http://www.jstor.orgURL:http://www.jstor.org/stable/3673438http://www.jstor.org/stable/3673438?seq=1&cid=pdf-reference#references_tab_contents
- J. G. Arnold, R. S. R. S. M. and J. R. W. (1998). LARGE AREA HYDROLOGIC MODELING AND ASSESSMENT. PART I: MODEL DEVELOPMENT. *Journal of the American Water Resources Association*.
- Jakab, G., Madarász, B., Szabó, J. A., Tóth, A., Zacháry, D., Szalai, Z., Kertész, Á., & Dyson, J. (2017). Infiltration and Soil Loss Changes during the Growing Season under Ploughing and Conservation Tillage. *Sustainability 2017, Vol. 9, Page 1726*, 9(10), 1726. <https://doi.org/10.3390/SU9101726>
- Jakab, G., Szabó, J., Szalai, Z., Mészáros, E., Madarász, B., Centeri, C., Szabó, B., Németh, T., & Sipos, P. (2016). Changes in organic carbon concentration and organic matter compound of erosion-delivered soil aggregates. *Environmental Earth Sciences 2015 75:2*, 75(2), 144-. <https://doi.org/10.1007/S12665-015-5052-9>
- Jetten, V., Govers, G., & Hessel, R. (2003). Erosion models: Quality of spatial predictions. *Hydrological Processes*, 17(5), 887–900. <https://doi.org/10.1002/hyp.1168>
- Jimenez-Munoz, J. C., Sobrino, J. A., Skokovic, D., Mattar, C., & Cristobal, J. (2014). Land surface temperature retrieval methods from landsat-8 thermal infrared sensor data. *IEEE Geoscience and Remote Sensing Letters*, 11(10), 1840–1843. <https://doi.org/10.1109/LGRS.2014.2312032>
- Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P., Mitchard, E. T. A., Reiche, J., Ryan, C. M., & Waske, B. (2016). A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sensing*, 8(1), 1–23. <https://doi.org/10.3390/rs8010070>
- Keller, B., Centeri, C., Szabó, J. A., Szalai, Z., & Jakab, G. (2021a). Comparison of the Applicability of Different Soil Erosion Models to Predict Soil Erodibility Factor and Event Soil Losses on Loess Slopes in Hungary. *Water 2021, Vol. 13, Page 3517*, 13(24), 3517. <https://doi.org/10.3390/W13243517>
- Keller, B., Centeri, C., Szabó, J. A., Szalai, Z., & Jakab, G. (2021b). Comparison of the Applicability of Different Soil Erosion Models to Predict Soil Erodibility Factor and Event Soil Losses on Loess Slopes in Hungary. *Water 2021, Vol. 13, Page 3517*, 13(24), 3517. <https://doi.org/10.3390/W13243517>
- Kenneth G. Renard, J. R. F. (1994). Using monthly precipitation data to estimate the R-factor in the revised USLE. *Journal of Hydrology*, 169(4)(93).
- K.G. Renard, G. R. F. G. A. W. D. K. M. and D. C. Y. (1997). *Predicting Soil Erosion by Water: A Guide to Conservation Planning With the Revised Universal Soil Loss Equation (RUSLE)*.

- Kinnell, P. I. A. (2005). Raindrop-impact-induced erosion processes and prediction: A review. *Hydrological Processes*, 19(14), 2815–2844. <https://doi.org/10.1002/hyp.5788>
- Lal, R. (2004). Soil Carbon Sequestration Impacts on Global Climate Change and Food Security. *Science*, 304(5677), 1623–1627. <https://doi.org/10.1126/SCIENCE.1097396>
- Levizzani, V., Kidd, C., Kirschbaum, D. B., Kummerow, C. D., Nakamura, K., & Turk, F. J. (n.d.). *Satellite Precipitation Measurement: Advances in Global Change Research* 67. Retrieved <http://www.springer.com/series/5588>
- Madarász, B., Jakab, G., Szalai, Z., Juhos, K., Kotroczó, Z., Tóth, A., & Ladányi, M. (2021). Long-term effects of conservation tillage on soil erosion in Central Europe: A random forest-based approach. *Soil and Tillage Research*, 209, 104959. <https://doi.org/10.1016/J.STILL.2021.104959>
- McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1–2), 3–52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Measho, S., Chen, B., Trisurat, Y., Pellikka, P., Guo, L., Arunyawat, S., Tuankruea, V., Ogbazghi, W., & Yemane, T. (2019). Spatio-temporal analysis of vegetation dynamics as a response to climate variability and drought patterns in the Semiarid Region, Eritrea. *Remote Sensing*, 11(6). <https://doi.org/10.3390/RS11060724>
- 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>
- 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>
- Merritt, W. S., Letcher, R. A., & Jakeman, A. J. (2003). A review of erosion and sediment transport models. *Environmental Modelling and Software*, 18(8–9), 761–799. [https://doi.org/10.1016/S1364-8152\(03\)00078-1](https://doi.org/10.1016/S1364-8152(03)00078-1)
- Meyer, W. B., & Turner, B. L. (1996). Land-use/land-cover change: Challenges for geographers. *GeoJournal*, 39(3), 237–240. <https://doi.org/10.1007/BF00188373>
- Montanarella, L., Pennock, D. J., McKenzie, N. J., Badraoui, M., Chude, V., Baptista, I., Mamo, T., Yemefack, M., Singh Aulakh, M., Yagi, K., Young Hong, S., Vijarnsorn, P., Zhang, G.-L., Arrouays, D., Black, H., Krasilnikov, P., Sobocká, J., Alegre, J., Henriquez, C. R., ... Vargas, R. (2015). *World's soils are under threat*. <https://doi.org/10.5194/soild-2-1263-2015>
- Montgomery, D. R. (2007). Soil erosion and agricultural sustainability. *Proceedings of the National Academy of Sciences of the United States of America*, 104(33), 13268–13272. <https://doi.org/10.1073/pnas.0611508104>
- Montgomery, D. R., & Matson, P. A. (2007). *Soil erosion and agricultural sustainability*. www.pnas.org/cgi/content/full/
- Moore, I. A. N. D., & Burch, G. J. (1986). Physical Basis of the Length-slope Factor in the Universal Soil Loss Equation. *Soil Conservation*, 50(1986), 1294–1298.
- Morgan, R. P. C. (2005). *Soil erosion & Conservation* (Third). Victoria:Blackwell Publishing.
- Morgan, R. P. C., Quinton, J. N., Smith, R. E., Govers, G., Poesen, J. W. A., Auerswald, K., Chisci, G., Torri, D., & Styczen, M. E. (1998). The European soil erosion model (EUROSEM): a

