land

MDPI

Article

Exploring the Effect of Sampling Density on Spatial Prediction with Spatial Interpolation of Multiple Soil Nutrients at a Regional Scale

Prava Kiran Dash [1,2,*], Bradley A. Miller [3], Niranjan Panigrahi [4] and Antaryami Mishra [1,5]

[1] Department of Soil Science, College of Agriculture, Odisha University of Agriculture and Technology, Bhubaneswar 751003, Odisha, India; antar199@yahoo.co.in
[2] Regional Research and Technology Transfer Station, Mahisapat, Odisha University of Agriculture and Technology 759013, Odisha, India
[3] Department of Agronomy, College of Agriculture and Life Sciences, Iowa State University, Ames, IA 50010, USA; millerba@iastate.edu
[4] Center for Water, Environment, and Development, Cranfield University, Cranfield MK43 0AL, UK; niranjan.panigrahi@cranfield.ac.uk
[5] Department of Soil Science and Agricultural Chemistry, Institute of Agricultural Sciences, Siksha 'O' Anusandhan University, Bhubaneswar 751029, Odisha, India
* Correspondence: pravakirandash@ouat.ac.in

Abstract: Essential soil nutrients are dynamic in nature and require timely management in farmers' fields. Accurate prediction of the spatial distribution of soil nutrients using a suitable sampling density is a prerequisite for improving the practical utility of spatial soil fertility maps. However, practical research is required to address the challenge of selecting an optimal sampling density that is both cost-effective and accurate for preparing digital soil nutrient maps across regional extents. This study examines the impact of sampling density on spatial prediction accuracy for a range of soil fertility parameters over a regional extent of 8303 km^2 located in eastern India. Surface soil samples were collected from 1024 sample points. The performance of six levels of sampling densities for spatial prediction of 14 soil properties was compared using ordinary kriging. From the sample points, randomization was used to select 224 points for validation and the remaining 800 for calibration. Goodness-of-fit for the semi-variograms was evaluated by R^2 of model fit. Lin's concordance correlation coefficient (CCC) and root mean square error (RMSE) were evaluated through independent validation as spatial prediction accuracy parameters. Results show that the impact of sampling density on prediction accuracy was unique for each soil property. As a common trend, R^2 of model fit and CCC scores improved, and RMSE values declined with the increasing sampling density for all soil properties. On the other hand, the rate of gain in the accuracy metrics with each increment in the sampling density gradually decreased and ultimately plateaued. This indicates that there exists a sampling density threshold beyond which the extra effort on additional sampling adds less to the spatial prediction accuracy. The findings of this study provide a valuable reference for optimizing soil nutrient mapping across regional extents.

Keywords: nutrient management; spatial interpolation; kriging; semi-variograms; sampling density; prediction accuracy; maps

check for updates

Citation: Dash, P.K.; Miller, B.A.; Panigrahi, N.; Mishra, A. Exploring the Effect of Sampling Density on Spatial Prediction with Spatial Interpolation of Multiple Soil Nutrients at a Regional Scale. *Land* **2024**, *13*, 1615. https://doi.org/10.3390/land13101615

Academic Editors: Zamir Libohova, Kabindra Adhikari, Michele Duarte De Menezes and Subramanian Dharumarajan

Received: 1 September 2024
Revised: 27 September 2024
Accepted: 29 September 2024
Published: 4 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In the present era of precision agriculture, site-specific nutrient assessment is essential for correcting nutrient deficiencies and optimizing crop yield by adopting variable-rate fertilizer applications [1]. However, adopting site-specific soil management is much more challenging in small land-holding settings, where farmers regularly lack the resources to test their farm soils. For instance, Indian agriculture is dominated by small and marginal land holdings, with 86% of individual land sizes being less than 2 ha and

an average of 1.08 ha [2]. Moreover, NPK fertilizer consumption in India has reached 30.6 million Metric Tons during 2023-24 [3]. Correspondingly, the subsidy bill on fertilizers in India has reached 2.7 billion USD during 2023-24 [4]. In these scenarios, various stakeholders, including farmers, researchers, and policymakers, ask for soil nutrient information in digital formats to make informed decisions about nutrient application and subsidy management [5].

