SOIL SCIENCE ANNUAL
2025, 76(3), 207765

https://doi.org/10.37501/soilsa/207765

High-resolution baseline digital mapping of soil fertility in the Euphrates basin, western Iraq

Jamal Abed Hammad[1], Soumaia M'nassri*[1], Balkis Chaabane[2], Ali-Hussein Ibraheem Al Bayati[3], Rajouene Majdoub[1]

[1] Laboratory of Research in Management and Control of Animal and Environmental Resources in Semi-aride Ecosystem, Higher Agronomic Institute of Chott Meriem, University of Sousse, BP 42,4042 Chott Meriem, Sousse, Tunisia
[2] Department of Horticultural Sciences and Landscape, Higher Agronomic Institute of Chott Meriem, University of Sousse, BP 42,4042 Chott Meriem, Sousse, Tunisia
[3] College of Agriculture, University of Al-Anbar, BP 55431 Baghdad, 55 Ramadi, Anbar, Iraq

* Corresponding author: Soumaia M'nassri, mnassrisoumaia@gmail.com, ORCID iD:https://orcid.org/0009-0006-4249-1145

Received: 2024-09-15
Accepted: 2025-07-01
Published online: 2025-07-01
Associated editor: Agota Horel

Keywords:

Digital soil mapping
Soil fertility
Multiple linear regression
Remote sensing data
Euphrates basin

Abstract

Digital Soil Mapping is a vital tool used to produce outputs for Digital Soil Assessment, facilitating analyses and recommendations for various environmental practices. Despite its significance, DSM has not been widely applied to soil fertility assessment due to the challenges associated with integrating multiple soil properties into a single map. Therefore, this study aimed to develop a new digital soil fertility map using higher-resolution data. The research was conducted in the Al-Anbar district, located in the Euphrates basin in western Iraq. Spectral indices and signature data from Sentinel-2 were utilized for soil sampling, supplemented by local ground measurements. Principal Component Analysis (PCA) was then employed to identify the most relevant soil fertility indicators. Subsequently, a digital soil map was generated using a Multiple Linear Regression (MLR) model. The results indicated that soil fertility was significantly represented by soil total nitrogen and phosphorus content. Moreover, the MLR model for soil fertility included soil moisture (SM), brightness index (BI), colour index (VI), green chlorophyll index (CIgreen), and B8A. Our approach demonstrated the potential of remote sensing and multiple linear modelling for soil fertility mapping, with an RMSE, R_c, and MAE of 1.11, 0.98, and 0.68, respectively. These findings suggest that integrating remote sensing with multiple linear regression modelling provides an effective method for accurately estimating soil properties over large areas, offering considerable benefits in terms of cost, efficiency, coverage, and the ability to monitor changes over time.

1. Introduction

Soil is a dynamic and integral ecosystem component, supporting various ecological processes and functions required to sustain life on Earth. Protecting and conserving soil quality is critical for maintaining ecosystem resilience and biodiversity, as well as ensuring food security and environmental sustainability (Zhou et al., 2020; Poppiel et al., 2021). The potential soil fertility assessment is the key element in monitoring, improving, and conserving land productivity (Song et al., 2020). Hence, macronutrient availability, soil organic matter, salinity pH and texture are important attributes of soil quality and fertility, and they are closely related to crop growth (Ozayzici et al., 2017; Kumar et al., 2018). Therefore, rapid and precise estimation techniques, as well as a thorough understanding of the spatial variability of these indicators, are imperative to maintain food security (Xue et al., 2023). However, conventional soil analysis and mapping methodologies, which rely on ground-based surveys, are labor-intensive and time-consuming (Srisomkiew et al., 2022). Thus, achieving precise estimations of soil properties through resilient and economical approaches becomes imperative (Yuzugullu et al., 2024).

In this realm, remote sensing emerges as a promising and cost-effective approach for rapidly acquiring relevant data regarding soil properties, which permits useful insights for policy development (Kasampalis et al., 2018; Yuan et al., 2024). One notable application of this approach in the agro-industrial system is the deployment of precise digital soil mapping (DSM) through spectral remote sensing technology (Bao et al., 2024). This advanced approach enables the acquisition of highly accurate estimates of soil properties. Consequently, it provides evidence-based insights critical for developing effective soil management strategies (Arrouays et al., 2020). For instance, Hounkpatin et al. (2018) used Digital Soil Mapping (DSM) to assess soil organic carbon stocks across different land use ecosystems and soil groups in Burkina Faso's

savannah zone. Dasgupt et al. (2023) developed a regional soil micronutrient management strategy that used DSM to improve agronomic biofertilization practices, proving the potential efficacy of the methodology. Zhang et al. (2024) applied DSM projections to investigate soil carbon dynamics at both spatial and temporal scales. Nonetheless, certain challenges exist in accurately assessing carbon dynamics using this methodology, owing to differences in the scales of influential variables over time. Their findings deem for vigilant adoption of DSM projections. Moreover, DSM has been used in further studies focusing on various aspects of soil quality and health, such as soil pH and carbonates (Lu et al., 2023), soil erodibility (Sun et al., 2024), and various soil biological properties including urease, acid phosphatase, alkaline phosphatase enzymes, metabolic quotient, and soil microbial biomass (Esmaeilizad et al., 2024). However, DSM based on remote sensing spectral indices relies heavily on statistical models combined with machine learning techniques. (Chlingaryan et al., 2018; Liakos et al., 2018; Wang et al., 2019; Du et al., 2020). Multiple Linear Regression (MLR) is widely used among these techniques due to its easily understood computational process and clear outcomes. Several studies have highlighted the potential of MLR, emphasizing its comprehensible methodology (Guo et al., 2015; Mondejar and Tongco, 2021; Srisomkiew et al., 2021; Dvornikov et al., 2022).

As seen from the literature review, previous studies have demonstrated the efficacy of DSM methodologies that use high-resolution satellite imagery combined with Multiple Linear Regression (MLR) techniques for delineating soil physical and chemical properties, achieving satisfactory results. However, the application of this approach to assess soil fertility has been rarely reported. In previous research in the El Euphrates basin (El Hammdane et al., 2024), we mostly concentrated on soil fertility assessment using conventional soil analysis methods. Despite these efforts, there is an intriguing imperative to advance towards a digital soil mapping framework, given the pressing demand for comprehensive soil property information to address agricultural productivity challenges and ensure food security. Consequently, our research aims to develop a DSM of soil fertility using remote sensing datasets. Therefore, the objectives of this study are: (1) to compute several spectral indices based on Sentinel-2 bands, (2) to use principal component analysis to identify an appropriate soil fertility indicator, and (3) to create a digital soil fertility map using an MLR model.

