Prediction of Soil Properties Using Quantile Regression Forest Machine Learning Algorithm – A Case Study of Salem and Rasipuram Block, Tamil Nadu, India.

ABSTRACT:

Digital soil mapping is a growing technology for mapping soil properties instead of conventional soil mapping. Especially for what are all countries have large geographical areas and human, not accessible areas. Compared to conventional soil mapping it is cost-wise less and more accurate. At the world level, globalsoilmap.net has taken the initiative for creating digital soil maps. In India like countries very much needed for digital soil mapping, is essential for agricultural planning, and decision-makers decide on it. This study predicted the soil properties such as sand, silt, clay, pH, and OC using the Quantile Regression Forest machine learning algorithm also provides uncertainty. The main aim of this study was to predict the soil properties in the top two depth intervals such as surface and subsurface. For achieving this goal, 56 soil samples were collected across the study area, and many environmental covariates were used for that such as DEM derivatives, satellite imagery, and Climatic Data. This study, using 56 soil samples data taken from the traditional soil survey, is a limited number of soil samples this tried to achieve a higher accuracy result using QRF.

***Keywords**: Soil Properties, Environmental covariates, Quantile Regression Forest, Soil samples.*

1. INTRODUCTION:

Soils are an important part of the environment and necessary for life, it's a habitat for many living organisms, provide nutrients and a medium for plant growth, work as a filtration system for surface water, store Carbon and maintain atmospheric gases. So, it should be maintained and monitored properly. Continuous digital soil information is needed to make most environment models, especially for the larger-scale area [3, 6, 10, 13]. This digital soil information is always not available at the needed scale [4] and high-accuracy mapping needs more time, and cost and is always challenging [21].

“Digital Soil Mapping (DSM) also known as Predictive Soil mapping is the creation of a digital soil information system using numerical models, soil observation, and their related environmental variables” [9]. “Digital maps of soil carbon stock are replacing conventional polygon-based soil maps. DSM techniques can be used not only to estimate SOC stock but also to quantify associated uncertainties in these estimates. DSM techniques can therefore be used to map new areas and upgrade the quality of the previous mapping. As these digital soil maps are stored in digital spatial formats, they have the advantage that numerous geospatial data handling tools in a spatial soil information system can be used to analyze and interpret large volumes of data” [2, 11, 12]. In DSM, soil properties are predicted through the quantitative relationship between measured soil attributes and soil environmental covariates based on the SCORPAN model [10]. The Digital Soil Mapping working group of the International Union of Soil Sciences (IUSS) has taken the initiative of the GlobalSoilMap project, the main objective of the project is to make new digital soil map at the world level using state-of-the-art and advanced technologies for mapping soil and predicting soil attributes at a fine resolution [1]. According to the GlobalSoilMap scenario, can be predicted 12 soil properties along with uncertainties over the six depth intervals.

“Recent day usage of machine learning approaches in DSM has increased. Compared to multiple machine learning techniques RF gives better results in the prediction of soil properties. Soil organic carbon concentrations, clay content, and pH were predicted with RF” [5, 22]. RTs on the other hand are widely applied. McKenzie and Ryan (1999) used them to predict soil properties from terrain attributes and gamma radiometric surveys. QRF is an expansion of the Random Forest technique; the main objective of this study is to predict a surface soil property (sand. silt, clay, and pH) along with uncertainty using Quantile Regression Forest techniques with the use of limited soil sample data. Similar work was discussed and carried out by several researchers, proving the efficiency of the model in DSM [17, 18, 19]. DSM techniques are very much needed for large countries like India because conventional soil mapping needs more money, is time-consuming, and has less accuracy compared to Digital Soil Mapping.

2. STUDY AREA:

The paper discusses, derivative work of digital soil mapping, wherein soil properties were mapped using the SCORPION model for Salem and Rasipuram blocks (figure 1). Salem lies in the foothills of Shevaroy hills which houses the famous hill Station 'Yercaud'. The soil in the district can be broadly classified into six major soil types viz., Red in-situ, Red Colluvial Soil, Black Soil, Brown Soil, Alluvial and Mixed Soil. The major part of the district is covered by Red in-situ and Red Colluvial soils. Black soils are mostly seen in Salem, Attur, Omalur and Sankari taluks. Brown Soil occupies a major portion of Yercaud and parts of Salem and Omalur taluks. Rasipuram block is in the Namakkal district of Tamil Nadu. The district identifies red loam, black soil, laterite soil, sandy coastal alluvium, red sandy soil, and clay loam. The black or regar lands are considered the most fertile. It absorbs moisture from the atmosphere, retains it, and helps the cultivation of different crops.

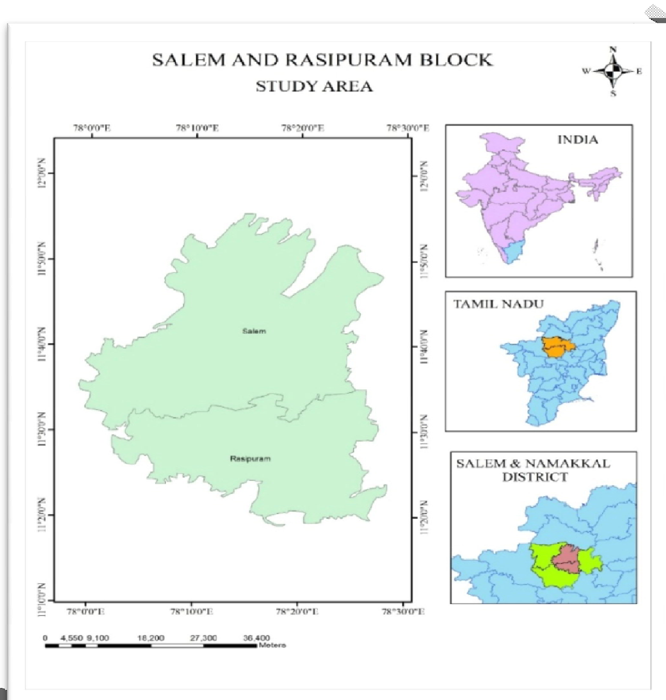

Figure 1. Study area

3. MATERIALS AND METHODS:

Figure 2 explains the workflow in the form of a flow chart. The data necessary for digital soil classification were collected from secondary sources provided by Tamil Nadu for Salem and Rasipuram blocks (Table 1). Then the necessary parameters for the modelling are derived from the data. Finally using QRF, soil properties are mapped.

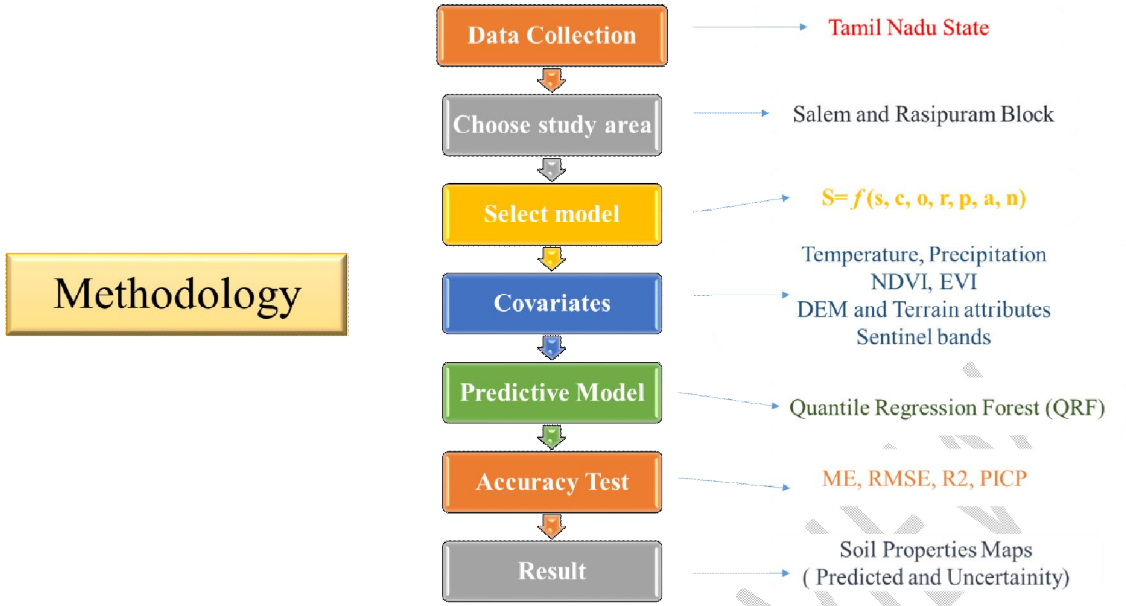

Figure 2. Methodology flowchart

Table 1. Environmental variables used for prediction and their sources

Predictor	Source	Resolution	Type	Range
		Terrain attributes		
• Elevation	ALOS PALSAR DEM	12.5m	Q	101 – 1546
• Slope	ALOS PALSAR DEM	12.5m	Q	0 - 79.04
• Aspect	ALOS PALSAR DEM	12.5m	Q	0.04 - 6.28
• TPI	ALOS PALSAR DEM	12.5m	Q	-52.1 – 48.4
• TWI	ALOS PALSAR DEM	12.5m	Q	0.10– 23.32
• Plan curvature	ALOS PALSAR DEM	12.5m	Q	-0.20 – 0.17
• Profile curvature	ALOS PALSAR DEM	12.5m	Q	-0.15– 0.16
• MrVBF	ALOS PALSAR DEM	12.5m	Q	0 – 6.73
• MrRTF	ALOS PALSAR DEM	12.5m	Q	0 – 5.75
• LS-Factor	ALOS PALSAR DEM	12.5m	Q	0 – 86.37
		Vegetation attributes		
• NDVI	MOD13Q1(2019)	250m_16days	Q	380-9790
• EVI	MOD13Q1(2019)	250m_16days	Q	383 - 9712
		Climate		
• Annual mean temperature	worldclime2	340sq.km_(1970-2000)	Q	24.8- 28.6
• Isothermality	worldclime2	340sq.km_(1970-2000)	Q	53.9- 59.5
• Precipitation	worldclime2	340sq.km_(1970-2000)	Q	783-906
		Remote sensing imagery		
• Band (1-13)	Sentinal-2 (2019)	10m	Q	

3.1. Estimation of model parameters

A different set of environmental covariates were used for predicting the soil properties. In arid and semi-arid regions relief or topography is one of the main soil farming factors [17].

