Geoderma Regional 20 (2020) e00250

Contents lists available at ScienceDirect

Geoderma Regional

journal homepage: www.elsevier.com/locate/geodrs

Digital soil mapping of key GlobalSoilMap properties in Northern Karnataka Plateau

S. Dharumarajan [a,*], B. Kalaiselvi [a], Amar Suputhra [a], M. Lalitha [a], Rajendra Hegde [a], S.K. Singh [b], Philippe Lagacherie [c]

[a] ICAR-National Bureau of Soil Survey and Land Use Planning, Regional Centre, Hebbal, Bangalore 560024, India
[b] ICAR-National Bureau of Soil Survey and Land Use Planning, Amaravati Road, Nagpur-10, India
[c] LISAH, Univ. Montpellier, INRA, IRD, Montpellier SupAgro, Montpellier, France

A R T I C L E I N F O

Article history:
Received 5 August 2019
Received in revised form 22 November 2019
Accepted 30 November 2019

Keywords:
Digital soil mapping
Vertisols
Alfisols
Quantile regression forest
Soil properties
Field capacity
Permanent wilting point

A B S T R A C T

Accurate and quantitative information on soil properties of each and every location is essential for site specific sustainable management of land resources. A study was conducted to predict the different key soil properties of Northern Karnataka as per GlobalSoilMap specifications using Quantile Regression Forest (QRF) Model. Along with Sentinel-2 data, terrain attributes such as elevation, slope, aspect, topographic wetness index, topographic position index, plan and profile curvature, multi-resolution index of valley bottom flatness, multi-resolution ridge top flatness and vegetation factors like NDVI and EVI were used as covariates. Equal-area quadratic splines were fitted to soil profile datasets to estimate soil properties viz. pH, OC, CEC, clay, sand, silt, field capacity and permanent wilting point at six standard soil depths (0–5, 5–15, 15–30, 30–60, 60–100 and 100–200 cm) as per GlobalSoilMap specifications. The coefficient of determination (R^2), mean error (ME) and root mean square error (RMSE) were calculated in order to assess model performance. Prediction interval coverage percentage (PICP) was calculated to evaluate the associated uncertainty predictions. The predicted soil properties are reliable with minimum errors and the QRF model captured maximum variability for most of the soil properties.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Soil properties are assessed through resource inventorisation with the main objective to delineate areas which need uniform management practices and provide users with information on soil properties. Assessment of spatial distribution of soil properties for each location is important for site-specific land management, land evaluation and land suitability analysis (Gessler et al., 2000; McBratney et al., 2003). Although several spatial soil databases are developed throughout the world, they are neither exhaustive nor precise enough for ensuring enlightened decisions. For example, though digitized soil maps are available for most of the world (Grunwald et al., 2011), those information are at very small scale (1:1 million or coarser) for many areas and do not adequately represent soil variability in a format that is useful for a non-pedologists (Sanchez et al., 2009). Digital soil mapping (DSM) represents a ground-breaking solution compared to conventional soil survey by its ability to exploit large sets of spatial data, to produce uncertainty estimates associated with soil predictions and can be revised once new data are collected (Lagacherie and McBratney, 2007). Soil database generated through field sampling and laboratory analysis are used to feed a DSM model that predicts soil properties in areas not sampled. Digital soil maps also provide the uncertainties associated with such predictions. The overall uncertainty of the prediction is estimated by combining uncertainties of input data, spatial inference, and soil functions (Dharumarajan et al., 2019a). Uncertainties are essential for understanding and dealing with risk in decision-making.

DSM has moved from a largely academic towards an operational activity through GlobalSoilMap project (http://www.globalsoilmap.net/). The project aims to map the several key soil properties of globe onto a three-dimensional grid at fine spatial resolution with local uncertainty estimates (Arrouays et al., 2014). The first versions of GlobalSoilMap products have already been produced in various countries (Mulder et al., 2016; Grundy et al., 2015; Adhikari et al., 2014; Poggio and Gimona, 2017) with spatial inference functions using globally available landscape parameters such as Digital Elevation Models, multispectral remote sensing, geology maps, and legacy soil maps as inputs.

In India, ICAR-National Bureau of Soil Survey and Land Use Planning (ICAR-NBSS&LUP), Nagpur has recently launched an ambitious program

* Corresponding author.
E-mail address: Dharumarajan.S@icar.gov.in (S. Dharumarajan).

https://doi.org/10.1016/j.geodrs.2019.e00250
2352-0094/© 2019 Elsevier B.V. All rights reserved.

called "IndianSoilGrids" with the objective to develop soil properties map as per GlobalSoilMap Specifications. In recent past, effort has been made to compile the legacy soil data in the form of harmonized databases and stored in NBSS&LUP Geoportal. Besides pursuing the storage effort, IndianSoilGrids project paved the ways to exploit legacy soil data through DSM models. In this context, the present exploratory study was carried out to produce a fine resolution map of major GlobalSoilMap soil properties such as organic carbon, pH, CEC, clay, sand, silt, field capacity and permanent wilting point in part of Northern Karnataka Plateau region representing semi arid tropics of south India using Quantile Regression Forest Model techniques.

2. Materials and methods

2.1. Study area

The present study was carried out in part of Koppal and Gadag districts of Northern Karnataka Plateau (Fig. 1). The study area is located in 14° 56′ to 15° 37′ N latitude and 75°23′ to 76° 25′ longitude with an area of 3655 km^2. The study region represents hot semi-arid climate with rainfall range of 600–750 mm and potential evapo-transpiration (PET) of 1600–1700 mm. The average annual rainfall is 672 mm. This area includes mountainous, expansive plateau with substantial area is underlined by basalts with continuation of Deccan trap of Maharashtra. The major area comes under rainfed cultivation with crops like Sorghum, Pigeon pea and Pearl millet. The major soils represented by shallow to deep vertisols, alfisols and inceptisols.

Table 1
Different covariates used in the model.