2. Materials and methods

2.1. Study area

Al-Anbar region, located in the Euphrates basin in the western Iraq was selected as the study area for this study. The geographical range spans from 43°10'0" to 43°34'0"E and from 33°28'0" to 33°28'30"N (Fig. 1), covering an area of approximately 18.16 ha. The study area has a predominately arid climate, with hot summers and mild winters. The annual total rainfall is 110 mm. The temperature ranged from 13.9 and 30.1°C. The dominant soil type in this area ranges from moderately coarse, typically including sandy loam, to moderately fine, such as silty clay loam and clay loam soils. The predominant vegetation comprises corn (*Zea mays)*, alfalfa (*Medicago sativa*) and clover (*Trifolium spp*).

2.2. Dataset

2.2.1. Soil data

In April 2022, 50 soil samples were systematically collected from the chosen study area. Each sample was carefully extracted from a depth of 0 to 30 cm. The collected samples were then thoroughly air-dried and sieved before the analysis. The analytical protocol included a thorough analysis of physical and chemical parameters such as soil texture, soil organic matter (SOM) content, calcium carbonate concentration ($CaCO_3$), total nitrogen (N_T), available phosphorus (P) content, exchangeable potassium (K) levels, pH, and electrical conductivity (EC). Furthermore, the Soil Fertility Index (SFI) was calculated using Equation 1, following the steps within the methodology explained by Abed

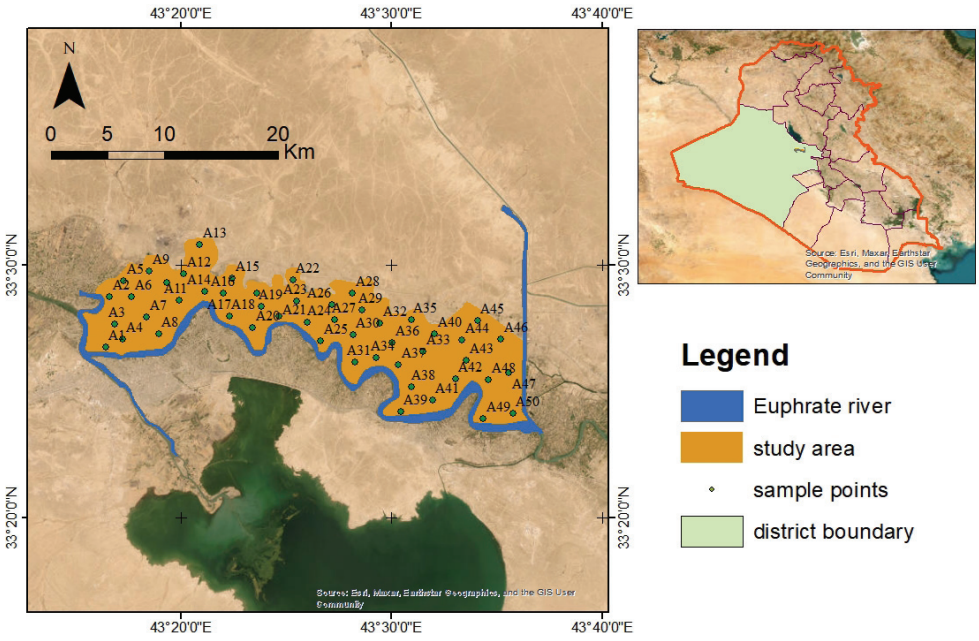

Fig. 1. Location of the study area

Hammad et al. (2024). A summary of the physicochemical analyses carried out following internationally recognized soil analysis methodologies is shown in Table 1.

$$SFI = [R_{max} \times \sqrt{\frac{A}{100} \times \frac{B}{100} \times \frac{C}{100} \ldots \ldots .]} \times 100 \quad \text{Eq. 1}$$

R_{max} is the maximum ratio of (A+B+C+.....N)/8, and A, B, C.....N are the rating values for each diagnostic parameter (N_{tot}, P_{ava}, K_{ava}, pH, $CaCO_3$, EC, SOM and texture).

2.2.2. Remote sensing data

A cloud-free Sentinel-2 image, processed at Level-1C, was obtained from the USGS Earth Explorer website and selected to closely match the date of field sampling. Level-1C processing includes geometric and radiometric corrections. Following that, atmospheric correction was carried out using the Semi-Automatic Classification tool plugin (Congedo, 2016) in a geographic information system software QGIS version 3.16 (QGIS Development Team, 2020). To compensate for the limited spatial coverage and cloud cross-contamination inherent in individual images, a mosaicking procedure was used to combine two Sentinel-2 images. The cloud cover across all satellite images used in this study was consistently less than 10%. Furthermore, our analysis included 10 Sentinel-2 bands (Gholizadeh et al., 2018; Grinand et al., 2017), as well as 18 spectral indices commonly employed in previous research for exploring soil properties (Wei et al., 2024; Geng et al., 2024; Diaz-Gonzalez et al., 2022; Khanal et al., 2018).

The spectral indices utilized in this study encompass a comprehensive suite of metrics including the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Transformed Vegetation Index (TVI), Soil Adjusted Total Vegetation Index (SATVI), Soil Adjusted Vegetation Index (SAVI), Moisture Stress Index (MSI), Green Normalized Difference Index (GNDI), Brightness Index (BI), Color Index (CI), Green Chlorophyll Index (CIgreen), Atmospherically Resistance Vegetation Index (ARVI), Saturated Index (S), Soil Organic Carbon Index (SOCI), Canopy Response Salinity Index (CRSI), Salinity Index (S1), Salinity Index (S2), Clay Index (ClI), and Carbonate Index (CarI). Comprehensive details regarding the spectral band coverage and corresponding spatial resolutions of the Sentinel-2 sensor employed in this study are provided in Table 2, while the formulas used to derive the spectral indices are highlighted in Table 3.

Table 1
Analytical methods used for the analysis of soil physical and chemical parameter

Physical and chemical parameters	Analytical method	Reference
Texture	International Pipette	(Wang et al., 2019)
SOM	Walkley-Black	(Nelson et al., 1999)
$CaCO_3$	Titration	(Soil Survey Staff, 1992)
N_T	Kjeldahl	(Tunçay et al., 2023)
P	Colorimetric method	(Soil Survey Staff, 1999)
K	Spectrophotometer flame	(Zhang et al., 2016)
pH	Potentiometric	(Huang et al., 2010)
EC	Potentiometric	(Huang et al., 2010)

Table 2
Details of Sentinel-2 spectral bands

Spectral band number	Spectral band description	Spectral range (nm)	Spectral position (nm)	Bandwidth (nm)	Spectral resolution (m)
B2	B	458–523	490	65	10
B3	G	543–578	560	35	10
B4	R	650–680	665	30	10
B5	RE1	698–713	705	15	20
B6	RE2	733–748	740	15	20
B7	RE3	773–793	783	20	20
B8	NIR	785–900	842	115	10
B8A	RE4	855–875	865	20	20
B11	SWIR1	1565–1655	1610	90	20
B12	SWIR2	2100–2280	2190	180	20

(ESA, 2010)
Explanations: B: blue; G: green; R: red; RE: red-edge; NIR: near infrared;. SWIR: shortwave infrared.

