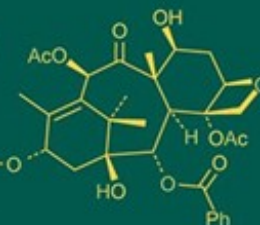
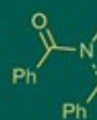
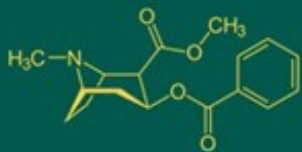


International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
 ISSN Online: 2617-4707
 NAAS Rating (2026): 5.29
 Indexing: Embase
 IJABR 2026; SP-10(10): 571-576
www.biochemjournal.com
 Received: 06-08-2026
 Accepted: 24-09-2026
 Published: 07-10-2026

The names and affiliations of all authors are provided below, following the references.

Digital mapping of DTPA-extractable zinc in agricultural soils of Tamil Nadu using Random Forest and environmental covariates

B Bhanukiran Reddy¹, S Maragatham², R Santhi³, D. Balachandar⁴, D Vijaya Lakshmi⁵, V Davamani⁶, D Vasu⁷, M Gopalakrishnan⁸

DOI: <https://www.doi.org/10.33545/26174693.2026.v10.i10Sg.10343>

ABSTRACT

Spatial information on plant-available zinc (Zn) is important for identifying areas with contrasting micronutrient status and for improving location-specific nutrient management. This study assessed the spatial variability of DTPA-extractable Zn in agricultural soils of Tamil Nadu, India, and predicted its distribution using Random Forest (RF) modelling and environmental covariates. A total of 250 georeferenced surface soil samples were collected from 0-15 cm depth using conditioned Latin Hypercube Sampling to represent the major agro-ecological and soil conditions of the state. Micronutrient cations were extracted with DTPA at pH 7.3 and Zn was quantified using atomic absorption spectrophotometry. Spectral and terrain covariates derived from Landsat 8 and Shuttle Radar Topography Mission data were screened using the *Boruta* feature-selection algorithm before RF modelling. DTPA-extractable Zn ranged from 0.01 to 4.35 mg kg⁻¹, with a mean of 0.97 mg kg⁻¹. *Boruta* identified slope, normalized difference vegetation index and topographic wetness index as the principal predictors. The optimized RF model, with mtry = 3 and ntree = 1000, achieved an *R*² of 0.69, RMSE of 0.44 mg kg⁻¹ and MAE of 0.32 mg kg⁻¹. Predicted Zn concentrations were comparatively higher in parts of southern and western Tamil Nadu and lower in portions of northern and central Tamil Nadu. The results show that georeferenced soil observations combined with terrain and remotely sensed covariates can delineate regional-scale variation in DTPA-extractable Zn and provide a spatial basis for targeted soil testing and zinc management.

Keywords: *Boruta*, DTPA, Digital soil mapping, Micronutrient, NDVI, Random Forest, Tamil Nadu, Zinc

1. Introduction

Zinc (Zn) is an essential micronutrient required for enzyme activation, protein synthesis, growth regulation and several physiological processes in crop plants. Although required in small quantities, inadequate Zn availability can restrict crop growth and nutritional quality. The concentration of plant-available Zn in soil is controlled by the amount of Zn present in reactive soil pools and by soil properties that regulate sorption, precipitation, complexation and movement within the profile.

Soil reaction is an important control on Zn availability. Increasing pH generally reduces soluble and exchangeable Zn through stronger adsorption and the formation of less soluble Zn compounds. Organic matter can modify Zn behaviour through complexation and by influencing the retention and release of micronutrients. These processes are particularly relevant in regions containing both acidic and alkaline or calcareous soils.

Tamil Nadu contains highly heterogeneous agricultural soils developed under contrasting parent materials, rainfall regimes, topographic settings and cropping systems. Major soil orders represented in the state include Alfisols, Vertisols, Inceptisols, Entisols and Ultisols. Their differences in texture, organic carbon, soil reaction, calcium carbonate and landscape position can produce substantial regional variation in DTPA-extractable Zn. Point-based soil analyses alone cannot provide a continuous representation of this variability across unsampled areas.