The digital elevation model was obtained from the Phased Array type L-band Synthetic Aperture Radar (ALOS PALSAR) data. The Derivatives of DEM like Elevation, Slope, Aspect, Topographic wetness index, Topographic position index, Curvatures (Plan and Profile), Multiresolution valley bottom flatness (MrVBF), and Multiresolution ridge flatness (MrRTF) were derived using SAGA GIS. Vegetation/organism is important for predicting soil organic carbon. So, the Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) were derived from MODIS 13Q data. In additional remote sensing imageries sentinel 2 all bands download from USGS and climatic data such as mean temperature, and precipitation data downloaded from Worldclimate.net are used for prediction. Also, the soil properties of sample data of the region of interest (ROI) are incorporated as part of the model. Totally 56 soil samples are collected in the study area. Surface and sub-surface soil samples (0-30 and 30-60cm respectively) of the study area were used for mapping of soil properties of the study area (USDA 2017). Summary statistics of surface and subsurface soil properties are shown in Tables 2 and 3. This can be further used in validating the prediction results by comparing them with statistics post prediction.

Table 2. Summary statistics on the surface properties of the samples collected

Statistics	PH	OC (%)	sand	clay	silt
Min	6.4	0.11	40	5	1.8
Max	9.1	1.3	93	48	15
Mean	8.15	0.498	61.132	30.181	8.358
S.D	0.612	0.234	12.43	9.108	3.633
Skewness	-0.990	0.91	0.165	-0.188	0.009
Kurtosis	0.311	1.459	-0.583	-0.101	-0.726

Table 3. Summary statistics on the sub-surface properties of the samples collected

Statistics	OC (%)	sand	clay	silt
Min	0.07	40	16	2.8
Max	0.93	78	45	18
Mean	0.356	56.148	33.892	10.194
S.D	0.172	9.649	6.716	3.93
Skewness	0.885	0.268	-0.557	0.079
Kurtosis	1.689	-1.06	-0.497	-0.785

3.2. SCORPON model

Digital soil mapping is performed using the SCORPAN model. The word SCORPAN is a mnemonic of the parameters of the empirical relationship between soil and environmental factors to use these as soil spatial prediction functions for DSM. [10] proposed the SCORPAN model (figure 3), where soil (as either soil classes, S_c, or soil attributes, S_a) at a point in space and time is an empirical quantitative function of seven environmental covariates: soil (s), climate (c), organisms (o), relief (r), parent material (p), age (a), and spatial location (n):

$$S_{c,a} = f(s, c, o, r, p, a, n)$$

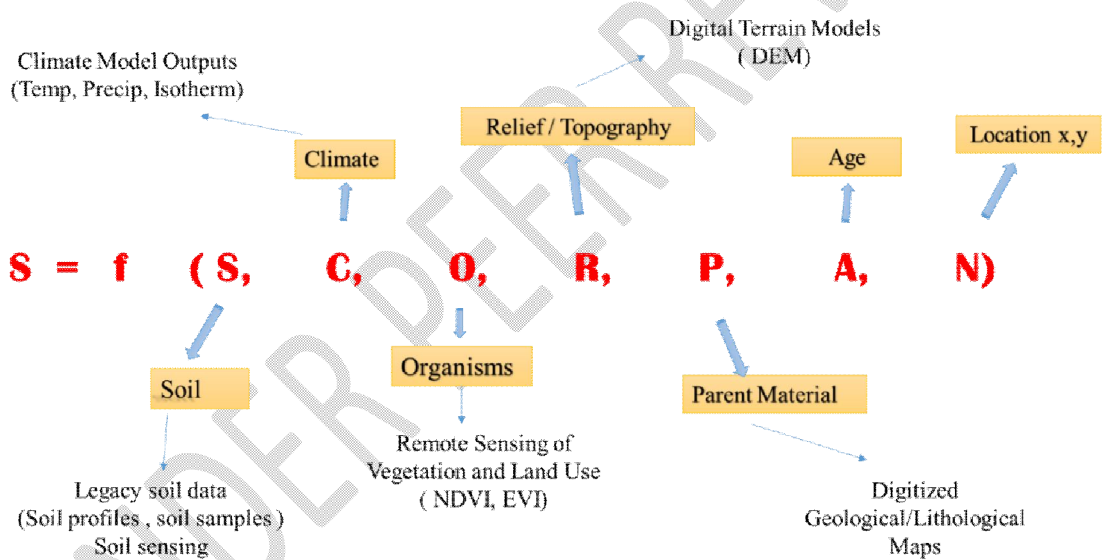

Figure 3. Illustration of SCORPAN model

In the commonly used SCORPAN model of DSM, predictor variables are represented implicitly or explicitly by a combination of one or more categorical or continuous variables [10]. For example, [1,8] used a range of surrogate measures of climate – humidity indices, mean/max/median rainfall, and air temperature which could be used to represent 'C' factors of the SCORPAN model. Because the relationships are based on the soil observations, the quality of the resulting soil map depends also on the soil observation quality. Usually, a digital soil mapper tries to optimize the accuracy of the models and minimize errors [7].

3.3 Quantile regression forest (QRF)

The quantile regression forest (QRF) model was used for the prediction of soil properties and uncertainty estimates in the study area. QRF is an extension of the Random Forest model and the advantage of QRF over the 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 this information. Many studies proved RF algorithm performed better in the prediction of soil properties. The random forest model basic assumption further improved as a quantile regression forest [15]. The main difference between QRF and RF is Random Forest keeps only the observations of mean values for each node in each tree, that falls into this node and ignores all other information whereas QRF keeps the all-observation values in this node, not just their mean and assesses the conditional distribution based on this information. In this study randomforesSRC and quantregForest packages are used for prediction in R studio. After the execution of the model, model performance was evaluated through the values of coefficient of determination (R2), Mean Error, Root Mean Square Error (RMSE), and uncertainty accuracy estimated via PICP.

4. RESULTS AND DISCUSSIONS:

4.1. Environmental variables

Different environmental parameters were derived from DEM data (figure 4) and are used in the model. The slope is the measure of steepness or the degree of inclination of a feature relative to the horizontal plane. This map provides a colorized representation of the slope (figure 5). Aspect is the compass direction that a slope faces. Its values indicate the directions the physical slopes face. Aspect direction is classified based on slope angle with a descriptive direction. An output aspect raster will typically result in several slope direction classes as shown in figure 6.

The topographic Position Index (TPI)compares the elevation of each cell in a DEM to the mean elevation of a specified neighbourhood around that cell (figure 7). Topographic Wetness Index (figure 8) also known as the compound topographic index is a steady state wetness index. It is commonly used to quantify topographic control on hydrological processes. The index is a function of both the slope and the upstream contributing area per unit width orthogonal to the flow direction.

Profile curvature is parallel to the direction of the maximum slope (figure 9). A negative value indicates that the surface is upwardly convex at that cell. A positive profile indicates that the surface is upwardly concave at that cell. A value of zero indicates that the surface is linear. Profile curvature affects the acceleration or deceleration of flow across the surface. Planform curvature (commonly called plan curvature) is perpendicular to the direction of the maximum slope. A positive value indicates the surface is sideward convex at that cell. A negative plan indicates the surface is sideward concave at that cell. A value of zero indicates the surface is linear. Profile curvature relates to the convergence and divergence of flow across a surface. (figure 10). Other variables are multiresolution ridge top flatness (MrRTF), multiresolution index of valley bottom flatness (MrVBF), and Slope length and steepness (LS) factor as shown in figures 11, 12, and 13.