Table 3
Details of derived indices

Index	Formulation	Reference
NDVI	$\frac{B8-B4}{B8+B4}$	Rouse et al., 1973
EVI	$G \times \frac{B8-B4}{B8+C1\times B4-C2\times B2+L}$	Buete et al., 2002
TVI	$(\frac{B8-B4}{B8+B4}+0.5)^{0.5}$	Nellis and Briggs 1992
SATVI	$\frac{B8-B4}{B8+B4+L\times B2}$	Marsett et al., 2006
SAVI	$\frac{(B8-B4)\times(1+L1)}{B8-B4+L1}$	Huete et al., 1988
SMI	$\frac{B11}{B8}$	Rock et al., 1985
GNDI	$\frac{B8-B3}{B8+B3}$	Gitelson et al., 1996
BI	$\frac{\sqrt{(B4\times B4)+(B3\times B3)+}}{2}$	Escadafal, 1989
CI	$\frac{B4-B3}{B4+B3}$	Ray et al., 2004
GI_{green}	$(\frac{B8}{B3})-1$	Nguy-Robertson et al., 2014
ARVI	$\frac{B8-2\times B4+B2}{B8+2\times B4-B2}$	Kaufman and Tanre, 1992
S	$\frac{B8}{B4}$	Ray et al., 2004
SOCI	$\frac{B2}{B4\times B3}$	Thaler et al., 2019
CRSI	$[\frac{(B8\times B4)-(B3\times B2)}{(B8\times B4)+(B3\times B2)}]^{0.5}$	Scudiero et al., 2014
S1	$(B3+B2)^{0.5}$	Douaoui et al., 2006
S2	$\frac{B4}{B8\times 100}$	Tripathi et al., 1997
ClI	$\frac{B11}{B12}$	Taghizadeh-Mehrjardi et al., 2014)
CarI	$\frac{B3}{B2}$	Taghizadeh-Mehrjardi et al., 2014)

Explanations: G= 2.5; C1=6; C2=7.5; L=1; L1=0.5.

2.2.3. Multivariate analysis

To establish relationships between spectral bands, indices, and indicators of soil properties, a Principal Component Analysis (PCA) was used in order to identify representative indicators for Digital Soil Mapping (DSM) through model equations. In this study, ten soil physico-chemical properties, such as silt, clay, sand, SOM, $CaCO_3$, N_t, P, K, pH, EC and SFI (Table 4), and twenty seven variables, including spectral bands and indices derived from remotely sensed satellite data were applied to the PCA using IBM Statistical Package for the Social Sciences (SPSS, version 2020). Extracted PCA was evaluated by factor loadings and eigenvalues. In fact, the total numbers of factors generated from the factor analysis that have an eigenvalue greater than 1 (M'nassri et al., 2019). After that, a Pearson matrix correlation was used to measure the linear relationship between soil indicators and environmental variables. Based on the obtained results, variables with significant correlations were chosen as critical predictors for developing a predictive model of soil properties using MLR approach.

Table 4
Summary statistics of soil physico-chemical properties (n=50)

Indicators	Minimum	Maximum	Mean	Standard deviation
Clay (%)	15.51	43.32	30.32	9.86
Silt (%)	41.62	74.40	54.32	8.81
Sand (%)	9.00	41.50	15.70	9.71
pH [-]	7.3	8.00	7.68	0.19
EC (dS/m)	4.20	52.30	16.01	12.31
CaCO3 (%)	1.71	2.25	1.95	1.25
SOM (%)	0.60	1.40	0.91	0.22
N_{tot} (%)	0.01	0.43	0.09	0.07
P_{av} (mg/kg)	2.50	15.60	6.12	4.18
K_{av} (mg/kg)	188.00	372.00	248.68	53.70
SFI	0.55	12.07	3.62	1.68

2.2.4. Modelling approach

The MLR algorithm is a statistical tool used to investigate the relationship between two or more independent variables and an outcome (Srisomkiew et al., 2021). Our goal in the current study is to develop optimal linear models for generating digital maps elucidating soil physicochemical properties and fertility using spectral bands and indices (Liu et al., 2022; Geng et al., 2024). Therefore, spectral bands and indices were chosen as explanatory variables, while physicochemical soil properties were identified as dependent variables. The MLR algorithm assumes a linear relationship between dependent and independent variables, with residuals having a normal distribution. Additionally, multicollinearity among predictor variables was assessed using the variance inflation factor (VIF) and Principal Component Analysis (PCA). VIF was calculated for each predictor, and variables with VIF values exceeding 10 were considered highly collinear and subject to removal. Then, PCA was applied to transform correlated variables into independent components, reducing redundancy in the dataset. The first few principal components explaining the majority of variance were retained for further analysis, ensuring a stable MLR model. After verifying multicollinearity, the obtained PC values were split into two groups: a training dataset (35 samples; 70%) and a testing dataset (15 samples; 30%). A 10-fold cross-validation was performed on the training dataset to assess the MLR model, as this method is commonly used for evaluating predictive models. Finally, the testing dataset was used to assess the performance of the predictive MLR model. MLR equations were developed using the SPSS software, allowing it easier to predict spatial distribution patterns of soil properties. The linear regression model can be represented as follows:

$$\gamma = a + x_1\beta_1 + x_2\beta_2 + \cdots \ldots . + x_n\beta_n \quad \text{Eq. 2}$$

Where γ represents the dependent variables, x_1, x_2, x_n are the coefficients representing the relationship between each dependent and independent variables, a is the intercept, and β_1, β_2, and β_n are the independent variables.

In the final stage, the models' performance and accuracy were evaluated using a variety of statistical criteria, including root mean square error (RMSE), coefficient of determination (R2), and mean absolute error (MAE), as shown in equations 3, 4, and 5. The detailed workflow is illustrated in Fig. 2

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum(X_0 - X_t)^2} \quad (3)$$

$$R^2 = 1 - \frac{\sum(X_0 - X_t)^2}{\sum(X_t)^2} \quad (4)$$

$$\text{MAE} = \frac{\sum|X_0 - X_t|}{n} \quad (5)$$

Where X_0 and X_t denote the predicted and observed values, respectively, and n denotes the total number of observed data.

