# Supplementary Information

## Mapping soil organic carbon with Soil Health Card observations and multisensor satellite data in Odisha, India

Pavan Kumar Bellam[1,2], Murali Krishna Gumma[1], Venkata Reddy Keesara[2], Mukund Dhavaji Patil[1], Narayanarao Bhogapurapu[2*]

[1]Digital Agriculture, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
[2]Department of Civil Engineering, National Institute of Technology Warangal, Warangal, India.

*Corresponding author(s). E-mail(s): bnarayanarao@nitw.ac.in;

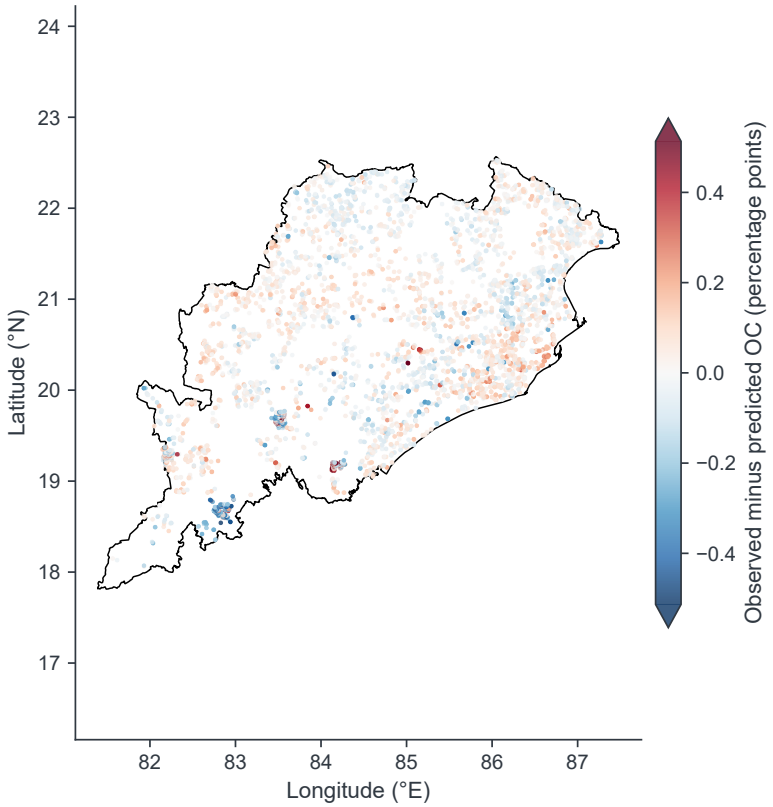
**Supplementary Table S1** The 17 environmental covariates retained in the random-forest model. TWI, topographic wetness index; TPI, topographic position index; SD, standard deviation; BIO12, mean annual precipitation; BIO15, precipitation seasonality; NIR, near-infrared; SWIR1, first short-wave infrared band; NDTI, normalised difference tillage index; HLS, Harmonized Landsat and Sentinel-2. Radar polarisation labels specify the transmitted and received polarisations: H, horizontal; V, vertical.

| Group | Retained predictors | Data source |
|---|---|---|
| Terrain | Elevation, slope, aspect, TWI, TPI, local elevation SD, local slope SD | MERIT DEM and derivatives |
| Climate | BIO12 mean annual precipitation, BIO15 precipitation seasonality | WorldClim v2.1 |
| Optical | Blue, NIR, SWIR1 and NDTI | HLS |
| C-band radar | VV, VH | Sentinel-1 |
| L-band radar | HH, HV | ALOS-2/PALSAR-2 |