Digital soil mapping provides a framework for extending field observations across landscapes by linking measured soil properties with spatially exhaustive environmental covariates. Satellite-derived vegetation and spectral indices can capture differences in land-surface conditions, while terrain attributes represent landscape processes affecting runoff, erosion, deposition and moisture redistribution. Machine-learning methods are useful for integrating these covariates because soil-environment relationships are commonly nonlinear and interactive.

Corresponding Author:
 B Bhanukiran Reddy
 Tamil Nadu Agricultural
 University, Coimbatore,
 Tamil Nadu, India

Random Forest (RF) is an ensemble-learning algorithm that combines multiple decision trees and can accommodate complex relationships without requiring strict assumptions about the statistical distribution of predictors or response variables (Breiman, 2001) [1]. RF has been widely applied in digital soil mapping because of its ability to integrate numerous environmental covariates and represent nonlinear interactions (Belgiu *et al.*, 2016; Bhunia *et al.*, 2018) [2, 3]. *Boruta* feature selection can further reduce model complexity by retaining predictors that contribute useful information relative to randomized shadow variables.

Previous work from the broader soil survey used RF to assess soil pH and soil organic matter (Reddy *et al.*, 2024) [4]. The present study focuses specifically on DTPA-extractable Zn as an independent micronutrient endpoint. The objectives were to (i) characterize the variability of DTPA-extractable Zn in agricultural soils of Tamil Nadu, (ii) identify the principal environmental covariates associated with its regional spatial distribution, and (iii) generate a continuous prediction of DTPA-extractable Zn using RF modelling.

2. Materials and Methods

2.1 Study area

The study covered agricultural soils distributed across Tamil Nadu, southern India. The region includes coastal plains, river valleys, semi-arid interior landscapes and hilly terrain associated with the Western Ghats. These physiographic settings are accompanied by substantial variation in rainfall, land use, parent material and agricultural production systems. The soil survey represented the major soil orders occurring in the study area, including Alfisols, Inceptisols, Vertisols, Entisols and Ultisols.

2.2 Sampling design and soil collection

Sampling locations were selected using conditioned Latin Hypercube Sampling (cLHS) to represent the environmental variability of the study region. Variables considered during sampling-point selection included land use and land cover, agro-ecological information, soil sub-order, digital elevation information, terrain parameters such as slope and aspect, road-network data and normalized difference vegetation index. This procedure resulted in 250 sampling locations distributed across the study area.

Surface soil samples were collected from 0-15 cm depth. At each location, five subsamples were collected within approximately a 1-m radius and combined to obtain one composite sample. Each sampling point was georeferenced using GPS. Samples were transported to the laboratory, air-dried and passed through a 2-mm sieve before chemical analysis.

2.3 Determination of DTPA-extractable zinc

Available micronutrient cations were extracted using DTPA solution buffered at pH 7.3 following Lindsay *et al.* (1978) [5]. The soil-extractant suspension was shaken for approximately 2 h, and Zn in the extract was measured using atomic absorption spectrophotometry with an appropriate hollow cathode lamp. DTPA-extractable Zn was expressed as mg kg⁻¹ soil.

2.4 Environmental covariates

Environmental covariates were derived from satellite imagery and terrain data. Spectral information was obtained from Landsat 8, while terrain attributes were derived from Shuttle Radar Topography Mission (SRTM) data. The broader covariate set included spectral bands and vegetation-, moisture- and soil-related indices together with terrain variables.

Variables represented in the original covariate dataset included blue, green, red and near-infrared reflectance; normalized difference vegetation index (NDVI); soil-adjusted vegetation index; transformed NDVI; ratio vegetation index; difference vegetation index; slope; aspect; topographic wetness index (TWI); valley-bottom flatness; and ridge-top flatness.

2.5 Feature selection using Boruta

The *Boruta* algorithm was used to identify environmental variables that contributed useful information for predicting DTPA-extractable Zn. *Boruta* generates randomized shadow variables from the original predictors and compares their RF-derived importance with that of the actual environmental covariates. Variables consistently outperforming the shadow predictors are retained for modelling, while less informative variables are rejected.