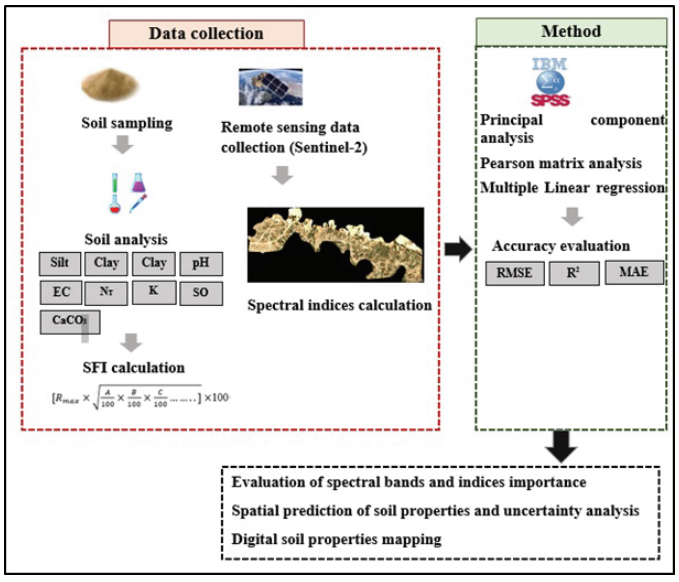

Fig. 2. Flow chart of this study

2.2.5. Digital mapping of soil fertility

The predicted values from the MLR model were used to create a digital soil fertility map for the study area. These values were calculated for each grid cell of the satellite image across the region. The map was then generated using QGIS software (3.16). To classify soil fertility, we applied the Soil Survey Staff (1999) classification system, which categorizes soil into five classes: very low, low, moderate, high, and very high.

3. Results and discussion

3.1. Soil fertility attributes

The principal component analysis identified four distinct factors in the soil quality dataset, each with an eigenvalue greater than one, accounting for 75.95% of the variance (Table 5). The variance distribution across the four principal components is as follows: 31.45% for PC1, 21.99% for PC2, 11.92% for PC3, and 10.59% for PC4. The varimax rotation for the PCA, as shown in Table 6, reveals the contribution of each parameter to the respective principal component. According to loading component theory, a factor loading greater than 0.5 indicates a relatively strong correlation between the variables and the underlying factor, making it significant for component explanation (Liu et al., 2023).

The results from PC1 exhibited high positive factor loadings for Kav (0.87), CaCO3 (0.95), EC (0.79), and pH (0.92), with component scores equal to or exceeding 0.5. In contrast, PC2 revealed high positive factor loadings for soil Ntot (0.65), Pav (0.86), and SFI (0.81). Meanwhile, PC3 showed high positive factor loadings for sand (0.95) and negative factor loadings for silt (–0.74), whereas PC4 exhibited high positive factor loadings for clay (0.97) and negative factor loadings for silt (–0.63). Consequently, PC1 indicates common behaviors related to soil salinity and alkalinity, while PC2 serves as a soil fertility indicator. Thus, the parameters associated with PC2 are identified as the most suitable indicators for soil fertility.

3.2. Determination of potential variables

Figure 3 depicts the Pearson correlation analysis of the spectral variables (indices and bands) with PC2. The results showed that the soil indicator from PC1 correlated significantly with the spectral indices of SM (R=0.589), BI (R=0.674), CI (R=0.542), CI green (R=0.500), and B8A (R=0.501). Other indices and bands, including NDVI, EVI, TVI, SATVI, GNDI, SOCI, Clay I, CARL, CRSI, S, ARVI, S1, S2, B2, B3, B4, B8, B5, B6, B8A, B11, and B12, had less

Table 5
Percentage of variance of selected factors using the principal component analysis extraction method

Component	Eigenvalue	% of total variance	% of cumulative variance
PC1	3.46	31.45	31.45
PC2	2.41	21.99	53.44
PC3	1.31	11.92	65.37
PC4	1.16	10.59	75.45

Table 6
Principal component analysis of soil properties

	Component			
	PC1	PC2	PC3	PC4
N_{tot}	0.067	0.651	–0.089	0.092
P_{av}	–0.241	0.861	–0.037	0.031
Kav	0.871	0.121	0.183	–0.111
CaCO3	0.953	0.111	–0.062	0.087
EC	0.791	–0.130	0.155	0.183
pH	0.929	0.077	–0.142	0.094
SOM	0.290	0.421	–0.234	–0.009
Clay	0.126	0.024	0.018	0.973
Silt	–0.121	0.114	–0.745	–0.634
Sand	0.043	–0.171	0.958	–0.055
SFI	0.135	0.812	0.016	–0.379

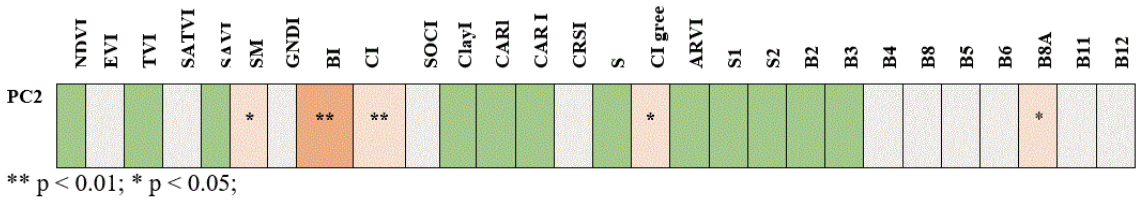

Fig. 3. Pearson correlation matrix between spectral variables and PC2

significant correlations with soil fertility. The MRL model equation was developed using spectral predictors with significant correlations to PC2 ($p < 0.01$ and $p < 0.05$). Thus, four variables such as SM, BI, CI, CI green, and B8A were identified as key predictors of soil fertility in this study.

According to Soares et al. (2023) and Srismokiew et al. (2022), the Brightness Index (BI) is a quantitative measure of soil color that is frequently used to assess soil properties and conditions, particularly organic matter and soil moisture levels. Zhang et al. (2023) also identified the Color Index (CI) as a significant predictor of soil organic carbon. Li et al. (2024) and Zhang et al. (2024) discovered that soil moisture (SM) has a significant correlation with soil quality indicators, particularly soil organic carbon (SOC). They discovered that improved SM accelerates vegetation growth and nutrient uptake, increasing crop yield. According to Wolanin et al. (2019) and Dong et al. (2020), the Green Chlorophyll Index (CI green) is related to crop biomass and yield, which are important factors in determining soil fertility. Furthermore, Andreatta et al. (2022) and Xu et al. (2023) explained that the red-edge spectral region provides precise information for identifying vegetation indices and mapping crop types on a cultivated land parcel scale. As a result, these indices are important variables that effectively express soil fertility in both this study area and other agricultural fields.

3.3. Multiple linear regression model

Table 7 presents the estimated soil fertility based on key variables, including SM, BI, CI, CI green, and B8A. The developed multiple linear regression (MLR) model is mathematically described by Equation 6. This model demonstrates strong predictive performance, with statistical metrics showing a low RMSE (1.11), a low MAE (0.68), and a very high $R_{\text{ˌ}}$ (0.98). The accuracy of the quantitative mapping was assessed by calculating the $R_{\text{ˌ}}$ value from a linear regression analysis between the test data and predicted values. The results revealed a statistically significant relationship, with an $R_{\text{ˌ}}$ of 0.79, indicating the model's capability to predict soil fertility effectively.