Predictor	Source	Resolution
Elevation (m)	SRTM DEM	30 m
Slope (%)	SRTM DEM	30 m
Aspect	SRTM DEM	30 m
TPI	SRTM DEM	30 m
TWI	SRTM DEM	30 m
Plan curvature	SRTM DEM	30 m
Profile curvature	SRTM DEM	30 m
MrVBF	SRTM DEM	30 m
MrRTF	SRTM DEM	30 m
NDVI	MOD13Q1(2011–2015)	250 m _16 days
EVI	MOD13Q1(2011–2015)	
Sentinel-2	13 bands of Sentinel 2 data	10-60 m

2.2. Sampling methodology

The profiles studied under Sujala III (Karnataka Watershed Development Project II) project were used for mapping of soil properties. Sixty soil profiles were studied upto 2 m or hard rock based on variability in landform and land use. The soil horizons were demarcated and from the representative soil horizons, soil samples were collected for laboratory analysis. Collected soil samples were air dried in shade and passed through 2 mm sieve by gently ground with a wooden mallet. The samples were analysed for particle-size distribution following International Pipette method (Richards, 1954), pH and electrical conductivity (EC) in 1:2.5 soil:water suspension (Jackson, 1962). Organic carbon was estimated by Walkley and Black (1934) method. The cation exchange capacity (CEC) and exchangeable cations were determined as described

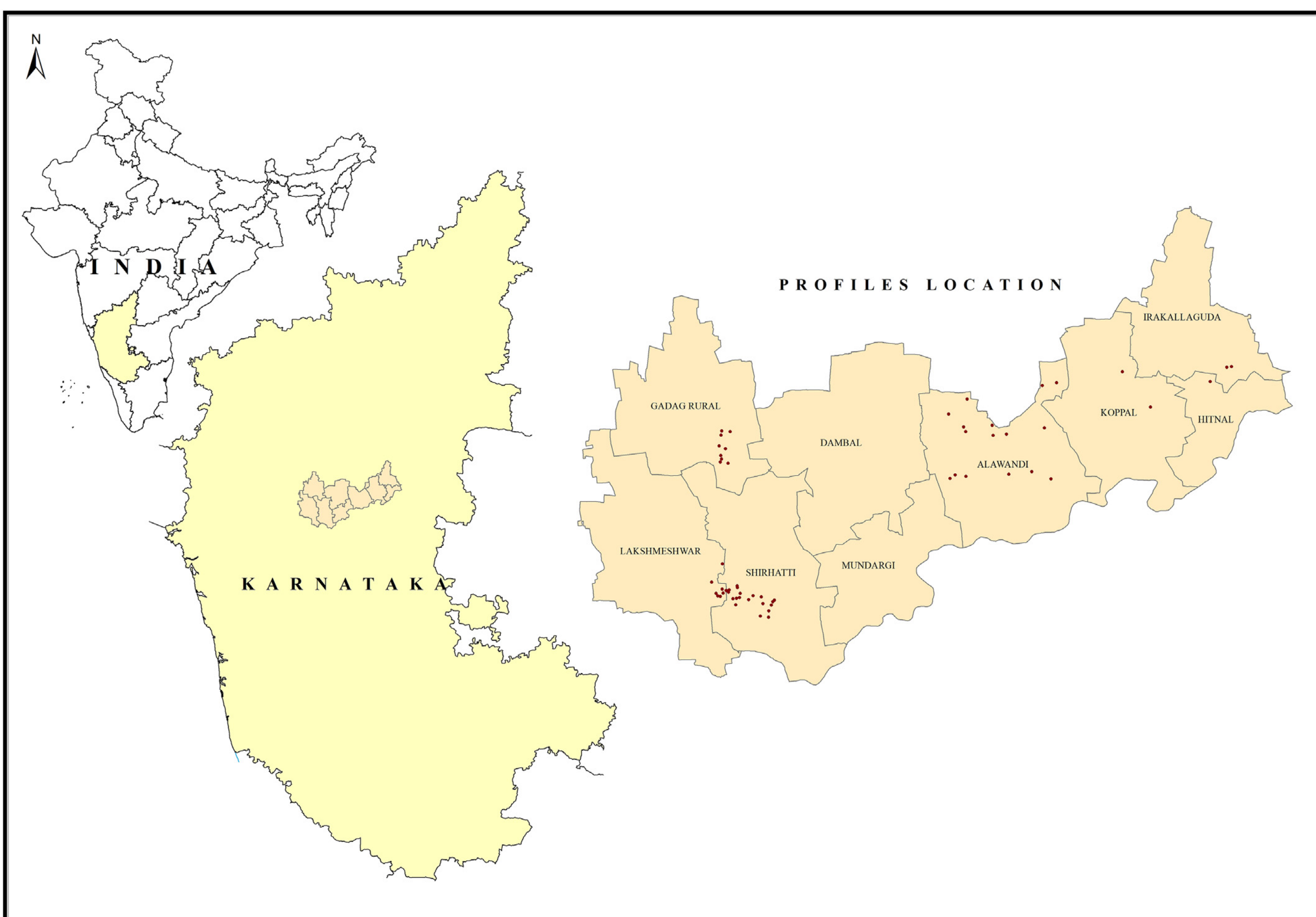

Fig. 1. Study area with profile locations.

Table 2
Statistical results of soil properties.

Properties	Mean	Min	Max	Std dev.	Skewness	Kurtosis
pH	8.2	4.7	9.9	1.1	−1.2	0.9
OC (%)	0.5	0.11	1.16	0.23	0.5	−0.35
Clay(%)	42.6	4.1	75.8	18.2	−0.1	−0.8
Sand(%)	41.3	8.7	87.6	22	0.3	−1
Silt(%)	16.1	4.7	40.7	6.9	0.6	0.3
CEC (c mol(+) kg^{-1})	32.3	2	80.9	20.5	0.3	−1.1
FC(%)	29.2	6	60	12.2	0.3	−0.6
PWP(%)	18.5	1.5	43.7	10.5	0.5	−0.6

by Jackson (1973). Field capacity (FC) and permanent wilting point (PWP) were estimated using pressure plate apparatus (Richards et al., 1956). The profile soil properties were pre-processed by harmonization of soil depth interval (GlobalSoilMap depth specification) predictions using equal-area spline functions (Bishop et al., 1999).

2.3. Environmental covariates and models used

A Digital elevation model (DEM) with 30 m resolution was obtained from SRTM and processed using ArcGIS10 data management tool box. The primary and secondary derivates of DEM like elevation, slope, aspect, curvatures (plan and profile), topographic wetness index (TWI) and topographic position index (TPI), LS factor, Multi-resolution Ridge Top Flatness (MrRTF) and Multi-resolution Index of Valley Bottom Flatness (MrVBF) were derived by using Saga-GIS 6.3.0 version. Along with DEM attributes, all the bands of Sentinel- 2 imagery (13 bands), Normalized Difference Vegetation Index (NDVI) and Enhanced vegetation index (EVI) (MOD13Q1) were used as covariates for prediction of soil properties (Table.1). The environmental variables were intersected for all the sampling points for prediction of soil properties.