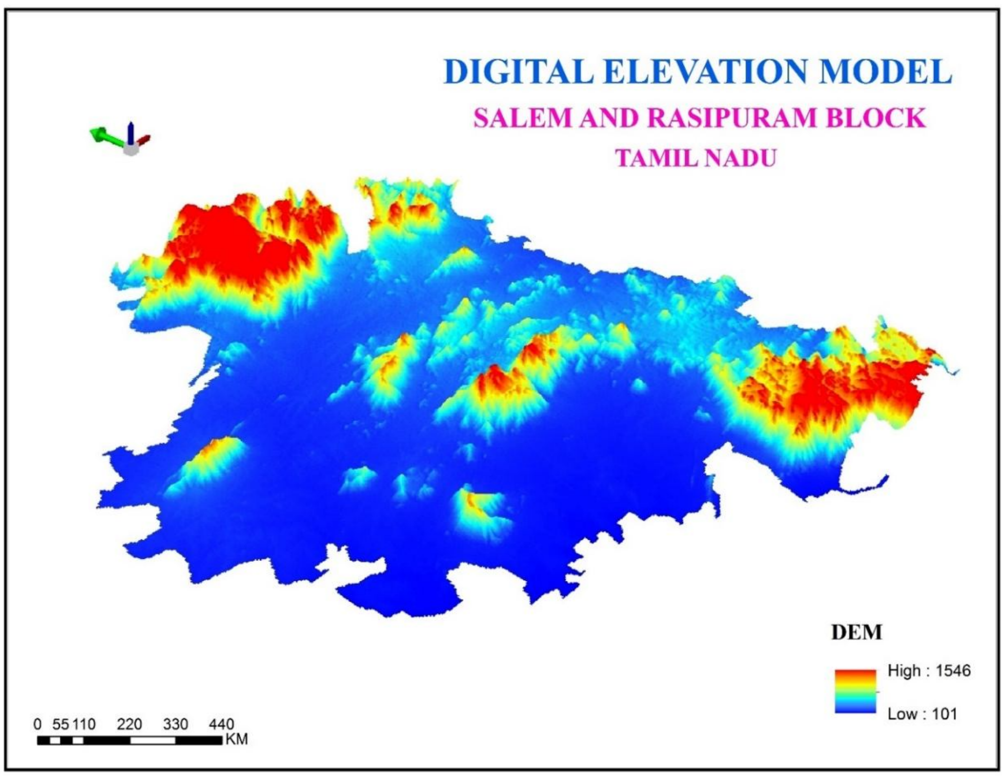

Figure 4. DEM visualization of Salem and Rasipuram block

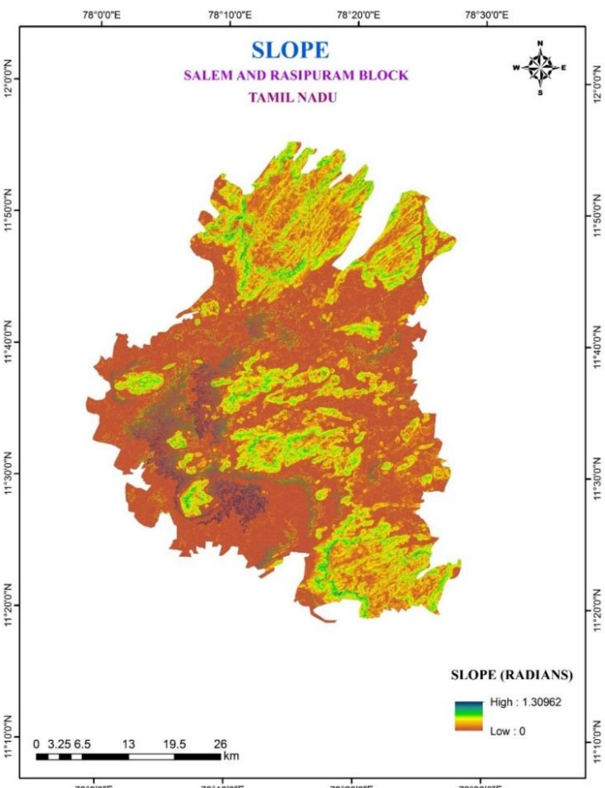

Figure 5. Slope map

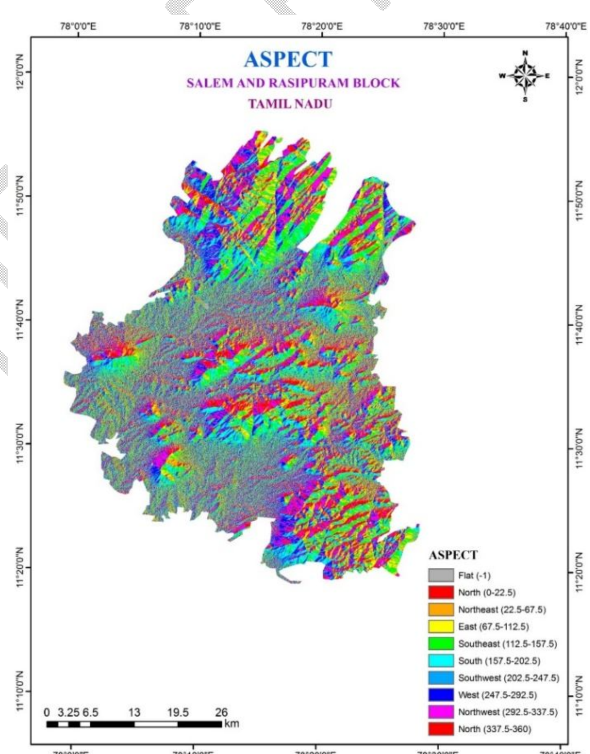

Figure 6. Aspect map

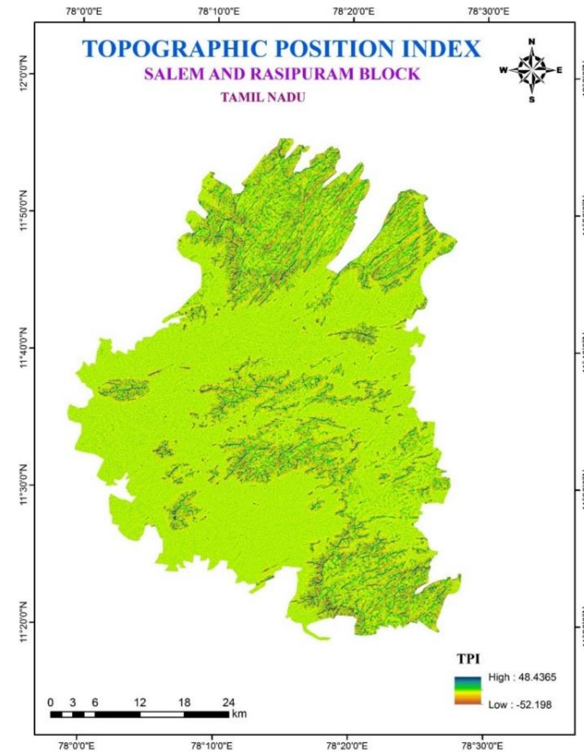

Figure 7. TPI map

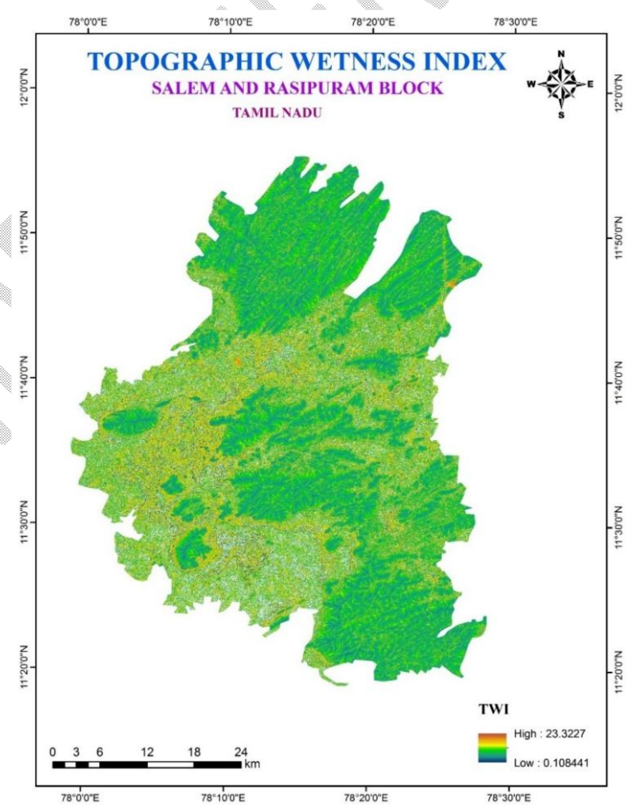

Figure 8. TWI map

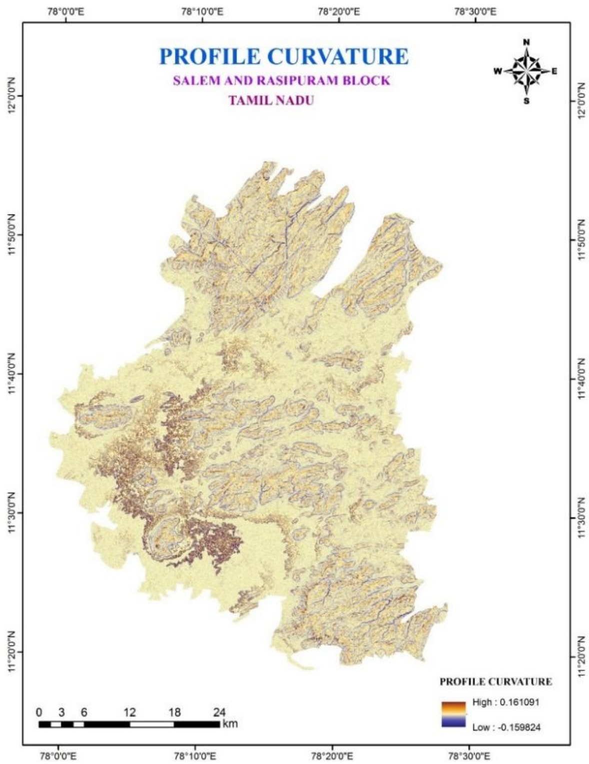

Figure 9. Profile curvature from DEM

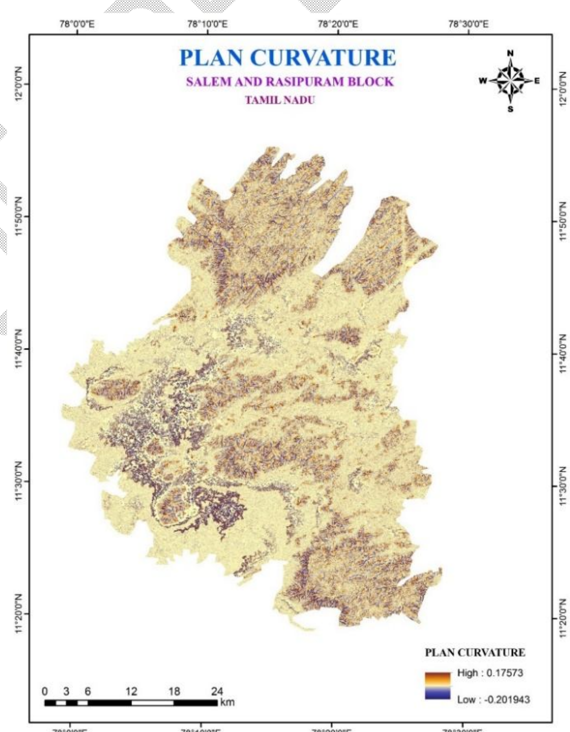

Figure 10. Plan curvature from DEM

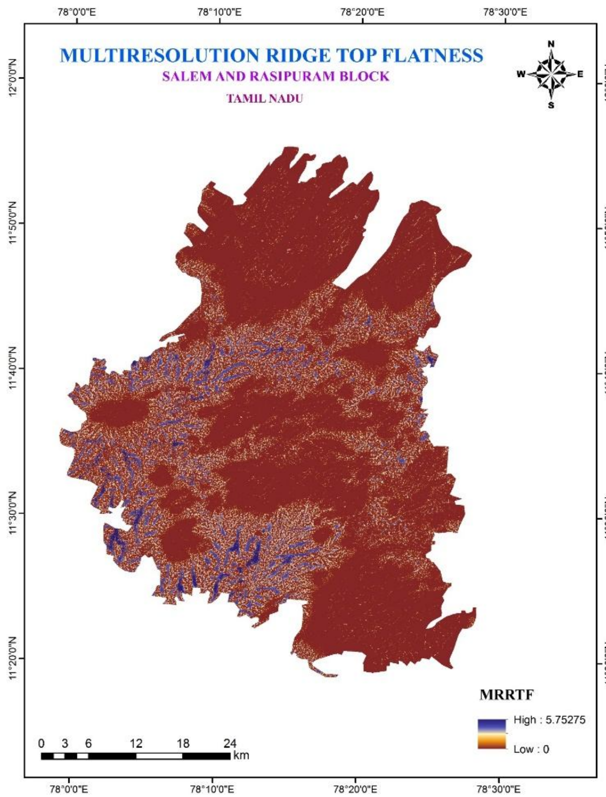

Figure 11. MrRTF derived from DEM for Salem and Rasipuram block

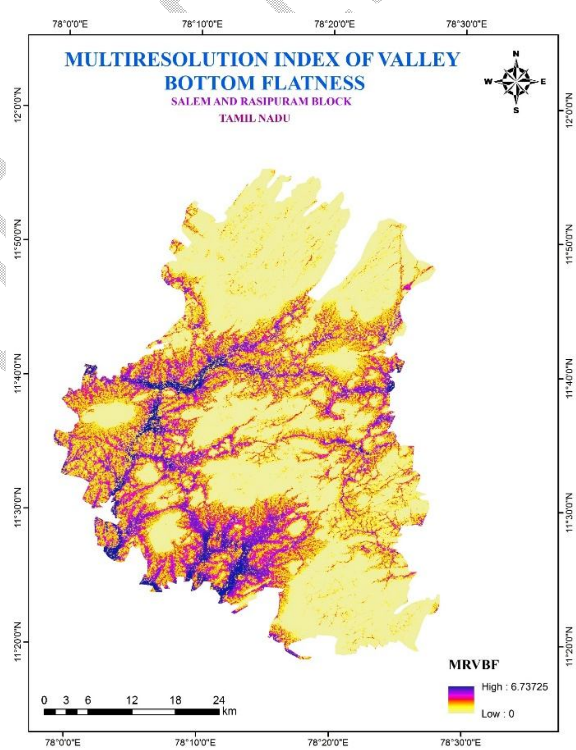

Figure 12. MrVBF derived from DEM for Salem and Rasipuram block

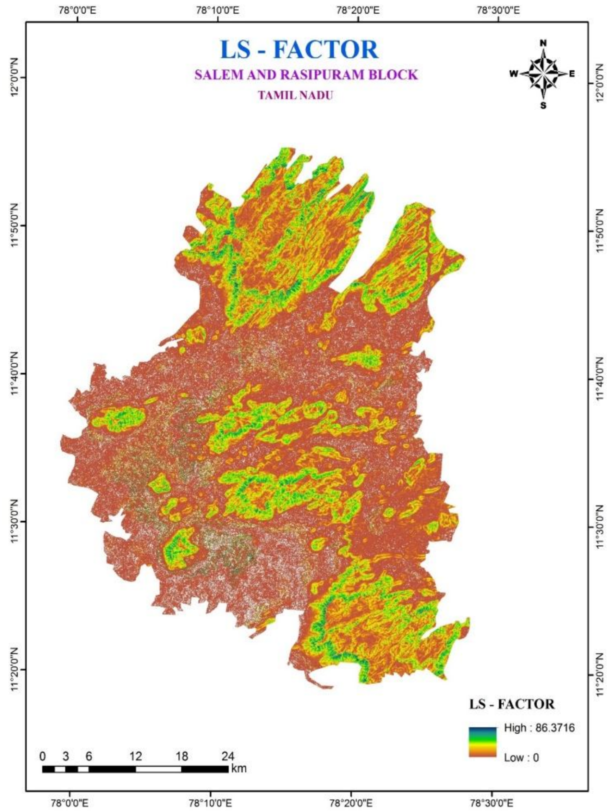

Figure 13. Slope length and steepness factor

4.2. Other Covariates

In addition to the environmental variables, other parameters act as part of the model in predicting soil properties. These include Vegetation attributes namely NDVI, EVI, and climate variables like annual mean temperature, isothermally, and precipitation. Based on NDVI and