Soil organic carbon measurement and modelling at regional scale using Sentinel-1/2 data based on machine learning approaches

Smita Rani Swain*, Naveen Kumar Purushothaman and Bhabani Sankar Das

Agricultural and Food Engineering Department, Indian Institute of Technology, Kharagpur, India

Authors e-mail id: smitarani09@iitkgp.ac.in, nk1098@kgpian.iitkgp.ac.in, bsdas@agfe.iitkgp.ac.in

Introduction

- Remote sensors are emerging as powerful sources of soil spectral information.
- The freely available multispectral imagery is an effective solution to address the problems concerning hyperspectral images in soil properties monitoring
- Sentinel-2 (S2) launched by ESA has great potential in estimating soil properties compared to other multispectral images.
- Accurate prediction of soil properties is necessary to promote sustainable soil management.
- Multi-source data fusion technology is a promising way to improve prediction accuracy of models compared to single sensor technology.
- And application of models to real world problems such as crop and soil management with the help of high resolution soil properties mapping is helpful to the farming community.

Objectives

- To assess the prediction performance of various machine learning modeling approaches for the estimation of soil organic carbon using Sentinel-2 (S2) Multispectral imagery (MSI) data over the study area
- To evaluate the ability of multisource (Sentinel-1 and Sentinel-2) remote sensing data for the estimation of soil organic carbon

Material and Methods

Description of study area

- Study area: Western catchment of Chilika lagoon
- Latitude:19.732° N to 19·858° N
- Longitude: 85.113° E to 85.348° E
- Total area: 2712 km²
- No. of samples: 167 (0-15cm depth)
- Date of sampling: 18th- 22nd Dec, 2023
- Annual rainfall:1200 mm
- Climate: Tropical
- Maximum temperature: 39.9° C
- Minimum temperature: 14° C
- Main crop: Rice (kharif season)

Chemometric modelling

- Models used: RF, Cubist, SVR, and PLSR
- Spectral pre-processing techniques: un-transformed (UT), log absorbance (ABS), first derivative (FD), second derivative (SD), standard normal variate (SNV) and Kubelka- Munk transformation (KM)
- Software used for modelling: RStudio and MATLAB
- Sentinel-2 data acquisition dates: 13 dates from 2023 and 2024
- Sentinel-2 data acquisition date: 26th Dec, 2023

Figure 1. Study area map

Figure 2. Analytical flowchart of the Soil organic carbon estimation

Dates of MSI acquisition
1. 7th Apr. 2023
2. 12th Apr. 2023
3. 22nd May 2023
4. 28th Nov. 2023
5. 7th Jan 2024
6. 6th Feb 2024
7. 16th Feb 2024
8. 21st Feb 2024
9. 2nd Mar 2024
10. 12th Mar 2024
11. 16th Apr 2024
12. 26th Apr 2024
13. 11th May 2024

Figure 3. Soil organic carbon measurement by Walkley-Black chromic acid oxidation method (Akhter et al.,2023)

Modelling Approaches

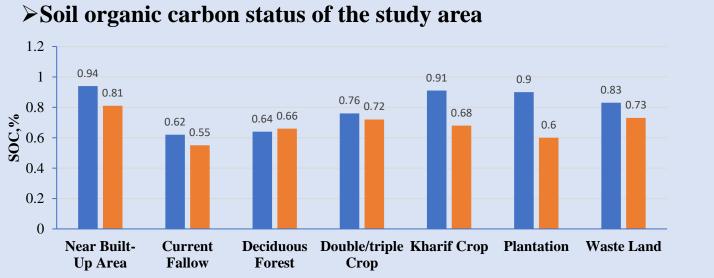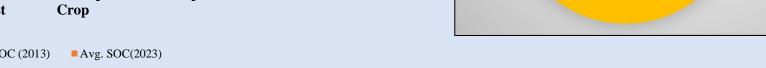| Scenario 1	Scenario 2	Scenario 3	Scenario 4
• Sentinel-2 spectral bands	• Sentinel-2 spectral bands • Sentinel-2 derived indices	• Sentinel-2 spectral bands • Topographic factors	• Sentinel-2 spectral bands • Topographic factors • Sentinel-1 backscatter coefficients