For DTPA-extractable Zn, *Boruta* identified slope, NDVI and TWI as the most important covariates, and these variables were retained for final RF modelling.

2.6 Random Forest modelling and performance assessment

Random Forest regression was used to predict DTPA-extractable Zn from the selected environmental covariates. Model configurations were evaluated by varying the number of predictor variables available at each tree split (mtry) and the number of trees (ntree). The optimum model used mtry = 3 and ntree = 1000.

Model performance was summarized using the coefficient of determination (R²), root mean squared error (RMSE) and mean absolute error (MAE), as reported in the original RF analysis. The fitted RF model was then applied to the environmental-covariate layers to generate a continuous spatial prediction of DTPA-extractable Zn across the study region.

$$RMSE = \sqrt{[(1/n) \sum (y_i - \hat{y}_i)^2]}$$

$$MAE = (1/n) \sum |y_i - \hat{y}_i|$$

$$R^2 = 1 - [\sum (y_i - \hat{y}_i)^2 / \sum (y_i - \bar{y})^2]$$

where y_i is the observed value, $\hat{y}_i$ is the predicted value, $\bar{y}$ is the mean of the observed values and n is the number of observations.

3. Results

3.1 Variability of DTPA-extractable zinc

DTPA-extractable Zn varied markedly among the 250 sampled soils. Values ranged from 0.01 to 4.35 mg kg⁻¹, with a mean of 0.97 mg kg⁻¹ and a median of 0.73 mg kg⁻¹. The mode was 0.51 mg kg⁻¹, the standard deviation was 0.80 mg kg⁻¹ and the standard error was 0.06 mg kg⁻¹. The distribution was positively skewed (1.64) with a kurtosis of 2.90, indicating that most observations occurred toward the lower portion of the measured range while a smaller number of sites had relatively high Zn concentrations (Table 1).

3.2 Relationships with selected soil properties

DTPA-extractable Zn showed a negative relationship with soil pH ($r = -0.33$) and a positive relationship with organic carbon ($r = 0.44$). These associations indicate that lower Zn availability tended to occur with increasing soil reaction, whereas soils with greater organic carbon generally supported higher DTPA-extractable Zn. The correlations were moderate rather than deterministic, confirming that regional Zn variability reflected multiple interacting soil and landscape controls.

3.3 Selection of environmental predictors

Boruta feature selection identified slope, NDVI and TWI as the most important environmental variables for predicting DTPA-extractable Zn. Slope represents terrain-related redistribution processes, whereas NDVI and TWI provide information related to vegetation cover and landscape moisture conditions. These three variables were retained for the final RF model.

3.4 Random Forest model performance

The optimized RF model achieved an R² of 0.69, with an RMSE of 0.44 mg kg⁻¹ and an MAE of 0.32 mg kg⁻¹. The best-performing configuration consisted of mtry = 3 and ntree = 1000 (Table 1). Increasing mtry beyond 3 did not produce a meaningful improvement in model performance.

The observed-versus-predicted relationship showed reasonable agreement across the measured range. Greater uncertainty occurred among

soils with lower Zn concentrations, where the model tended to slightly overpredict some observations (Fig. 2c).

3.5 Spatial distribution of predicted DTPA-extractable zinc

The RF prediction revealed clear regional variation in DTPA-extractable Zn. Comparatively higher predicted concentrations occurred in low-lying areas and in parts of southern and western Tamil Nadu, while lower values were evident across portions of the northern and central regions (Fig. 2d). The prediction surface therefore delineated broad zones of contrasting Zn availability across the agricultural landscape.

4. Discussion

4.1 Regional variability in DTPA-extractable zinc

The range of 0.01–4.35 mg kg⁻¹ demonstrates substantial spatial heterogeneity in DTPA-extractable Zn across the sampled soils. This variability is consistent with the diversity of soil orders and environmental settings represented in Tamil Nadu. Differences in parent material, texture, soil reaction, organic matter, carbonate content and management can alter both the amount of Zn retained in the soil and the fraction extracted by DTPA.