- dynamic approach for predicting sediment transport from fields and small catchments. *Earth Surface Processes and Landforms*, 23(6), 527–544. [https://doi.org/10.1002/\(SICI\)1096-9837\(199806\)23:6<527::AID-ESP868>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1096-9837(199806)23:6<527::AID-ESP868>3.0.CO;2-5)
- Mulder, V. L., de Bruin, S., Schaepman, M. E., & Mayr, T. R. (2011). The use of remote sensing in soil and terrain mapping - A review. In *Geoderma* (Vol. 162, Numbers 1–2, pp. 1–19). <https://doi.org/10.1016/j.geoderma.2010.12.018>
- Nelson D.W, S. L. E. (1982). Total carbon, organic carbon and organic matter. In: Page, A.L. (Ed.), *Methods of Soil Analysis. Part 2. Agron. Monogr. In ASA an SSSA, Madison, WI* (2nd ed., Vol. 9, pp. 539–580.).
- Nyssen, J., Poesen, J., Moeyersons, J., Deckers, J., Haile, M., & Lang, A. (2004a). Human impact on the environment in the Ethiopian and Eritrean highlands - A state of the art. *Earth-Science Reviews*, 64(3–4), 273–320. [https://doi.org/10.1016/S0012-8252\(03\)00078-3](https://doi.org/10.1016/S0012-8252(03)00078-3)
- Nyssen, J., Poesen, J., Moeyersons, J., Deckers, J., Haile, M., & Lang, A. (2004b). Human impact on the environment in the Ethiopian and Eritrean highlands - A state of the art. *Earth-Science Reviews*, 64(3–4), 273–320. [https://doi.org/10.1016/S0012-8252\(03\)00078-3](https://doi.org/10.1016/S0012-8252(03)00078-3)
- OCHA. (2023). *OCHA Eritrea – Subnational Administrative Boundaries. Humanitarian Data Exchange (HDX)*.
- Padarian, J., Minasny, B., & McBratney, A. B. (2020). Machine learning and soil sciences: A review aided by machine learning tools. In *SOIL* (Vol. 6, Number 1, pp. 35–52). Copernicus GmbH. <https://doi.org/10.5194/soil-6-35-2020>
- Panagos, P., Ballabio, C., Borrelli, P., & Meusburger, K. (2016). Spatio-temporal analysis of rainfall erosivity and erosivity density in Greece. *Catena*, 137, 161–172. <https://doi.org/10.1016/j.catena.2015.09.015>
- Panagos, P., Ballabio, C., Borrelli, P., Meusburger, K., Klik, A., Rousseva, S., Tadić, M. P., Michaelides, S., Hrabalíková, M., Olsen, P., Aalto, J., Lakatos, M., Rymaszewicz, A., Dumitrescu, A., Beguería, S., & Alewell, C. (2015). Rainfall erosivity in Europe. *Science of the Total Environment*, 511, 801–814. <https://doi.org/10.1016/j.scitotenv.2015.01.008>
- Panagos, P., Ballabio, C., Meusburger, K., Spinoni, J., Alewell, C., & Borrelli, P. (2017). Towards estimates of future rainfall erosivity in Europe based on REDES and WorldClim datasets. *Journal of Hydrology*, 548, 251–262. <https://doi.org/10.1016/j.jhydrol.2017.03.006>
- Panagos, P., Borrelli, P., Meusburger, K., Alewell, C., Lugato, E., & Montanarella, L. (2015). Estimating the soil erosion cover-management factor at the European scale. *Land Use Policy*, 48, 38–50. <https://doi.org/10.1016/j.landusepol.2015.05.021>
- Pásztor, L., Waltner, I., Centeri, C., Belényesi, M., & Takács, K. (2016). Soil erosion of Hungary assessed by spatially explicit modelling. *Journal of Maps*, 12, 407–414. <https://doi.org/10.1080/17445647.2016.1233913>
- (PDF) *Assessment of Soil and Water Conservation Options in Eritrea*. (n.d.). Retrieved September 14, 2026, from https://www.researchgate.net/publication/313478758_Assessment_of_Soil_and_Water_Conservation_Options_in_Eritrea
- Phinzi, K., & Ngetar, N. S. (2019). The assessment of water-borne erosion at catchment level using GIS-based RUSLE and remote sensing: A review. In *International Soil and Water Conservation*

- Research* (Vol. 7, Number 1, pp. 27–46). International Research and Training Center on Erosion and Sedimentation and China Water and Power Press.
<https://doi.org/10.1016/j.iswcr.2018.12.002>
- Pimentel, D., Harvey, C., Resosudarmo, P., Sinclair, K., Kurz, D., McNair, M., Crist, S., Shpritz, L., Fitton, L., Saffouri, R., & Blair, R. (1995). Environmental and economic costs of soil erosion and conservation benefits. *Science*, 267(5201), 1117–1123.
<https://doi.org/10.1126/science.267.5201.1117>
- P.J.J. Desmet and G. Govers. (1996). A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape. *J. Soil and Water Cons.*
- Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *SOIL*, 7(1), 217–240. <https://doi.org/10.5194/soil-7-217-2021>
- Predicting the spatial distribution of soil erodibility factor using USLE nomograph in an agricultural watershed, Ethiopia | Request PDF.* (n.d.). Retrieved July 21, 2026, from https://www.researchgate.net/publication/286903634_Predicting_the_spatial_distribution_of_soil_erodibility_factor_using_USLE_nomograph_in_an_agricultural_watershed_Ethiopia
- Renard *’a, K. G., & Freimund, J. R. (1994). Using monthly precipitation data to estimate the R-factor in the revised USLE. In *Journal of Hydrology ELSEVIER* (Vol. 1).
- Riley, S.J., De Gloria, S.D. and Elliot, R. (1999) *A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity. Intermountain Journal of Sciences*, 5, 23-27. - *References - Scientific Research Publishing.* (n.d.). Retrieved September 14, 2026, from <https://www.scirp.org/reference/referencespapers?referenceid=3355397>
- Rouse, W. and Haas, R.H. (1974) *Monitoring Vegetation Systems in the Great Plains with Ert.s.* - *References - Scientific Research Publishing.* (n.d.). Retrieved September 14, 2026, from <https://www.scirp.org/reference/referencespapers?referenceid=3587828>
- Schmidt, S., Alewell, C., Panagos, P., & Meusburger, K. (2016). Regionalization of monthly rainfall erosivity patterns in Switzerland. *Hydrology and Earth System Sciences*, 20(10), 4359–4373.
<https://doi.org/10.5194/hess-20-4359-2016>
- Sharpley, A. N. , and J. R. W. (1985). *EPIC-Erosion/Productivity Impact Calculator 1. Model Documentation.*
- Simms, A. D., Woodroffe, C. D., & Jones, B. G. (2003). Application of RUSLE for erosion management in a coastal catchment, southern NSW. *Faculty of Science, Medicine and Health*, 2, 678–683.
/articles/conference_contribution/Application_of_RUSLE_for_erosion_management_in_a_coastal_catchment_southern_NSW/27702651/1
- Stillhardt, Brigitta., Herweg, Karl., & Hurni, Hans. (2002). *Long-term monitoring of soil erosion and soil and water conservation in Afdeyu, Eritrea (1984-1998) : soil erosion and soil and water conservation database.* Centre for Development and Environment (CDE).
- TEFERI2, W. B. A. E. (2009). ASSESSMENT OF SOIL EROSION HAZARD AND PRIORITIZATION FOR TREATMENT AT THE WATERSHED LEVEL: CASE STUDY IN THE CHEMOGA WATERSHED, BLUE NILE BASIN, ETHIOPIA. *Land Degradation & Development*, 20(June), 609–622. <https://doi.org/10.1002/ldr.944> ASSESSMENTr

