EEET ECOLOGICAL ENGINEERING & ENVIRONMENTAL TECHNOLOGY

Ecological Engineering & Environmental Technology, 2026, 27(7), 235–250
https://doi.org/10.12912/27197050/223924
ISSN 2719–7050, License CC-BY 4.0

Received: 2026.05.17
Accepted: 2026.06.15
Published: 2026.07.01

Spatial variability and assessment of soil organic matter in the El Maleh sub-watershed (northeastern Algeria)

Nouha Menadjlia[1*], Ibtissem Samai[1], Mohamed Benslama[1]

[1] Laboratory of Soil and Sustainable Development Research, Department of Biology, Faculty of Science, Badji Mokhtar University, BP 12, 23000 Annaba, Algeria
* Corresponding author's e-mail: nouha.menadjlia@hotmail.fr

ABSTRACT

This study investigates the spatial variability of soil organic matter (SOM) in the El Maleh sub-watershed (northeastern Algeria), a semi-arid Mediterranean agroecosystem subject to significant soil degradation and erosion pressures. SOM is a fundamental indicator of soil quality, playing a central role in nutrient cycling, carbon sequestration, aggregate stability, and ecosystem resilience. Fifty-two georeferenced composite soil samples were collected from the 0–20 cm horizon using a systematic 1 km × 1 km sampling grid. SOM content was determined by the Walkley–Black wet oxidation method. Spatial structure was characterized through descriptive statistics and ordinary kriging, with a Gaussian semivariogram model selected based on residual sum of squares minimization and leave-one-out cross-validation. Slope and NDVI covariates were incorporated to assess environmental controls on SOM distribution. SOM ranged from 0.302 to 4.401 g/kg, with a mean of 2.274 ± 0.908 g/kg and a coefficient of variation of 40.0%, indicating marked spatial heterogeneity. The Gaussian semivariogram model provided the best fit (nugget = 0.671, range = 14.69 km, SPD = 35.8%), reflecting moderate spatial dependence. Cross-validation confirmed acceptable prediction accuracy (ME = −0.002, RMSE = 0.971 g/kg). Neither slope nor NDVI emerged as significant predictors of SOM distribution (Spearman $\rho = 0.186$, $p = 0.264$ and $\rho = -0.026$, $p = 0.858$, respectively), with principal component analysis confirming that SOM variability is largely independent of these environmental covariates. The study is limited by the moderate sampling density (52 points over 521 km²), the absence of seasonal monitoring, and the restriction to SOM as the sole soil quality indicator. These findings provide the first spatially explicit SOM baseline for the El Maleh sub-watershed, supporting soil conservation planning and sustainable land management in semi-arid Mediterranean agroecosystems.

Keywords: soil organic matter, spatial variability, ordinary kriging, geostatistics, semi-arid Mediterranean, Algeria.

INTRODUCTION

In Mediterranean regions, soils are particularly vulnerable to degradation under the combined pressure of climatic variability and intensifying land use. Soil organic matter (SOM) is a fundamental indicator of soil quality, regulating nutrient cycling, water retention, aggregate stability, and carbon sequestration (Brady and Weil, 2017; Wiesmeier et al., 2019). In semi-arid Mediterranean agroecosystems, SOM spatial distribution is controlled by the interplay of topography, vegetation cover, erosion dynamics, and land-use history (Lal, 2004; Fan et al., 2020), making its spatially explicit assessment a prerequisite for sustainable soil management.

Geostatistical approaches to SOM mapping have gained traction in North Africa over the past decade. In the adjacent Rhirane catchment – a sub-basin of the Mellah watershed sharing similar lithological and climatic conditions with El Maleh – Othmani et al. (2023) mapped soil erodibility and organic matter using ordinary kriging on 132 samples, demonstrating significant SOM differences between chromic cambisols (mean = 3.4%) and calcic cambisols (mean = 2.2%), and identifying land management as a primary driver of SOM spatial variability. However, their study focused on erodibility rather than SOM distribution per se, was limited to a single soil horizon (0–20 cm), and did not integrate recent high-resolution

remote sensing covariates. More broadly, available soil information for the El Maleh sub-watershed relies on coarse national maps that lack the spatial resolution required for sub-watershed scale management. To date, no spatially explicit geostatistical assessment of SOM has been conducted in this basin.

This gap has three concrete consequences for land management in the region: SOM depletion zones cannot be identified at the field scale, the relative contribution of topographic and vegetation drivers to SOM distribution remains unquantified, and soil conservation planning lacks the spatial foundation it requires.

This study therefore aims to produce the first high-resolution geostatistical map of SOM distribution in the El Maleh sub-watershed, using a reproducible R-based workflow integrating field sampling, Walkley–Black analysis, and ordinary kriging with Gaussian semivariogram modeling. Unlike previous regional studies, this work explicitly tests whether slope gradient and NDVI — derived from SRTM DEM and Landsat 8 OLI imagery — are significant predictors of SOM spatial variability at the watershed scale, a question that remains unresolved for semi-arid agroecosystems of northeastern Algeria. Three testable hypotheses were formulated based on established soil-formation principles. First, SOM will exhibit significant spatial variability across the watershed, driven by heterogeneity in land use, vegetation cover, and topographic position (CV > 30% expected). Second, SOM concentrations will be significantly higher in low-slope depositional zones than on steep erosion-prone slopes, reflecting reduced erosion and greater organic matter accumulation (Moore et al., 1993; Fan et al., 2020). Third, NDVI will be positively associated with SOM, as vegetation productivity constitutes the primary source of organic inputs to the soil (Lal, 2004; Wiesmeier et al., 2019). Testing these hypotheses in a semi-arid Mediterranean context where environmental drivers may be overridden by anthropogenic factors constitutes the scientific contribution of this work.

MATERIALS AND METHODS

Study area

The study was conducted in the El Maleh sub-watershed, located in the upstream portion of the Seybouse catchment in northeastern Algeria (Guelma province), between 36.286°–36.512° N and 6.841°–7.164° E, covering an area of approximately 521 km². The watershed includes several predominantly agricultural municipalities, such

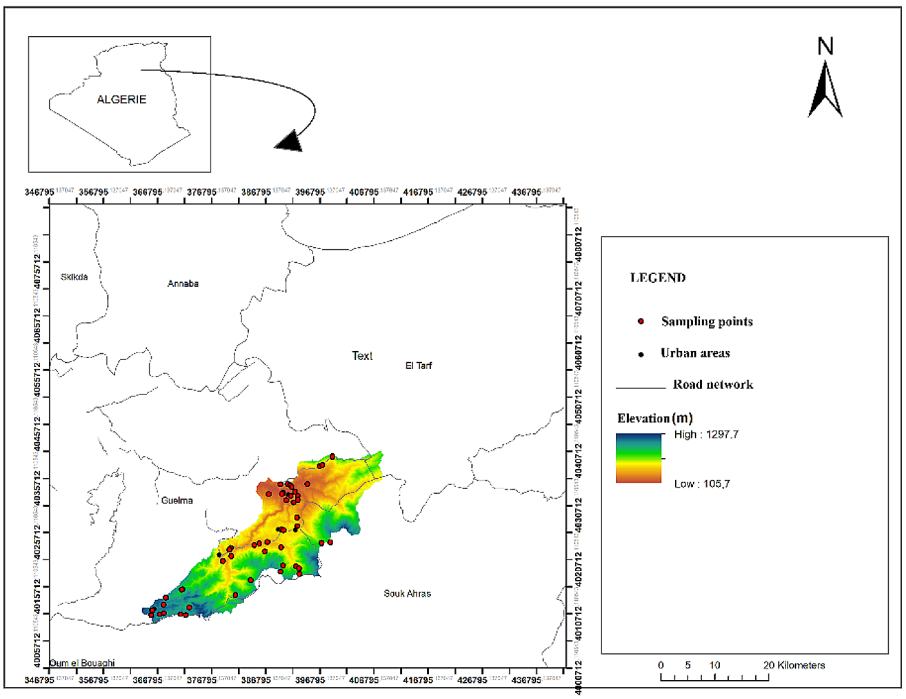

Figure 1. Location map of the El Maleh sub-watershed within the Seybouse catchment, northeastern Algeria, showing the main administrative boundaries and hydrographic network

as Medjez Sfa, Mechroha, Dahouara, and Hammam N'Bails (Figure 1).

