Digital Mapping of Soil pH Using Machine Learning, Remote Sensing, and Field Data for Syrian Soils

Alaa Khallouf[1]
Younis Idries[2]
Muhammad Manhal Al-Zoubi[1]
Wassim Al-Mesber[3]
Maan Daoud[1]

1: General Commission for Scientific Agricultural Research (GCSAR), Damascus, Syria, alaakhallouf@gmail.com
2: General Organization of Remote Sensing (GORS), Damascus, Syria
3: Soil Department, Faculty of Agriculture, Damascus, Syria

1. Introduction

The world is facing numerous issues and challenges in the 21st century, such as population explosion, food security, environmental degradation, water scarcity, threats to biodiversity, climate change, and sustainable development issues (FAO, 2011). In confronting these local and global challenges, the urgent and growing need to update soil data becomes evident (Sanchez et al., 2009). More than 30% of the world's lands are becoming increasingly acidic due to human activities.

There is a strong correlation between pH and soil quality, as it affects the soil's natural environment, microbial activity, and structure. Additionally, it has a significant impact on crop productivity by influencing nutrient availability and the soil's buffering capacity (Neina, 2019). Therefore, monitoring spatial changes in pH values is particularly important for determining soil quality and land use patterns. This study aimed to use remote sensing and field data to predict soil pH values across the entire territory of the Syrian Arab Republic using machine learning. It also aimed to compare the best models for predicting pH at two depths: 0-30 cm and 30-60 cm.

2. Methodology:

2. 1. Soil data

GCSAR (2020) and Khallouf (2022) soil pH data were used (Fig 2.)

2.2. Environmental and Remote Sensing Covariates

The SCORPAN (climate, topography, parent material, vegetation cover, human activity, and time) methodology was used to predict soil pH values, along with field data and geographical location. The predictive factors included 30 variables as shown in Table (1). These data were downloaded via the Google Earth Engine platform and resampled to 250 m resolution.

Fig. 1 Location of Syrian Arab republic

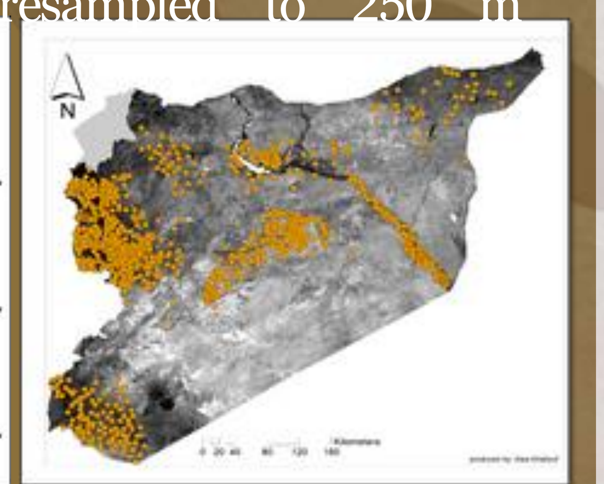

Fig. 2 Soil samples locations

Data	source	Resolution (m)
Climate		
Mean annual temperature (MAT)	Climate AP	1000
Mean annual precipitation (MAP)		
Annual humidity-heat index (AHM)		
Remote sensing image		
band 1	MODIS	250
band2		
band3		
band4		
band5		
Vegetation Index		
NDVI (normalized difference vegetation index)	NIR − Red/NIR + Red	250
RVI (ratio vegetation index)	NIR/Red	250
MSAVI (modified soil-adjusted vegetation index)	0.5*(2*NIR+1-SQRT((2*NIR)2-8*(NIR-Red)))	250
EVI (enhanced vegetation index)	2.5 × (NIR−Red)/(NIR+6*Red-7.5*Blue+1)	250
RGRI (ratio green-red index)	Red/Green	250
GVI	NIR/Green	250
BI (brightness index)	0.5*SQRT(Red*Red)*(Green*Green))	250
B12 (brightness index2)	0.5*SQRT(Red*Red)*(Green*Green)*(NIR*NIR))	250
RI (redness index)	(Red*Red)/(Green*Green*Green)	250
CI (color index)	(Red-Green)/(Red + Green)	250
NDWI (normalized difference water index)	(Green-NIR)/(Green + NIR)	250
DEM derivatives		
Elevation		250
Slope	The degree of steepness of a surface element.	250
Aspect	The degree to which the ground tilts.	250
Plane curvature	The surface shape is viewed in a horizontal plane that has sliced through the surface at the target point.	250

