Digital Mapping of Soil Organic Matter Using Open Source Accessible Products of ESA® in Arable Plain

Fuat Kaya*, Levent BAŞAYİĞİT Isparta University Of Applied Sciences, Faculty Of Agriculture, Department Of Soil Science And Plant Nutrition (*fuatkaya@isparta.edu.tr)

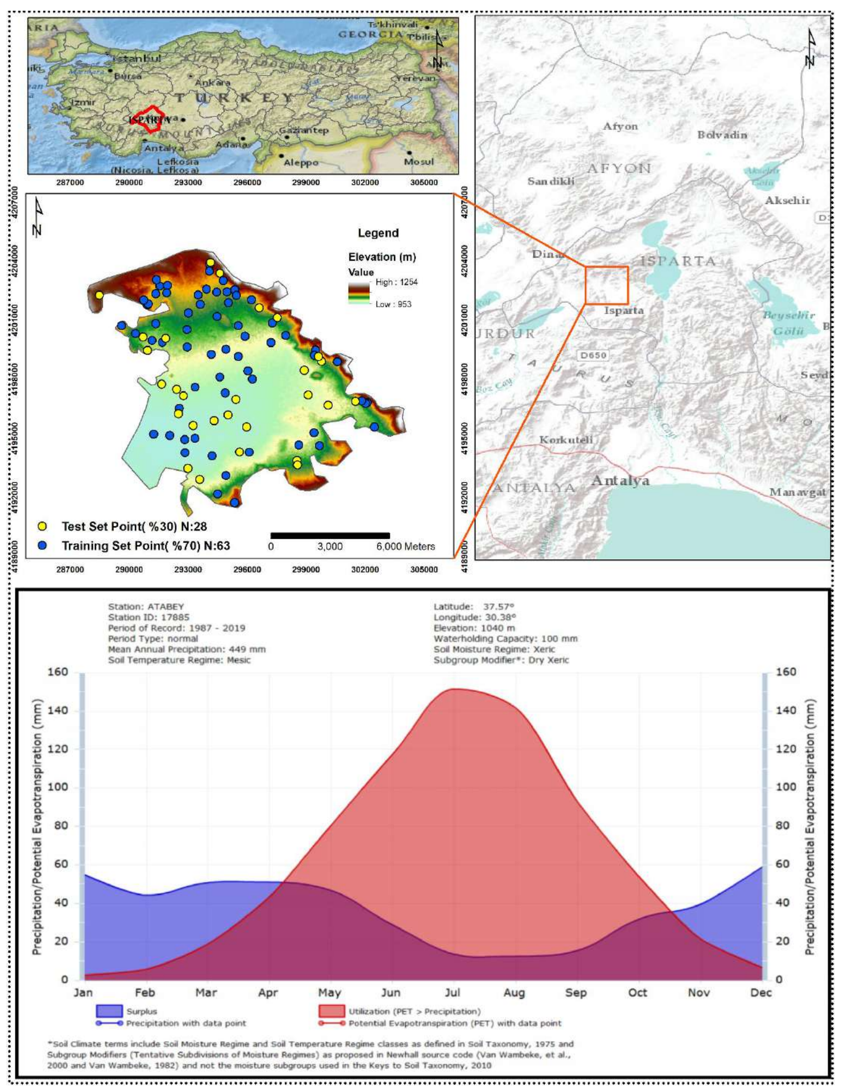
INTRODUCTION

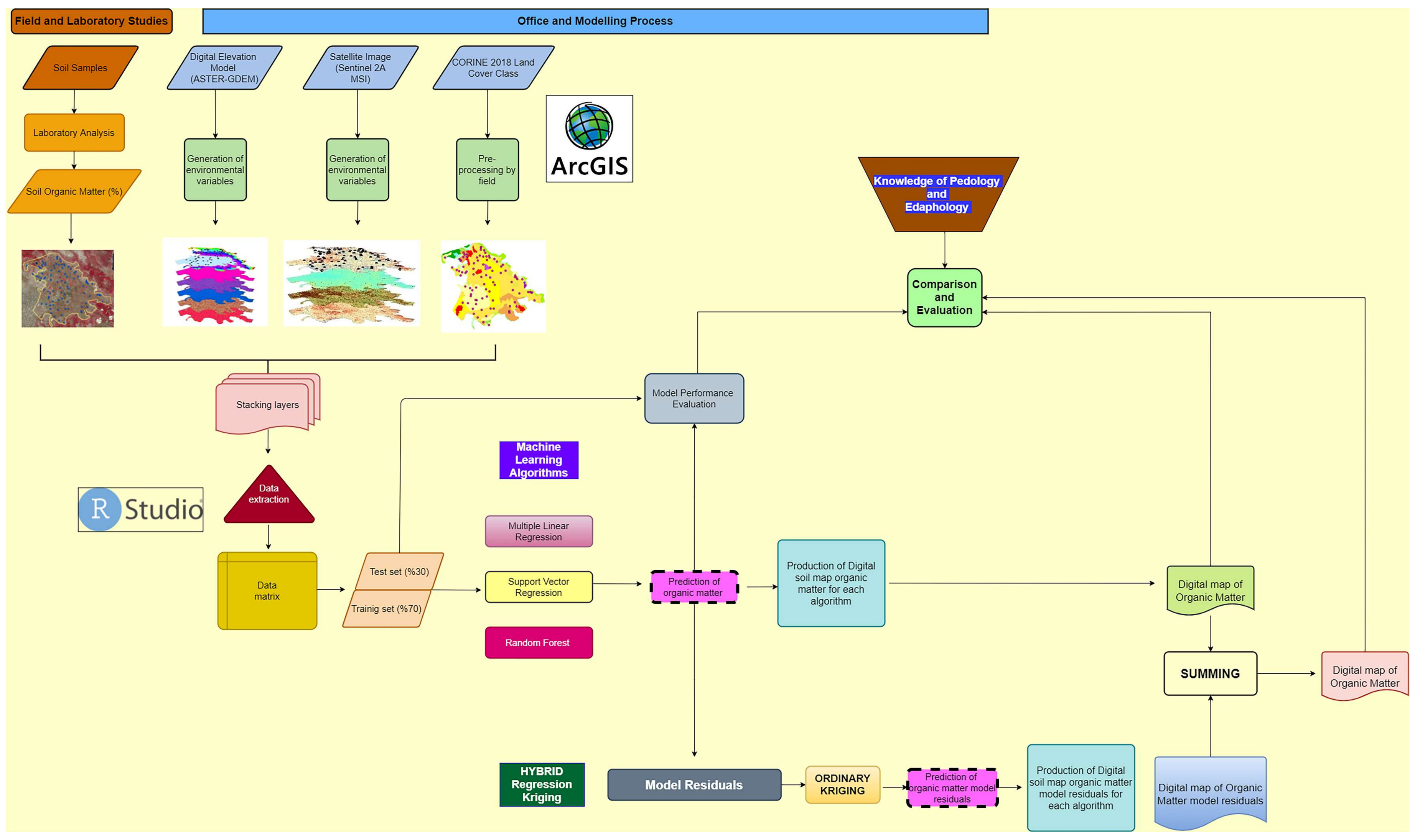
The soil organic matter represents a key function affecting other soil properties (Wiesmeier et al., 2019). Soil scientists are interested in spatially accurate prediction of the organic matter content of soils, which are being greatly interfered with by humans. Irrigation applications and additional organic and chemical fertilizer applications are applied to the soil in most of the arid and semi-arid areas. Model-based digital soil mapping is seen as a promising alternative to the surrogate-based approach for quantifying organic matter content. It ensures some advantage time- and cost intensive measurements. In digital soil mapping, the usability of soil organic matter information obtained from geostatistical approaches is limited in areas where there are socio-ecologically heterogeneous farms, that is, different land uses (Mponella et al., 2020). However, the spatial location of training and test points is often neglected in the process of mapping studies with machine learning algorithms currently used to generate spatial predictions. As a result, only data-driven modeling results may be oversensitive and overfitting problems may be encountered.