Key morphometric characteristics were derived from a 30 m resolution SRTM Digital Elevation Model (DEM) using GIS-based watershed delineation in ArcGIS 10.4. The watershed has an elongated southwest–northeast shape, a hydrological length of approximately 32 km, and a mean elevation of around 650 m a.s.l., with higher elevations upstream and broader valleys downstream. These topographic features, combined with a semi-arid Mediterranean climate (hot, dry summers; mild, wet winters; annual precipitation 500–600 mm, mostly from October to April), influence soil formation and contribute to the spatial variability of soil organic matter (SOM) across the watershed.

Soils in the study area are dominated by Calcaric Regosols and Cambisols on slopes, with Fluvisols and Vertisols present in valley bottoms and agricultural fields. This diversity of soil types, shaped by the interaction of climate, topography, and land use, underlies the expected heterogeneity in SOM distribution.

Soil sampling design and field protocol

Sampling grid

A systematic 1 × 1 km sampling grid was generated in ArcGIS 10.4, resulting in 57 candidate locations across the 521 km² watershed. Five locations were inaccessible due to terrain constraints, leaving 52 effective sampling points after QA/QC. The choice of a 1 km grid was a compromise between logistical feasibility and coverage of the entire watershed. Finer resolutions (e.g., 500 m) would have required substantially more field effort, while stratified or cluster sampling was not feasible given the available resources and the need to cover all major geomorphological units (upland slopes, agricultural plains, valley bottoms). We acknowledge that a 1 km × 1 km grid represents a coarse resolution for soil organic matter (SOM) mapping, and is expected to produce a high nugget effect in semivariogram analysis. The observed nugget-to-sill ratio of 64.2% confirms that fine-scale variability is underrepresented at this grid spacing. While this grid allows identification of broad-scale spatial patterns of SOM, denser or stratified sampling would be necessary to capture small-scale heterogeneity

Field sampling procedure

Fieldwork was conducted during a single campaign in spring 2023, at the end of the wet season, when soil moisture is relatively stable, to minimize its confounding effect on SOM determination (Lal, 2004). A single campaign was considered acceptable because SOM in the topsoil of semi-arid Mediterranean croplands shows limited short-term seasonal variation, especially after the main rainy period, which dominates organic matter inputs and mineralization dynamics.

Each location was georeferenced using a Garmin eTrex GPS (±3 m accuracy). At each site, three sub-samples were collected within a 1 m radius at a standardized depth of 0–20 cm. This layer was chosen as it corresponds approximately to the Ap horizon in cultivated soils, representing the main zone of organic matter accumulation and tillage influence. Sub-samples were homogenized into a single composite sample (~500 g), stored under cool conditions (< 10 °C), and transported to the laboratory within 24 hours.

Land-use type, vegetation cover, surface stoniness, erosion evidence, and soil color (Munsell notation) were recorded at each sampling point. Sampling was designed to cover the main land-use categories (cropland, pasture, and fallow areas) proportionally across the watershed; however, stratified sampling within each category was not applied, which may contribute to residual bias in fine-scale SOM variability. Field data are recorded on standardized sheets and are available from the corresponding author upon request (Figure 2).

Laboratory analysis

Sample preparation

All soil samples were air-dried at room temperature (20–25 °C) for 5–7 days until mass stabilization, verified by weighing at 24-hour intervals. Dried samples were gently disaggregated using a ceramic mortar and sieved through a 2 mm stainless-steel mesh to obtain the fine-earth fraction (< 2 mm) used for all physicochemical analyses.

Soil organic matter (SOM) was selected as a key indicator of soil quality in semi-arid Mediterranean agroecosystems (Wiesmeier et al., 2019). SOM was determined using the Walkley–Black wet oxidation method (Walkley and Black, 1934; FAO, 2019), which estimates primarily the labile (oxidizable) fraction of soil organic carbon

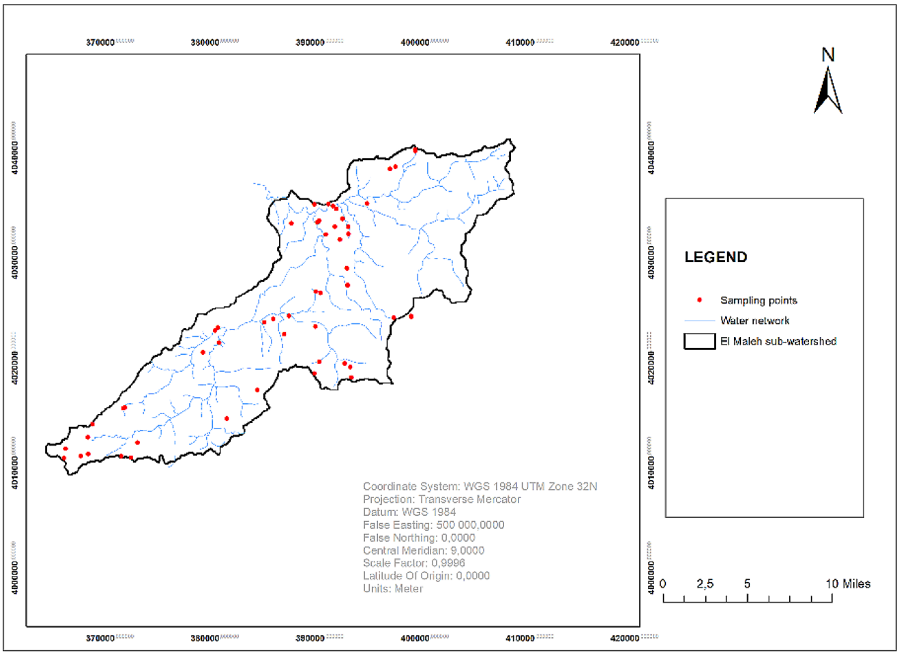

Figure 2. Spatial distribution of the 52 georeferenced soil sampling points across the El Maleh sub-watershed, overlaid on a digital elevation model (SRTM 30 m). Points were selected using a systematic 1 × 1 km grid design

(SOC) rather than total SOC, due to incomplete oxidation of more stable organic compounds.

SOC values were converted to SOM using the Van Bemmelen factor (SOM = SOC × 1.724), based on the conventional assumption that soil organic matter contains 58% carbon (Pribyl, 2010). Although widely used for comparability across studies, this conversion introduces uncertainty in semi-arid soils where SOM composition may deviate from this assumption due to higher proportions of resistant organic fractions.

The Walkley–Black procedure is known to underestimate total SOC because of incomplete oxidation, typically ranging from 10–30% depending on soil type, texture, and organic matter quality. To partially correct for this bias, the empirical correction factor of 1.33 was applied, although it assumes a relatively uniform oxidation efficiency and may not fully account for variability in semi-arid Mediterranean soils.

The analytical procedure followed standard protocols: 1.0 g of sieved soil was placed in a 500 mL Erlenmeyer flask; 10 mL of 1 N potassium dichromate ($K_2Cr_2O_7$) and 20 mL of concentrated sulfuric acid (H_2SO_4) were added and mixed. After a 30-minute digestion period at room temperature, 200 mL of distilled water was added, followed by 10 mL of 85% phosphoric acid (H_3PO_4) and 1 mL of diphenylamine indicator. Excess dichromate was back-titrated with 0.5 N ferrous sulfate ($FeSO_4$). SOM was calculated using the following equation:

$$SOM\ (g/kg) = [(B - S) \times N \times 0.003 \times \times 1.33 \times 1.724 / W] \times 1000 \quad (1)$$

where: B is the blank titration volume (mL), S is the sample titration volume (mL), N is the normality of $FeSO_4$, 0.003 is the milliequivalent weight of carbon (g/meq), 1.33 is the empirical correction for incomplete oxidation, 1.724 is the Van Bemmelen factor, and W is the sample weight (g).

All titrations were performed in triplicate and the mean value retained. *SOM* values are reported in g/kg of oven-dry soil equivalent.