Soil is heterogeneous at all scales due to dynamic interactions among natural and human-induced factors [6]. Measuring soil properties everywhere and all the time is impractical. Over the past decades, digital soil mapping (DSM) techniques have been extensively used to prepare spatial soil maps based on limited soil observations [7]. Among different mapping approaches, kriging [8] is a long-established geostatistical method for spatial interpolation, which produces unbiased estimates at unsampled locations with minimum error variances. Spatial interpolation adheres to the fundamental principle of geography, i.e., spatial autocorrelation encompassing 'stationarity' and 'linearity' [9,10]. It considers that all things are related, but near things are more related than farther apart [11]. Among different forms of kriging, ordinary kriging is regularly employed for soil mapping, especially when maps need to be prepared without any auxiliary information or covariates [12,13].

Because kriging works on the principle of spatial autocorrelation, sample locations must fully cover the map area. In addition, predictions perform best overall when the distances among sample points are minimized, providing sufficient support for the semivariogram [14,15]. Ordinary kriging interpolation is especially sensitive to the number of sample points or sampling density [16,17]. A suitable sampling density is a critical prerequisite for accurate prediction of the spatial distribution of soil properties [18,19]. In general, more samples would be expected to capture the variation of soil properties more accurately [20]. However, operationally, the quantity of samples cannot be increased infinitely as scientists and policymakers must work with limited budgets [21–23]. Therefore, practical research is required to address the challenge of selecting an optimal sampling density that balances both economic efficiency and map accuracy [24–26].

Several studies have reported their observed threshold sample sizes at various mapping extents. While predicting soil organic carbon (SOC) at a small spatial extent of 0.24 km^2, Vašát et al. observed that accuracy increased up to an optimal sample size of 37 samples (156,118 samples per 1000 km^2) [27]. With a mapping area nearly twice the size (0.44 km^2), Wang et al. concluded that 40 samples (90,909 samples per 1000 km^2) were optimal [28]. Although somewhat different from SOC concentration, Sherpa et al. found 160 samples (6897 samples per 1000 km^2) to be optimal for mapping SOC stock for a 23 km^2 mapping area [29]. Zhang et al. while predicting SOC stock in a geographical extent of 72 km^2, observed that the prediction accuracy initially decreased and then increased with the number of samples up to 357 (4958 samples per 1000 km^2) [30]. Similarly, Yu et al. observed the accuracy of SOC prediction in a 927 km^2 area improving with each increment of sample size to the maximum of 525 (566 samples per 1000 km^2) [31]. Continuing this trend, Ye et al. found 669 samples (41 samples per 1000 km^2) to be optimal for mapping SOC over a 16,400 km^2 area [32]. Finally, in a study with a much larger mapping extent (180,000 km^2), Sun et al. concluded that a sample size of 800 samples (4.4 samples per 1000 km^2) was optimal for mapping SOC [33]. Trends suggest that optimal sampling densities are situation-specific, and the ranges vary significantly across the mapping extents. However, most studies so far on the impact of sampling density on spatial prediction accuracy are around smaller mapping extents ranging from a few ha to less than 100 km^2. And, we have less certainty that these patterns continue to larger extents. Therefore, there is a need to unfold these results for regional extents to facilitate large-scale soil management for small land holdings.

The sampling requirement for one soil property might not be the same for another [27,34–36]. Information on the impact of sampling density on spatial interpolation accuracy has been reported only for a few soil properties, viz., SOC [27,28,31–33,37];

SOC stock [29,30], soil organic matter (SOM) [26,35,38–40], soil depth [41], clay [38], texture [16], granularity [35], pH [27,35], cation exchange capacity [26], available P, K [38,42], and available micronutrients [43]. This indicates that the literature so far is limited to only a few target soil properties, majorly SOC and SOM. Information on the impact of sampling density on spatial prediction accuracy for soil fertility parameters, especially soil nutrients, is very limited. These warrants experimenting on multiple soil properties, especially soil fertility parameters, which need to be managed through the application of external inputs.