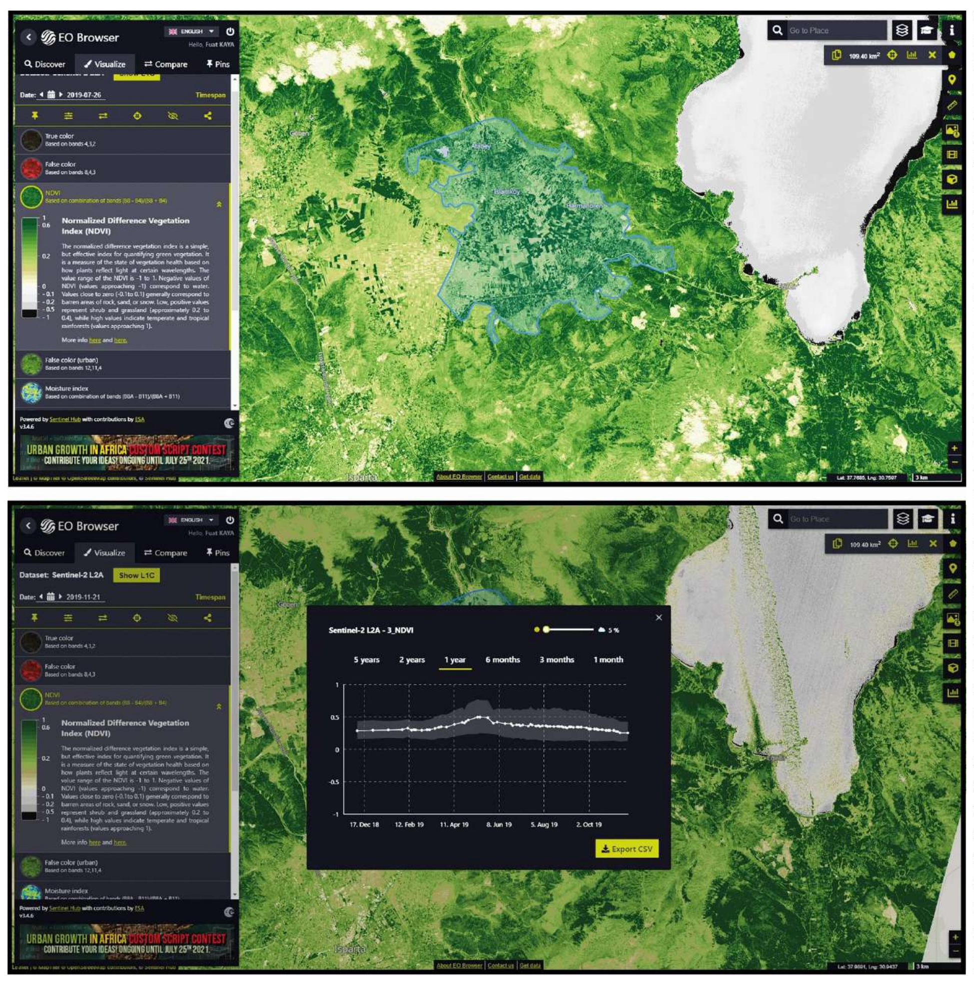
MATERIAL AND METHODS

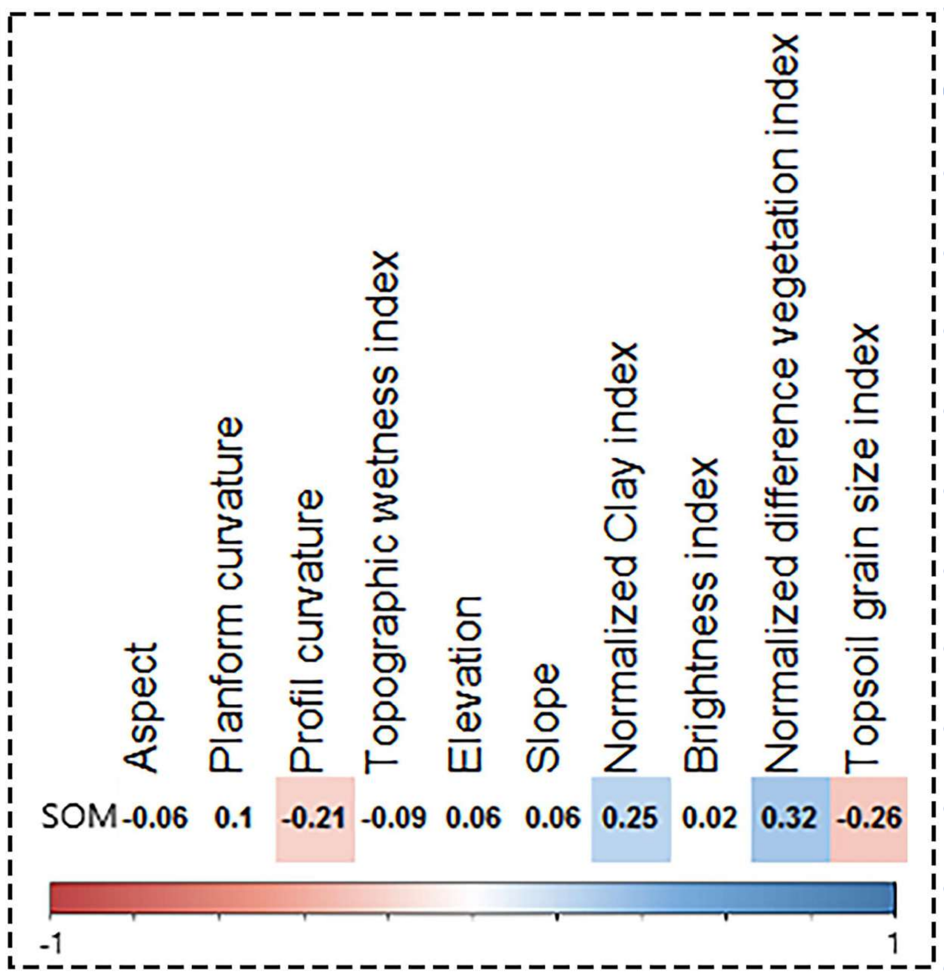
This study was carried out on an area of approximately 10000 ha in the Southwest of Turkey. The study area was monitored using the Sentinel Hub EO Browser tool. Soil samples were taken during the period when the NDVI value was lower than average for that year. We were used 91 topsoil observations(%70 training, N:63; % 30 test, N:28) (R Core Team, 2021), indices produced from Sentinel 2A satellite imagery, topographical variables produced from the digital elevation model, and the latest published CORINE land cover classes map which show the effectiveness of agricultural activities for many years. Soil organic matter (SOM) in the soil samples was determined in the laboratory as % (FAO, 2019). Lin's concordance correlation coefficient (LCCC) and Root mean square error (RMSE) were used as model accuracy criteria.

In this study, we determined the pure model prediction accuracy of multiple linear regression, which can only detect linear relationships, and random forest and support vector machine algorithms that can detect nonlinear relationships. In the hybrid (Regression Kriging- RK) approach (Hengl et al., 2018), explanatory variation is estimated by 3 different algorithms and the process is carried out by collecting the regression value of soil organic matter and the kriging values of model residuals in non-sampled locations.