Quantile regression forest (QRF) model was used for prediction of soil properties and uncertainty estimates in the study area. QRF is an extension of Random forest model and the advantage of QRF over Random Forest model (RFM) is for each node in each tree, RFM keeps only the mean of the observations that fall into this node and neglects all other information whereas QRF keeps the value of all observations in this node, and assesses the conditional distribution based on the information (Meinshausen, 2006; Vaysse and Lagacherie, 2017; Dharumarajan et al., 2019a). For the present study, ranger package was used for running the QRF algorithm in R environment. Ranger package helps to identify the best RF properties for running the model. Ten folds cross

Table 3
Performance of Quantile Regression Forest model for prediction of soil properties.

		Mean error	RMSE	R^2(%)	PICP
pH	0–5 cm	−0.19 ± 0.02	0.96 ± 0.02	10 ± 5	88.7 ± 1.3
	5–15 cm	−0.18 ± 0.02	0.94 ± 0.03	13 ± 6	89.0 ± 0.78
	15–30 cm	−0.18 ± 0.02	0.95 ± 0.03	8 ± 9	90.0 ± 1.1
	30-60 cm	−0.14 ± 0.03	1.00 ± 0.03	9 ± 6	88.3 ± 2.0
	60–100 cm	−0.3 ± 0.02	1.02 ± 0.02	23 ± 3	89.3 ± 2.1
	100-200 cm	−0.23 ± 0.01	0.88 ± 0.03	4 ± 8	86.4 ± 3.2
OC (%)	0–5 cm	−0.02 ± 0.0	0.22 ± 0.01	08 ± 6	87.5 ± 1.7
	5–15 cm	−0.02 ± 0.01	0.22 ± 0.01	07 ± 5	86.2 ± 1.9
	15–30 cm	0.0 ± 0.00	0.21 ± 0.01	10 ± 4	86.8 ± 2.3
	30-60 cm	0.02 ± 0.01	0.20 ± 0.01	27 ± 4	88.6 ± 1.5
	60–100 cm	0.03 ± 0.01	0.19 ± 0.00	0 ± 2	89.4 ± 1.7
	100-200 cm	0.03 ± 0.00	0.20 ± 0.00	5 ± 3	84.8 ± 2.1
Clay (%)	0–5 cm	0.19 ± 0.2	6.10 ± 0.29	37 ± 6	88.4 ± 2.3
	5–15 cm	2 0.22 ± 0.17	6.04 ± 0.22	39 ± 5	88.3 ± 2.0
	15–30 cm	0.59 ± 0.16	6.09 ± 0.20	39 ± 4	88.2 ± 1.2
	30-60 cm	−0.04 ± 0.51	12.39 ± 0.42	43 ± 4	87.2 ± 2.3
	60–100 cm	0.0 ± 0.09	4.95 ± 0.1	18 ± 3	88.3 ± 2.1
	100–200 cm	0.05 ± 0.21	5.2 ± 0.22	0 ± 8	83.2 ± 1.6
Sand (%)	0–5 cm	1.54 ± 0.56	17.24 ± 0.59	48 ± 4	86.6 ± 1.7
	5–15 cm	2.25 ± 0.56	16.92 ± 0.72	49 ± 4	87.6 ± 2.0
	15–30 cm	1.37 ± 0.44	16.10 ± 0.34	45 ± 2	85.3 ± 1.4
	30-60 cm	40.04 ± 0.45	17.86 ± 0.43	42 ± 3	92.2 ± 1.2
	60–100 cm	−0.24 ± 0.52	15.43 ± 1.18	45 ± 9	93.6 ± 1.3
	100-200 cm	0.1 ± 0.92	16.66 ± 0.94	41 ± 7	88 ± 0.0
Silt (%)	0–5 cm	−0.20 ± 0.39	13.47 ± 0.5	49 ± 4	91.8 ± 1.5
	5–15 cm	0.05 ± 0.32	13.90 ± 0.44	45 ± 4	90.8 ± 1.6
	15–30 cm	−0.08 ± 0.31	13.45 ± 0.40	40 ± 4	89.5 ± 1.2
	30-60 cm	0.57 ± 0.17	6.30 ± 0.21	29 ± 5	84.0 ± 1.8
	60–100 cm	0.53 ± 0.49	11.78 ± 0.58	52 ± 5	90.8 ± 1.3
	100-200 cm	0.57 ± 0.17	6.30 ± 0.21	29 ± 5	84.0 ± 1.8
CEC (C mol kg^{-1})	0–5 cm	1.26 ± 0.27	13.48 ± 0.66	51 ± 5	87.6 ± 1.2
	5–15 cm	1.19 ± 0.51	13.47 ± 0.42	51 ± 3	86.9 ± 1.5
	15–30 cm	1.03 ± 0.38	12.27 ± 0.37	56 ± 3	87.6 ± 1.9
	30-60 cm	1.19 ± 0.67	14.88 ± 0.55	43 ± 4	87.1 ± 2.3
	60–100 cm	1.52 ± 0.61	15.98 ± 0.66	40 ± 5	88.8 ± 2.3
	100-200 cm	1.55 ± 0.69	17.1 ± 0.71	31 ± 6	87.1 ± 2.3
FC (%)	0–5 cm	0.03 ± 0.25	8.42 ± 0.29	36 ± 4	88 ± 2.9
	5–15 cm	0.12 ± 0.28	8.69 ± 0.21	30 ± 3	88.0 ± 2.5
	15–30 cm	−0.04 ± 0.2	7.79 ± 0.15	37 ± 3	86.0 ± 1.6
	30-60 cm	0.26 ± 0.16	8.49 ± 0.15	38 ± 2	85.7 ± 2.1
	60–100 cm	0.66 ± 0.37	10.1 ± 0.41	38 ± 5	89.6 ± 1.5
	100-200 cm	−0.29 ± 0.31	10.7 ± 0.41	40 ± 5	84.6 ± 2.4
PWP (%)	0–5 cm	0.26 ± 0.21	6.8 ± 0.2	41 ± 3	92.0 ± 1.5
	5–15 cm	0.3 ± 0.21	6.74 ± 0.25	41 ± 4	91.9 ± 1.4
	15–30 cm	0.3 ± 0.16	6.31 ± 0.13	43 ± 2	89.9 ± 1.7
	30-60 cm	0.50 ± 0.17	6.91 ± 0.16	47 ± 3	91.6 ± 2.1
	60–100 cm	0.79 ± 0.19	8.62 ± 0.39	42 ± 5	90.8 ± 1.8
	100-200 cm	0.84 ± 0.37	8.89 ± 0.53	49 ± 6	90.6 ± 2.7

