Vol. 51, No. 2, pp 154-163, 2023
Indian Journal of Soil Conservation
https://iaswc.com/ijsc-2023.html

Mapping of soil organic carbon stock and carbon sequestration potential in Vemagal Hobli, Kolar district, Karnataka, India

G.K. Harikaran[1], S. Dharumarajan[2,*], R. Vasundhara[2], M. Lalitha[2], B. Kalaiselvi[2], S. Parvathy[2] and G. Ushakiran[2]

[1]*Department of Soil Science and Agricultural Chemistry, University of Agricultural Sciences, Bangalore;* [2]*ICAR-National Bureau of Soil Survey and Land Use Planning, Regional Centre, Bangalore.*

**Corresponding author:*
E-mail: sdharmag@gmail.com (S. Dharumarajan)

ARTICLE INFO

doi: 10.59797/ijsc.v51.i2.131

Article history:
Received : December, 2022
Revised : June, 2023
Accepted : August, 2023

Key words:
Carbon sequestration potential
Digital soil mapping
Machine learning
Random forest
Soil organic carbon stock

ABSTRACT

Soil organic carbon (SOC) pool has declined to a greater extent due to intensive tillage practices, continuous use of chemical fertilizers, removal of crop residues, etc. This creates the need of increasing SOC stocks and to improve carbon sequestration potential (CSP). It is important to know the spatial distribution of SOC stocks and CSP, which can be achieved with the help of digital soil mapping (DSM). In the present study, CSP was calculated based on the maximum potential of fine fraction (clay + silt), which can store the carbon to the actual carbon associated with the fine fraction. The mean observed SOC stock and CSP in different depths were in the range of 0.83-1.54 kg m^{-2} and 5.52-6.51 kg m^{-2}, respectively. Four machine learning (ML) algorithms [random forest (RF), support vector machine (SVM), cubist and artificial neural network (ANN)] were compared for prediction of SOC stock and CSP using 99 soil profiles data collected in Vemagal Hobli, Kolar district, Karnataka. The model performances were evaluated using uncertainty indicators such as coefficient of determination (R^2) and root mean squared error (RMSE). RF model performed better than SVM, cubist and ANN by explaining variability of 28-32% and 28-33% in prediction of SOC stock and CSP, respectively. The predicted SOC stock and CSP in 100 cm soil profile were found to be 3.0-5.8 kg m^{-2} and 19.1-27.3 kg m^{-2}, respectively. These digital maps help in understanding the spatial distribution of SOC stock and CSP in the study area.

1. INTRODUCTION

The element carbon plays an essential role in the survival of all life forms on earth. It is vital for all life forms to take in and release carbon, whether for the manufacture of food or for respiration (FAO, 2016). The carbon pool in soil makes up the largest portion of the carbon pool in the atmosphere. The soil carbon pool is approximately 3.1 times larger than the atmospheric pool of 800 GT (Oelkers and Cole, 2008). It is estimated that one-third of the world's soil resources have been degraded. If these soils are restored, they will remove about 63 billion tonnes of carbon from the atmosphere (FAO, 2016). Combustion of fossil fuels and change in land use patterns are the principal anthropogenic activities that influence global carbon cycle (Adhikari *et al.*, 2020). About 20 Pg carbon (C) can be sequestered by soil in 25 years, which is more than 10% of the anthropogenic emissions (FAO, 2016). As per Intergovernmental Panel on Climate Change (2007), the top 30 cm of soil contains about 697 Pg C and the top 1 m contains about 1500 Pg C (Bernstein *et al.,* 2008). The distribution of SOC stock in the ecosystem is highly influenced by the change in land use-land cover (Sharma and Sharma, 2022). Intensive tillage practices, continuous use of chemical fertilizers, removal of crop residues results in rapid mineralization which causes the depletion of SOC pool (Sarma *et al.*, 2013). SOC plays an important role in maintaining soil fertility and in sustaining the productivity of agro-ecosystems (Das *et al.*, 2019). It improves soil structure, hydraulic properties, aeration, water holding capacity, nutrient holding capacity and microbial activities. Hence, it is important to increase SOC stocks and to improve soil organic CSP. An important aspect of SOC management is to know the spatial distribu-

tion of carbon stocks and CSP which can be achieved with the help of DSM.

This DSM technique creates quantitative connections between soil properties using a collection of environmental factors and then regression equations are created to predict soil characteristics. These links can be combined to produce a computerised soil map of the study site (McBratney *et al.*, 2003). The digital soil map is a raster made up of two-dimensional cells (pixels) arranged in a grid, each of which is assigned a specific geographic location and is filled with soil information (Carré *et al.*, 2007). This approach yields accurate quantitative soil information at a faster rate through large scale dynamic monitoring of soil resources and by considering both spatial and temporal variability of soil properties. It makes use of different machine learning (ML) algorithms for accurate prediction of various soil properties. These algorithms can be divided into supervised or unsupervised algorithms. In supervised learning techniques, both inputs and outputs are labelled in the process of training the model, whereas in unsupervised learning techniques, only the input is fed for training the model. The model will identify the hidden patterns and gives the desired output.

Several researchers carried out the digital mapping of SOC stock and organic CSP by prediction using different machine learning algorithms. Ließ *et al.* (2016) carried out the spatial prediction of SOC stocks in a complex tropical mountain landscape by comparing five ML algorithms (RF, ANN, multivariate adaptive regression splines, boosted regression tree (BRT) and SVM) using five times repetition of a tenfold cross-validation. They revealed that BRT algorithm resulted in the overall best model with a mean RMSE of 0.12%. Rostaminia *et al.* (2021) predicted SOC stocks in Gavi plain located in the western Iran by comparing the performance of RF, cubist (Cu) along with random forest-ordinary kriging (RF-OK) and cubist-ordinary kriging (Cu-OK) hybrid ML models using 10-fold cross validation. They revealed that RF-OK showed better performance ($R^2 = 0.75$, RMSE = 6.33 t ha^{-1}). Sreenivas *et al.* (2016) carried out estimation and mapping of the SOC and soil inorganic carbon (SIC) densities up to 100 cm depth or paralithic contact of whole India using RF model at 250 m spatial resolution. Using quantile regression forest, Dharumarajan *et al.* (2021) mapped the SOC stock of the western ghats over an area of 56,763 km^2 and found that SOC stock prediction performed better for the top layer (R^2 = 31-43%) followed by a decline with depth (R^2 = 7-21%).