3.4. Digital soil fertility map

The MLR model's predicted values were used to generate a digital soil fertility map with 20 and 10 meter resolutions (Fig.4a). As shown in Fig. 4a, the predicted values varied from 0.29 to 11. As a result, four soil fertility classes (Fig. 4b) were identified: extremely low, low, medium, and very high. The SFI threshold values and classes used in this study were defined based on Soil Survey Staff (1999) and Tunçay et al. (2021). Figure 5 shows the distribution of predicted classes, including

Table 7
Summary of the developed multiple linear regression model parameters

MLR	RMSE	R^2	MAE
$SF = 1.31 + 1.92SM - 4.97BI + 16.04CI - 1.48CI_{green} + 1.36B8A$	1.11	0.98	0.68

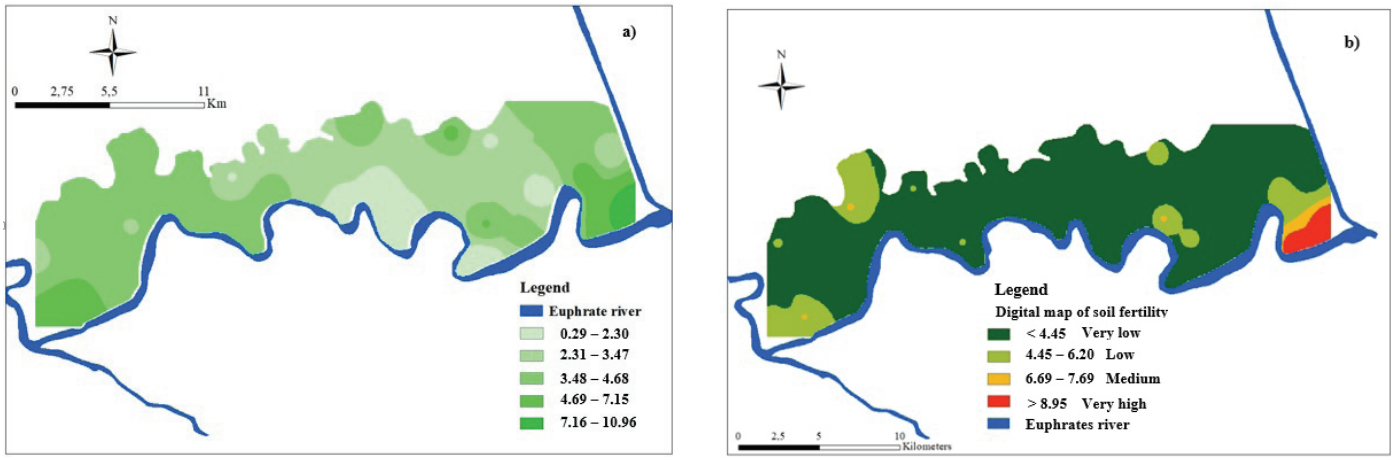

Fig. 4. Digital map of soil fertility in the Euphrates basin: a) map of predicted fertility values; b) map of fertility classes

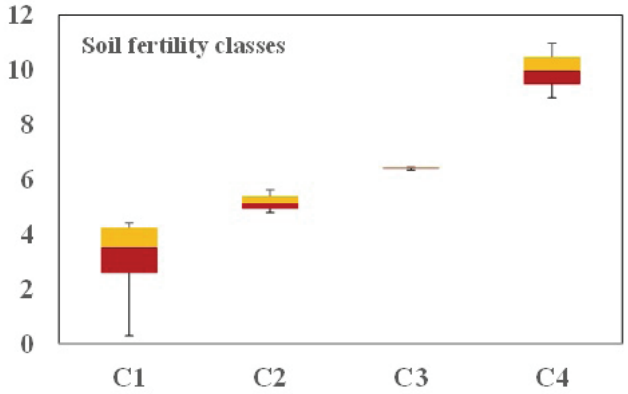

Fig. 5. Distribution of predicted soil fertility classes at Euphrates basin

minimum, maximum, mean, and standard deviation values: very low (0.29–4.45, 3.53±1.18), low (4.45–6.20, 5.14±0.36), medium (6.20–7.69, 6.43±0.05) and very high (8.95–10.79, 9.97±1.40).

Overall, the majority of the study area's soil was rated as very low fertility. The spatial distribution of soil fertility revealed that the eastern parts of the study area were moderate to very fertile, with moderate values also found in the western parts. However, the central region of the survey area had relatively low soil fertility. This indicates that a large portion of the agricultural soil in the study area is deficient in fertility, highlighting severe production constraints.

The creation of a soil fertility map is a useful and effective method for determining land suitability for agriculture. Previous research has identified important predictors of soil properties for this purpose. Several global digital soil property maps have been successfully developed at fine resolutions for various applications, including hydrological modeling (Lopez-Ballestero et al., 2023), guiding future field soil sample collections (Blackford et al., 2022), and providing national-scale site-specific development recommendations (Gagkas et al., 2021).

Remote sensing data has been successfully used in digital soil mapping studies on a regional and global scale (Zhang et al., 2024; Tiangi et al., 2024; Nasser et al., 2024; Justin Fagnombo et al., 2022). The majority of these studies have focused on cultivated land to assess soil fertility, emphasizing the value of remote sensing in predicting soil properties. Our findings show that surfaces without vegetation cover have an advantage in terms of facilitating the collection of spectral reflectance for soil property evaluation.

4. Conclusions

This study focuses on digitally assessing soil fertility using a digital soil map in the Al-Anbar district of the Euphrates basin. The map was created with high-resolution remote sensing data and multiple linear regression models. Total soil nitrogen and available phosphorus content were found to be significant indicators of soil fertility. Our quantitative analysis using RMSE, MAE, and $R_{\text{¸}}$ metrics showed that digital soil fertility maps were more accurate than measured data maps. This demonstrates the potential advantages of using remote sensing data with resolutions of 20 and 10 meters in conjunction with multiple linear regression models. Future research should look into the relationship between remote sensing data and soil fertility across a variety of crops, as well as environmental covariates like precipitation.

Conflict of interest

The authors declare no conflict of interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. This research did not involve human or animal subjects.

Author Contributions

Jamal Abed Hammad: Conceptualization, Data curation, Investigation, Writing – original draft. **Soumaia M'nassri**: Writing – review & editing, Visualization, Data curation. **Balkis Chaabane**: Data curation, Visualization. **Ali Hussein Ibraheem Al-Bayati**; Conceptualization, Methodology, Supervision. **Rajouene Majdoub:** Supervision, Validation, Writing – review & editing. All authors read and approved the final manuscript.