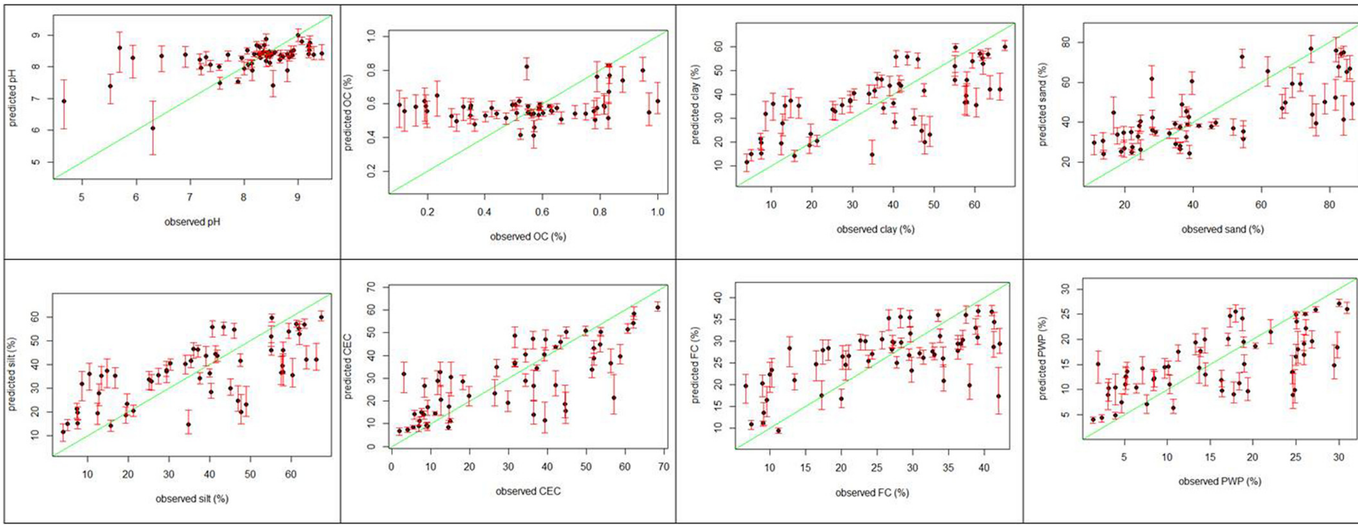

Fig. 2. (a–h). Observed soil properties Vs Predicted soil properties in 0-5 cm depth.

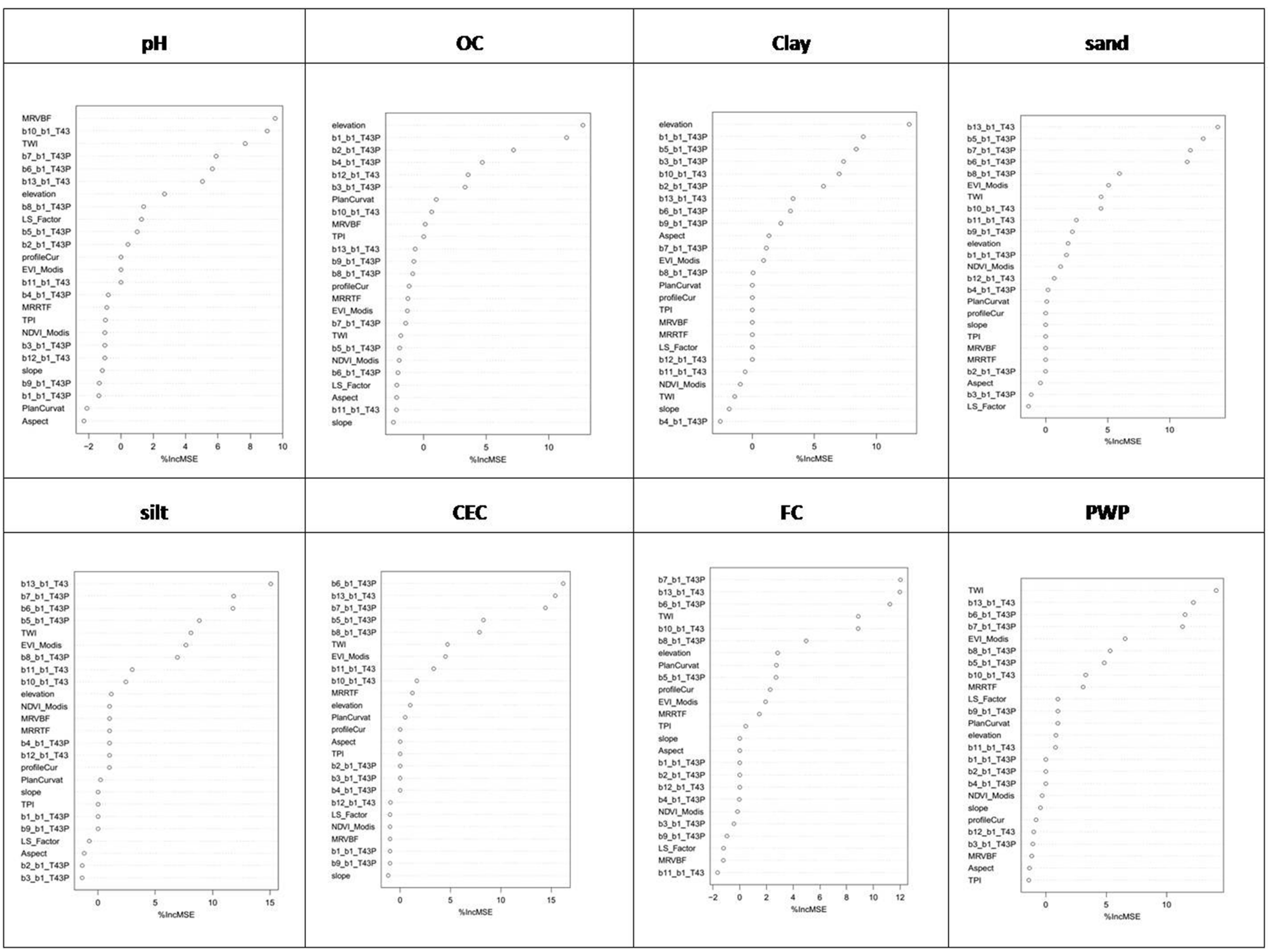

Fig. 3. (a–h). Variation importance ranking of Random forest model in prediction of different soil properties.

Table 4
Summary statistics of predicted soil properties.