Weismeier *et al.* (2014) estimated the CSP of soils in south-east Germany by calculating the potential SOC saturation of silt and clay particles based on 516 soil profiles. The findings indicated that farmland soils might hold significant amounts of extra SOC despite having a low degree of carbon saturation of just about 50% and reasonably high CSP was found in grassland soils, on the other hand, forest soils, were nearly saturated in carbon and had a low potential for carbon sequestration. Hammad *et al.* (2020) compared the CSP of various land use systems such as forests land, croplands, agroforests and orchards in the arid region of Pakistan at different depths (0-20, 20-40, 40-60 and 60-80 cm). They revealed that the highest CSP (64.54 Mg ha^{-1}) noticed in forest land and the lowest was noticed in crop land (33.50 Mg ha^{-1}). Padarian *et al.* (2022) reported that SOC storage potential in the topsoil of global croplands is ranged from 29 to 65 Pg C.

Most of the works across the globe focussed only on the estimation of soil organic CSP rather than mapping of its spatial distribution. In this context, the present study was carried out in Vemagal Hobli, Kolar district of Karnataka with an aim to compare four ML algorithms (RF, SVM, cubist and ANN) for prediction of SOC stock and organic CSP and to map them along with uncertainty in four different depths (0-25, 25-50, 50-75 and 75-100 cm) using best ML algorithm at 30 m resolution.

2. MATERIALS AND METHODS

Study Area

The study area constitutes Vemagal Hobli (block), which is located in Kolar district, Karnataka. Location map of the study area is shown in Fig. 1. It covers an area of about 13,948 ha and is geographically located at 13°9'27.47"N to 13°18'17.13"N latitudes and 77°56'5.32"E to 78°5'10.43"E longitudes. It falls under eastern dry zone of Karnataka according to National Agriculture Research Project (NARP). According to ICAR agro ecological sub region, it lies in Central Karnataka plateau, hot, moist, semi-arid eco-sub region. Major geology in the region is granite and granite-gneiss. The elevation is ranging from 820 to 1100 m above the mean sea level. The maximum temperature in the study area ranging from 30 to 36°C and minimum temperature ranging from 15 to 22°C. The area receives rainfall of 600 to 700 mm annually. Major soil orders in the study area are *Alfisols* and *Inceptisols*. Area is classified under ustic soil moisture regime and iso hyperthermic soil temperature regime. The main cropping season is *kharif* and major crops of the study area are mango (*Mangifera indica*), finger millet (*Eleusine coracana*), red gram (*Cajanus cajan*), tomato (*Solanum lycopersicum*), mulberry (*Morus albus*), etc.

Soil Sampling and Analysis

High resolution satellite imagery (Sentinel-2) was used for identifying the soil profile locations. Field examination was carried out from 28.12.2021 to 19.02.2022 and about 46 profile locations were identified for study. Profile locations were selected in transects based on variations in physiogra-

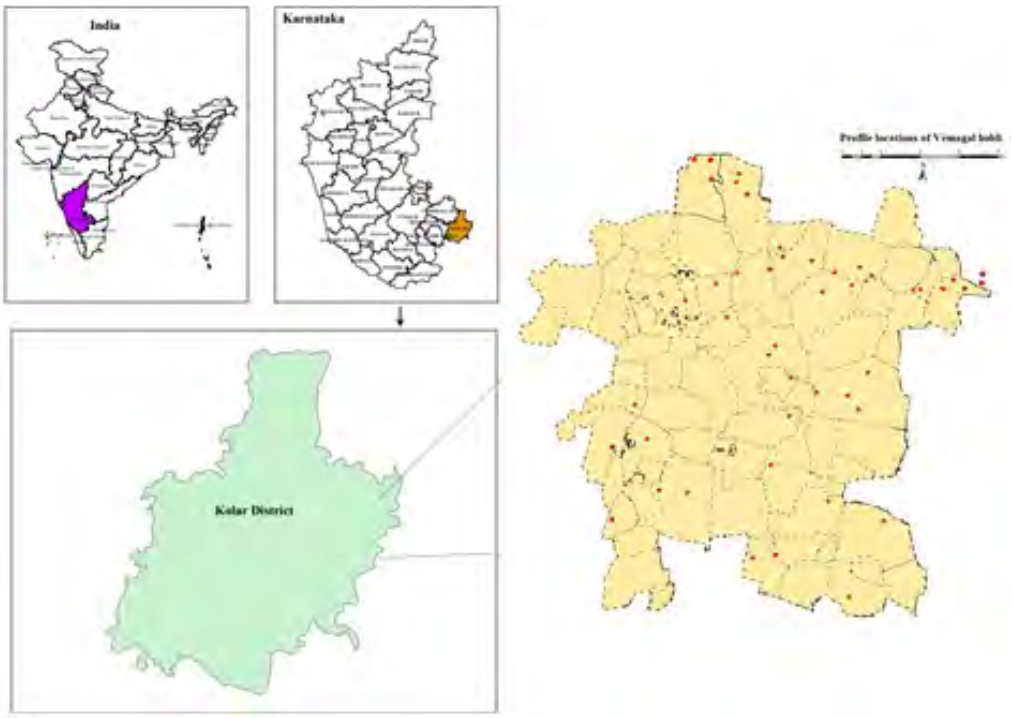

Fig. 1. Location map of study area - Vemagal hobli, Kolar district, Karnataka, India

phy and cropping pattern of the study area. These soil profiles were studied to the depth of 200 cm or to the depth limited by bedrock. Representative horizon wise samples from each soil profile were collected covering the variability of the study area. A total of 183 samples were collected from 46 identified profiles. Samples, to be determined for bulk density, had been collected in a metallic core of known volume, which was inserted in the soil horizon. The collected samples were shade dried in laboratory and then grounded using wooden pestle and mortar. After grinding, the samples were passed through a 2 mm sieve to separate coarse fragments (gravel, pebbles, roots, etc). For organic carbon (OC), the soil samples were further grinded finer and then sieved using 0.2 mm sieve.