References

Abed Hammad, J., M'nassri, S., Chaabane, B., Al-Bayati, A.H., Majdoub, R., 2024. Assessing agricultural potential of abandoned land in the Eupherates basin: Soil fertility modelling and geostatistical analysis. Modelling Earth Systems and Environment 10, 4627–4639. https://doi.org/10.1007/s40808-024-01982-9

Andreatta, D., Gianelle, D., Scotton, M., Dalponte, M., 2022. Estimating grassland vegetation cover with remote sensing: A comparison between Landsat-8, Sentinel-2 and PlanetScope imagery. Ecological Indicators 141, 109102. https://doi.org/10.1016/j.ecolind.2022.109102

Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A.C., Morgan, C.L.S., Roudier, P., Poggio, L., Mulder, V.L., 2020. Impressions of digital soim maps: The goods, the not so good, and making them ever better. Geoderma Regional 20, e00255. https://doi.org/10.1016/j.geodrs.2020.e00255

Bao, Y., Yao, F., Meng, X., Wang, J., Liu, H., Wang, Y., Liu, Q., Zhang, J., Mouazen, A.M., 2024. A fine digital soil mapping by integrating remote-based process model and deep learning method in Northeast China. Soil and Tillage Research 238, 106010. https://doi.org/10.1016/j.still.2024.106010

Blackford, C., Heung, B., Webster, K.L., 2022. Incorporating spatial uncertainty maps into soil sampling improves digital soil mapping classification accuracy in Ontario, Canada. Geoderma Regional 29, e00495. https://doi.org/10.1016/j.geodrs.2022.e00495

Chlingaryan, A., Sukkarieh, S., Whelan, B., 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. Computers and Electronics in Agriculture 1511, 61–69. https://doi.org/10.1016/j.compag.2018.05.012

Congedo, L., 2016. Semi-Automatic Classification Plugin Documentation Release 4.8.0.1.

Dasgupto, S., Debnath, S., Das, A., Biswas, A., Weindorf, D.C., Li, B., Shukla, A.K., Das, S., et al., 2023. Developping regional soil micronutrient management strategies through ensemble learning based digital soil mapping. Geoderma 433, 116457. https://doi.org/10.1016/j.geoderma.2023.116457.

Diaz-Gonzalez, F., Vuelvas, J., Correa, C.A., Vallejo, V.E., Patino, D., 2022. Machine learning and remote sensing techniques applied to estimate soil indicators-Review. Ecological Indicators 135, 108517. https://doi.org/10.1016/j.ecolind.2021.108517

Douaoui, A.K., Nicolas, H., Walter, Ch., 2006. Detecting salinity hazard within a semiarid context by means of combining soil and remote sensing data. Geoderma 134 (1–2), 217–230. https://doi.org/10.1016/j.geoderma.2005.10.009

Du, P., Bai, X., Tan, K., Xue, Z., Samat, A., Xia, J., et al. 2020. Advances of four machine learning methods for spatial data handling: a review. Journal of Geovisualization and Spatial Analysis 4, 13. https://doi.org/10.1007/s41651-020-00048-5

Dvornikov, Y., Slukovskaya, M.V., Yaroslavtsev, A., Meshalkina, J., Ryazanov, A.V., Sarzhanov D., Vasenev, V., 2022. High-resolution mapping of soil pollution by Cu and Ni at a popular industrial barren area using proximal and remote sensing. Land Degradation & Development 33, 1731–1744. https://doi.org/10.1002/ldr.4261

Dong, E., Liu, J., Qian, B., He, L., Liu, J., Jing, Q., Champagne, C., McNairn, H., Powers, J., Shi, Y., Chen, J., Shang, J., 2020. Estimating crop biomass using leaf area index derived from Landsat 8 and Sentinel-2 data. ISPRS Journal of Photogrammetry and Remote Sensing 168, 236–250. https://doi.org/10.1016/j.isprsjprs.2020.08.003

Escadafal, R., 1989. Remote sensing of arid soil surface color with Lanat Thematic Mapper. Advances in Space Research 9(1), 159–163. https://doi.org/10.1016/0273-1177(89)90481-X

Esmaeilizad, A., Shokri, R., Davatgar, N., Dolatabad, H., 2024. Exploring the driving forces and digital mapping of soil biological properties in semi-arid regions. Computers and Electronics in Agriculture 220, 108891. https://doi.org/10.1016/j.compag.2024.108831

European Space Agency, 2010. GMES Sentinel-2 mission requirements document. In: Technical Report Issue 2 Revision 1.

Gagkas, Z., Lilly, A., Baggaley, N.J., 2021. Digital soil maps can perform as well as large-scale conventional soil maps for the prediction of catchment baseflows. Geoderma 400, 115230. https://doi.org/10.1016/j.geoderma.2021.115230

Geng, J., Tan, Q., Lv, J., Fang, H., 2024. Assessing spatial variations in soil organic carbon and C:N in Northeast China's black soil region: Insights from Landsat-9 satellite and crop growth information. Soil Tillage and Research 235, 105897. https://doi.org/10.1016/j.still.2023.105897

Grinand, C., Le Maire, G., Vieilledent, G., Razakamanarivo, H., Razafimbelo, T., 2017. Estimating temporal changes in soil carbon stocks at ecoregional scale in Madagascar using remote-sensing. International Journal of Applied Earth Observation and Geoinformation 54, 1–14. https://doi.org/10.1016/j.jag.2016.09.002

Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment 58, 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gholizadeh, A., Zizala, D., Saberioon, M., Boruvka, L., 2018. Soil organic carbon and texture retrieving and mapping using proximal airborne and Sentinel-2 spectral imaging. Remote Sensing of Environment 218, 89–103. https://doi.org/10.1016/j.rse.2018.09.015

Guo, P.T., Li, M.F., Luo, W., Tang, Q.F., Liu, Z.W., Lin, Z.M., 2015. Digital of soil organic matter for rubber plantation at regional scale: an application of random forest plus residuals kriging approach. Geoderma 237, 49–59. https://doi.org/10.1016/j.geoderma.2014.08.009

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Ferreira, L.G., 2002. Overview of the radiometric and physiological performance of the MODIS vegetation indices. Remote Sensing of Environment 83 (1–2), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2.