The positively skewed distribution indicates that many sampling locations contained relatively low to moderate concentrations, whereas a smaller number of soils had substantially higher extractable Zn. Such heterogeneity limits the usefulness of generalized assumptions regarding micronutrient status over large geographic areas and supports spatially explicit soil testing.

4.2 Influence of soil reaction and organic carbon

The negative correlation between DTPA-extractable Zn and soil pH ($r = -0.33$) is consistent with the known decline in Zn solubility as soil reaction increases. Under alkaline and calcareous conditions, stronger adsorption and precipitation reactions can reduce the fraction of Zn remaining in readily extractable forms. DTPA was originally developed as a chelating soil test for Zn and related micronutrients and is widely used for evaluating micronutrient availability (Lindsay *et al.*, 1978) [5].

Organic carbon showed a positive correlation with DTPA-extractable Zn ($r = 0.44$). Organic matter can contribute to Zn retention and complexation and can influence the balance between soluble, adsorbed and organically associated forms. The moderate magnitude of both correlations indicates that neither pH nor organic carbon alone explained regional Zn variability; rather, their effects occurred together with terrain, moisture and other soil-forming factors.

4.3 Role of slope, vegetation and landscape moisture

Slope was retained as an important predictor of DTPA-extractable Zn. Terrain position affects erosion, deposition and redistribution of fine soil particles and organic matter, both of which can influence micronutrient storage. Lower landscape positions may accumulate sediment and associated nutrients, whereas erosional positions can lose nutrient-bearing surface material.

NDVI and TWI were also selected by *Boruta*. NDVI is not a direct measure of soil Zn, but it provides information on vegetation cover and land-surface condition that may co-vary with soil fertility and management. TWI represents the tendency of landscape positions to accumulate or shed water and can therefore capture differences in soil moisture and depositional environment. The combined importance of slope, NDVI and TWI suggests that regional Zn patterns were associated with interacting terrain, vegetation and moisture conditions rather than a single soil factor.

4.4 Predictive performance and uncertainty

The RF model explained 69% of the observed variation in DTPA-extractable Zn. Although this performance was lower than that obtained for some other soil properties in the broader survey, it indicates that

the selected environmental covariates contained substantial information relevant to regional Zn distribution. The RMSE of 0.44 mg kg⁻¹ and MAE of 0.32 mg kg⁻¹ should be interpreted relative to the observed concentration range of 0.01–4.35 mg kg⁻¹.

Prediction uncertainty was greater at the lower end of the Zn range, where the model tended to overestimate some low observations. Localized effects of fertilizer or micronutrient inputs, crop history, carbonate content and other soil chemical conditions may contribute to Zn availability but were not represented explicitly in the selected environmental covariates. The prediction should therefore be interpreted as a regional surface rather than as a substitute for field-specific micronutrient analysis.

4.5 Spatial pattern and implications for zinc management

The predicted distribution showed comparatively greater DTPA-extractable Zn in parts of southern and western Tamil Nadu, including low-lying areas, and lower predicted concentrations in portions of northern and central Tamil Nadu. The original thesis interpretation associated the higher values with greater organic matter and moisture retention and the lower values with comparatively drier, coarser-textured environments. These interpretations are consistent with the selected covariates but should be treated as explanatory mechanisms rather than direct causal tests.

From a nutrient-management perspective, the map can be used to prioritize locations for confirmatory soil testing. Areas predicted to have relatively low DTPA-extractable Zn can be targeted for more intensive sampling and crop-specific assessment before Zn fertilizer recommendations are made. Conversely, areas with greater predicted Zn can be evaluated to avoid unnecessary supplementation where soil tests indicate adequate availability.

The value of the spatial prediction is therefore in supporting more efficient allocation of soil-testing resources. Field-level nutrient recommendations should continue to rely on locally collected soil samples, crop requirement and established interpretation criteria.

4.6 Scope and limitations

This study was intentionally restricted to DTPA-extractable Zn and the environmental and modelling outputs available from the original soil survey. The RF model captured broad regional variability but did not explicitly incorporate historical Zn fertilization, crop-specific nutrient inputs or detailed management records. The available environmental layers also represent regional-scale information and may not capture fine-scale variability within individual fields.