Physicochemical parameters

SOM was the primary physicochemical parameter determined in this study, selected as a key indicator of soil quality in semi-arid Mediterranean agroecosystems (Wiesmeier et al., 2019). SOM was quantified using the Walkley–Black wet oxidation method (Walkley and Black, 1934; FAO, 2019), which measures SOC through dichromate oxidation and back-titration, subsequently converted to SOM using the Van Bemmelen factor (SOM = SOC × 1.724), based on the assumption that organic matter contains 58% carbon (Pribyl, 2010). The analytical procedure was as follows:

1.0 g of sieved soil was placed in a 500 mL Erlenmeyer flask; 10 mL of 1 N potassium dichromate ($K_2Cr_2O_7$) and 20 mL of concentrated sulfuric acid (H_2SO_4) were added and mixed. After a 30-minute digestion period at room temperature, 200 mL of distilled water was added, followed by 10 mL of 85% phosphoric acid (H_3PO_4) and 1 mL of diphenylamine indicator. Excess dichromate was back-titrated with 0.5 N ferrous sulfate ($FeSO_4$). SOM content was calculated as:

$$SOM\ (g/kg) = [(B - S) \times N \times 0.003 \times \times 1.33 \times 1.724 / W] \times 1000 \quad (2)$$

where: B is the volume of $FeSO_4$ used in the blank titration (mL), S is the volume used in the sample titration (mL), N is the normality of $FeSO_4$, 0.003 is the milliequivalent weight of carbon (g/meq), 1.33 is the Walkley–Black correction factor for incomplete oxidation, 1.724 is the Van Bemmelen factor, and W is the sample weight (g).

All titrations were performed in triplicate and the mean value retained. *SOM* values are reported in g/kg of oven-dry soil equivalent.

Quality assurance and quality control (QA/QC)

Analytical precision was assessed through triplicate titrations for each sample (n = 3), yielding a mean relative standard deviation (RSD) typically below 5% for the Walkley–Black method, consistent with established analytical performance criteria for this procedure (FAO, 2019). Systematic analytical consistency was further supported by periodic standardization of the $FeSO_4$ titrant against $K_2Cr_2O_7$, ensuring stability of reagent concentration across analytical batches.

While the Walkley–Black and Van Bemmelen correction factors (1.33 and 1.724, respectively) were applied for comparability with established SOM datasets, they introduce known methodological uncertainty, particularly in heterogeneous semi-arid soils where oxidation efficiency and organic matter composition may vary.

Outlier detection was performed using the interquartile range (IQR) method. Given the relatively small sample size (n = 52), IQR was preferred over parametric outlier tests because it is non-parametric and less sensitive to deviations from normality, which is common in spatially structured environmental datasets. Nevertheless, its sensitivity is reduced in small to moderate datasets, and results should therefore be interpreted cautiously. One sample (OBJECTID 12, Value = −0.98 g/kg) was was identified as a clear analytical artefact (negative SOM value) and removed following verification, reducing the final dataset to 52 samples. The QA/QC procedures applied are summarized in Table 1.

Environmental covariates

Three environmental covariates were derived from remote sensing and terrain analysis to support interpretation of SOM spatial patterns. All raster layers were reprojected to WGS 1984 UTM Zone 32N (EPSG:32632) prior to extraction to ensure geometric consistency.

Slope (%)

Slope values were derived from the SRTM 30 m DEM using the Slope tool in ArcGIS 10.4 Spatial Analyst and expressed as percentage rise, reprojected to UTM Zone 32N. Values were extracted at the 52 sampling locations using the Extract Multi Values to Points tool.

Due to the spatial extent of the DEM being slightly smaller than the full watershed boundary, slope values could not be retrieved for 21 sampling points located near the watershed edge. This reflects a data processing limitation rather than

Table 1. Quality assurance and quality control (QA/QC) procedures applied during SOM determination by the Walkley–Black wet oxidation method (El Maleh sub-watershed, 2023)

QA/QC procedure	Protocol details	Status
Walkley–Black correction factor	Factor 1.33 applied to all samples	☑
Van Bemmelen conversion factor	SOC × 1.724 = SOM (Pribyl, 2010)	☑
Triplicate titrations	Mean of 3 titrations retained per sample	☑
$FeSO_4$ standardization	Standardized against $K_2Cr_2O_7$ before each session	☑
Outlier screening	IQR method; 1 sample excluded (Value = −0.98 g/kg)	☑

Note: SOM – soil organic matter; SOC – soil organic carbon; IQR – interquartile range; $FeSO_4$ – ferrous sulfate; $K_2Cr_2O_7$ – potassium dichromate; ☑ – procedure applied.

natural absence, and these points were therefore excluded from slope-based analyses (N = 31).

Excluding ~40% of the slope dataset reduces the statistical power for detecting relationships between slope and SOM, and may limit the ability to capture fine-scale heterogeneity at watershed margins, even though the Mann–Whitney U test showed no significant difference in SOM between excluded (mean = 2.34 ± 0.95 g/kg) and included points (mean = 2.23 ± 0.89 g/kg; U = 299.0, p = 0.628).

Furthermore, the 30 m DEM resolution introduces a scale mismatch with the 1 × 1 km sampling grid. Local slope variability within each grid cell is averaged in the DEM, likely contributing to underestimation of slope extremes and reducing the sensitivity of slope–SOM correlations at fine scales. This limitation should be considered when interpreting slope-related results.

Vegetation index (NDVI)

NDVI was computed from Landsat 8 OLI multispectral imagery (30 m spatial resolution, Level-1 Terrain Precision, path 193/row 035) acquired on 26 November 2023 (product ID: LC08_L1TP_193035_20231126). The use of a single-date scene introduces a clear temporal mismatch with field soil sampling (spring 2023), which constitutes a methodological limitation of this study, as vegetation conditions at the time of image acquisition do not directly correspond to in-situ soil measurements. This mismatch may introduce noise in NDVI–SOM relationships and limits the interpretation of NDVI as a direct proxy for soil organic matter inputs.

Despite this limitation, the selected image represents the only cloud-free acquisition closest to the study area within the same annual cycle, and was therefore used as a pragmatic compromise. A single-date NDVI was preferred over seasonal composites to avoid additional uncertainties associated with phenological averaging in semi-arid systems, where short but intense vegetation growth periods can be smoothed out in multi-temporal products, potentially masking spatial gradients relevant to SOM distribution.

NDVI was calculated from top-of-atmosphere (TOA) reflectance using bands B5 (NIR, 0.85–0.88 µm) and B4 (Red, 0.64–0.67 µm) in ArcGIS 10.4 Spatial Analyst using the Raster Calculator tool. It is further noted that the use of TOA reflectance (rather than surface reflectance) introduces additional atmospheric effects, which may reduce direct comparability with ground-based soil processes and slightly increase uncertainty in NDVI-derived correlations.

NDVI values were extracted at all 52 sampling locations and used as a proxy for vegetation productivity and potential organic matter inputs to soil (Figures 3, 4).

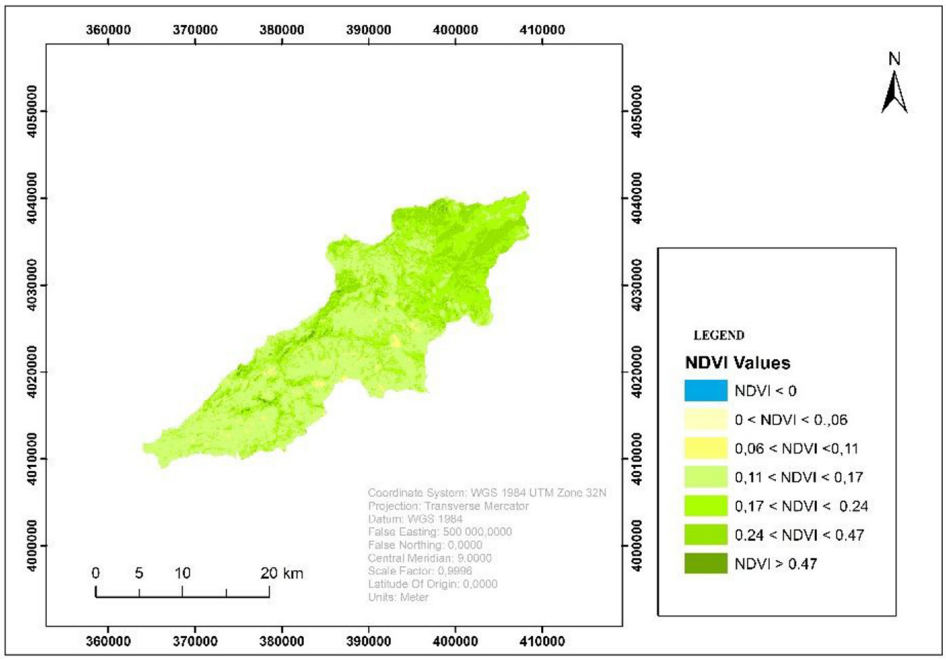

Figure 3. Spatial distribution of the normalized difference vegetation index (NDVI) derived from Landsat 8 OLI imagery (26 November 2023, 30 m resolution, top-of-atmosphere reflectance, path 193/row 035) across the El Maleh sub-watershed