In this context, this study was conducted to examine the effect of varying sampling densities on (a) goodness-of-fit for semi-variograms and (b) spatial prediction accuracy of ordinary kriging interpolation at a regional scale. A range of relatively static to dynamic soil properties were considered as target soil properties to identify an optimal sampling density that would suffice to generate reasonably accurate and economically efficient spatial soil maps across the parameters. The evaluation was performed in two stages. First, the goodness-of-fit for the resulting semi-variograms was measured quantitatively i.e., R^2 of semi-variogram model fit. Then, the prediction accuracies in terms of Lin's concordance correlation coefficient (CCC) and root mean square error (RMSE) of the resulting maps were evaluated using a consistent independent validation data set. Briefly, this study is unique in terms of (1) its larger mapping extent, (2) consideration of multiple target soil properties, especially soil nutrients, (3) reporting the results in terms of sampling density instead of sample size, (4) exploring quantitative relationships between goodness-of-fit for semi-variograms and prediction accuracies of the resulting maps, and (5) performing a brief literature review of the previous works to identify a general trend of optimal sampling density with mapping extent.

2. Materials and Methods

A methodology was developed using evidence synthesis to support the development of a conceptual framework that would assess the relationship between sampling density and prediction accuracy of spatial soil maps for a suite of 14 soil properties at a regional extent. A brief description of the study area and conceptual framework are outlined below.

2.1. Study Area

The study area is Keonjhar district of Odisha state, located in eastern India, extending from 21.01° N to 22.16° N latitude and 85.18° E to 86.37° E longitude with a geographical extent of 8303 km^2 (Figure 1). The region experiences a hot and moist sub-humid tropical to sub-tropical climate with hot summers and cool winters. May is the warmest month of the year, with mean minimum and maximum temperatures of 25 °C and 46 °C, respectively. December is the coolest month of the year, with mean minimum and maximum temperatures of 12 °C and 25 °C, respectively. The mean annual rainfall is 1488 mm, more than half of which is received in June, July, and August during the movement of the SouthWest monsoon [44]. Common soil types of this region include red, yellow, alluvial, and lateritic with sandy loam to silty clay loam soil textures [45]. The parent material of the study area is predominately granite-gneiss. Soils belong to *Alfisols*, *Inceptisols*, and *Entisols* soil orders as per the USDA Soil Taxonomy [46].

2.2. Conceptual Framework

A conceptual framework was developed using literature review and field visits. A diagrammatic representation of the conceptual framework is presented in Figure 2. The components of this framework include soil sampling, laboratory analysis, sampling strategy for modeling, geostatistical modeling, soil mapping, and evaluation of spatial prediction accuracy. Individual components of this workflow are described below.

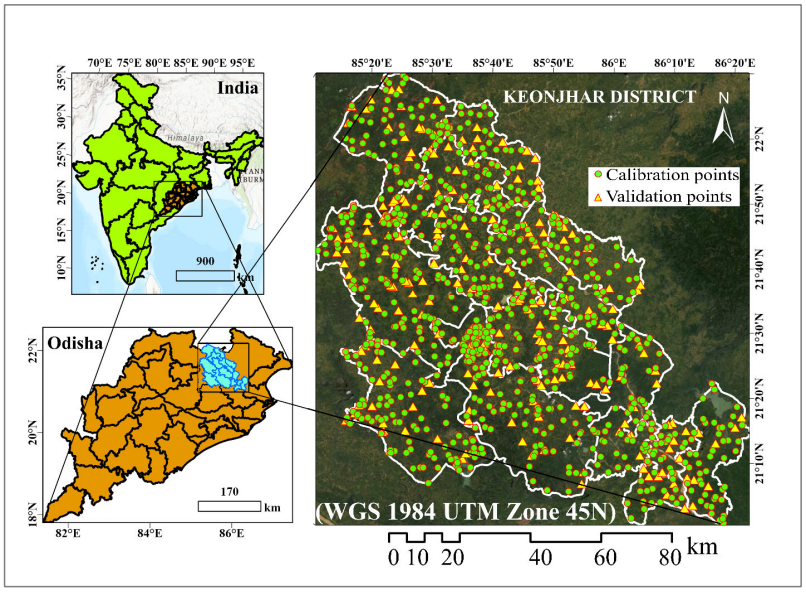


Figure 1. The study area with sample points. White boundaries of the study area represent the administrative block boundaries of the district.