		Mean	Min	Max	stdev	Kurtosis	Skewness
pH	0–5 cm	8.1	5.5	9.2	0.6	2.2	−1.5
	5–15 cm	8.2	5.9	9.2	0.4	1.8	−0.8
	15–30 cm	8.2	6.0	9.2	0.3	5.7	−1.6
	30-60 cm	8.2	6.4	9.3	0.4	3.8	−0.9
	60–100 cm	8.5	6.5	9.2	0.3	5.3	−1.5
	100-200 cm	8.6	8.0	9.1	0.2	1.7	−1.4
OC (%)	0–5 cm	0.61	0.35	0.83	0.08	0.9	1.3
	5–15 cm	0.60	0.40	0.84	0.07	3.0	1.7
	15–30 cm	0.59	0.37	0.83	0.10	−0.5	0.5
	30-60 cm	0.59	0.28	0.80	0.11	−0.7	0.0
	60–100 cm	0.4	0.3	0.6	0.1	−0.9	−0.1
	100-200 cm	0.4	0.3	0.6	0.0	9.9	0.9
Clay (%)	0–5 cm	16.2	5.5	32.6	5.1	−1.0	0.0
	5–15 cm	36.6	7.7	62.1	11.3	−0.7	−0.2
	15–30 cm	39.9	7.3	64.1	8.6	0.0	0.2
	30-60 cm	45.7	8.5	66.5	8.6	0.1	0.6
	60–100 cm	15.2	8.0	19.0	2.4	−1.3	−0.2
	100-200 cm	17.2	12.1	21.7	1.7	1.6	−0.5
Sand (%)	0–5 cm	45.5	17.8	82.6	15.4	−0.7	0.7
	5–15 cm	45.3	17.7	82.9	15.0	−0.8	0.7
	15–30 cm	41.7	15.5	76.5	13.5	−1.0	0.5
	30-60 cm	36.1	13.4	80.5	11.9	−0.6	0.3
	60–100 cm	38.4	12.4	57.8	13.9	−1.3	−0.3
	100-200 cm	38.0	11.2	62.6	15.6	−1.4	−0.4
Silt (%)	0–5 cm	36.5	7.7	61.9	12.0	−0.8	−0.2
	5–15 cm	16.7	6.2	32.6	5.6	−0.7	0.2
	15–30 cm	17.5	7.1	35.6	6.4	−1.0	0.3
	30-60 cm	15.8	8.7	27.5	4.2	−0.4	0.3
	60–100 cm	46.9	20.2	69.0	10.7	−1.2	0.4
	100-200 cm	45.0	24.1	71.4	12.9	−1.5	0.3
CEC (C mol kg^{-1})	0–5 cm	27.5	4.7	62.3	14.0	−1.3	0.1
	5–15 cm	27.5	5.2	62.3	14.5	−1.3	0.2
	15–30 cm	29.9	5.0	64.0	14.6	−1.3	0.2
	30-60 cm	32.9	7.0	62.3	13.3	−1.2	0.1
	60–100 cm	32.4	10.6	65.3	14.9	−1.6	0.1
	100-200 cm	31.0	13.9	54.5	14.7	−1.5	0.3
FC (%)	0–5 cm	24.9	9.0	41.0	6.8	−0.5	−0.7
	5–15 cm	25.6	10.1	39.0	6.3	−0.4	−0.6
	15–30 cm	26.7	12.1	39.9	5.7	−0.8	−0.2
	30-60 cm	29.3	10.4	46.8	5.9	−1.1	0.4
	60–100 cm	32.2	18.8	45.7	7.9	−1.5	−0.2
	100-200 cm	30.5	15.2	50.4	8.3	−1.4	0.3
PWP (%)	0–5 cm	14.6	3.2	29.5	5.7	−0.9	0.1
	5–15 cm	15.1	3.5	27.3	5.2	−0.9	0.3
	15–30 cm	16.3	4.8	27.9	4.7	−1.1	0.3
	30-60 cm	18.5	2.9	36.1	5.7	−1.0	0.3
	60–100 cm	19.8	9.0	37.5	7.0	−1.3	0.1
	100-200 cm	19.8	7.8	39.1	8.0	−1.5	−0.1

validation techniques with 20 times repetition was used to evaluate the performance of QRF model. The performance of QRF was evaluated using indicators such as Coefficient of determination (R^2), Root Mean Square Error (RMSE), mean error (ME). Prediction interval coverage percentage (PICP) was used to evaluate the uncertainty of prediction.

3. Results and discussion

3.1. Summary statistics of soil properties

Summary of the soil properties are presented in Table 2. The soil pH ranged from 4.7 to 9.9 with a mean and standard deviation of 8.2 and 1.1, respectively. The organic carbon content ranged between 0.11 and 1.16% with mean of 0.5% and standard deviation of 0.23%. The organic carbon skewed positively whereas pH skewed negatively showed that asymmetrical distribution. The higher variability in pH is mainly attributed to soil pedological factors and land management. The soil hydraulic properties such as field capacity and permanent wilting point were ranged from 6 to 60% and 1.5 to 43.7% with mean and standard deviation of 29.2, 18.5 and 12.2, 10.5% respectively. Cation exchange capacity of the soil varied from 2.0 to 80.9 cmol(+) $kg^{-1}$with mean and SD of 32.3 and 20.5 cmol(+) kg^{-1} respectively. Except pH and silt content, all other soil properties had registered negative kurtosis. Similar way except, clay content and pH, all other properties showed positive skewness. The correlation analysis showed that field capacity and permanent wilting point has significant positive correlation with clay and silt and negative correlation with sand content.

3.2. Performance of Quantile Regression Forest Model in predicting soil properties

The performance of Quantile Regression Forest model was evaluated by calculating statistical indicators viz., Coefficient of determination (R^2), Mean error (ME) and Root Mean Square Error (RMSE). The cross validation results (Table 3 and Fig. 2a–h) showed that the combination of different covariates explained the varaibilities of predicted soil properties viz., pH, organic carbon, CEC, clay, sand, silt, FC and PWP. The model could capture low to medium variability ($R^2 = 0$–56%) while predicting pH, Organic carbon and CEC for different depth ranges.