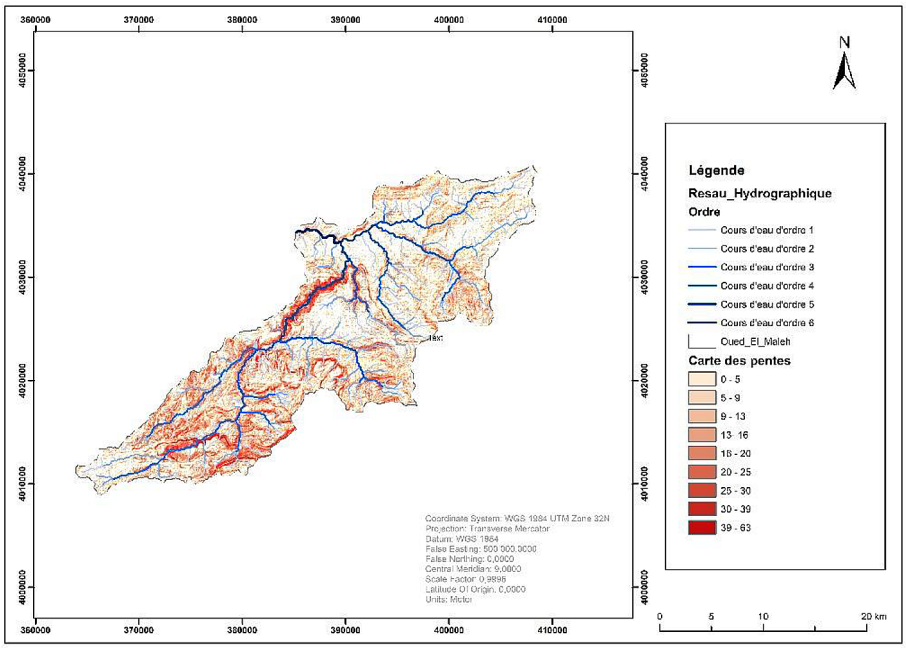

Figure 4. Slope map of the El Maleh sub-watershed derived from the SRTM 30 m Digital Elevation Model, reprojected to WGS 1984 UTM Zone 32N (EPSG:32632). Values expressed as percentage rise

GEOSTATISTICAL AND STATISTICAL ANALYSIS

Software and reproducibility

All statistical and geostatistical analyses were performed in R version 4.3.0 (R Core Team, 2023). Package versions were specified consistently across all analyses as follows: FactoMineR 2.9 (Lê et al., 2008) for principal component analysis, ggplot2 3.4.0 for data visualization, and gstat 2.1-1 (Pebesma, 2004) for semivariogram modelling and ordinary kriging. Base R functions were used for descriptive statistics and correlation analyses to maintain methodological transparency.

To ensure reproducibility of the geostatistical workflow, a fixed random seed (set.seed(1234)) was applied prior to all stochastic procedures involved in variogram fitting and kriging interpolation. This guarantees that model initialization and any internally random optimization steps in gstat produce identical outputs upon re-running the analysis.

All spatial data were projected to EPSG:32632 (WGS 1984 UTM Zone 32N) to enable distance-based geostatistical computations in metres. This projection was selected as a compromise between minimizing spatial distortion and maintaining metric consistency across the north–south extent of the Seybouse catchment. Although the study area lies within Algeria, its longitudinal span (~6.8°–7.2°E) remains well within the acceptable distortion limits of UTM Zone 32N, making it suitable for local-to-catchment scale spatial analysis.

The complete annotated R scripts are available from the corresponding author upon request.

Descriptive statistics

Descriptive statistics were computed for all 52 SOM values using base R functions, including mean, median, standard deviation (Bessel's correction), coefficient of variation, skewness, and Pearson kurtosis (excess kurtosis + 3). Outliers were screened using the interquartile range method (bounds: Q1 − 1.5 × IQR and Q3 + 1.5 × IQR). Normality was assessed using the Shapiro–Wilk test, as a prerequisite for ordinary kriging.

Correlation and regression analyses

Monotonic associations between SOM and each covariate (slope, NDVI) were assessed using Spearman's rank correlation (cor.test(), method = "spearman") and Pearson's r (cor.test(), method = "pearson"). The joint effect of slope and NDVI on SOM was assessed through ordinary least squares (OLS) multiple regression using the lm() function in R. Standard errors, t-statistics, and two-tailed p-values were extracted from the model summary.

Principal component analysis

To explore the multivariate structure of the SOM–slope–NDVI dataset, a principal component analysis (PCA) was performed using the PCA() function from the FactoMineR package 2.9 (Lê et al., 2008) in R 4.3.0, after standardization to zero mean and unit variance. The analysis was conducted on the subset of complete cases (n = 31), as slope values were unavailable for a portion of the dataset due to DEM edge effects, and no imputation was applied to avoid introducing artificial spatial structure into a small geospatial dataset. This decision prioritizes data integrity over sample size preservation, although it results in reduced statistical power and may affect the stability of eigenvector estimation. Given the relatively small sample size, PCA results should be interpreted with caution, as eigenvectors and component loadings may exhibit sensitivity to sampling variability. This limitation is particularly relevant in multivariate environmental datasets where n/p ratios are low, potentially leading to less stable ordination structures. Prior to PCA, multicollinearity among predictors was assessed using variance inflation factors (VIF), confirming no severe collinearity issues (Slope: VIF = 1.52; NDVI: VIF = 1.52), and supporting the suitability of variables for multivariate decomposition. Results are reported as eigenvalues, percentage variance explained, cumulative variance, and variable loadings (correlation of each variable with each principal component). The biplot was generated using ggplot2 3.4.

Experimental semivariogram

Spatial autocorrelation was characterized through the experimental semivariogram, computed using the classical Matheron estimator implemented in the gstat package (Pebesma, 2004) in R 4.3.0:

$$\gamma(h) = 1 / 2N(h) \times \Sigma[Z(x_i) - Z(x_i + h)]^2 \qquad (3)$$

where: $\gamma(h)$ is the semivariance at lag distance h, $N(h)$ is the number of sample pairs at that lag, and $Z(x_i)$ is the SOM value at location x_i.

The semivariogram was computed with 19 lag classes over a maximum distance of 25,000 m.

Semivariogram model fitting and selection

Four theoretical models were fitted to the experimental semivariogram using the fit.variogram() function in the gstat package (Pebesma, 2004): spherical, exponential, Gaussian, and linear. Model selection was based on two independent criteria applied in parallel: (i) minimization of the residual sum of squares (RSS) between experimental and theoretical semivariances; and (ii) minimization of the LOOCV RMSE. The spatial dependence index (SPD) was calculated as:

$$SPD\ (\%) = C_1 / (C_0 + C_1) \times 100 \qquad (4)$$

where: C_0 is the nugget and C_1 is the partial sill (Cambardella et al., 1994). SPD < 25% = weak spatial dependence; 25–75% = moderate; > 75% = strong.

Prior to isotropic semivariogram fitting, directional semivariograms were computed in four directions (0°, 45°, 90°, 135°) with a tolerance angle of 30° to assess potential geometric anisotropy (Figure 9). An isotropic Gaussian semivariogram model was retained as the most parsimonious choice for this dataset (see Discussion for full anisotropy assessment).

Ordinary kriging and cross-validation

Spatial interpolation was performed using ordinary kriging based on the fitted semivariogram model. The output prediction raster was generated at 100 m resolution. Prediction accuracy was assessed using leave-one-out cross-validation (LOOCV), implemented using the krige.cv() function in the gstat package (Pebesma, 2004) in R 4.3.0. In LOOCV, each of the 52 samples is withheld in turn; the remaining 51 points are used to krige the withheld location; the predicted value is compared to the observed value. The following statistics were computed: Mean error (ME) = mean of (predicted − observed), measuring systematic bias (values close to 0 indicate unbiased prediction); root mean square error (RMSE) = √[mean of (predicted − observed)²], measuring overall prediction error; and Pearson R^2 between predicted and observed values, measuring directional agreement. These statistics were computed directly in R 4.3.0 using base R functions. No seasonal corrections were applied, as sampling was conducted during a single campaign at the end of the wet season — the period of maximum SOM stability in semi-arid Mediterranean soils (Lal, 2004).