Future work could combine higher-resolution remote sensing, management covariates and independent spatial validation. Nevertheless, the present 250-point soil survey and RF framework were sufficient to delineate meaningful regional patterns in DTPA-extractable Zn across Tamil Nadu.

5. Conclusions

DTPA-extractable Zn showed substantial spatial variability across agricultural soils of Tamil Nadu, ranging from 0.01 to 4.35 mg kg⁻¹ with a mean of 0.97 mg kg⁻¹. Zn was negatively associated with soil pH and positively associated with organic carbon, while *Boruta* feature selection identified slope, NDVI and TWI as the principal environmental predictors. The optimized Random Forest model, with $mtry = 3$ and $ntree = 1000$, achieved an R^2 of 0.69, RMSE of 0.44 mg kg⁻¹ and MAE of 0.32 mg kg⁻¹. The resulting prediction indicated comparatively higher DTPA-extractable Zn in parts of southern and western Tamil Nadu and lower values in portions of northern and central Tamil Nadu. These findings show that georeferenced soil observations integrated with terrain and remotely sensed covariates can provide a useful regional representation of Zn availability. The generated map can support targeted soil testing and spatially focused zinc management, while field-specific fertilizer decisions should continue to be based on confirmatory soil analysis.

Acknowledgements

The authors acknowledge the institutional facilities and technical support provided for soil sampling, laboratory analysis, remote-sensing data processing and spatial modelling.

Conflict of interest

The authors declare no conflict of interest.

Data availability

The data used in this study were generated during the soil survey, laboratory analysis and spatial modelling undertaken as part of the doctoral research programme.

Tables

Table 1. Descriptive statistics and Random Forest model performance for DTPA-extractable zinc in agricultural soils of Tamil Nadu.

Parameter	Value
Number of soil samples	250
Sampling depth	0-15 cm
Minimum DTPA-Zn	0.01 mg kg ⁻¹
Maximum DTPA-Zn	4.35 mg kg ⁻¹
Mean	0.97 mg kg ⁻¹
Median	0.73 mg kg ⁻¹
Mode	0.51 mg kg ⁻¹
Standard deviation	0.80 mg kg ⁻¹
Standard error	0.06 mg kg ⁻¹
Skewness	1.64
Kurtosis	2.90
Correlation with soil pH	$r = -0.33$
Correlation with organic carbon	$r = 0.44$
Selected RF predictors	Slope, NDVI, TWI
Optimal mtry	3
Optimal ntree	1000
R ²	0.69
RMSE	0.44 mg kg ⁻¹
MAE	0.32 mg kg ⁻¹

Note: DTPA-Zn values, correlation coefficients and model metrics are reported as in the thesis tables and Random Forest outputs. NDVI = normalized difference vegetation index; TWI = topographic wetness index; RMSE = root mean squared error; MAE = mean absolute error.