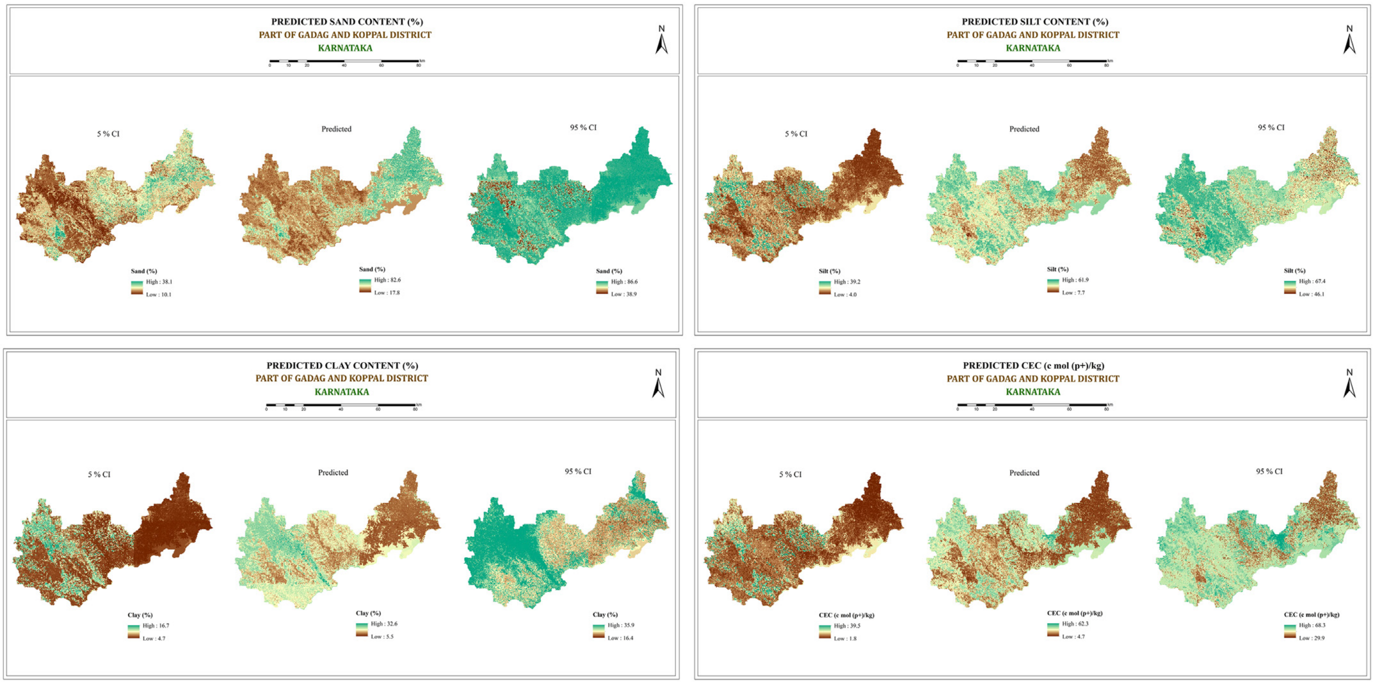

Fig. 4. Predicted sand, silt, clay and CEC content in 0–5 cm depth.

Among these soil properties, CEC prediction was good compared to pH and Organic carbon. The present model explained 31–56% of variation for prediction of CEC in different depth intervals. Similar results were observed by different researchers (Gallo et al., 2018, $R^2 = 40\%$; Chagas et al., 2018, $R^2 = 47\%$; Ghaemi et al., 2013, $R^2 = 45–65\%$). In case of pH, only 8–23% of variability was captured by the model. The poor prediction may be attributed to more variability in pH influenced by soil intrinsic (pedogneic) and extrinsic (land management) factors. Like, pH, the performance of the model for prediction of organic carbon is also very low ($R^2 = 0–27\%$). The poor performance may be related to the low levels of soil organic carbon compared to soils having high organic carbon (Lo Seen et al., 2010; De et al., 2014; Gastaldi et al., 2012; Dharumarajan et al., 2017; Dharumarajan et al., 2019a). The prediction of particle size quantities viz., clay, sand and silt content were fairly good. Prediction accuracy for sand is 41–49% with RMSE of 15.4–17.9%.R^2 of silt varied from 29 to 49%for different depth intervals. Similar results were observed by Akpa et al. (2014) who recorded R^2 value of 16–56% for prediction of particle size fractions in Nigeria using RFM whereas Santra et al. (2017) found only 21–28% of variation in sand content captured by Random forest algorithm.

Soil hydraulic properties are important for irrigation scheduling and proper landuse planning (Dharumarajan et al., 2019b). Soil hydraulic properties such as field capacity and permanent wilting point determines the availability and retention of the water for crop growth. Field capacity and permanent wilting point were well predicted by QRF model. Compared to field capacity ($R^2 = 30–38\%$), permanent wilting point was predicted with high accuracy ($R^2 = 41–49\%$). Hong et al. (2013) recorded digital soil mapping approach for prediction of soil hydraulic property with maximum accuracy (R^2–61%) whereas Román Dobarco et al. (2019) reported prediction accuracy (R^2) of FC and PWP were 21 and 29% respectively.

Prediction interval coverage probability (PICP) is an indication of efficiency of uncertainty measurements. The present predictions found that the PICP values ranged from 83.2 to 92.2%. Overall, the prediction performance of this model was high for soil hydraulic properties. Higher sample density is required for better results in tropical countries where soil pattern is complex due to the geological uplift than other regions (De et al., 2014).

3.3. Importance of predictor variables for predicting soil properties

RFM model estimates the importance of covariates based on how best or worse the prediction would be if one or more variable is removed and also it protects elimination of good predictor variables which are important for the model (Prasad et al., 2006). Fig. 3a–h shows the variable importance rankings of Random Forest model for pH, OC, clay, sand, silt, CEC, FC and PWP. Elevation is emerged as top predictor for prediction of clay and organic carbon. MRVBF and TWI are ranked as most important predictor for prediction of pH and PWP. Different bands of Sentinel −2 imagery occupies in the top position for prediction of majority of soil properties. Different researchers recorded usefulness of Sentinel-2 imageries in prediction of different soil properties (Castaldi et al., 2019; Gholizadeh et al., 2018; Vaudour et al., 2019). Recently, Gomez et al. (2019) showed good discrimination ability of time series Sentinel-2 images in identifying different texture class and associated uncertainty.