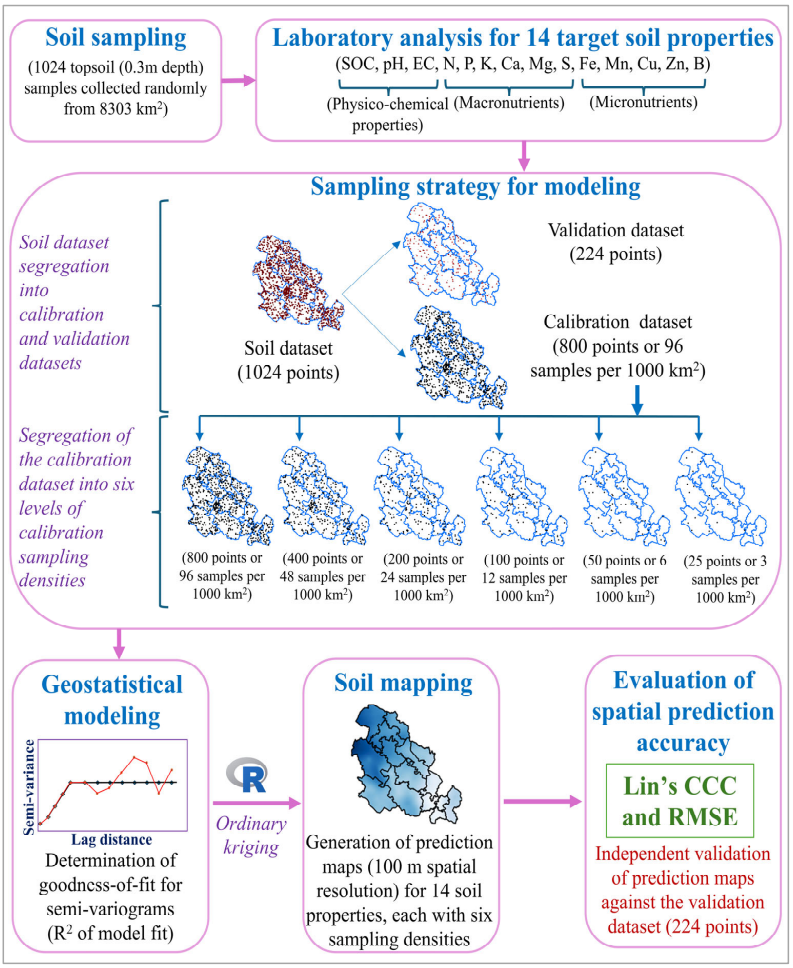


Figure 2. Conceptual framework of the methodology followed in this study. Corresponding abbreviations are SOC: soil organic carbon, EC: electrical conductivity, N: nitrogen, P: phosphorus, K: potassium, Ca: calcium, Mg: magnesium, S: sulfur, Fe: iron, Mn: manganese, Cu: copper, Zn: zinc, B: boron, R^2: coefficient of determination, CCC: Lin's concordance correlation coefficient, RMSE: root mean square error.

2.2.1. Soil Sampling

Topsoil samples (0.30 m depth) were collected randomly from 1024 sample points during the year 2019–2020 (Figure 1). Inaccessible areas were avoided for sampling. Local differentiations in soil colour and texture were given importance while choosing the sample points to capture the local variabilities. Approximately 1 kg of soil was collected from each point. The geographical coordinates (latitude and longitude) of the sample points were recorded using a GPS instrument (Garmin 76MAPCSx). The collected soil samples, after processing, were stored within properly levelled polyvinyl bottles for further laboratory analysis.