EVI, the study area was classified of sparse to dense vegetation. The health of the vegetation is also an indirect indicator of the presence of organic carbon content, one of the soil properties to be predicted. From figures 14 and 15, it can be inferred that both the index provided a similar pattern of vegetation cover.

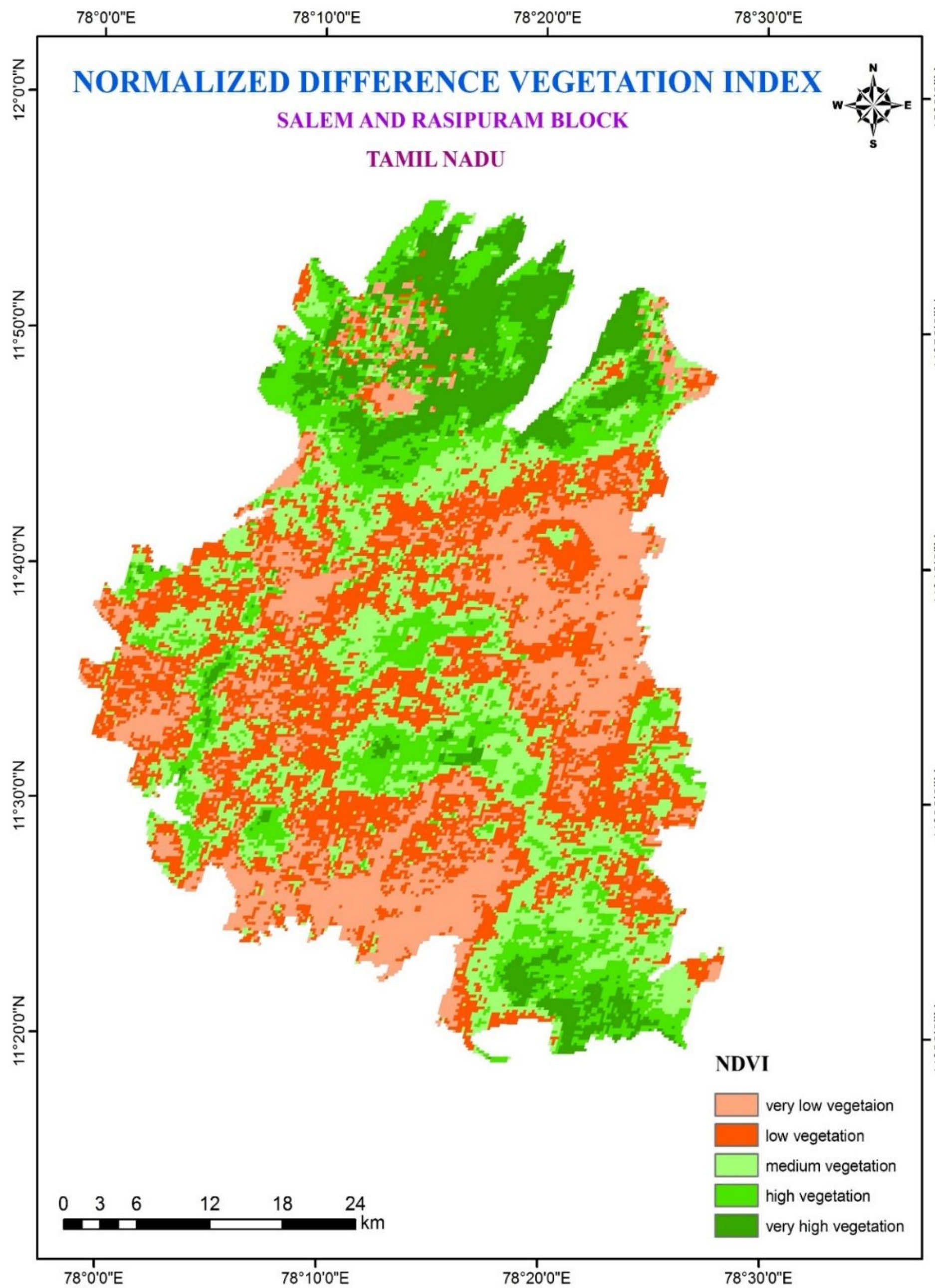

Figure 14. NDVI map

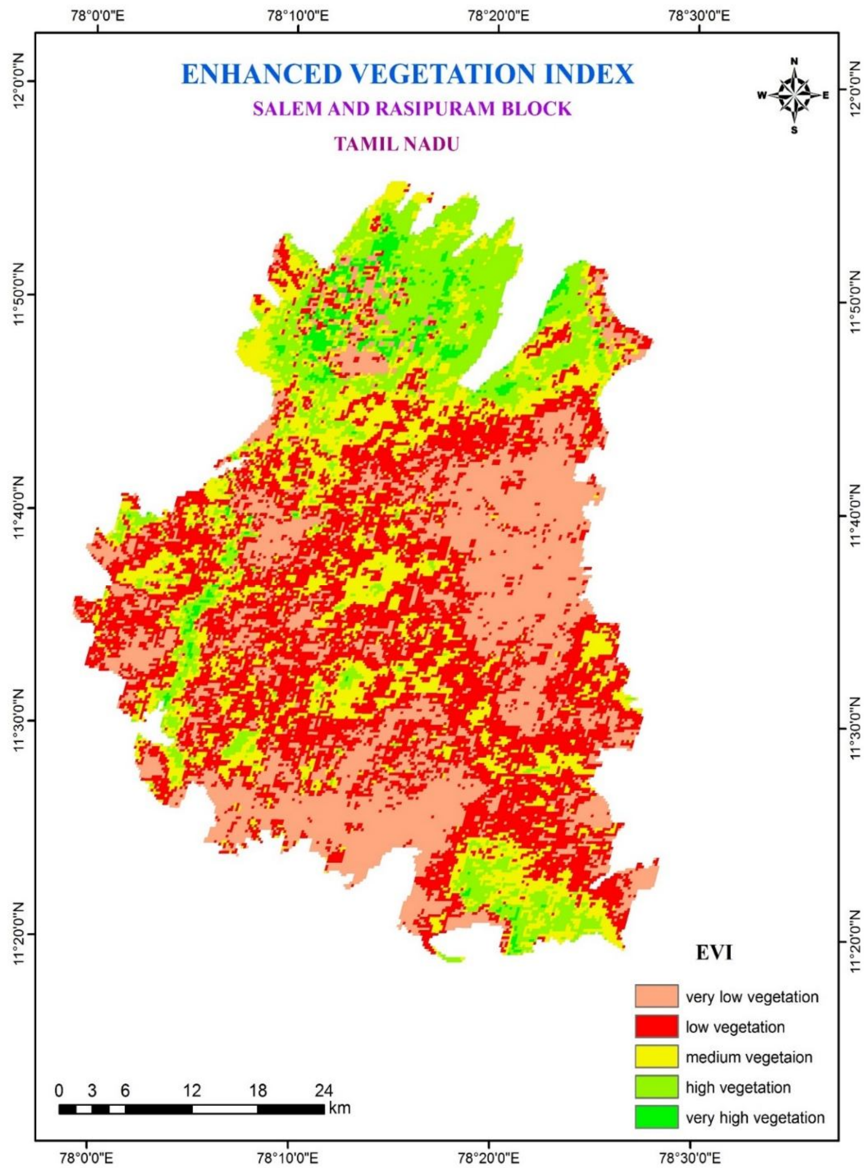

Figure 15. EVI map

4.4. Prediction of Soil properties

The model was run to predict the composition of sand, silt, and clay, and pH, Organic carbon for surface and composition of sand, silt, and clay, and organic carbon for sub-surface. The prediction results and uncertainty of prediction are mapped as shown in figures 16, 19, 22, 25, 28, 31, 34, 37, and 40. Scatterplots of observed against predicted were plotted for surface and sub-surface properties as shown in figures 17, 20, 23, 26, 29, 32, 35, 38, and 41. The importance of variables against each soil property was also plotted in the form of a bar chart as shown in figures 18, 21, 24, 27, 30, 33, 36, 39, and 42.