The processed samples were analysed in laboratory using standard analytical methods. Bulk density was determined gravimetrically by using core method. SOC was determined by wet digestion method (Walkley and Black, 1934). Particle size analysis was carried out by the international pipette method (Jackson, 1973).

SOC Stock

SOC stock was estimated for each layer of the soil profile by using the equation below (Dharumarajan *et al.*, 2021).

$$SOC\,stock\,(kg\,m^{-2}) = (OC\,\%/100) \times BD \times D \times (1-G) \times 10.$$

Where, *SOC stock* = soil organic carbon stock in kg m^{-2}; *OC* = organic carbon in %; *BD* = bulk density in Mg m^{-3}; *D* = depth of each horizon in cm; *G* = % Gravel content / 100.

Soil Organic CSP

CSP was calculated by using the equations given below:

Potential C saturation (C_{sat}) = $4.09 + 0.37 \times (clay\,(\%) + silt\,(\%)$ (Hassink, 1997).

C saturation deficit ($C_{sat\text{-}def}$) = $C_{sat} - C_{cur}$

Where, $C_{sat\text{-}def}$ = *C* saturation deficit (mg g^{-1}); C_{sat} = Potential *C* saturation (mg g^{-1}); C_{cur} = Current *C* concentration of the clay and silt fraction.

$$CSP\,(kg\,m^{-2}) = C_{sat\text{-}def} \times BD \times (1-G) \times D \times 10^{-2}$$

Where, *CSP* = Carbon sequestration potential (kg m^{-2}); $C_{sat\text{-}def}$ = C saturation deficit (mg g^{-1}); *BD* = Soil bulk density (Mg m^{-3}); *D* = Depth of the sampled soil layer (cm); *G* = % Gravel content / 100.

Preparation of Basic Satasets

In addition to 46 soil profiles studied during the survey, legacy datasets from 53 profiles of study area were also collected from National Bureau of Soil Survey and Land Use Planning (NBSS&LUP), Regional Centre, Hebbal, Bangalore. The data collected during present study was compiled with the legacy data. A total of about 99 profiles were considered for prediction of SOC stock and CSP.

Table: 1
Different covariates used for the prediction of soil properties

Covariates	Source	Resolution
Sentinel - 2 bands (b2-b12)	11 bands	10-60 m
Elevation (m)	SRTM DEM	30 m
Aspect	SRTM DEM	30 m
Channel network base level	SRTM DEM	30 m
Channel network distance	SRTM DEM	30 m
Analytical hill shading	SRTM DEM	30 m
Closed depressions	SRTM DEM	30 m
Convergence index	SRTM DEM	30 m
MrRTF	SRTM DEM	30 m
MrVBF	SRTM DEM	30 m
Plan curvature	SRTM DEM	30 m
Profile curvature	SRTM DEM	30 m
Relative slope position	SRTM DEM	30 m
Slope	SRTM DEM	30 m
TPI	SRTM DEM	30 m
TWI	SRTM DEM	30 m
Total catchment	SRTM DEM	30 m
Valley depth	SRTM DEM	30 m
NDVI	MOD13Q1	250 m_16 days
EVI	MOD13Q1	250 m_16 days
Land surface temperature	MOD11A1	1 km
Precipitation (mm)	WorldClim 2 data	30 seconds
Average temperature (°C)	WorldClim 2 data	30 seconds

MrRTF - Multi-resolution index of Ridge Top Flatness, MrVBF - Multi-resolution index of Valley Bottom Flatness, TPI - Topographic Position Index, TWI - Topographic Wetness Index, NDVI - Normalized Difference Vegetation Index, EVI - Enhanced Vegetation Index, SRTM DEM - Shuttle Radar Topography Mission Digital Elevation Model

Depth wise (0-25 cm, 25-50 cm, 50-75 cm and 75-100 cm) distribution of SOC stock and organic CSP were assessed by using weighted average method.

Environmental Attributes Considered

Different set of environmental covariates (Table 1) were used for prediction of SOC stock and organic CSP. Digital elevation model (DEM) was obtained from shuttle radar topography mission (SRTM) having 30 m resolution. The derivatives of DEM like multi-resolution index of valley bottom flatness (MrVBF), multi-resolution index of ridge top flatness (MrRTF), topographic position index (TPI), topographic wetness index (TWI), aspect, channel network base level, channel network distance (CND), analytical hill shading, closed depressions, convergence index, plan curvature, profile curvature, relative slope position (RSP), slope, total catchment and valley depth were derived by using system of automated geoscientific analyses *i.e.* Saga-GIS ver. 2.3.1. Normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) were obtained from moderate resolution imaging spectroradiometer (MODIS) data (MOD13Q1). Land surface temperature (LST) imagery was obtained from MODIS data (MOD11A1). Sentinel 2 bands namely band 2 to band 12 were also used as variable input. Precipitation and average temperature datasets were obtained from world climate version 2 data. All of these covariates were resampled to a spatial resolution of 30 m and also intersected to the sampling points for the prediction of SOC stock and organic CSP.

Modelling and Mapping

For the present study, four ML algorithms - RF, SVM, cubist and ANN were compared for the prediction of SOC stock and organic CSP.

RF model is an extension of regression tree model which works based on assemblage of a number of classification and regression trees by means of two levels of randomization for each tree in the forest (Breiman, 2001). It is the most used model at present, since it improves the prediction accuracy and reduces model over fitting and it is non-sensitive to missing data and has capacity to handle large number of both quantitative and categorical data (Dharumarajan *et al.*, 2017). The number of tree (n_{tree}), minimum number of samples at terminal node (n_{min}) and number of predictors used for fitting the tree (M_{try}) will decide the fitting of model. For the present study, random forest package (ver. 4.6-14) was used for running the RF model in R environment. This package helps in performing classification and regression operations.

SVM maintains all covariates to define a maximal margin which is the margin of tolerance using the support vectors (observations) and to separate or fit data linearly. The margin is the distance from the decision surface which ensures high generalization ability of the algorithm, making the results more applicable to the unseen data (Pradhan, 2013). In addition, this approach applies kernel functions to map non-linear vectors to a very high dimensional space for solving non-linear problems. For performing statistical functions, e1071 package (ver. 1.7-9) was used in R environment (Cortes and Vapnik, 1995).