- Tesfay, T., Mohamed, E. S., Savin, I. Y., Kucher, D. E., Rebouh, N. Y., & Ogbazghi, W. (2025). Soil Organic Carbon Modelling with Different Input Variables: The Case of the Western Lowlands of Eritrea. *Sustainability (Switzerland)*, 17(21). <https://doi.org/10.3390/su17219884>
- Tiruneh, G. A., Ayalew, M., Ayalew, M., Mequanent, T. D., Chandrakala, M., Tibebe, D., & Reichert, M. (2025). Spatial Soil Erosion Modeling for Improved Land Management in Gumara Watershed , Ethiopia. *WILEY*, 2025. <https://doi.org/10.1155/aess/8547245>
- Van Der Knijff, J. M., Jones, R. J. A., & Montanarella, L. (2000). *European Soil Bureau Soil Erosion Risk Assessment Italy Soil Erosion Risk Assessment in Europe*.
- Vaudour, E., Gomez, C., Fouad, Y., & Lagacherie, P. (2019). Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems. *Remote Sensing of Environment*, 223, 21–33. <https://doi.org/10.1016/J.RSE.2019.01.006>
- Vrieling, A. (2006). Satellite remote sensing for water erosion assessment: A review. In *Catena* (Vol. 65, Number 1, pp. 2–18). <https://doi.org/10.1016/j.catena.2005.10.005>
- Vrieling, A., Sterk, G., & de Jong, S. M. (2010). Satellite-based estimation of rainfall erosivity for Africa. *Journal of Hydrology*, 395(3–4), 235–241. <https://doi.org/10.1016/j.jhydrol.2010.10.035>
- Wadoux, A. M. J. C., Minasny, B., & McBratney, A. B. (2020). Machine learning for digital soil mapping: Applications, challenges and suggested solutions. In *Earth-Science Reviews* (Vol. 210). Elsevier B.V. <https://doi.org/10.1016/j.earscirev.2020.103359>
- Walkley, A., science, I. B.-S., & 1934, undefined. (n.d.). An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Journals.Lww.Com A Walkley, IA Black Soil Science, 1934•journals.Lww.Com*. Retrieved July 14, 2026, from https://journals.lww.com/soilsci/citation/1934/01000/AN_EXAMINATION_OF_THE_DEGTJAREFF_METHOD_FOR.3.aspx.
- Walling, D. E. (1983). The sediment delivery problem. *Journal of Hydrology*, 65(1–3), 209–237. [https://doi.org/10.1016/0022-1694\(83\)90217-2](https://doi.org/10.1016/0022-1694(83)90217-2)
- Weiss, A. (2001). (n.d.).
- Were, K., Bui, D. T., Dick, Ø. B., & Singh, B. R. (2015). A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afrotropical landscape. *Ecological Indicators*, 52, 394–403. <https://doi.org/10.1016/j.ecolind.2014.12.028>
- Williams J., Shaffer J, Renard, K., Foster, G., Laflen, J., Lyles, L., Onstad, C., Sharpley A, Nicks, A., Jones, C., Richardson, C., Dyke, P., Cooley, K., & Smith. S. (1981). EPIC-Erosion/Productivity Impact Calculator 1. Model Documentation. *USDA Techn. Bull. 1759*, 235. <https://cir.nii.ac.jp/crid/1570854174228462720.bib?lang=en>
- Wischmeier, W. H. , and S. D. D. 1978. (1978). *PREDICTING RAINFALL EROSION LOSSES: A guide to conservation planning*.
- Wischmeier, W. H., & Smith, D. D. (1978). *Predicting Rainfall Erosion Losses: A guide to conservation planning*. USDA Agriculture Handbook no. 537, Washington, DC.

- Wischmeir, W. H. and Smith, DD. (1978). Predicting rainfall-Erosion Losses From Cropland East of the Rocky Mountains. In *A GUIDE TO CONSERVATION PLANNING* (Vol. 2). A guide to conservation planning “Science and Education Administration.”
- Zerihun, M., Mohammedyasir, M. S., Sewnet, D., Adem, A. A., & Lakew, M. (2018). Assessment of soil erosion using RUSLE, GIS and remote sensing in NW Ethiopia. *Geoderma Regional*, 12, 83–90. <https://doi.org/10.1016/J.GEODRS.2018.01.002>
- Zerizghi, T., Guo, Q., Tian, L., Wei, R., & Zhao, C. (2022). An integrated approach to quantify ecological and human health risks of soil heavy metal contamination around coal mining area. *Science of The Total Environment*, 814, 152653. <https://doi.org/10.1016/J.SCITOTENV.2021.152653>
- Zevenbergen, L. W., & Thorne, C. R. (1987). QUANTITATIVE ANALYSIS OF LAND SURFACE TOPOGRAPHY. In *EARTH SURFACE PROCESSES AND LANDFORMS* (Vol. 12).

APPENDIX

		Soil Particle size, soil texture and OM from the MOA (NARI) (USDA System)								
ID	X_UTM	Y_UTM	Region/Zoba	Sub Regions	Sam N.	Sand %	Clay %	Silt %	Textural Class	OM, %
1	484021.33	1695524.55	Maekel	Berik	ZB001	19.4	44.2	36.4	Clay	1.83
2	484018.42	1695538.46	Maekel	Berik	ZB002	45	20.9	34.1	Loam	1.02
3	487662.07	1710615.32	Maekel	Serejaka	ZB003	37.5	15.8	46.7	Loam	1.47
4	487657.19	1710621.80	Maekel	serejaka	ZB004	46.2	10.6	43.2	Loam	1.11
5	483657.29	1712987.07	Maekel	Serejaka	ZB005	35.0	17.7	47.3	Loam	1.18
6	471324.69	1686824.55	Maekel	Gala Nefhi	ZB006	22.5	27.9	49.6	Clay Loam	1.32
7	483669.56	1712967.08	Maekel	Serejaka	ZB007	32.5	40.6	26.9	Clay	0.64
8	471338.31	1686839.75	Maekel	Gala Nefhi	ZB008	54.4	10.5	35.1	Sandy Loam	0.91
9	493063.98	1692384.81	Maekel	Asmara	ZB009	54.0	19.4	26.6	Sandy Loam	1.86
10	493162.06	1692324.53	Maekel	Asmara	ZB010	49.4	29.5	21.1	Sandy Loam	1.29
11	493127.08	1692354.41	Maekel	Asmara	ZB011	20.0	61.3	18.7	Clay	1.44
12	493176.67	1692302.78	Maekel	Asmara	ZB012	38.2	41.4	20.4	Clay	2.55
13	493104.04	1692394.15	Maekel	Asmara	ZB013	20.0	67.3	12.7	Clay	1.93
14	493063.98	1692384.78	Maekel	Asmara	ZB014	23.3	58.1	18.6	Clay	2.28