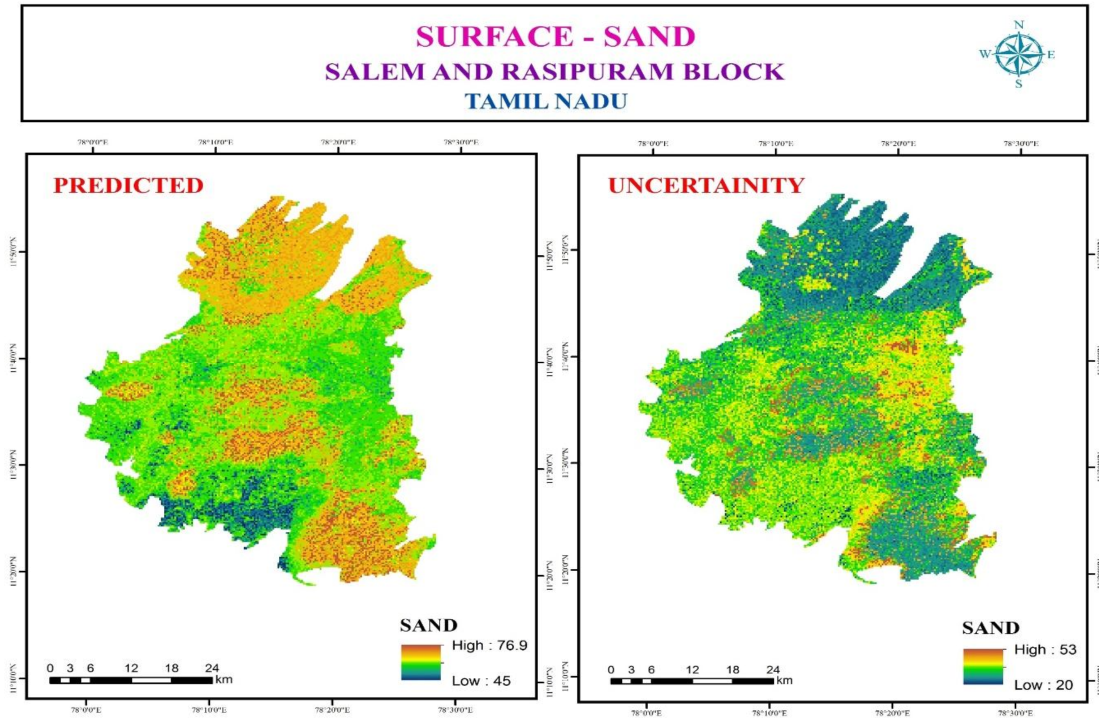

Figure 16. Prediction of surface – sand

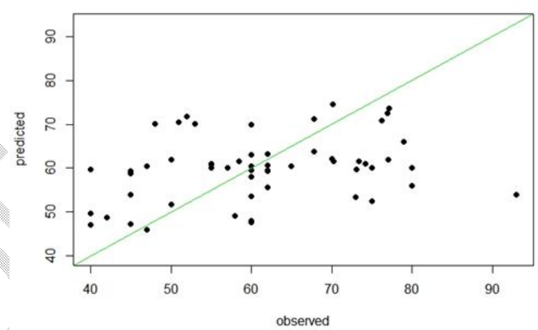

Figure 17. Scatterplot of surface – sand

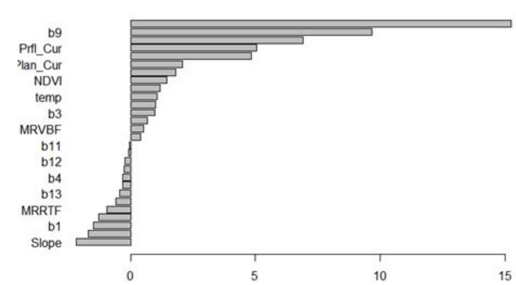

Figure 18. Variable importance of surface - sand

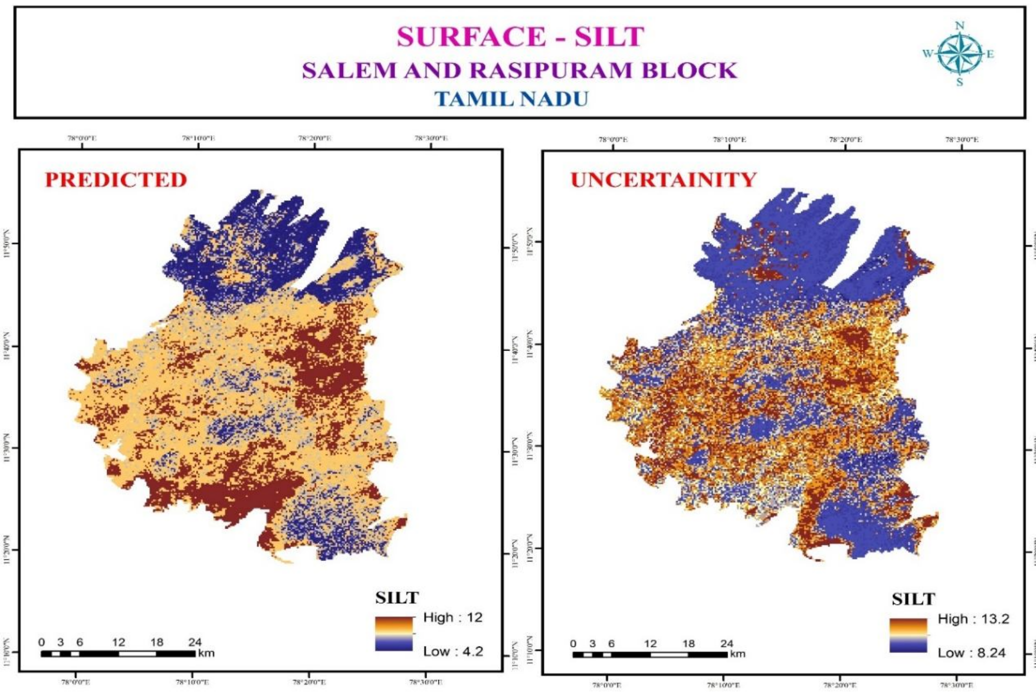

Figure 19. Prediction of Surface – silt

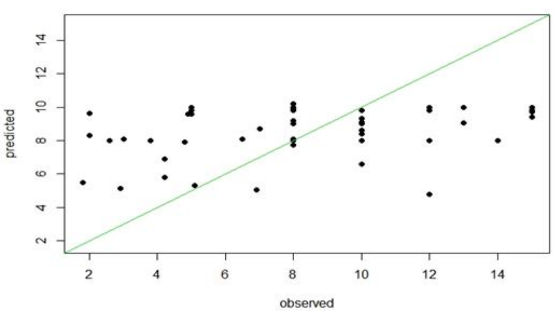

Figure 20. Scatterplot of surface – silt

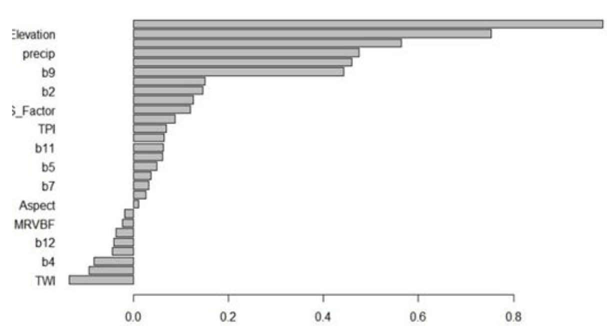

Figure 21. Variable importance of surface - silt

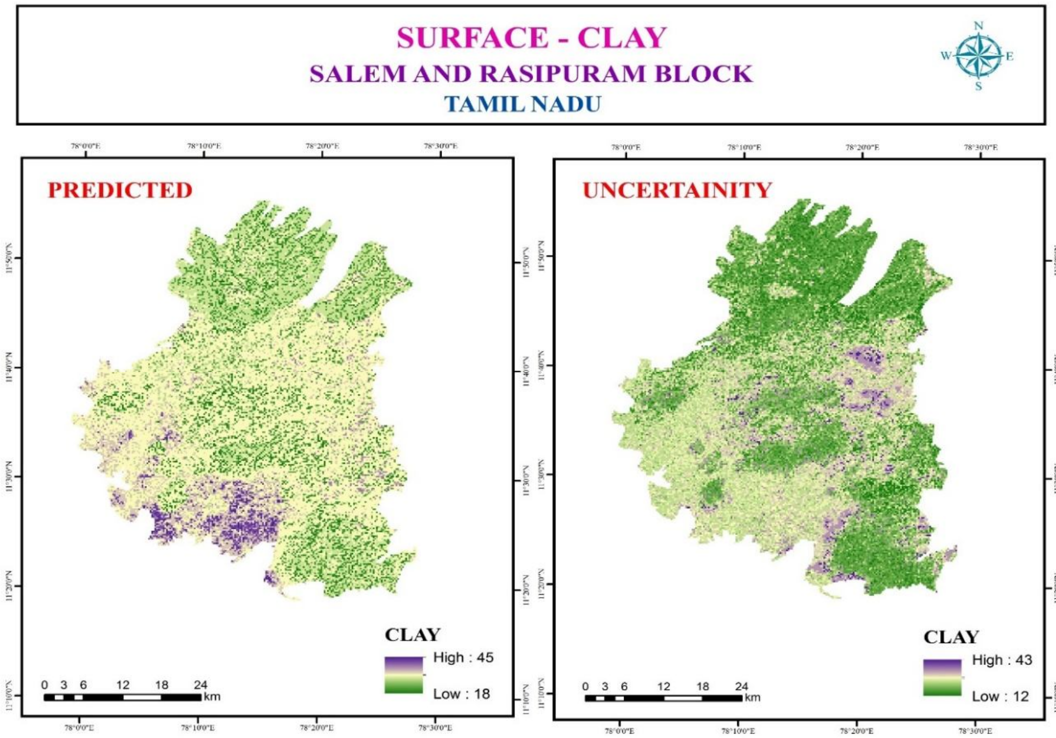

Figure 22. Prediction of Surface – clay

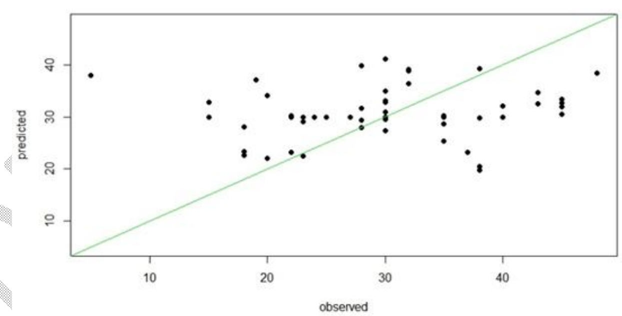

Figure 23. Scatterplot of surface–clay

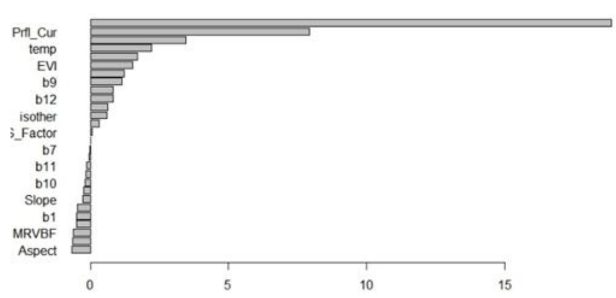

Figure 24. Variable importance of surface - clay

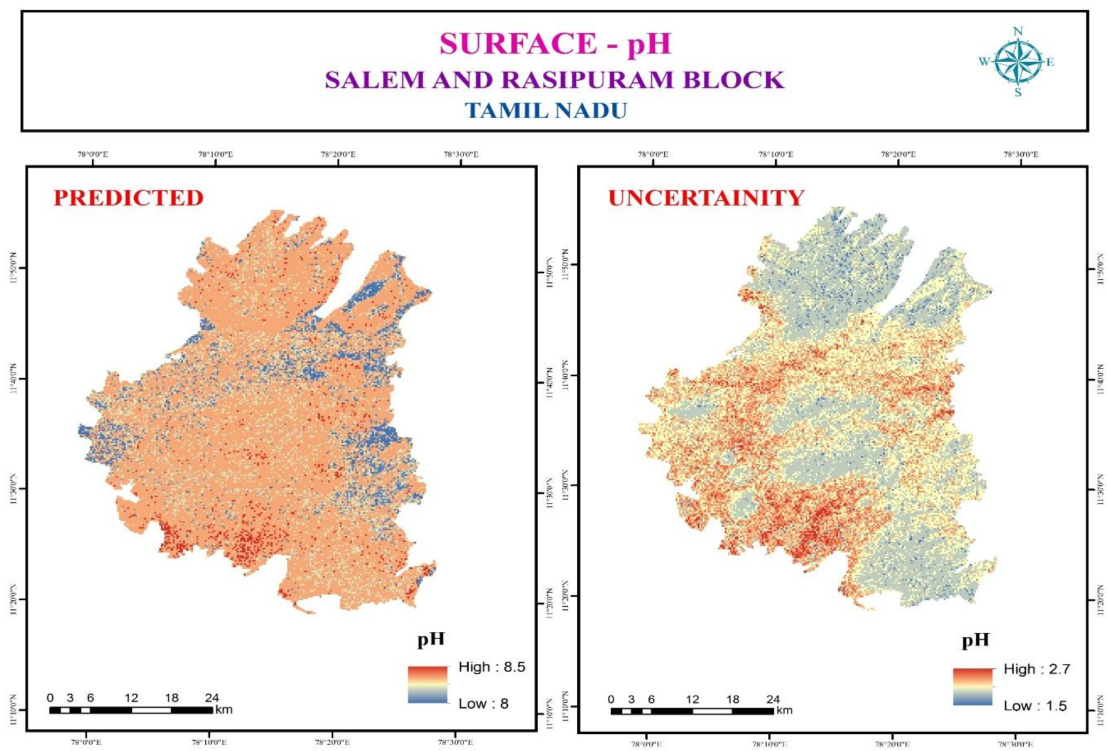

Figure 25. Prediction of Surface – pH

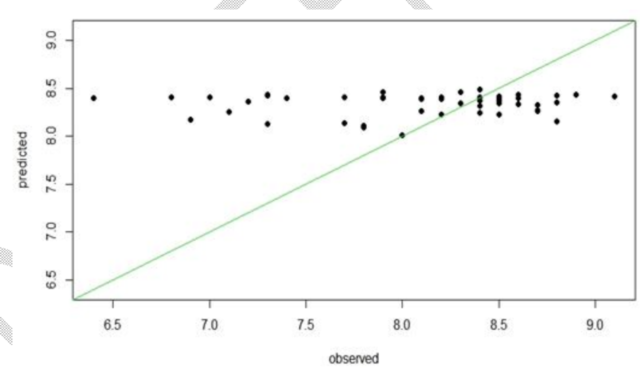

Figure 26. Scatterplot of surface – pH

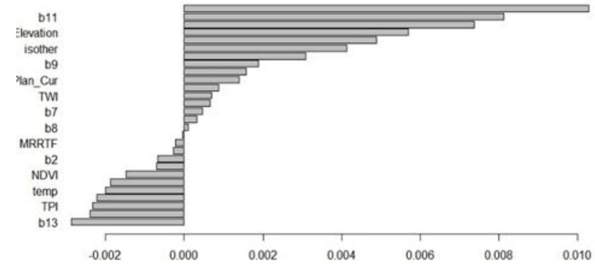

Figure 27. Variable importance of surface - pH

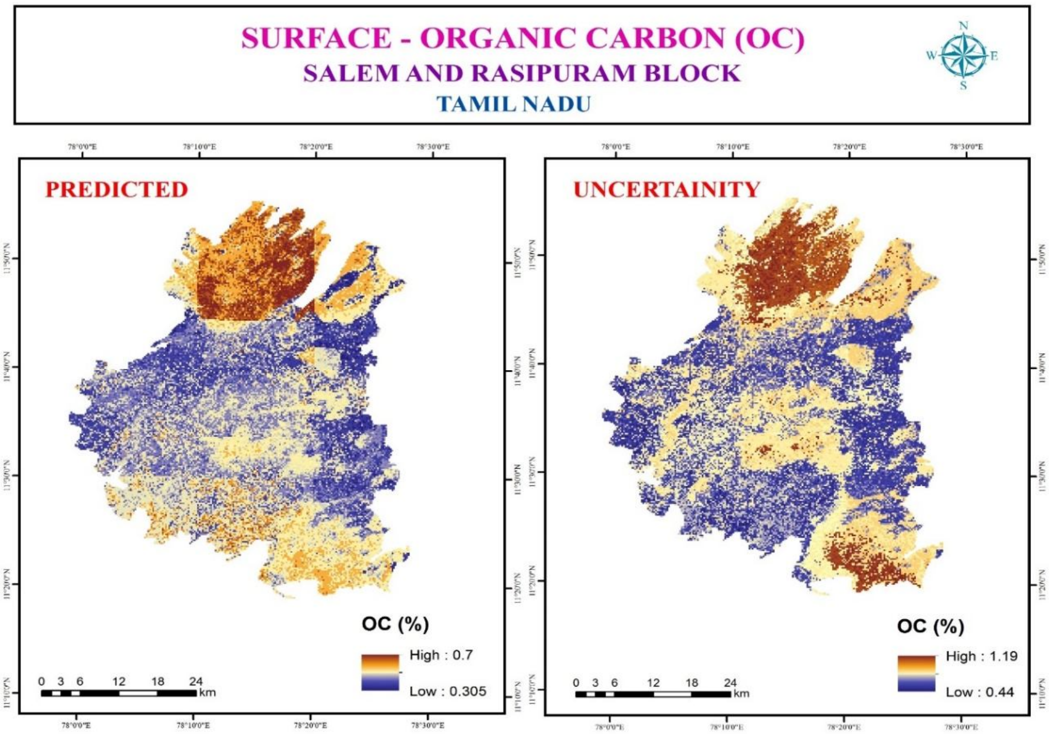

Figure 28. Prediction of Surface-Organic carbon

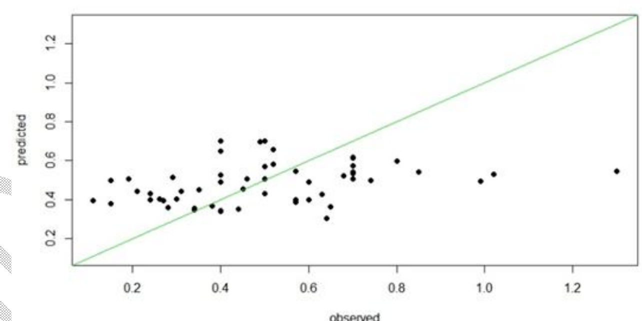

Figure 29. Scatterplot of surface – organic carbon

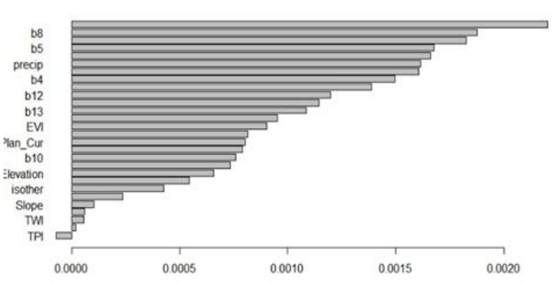

Figure 30. Variable importance of surface – organic carbon

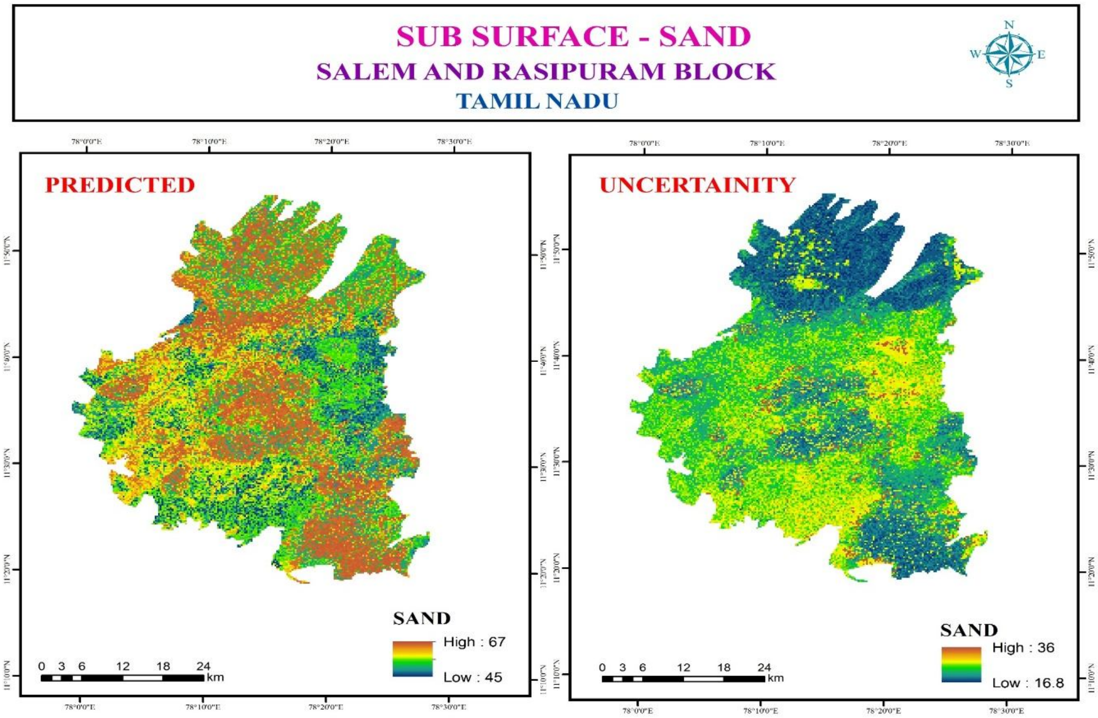

Figure 31. Prediction of Sub surface – sand

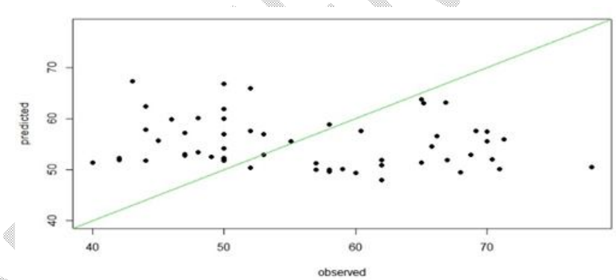

Figure 32. Scatterplot of Sub surface – sand

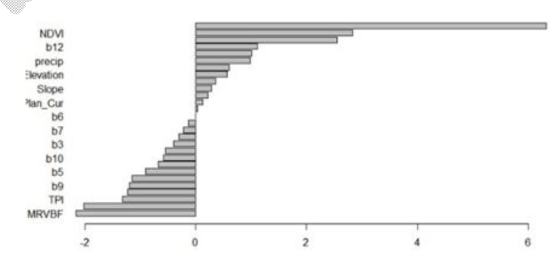

Figure 33. Scatterplot of Sub surface - sand

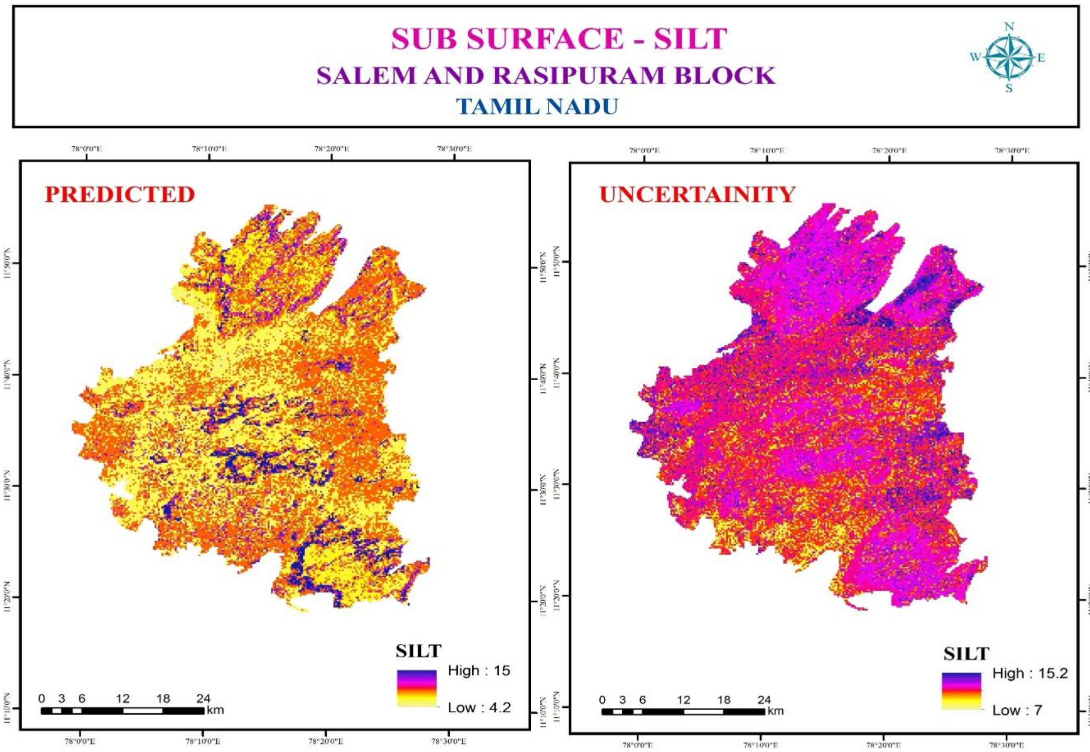

Figure 34. Prediction of Sub surface – silt

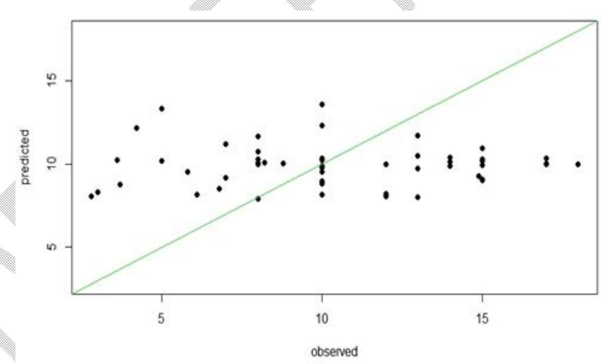

Figure 35. Scatterplot of Sub surface – silt

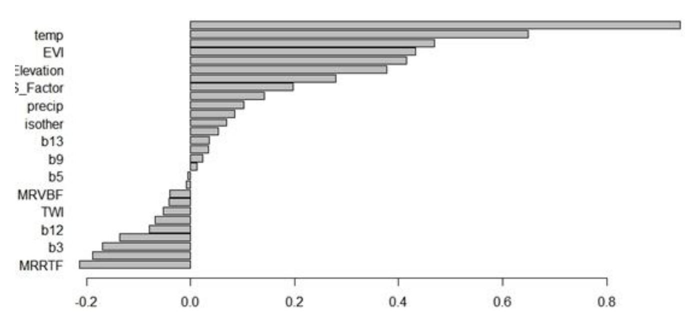

Figure 36. Variable importance of Sub surface - silt

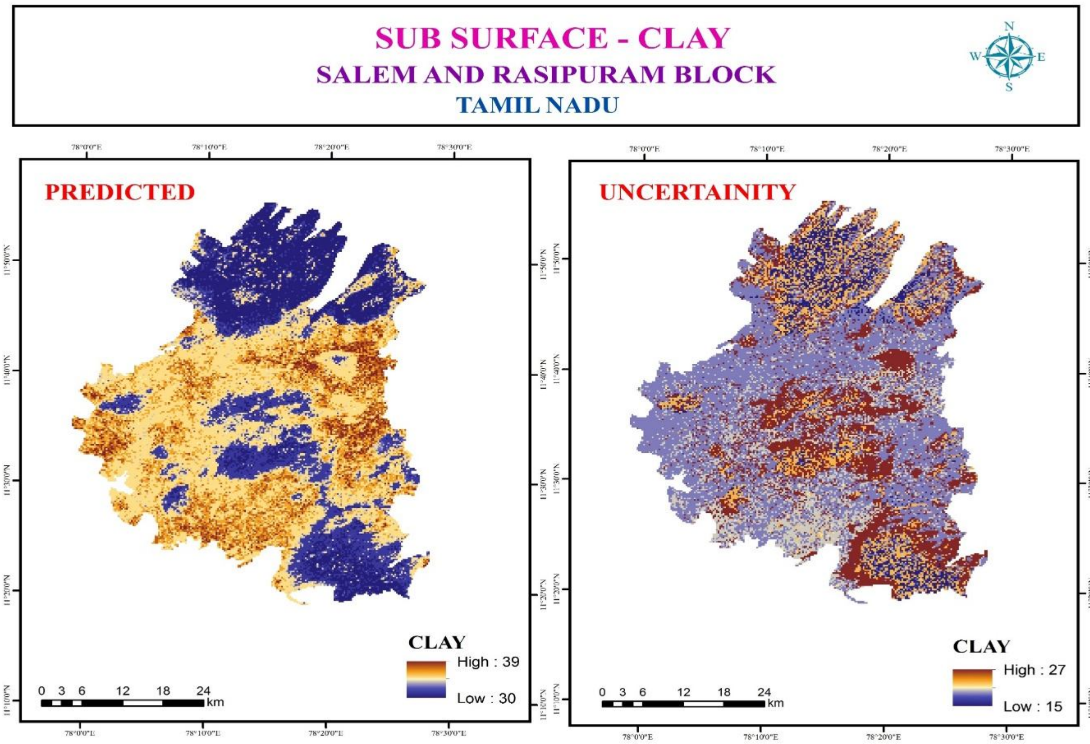

Figure 37. Prediction of Sub surface – clay

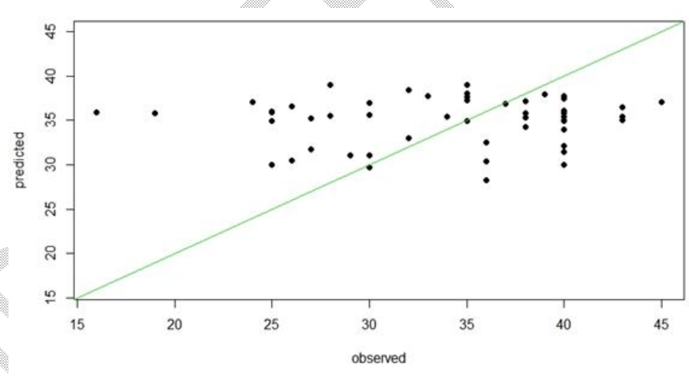

Figure 38. Prediction of Sub surface – clay

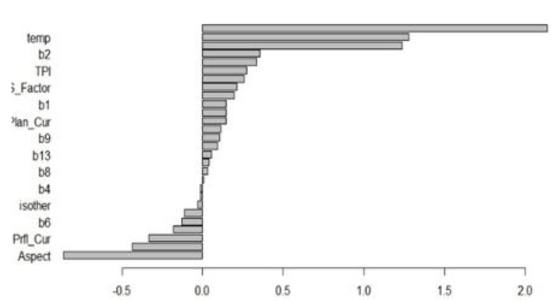

Figure 39. Prediction of Sub surface – clay

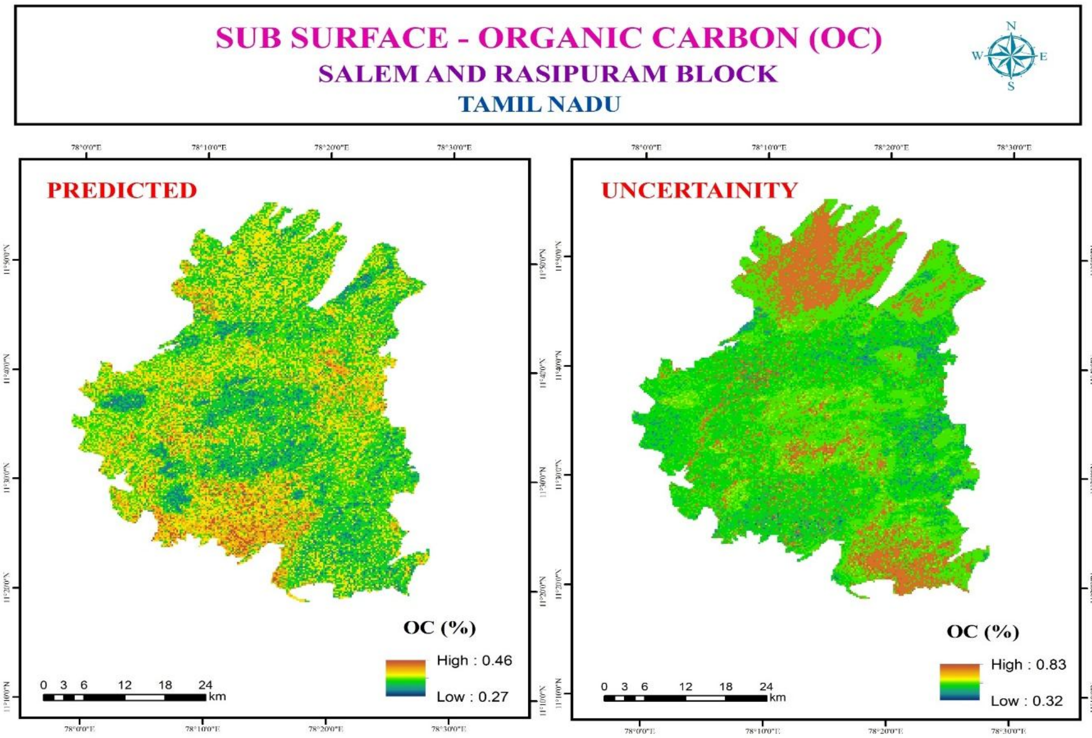

Figure 40. Prediction of Sub surface – Organic carbon

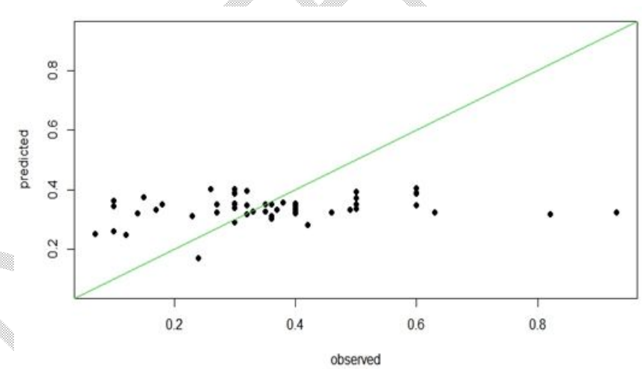

Figure 41. Scatterplot of Sub surface – Organic carbon

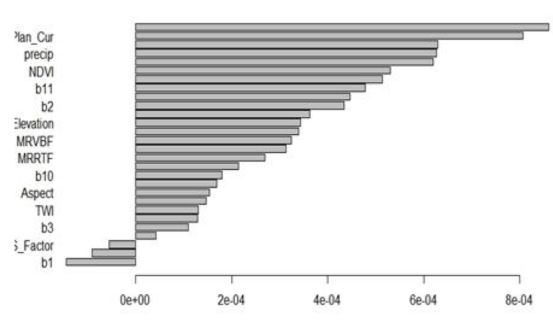

Figure 42. Variable importance of Sub surface – Organic carbon

At the surface level, uncertainty is low in the prediction of silt composition when compared with the prediction of sand and silt. The pH value at the surface is around 8 with a very low uncertainty range of 1.5 to 2.7. At sub-surface level, once again the slit prediction uncertainty is low, proving that the samples and the parameters of the model aid sufficiently in silt prediction.