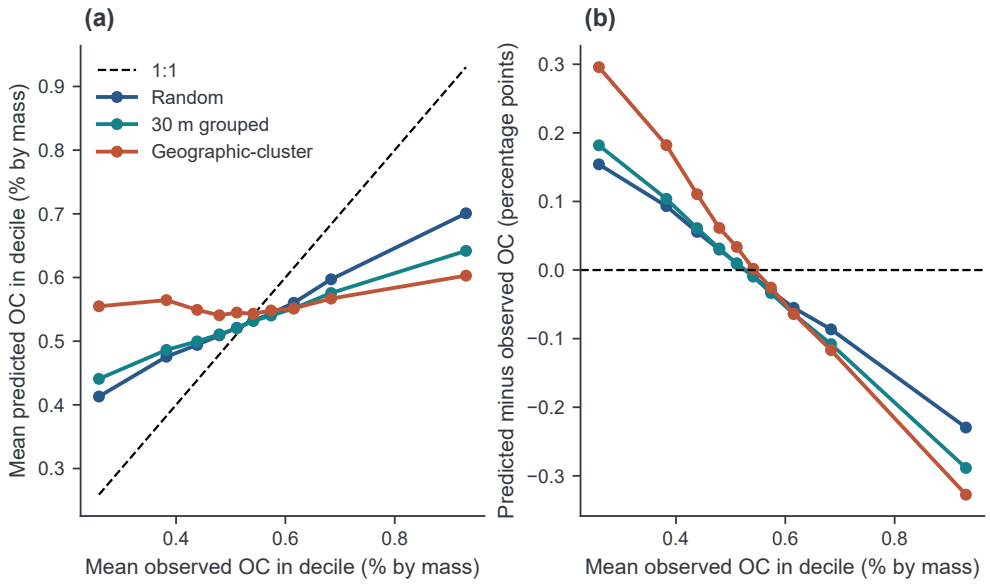
Carbon stocks were compared between the full mapped extent and cropland within the AOA under two Walkley–Black recovery assumptions (Table S2). These are analytical scenarios rather than uncertainty bounds.

**Supplementary Table S2** Scenario-based estimates of soil carbon stocks in the upper 15 cm by reporting domain and Walkley–Black recovery factor. AOA denotes area of applicability; Mha denotes million hectares; SD is the standard deviation of pixel-level carbon densities within each domain. The recovery factors define analytical scenarios rather than statistical uncertainty bounds.

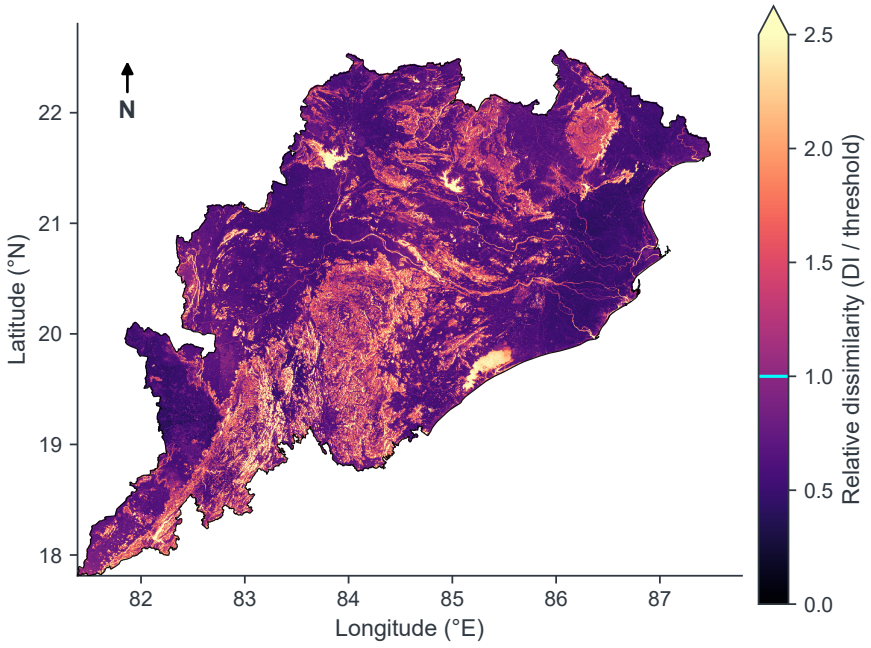
| Domain | Factor | Area (Mha) | Mean | SD (t C $ha^{-1}$) | Range | Total (Tg C) |
|---|---|---|---|---|---|---|
| Full mapped extent | 1.00 | 15.607 | 11.581 | 1.155 | 4.438–28.955 | 180.74 |
| Full mapped extent | 1.32 | 15.607 | 15.286 | 1.524 | 5.858–38.221 | 238.57 |
| Cropland within AOA | 1.00 | 5.317 | 11.064 | 0.915 | 4.438–25.843 | 58.83 |
| Cropland within AOA | 1.32 | 5.317 | 14.604 | 1.208 | 5.858–34.113 | 77.66 |