2.2.2. Laboratory Analysis

Soil samples were analyzed for 14 soil properties viz., SOC, pH, electrical conductivity (EC), available nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), sulfur (S), iron (Fe), manganese (Mn), copper (Cu), zinc (Zn), and boron (B). SOC was determined by the wet digestion method using the Ferroin indicator, and the carbon oxidized was multiplied by the efficiency factor of the experiment, i.e., 77% [47]. Soil pH at 1:2.5 w/v was determined using a glass electrode digital pH meter, and EC was estimated at 1:2.5 w/v by an electronic conductivity meter [48]. Available N was determined by the alkaline potassium permanganate method using an automatic nitrogen auto analyzer [49]. Available P was estimated by Bray's No. 1 method using a spectrophotometer at 660 nm [50]. Available K was determined using the neutral normal ammonium acetate extraction method using a flame photometer [51]. Exchangeable Ca and Mg were determined by extraction with neutral normal ammonium acetate followed by titration against 0.01N EDTA (Ethylenediamine tetraacetic acid) using Eriochrome Black T and Calcon indicators [52]. Available S was determined using 0.15% $CaCl_2.2H_2O$ extractant and $BaCl_2$ reagent followed by measuring the turbidity at 410 nm using a spectrophotometer [53]. Available Fe, Mn, Cu, and Zn were determined by the DTPA (diethylenetriaminepentaacetic acid) extraction method using an atomic absorption spectrophotometer [54]. Available B was estimated by the hot water extraction method using the Azomethine-H indicator [55].

Individual soil properties were grouped into three categories of soil properties, i.e., (i) physico-chemical soil properties including SOC, pH, EC, (ii) macronutrients including N, P, K, Ca, Mg, S, and (iii) micronutrients including Fe, Mn, Cu, Zn, and B. Descriptive statistics for all the soil properties at different sampling densities is presented in Table 1. The mean, standard deviation (SD), coefficient of variation (CoV), and skewness values of the soil properties did not differ remarkably across the sample subsets. Overall, CoV for different soil properties of the whole dataset varied between 16% (pH) and 59% (B). The CoVs for the physico-chemical soil properties, macronutrients, and micronutrients varied in the range of 16–44%, 31–49%, and 39–59%, respectively.

2.2.3. Sampling Strategy for Modeling

A soil dataset of 14 target soil properties for the 1024 sample points was generated by laboratory analysis. At first, 224 points were randomly selected out of the 1024 sample points as the independent validation dataset. The remaining 800 sample points were considered as the calibration dataset. The selection of samples was random using the =RAND () function in Microsoft Excel Version 2408, which returned evenly distributed random real numbers from the dataset. Using this function, new random real numbers were returned every time the worksheet was calculated. During each random selection of sample points, the function was repeated for 10 iterations to avoid the possible instability caused by a single random sample and to ensure good coverage of sample points for the entire study area. A similar procedure was also followed by Lai et al. [40]. The distribution of calibration and validation sample points has been presented in Figure 1.

Secondly, the calibration dataset was segregated into six levels of calibration sampling densities (Figure 2). The calibration sample sizes have been expressed in terms of sampling densities since sampling density also represents the mapping extent. The original calibra-

tion dataset of 800 samples (per 8303 km^2) corresponds to a sampling density of 96 samples per 1000 km^2. This original sampling density was subjected to several consecutive random selections to generate six levels of calibration sampling densities. The original sampling density of 96 samples per 1000 km^2 (800 points) was considered as the first calibration dataset. For generating the second calibration subset, only 50% of the samples of the first calibration dataset were selected randomly to generate a sampling density of 48 samples per 1000 km^2 (400 points) using the =RAND () function in Microsoft Excel (Version 2408) with 10 iterations. This process of randomly selecting 50% of the sample points was carried out subsequently four more times to generate a total of six overlapping sampling densities viz., 96, 48, 24, 12, 6, and 3 samples per 1000 km^2 with corresponding sample sizes of 800, 400, 200, 100, 50, and 25, respectively.

Table 1. Descriptive statistics of the soil properties at different sampling densities.