15	493218.83	1692390.78	Maekel	Asmara	ZB015	29.6	26.6	43.8	Loam	2.7
16	493010.40	1692486.06	Maekel	Asmara	ZB016	34.4	30.9	35.7	Clay Loam	2.94
17	492879.39	1692253.65	Maekel	Asmara	ZB017	27.9	36.1	36.0	Clay Loam	2.04
18	493042.33	1692260.42	Maekel	Asmara	ZB018	48.5	24.8	26.8	Sandy Clay Loam	2.31
19	491689.07	1687117.03	Maekel	Gala Nefhi	ZB019	23.4	32.9	43.6	Clay Loam	2.29
20	491750.14	1687074.60	Maekel	Gala Nefhi	ZB020	29.5	35.0	35.5	Clay Loam	1.91
21	491603.17	1687036.91	Maekel	Gala Nefhi	ZB021	90.0	3.3	6.7	Sand	1.36
22	491706.80	1687051.98	Maekel	Gala Nefhi	ZB022	47.5	23.0	29.5	Loam	1.96
23	491718.11	1687242.29	Maekel	Gala Nefhi	ZB023	71.8	10.0	18.3	Sandy Loam	1.82
24	491708.81	1687119.82	Maekel	Gala Nefhi	ZB024	46.4	16.9	36.7	Loam	3.09
25	491623.90	1687138.66	Maekel	Gala Nefhi	ZB025	24.0	36.2	39.8	Clay Loam	1.8
26	491678.54	1687148.06	Maekel	Gala Nefhi	ZB026	44.1	25.0	30.9	Loam	1.67
27	491755.79	1687496.56	Maekel	Gala Nefhi	ZB027	61.2	15.0	23.7	Sandy Loam	1.74
28	491036.01	1687085.43	Maekel	Gala Nefhi	ZB028	63.2	16.5	20.3	Sandy Loam	1.45
29	491961.17	1687321.43	Maekel	Gala Nefhi	ZB029	36.6	28.9	34.6	Loam	1.99
30	492258.88	1686526.28	Maekel	Gala Nefhi	ZB030	39.7	28.7	31.6	Clay Loam	1.72
31	496844.08	1685095.64	Maekel	Gala Nefhi	ZB031	20.0	45.7	34.3	Clay	1.56
32	496847.21	1685048.69	Maekel	Gala Nefhi	ZB032	35.6	21.6	42.8	Loam	1.34
33	496806.07	1685069.07	Maekel	Gala Nefhi	ZB033	30.0	43.5	26.5	Clay	1.62

34	496827.98	1685082.23	Maekel	Gala Nefhi	ZB034	74.3	10.3	15.4	Sandy Loam	1.25
35	496911.15	1685060.77	Maekel	Gala Nefhi	ZB035	59.7	15.2	25.1	Sandy Loam	1.09
36	496904.44	1685171.66	Maekel	Gala Nefhi	ZB036	75.5	10.8	13.7	Sandy Loam	1.72
37	496816.36	1685135.44	Maekel	Gala Nefhi	ZB037	65.8	9.7	24.5	Sandy Loam	1.34
38	496848.10	1684970.89	Maekel	Gala Nefhi	ZB038	78.7	7.6	13.7	Loamy Sand	1.5
39	496860.18	1685044.22	Maekel	Gala Nefhi	ZB039	93.8	3.3	2.9	sand	1.29
40	496873.59	1684976.26	Maekel	Gala Nefhi	ZB040	90.0	3.8	6.2	sand	1.02
41	496814.12	1685113.98	Maekel	Gala Nefhi	ZB041	82.9	3.2	14.0	Loamy Sand	1.54
42	496999.68	1685086.25	Maekel	Gala Nefhi	ZB042	68.1	11.6	20.3	Sandy Loam	2.01
43	496945.58	1685021.87	Maekel	Gala Nefhi	ZB043	92.9	3.8	3.4	sand	1.69
44	496885.66	1685032.60	Maekel	Gala Nefhi	ZB044	80.8	6.1	13.1	Loamy Sand	2.55
45	496845.87	1685091.02	Maekel	Gala Nefhi	ZB045	52.9	16.4	30.7	Sandy Loam	3.18
46	496925.46	1685104.59	Maekel	Gala Nefhi	ZB046	36.7	30.1	33.2	Clay Loam	2.2
47	496873.60	1685082.68	Maekel	Gala Nefhi	ZB047	62.9	12.6	24.5	Sandy Loam	2.13
48	496870.46	1685059.43	Maekel	Gala Nefhi	ZB048	91.3	4.6	4.2	sand	2.17
49	499216.18	1682582.84	Maekel	Gala Nefhi	ZB049	83.4	7.6	9.0	Loamy Sand	1.62
50	499220.04	1682577.81	Maekel	Gala Nefhi	ZB050	46.8	25.0	28.2	Loam	1.82
51	499229.43	1682571.78	Maekel	Gala Nefhi	ZB051	86.4	3.4	10.2	Loamy Sand	1.53
52	499210.65	1682580.83	Maekel	Gala Nefhi	ZB052	65.0	10.2	24.9	Sandy Loam	1.77

53	499220.04	1682564.06	Maekel	Gala Nefhi	ZB053	80.6	5.9	13.6	Loamy Sand	1.81
54	499210.65	1682003.30	Maekel	Gala Nefhi	ZB054	38.2	38.3	23.5	Clay Loam	1.88
55	499264.97	1682602.96	Maekel	Gala Nefhi	ZB055	43.9	34.8	21.3	Clay Loam	1.85
56	499232.78	1682588.21	Maekel	Gala Nefhi	ZB056	79.8	7.6	12.6	Loamy Sand	3.04
57	499253.24	1682508.73	Maekel	Gala Nefhi	ZB057	79.5	8.7	11.8	Loamy Sand	1.57
58	499144.25	1682636.83	Maekel	Gala Nefhi	ZB058	71.7	11.5	16.8	Sandy Loam	1.41
59	499116.08	1682588.88	Maekel	Gala Nefhi	ZB059	67.2	13.3	19.5	Sandy Loam	1.17
60	499117.42	1682524.49	Maekel	Gala Nefhi	ZB060	89.8	5.1	5.1	Sand	2.46
61	499177.11	1682585.53	Maekel	Gala Nefhi	ZB061	90.0	6.1	3.9	Sand	3.31
62	495168.00	1689931.00	Maekel	Gala Nefhi	ZB062	58.4	7.6	34	Sandy loam	0.85
63	494758.00	1689800.00	Maekel	Gala Nefhi	ZB063	31.5	28.2	40.3	Clay loam	3.45
64	494575.00	1689720.00	Maekel	Gala Nefhi	ZB064	34.9	26	39.1	Loam	2.3
65	494662.00	1689999.00	Maekel	Gala Nefhi	ZB065	16.7	57.6	25.7	Clay	2.37
66	494412.00	1689950.00	Maekel	Gala Nefhi	ZB066	25.5	36	38.5	Clay loam	2.27
67	494014.00	1690332.00	Maekel	Gala Nefhi	ZB067	15.4	52.7	31.9	Clay	2.64
68	494530.00	1690384.00	Maekel	Gala Nefhi	ZB068	25.8	34.3	39.9	Clay loam	2.6
69	494257.00	1690434.00	Maekel	Gala Nefhi	ZB069	28.8	36.4	34.8	Clay loam	2.64
70	492959.00	1689949.00	Maekel	Gala Nefhi	ZB070	32.4	36.4	31.2	Clay loam	2.54
71	492646.00	1689874.00	Maekel	Gala Nefhi	ZB071	28.5	41.3	30.2	Clay	2.38