Table 1. Remote sensing indices and covariates that used in modeling process

2.2. Modeling Process

Machine Learning (ML) algorithms were used in the R programming environment to predict soil pH in Syrian soils. The study tested five mathematical models: Multiple Linear Regression, Random Forest (RF), Support Vector Machine (SVM), Cubist, and CART models.

2.3. Model Evaluation The data were split into 70% for model calibration and 30% for accuracy assessment based on many indicators: ME, RMSE, R2, and NSE.

3. Results:

3.1. Testing Model Performance in the Machine Learning Algorithm

From Table (3), it is evident that the best model for predicting soil pH values in Syrian soils is the Random Forest (RF) model. This model has the lowest ME and RMSE values of 0.28213 and 0.48861, respectively, compared to the other models. Additionally, its R2 and NSE values are 0.9531 and 0.99425, respectively, indicating it is the most accurate.

Table 2. Performance Indicators of Mathematical Models within the Machine Learning Algorithm.

NSE	R^2	RMSE	ME	Model
-5.1464	0.11334	1.88351	1.4367	**MLR**
0.99425	0.953126	0.48801	0.28213	**RF**
-4.509335	0.09073	1.9288	1.442207	**SVM**
0.9998	0.94554	0.43466	0.15907	**Cubist**
-5.12533	0.090829	1.92951	1.47827	**CART**

3.2. Spatial Prediction of Soil pH Values in Syrian Soils:

From Figure (2), it can be observed that the pH values ranged between 5.51 and 8.92. Lower values were concentrated in the Shin Plateau in Homs Governorate. Generally, soil pH is high in most Syrian soils, with values exceeding 7.5.