Soil Property	Sampling Density (Number of Samples/1000 km²)	Minimum	Maximum	Mean	SD	CoV (%)	Skewness
SOC (%)	3	0.30	1.38	0.76	0.28	37.40	0.17
	6	0.10	1.38	0.72	0.29	40.82	0.26
	12	0.10	1.38	0.72	0.29	40.30	0.02
	24	0.09	1.38	0.72	0.29	39.83	0.03
	48	0.09	1.38	0.72	0.29	39.67	−0.03
	96	0.09	1.38	0.73	0.29	39.25	−0.09
	Validation dataset (N = 224)	0.10	1.32	0.75	0.29	37.92	−0.08
pH	3	4.15	6.90	5.58	0.86	15.48	0.15
	6	4.15	7.35	5.65	0.88	15.63	0.09
	12	4.15	7.35	5.70	0.88	15.47	0.12
	24	4.15	7.37	5.67	0.89	15.62	0.15
	48	4.15	7.37	5.63	0.89	15.76	0.23
	96	4.15	7.38	5.63	0.88	15.58	0.26
	Validation dataset (N = 224)	4.18	7.35	5.57	0.88	15.77	0.33
EC (dS/m)	3	0.10	0.50	0.30	0.11	35.80	−0.16
	6	0.03	0.50	0.27	0.12	43.86	−0.20
	12	0.03	0.50	0.27	0.11	40.71	−0.26
	24	0.03	0.50	0.26	0.10	39.85	−0.18
	48	0.03	0.50	0.27	0.11	39.82	−0.18
	96	0.03	0.50	0.27	0.11	39.56	−0.11
	Validation dataset (N = 224)	0.04	0.50	0.27	0.10	38.34	0.03
N (kg/ha)	3	95.28	531.42	295.49	93.12	31.52	0.28
	6	53.72	531.42	272.83	100.38	36.79	0.19
	12	50.99	531.42	278.69	92.97	33.36	−0.05
	24	50.99	531.42	275.62	90.93	32.99	0.07
	48	50.99	531.42	277.85	90.22	32.47	−0.08
	96	50.99	533.29	278.92	88.29	31.65	0.09
	Validation dataset (N = 224)	77.00	501.19	277.19	82.41	29.73	0.35
P (kg/ha)	3	8.53	44.48	22.64	9.38	41.41	0.31
	6	5.87	44.48	21.60	9.69	44.85	0.42
	12	5.87	44.48	21.99	8.62	39.20	0.22
	24	5.76	44.48	22.12	8.73	39.48	0.19
	48	5.76	44.48	21.93	8.87	40.44	0.20
	96	5.68	44.48	22.11	8.81	39.86	0.17
	Validation dataset (N = 224)	5.87	41.15	23.44	9.16	39.08	−0.04