Cubist is a rule-based algorithm that has recently increased in popularity among digital soil mappers (Miller *et al.*, 2015). It initially creates a tree structure from a pool of provided covariates and then it collapses paths through the tree to create rules using boosting training. Each rule contains a multiple linear regression (MLR) model for predicting the target variable under the conditions of the respective rule. The final model located at the terminal nodes shows a collection of MLR models for calculating predicted values. In addition, this ensemble model adds boosting to improve the prediction accuracy. Ensemble learning combines models produced by multiple repetitions of the same algorithm which usually obtains stronger

predictive performance than results produced from any of the models individually. In the present study, Cubist package (ver. 0.3.0) was used for running the cubist model in R environment. This package performs rule and instance-based regression modeling.

ANN algorithm was inspired from the functioning of human brain through a network of neurons, this algorithm is composed of connections of components operating in parallel. It can be trained to perform a specific function and obtain certain output by adjusting the values of the connection weights between the components. The network is adjusted, based on a comparison of the secured output and targeted outcome, till the sum of square differences between them become minimum. The attractiveness of ANN comes from the remarkable information processing characteristics of the biological system such as non-linearity, high parallelism, robustness, fault and failure tolerance, learning, ability to handle imprecise and fuzzy information and their capability to generalize. An ANN can be regarded as a class of universal approximators that implements a non-linear, input-output mapping of a general nature (Prieto *et al.*, 2013). For training the ANN model in R environment, neuralnet package (ver. 1.44.2) was used.

All the 4 models considered for study were validated using 10-fold cross validation techniques with 20 times repetition. The predictive performance of algorithms was evaluated using uncertainty indicators such as coefficient of determination (R^2) and RMSE. For good models, R^2 should be equal or close to 1, whereas RMSE should be low or near to 0.

$$\text{Coefficient of determination } (R^2) = 1 - \frac{\sum_{i=1}^{n}(pi - oi)^2}{\sum_{i=1}^{n}(oi - \overline{oi})^2}$$

$$\text{Root Mean Squared Error (RMSE)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(oi - pi)^2},$$

Where, *pi* and *oi* are predicted and observed values, respectively.

Uncertainty analysis was performed at 90% confidence interval using quantile random forest (QRF) model. The prediction uncertainty was evaluated by using prediction interval coverage percentage (PICP). A 30 m grid was created and all the environmental covariates were extracted for that grid. Among different algorithms used for study, the best algorithm was used to predict SOC stock and organic CSP at 30 m grid. Predicted properties were then mapped using point to raster tool of Arc-GIS ver. 10.7.1.

3. RESULTS AND DISCUSSION

Summary Statistics

Table 2 shows the summary statistics of observed soil bulk density and textural fractions. The mean observed BD was found to increase with depth with coefficient of variation ranging from 10.13 to 12.50%. Higher variability in bulk density was observed in 50-75 cm depth. The mean of observed sand and silt content was found to decrease and increase with depth, respectively. The mean clay content exhibited an irregular trend with depth. The kurtosis was found to be positive in all depths in case of sand and clay content. The sand content was negatively skewed with depth, whereas silt and clay content were found to be positively skewed with depth. The variability in silt content was found to be decreasing with depth.

The mean observed SOC stock was found to be 1.54, 1.08, 0.93 and 0.83 kg m^{-2} in 0-25, 25-50, 50-75, 75-100 cm

Table: 2
Summary statistics of observed soil bulk density and textural fractions

Property	Depth (cm)	Min	Max	Mean	SD	Kurtosis	Skewness	CV (%)
BD (Mg m^{-3})	0-25	1.08	1.59	1.58	0.16	1.07	-0.12	10.13
	25-50	1.24	1.65	1.59	0.17	-0.79	-0.32	10.69
	50-75	1.28	1.70	1.60	0.20	2.07	0.74	12.5
	75-100	1.40	1.75	1.61	0.18	-0.71	0.12	11.18
Sand (%)	0-25	21.17	85.73	59.66	12.02	0.58	-0.74	20.15
	25-50	10.73	73.64	49.51	10.88	2.06	-0.63	21.98
	50-75	10.85	70.40	46.55	10.16	2.41	-0.84	21.83
	75-100	11.35	64.57	45.70	10.53	1.56	-0.74	23.04
Silt (%)	0-25	2.50	36.79	13.51	7.25	0.18	0.62	53.66
	25-50	2.20	35.47	14.37	6.91	-0.12	0.48	48.09
	50-75	3.75	36.97	15.03	6.96	0.13	0.44	46.31
	75-100	3.48	36.97	15.56	7.07	0.08	0.35	45.44
Clay (%)	0-25	5.17	61.69	26.79	11.34	0.43	0.44	42.33
	25-50	7.09	65.24	36.14	11.27	0.70	0.32	31.18
	50-75	18.34	68.72	38.20	10.98	0.51	0.70	28.74
	75-100	10.18	67.80	37.98	11.65	1.11	0.29	30.67

Table: 3
Summary statistics of observed SOC stock and CSP

Property	Depth (cm)	Min	Max	Mean	SD	Kurtosis	Skewness	CV (%)
SOC stock (kg m^{-2})	0-25	0.35	4.64	1.54	0.97	2.17	1.54	62.99
	25-50	0.23	3.62	1.08	0.67	4.73	1.87	62.04
	50-75	0.15	2.99	0.93	0.61	3.34	1.63	65.59
	75-100	0.15	2.92	0.83	0.59	2.70	1.54	71.08
CSP (kg m^{-2})	0-25	2.09	8.37	5.52	1.44	-0.13	-0.26	26.09
	25-50	2.27	9.92	6.51	1.90	-0.62	-0.37	29.19
	50-75	2.61	10.10	6.38	1.89	-0.54	-0.41	29.62
	75-100	2.35	9.70	5.98	1.87	-0.76	-0.04	31.27

soil depths, respectively. It was found to decrease with an increase in soil depth. The observed SOC stock exhibited positive kurtosis and skewness at all depths. The variation in carbon stock was found to be higher in 75-100 cm depth. On the other hand, the mean observed CSP showed an irregular trend with depth ranging from 5.52 to 6.51 kg m^{-2}. The kurtosis and skewness were found to be negative at all depths. The variability in CSP was found to increase with depth (Table 3).