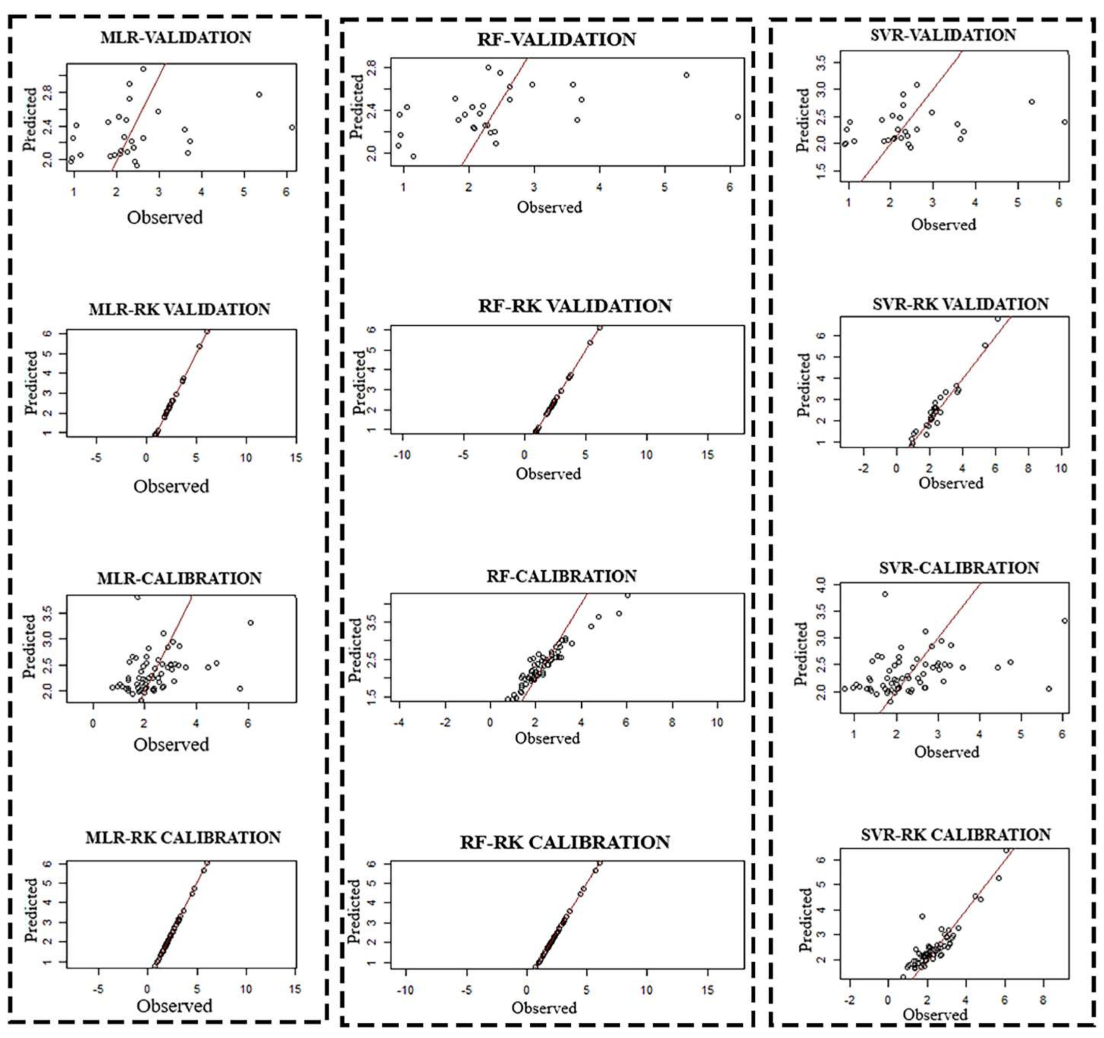


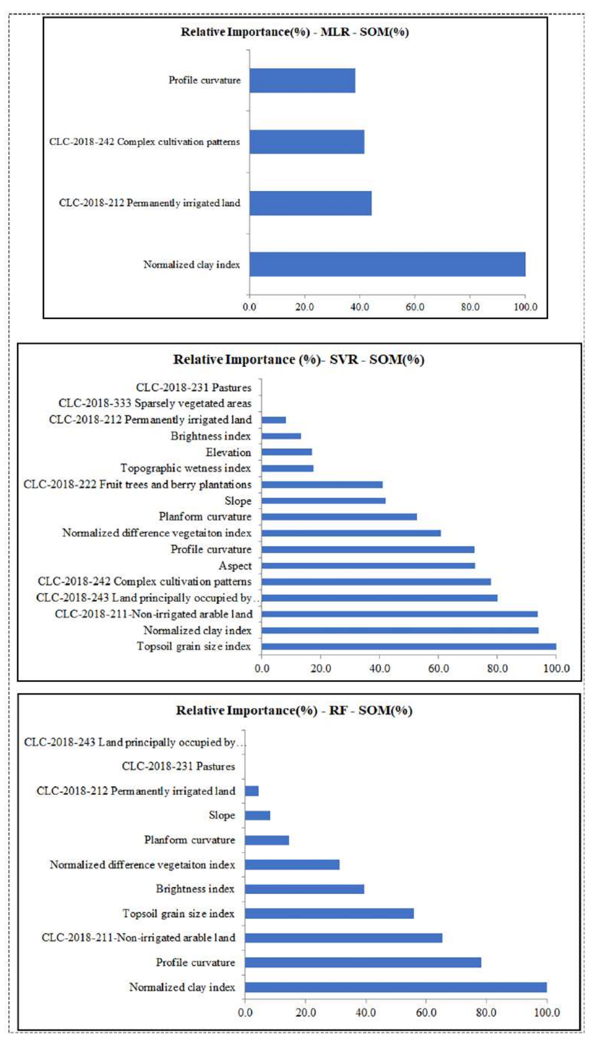


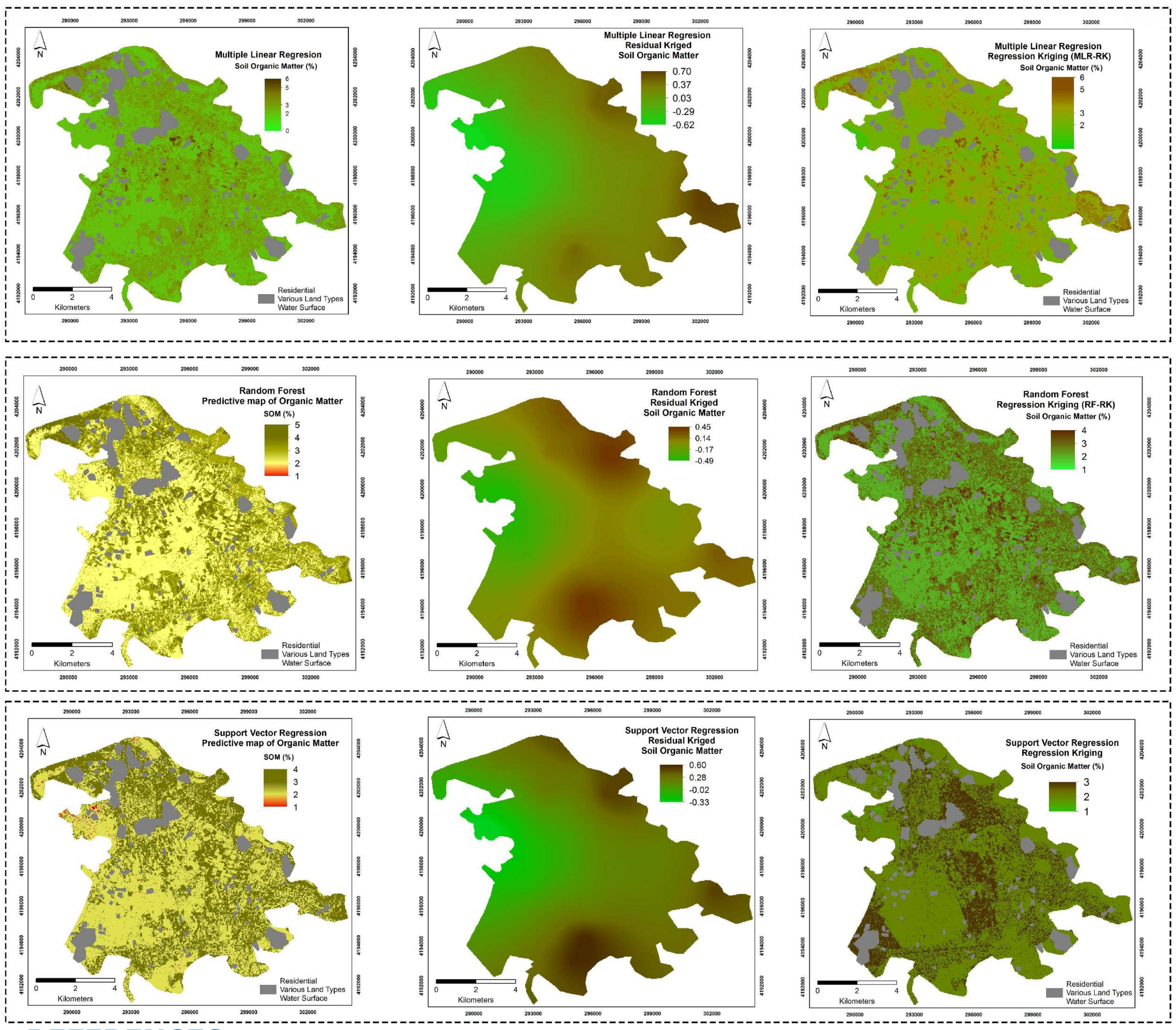
Variable	SOM (%)
Mean	2.37
StDev	1.06
Variance	1.12
CoefVar	44.69
Minimum	0.76
Q1	1.77
Median	2.16
Q3	2.70
Maximum	6.11
Skewness	1.57
Kurtosis	3.33