72	492362.00	1689669.00	Maekel	Gala Nefhi	ZB072	14.7	56.8	28.5	Clay	2.27
73	491855.00	1690146.00	Maekel	Gala Nefhi	ZB073	22.9	48.8	28.3	Clay	1.82
74	491861.00	1689205.00	Maekel	Gala Nefhi	ZB074	24.2	47.4	28.4	Clay	1.75
75	491411.00	1689260.00	Maekel	Gala Nefhi	ZB075	31.4	40.1	28.5	Clay	2.81
76	490853.00	1689275.00	Maekel	Gala Nefhi	ZB076	18.8	43.8	37.4	Clay	1.62
77	495164.75	1689978.73	Maekel	Gala Nefhi	ZB077	64.0	3.1	32.9	Sandy Loam	1.44
78	494785.03	1689748.07	Maekel	Gala Nefhi	ZB078	62.2	14.0	23.8	Sandy Loam	0.88
79	494307.38	1690195.00	Maekel	Gala Nefhi	ZB079	50.0	24.0	26.0	Sandy Clay Loam	1.82
80	494268.99	1690402.04	Maekel	Gala Nefhi	ZB080	67.6	8.2	24.2	Sandy Loam	1.59
81	493428.35	1690754.44	Maekel	Gala Nefhi	ZB081	52.4	19.1	28.5	Sandy Loam	2.12
82	493084.34	1690657.65	Maekel	Gala Nefhi	ZB082	55.7	15.7	28.6	Sandy Loam	2.41
83	493042.80	1688818.85	Maekel	Gala Nefhi	ZB083	68.8	4.1	27.1	Sandy Loam	2.03
84	492884.68	1688898.62	Maekel	Gala Nefhi	ZB084	59.1	13.7	27.2	Sandy Loam	2.04
85	492698.92	1689040.62	Maekel	Gala Nefhi	ZB085	58.1	8.2	33.7	Sandy Loam	2.00
86	492860.40	1689418.63	Maekel	Gala Nefhi	ZB086	36.6	25.5	37.9	Loam	1.96
87	492791.78	1689626.47	Maekel	Gala Nefhi	ZB087	60.8	5.0	34.2	Sandy Loam	2.10
88	492724.27	1690020.70	Maekel	Gala Nefhi	ZB088	44.0	25.5	30.5	Loam	2.04
89	492527.20	1690452.11	Maekel	Gala Nefhi	ZB089	43.9	25.6	30.5	Loam	1.78
90	492090.23	1690586.00	Maekel	Gala Nefhi	ZB090	53.1	24.0	22.9	Sandy Clay Loam	2.26

91	492019.68	1690198.45	Maekel	Gala Nefhi	ZB091	60.8	14.0	25.2	Sandy Loam	2.48
92	491623.24	1690082.23	Maekel	Gala Nefhi	ZB092	52.4	20.5	27.1	Sandy Clay Loam	1.96
93	491548.07	1689881.63	Maekel	Gala Nefhi	ZB093	50.7	19.0	30.3	Loam	2.26
94	491426.99	1689815.77	Maekel	Gala Nefhi	ZB094	52.8	15.7	31.5	Sandy Loam	1.88
95	491311.05	1689703.00	Maekel	Gala Nefhi	ZB095	49.0	20.7	30.3	Loam	1.75
96	490713.37	1689610.67	Maekel	Gala Nefhi	ZB096	51.3	22.4	26.3	Sandy Clay Loam	1.67
97	490898.00	1689462.98	Maekel	Gala Nefhi	ZB097	40.4	25.6	34.0	Loam	1.99
98	491352.19	1689284.80	Maekel	Gala Nefhi	ZB098	43.9	24.0	32.1	Loam	1.12
99	491618.50	1689217.51	Maekel	Gala Nefhi	ZB099	26.8	37.6	35.6	Clay Loam	1.21
100	491751.00	1688976.47	Maekel	Gala Nefhi	ZB100	24.9	42.6	32.5	Clay	1.53
101	491950.26	1688948.58	Maekel	Gala Nefhi	ZB101	43.5	23.4	33.1	Loam	1.66
102	492123.90	1688787.39	Maekel	Gala Nefhi	ZB102	55.1	14.0	30.9	Sandy Loam	2.12
103	492370.80	1688804.10	Maekel	Gala Nefhi	ZB103	56.4	18.4	25.2	Sandy Clay Loam	0.66
104	491756.60	1688645.35	Maekel	Gala Nefhi	ZB104	23.0	40.3	36.7	Clay	1.36
105	490834.04	1687404.90	Maekel	Gala Nefhi	ZB105	34.2	32.0	33.8	Clay Loam	1.19
106	490716.92	1687711.00	Maekel	Gala Nefhi	ZB106	46.5	24.9	28.6	Loam	1.71
107	490935.00	1687857.01	Maekel	Gala Nefhi	ZB107	54.0	23.4	22.6	Sandy Clay Loam	1.60
108	491009.67	1688523.94	Maekel	Gala Nefhi	ZB108	43.5	26.8	29.7	Loam	1.73
109	491070.74	1688831.34	Maekel	Gala Nefhi	ZB109	31.3	32.0	36.7	Clay Loam	1.24

110	490870.21	1689246.64	Maekel	Gala Nefhi	ZB110	41.7	30.3	28.0	Clay Loam	1.34
111	490538.51	1689306.09	Maekel	Gala Nefhi	ZB111	49.2	21.7	29.1	Loam	0.65
112	490238.99	1689090.50	Maekel	Gala Nefhi	ZB112	45.4	28.8	25.8	Sandy Clay Loam	0.99
113	489963.52	1688814.00	Maekel	Gala Nefhi	ZB113	46.5	24.9	28.6	Clay Loam	1.14
114	489403.50	1688379.33	Maekel	Gala Nefhi	ZB114	36.1	35.1	28.8	Clay Loam	1.41
115	489144.70	1688606.63	Maekel	Gala Nefhi	ZB115	39.6	26.5	33.9	Loam	2.28
116	489489.50	1688280.33	Maekel	Gala Nefhi	ZB116	44.0	35.0	21.0	Loam	1.26
117	489234.76	1688306.02	Maekel	Gala Nefhi	ZB117	25.6	36.8	37.6	Clay	1.06
118	486081.00	1713393.00	Maekel	Serejaka	ZB118	55.1	14	30.9	Sandy loam	1.55
119	486283.00	1713630.00	Maekel	Serejaka	ZB119	44.0	35.0	21.0	Loam	0.48
120	485643.00	1714586.00	Maekel	Serejaka	ZB120	23.8	25.4	50.8	Silt loam	0.38
121	485614.00	1714313.00	Maekel	Serejaka	ZB121	44.0	35.0	21.0	Loam	1.78
122	481710.50	1716473.00	Maekel	Serejaka	ZB122	78.2	0	21.8	Loamy sand	0.94
123	481592.30	1716713.00	Maekel	Serejaka	ZB123	46.7	16.4	37	Loam	1.87
124	481199.20	1716727.00	Maekel	Serejaka	ZB124	62.1	8.3	29.6	Sandy loam	1.62
125	480918.20	1716614.00	Maekel	Serejaka	ZB125	48.2	14.7	37.1	Loam	1.92
126	481015.00	1715988.00	Maekel	Serejaka	ZB126	64	6.6	29.4	Sandy loam	1.43
127	481130.60	1716137.00	Maekel	Serejaka	ZB127	62.2	10.1	27.7	Sandy loam	1.31
128	482447.70	1715576.00	Maekel	Serejaka	ZB128	49.1	9.1	41.8	Loam	1.29