RESULTS

Descriptive statistics of SOM

Following the data cleaning procedure described in the Laboratory Analysis section, one sample (OBJECTID 12) was excluded due to a physically impossible negative SOM value (−0.98 g/kg), attributed to a laboratory titration error during the Walkley–Black back-titration. To evaluate the sensitivity of the dataset to this exclusion, descriptive statistics were computed both including (N = 53) and excluding (N = 52) this observation. Inclusion of the erroneous value resulted in a negligible reduction in mean SOM from 2.274 to 2.252 g/kg and a moderate increase in the coefficient of variation from 40.0% to 42.3%. The outcome of the Shapiro–Wilk normality test remained unchanged (W = 0.968, p = 0.187). Given its non-physical nature and minimal influence on the overall distributional structure, this observation was excluded from all subsequent analyses. All analyses reported hereafter are therefore based on N = 52 samples. This exclusion does not alter the statistical behaviour of the dataset nor the performance of subsequent statistical and geostatistical models.

Descriptive statistics were computed in R 4.3.0 using base R functions. The mean was calculated as the arithmetic mean, while the standard deviation was computed using Bessel's correction. The coefficient of variation was expressed as (standard deviation divided by mean) multiplied by 100. Skewness and kurtosis were estimated using the moments package, with kurtosis reported according to Pearson's definition. Outliers were screened using the IQR method, with bounds defined as Q1 − 1.5 × IQR and Q3 + 1.5 × IQR. Normality was assessed using the Shapiro–Wilk test (shapiro.test() in base R).

SOM values ranged from 0.302 to 4.401 g/kg, with a mean of 2.274 ± 0.908 g/kg and a median of 2.108 g/kg. The coefficient of variation of 40.0% indicates substantial spatial variability, which is commonly interpreted in soil studies as moderate to high heterogeneity (Wilding, 1985; Cambardella et al., 1994). The distribution exhibited slight positive skewness (0.467) and near-normal kurtosis (2.806). No additional outliers were identified within the IQR-defined limits of 0.016 to 4.433 g/kg. The Shapiro–Wilk test indicated no significant departure from normality (W = 0.971, p = 0.241), supporting the assumption of normality required for ordinary kriging. All descriptive statistics are presented in Table 2.

Table 2. Descriptive statistics of soil organic matter (SOM, g/kg) for the 52 sampling points in the El Maleh sub-watershed

Parameter	Value
N	52
Minimum	0.302 g/kg
Maximum	4.401 g/kg
Mean ± SD	2.274 ± 0.908 g/kg
Median	2.108 g/kg
Q1 / Q3	1.672 / 2.777 g/kg
CV (%)	40.0
Skewness	0.467
Kurtosis (Pearson)	2.806
Shapiro–Wilk W	0.971
Shapiro–Wilk p-value	0.241
Outliers detected	0

Note: SD = standard deviation; CV = coefficient of variation; Q1, Q3 = first and third quartiles.

Semivariogram analysis and model selection

All coordinates were first reprojected from WGS84 (EPSG:4326) to WGS 1984 UTM Zone 32N (EPSG:32632) using the sp and rgdal packages in R, ensuring all distance calculations were expressed in metres. The experimental semivariogram was then computed using the variogram() function from the gstat package (Pebesma, 2004), with 19 lag classes over a maximum distance of 25,000 m.

Four theoretical models (spherical, exponential, Gaussian, and linear) were fitted to the experimental semivariogram. Model selection was based on two complementary criteria: minimization of the residual sum of squares (RSS) between experimental and theoretical semivariances, and minimization of leave-one-out cross-validation (LOOCV) root mean square error (RMSE). The comparative performance of all models is summarized in Table 3, and the fitted Gaussian model is presented in Figure 5.

The Gaussian model was retained due to slightly superior cross-validation performance (RMSE_CV = 0.971 g/kg; R^2_CV = 0.094) and theoretical appropriateness for SOM data exhibiting smooth, continuous spatial variation consistent with the agricultural mosaic of the El Maleh watershed (Webster and Oliver, 2007).

The estimated range of 14.69 km should be interpreted cautiously: given the 1 × 1 km sampling grid, this value provides only an approximate indication of spatial correlation extent rather than a precise process-scale measurement.

The spatial dependence index (SPD) of 35.8% indicates moderate spatial dependence according to the classification of Cambardella et al. (1994). The high nugget-to-sill ratio (64.2%) reveals substantial fine-scale SOM variability operating below the 1 km sampling resolution, implying that the predictive capacity of the kriging model is limited at short distances. Anisotropy tests did not detect significant directional trends, suggesting that any apparent directional variability in semivariance is likely an artefact of the reduced sample size (N = 52) rather than a true geostatistical feature.

Overall, these semivariogram results highlight the limited spatial structure of SOM at the sampled resolution and emphasize the need for higher-resolution or denser sampling in future studies to capture fine-scale heterogeneity.

Kriging prediction accuracy

Following the LOOCV procedure described in the Ordinary Kriging and Cross-validation subsection, each of the 52 samples was withheld in turn, kriged from the remaining 51 points using the fitted Gaussian model, and the predicted value compared to the observed value. This procedure was executed using the krige.cv() function in the gstat package (Pebesma, 2004) in R 4.3.0. Cross-validation statistics are reported in Table 4 and the predicted vs. measured plot is presented in Figure 6.

The ME was virtually zero, confirming negligible systematic bias. The RMSE (0.971 g/kg) represents 42.7% of the observed SOM range (4.099 g/kg), reflecting the substantial nugget effect identified in the Semivariogram Analysis subsection. The low R^2 (0.094) reflects the dominance of the nugget effect (64.2% of total variance), indicating that a substantial proportion of SOM variability occurs at spatial scales below the 1 × 1 km sampling grid resolution. This short-range variability is not a model failure but rather an inherent characteristic of SOM distribution in heterogeneous agroecosystems, where field-scale management decisions — tillage patterns, organic amendments, micro-topographic redistribution — operate at scales finer than the sampling interval (Webster and Oliver, 2007). These findings point to the value of denser sampling designs in future studies. The prediction error distribution was approximately symmetrical around zero (Figure 6), confirming absence of directional bias.

Spatial distribution of SOM

The ordinary kriging prediction map (Figure 7) was generated at 100 m resolution using the fitted Gaussian semivariogram model in ArcGIS 10.4 Geostatistical Analyst. The map suggests the presence of three broad SOM zones, interpreted

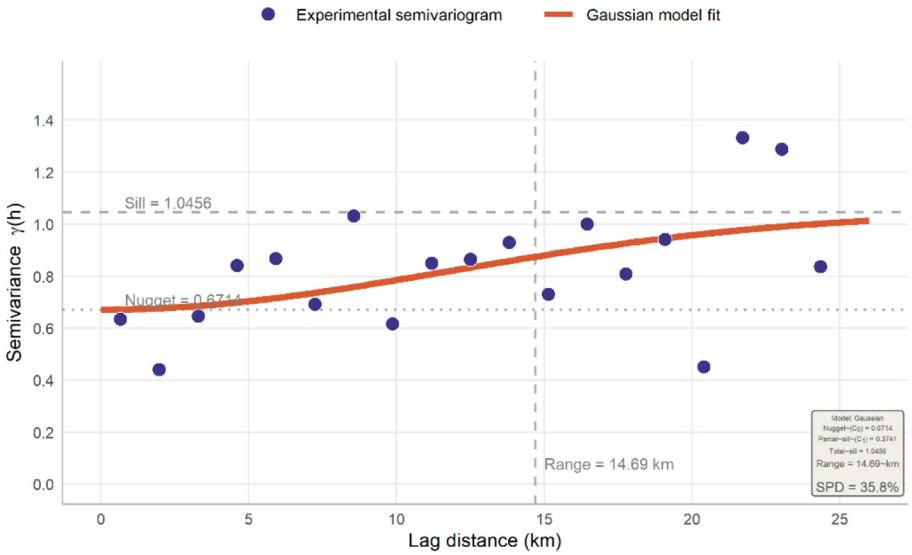

Figure 5. Experimental semivariogram (filled circles) and fitted Gaussian theoretical model (solid line) for soil organic matter (SOM) in the El Maleh sub-watershed (N = 52). Model parameters: nugget C_0 = 0.671, partial sill C_1 = 0.374, range = 14.69 km, SPD = 35.8%

Table 3. Comparative performance of four theoretical semivariogram models fitted to the experimental SOM data, assessed by residual sum of squares (RSS) and leave-one-out cross-validation (LOOCV) statistics. The retained model is highlighted in bold

Model	Nugget C_0	Part. sill C_1	Total sill	Range (m)	SPD (%)	RSS	RMSE_CV	R^2_CV
Spherical	0.618	0.520	1.139	42,893	45.7	0.742	0.973	0.078
Exponential	0.574	0.542	1.116	16,952	48.6	0.737	0.977	0.043
Gaussian	**0.671**	**0.374**	**1.046**	**14,694**	**35.8**	**0.763**	**0.971**	**0.094**
Linear	0.629	0.805	1.433	49,185	56.2	0.742	0.971	0.093

Note: SPD = spatial dependence index = $C_1/(C_0+C_1) \times 100$ (Cambardella et al., 1994); RMSE_CV = root mean square error from LOOCV; R^2_CV = Pearson R^2 from LOOCV. Bold = retained model.