RESULTS AND DISCUSSION

	Machine Learning Model						
		Training			Test		
	Model	R2	Lin's concordance	RMSE	R2	Lin's concordance	RMSE
SOM (%)	MLR	0.13	0.227	0.92	0.07	0.136	1.13
	SVM	0.519	0.591	0.776	0.018	0.138	0.868
	RF	0.743	0.79	0.501	0.11	0.135	1.11
	Hybrid Model						
		Training			Test		
	Model	R2	Lin's concordance	RMSE	R2	Lin's concordance	RMSE
SOM (%)	MLR-RK	1	0.984	1.29E-15	1	0.964	1.81E-15
	SVM -RK	0.796	0.869	0.44	0.932	0.934	0.3049
	RF -RK	1	0.984	2.00E-16	1	0.964	2.19E-16







CONCLUSIONS

The most important environmental variables in the model were the Normalized Clay Index and the Topsoil Grain Index of the Sentinel 2A image. Among the categorical environmental variables, the CLC-2018-212 Permanently irrigated land class is one of the important variables. ESA's accessible resources and hybrid modeling approaches improve spatial estimation of soil organic matter and can provide more accurate insights from the maps obtained

REFERENCES

FAO. (2019). Standard operating procedure for soil organic carbon Walkley-Black method Titration and colorimetric method. GLOSOLAN-SOP-02. Version number : 1 Page 1 of 25. Accessed Url: https://www.fao.org/3/ca7471en/ca7471en.pdf

Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B., & Gräler, B. (2018). Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ, 6, e5518.

Mponela, P., Snapp, S., Villamor, G., Tamene, L., Le, Q. B., & Borgemeister, C. (2020). Digital soil mapping of nitrogen, phosphorus, potassium, organic carbon and their crop response thresholds in smallholder managed escarpments of Malawi. Applied Geography, 124, 102299.

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R project.org/

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., ... & 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.