4.5. Performance measure of QRF

The performance of the Random Forest model was evaluated for each property by calculating uncertainty indicators. Ten folds cross-validation techniques with 20 times repetition were used to evaluate the performance of the QRF model. The performance of models was evaluated using classic indicators such as Coefficient of determination (R2), Root. Mean Square Error (RMSE), mean error (ME), and Prediction interval coverage percentage (PICP). The accuracy assessment report is provided in below table 4.

Table 4. Model validation for prediction of Soil properties

MODEL ACCURACY ASSESMENT									
Parameters	**Layer**	**Mean error**		**RMSE**		R^2		**PICP**	
			SD		**SD**		**SD**		**SD**
Sand	Surface	1.22	0.34	12.61	0.50	-0.04	0.08	84.30	2.30
	subsurface	1.01	0.49	11.81	0.31	-0.51	0.09	86.30	1.95
Silt	Surface	-0.12	0.08	3.48	0.09	0.07	0.05	89.30	2.12
	subsurface	0.26	0.12	4.14	0.10	-0.12	0.06	85.00	1.58
Clay	Surface	-0.50	0.34	9.96	0.31	-0.22	0.08	87.0	1.48
	subsurface	-1.14	0.14	7.04	0.14	-0.09	0.04	92.70	1.63
pH	Surface	-0.19	0.01	0.63	0.01	0.02	0.03	89.70	2.38
OC	Surface	0.02	0.01	0.22	0.01	0.09	0.04	87.90	1.80
	subsurface	0.02	0.00	0.17	0.00	0.03	0.04	90.50	2.10

The performance of the model is also depending on sampling density. The sampling density for this study is very low. Whereas higher sample density is required for better results in tropical countries where soil pattern is complex due to the geological uplift than in other regions. Overall, complex regional soil patterns with changing land use are the causes for weak prediction of these dynamic soil properties. The prediction accuracy of the RFM model was low in surface soil compared to subsurface.

Further assessment was done by computing statistics on the predicted properties of soil at the surface and sub-surface levels (Tables 5 and 6). These values are compared with original values and found that there is much deviation in values except for the mean.

Table 5. Summary statistics on the predicted surface properties of the soil

Statistics	**PH**	**OC (%)**	**sand**	**clay**	**Silt**
Min	8.0	0.305	45	18	4.2
Max	8.5	0.7	76.9	45	12
Mean	8.37	0.495	63.323	29.15	7.99
S.D	0.079	0.074	5.784	4.924	1.301
Skewness	-1.884	0.451	-0.143	-0.297	-0.322

Table 6. Summary statistics on the predicted sub-surface properties of the soil

Statistics	**OC (%)**	**sand**	**clay**	**silt**
Min	0.27	45	30	4.2
Max	0.46	67	39	15
Mean	0.340	56.76	34.272	9.86
S.D	0.019	4.80	2.761	1.520
Skewness	0.542	0.345	-0.497	0.968

5. CONCLUSION