Huan, S., Zhang, W. J., Yu, X.C., Huang, Q.R., 2010. Effects of long-term fertilizer on corn productivity and its sustainability in an Ultisol of southern China. Agriculture, Ecosystems and Environment 138 (1–2), 44–50. https://doi.org/10.1016/j.agee.2010.03.015

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 25(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Justin Fagnombo D., Jean-Martial, J., Saito, K., 2022. Ca soil fertility properties in rice fields in ssub-Saharan Africa be predicted by digital soil information? A case study of AsoilGrids250m. Geoderma Regional, e00563. https://doi.org/10.1016/j.geodrs.2022.e00563

Kasampalis, D., Alexandridis, T., Deva, C., Challinor, A., Moshou, D., Zalidis, G., 2018. Contribution of remote sensing on crop models: a review. Journal of Imaging 4(4), 52. https://doi.org/10.3390/jimaging4040052

Kaufman, Y., Tanre, D., 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. IEEE Transactions on Geoscience and Remote Sensing 30, 261–270. https://doi.org/10.1109/36.134076

Hounkpatin, O., De Hipt, F., Bossa, A., Welp, G., Amelung, W., 2018. Soil organic carbon stocks and their determining factors in the Dano catchment (Southwest Burkin Faso). Catena 166, 298–309. https://doi.org/10.1016/j.catena.2018.04.013

Kumar, U., Nayak, A.K., Shahid, M., Gupta, V., Panneerslvan, P., Mohantry, S., Kaviraj, M., Kumar, A., Chatterjee, D., Lal, B., Gautam, P., Tripathi, R., Panda, B., 2018. Continuous application of inorganic and organic fertilizers over 47 years in paddy soil alters the bacterial community structure and its influence on rice production. Agriculture, Ecosystems and Environment 262, 65–75. https://doi.org/10.1016/j.agee.2018.04.016

Lopez-Ballesteros, A., Nielson, A., Gastellanos-Osorio, G., Trolle, D., Senent-Aparicio, J., 2023. DSOLMap, a novel high-resolution global digital soil property map for the SWAT+model: Development and hydrological evaluation. Catena 231, 107339. https://doi.org/10.1016/j.catena.2023.107339

Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., Shearer S., 2018. Integration of high-resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. Computers and Electronics in Agriculture 153, 213–225. https://doi.org/10.1016/j.compag.2018.07.016

Marsett, R.C., Qi, J., Heilman, P., Biedenbender, S.H., Watson, M.C., Amer, S., Weltz, M., Goodrich, D., Marsett, R., 2006. Remote sensing for grassland management in the arid southwest. Rangeland Ecology and Management 59, 530–540. https://doi.org/10.2111/05-201R.1

Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine learning in agriculture: a review. Sensors 18, 1–29. https://doi.org/10.3390/s18082674

Liu, F., Zhang, J., Wang, S., Zou, J., Zhen, P., 2023. Multivariate statistical analysis of chemical and stable isotopic data as indicative of groundwater evolution with reduced exploitation. Geoscience Frontiers 14, 101476. https://doi.org/10.1016/j.gsf.2022.101476

Li, Y., Feng, X., Huai, Y., Hassan, M., Cui, Z., Ning, P., 2024. Enhancing crop productivity and resilience by promoting soil organic carbon and moisture in wheat and maize rotation. Agriculture, Ecosystems and Environment 368, 109021. https://doi.org/10.1016/j.agee.2024.109021

Lu, Q., Tian, S., Wei, L., 2023. Digital mapping of soil pH and carbonates at the European scale using environmental variables and machine learning. Science of the Total Environment 856, 159171. https://doi.org/10.1016/j.scitotenv.2022.159171

M'nassri, S., Dridi, L., Schafer, G., Hachicha, M., Majdoub, R. 2019. Groundwater salinity in a semi-arid region of central-eastern Tunisia: insights from multivariate statistical techniques and geostatistical modelling. Environmental Earth Sciences 78, 288. https://doi.org/10.1007/s12665-019-8270-8

Mondejar, J.P., Tongco, A.F., 2019. Estimating topsoil texture fractions by digital soil mapping: A response to the long-outdated soil map in the Philippines. Sustainable Environment Research 29, 31. https://doi.org/10.1186/s42834-019-0032-5

Nasser, F.C., De Mello, D.C., Francelino, M.R., Krause, M.B., Soares, H., Dematte, J., 2024. Mapping desactivated mine areas in the amazon for-

est impacted by seasonal flooding: Assessing soil-hydrological processes and quality dynamics by remote and geophysical techniques. Remote Sensing Applications: Society and Environment 34, 101148. https://doi.org/10.1016/j.rsase.2024.101148

Nelson, D.W., Sommers, L.E., Sparks, D.L., Page, A.L., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. 1996. Total carbon, organic carbon and organic matter. Methods Soil Analysis 9, 961–1010. https://doi.org/10.2136/sssabookser5.3.c34

Nellis, M.D., Briggs, J.M., 1992.Transformed vegetation index for measuring spatial variation in drought impacted biomass on Konza Prairie, Kansas. Transactions of Kansas Academy of Science 1903(95), 93–99. https://doi.org/10.2307/3628024

Nguy-Robertson, A., Peng, Y., Gitelson, A., Arkebauer, T., Pimstein, A., Herrmann, I., Karnieli, A., Rundquist, D., Bonfil, D., 2014. Estimating green LAI in four crops: Potential of determining optimal spectral bands for a universal algorithm. Agricultural and Forest Meteorology 192-193, 140–148. https://doi.org/10.1016/j.agrformet.2014.03.004

Ozayzici, M.A., Dengiz, O., Saglam, M., Erkoçak, A., Turkmen, F., 2017. Mapping and assessment-based modelling of soil fertility differences in the central and eastern parts of the black sea region using GIS and geostatistical approaches. Arabian Journal of Geosciences 1045, 1–9. https://doi.org/10.1007/s12517-016-2819-6

Poppiel, R.R., Dematte, J.A., Rosin, N.A., Camps, L.R., Tayebi, M., Bonfatti, B.R., Ayoubi, S., et al. 2021. High-resolution middle eastern soil attributes mapping via open data and cloud computing. Geoderma 385, 114890. https://doi.org/10.1016/j.geoderma.2020.114890

QGIS Development Team (2020): QGIS geographic system. https://qgis.org.

Ray, S.S., Singh, J.P., Das, G., Panigrahy, S., 2004. Use of High resolution remote sensing data for generating site-specific soil management plant. International Society for Photogrammetry and Remote Sensing 35, 127–132. https://pdfs.semanticscholar.org/c69a/f13fbfd1795dafb962619f1 b7a673fdbb013.pd

Rock, B.N., Williams, D.L., Vogelmann, J.E., 1985. Field and airborne spectral characterization of suspected acid deposition damage in red spruce (Picea ribens) from Vermont. In: Proceeding of Sumposium on Machine Processing of Remotely Sensed Data. Wesr Ladayette, Indiana, USA.

Rouse, J.W., Hass, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the great plains with ERTS. In: Third Earth Resources Technology Satellite-1 Symposium. pp. 309–317.

Scudiero, E., Skaggs, T.H., Corwin, D.L., 2014. Regional scale soil salinity evaluation using Landsat 7, western San Joaquin Valley, California, USA. Geoderma Regional 2-3, 82–90. https://doi.org/10.1016/j.geodrs.2014.10.004

Soil Survey Staff, 1992. Keys to soil Taxonomy, fifth ed. Pocahontas Press, Blacksburg, VA.