129	482595.60	1716082.00	Maekel	Serejaka	ZB129	50.8	10.8	38.4	Loam	0.84
130	482513.60	1715224.00	Maekel	Serejaka	ZB130	47.8	6.1	46.1	Sandy loam	0.46
131	480832.3	1716821.0	Maekel	Serejaka	ZB131	64	3.3	32.7	Sandy loam	0.99
132	480872.6	1716662.0	Maekel	Serejaka	ZB132	64	3.3	32.7	Sandy loam	1.27
133	481003.4	1716700.0	Maekel	Serejaka	ZB133	56.7	5.1	38.2	Sandy loam	2.17
134	481109.1	1716592.0	Maekel	Serejaka	ZB134	59.2	6.1	34.8	Sandy loam	1.54
135	481314.7	1716674.0	Maekel	Serejaka	ZB135	58.1	6.1	35.8	Sandy loam	1.63
136	481416.2	1716830.0	Maekel	Serejaka	ZB136	56.1	10.2	33.7	Sandy loam	1.96
137	481547.9	1716770.0	Maekel	Serejaka	ZB137	56.1	9.6	34.3	Sandy loam	1.54
138	481564.6	1716526.0	Maekel	Serejaka	ZB138	53	9.6	37.4	Sandy loam	3.5
139	481851.7	1716403.0	Maekel	Serejaka	ZB139	62.5	6.1	31.4	Sandy loam	1.89
140	481906.1	1716248.0	Maekel	Serejaka	ZB140	66.8	4.3	28.9	Sandy loam	1.24
141	481762.5	1717160.0	Maekel	Serejaka	ZB141	59.2	7.8	33	Sandy loam	1.64
142	481995.3	1716286.0	Maekel	Serejaka	ZB142	65.5	4.3	30.2	Sandy loam	1.03
143	482549.4	1715172.0	Maekel	Serejaka	ZB143	46.3	6.8	46.9	Sandy loam	1.16
144	482542.1	1715458.0	Maekel	Serejaka	ZB144	58.1	1.5	40.4	Sandy loam	0.61
145	482123.3	1715675.0	Maekel	Serejaka	ZB145	60.7	3.3	36	Sandy loam	2.07
146	482882.1	1715542.0	Maekel	Serejaka	ZB146	55.2	3.3	41.5	Sandy loam	0.86
147	483125.8	1715609.0	Maekel	Serejaka	ZB147	46.1	6.1	47.9	Sandy loam	1.44

148	482920.8	1715843.0	Maekel	Serejaka	ZB148	41.3	10.5	48.2	Loam	1.32
149	482827.6	1716093.0	Maekel	Serejaka	ZB149	57.6	3.5	38.9	Sandy loam	0.41
150	482646.7	1716363.0	Maekel	Serejaka	ZB150	69.8	1.7	28.5	Sandy loam	0.88
151	482800.0	1716497.0	Maekel	Serejaka	ZB151	63.3	6.3	30.3	Sandy loam	1.85
152	482512.9	1716726.0	Maekel	Serejaka	ZB152	67.9	3.5	28.7	Sandy loam	0.82
153	482214.8	1716643.0	Maekel	Serejaka	ZB153	60.5	3.5	36	Sandy loam	0.55
154	482383.1	1715932.0	Maekel	Serejaka	ZB154	57.2	6.3	36.5	Sandy loam	1.36
155	482104.0	1715903.0	Maekel	Serejaka	ZB155	69.8	3.5	26.7	Sandy loam	0.8
156	482124.6	1716209.0	Maekel	Serejaka	ZB156	57.5	6.3	36.2	Sandy loam	1.8
157	481803.0	1715889.0	Maekel	Serejaka	ZB157	66.3	1.7	32	Sandy loam	1.01
158	481466.0	1716003.0	Maekel	Serejaka	ZB158	66.3	5.3	28.4	Sandy loam	0.98
159	481428.3	1716142.0	Maekel	Serejaka	ZB159	71	1.7	27.4	Sandy loam	0.67
160	481204.6	1715890.0	Maekel	Serejaka	ZB160	57.2	1.7	41.2	Sandy loam	0.74
161	481194.8	1716175.0	Maekel	Serejaka	ZB161	57.2	6.3	36.5	Sandy loam	1.4
162	480902.0	1716200.0	Maekel	Serejaka	ZB162	61.8	8.1	30.2	Sandy loam	1.18
163	482604.18	1685327.15	Maekel	Gala Nefhi	ZB163	52.4	20.14	27.55	Loam	0.88
164	484452.45	1681689.38	Maekel	Gala Nefhi	ZB164	64.3	7.89	27.81	Sandy Loam	0.49
165	477611.63	1686022.91	Maekel	Gala Nefhi	ZB165	28.3	39.99	31.71	Clay	0.76
166	469835.49	1666456.19	Maekel	Gala Nefhi	ZB166	42.1	18.74	39.16	Loam	1.82

167	491671.00	1705714.00	Maekel	Serejeka		42.65	8.795	48.56	Loam	2.66831
168	488230.00	1712664.00	Maekel	Serejeka		45.83	13.08	41.09	Loam	4.132667
169	489989.00	1711905.00	Maekel	Serejeka		36.45	11.03	52.52	Silt Loam	3.392034
170	497272.00	1705905.00	Maekel	Asmera		38.36	8.795	52.84	Silt Loam	3.084282
171	498665.00	1705125.00	Maekel	Asmera		46.24	10.5	43.26	Loam	2.123826
172	499483.00	1705290.00	Maekel	Asmera		41.05	10.5	48.45	Loam	2.252337
173	491220.00	1703730.00	Maekel	Serejeka		39.86	16.65	43.49	Loam	1.488031
174	484318.00	1710907.00	Maekel	Serejeka		38.25	10.5	51.25	Silt Loam	2.252337
175	483512.00	1709694.00	Maekel	Serejeka		39.83	26.51	33.66	Loam	3.479963
176	485038.00	1706527.00	Maekel	Berik		45.83	8.066	46.1	Loam	2.45527
177	488032.00	1707414.00	Maekel	Serejeka		46.71	8.795	44.5	Loam	2.135018
178	488809.00	1705048.00	Maekel	Serejeka		57.6	8.066	34.33	Sandy Loam	1.631287
179	470628.00	1687158.00	Maekel	Galane fhi		44.66	16.38	38.96	Loam	2.535333
180	473390.00	1688203.00	Maekel	Galane fhi		39.9	15.01	45.09	Loam	2.648756
181	475820.00	1685981.00	Maekel	Galane fhi		63.33	12.7	23.96	Sandy Loam	0.503731
182	482385.00	1683280.00	Maekel	Galane fhi		71.39	0	28.61	Loamy Sand	2.008251
183	485826.00	1683799.00	Maekel	Galane fhi		50.96	5.159	43.88	Sandy Loam	1.447809
184	487797.00	1686368.00	Maekel	Galane fhi		50.96	2	47.04	Sandy Loam	1.984258
185	491577.00	1683037.00	Maekel	Galane fhi		54.07	2	43.93	Sandy Loam	2.335939