Soil is the habitat of plants and hence it is a must to conserve it. In order to manage soil resources, an inventory of soil properties has to be made. Traditional techniques are quite laborious and time-consuming. With the advancements in remote sensing technology, we could map soil properties digitally. This study has adopted SCORPAN model to predict five major soil properties for surfaces such as sand, silt, clay, pH, OC, and four major soil properties for subsurfaces such as sand, silt, clay, and OC. The model is quite flexible with the input of covariates and soil samples. For this study, different covariates including vegetation, terrain, legacy soil information, and derivatives of DEM and satellite imagery were used for prediction. Results show that with very low sampling density and freely available limited resources, we achieved moderately accurate results. The model works well at a global level but the study applies the model at the local level and hence the lack in achieving high accuracy. This can be overcome with an increase in the soil samples, and environmental covariates with high-resolution satellite imagery and DEM data.

ACKNOWLEDGEMENTS:

The authors are grateful to The ICAR – NBSS&LUP, Bangalore and thankful to Dr.S.Dharumarajan Ph D (IARI), Post Doc (Sydney Uni), Senior Scientist (Soil Science) for his guidance and support.

COMPETING INTERST:

Authors has declared that no competing interest exist.

REFERENCES

1. Gray, J.M., Bishop, T.F.A., Wilson, B.R., (2015). Factors controlling soil organic carbon stocks with depth in Eastern Australia. Soil Sci. Soc. Am. J. 79 (6), 1741–1751.
2. Grunwald, S. (2009). Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma 152 (3), 195–207.
3. Gessler, P.E., I.D. Moore, N.J. McKenzie, and P.J. Ryan. (1995). Soil-landscape modeling and spatial prediction of soil attributes. International Journal of Geographical Information Systems. 9, 421 – 432.
4. Greve, M.H., Kheir, R.B., Greve, M.B., and Bøcher, P.K. (2012). Using digital elevation models as an environmental predictor for soil clay contents. Soil Sci. Soc. Am. J. 76: 2116–2127. doi:10.2136/sssaj2010.0354.
5. Grimm, R., Behrens, T., Märker, M., Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island — digital soil mapping using Random Forests analysis. Geoderma 146, 102–113.

6. Hartemink, A. E., & McBratney, A. (2008). A soil science renaissance. Geoderma, 148(2), 123–129. https://doi.org/10.1016/J.GEODERMA.2008.10.006
7. Heuvelink, G., Brus, D., de Gruijter, J. (2007). Optimisation of sample configurations for digital soil mapping with universal kriging. In: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), Digital Soil Mapping: An Initial Perspective. Developments in Soil Science 31, Elsevier Amsterdam.