Topographic Factors

Topographic wetness index (TWI), Terrain ruggedness index (TRI), Slope, Relative slope position (RSP), LS-factor, Flow accumulation (FA), Elevation, Cross-sectional curvature, Aspect, Valley depth, Long curve, Hill shading, Channel network base level (CNBL), Convergence index (CI), Vertical distance to channel network (VDCN)
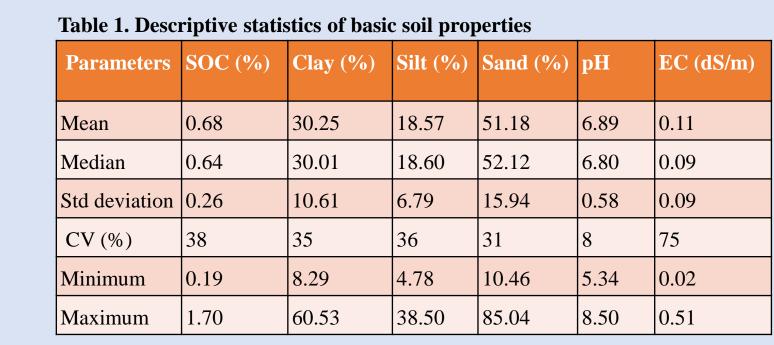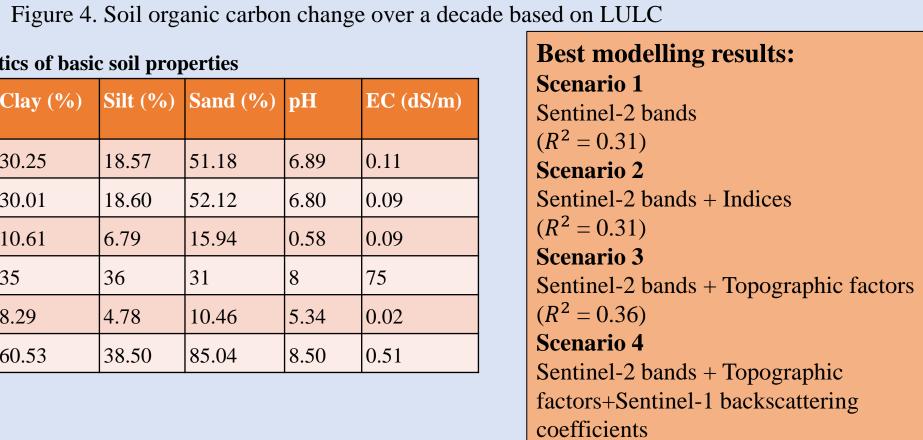
Results and Discussion

Soil organic carbon status of the study area

Figure 4. Soil organic carbon change over a decade based on LULC

Table 1. Descriptive statistics of basic soil properties

Parameters	SOC (%)	Clay (%)	Silt (%)	Sand (%)	pH	EC (dS/m)
Mean	0.68	30.25	18.57	51.18	6.89	0.11
Median	0.64	30.01	18.60	52.12	6.80	0.09
Std deviation	0.26	10.61	6.79	15.94	0.58	0.09
CV (%)	38	35	36	31	8	75
Minimum	0.19	8.29	4.78	10.46	5.34	0.02
Maximum	1.70	60.53	38.50	85.04	8.50	0.51

Best modelling results:

Scenario 1
Sentinel-2 bands
($R^2 = 0.31$)
Scenario 2
Sentinel-2 bands + Indices
($R^2 = 0.31$)
Scenario 3
Sentinel-2 bands + Topographic factors
($R^2 = 0.36$)
Scenario 4
Sentinel-2 bands + Topographic factors+Sentinel-1 backscattering coefficients
($R^2 = 0.38$)
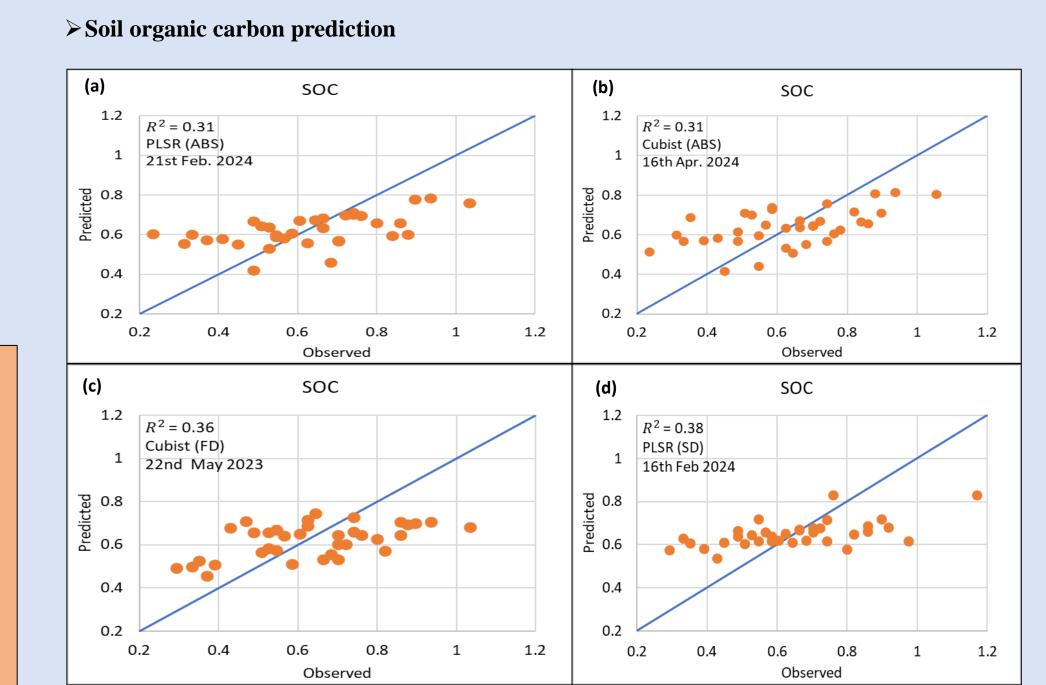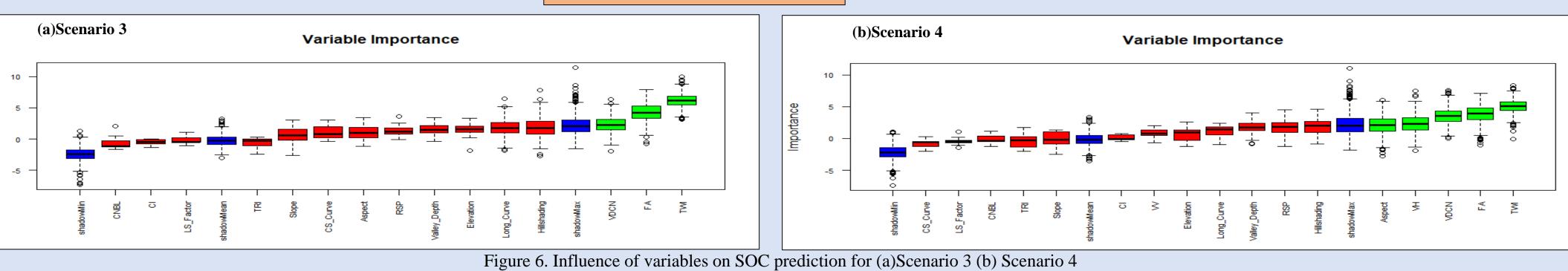
Soil organic carbon prediction

Figure 5. The top performing models of SOC using (a) S2 spectral bands (b) S2 spectral bands and indices (c)S2 spectral bands and environmental variables (d) S2 spectral bands, environmental variables and S1 bands

Figure 6. Influence of variables on SOC prediction for (a)Scenario 3 (b) Scenario 4

Conclusions

- This study highlights the applicability of multitemporal Sentinel-2 remote sensing for effective Soil organic carbon prediction
- The accuracy obtained using Sentinel-2 data combined with terrain attributes and Sentinel-1 data was improved slightly ($R^2 =0.38$) compared to that obtained using only Sentinel-2 data ($R^2 =0.31$), S2 data with spectral indices ($R^2 =0.31$) and Sentinel-2 data combined with terrain attributes ($R^2 =0.36$)
- Among all the models, Cubist and PLSR show comparatively better performance

Acknowledgement

- The authors would like to thank the Indian Institute of Technology, Kharagpur, West Bengal, India for the required lab and library facilities.
- The first and second authors acknowledge the Ministry of Human Resource Development, Government of India for providing PhD fellowships.
- The authors would also like to thank the Copernicus open access data hub (European space agency) for providing the required remote sensing data freely for conducting this study.

References

- Swain, S. R., Chakraborty, P., Panigrahi, N., Vasava, H. B., Reddy, N. N., Roy, S., & Das, B. S. (2021). Estimation of soil texture using Sentinel-2 multispectral imaging data: An ensemble modeling approach. *Soil and Tillage Research*, *213*, 105134.
- Vaudour, E., Gomez, C., Lagacherie, P., Loiseau, T., Baghdadi, N., Urbina-Salazar, D., & Arrouays, D. (2021). Temporal mosaicking approaches of Sentinel-2 images for extending topsoil organic carbon content mapping in croplands. *International Journal of Applied Earth Observation and Geoinformation*, *96*, 102277.
- Castaldi, F., Koparan, M. H., Wetterlind, J., Žydelis, R., Vinci, I., Savaş, A. Ö., ... & Vaudour, E. (2023). Assessing the capability of Sentinel-2 time-series to estimate soil organic carbon and clay content at local scale in croplands. *ISPRS Journal of Photogrammetry and Remote Sensing*, *199*, 40-60.