186	490416.00	1680935.00	Maekel	Galane fhi		50.96	5.159	43.88	Sandy Loam	2.890886
187	491602.00	1678769.00	Maekel	Galane fhi		54	3	43	Sandy Loam	2.335939
188	495590.00	1685680.00	Maekel	Galane fhi		47.32	5.159	47.52	Sandy Loam	2.745696
189	493038.00	1692199.00	Maekel	Asmera		47.32	2	50.68	Silt Loam	2.565016
190	493594.00	1693709.00	Maekel	Asmera		42.65	2	55.35	Silt Loam	2.726338
191	494923.00	1694896.00	Maekel	Asmera		36.45	12.03	51.52	Silt Loam	3.339361
192	498598.00	1696675.00	Maekel	Asmera		52.52	2	45.48	Sandy Loam	2.432732
193	500495.00	1696103.00	Maekel	Asmera		75.87	2	22.13	Loamy Sand	3.187718
194	502106.00	1687322.00	Maekel	Asmera		75	3	22	Loamy Sand	3.2
195	492394.00	1702931.00	Maekel	Asmera		46.89	5.159	47.95	Sandy Loam	3.187718
196	492598.00	1700823.00	Maekel	Asmera		43.06	12.03	44.91	Loam	2.87798
197	512424.81	1631925.71	Debub	Mai Aini	HP1-A1-	54.8	16.4	28.8	Sandy Loam	1.02
198	509195.82	1625369.87	Debub	Mai Aini	HP1-A2	63.4	21.1	15.5	Sandy Clay Loam	0.42
199	510853.84	1629107.67	Debub	Mai Aini	HP2 – B1	63.4	19.4	17.2	Loam	0.69
200	507398.16	1627754.57	Debub	Mai Aini	HP5 – B2	46.4	25.1	28.5	Loam	0.82
201	510779.00	1627275.33	Debub	Tsorona	HP3 – C1	45.3	20.4	34.3	Clay loam	0.74
202	513726.92	1629050.14	Debub	Tsorona	HP6 – C2	39.6	30	30.4	Sandy Loam	0.49
203	509185.95	1626179.11	Debub	Mai Aini	HP4 – D1	66.9	19.4	13.7	Clay	0.7
204	505572.81	1628528.30	Debub	Mai Aini	HP9 – D2	50.8	31.7	17.5	Clay	0.74

205	512778.08	1629292.95	Debub	Mai Aini	HP7 – E1	30.5	48.6	20.9	Sandy Loam	0.5
206	479119.47	1665226.97	Debub	Dibarwa		17.14	37.46	45.4	Silty Clay Loam	1.38
207	478926.16	1665383.09	Debub	Dibarwa		17.14	34.29	48.57	Silty Clay Loam	1.68
208	479609.57	1665232.07	Debub	Dibarwa		22.15	49.58	28.27	Clay	2.2
209	479358.12	1665278.74	Debub	Dibarwa		23.48	32.43	44.09	Clay Loam	1.69
210	479024.11	1665555.55	Debub	Dibarwa		54.07	12.5	33.43	Sandy Loam	1.6
211	479668.69	1665238.66	Debub	Dibarwa		8.21	50.06	41.74	Silty Clay	1.69
212	479080.89	1665343.14	Debub	Dibarwa		47.21	11.45	41.34	Loam	1.02
213	479452.13	1665859.34	Debub	Dibarwa		22.96	36.37	40.67	Clay Loam	2.7
214	480227.57	1665234.86	Debub	Dibarwa		23.72	24.97	51.31	Silt Loam	1.75
215	480400.63	1665259.05	Debub	Dibarwa		28.71	40.1	31.18	Clay Loam	1.8
216	480336.18	1674340.97	Debub	Dibarwa		14.5	38.4	47.2	Silty Clay Loam	1.1
217	480672.57	1665298.65	Debub	Dibarwa		19.02	34.11	46.87	Silty Clay Loam	2
218	480323.02	1664991.45	Debub	Dibarwa		31	21.27	47.73	Loam	1
219	479892.98	1664839.18	Debub	Dibarwa		22.96	28.35	48.69	Clay Loam	1.6
220	479892.98	1664839.18	Debub	Dibarwa		15.46	37.9	46.64	Silty Clay Loam	2.4
221	480122.92	1664764.88	Debub	Dibarwa		26.35	32.28	41.37	Clay Loam	2.2
222	481319.07	1664646.66	Debub	Dibarwa		48.6	16	35.4	Loam	0.06
223	481883.12	1664360.86	Debub	Dibarwa		14.9	56	29.1	Clay	1

224	482035.78	1664416.05	Debub	Dibarwa		24.84	33.29	41.87	Clay Loam	1.02
225	481719.50	1664047.98	Debub	Dibarwa		24.46	47.91	27.63	Clay	10.5
226	481680.71	1663918.60	Debub	Dibarwa		17.3	54.9	27.8	Clay	1.78
227	481853.70	1663842.15	Debub	Dibarwa		25.5	47	27.5	Clay	1.4
228	482307.43	1664040.89	Debub	Dibarwa		20.56	44.27	35.17	Clay	1.21
229	482307.43	1664040.89	Debub	Dibarwa		33.16	40.68	26.17	Clay	0.91
230	481927.68	1663598.76	Debub	Dibarwa		12.6	44.2	43.4	Silt Clay	1.59
231	481927.68	1663598.76	Debub	Dibarwa		19.4	26.2	54.4	Silt Clay	1.75
232	481584.95	1663792.58	Debub	Dibarwa		25.2	30.83	43.96	Clay Loam	0.91
233	470151.20	1672059.41	Debub	Dibarwa	# 5	40.8	12.29	46.91	Silt Loam	3.15
234	476315.03	1665897.36	Debub	Dibarwa	# 6	26.9	17.64	55.46	Silt Loam	0.66
235	476194.10	1665783.22	Debub	Dibarwa	# 7	25.9	27.54	46.56	Clay Loam	0.14
236	484731.37	1674731.71	Debub	Dibarwa	# 22	31.4	13.49	55.11	Silt Loam	2.19
237	472593.82	1660294.31	Debub	Dibarwa	# 102	10.7	35.34	54.06	Silty Clay Loam	0.98
238	484620.15	1676371.14	Debub	Dibarwa	# 118	29.2	8.84	61.96	Silt Loam	1.48
239	476958.04	1666575.95	Debub	Dibarwa	# 317	20.4	30.34	49.26	Silty Clay Loam	1.54
240	469280.25	1661850.93	Debub	Dibarwa	# 401	21.6	21.59	56.81	Silty Loam	0.86
241	469514.69	1661513.74	Debub	Dibarwa	# 403	56.5	5.99	37.51	Sandy Loam	0.56
242	484874.04	1669616.83	Debub	Dibarwa	# 410	41.3	17.69	41.01	Loam	1.18