Figures

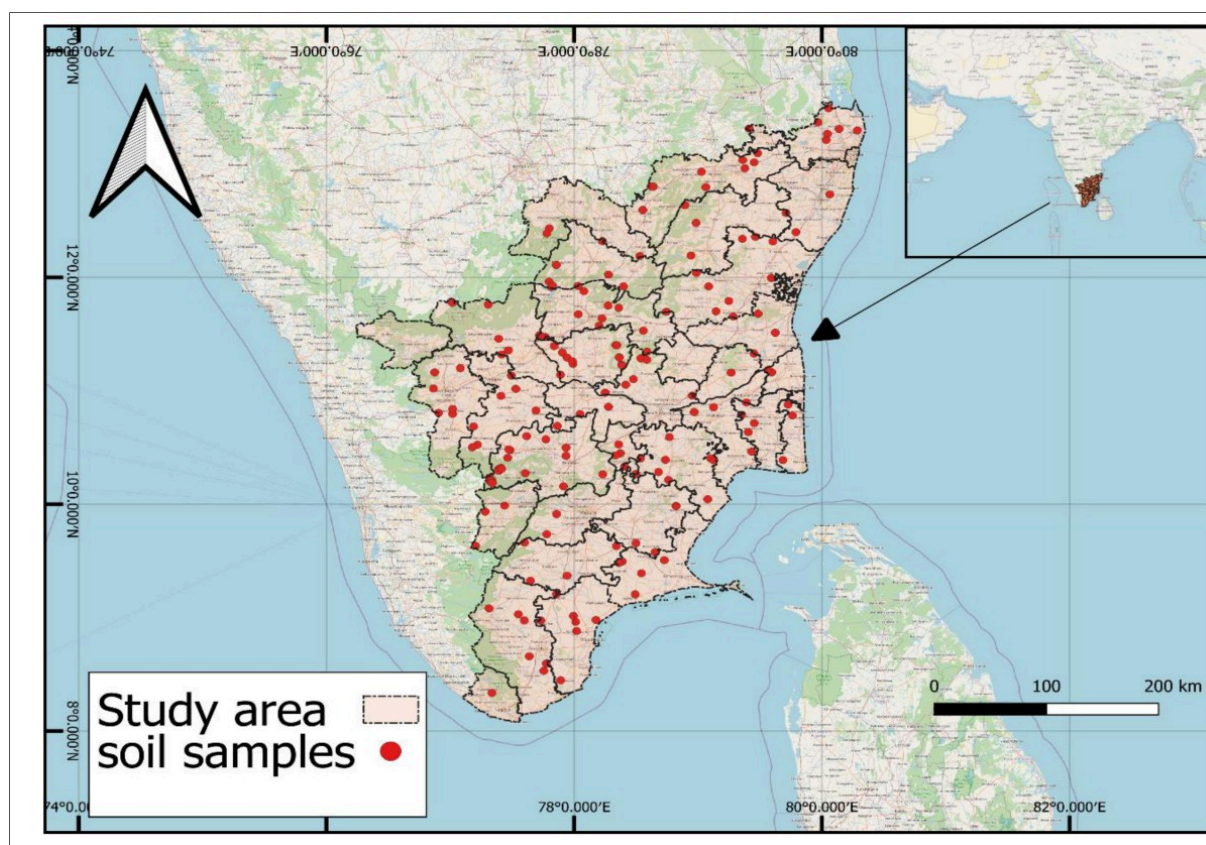


Figure 1. Study area and distribution of the georeferenced soil sampling locations across Tamil Nadu.