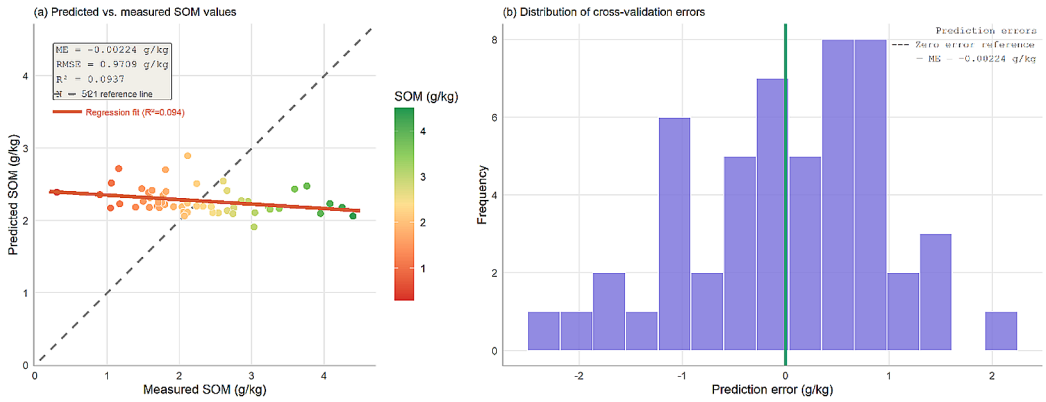

Figure 6. Leave-one-out cross-validation results for the Gaussian ordinary kriging model: (a) predicted vs. measured SOM values (dashed line = 1:1 reference; solid line = regression fit); (b) distribution of prediction errors. ME = −0.002 g/kg, RMSE = 0.971 g/kg, N = 52

Table 4. Leave-one-out cross-validation statistics for the Gaussian ordinary kriging model (N = 52).

Statistic	Value	Interpretation
Mean Error (ME)	−0.000002 g/kg	Negligible systematic bias
RMSE	0.971 g/kg	42.7% of observed SOM range
Pearson r	−0.306	Weak directional agreement
R^2 (LOOCV)	0.094	Reflects dominant nugget effect

Note: ME = mean error; RMSE = root mean square error.

with caution given the moderate prediction accuracy (RMSE = 0.971 g/kg, R^2_CV = 0.094). These zones represent indicative spatial trends rather than precise delineations, and field verification is recommended before operational use. SOM-depleted areas (< 2.0 g/kg) were concentrated in the northern and steep-slope sectors, associated with active erosion processes and sparse vegetation. Moderately endowed zones (2.0–2.5 g/kg) dominated the central agricultural plains, representing the most spatially extensive class. Locally elevated SOM concentrations (> 2.5 g/kg) were observed in valley bottoms and areas with denser vegetation cover, where reduced erosion and greater biomass inputs favor organic matter accumulation.

Statistical relationships between SOM, slope, and NDVI

All statistical analyses described in this section were performed in R 4.3.0 using base R functions and the FactoMineR 2.9 package, following the procedures described in the Geostatistical and Statistical Analysis section. Slope values were extracted for 31 of the 52 sampling points from the SRTM-derived slope raster reprojected to UTM Zone 32N; the remaining 21 points fell outside the raster extent and were excluded from slope-based analyses. NDVI values were available for all 52 points. Correlation analysis. Monotonic associations between SOM and each covariate were

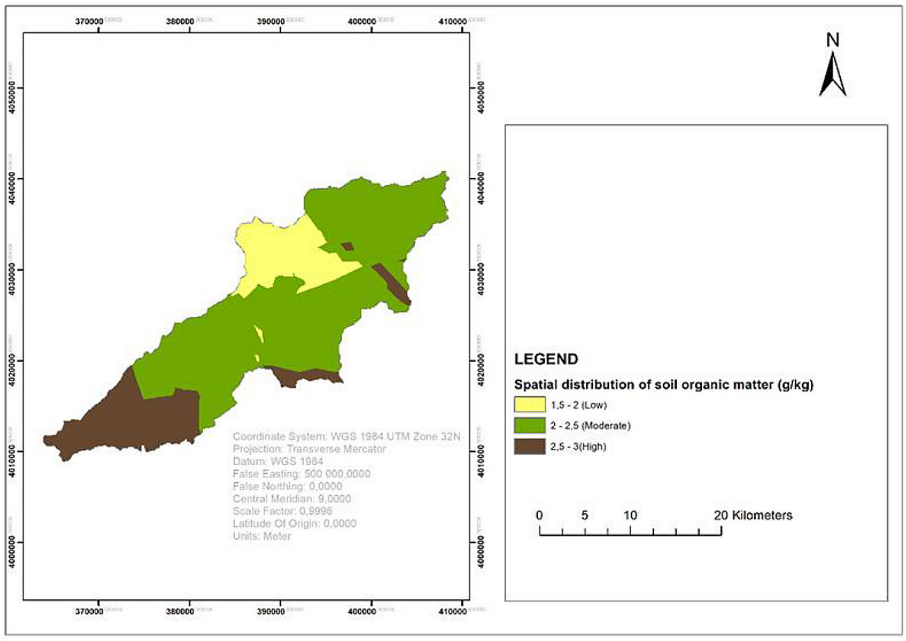

Figure 7. Spatial distribution of soil organic matter (SOM, g/kg) in the El Maleh sub-watershed predicted by ordinary kriging (Gaussian model, 100 m resolution). Three SOM zones are delineated: low (1.5–2.0 g/kg), moderate (2.0–2.5 g/kg), and high (2.5–3.0 g/kg). Coordinate system: WGS 1984 UTM Zone 32N

assessed using Spearman's rank correlation (cor. test(), method = "spearman") and Pearson's r (cor. test(), method = "pearson") in base R. Results are presented in Table 5.

Multiple linear regression

The joint effect of slope and NDVI on SOM was assessed through OLS regression using the lm() function in R, with standard errors, t-statistics, and p-values extracted from summary(lm()). Results are presented in Table 6.

Neither slope nor NDVI was a significant predictor of SOM. The model explained only 7.4% of SOM variance ($R^2 = 0.074$, adjusted $R^2 = 0.021$).

Principal component analysis. PCA was performed on the standardized matrix of SOM, slope, and NDVI (N = 31 for points with valid slope data) using the PCA() function from the FactoMineR package 2.9 (Lê et al., 2008) in R 4.3.0. Results are presented in Table 7 and Figure 8. The first two components explained 87.8% of total variance (PC1 = 58.9%, PC2 = 28.9%). PC1 was primarily structured by slope (loading = 0.835) and NDVI (loading = 0.881), while SOM showed a weaker association with PC1 (loading = 0.326). PC2 was almost exclusively driven by SOM (loading = 0.945), with slope (loading = −0.327) and NDVI (loading = −0.040) contributing minimally. This structure demonstrates that SOM varies largely independently of the slope–NDVI environmental gradient, as confirmed by the near-orthogonal orientation of the SOM vector relative to slope and NDVI vectors in the biplot (Figure 8). This result provides multivariate confirmation that SOM distribution in the El Maleh sub-watershed is governed by factors independent of the two environmental covariates tested.

Table 5. Spearman rank correlation and Pearson correlation coefficients between SOM and environmental covariates (slope and NDVI)

Variable pair	Spearman ρ	p-value	Pearson r	p-value	Interpretation
SOM vs Slope (%, N=38)	+0.186	0.264	+0.027	0.848	Not significant
SOM vs NDVI (N=52)	−0.026	0.858	+0.177	0.208	Not significant
Slope vs NDVI (N=38)	+0.381	0.018	—	—	Significant

Note: slope analyses based on N = 31 points with valid slope data; NDVI analyses based on N = 52. ns = not significant ($p > 0.05$).