Table 4 shows the correlation analysis between SOC stock, CSP and environmental covariates. SOC stock and CSP were found to be significantly correlated with CND, convergence index, plan curvature at 5% level. CSP was found to exhibit significant negative correlation with RSP at 5% level. Strong positive correlation (0.36*) was observed between CSP and convergence index.

Comparison of Different Algorithms

SOC stock and CSP were predicted using different ML algorithms. The depth wise performance of four different algorithms in prediction of SOC stock and CSP is shown in Table 5. For prediction of SOC stock and CSP, the order of performance of algorithms is found to be RF > SVM > Cubist > ANN and RF > Cubist > SVM > ANN, respectively with respect to R^2. The dynamic nature and amount of OC in the study area determines the prediction performance of OC stock. The performance of RF in predicting SOC stocks has been slightly reduced in sub-surface layer which is indicated by decrease of R^2 from 32 to 28%. Dharumarajan *et al.* (2021) also found that the model's performance is reduced with an increase in depth (R^2 = 7-43%). The RF model predicted CSP better with R^2 and RMSE of 28-33% and 0.59-1.77 kg m^{-2} (Table 5). The prediction performance of CSP depends upon the distribution of SOC stock in the study area.

RF model outperformed other three models based on uncertainty indicators - R^2 and RMSE in prediction of SOC stock and CSP. Apart from this aspect, it also improved the accuracy of the prediction compared to other models

Table: 4
Correlation analysis between surface SOC stock, CSP and environmental covariates

Properties / Covariates	SOC stock	CSP
B2	-0.01	-0.01
B3	-0.01	-0.01
B4	-0.04	-0.04
B5	-0.05	-0.05
B6	-0.02	-0.02
B7	-0.04	-0.04
B8	0.10	0.11
B9	0.19	0.19
B10	-0.06	-0.06
B11	-0.09	-0.09
B12	-0.07	-0.07
NDVI	-0.01	-0.01
EVI	0.07	0.07
LST	-0.01	-0.01
DEM	-0.05	-0.05
Aspect	-0.05	-0.05
CNBL	0.09	0.09
CND	-0.31*	-0.31*
AHS	-0.08	-0.08
CD	-0.16	-0.02
CI	0.30*	0.36*
MrRTF	-0.01	0.10
MrVBF	-0.25	-0.08
Plan curvature	0.32*	0.33*
Profile curvature	0.25	0.16
RSP	-0.04	-0.33*
Slope	0.17	0.01
TPI	0.17	0.09
TWI	-0.08	0.05
Total catchment	-0.11	0.1
Valley depth	-0.02	0.04
Precipitation	0.03	-0.07
Average temperature	-0.07	-0.10

**Significant correlation at 5% level*

B2-B12: Bands of sentinel-2 satellite imagery, NDVI: Normalised Difference Vegetation Index, EVI: Enhanced Vegetation Index, LST: Land Surface Temperature, DEM: Digital Elevation Model, CNBL: Channel Network Base Level, CND: Channel Network Distance, AHS: Analytical Hill Shading, CD: Closed Depressions, CI: Convergence Index, MrRTF: Multi-resolution index of Ridge Top Flatness, MrVBF: Multi-resolution index of Valley Bottom Flatness, RSP: Relative Slope Position, TPI: Topographic Position Index, TWI: Topographic Wetness Index

Table: 5
Performance of different models for prediction of SOC stock and CSP

Properties	Depth (cm)	SVM		RF		Cubist		ANN	
		R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
SOC stock (kg m^{-2})	0-25	0.28	0.87	0.32	0.85	0.22	0.98	0.24	1.00
	25-50	0.27	0.66	0.31	0.59	0.29	0.61	0.23	0.73
	50-75	0.30	0.63	0.31	0.56	0.27	0.58	0.26	0.67
	75-100	0.25	0.59	0.28	0.53	0.26	0.56	0.23	0.66
CSP (kg m^{-2})	0-25	0.25	1.44	0.28	1.42	0.17	1.53	0.22	1.65
	25-50	0.25	1.85	0.31	0.59	0.29	0.61	0.23	0.73
	50-75	0.21	1.87	0.33	1.77	0.31	1.91	0.25	1.93
	75-100	0.28	1.71	0.31	1.69	0.24	1.89	0.25	1.93

Table: 6
Summary statistics of predicted OC stock and CSP

Property	Depth (cm)	Min	Max	Mean	SD	Kurtosis	Skewness	CV (%)
SOC stock (kg m^{-2})	0-25	0.91	2.51	1.56	0.27	-0.43	0.14	17.31
	25-50	0.68	1.76	1.04	0.13	6.08	1.11	12.50
	50-75	0.42	1.20	0.69	0.12	2.19	1.00	17.39
	75-100	0.25	0.98	0.48	0.08	5.89	1.94	16.67
CSP (kg m^{-2})	0-25	4.39	6.55	5.54	0.25	3.38	-0.41	4.51
	25-50	4.88	8.07	6.52	0.61	-0.91	-0.38	9.36
	50-75	4.19	7.54	6.02	0.51	0.59	-0.08	8.47
	75-100	4.47	6.83	5.58	0.22	3.34	0.34	3.94

(Breiman, 2001; Dharumarajan and Hegde, 2022). Akpa *et al.* (2016) found that the RF model performed better when compared to cubist and BRT in prediction of SOC stock. Gomes *et al.* (2019) also reported better predictive performance of RF model by comparison with cubist, generalized linear model boosting and SVM in prediction of SOC stock. Uncertainty can be indicated by an indicator called PICP which indicates % of all observed values fitting within their prediction limits. Commonly used confidence interval for PICP is 90% which means that the true value can be found 9 out of 10 times within the range of values (Dharumarajan *et al.*, 2020). The PICP values were derived from QRF model. The PICP range of OC stock and CSP in the study area was found to be 83.6 to 87.1 and 84.5 to 86.3, respectively.