8. Hobley, E., Wilson, B., Wilkie, A., Gray, J., Koen, T. (2015). Drivers of soil organic carbon storage and vertical distribution in Eastern Australia. Plant Soil 390 (1–2), 111–127.
9. Legacherie Ph, Baret F, Feret J, Netto JM, Robbez-Masson JM. (2008). Estimation of soil clay and calcium carbonate using laboratory, field, and airborne hyperspectral measurements. Remote Sensing of Environment 112, 825-835. http://dx.doi.org/10.1016/j.rse.2007.06.014
10. McBratney, A.B., M.L. Mendonça-Santos, and B. Minasny. (2003). On digital soil mapping. Geoderma 117:3-52.
11. Meersmans, J., van Wesemael, B., De Ridder, F., Van Molle, M. (2009). Modeling the three-dimensional spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 152 (1), 43–52.
12. Mueller, T.G., Pierce, F.J. (2003). Soil carbon maps: enhancing spatial estimates with simple terrain attributes at multiple scales. Soil Sci. Soc. Am. J. 67 (1), 258–267.
13. Minasny, Budiman & Mcbratney, Alex & McKenzie, Neil & Grundy, Mike. (2008). Predicting soil properties using pedotransfer functions and environmental correlation. Australian Soil and Land Survey Handbook Series. 349-367.
14. Mehnatkesh, Abdolmohammad & Ayoubi, Shamsollah & Jalalian, Ahmad & Sahrawat, Kanwar. (2013). Relationships between soil depth and terrain attributes in a semi arid hilly region in western Iran. Journal of Mountain Science. 10. 163-172. 10.1007/s11629-013-2427-9.
15. Meinshausen, Nicolai. (2006). Quantile Regression Forests. Journal of Machine Learning Research. 7. 983-999.
16. McKenzie, N.J., Ryan, P.J. (1999). Spatial prediction of soil properties using environmental correlation. Geoderma 89, 67–94.
17. S. Dharumarajan, Rajendra Hegde, S.K. Singh (2017). Spatial prediction of major soil properties using Random Forest Techniques- A case study in semi-arid tropics

of South India. Geoderma Regional 10, 154-162. https://dx.doi.org/10.1016/j.geodrs.2017.07.005

18. S. Dharumarajan, Rajendra Hegde, N. Janani, S.K. Singh, (2019). The need for digital soil mapping in India. Geoderma Regional 15, e00204. https://doi.org./j.geodrrs.2019.e00204
19. S. Dharumarajan, Philippe Lagacherie, Rajendra Hegde, S.K. Singh, (2018). Digital Soil Mapping Concepts and Applications- Manual. NBSS&LUP (ICAR), Bengaluru.
20. USDA, (2017). Soil Survey Manual. Handbook NO.18. Issue: March 2017 (ISBN: PS57-D36-39BA)
21. Vågen Tor-G, Leigh A. Winowiecki, Jerome E. Tondoh, Lulseged T. Desta, Thomas Gumbricht (2016). Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. Geoderma, Vol.263, pp: 216-225.
22. Viscara Rossel, R.A., Behrens, T. (2010). Using data mining to model and interpret soil 650 diffuse reflectance spectra. Geoderma 158 (1–2), 46–54.