Table 6. Ordinary least squares (OLS) multiple regression results for SOM as a function of slope and NDVI (N = 31)

Predictor	β	SE	t-statistic	p-value	Significance
Intercept	1.843	0.268	6.876	< 0.001	***
Slope (%)	+0.005	0.017	0.177	0.860	ns
NDVI	+3.755	2.934	1.280	0.209	ns
R^2 = 0.074; Adjusted R^2 = 0.021					

Note: SE = standard error. Significance codes: *** p < 0.001; ns = not significant (p > 0.05).

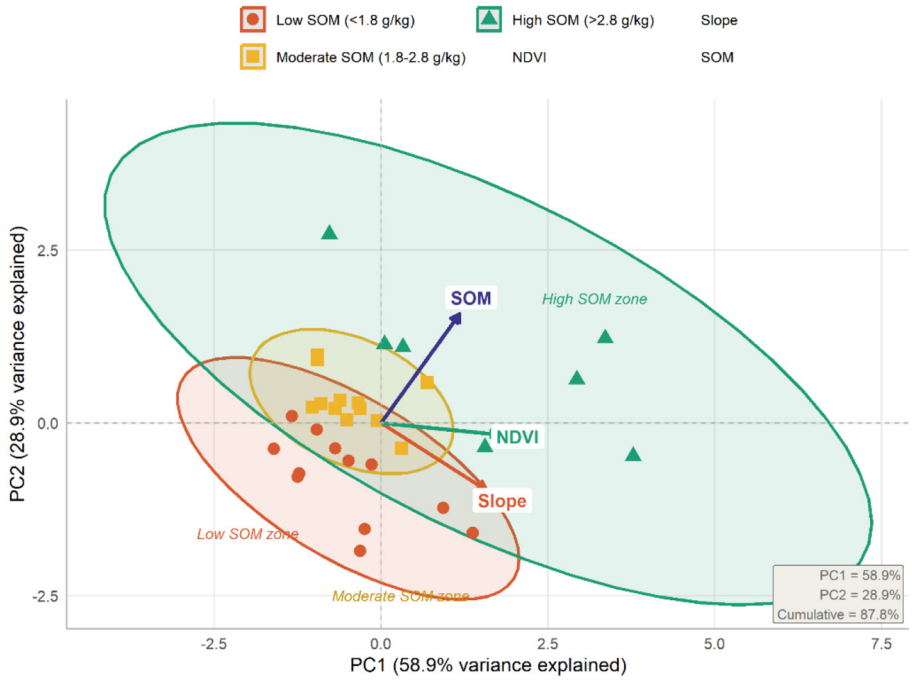

Figure 8. Principal component analysis (PCA) biplot for SOM, slope (%), and NDVI (N = 31 points with valid slope data). Points represent sampling locations grouped by SOM class: low (< 1.8 g/kg), moderate (1.8–2.8 g/kg), and high (> 2.8 g/kg). Ellipses represent 95% confidence regions for each group. Arrows represent variable loadings. PC1 explains 58.9% and PC2 explains 28.9% of total variance (cumulative = 87.8%)

Table 7. Principal component analysis (PCA) results for the standardized matrix of SOM, slope, and NDVI (N = 31 points with valid slope data). Loadings represent the correlation of each variable with each principal component

Component	Eigenvalue	Variance (%)	Cumulative (%)	SOM	Slope	NDVI
PC1	1.768	58.9	58.9	0.326	0.835	0.881
PC2	0.868	28.9	87.8	0.945	−0.327	−0.040
PC3	0.364	12.2	100.0	−0.140	−0.464	0.492

Note: SOM = soil organic matter; NDVI = Normalized Difference Vegetation Index. PC = principal component.

DISCUSSION

SOM levels in a regional and Mediterranean context

The mean SOM content recorded in the El Maleh sub-watershed (2.274 ± 0.908 g/kg, N = 52) is consistent with values reported in comparable semi-arid catchments of northeastern Algeria. In the adjacent Rhirane catchment, Othmani et al. (2023) reported mean SOM of 3.4% in Chromic Cambisols and 2.2% in Calcic Cambisols, with a CV of 54.01% — higher than the 40.0% observed here, likely reflecting the more heterogeneous land-cover mosaic of the Rhirane basin. More broadly, the SOM levels recorded fall within the lower range typical of Mediterranean semi-arid agroecosystems (1–5 g/kg), consistent with patterns reported

for comparable *environments* across North Africa and the Mediterranean Basin (Lal, 2004; Wiesmeier et al., 2019; Fan et al., 2020). For reference, Wiesmeier et al. (2019) reported mean topsoil SOC stocks of 20–40 t/ha for Mediterranean dryland agroecosystems, and Fan et al. (2020) documented SOM values of 1.5–4.0% in semi-arid Chinese catchments with similar climate and land-use characteristics, supporting the representativeness of the values obtained in the present study. The slightly right-skewed distribution (skewness = 0.467, Pearson kurtosis = 2.806) is a recurrent feature in heterogeneous landscapes (Cambardella et al., 1994). This pattern may reflect the influence of spatially variable organic inputs associated with differences in land use, vegetation cover, or erosion intensity across the watershed; however, as field management data were not collected in the present study, these factors remain hypothetical drivers rather than confirmed explanations.

Spatial structure and geostatistical performance

The Gaussian semivariogram model provided the best overall fit based on leave-one-out cross-validation (RMSE = 0.971 g/kg, R^2 = 0.094), with an estimated range of 14.69 km and a spatial dependence index (SPD) of 35.8%, indicating moderate spatial dependence (Cambardella et al., 1994). These parameters describe the fitted geostatistical structure of SOM without implying process causality.

The nugget-to-sill ratio (64.2%) indicates that a substantial proportion of total variance occurs at spatial scales smaller than the 1 × 1 km sampling interval. This reflects weakly resolved fine-scale variability in the dataset and implies a limited ability of the kriging model to resolve short-range spatial structure. Consequently, predictive accuracy is expected to decrease at local scales, particularly where abrupt changes in land management or micro-topography occur.

From a modelling perspective, this weak spatial structure constrains the precision of spatial prediction, especially in areas where observations are sparse or spatial gradients are steep.

Directional semivariograms were computed in four directions (0°, 45°, 90°, 135°) with a tolerance angle of 30° to assess potential anisotropy (Figure 9). The observed directional variability (anisotropy ratio = 11.99; range CV = 78.2%) is primarily attributed to unequal and limited pair counts per directional class, which is a known limitation in anisotropy assessment for small datasets (N = 52; Webster and Oliver, 2007). No

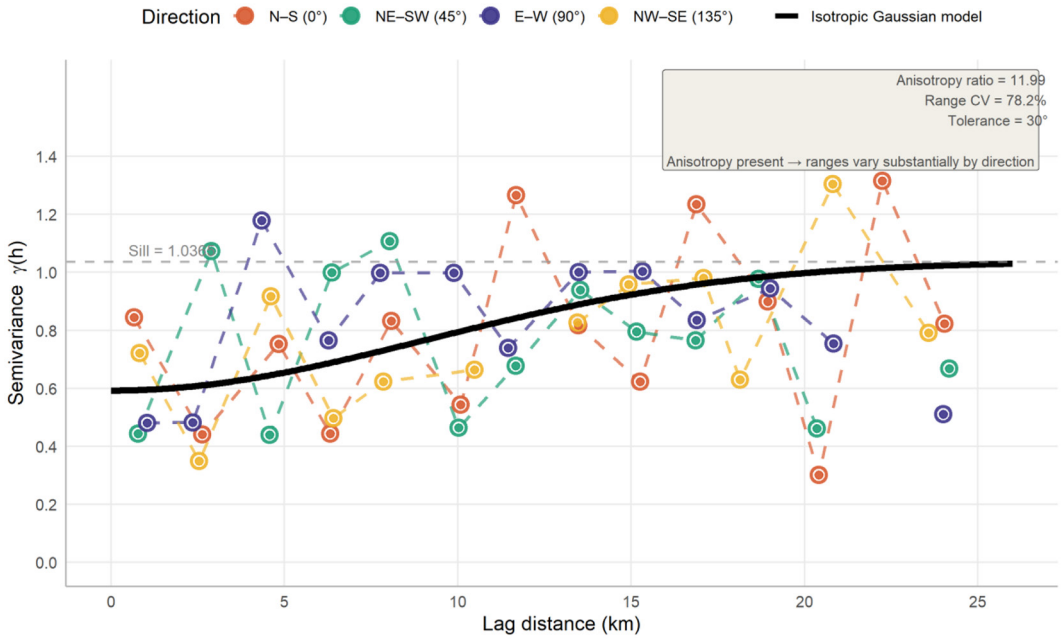

Figure 9. Directional semivariograms computed in four directions (0°, 45°, 90°, 135°; tolerance angle = 30°) for soil organic matter (SOM) in the El Maleh sub-watershed (N = 52). The solid black line represents the isotropic Gaussian model retained for ordinary kriging. The high variability across directions (anisotropy ratio = 11.99, range CV = 78.2%) reflects estimation uncertainty due to limited sample pairs per directional class rather than genuine geometric anisotropy

statistically robust evidence of geometric anisotropy was therefore retained, and an isotropic Gaussian model was selected as the most parsimonious representation of spatial dependence.