Environmental covariates play an important role in influencing the soil properties and in improving the prediction accuracy. Based on how accurate or inaccurate the prediction would be if one or more covariates were removed from the model, the RF model predicts the relative importance of various covariates (Prasad *et al.*, 2006). In case of prediction of OC stock, band 4 of sentinel-2 imagery and TPI were found as most important covariates. Hilly and mountainous areas have more OC stock due to the presence of dense forests which contributes to more biomass in the soil. Dharumarajan *et al.* (2021) found elevation as the most important variable for surface prediction of OC stock, whereas Sreenivas *et al.* (2016) found land cover as the most important covariate for the prediction of OC stock. Akpa *et al.* (2016) found soil type, climate, vegetation indices and terrain attributes as important covariates in prediction of SOC stock. Convergence index was found to be the top-most covariate in prediction of CSP. Gray *et al.* (2022) found that a combination of climate, parent material and vegetation cover influences soil CSP to a greater extent.

Prediction of SOC stock and CSP

Table 6 shows the summary statistics of predicted OC stock and CSP. The mean value of predicted OC stock was found to decrease with increase in depth (1.56 to 0.48 kg m^{-2}). The kurtosis was found to be negative at 0-25 cm depth, whereas SOC stocks at other depths shown positive kurtosis. The SOC stocks exhibited positive skewness at all depths. Higher OC stock in surface layer may be due to slow decomposition of organic matter because of reduced soil temperature and also due to more addition of biomass through litter deposition (Sreenivas *et al.*, 2016; Srinivasan *et al.*, 2019; Dharumarajan *et al.*, 2021). The mean value of predicted CSP ranged from 5.54 to 6.52 kg m^{-2} and the coefficient of variation was found to be higher (9.36%) in 25-50 cm depth. The kurtosis was found to be negative in 25-50 cm depth and CSP was positively skewed in 75-100 cm depth. Soils with lower OC stock generally have higher potential to sequester further carbon in them. Higher CSP

might also be due to the presence of more clay and silt fractions (Wiesmeier *et al.*, 2014).

The predicted and uncertainty OC stock and CSP maps at 0-100 cm depth are depicted in Fig's. 2 and 3, respectively. The predicted SOC stock and CSP for 100 cm soil profile were found to be 3.0-5.8 kg m^{-2} and 19.1-27.3 kg m^{-2}, respectively. From the predicted map of CSP, it was found that soils of hilly areas exhibited lower CSP. This may be due to the presence of higher SOC stocks. Uncertainty was found to be moderate in prediction of SOC stock. In case of prediction of CSP, moderate to high uncertainty was noticed. The reason for high uncertainty might be due to the existence of limited datasets and due to different landforms prevailing in the study area.

4. CONCLUSIONS

From the study, it is clear that the SOC stock is influenced by SOC content, bulk density, gravel content and soil depth. Land use system and cultivation practices such as tillage, manuring, etc. also influences SOC stock to a greater extent. The CSP was found to be inversely related with SOC stock *i.e.* soils with higher OC stock generally have lower potential to sequester further carbon in them and vice-versa. The presence of more fine soil fractions (clay and silt fractions) improves the CSP of soils. It is important to increase the OC stock levels in soil as it not only improves the soil fertility but also reduces the emission of carbon dioxide into the atmosphere to a greater extent. The results of the study show the effectiveness of RF model in prediction of SOC stock and CSP than other models. The prediction accuracy of RF model can still be improved by the addition of more datasets and more suitable environmental covariates. More samples need to be collected in areas having higher uncertainty for accurate prediction. Long term research should be carried out to assess the potential of new management practices in C sequestration and its stabilization on a long-term basis by the use of advanced ML algorithms.

REFERENCES

Adhikari, S., Pal, S., Barik, A., Chakraborty, S. and Mukhopadhyay, S.K. 2020. Assessing soil and sediment organic carbon sequestration potential of selected wetlands at different physiographic regions of West Bengal, India. *Indian J. Soil Cons.*, 48(3): 251-261.

Akpa, S.I., Odeh, I.O., Bishop, T.F., Hartemink, A.E. and Amapu, I.Y. 2016. Total soil organic carbon and carbon sequestration potential in Nigeria. *Geoderma,* 271: 202-215, doi.org/10.1016/j.geoderma.2016.02.021.

Breiman, L. 2001. Random forests. *Mach. Learn.*, 45(1): 5-32, doi.org/10.1023/A:1010933404324.

Cortes, C. and Vapnik, V. 1995. Support-vector networks. *Mach. Learn*, 20: 273-297.