4. Spatial prediction of soil properties

Mapping of soil properties is a preliminary step due its variability for decision making such as the delineation of suitable crop growing areas or identification of degraded areas. Summary statistics of predicted soil properties are presented in Table 4. Predicted maps of sand, silt, clay, CEC, FC and PWP in the surface (0–5 cm) along with uncertainty using Quantile Regression Forest are presented in Figs. 4 & 5. The predicted sand content in 0–5 cm varied from 17.8–82.6%. The predicted silt and clay content varied from 7.7–61.9% and 5.5–32.6% respectively. High sand content recorded in North-eastern part of study area and high clay and silt content recorded in north-western part. The high sand content of surface soils in North-eastern part might be due to severity of the erosion where finer particles are moved into the low lying areas. The predicted cation exchange capacity varied from 4.7–62.3% and recorded low CEC in north-eastern part. Predicted hydraulic properties viz., field capacity and permanent wilting point ranged from 9.1–41% and 3.2–29.5% respectively. The spatial prediction of soil properties suggested that distribution of soil properties on the surface are highly

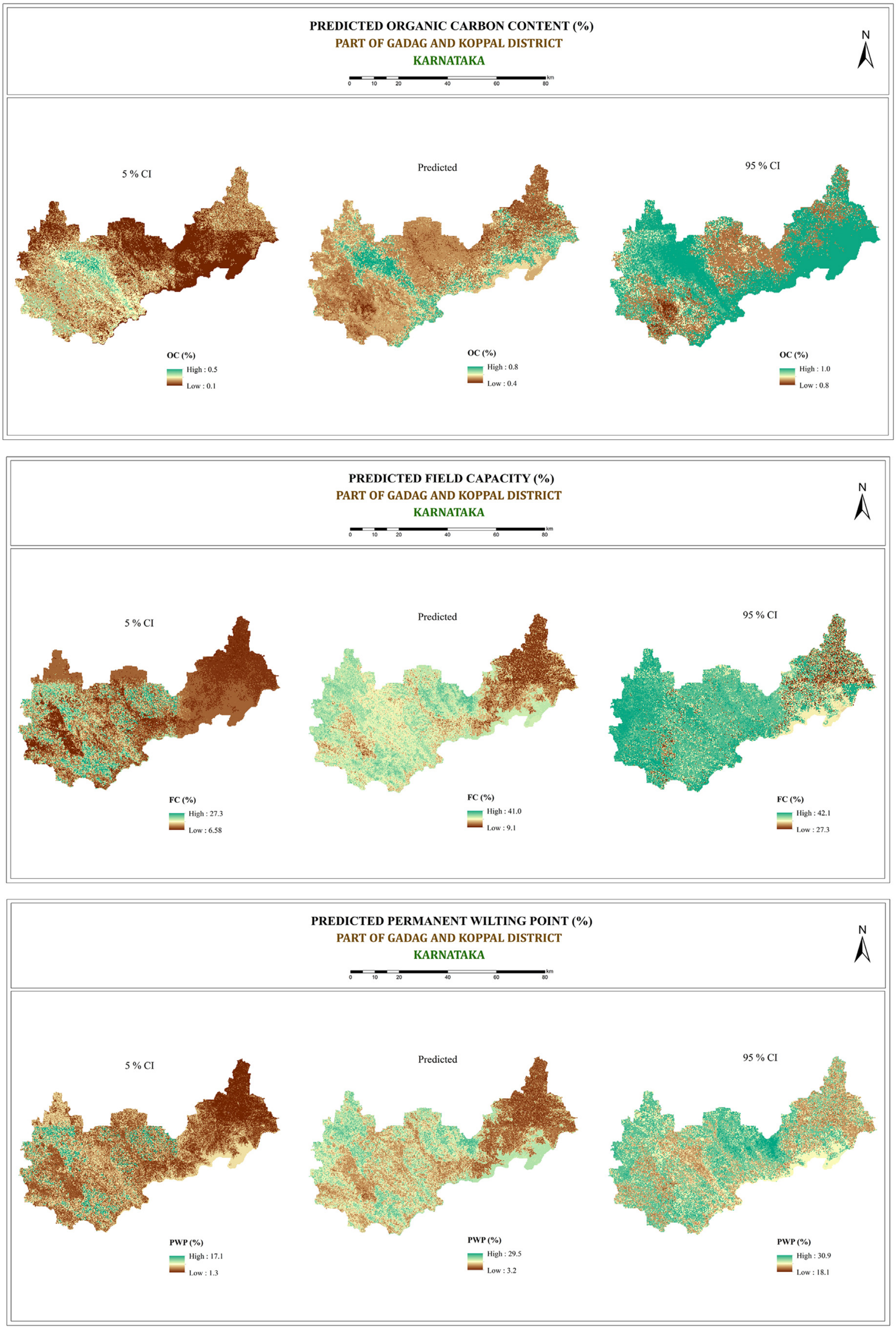

Fig. 5. Predicted Organic carbon content, Field capacity and permanent wilting point in 0–5 cm depth.

variable due to variations in environmental factors, land management and land use. The spatial resolution of the maps helps to assess and monitor the soil health and preparation of proper land use plan.

5. Conclusion

The prediction of soil properties and uncertainty by QRF model was reasonable and varied from 8 to 51% for surface and 0–56% for subsoil. Except pH and OC, the present model predicted better for most of the soil properties compared to previous studies. Weak variations in soil properties, mixed lithologic occurrence and sparse sample density are linked with performance of the model. The data augmentation certainly helps in reducing the uncertainty and over fitting and to improve model accuracy further. The prediction can also be improvised by increasing the environment covariates such as geology map and climatic datasets.

Acknowledgements

The funding received from the project ATCHA, ANR-16-CE03-0006 to attend and present this paper in the 2019 digital soil mapping and GlobalSoilMap workshop held in The University of Chile, Santiago is highly acknowledged. The authors also acknowledge Dr. Laurent Ruiz, Indo-French Cell for Water Sciences, Bangalore for his guidance in Indian digital soil mapping programme. PL is a collaborator of the consortium GLADSOILMAP supported by LE STUDIUM Loire Valley Institute for Advanced Studies through it's LE STUDIUM research consortium programme.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:https://doi.org/10.1016/j.geodrs.2019.e00250. These data include the Google map of the most important areas described in this article.

References

Lo Seen, D., Ramesh, B.R., Nair, K.M., Martin, M., Arrouays, D., Bourgeon, G., 2010. Soil carbon stocks, deforestation and land-cover changes in the Western Ghats biodiversity hotspot (India). Glob. Chang. Biol. 16, 1777–1792. https://doi.org/10.1111/j.1365-2486.2009.02127.x.

Adhikari, K., Minasny, B., Greve, M.B., Greve, M.H., 2014. Constructing a soil class map of Denmark based on the FAO legend using digital techniques. Geoderma 214–215, 101–113.

Akpa, S.I.C., Odeh, I.O.A., Bishop, T.F.A., Hartemink, A.E., 2014. Digital mapping of soil particle-size fractions for Nigeria. Soil Sci. Soc. Am. J. 78, 1953–1966.

Arrouays, D., Grundy, M.G., Hartemink, A.E., Hempel, J.W., Heuvelink, G.B.M., Hong, S.Y., Lagacherie, P., Lelyk, G., McBratney, A.B., McKenzie, N.J., Mendonça-Santos, M.D., Minasny, B., Montanarella, L., Odeh, I.O.A., Sanchez, P.A., Thompson, J.A., Zhang, G.L., 2014. GlobalSoilMap: towards a fine-resolution global grid of soil properties. Adv. Agron. 125, 93–134.