Edge effects were observed at higher lag distances, particularly beyond approximately 20 km, where the number of contributing point pairs decreases substantially. This leads to increased uncertainty and greater scatter in empirical semivariance estimates near the upper limit of the lag distance. Given that the effective range (14.69 km) corresponds to approximately 46% of the watershed length (32 km), predictions near watershed boundaries should be interpreted with caution, as edge-related instability may influence extreme interpolated values.

Overall, the combination of high nugget effect, weak directional structure, and edge-induced variability indicates that SOM spatial structure is only partially resolved at the current sampling density, limiting the precision of fine-scale spatial prediction.

Slope and NDVI as SOM predictors

Neither slope nor NDVI showed a statistically significant relationship with SOM at the watershed scale. Spearman correlations were weak and non-significant ($\rho = 0.186$, $p = 0.264$ for slope; $\rho = -0.026$, $p = 0.858$ for NDVI), and the multiple linear regression model explained a limited proportion of variance in SOM ($R^2 = 0.074$).

Principal component analysis further supported the weak coupling between SOM and the examined environmental covariates. PC1, explaining 58.9% of total variance, was mainly structured by slope (loading = 0.835) and NDVI (loading = 0.881), whereas PC2, accounting for 28.9% of variance, was dominated by SOM (loading = 0.945). The separation of SOM from the slope–NDVI gradient in the PCA space indicates limited multivariate association at the sampling resolution used in this study.

This pattern differs from findings in other Mediterranean environments where topographic position is reported as a primary control on SOM distribution (Moore et al., 1993; Fan et al., 2020). The divergence may be explained by a combination of methodological and environmental factors.

First, the relatively coarse 1 × 1 km sampling grid likely smooths local-scale topographic variability, thereby attenuating potential slope–SOM relationships that may operate at finer spatial scales. This scale effect is important, as studies with denser sampling designs (e.g., Othmani et al., 2023; 132 points over 80.82 km²) were better able to resolve topographic controls, suggesting that sampling resolution strongly conditions detectability of slope effects.

Second, land management practices (e.g., tillage, organic amendments, and crop rotation) may contribute to SOM redistribution; however, these effects should be interpreted cautiously here, as they are inferred rather than directly quantified in the present dataset.

Third, lithological differences across the study area may also contribute to spatial variability in SOM, although their role cannot be explicitly isolated within the current modelling framework due to the absence of soil unit-specific parameterization.

Regarding NDVI, its lack of statistical significance ($\rho = -0.026$, $p = 0.858$) is consistent with both the temporal mismatch between field sampling (spring 2023) and satellite acquisition (26 November 2023) and the use of a single-date image, which captures instantaneous vegetation conditions rather than cumulative productivity. Future studies should incorporate multi-temporal NDVI composites to better represent long-term vegetation dynamics and their relationship with SOM (Wiesmeier et al., 2019).

Spatial distribution and management implications

The ordinary kriging prediction maps revealed three broad SOM zones across the El Maleh sub-watershed. SOM-depleted areas (< 2.0 g/kg) were concentrated in the northern and steep-slope sectors, where active erosion processes and sparse vegetation limit organic matter inputs and promote SOM redistribution. Moderately endowed zones (2.0–2.5 g/kg) dominated the central agricultural plains under mixed land use, representing the most spatially extensive class. Locally elevated SOM concentrations (> 2.5 g/kg) were observed in valley bottoms and areas with denser vegetation cover, where reduced erosion and greater biomass inputs favor organic matter accumulation.

These spatial patterns broadly reflect the watershed geomorphology, despite the absence of a significant point-scale slope–SOM relationship — underscoring the limitation of correlation-based approaches in capturing landscape-scale

organization. SOM-depleted zones provide a direct basis for prioritizing organic amendment programs and erosion control, while enriched valley bottom areas should be protected from further intensification to preserve existing carbon stocks.

CONCLUSIONS

This study provides the first spatially explicit assessment of SOM in the El Maleh sub-watershed using field measurements, Walkley–Black analysis, and geostatistical modeling via ordinary kriging. SOM exhibited marked spatial heterogeneity (mean = 2.274 ± 0.908 g/kg; CV = 40.0%). Hypothesis 1, which predicted high variability (CV > 30%), was supported. In contrast, Hypotheses 2 and 3, proposing slope and NDVI as primary predictors of SOM distribution, were not supported at the watershed scale with the current 1 × 1 km sampling grid. Principal component analysis confirmed that SOM variability is largely orthogonal to the slope–NDVI environmental gradient, suggesting that unmeasured factors—potentially land management practices—may influence fine-scale SOM patterns.

The Gaussian semivariogram model provided the best fit (nugget = 0.671, range = 14.69 km, SPD = 35.8%), while the high nugget-to-sill ratio (64.2%) indicates that a substantial portion of SOM variability occurs below the sampling resolution. This finding highlights the limitation of coarse grids for capturing fine-scale heterogeneity and provides new insight into the spatial structure of SOM in semi-arid Mediterranean agroecosystems. Edge effects observed at lag distances >20 km further caution against overinterpreting predictions near watershed boundaries. The generated kriging maps offer an indicative trend of SOM distribution, serving as a preliminary baseline for soil conservation planning, but should not be interpreted as exact SOM values.

This work fills a critical gap in regional soil studies by providing the first spatially explicit SOM dataset for the El Maleh sub-watershed, quantifying fine-scale variability, and establishing a methodological framework for geostatistical soil assessment in data-limited semi-arid regions.

REFERENCES

1. Brady, N. C., Weil, R. R. (2017). *The nature and properties of soils* (15th ed.). Pearson.
2. Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F., Konopka, A. E. (1994). Field-scale variability of soil properties in Central Iowa soils. *Soil Science Society of America Journal*, *58*(5), 1501–1511.
3. Fan, M., Lal, R., Zhang, H., Margenot, A. J., Wu, J., Wu, P., Zhang, L., Yao, J., Chen, F., Gao, C. (2020). Variability and determinants of soil organic matter under different land uses and soil types. *Soil and Tillage Research*, *198*, 104544.
4. FAO. (2019). *Standard operating procedure for soil organic carbon: Walkley-Black method*. Food and Agriculture Organization of the United Nations.
5. Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science*, *304*(5677), 1623–1627.
6. Lê, S., Josse, J., Husson, F. (2008). FactoMineR: A package for multivariate analysis. *Journal of Statistical Software, 25*(1), 1–18.
7. Moore, I. D., Gessler, P. E., Nielsen, G. A., Peterson, G. A. (1993). Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal*, *57*(2), 443–452.
8. Othmani, O., Khanchoul, K., Boubehziz, S., Bouguerra, H., Benslama, A., Navarro-Pedreño, J. (2023). Spatial variability of soil erodibility at the Rhirane catchment using geostatistical analysis. *Soil Systems*, *7*(2), 32. https://doi.org/10.3390/soilsystems7020032
9. Pebesma, E. J. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, *30*(7), 683–691. https://doi.org/10.1016/j.cageo.2004.03.012
10. Pribyl, D. W. (2010). A critical review of the conventional SOC to SOM conversion factor. *Geoderma*, *156*(3–4), 75–83.
11. R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/
12. Walkley, A., Black, I. A. (1934). An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Science*, *37*(1), 29–38.
13. Webster, R., Oliver, M. A. (2007). *Geostatistics for environmental scientists* (2nd ed.). John Wiley & Sons.
14. Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H. J., Kögel-Knabner, I. (2019). Soil organic carbon storage as a key function of soils — A review of drivers and indicators at various scales. *Geoderma, 333*, 149–162.
15. Wilding, L. P. (1985). Spatial variability: its documentation, accommodation and implication to soil surveys. In D. R. Nielsen & J. Bouma (Eds.), *Soil Spatial Variability* (pp. 166–194). Pudoc.