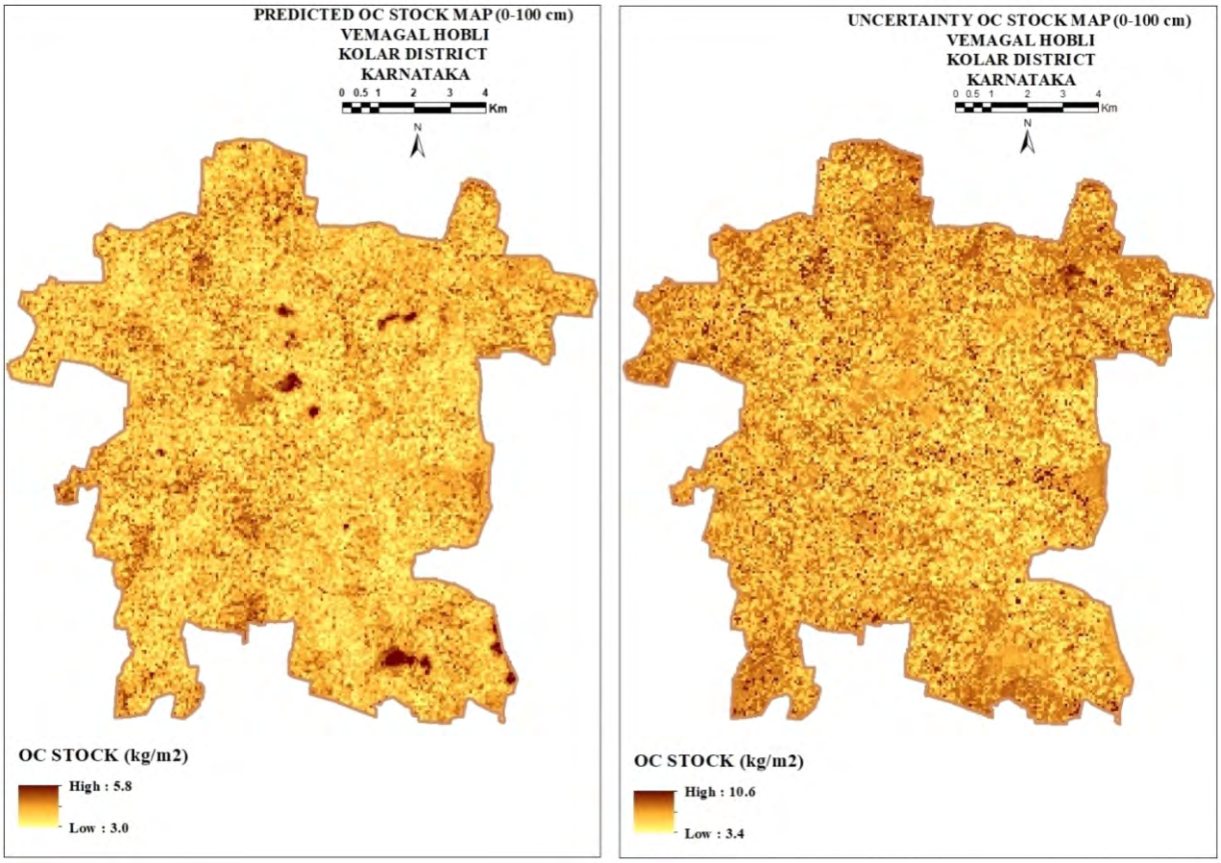

Fig. 2. Predicted and uncertainty maps of organic carbon stock (0-100 cm)

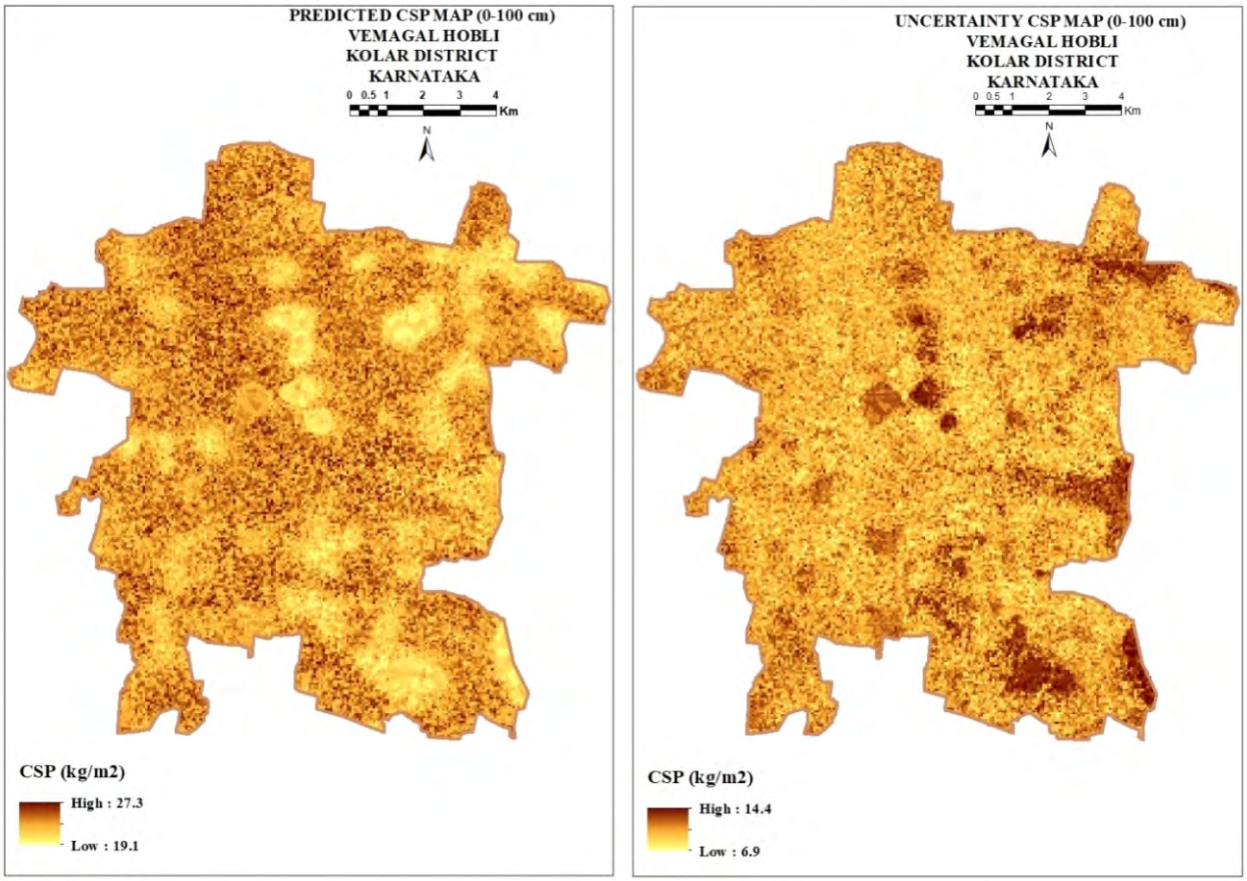

Fig. 3. Predicted and uncertainty maps of CSP (0-100 cm)

Carré, F., McBratney, A.B., Mayr, T. and Montanarella, L. 2007. Digital soil assessments: beyond DSM. *Geoderma,* 142: 69-79, doi.org/10.1016/j.geoderma.2007.08.015.

Das, M., Dash, P.K., Bhattacharyya, P., Munda, S., Padhi, S.R., Padhi, P.P., Das, M. and Nayak, A.K. 2019. Energy saving in relation to soil carbon pools and enzymatic activities under different conservation tillages and nutrient management in tropical rice. *Indian J. Soil Cons*, 47(2): 172-179.