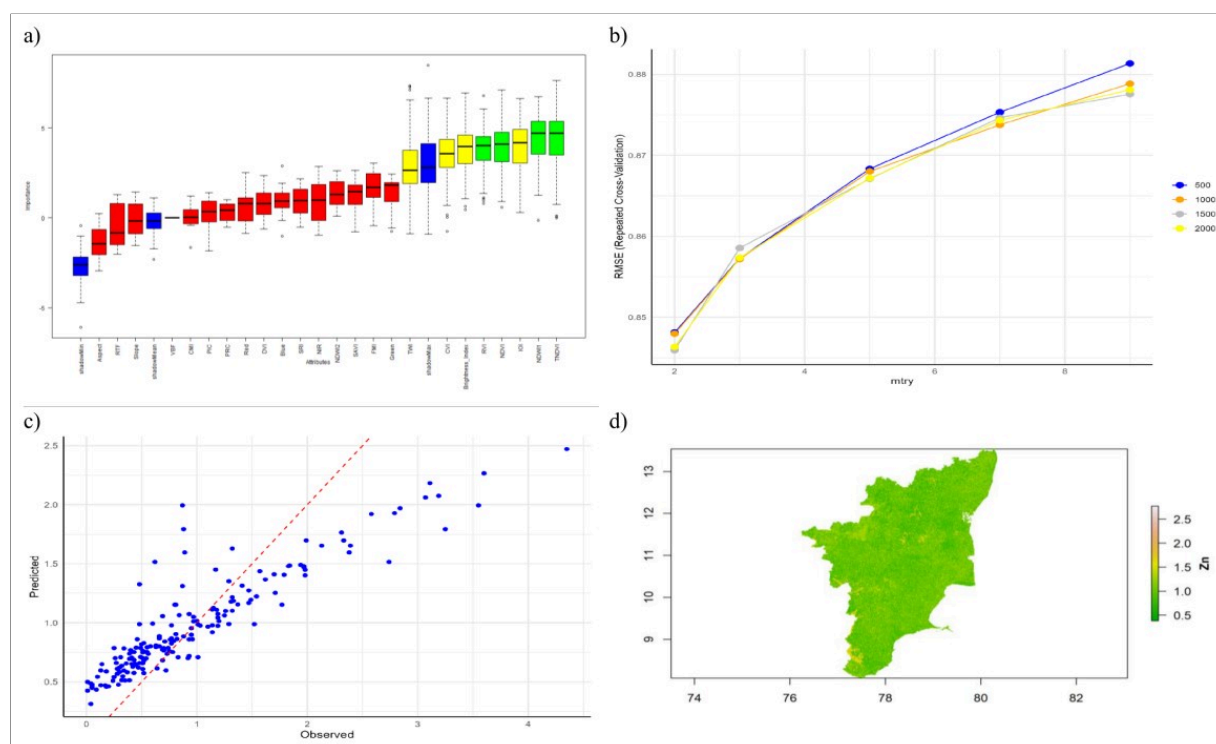


Figure 2. Random Forest analysis of DTPA-extractable zinc distribution: (a) *Boruta* feature selection, (b) Random Forest tuning, (c) observed versus predicted DTPA-extractable Zn (mg kg^{-1}), and (d) predicted spatial distribution of DTPA-extractable Zn (mg kg^{-1}) across Tamil Nadu.

References

1. Breiman L. Random forests. *Machine Learning*. 2001;45(1):5-32.
2. Belgiu M, Drăguț L. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2016;114:24-31.
3. Bhunia GS, Shit PK, Maiti R. Digital mapping of soil properties using Random Forests in a complex agro-ecological landscape. *Ecological Indicators*. 2018;93:493-503.

4. Reddy BB, Maragatham S, Santhi R, Balachandar D, Vijaya Lakshmi D, Davamani V, Vasu D, Gopalakrishnan M. Predictive soil mapping using random forest models: Applications in pH and soil organic matter assessment. *Plant Science Today*. 2024;11(4):463-474.
5. Lindsay WL, Norvell WA. Development of a DTPA soil test for zinc, iron, manganese, and copper. *Soil Science Society of America Journal*. 1978;42(3):421-428.
6. Grimm R, Behrens T, Märker M, Elsenbeer H. Soil organic carbon concentrations and stocks on Barro Colorado Island: Digital soil mapping using Random Forests. *Geoderma*. 2008;146(1-2):102-113.
7. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2(3):18-22.
8. Ma T, Brus DJ, Zhu AX, Zhang L, Scholten T. Comparison of conditioned Latin hypercube and feature space coverage sampling for predicting soil classes using simulation from soil maps. *Geoderma*. 2020;370:114366.
9. Maragatham S, Santhi R, Radhika K, Sivagnanam S, Rajeswari R, Hemalatha S, Kanimozhi A. An appraisal of available nutrients status and soil fertility mapping for Salem District of Tamil Nadu. *Madras Agricultural Journal*. 2014;101(1-3):51-58.
10. Taghizadeh-Mehrjardi R, Minasny B. Integrating remote sensing data with Random Forest to map soil properties in arid regions. *Geoderma*. 2021;406:115553.

How to cite this article: Reddy BB, Maragatham S, Santhi R, Balachandar D, Lakshmi DV, Davamani V, et al. Digital mapping of DTPA-extractable zinc in agricultural soils of Tamil Nadu using Random Forest and environmental covariates. *International Journal of Advanced Biochemistry Research*. 2026;SP-10(10):571-576. doi: <https://www.doi.org/10.33545/26174693.2026.v10.i10Sg.10343>

Authors

B Bhanukiran Reddy

Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

S Maragatham

Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

R Santhi

Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

D. Balachandar

Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

D Vijaya Lakshmi

Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

V Davamani

Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

D Vasu

Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

M Gopalakrishnan

Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India