Bishop, T.F.A., McBratney, A.B., Laslett, G.M., 1999. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. A comparison of prediction methods for the creation of field-extent soil property maps. Geoderma 91, 27–45.

Castaldi, F., Hueni, A., Chabrillat, S., Ward, K., Buttafuoco, G., Bomans, B., Vreys, K., Brell, M., van Wesemael, B., 2019. Evaluating the capability of the sentinel 2 data for soil organic carbon prediction in croplands. ISPRS J. Photogramm. Remote Sens. 147, 267–282.

Chagas, C.S., Carvalho Júnior, W., Pinheiro, H.S.K., Xavier, P.A.M., Bhering, S.B., Pereira, N.R., Calderano Filho, B., 2018. 2018. Mapping soil cation exchange capacity in a semiarid region through predictive models and covariates from remote sensing data. Rev. Bras. Cienc Solo 42, e0170183. https://doi.org/10.1590/18069657rbcs20170183.

Dharumarajan, S., Hegde, R., 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. 10C, 154–162. https://doi.org/10.1016/j.geodrs.2017.07.005.

Dharumarajan, S., Hegde, Rajendra, Janani, N., Singh, S.K., 2019a. The need for digital soil mapping in India. Geoderma Reg. 15. https://doi.org/10.1016/j.geodrs.2019.e00204.

Dharumarajan, S., Hegde, R., Lalitha, M., Kalaiselvi, B., Singh, S.K., 2019b. Pedotransfer functions for predicting soil hydraulic properties in semi-arid regions of Karnataka Plateau, India. Curr. Sci. 116 (7), 1237–1246.

Gallo, B.C., José, A.M., Rizzo, R., Safanelli, J.L., Mendes, W. de S., Igo, F., Marcus, V., Danilo, J., Lacerda, M.P.C., 2018. Multi-temporal satellite images on topsoil attribute quantification and the relationship with soil classes and geology. Remote Sens. 10, 1571. https://doi.org/10.3390/rs10101571.

Gastaldi, G., Minasny, B., McBratney, A.B., 2012. Mapping the occurrence and thickness of soil horizons within soil profiles. In: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), Digital Soil Assessments and beyond. CRC Press, Balkema, London, pp. 145–148.

Gessler, P.E., Chadwick, O.A., Chamran, F., Althouse, L., Holmes, K., 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Sci. Soc. Am. J. 64, 2046–2056.

Ghaemi, M., Astaraei, A.R., Sanaeinejad, S.H., Zare, H., 2013. Using satellite data for soil cation exchange capacity studies. Int. Agrophys 27, 409–417. https://doi.org/10.2478/intag-2013-0011.

Gholizadeh, A., Žižala, Daniel, Saberioon, Mohammadmehdi, Borůvka, Luboš, 2018. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. Remote Sens. Environ. 218, 89–103.

Gomez, C., Dharumarajan, S., Féret, J.B., Lagacherie, P., Ruiz, L., Sekhar, M., 2019. Use of sentinel-2 time-series images for classification and uncertainty analysis of inherent biophysical property: case of soil texture mapping. Remote Sens. 11, 565.

Grundy, M.J., Viscarra Rossel, R.A., Searle, R.D., Wilson, P.L., Chen, C., Gregory, L.J., 2015. Soil and landscape grid of Australia. Soil Res 53, 835–844.

Grunwald, S., Thompson, J.A., Boettinger, J.L., 2011. Digital soil mapping and modeling at continental scales: finding solutions for global issues. Soil Sci. Soc. Am. J. 75, 1201–1213.

Hong, S.Y., Minasny, Budiman, Han, Kyung Hwa, Kim, Yihyun, Lee, Kyungdo, 2013. Predicting and mapping soil available water capacity in Korea. Peer J. 1, e71. https://doi.org/10.7717/peerj.71.

Jackson, M.L., 1962. Soil Chemical Analysis. Prentice Hall, Englewood Cliffs, New York, U.S.A.

Jackson, M.L., 1973. Soil Chemical Analysis. Prentice Hall of India Pvt. Ltd, New Delhi.

De, C.J.W., Lagacherie, P., da Silva Chagas, C., Filho, B., Bhering, S.B., 2014. A regional-scale assessment of digital mapping of soil attributes in a tropical hillslope environment. Geoderma 232–234, 479–486.

Lagacherie, P., McBratney, A.B., 2007. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: Lagacherie, P., McBratney, A.B., McBratney, A.B., Mendonça-Santos, M.L., Minasny, B. (Eds.), 2003. On Digital Soil Mapping. Geoderma. 117, pp. 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4.

Meinshausen, N., 2006. Quantile regression forests. J. Mach. Learn. Res. 7, 983–999.

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

Mulder, V.L., Lacoste, M., Richer de Forges, A.C., Martin, M.P., Arrouays, D., 2016. National versus global modelling the 3D distribution of soil organic carbon in mainland France. Geoderma 263, 13–34.

Poggio, L., Gimona, A., 2017. 3D mapping of soil texture in Scotland. Geoderma Reg. 9, 5–16. https://doi.org/10.1016/j.geodrs.2016.11.003, 2017.

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

Richards, L.A., 1954. Diagnosis and improvement of saline and alkali soils. USDA, Hand Book No, 60, p. 101.

Richards, L.A., Gardner, W.R., Ogata, G., 1956. Physical processes determining water loss from soil. Soil Sci. Soc. Am. Proc. 20, 310–314.

Román Dobarco, M., Bourennane, H., Arrouays, D., Saby, N.P.A., Cousin, I., Martin, M.P., 2019. Uncertainty assessment of GlobalSoilMap soil available water capacity products: a French case study. Geoderma 344, 14–30. https://doi.org/10.1016/j.geoderma.2019.02.036.

Sanchez, P.A., Ahamed, S., Carr'e, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., Mendonc, De Lourdes, Santos, M., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vagen, T.G., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A., Zhang, G.L., 2009. Digital soil map of the world. Science 325, 680–681. https://doi.org/10.1126/science.1175084.

Santra, P., Kumar, M., Panwar, N., 2017. Digital soil mapping of sand content in arid western India through geostatistical approaches. Geoderma Reg. 9, 56–72.

Vaudour, E., Gomez, C., Fouad, Y., Lagacherie, P., 2019. Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems. Remote Sens. Environ. 223, 21–33.

Vaysse, K., Lagacherie, P., 2017. Using quantile regression forest to estimate uncertainty of digital soil mapping products. Geoderma 291, 55–64. https://doi.org/10.1016/j.geoderma.2016.12.017.

Walkley, A., 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.