Dharumarajan, S., Hegde, R. and Singh, S.K. 2017. Spatial prediction of major soil properties using random forest techniques - a case study in semi-arid tropics of South India. *Geoderma Reg.,* 10: 154-162, doi.org/10.1016/j.geodrs.2017.07.005.

Dharumarajan, S., Kalaiselvi, B., Suputhra, A., Lalitha, M., Hegde, R., Singh, S.K. and Lagacherie, P. 2020. Digital soil mapping of key Global Soil Map properties in northern Karnataka plateau. *Geoderma Reg.,* 20, doi.org/10.1016/j.geodrs.2019.e00250.

Dharumarajan, S., Kalaiselvi, B., Suputhra, A., Lalitha, M., Vasundhara, R., Anil Kumar, K.S., Nair, K.M., Hegde, R., Singh, S.K. and Lagacherie, P. 2021. Digital soil mapping of soil organic carbon stocks in Western Ghats, South India. *Geoderma Reg.,* 25: 1-10, doi.org/10.1016/j.geodrs.2021.e00387.

Dharumarajan, S. and Hegde, R. 2022. Digital mapping of soil texture classes using Random Forest classification algorithm. *Soil Use Manage.*, 38 (1): 135-149, doi.org/10.1111/sum.12668.

FAO. 2016. *State of the World's Forests.* Forests and Agriculture: Trend in Land-Use Change. 7-22.

Gomes, L.C., Faria, R.M., De Souza, E., Veloso, G.V., Schaefer, C.E.G. and Fernandes Filho, E.I. 2019. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma,* 340: 337-350, doi.org/10.1016/j.geoderma.2019.01.007.

Gray, J.M., Wang, B., Waters, C.M., Orgill, S.E., Cowie, A.L. and Ng, E.L. 2022. Digital mapping of soil carbon sequestration potential with enhanced vegetation cover over New South Wales, Australia. *Soil Use Manage.*, 38: 229-247, doi.org/10.1111/sum.12766.

Hammad, H.M., Nauman, H.M.F., Abbas, F., Ahmad, A., Bakhat, H.F., Saeed, S., Shah, G.M., Ahmad, A. and Cerdà, A. 2020, Carbon sequestration potential and soil characteristics of various land use systems in arid region. *J. Environ. Manage.,* 264: 1-9.

Hassink, J. 1997. The capacity of soils to preserve organic C and N by their association with clay and silt particles. *Plant Soil.,* 191: 77-87, doi.org/10.1023/A:1004213929699.

Bernstein, L., Bosch, P., Canziani, O., Chen, Z., Christ, R. and Riahi, K. 2008. *IPCC, 2007*: climate change synthesis report.

Jackson, M.L. 1973. *Soil Chem. Anal.,* Prentice Hall of India (Pvt) Ltd., New Delhi.

Ließ, M., Schmidt, J. and Glaser, B. 2016. Improving the spatial prediction of soil organic carbon stocks in a complex tropical mountain landscape by methodological specifications in machine learning approaches. *PLoS One,* 11: 1-22, doi.org/10.1371/journal.pone. 0153673.

McBratney, A.B., Mendonca Santos, M.L. and Minasny, B. 2003. On digital soil mapping. *Geoderma,* 117: 3-52, doi.org/10.1016/S0016-7061(03)00223-4.

Miller, B.A., Koszinski, S., Wehrhan, M. and Sommer, M. 2015. Impact of multi-scale predictor selection for modeling soil properties. *Geoderma*, 239: 97-106, doi.org/10.1016/j.geoderma.2014.09.018.

Oelkers, E.H. and Cole, D.R. 2008. Carbon dioxide sequestration: a solution to the global problem. *Elements,* 4(5): 305-310.

Padarian, J., Minasny, B., McBratney, A. and Smith, P. 2022. Soil carbon sequestration potential in global croplands. *Peer J.*, 10: 13740.

Pradhan, B. 2013. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput. Geosci.*, 51: 350-365, doi.org/10.1016/j.cageo.2012.08.023.

Prasad, A.M., Iverson, L.R. and Liaw, A. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosyst.*, 9: 181-199.

Prieto, A., Atencia, M. and Sandoval, F. 2013. Advances in artificial neural networks and machine learning. *Neurocomputing* 121: 1-4, doi.org/10.1016/j.neucom.2013.01.008.

Rostaminia, M., Rahmani, A., Mousavi, S.R., Taghizadeh-Mehrjardi, R. and Maghsodi, Z. 2021. Spatial prediction of soil organic carbon stocks in an arid rangeland using machine learning algorithms. *Environ. Monit. Assess.*, 193(12): 1-17.

Sarma, U.J., Chakravarty, M. and Bhattacharyya, H.C. 2013. Emission and sequestration of carbon in soil with crop residue incorporation. *J. Indian Soc. Soil Sci.,* 61(2): 117-121.

Sharma, G. and Sharma, L.K. 2022. Assessment of soil carbon stock and important physicochemical properties in relation to land use patterns in semi-arid region of Rajasthan, India. *J. Indian Soc. Soil Sci.,* 70(2): 191-203, dx.doi.org/10.5958/0974-0228.2022.00019.6.

Sreenivas, K., Dadhwal, V.K., Kumar, S., Sri Harsha, G., Mitran, T., Sujatha, G., Suresh, J.R.G., Fyzee, M.A. and Ravisankar, T. 2016. Digital mapping of soil organic and inorganic carbon status in India. *Geoderma*, 269: 160-173, doi.org/10.1016/j.geoderma.2016.02.002.

Srinivasan, R., Natarajan, A., Anil Kumar, K.S. and Lalitha, M. 2019. Carbon stocks in major cashew growing soils of coastal Karnataka, India. *J. Plant. Crops.,* 47(1): 64-70, doi: 10.25081/jpc.2019.v47.i1.5536.

Walkley, A. and Black I.A. 1934. An estimation of the method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Sci*., 37: 29-38.

Wiesmeier, M., Hubner, R., Sporlein, P., Geub, U., Hangen, E., Reischl, A., Schilling, B., Von-Lutzow, M. and Kogel-Knabner, I. 2014. Carbon sequestration potential of soils in southeast Germany derived from stable soil organic carbon saturation. *Global Change Biol.*, 20(2): 653-665, doi.org/10.1111/gcb.12384.