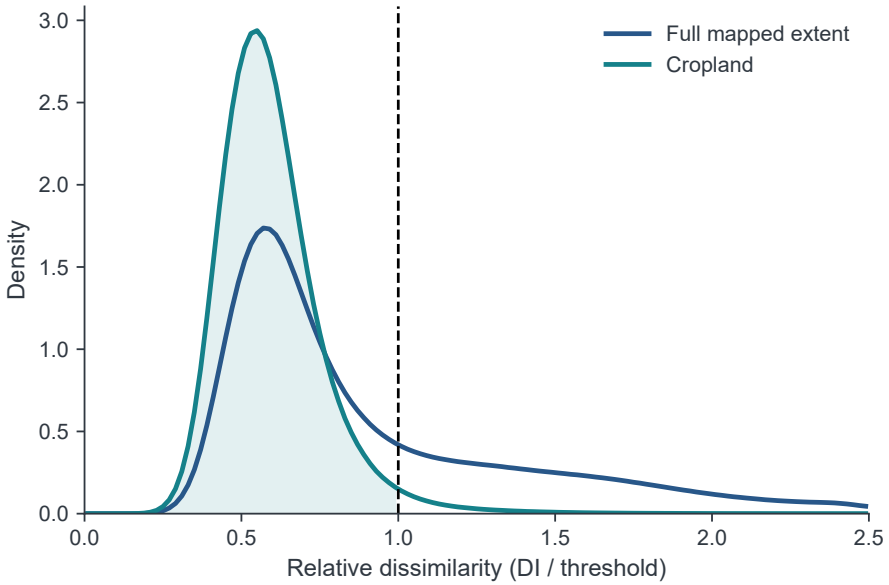

**Supplementary Fig. S1** Residuals at observations excluded from model fitting under geographic-cluster validation. Residuals are observed minus predicted OC; positive values indicate underprediction and negative values indicate overprediction.



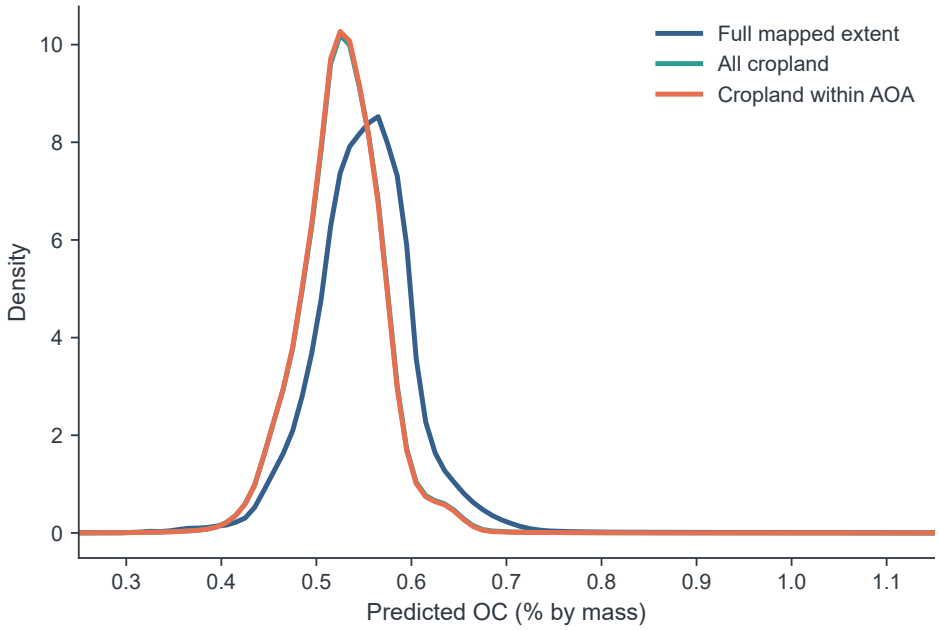

**Supplementary Fig. S2** Conditional bias by observed-OC decile under random, 30 m grouped and geographic-cluster five-fold validation. (a) Mean predicted versus mean observed OC within each decile; the dashed line indicates 1:1 agreement. (b) Mean prediction bias, expressed as predicted minus observed OC in percentage points. Positive values indicate overprediction and negative values indicate underprediction.




**Supplementary Fig. S3** Relative dissimilarity across Odisha, calculated as the dissimilarity index (DI) divided by its applicability threshold. Values at or below 1 fall inside the area of applicability (AOA); values above 1 indicate greater dissimilarity from the model-development observations.




**Supplementary Fig. S4** Relative-DI distributions for the full mapped extent and cropland. The dashed line at relative DI = 1 marks the AOA boundary, corresponding to an absolute DI threshold of 0.529. The displayed range ends at relative DI = 2.5.




**Supplementary Fig. S5** Distributions of predicted OC across the full mapped extent, all cropland and cropland inside the area of applicability.