Soil Survey Staff, 1999. Soil Taxonomy. A basic system of soil classification for making and interpreting soil survey. USDA Agriculture Handbook N 436. Washington D.C, USA.

Song, X.D., Rossiter, D.G., Liu, F., Wu, H.Y., Zhao, X.R., Cao, Q., Zhang, G.L., 2020. Can pedotransfer fucntions based on environmental variables improve soil total nutrient mapping at a regional scale? Soil and Tillage Research 202, 104672. https://doi.org/10.1016/j.still.2020.104672

Soares Vieira, A., Do Valle Junior, R.F., Rodrigues, V.S., Silva Quinaia, T., Mendes, R.G., Valera, C.A., Fernandes, L., Pacheco, F., 2023. Estimating water erosion from the brightness index of orbital images: A framework for the prognosis of degraded pastures. Science of the Total Environment 776(1), 146019. https://doi.org/10.1016/j.scitotenv.2021.146019

Srisomkiew, S., Kawahigashi, M., Limtong, P., Yuttum, O., 2022. Digital soil assessment of soil fertility for Thai jasmine rice in the Thung Kula Ronghi region, Thailand. Geoderma 409, 115597. DOI: 10.1016/j.geoderma.2021.115597

Srisomkiew, S., Kawahigashi, M., Limtong, P., 2021. Digital mapping of soil chemical properties with limited data in the Thung Kula Ronghai region, Thailand. Geoderma 389, 114942. https://doi.org/10.1016/j.geoderma.2021.114942

Sun, L., Liu, F., Zhu, X., Zhang, G., 2024. High-resolution digital mapping of soil erodibility in China. Geoderma 444, 116853. https://doi.org/10.1016/j.geoderma.2024.116853

Tiangi, Z., Ye, L., Wang, M., 2024. Remote sensing-based prediction of organic carbon in agricultural and natural soils influenced by salt and sand mining using machine learning. Journal of Environmental Management 352, 120107. https://doi.org/10.1016/j.jenvman.2024.120107.

Tunçay, T., Alaboz, P., Dengiz, O., Baskan, O., 2023. Application of regression kriging and machine learning to estimate soil moisture constants in a semi-arid terrestrial area. Computers and Electronics in Agriculture 212, 108118. https://doi.org/10.1016/j.compag.2023.108118

Tunçay, T., Kiliç, S., Dedeoglu, M., Dengiz, O., Baskan, O., Bayramin, I., 2021. Assessing soil fertility index based on remote sensing and gis techniques with field validation in a semiarid agricultural ecosystem. Journal of Arid Environments 190, 104525. https://doi.org/10.1016/j.jaridenv.2021.104525

Taghizadeh-Mehrjardi, R., 2014. Digital mapping of soil salinity region, central Iran. Geoderma 213, 15–28. https://doi.org/10.1016/j.geoderma.2013.07.020

Thaler, E.A., Larsen, I.J., Yu, Q., 2019. A new index for remote sensing of soil organic carbon based solely on visible wavelengths. Soil Science Society of America Journal 83, 1443–1450. https://doi.org/10.2136/sssaj2018.09.0318

Tripathi, N.K., Brijesh, K.R.D., 1997. Spatial modelling of soil alkalinity in GIS environment using IRS data. Paper presented at the 18th Asian conference in Remote Sensing.

Wolanin, A., Camps-Valls, G., Gomez-Choval, L., Mateo-Gracia, G., Van der Tol, C., Zhang, Y., Guanter, L., 2019. Estmating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations. Remote Sensing of Environment 225, 441–457. https://doi.org/10.1016/j.rse.2019.03.002

Wei, L., Tian, S., Lu, Q., Zhong, Y., Zheng, Y., Lu, Y., Xiao, Z., 2024. Estimating soil organic carbon content of multipe soil horizons in the middle and upper reaches of the Heihe River Basin. Catena 234, 107574. https://doi.org/10.1016/j.catena.2023.107574

Wang, H., Yao, L., Huang, B., Hu, W., Qu, M., Zhao, Y., 2019. An integrated approach to exploring soil fertility from the perspectiveof rice (Oryza sativa L.) Yields. Soil and Tillage Research 194, 104322. https://doi.org/10.1016/j.still.2019.104322

Xu, L., Ming, D., Zhang, L., Dong, D., Qong, Y., Yang, J., Zhou, C., 2023. Parcel level staple crop type identification based on newly defined red-edge vegetation indices and ORNN. Computers and Electronics in Agriculture 211, 108012. https://doi.org/10.1016/j.compag.2023.108012

Yuan J., Gao, J., Yu, B., Yan, Ch., Ma, Ch., Xu, J., Liu, Y., 2024. Estimating of soil organic matter content based on spectral indices constructed by improved Hapke model. Geoderma 443: 116823. https://doi.org/10.1016/j.geoderma.2024.116823

Yuzugullu, O., Fajraoui, N., Don, A., Liebisch, F., 2024. Satellite-based soil organic carbon mapping on European soils using available datasets and support sampling. Science of Remote Sensing 9, 100118. https://doi.org/10.1016/j.srs.2024.100118

Zhang, X., Xue, J., Chen, S., Zhuo, Z., Wang, Z., Chen, X., Xiao, Y., Shi, Z., 2024. Improving model performance in mapping cropland soil organic matter using time-series remote sensing data. Journal of Integrative Agriculture. https://doi.org/10.1016/j.jia.2024.01.015

Zhang, T., Huang, L.M., Yang, R.M., 2024. Evaluation of digital soil mapping projections in soil organic carbon change modelling. Ecological Informatics 79, 102394. https://doi.org/10.1016/j.ecoinf.2023.102394

Zhang, S., Tian, J., Lu, X., Tian, Q., 2023. Temporal and spatial dynamics distribution of organic carbon content of surface soil coastal wetlands

of Yancheng, China from 2000 to 2022 based on Landsat images. Catena 223, 106961. https://doi.org/10.1016/j.catena.2023.106961

Zhang, G., Bai, J., Xi, M., Zhao, Q., Lu, Q., Jia, J., 2016. Soil quality assessment of coastal wetlands in the Yellow River Delta of China based on the minimum data set. Ecological indicators 66, 458–466. https://doi.org/10.1016/j.ecolind.2016.01.046

Zhang, J., Cai, J., Xu, D., Wu, B., Chang, H., Zhang, B., Wei, Z., 2024. Soil salinization poses greater effects than soil moisture on field crop growth and yield in arid farming areas with intense irrigation. Journal of Cleaner Production 451, 142007. https://doi.org/10.1016/j.jclepro.2024.142007

Zhou, T., Geng, Y., Pan, J., Haase, D., Lausch, A., 2020. High-resolution digital mapping of soil organic carbon and soil total nitrogen using DEM derivatives, Sentinel-1 and Sentinel-2 data based on machine learning algorithms. Science of the Total Environment 729, 138244. https://doi.org/10.1016/j.scitotenv.2020.138244