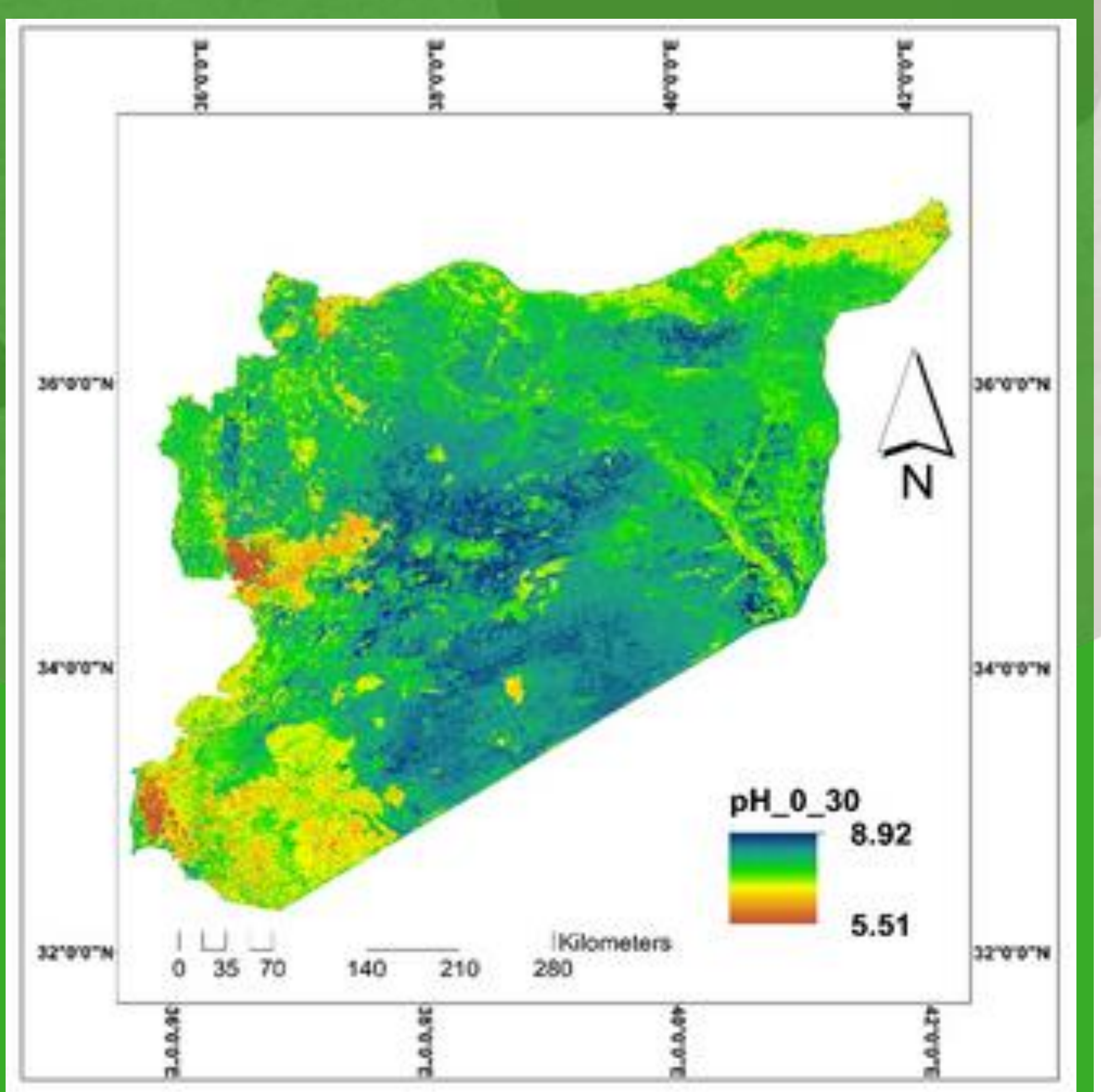

Fig. 3 Spatial distribution of soil pH at depth (0-30 cm)