243	487443.40	1674243.64	Debub	Dibarwa	# 501	50.3	13.19	36.51	Loam	1.38
244	484694.00	1675178.00	Debub	Debarwa		45.03	2	52.97	Silt Loam	2.221426
245	482593.00	1676488.00	Debub	Debarwa		38.25	17.89	43.86	Loam	2.796784
246	486870.00	1668435.00	Debub	Debarwa		56.78	15.91	27.32	Sandy Loam	0.869846
247	481092.00	1668912.00	Debub	Debarwa		32.98	27.99	39.03	Clay Loam	1.577978
248	478316.00	1669174.00	Debub	Debarwa		45.03	14.19	40.78	Loam	0.903891
249	478790.00	1671623.00	Debub	Debarwa		36.45	26.37	37.17	Loam	1.155822
250	475705.00	1674615.00	Debub	Debarwa		48.42	12.03	39.55	Loam	1.155822
251	474322.00	1668823.00	Debub	Debarwa		43.33	2	54.67	Silt Loam	1.125182
252	475194.00	1665181.00	Debub	Debarwa		47.31	12.03	40.67	Loam	1.094542
253	479157.00	1666496.00	Debub	Debarwa		73	7.033	19.97	Sandy Loam	1.370305
254	472359.00	1650776.00	Debub	Mendefera		69.79	2	28.21	Sandy Loam	1.063901
255	459251.00	1652894.00	Debub	Areza		45.03	24.69	30.28	Loam	0.934531
256	465095.00	1648900.00	Debub	Areza		62	2	36	Sandy Loam	0.849419
257	447452.00	1609659.00	Debub	Maimine		76.21	2	21.79	Loamy Sand	2.010348
258	446646.00	1609013.00	Debub	Maimine		49.39	33.68	16.93	Sandy Clay Loam	0.897513
259	466830.00	1610476.00	Debub	Adiquala		63.29	20.18	16.53	Sandy Clay Loam	1.285186
260	480138.00	1612768.00	Debub	Adiquala		36.63	39.51	23.86	Clay Loam	2.317896
261	484574.00	1619605.00	Debub	Adiquala		28.8	44.34	26.86	Clay	1.464363

262	483753.00	1623782.00	Debub	Adiquala		28.8	44.16	27.04	Clay	1.754304
263	483852.00	1626230.00	Debub	Emnihaili		43.65	21.84	34.51	Loam	0.868193
264	482939.00	1633795.00	Debub	Emnihaili		26.8	42.58	30.62	Clay	1.819459
265	482056.00	1638403.00	Debub	Emnihaili		26.93	44.34	28.73	Clay	1.532776
266	480057.00	1640808.00	Debub	Mendefera		32.53	41.64	25.84	Clay	1.816201
267	481682.00	1644769.00	Debub	Mendefera		21.31	46.62	32.07	Clay	1.588158
268	482930.00	1644024.00	Debub	Mendefera		21.31	46.13	32.57	Clay	1.552323
269	477832.00	1646367.00	Debub	Mendefera		36.24	39.33	24.43	Clay Loam	1.207
270	477938.00	1645180.00	Debub	Mendefera		20.69	46.44	32.87	Clay	1.467621
271	480236.00	1649689.00	Debub	Mendefera		29.49	36.75	33.75	Clay Loam	1.399208
272	479569.00	1651796.00	Debub	Mendefera		41.79	36.27	21.93	Clay Loam	2.076822
273	480228.00	1657407.00	Debub	Debarwa		36.45	36.4	27.14	Clay Loam	1.379661
274	482694.00	1660114.00	Debub	Debarwa		45.81	29.44	24.75	Sandy Clay Loam	1.594673
275	491440.00	1662227.00	Debub	Dekemhare		61.71	4	34.29	Sandy Loam	1.386177
276	495293.00	1664657.00	Debub	Dekemhare		63.33	4	32.67	Sandy Loam	1.685891
277	503743.00	1679055.00	Debub	Dekemhare		73	2	25	Loamy Sand	2.135462
278	515223.00	1677861.00	Debub	Dekemhare		66.57	4	29.43	Sandy Loam	0.548933
279	511098.00	1676063.00	Debub	Dekemhare		57.16	7.033	35.81	Sandy Loam	0.931638
280	508679.00	1632853.00	Debub	Maiaini		31.74	15.91	52.35	Silt Loam	1.282254

281	510535.00	1636925.00	Debub	Maiaini		56.41	7.033	36.56	Sandy Loam	1.228827
282	505106.00	1640247.00	Debub	Maiaini		39.86	11.47	48.67	Loam	0.747982
283	503482.00	1643462.00	Debub	Maiaini		55.12	9.667	35.21	Sandy Loam	1.278915
284	505789.00	1656426.00	Debub	Dekemhare		61.71	9.667	28.63	Sandy Loam	0.71459
285	504770.00	1657734.00	Debub	Dekemhare		38.36	13.78	47.85	Loam	1.101937
286	507277.00	1658858.00	Debub	Dekemhare		50.1	9.667	40.23	Loam	1.769778
287	511714.00	1664537.00	Debub	Dekemhare		74.14	4	21.86	Loamy Sand	1.515999
288	516574.00	1662508.00	Debub	Segeneiti		45.83	7.033	47.14	Loam	2.64131
289	522713.00	1662242.00	Debub	Segeneiti		57.16	7.033	35.81	Sandy Loam	2.400888
290	524420.00	1652042.00	Debub	Segeneiti		73	2	25	Loamy Sand	1.202113
291	526884.00	1649325.00	Debub	Segeneiti		45.03	14.38	40.6	Loam	1.873293
292	525212.00	1652061.00	Debub	Segeneiti		49.39	12.7	37.9	Loam	1.582783
293	526541.00	1652283.00	Debub	Segeneiti		48.42	14.38	37.21	Loam	0.778035
294	535047.00	1646413.00	Debub	Adikeih		46.73	12.7	40.57	Loam	1.515999
295	537007.00	1646348.00	Debub	Adikeih		50.96	10.5	38.54	Loam	3.402649
296	538544.00	1644494.00	Debub	Adikeih		46.24	8.066	45.69	Loam	1.84658
297	539344.00	1645840.00	Debub	Adikeih		46.73	8.066	45.21	Loam	1.736386
298	541726.00	1640484.00	Debub	Adikeih		59.58	9.824	30.6	Sandy Loam	1.489285
299	540868.00	1640167.00	Debub	Adikeih		55.62	12.7	31.68	Sandy Loam	1.873293

300	543852.00	1613591.00	Debub	Senafe		36.92	16.65	46.43	Loam	3.619697
301	545488.00	1612735.00	Debub	Senafe		56.78	12.7	30.52	Sandy Loam	4.731652
302	545734.00	1624149.00	Debub	Senafe		45.19	13.92	40.89	Loam	3.539556
303	550313.00	1624056.00	Debub	Senafe		66.57	8.066	25.36	Sandy Loam	2.227249
304	547521.00	1625027.00	Debub	Senafe		52.67	21.49	25.84	Sandy Clay Loam	2.112534
305	513032.00	1660681.00	Debub	Segeneiti		55.12	14.38	30.51	Sandy Loam	1.504151
306	500340.00	1675041.00	Debub	Dekemhare		49.14	21.75	29.11	Loam	1.366341