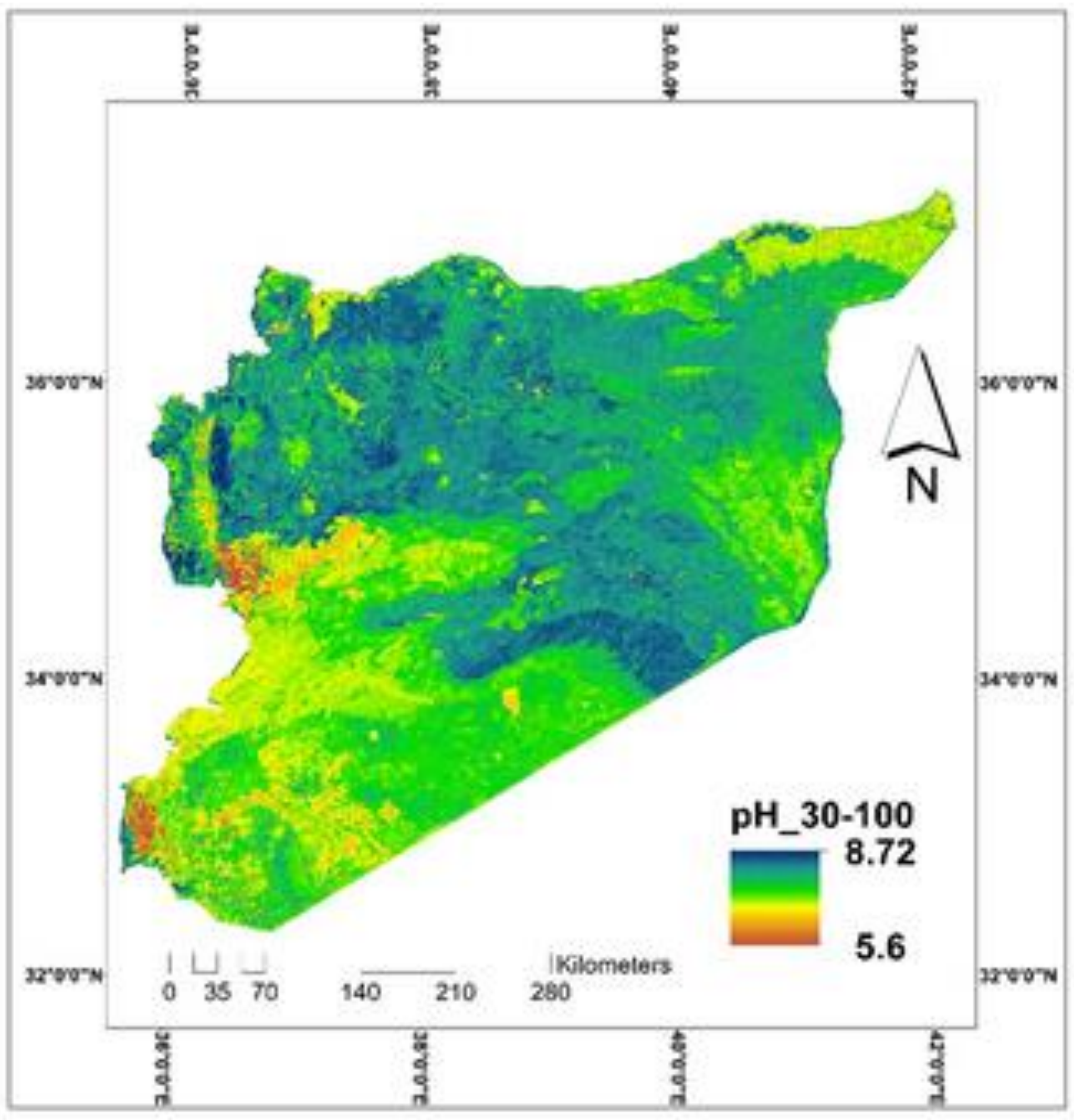

Fig. 4 Spatial distribution of soil pH at depth (30-100 cm)

3.3. Evaluation of Model Efficiency in Predicting pH Values:

The RMSE values (Table. 3) were 1.132 and 0.764 for the two depths, respectively. R² with values of 0.77 and 0.84 for the two depths., the NSE value was -5.692, with an NSE value of -1.8 for subsurface depth.

Table 3. Evaluating the Efficiency of the RF Model in Predicting Soil PH

NSE	R^2	RMSE	Bias	Soil depth(cm)
-5.692683	0.7705202	1.13288	0.05437999	**30-0**
-1.801086	0.8450531	0.7644287	0.03952767	**100-30**

4. Conclusion:

The study demonstrated that remote sensing data, and field data can be effectively used to predict soil pH values in Syrian soils. The pH values ranged from 5.51 to 8.92 in the surface layer and from 5.6 to 8.72 in the subsurface layer.

5. References:

FAO (2011): FAO in the 21st Century – Ensuring Food Security in a Changing World, Food and Agriculture Organization of the United Nations, Rome, www.fao.org/docrep/015/i2307e/i2307e.pdf.

GCSAR (General Commission for Scientific Agricultural Research)(2020). Syrian soil database, unpublished data source.

Khallouf (2022). Land use planning using remote sensing and GIS and linear programming in the western region og Hamah governorate . PhD. THISES .Faculty of AGRICULTURE, Al-Baath University, Homs, Syria

Neina, D. (2019): The Role of Soil pH in Plant Nutrition and Soil Remediation. Appl. Environ. Soil Sci. 2019, 5794869.

Sanchez, P.A.; Ahamed, S.; Carré, F.; Hartemink, A.E.; Hempel, J.; et al.,(2009): Digital Soil Map of the World: Increased demand and advanced techniques could lead to more refined mapping and management of soils. Science, 325(5941): 680-681. https://www.science.org/doi/abs/10.1126/science.1175084