Table 1. *Cont.*

Soil Property	Sampling Density (Number of Samples/1000 km²)	Minimum	Maximum	Mean	SD	CoV (%)	Skewness
K (kg/ha)	3	241.87	657.67	418.32	95.75	22.89	0.40
	6	173.22	724.68	428.84	124.42	29.01	0.51
	12	92.00	724.68	425.92	127.61	29.96	0.08
	24	76.90	728.24	425.91	130.02	30.53	0.01
	48	76.90	728.24	421.35	132.05	31.34	0.11
	96	76.90	728.24	419.37	130.17	31.04	0.14
	Validation dataset (N = 224)	98.24	722.22	411.48	125.06	30.39	−0.01
Ca (cmol(+)/kg)	3	0.49	5.24	2.90	1.48	50.95	0.05
	6	0.49	6.02	3.12	1.51	48.33	0.00
	12	0.49	6.09	3.23	1.41	43.72	−0.01
	24	0.49	6.39	3.27	1.39	42.67	−0.01
	48	0.49	6.60	3.24	1.40	43.38	0.12
	96	0.41	6.60	3.25	1.40	43.01	0.16
	Validation dataset (N = 224)	0.48	6.55	3.14	1.39	44.29	0.19
Mg (cmol(+)/kg)	3	0.13	3.23	1.71	0.91	53.07	0.05
	6	0.13	3.43	1.81	0.91	50.14	−0.04
	12	0.13	3.62	1.89	0.87	45.89	−0.01
	24	0.13	3.62	1.90	0.86	45.25	−0.07
	48	0.06	3.68	1.86	0.87	47.02	0.03
	96	0.06	3.68	1.88	0.87	46.13	0.03
	Validation dataset (N = 224)	0.07	3.68	1.82	0.87	47.72	0.07
S (mg/kg)	3	1.73	29.60	15.53	7.34	47.25	−0.05
	6	1.35	29.60	13.90	7.67	55.19	0.18
	12	1.35	29.60	13.96	6.96	49.86	0.12
	24	1.10	29.60	13.63	7.05	51.72	0.24
	48	1.06	29.60	13.93	6.90	49.57	0.07
	96	1.06	29.71	14.01	6.93	49.44	0.14
	Validation dataset (N = 224)	1.21	29.48	13.84	6.68	48.28	0.24
Fe (mg/kg)	3	35.55	180.60	81.59	35.09	43.01	0.85
	6	10.31	180.60	78.00	35.25	45.20	0.93
	12	10.31	180.60	76.09	31.60	41.53	0.56
	24	10.31	180.60	76.58	30.72	40.12	0.45
	48	10.05	180.60	76.44	32.95	43.10	0.37
	96	5.61	180.60	77.34	33.44	43.24	0.33
	Validation dataset (N = 224)	16.57	180.07	79.74	34.60	43.39	0.45
Mn (mg/kg)	3	17.49	85.99	44.92	17.45	38.85	0.42
	6	0.52	85.99	43.95	18.22	41.45	0.45
	12	0.52	85.99	43.46	17.69	40.71	0.07
	24	0.52	91.28	43.62	17.07	39.15	0.15
	48	0.52	91.28	42.65	16.75	39.28	0.17
	96	0.52	91.28	43.20	16.96	39.26	0.06
	Validation dataset (N = 224)	0.59	84.52	43.91	16.97	38.64	0.18
Cu (mg/kg)	3	1.29	5.51	3.21	1.16	36.10	0.09
	6	0.03	5.51	2.99	1.13	37.91	0.14
	12	0.03	5.86	2.90	1.23	42.35	0.29
	24	0.03	5.86	2.95	1.20	40.47	0.18
	48	0.03	5.86	3.01	1.19	39.42	0.05
	96	0.03	5.90	3.01	1.17	39.06	0.02
	Validation dataset (N = 224)	0.21	5.85	3.08	1.14	36.91	0.04

Table 1. *Cont.*

Soil Property	Sampling Density (Number of Samples/1000 km²)	Minimum	Maximum	Mean	SD	CoV (%)	Skewness
Zn (mg/kg)	3	0.30	2.32	1.00	0.52	51.97	0.57
	6	0.02	2.42	0.92	0.52	57.11	1.01
	12	0.02	2.42	0.89	0.52	58.65	1.04
	24	0.02	2.42	0.91	0.51	55.86	0.89
	48	0.02	2.49	0.93	0.50	54.45	0.74
	96	0.01	2.49	0.92	0.49	53.48	0.72
	Validation dataset (N = 224)	0.06	2.33	0.94	0.48	50.47	0.58
B (mg/kg)	3	0.25	1.63	0.78	0.40	51.44	0.57
	6	0.08	2.39	0.78	0.47	59.71	1.16
	12	0.07	2.43	0.77	0.46	60.08	1.13
	24	0.07	2.63	0.78	0.47	59.95	1.27
	48	0.06	2.63	0.83	0.51	61.11	1.15
	96	0.04	2.63	0.82	0.49	59.15	1.12
	Validation dataset (N = 224)	0.06	2.48	0.84	0.48	57.98	1.07

2.2.4. Geostatistical Modeling and Soil Mapping

Geostatistical analyses were performed in the 'gstat' package of R version 4.1.0 [56]. Considering the values of a soil property z at two different points viz., x and x + h, which are separated at a distance (lag) of h from each other, the semi-variance (γ) between the pair of points was calculated following Equation (1). While building the empirical variogram, the default lag provided by the R software (version 4.1.0) was used.

$$\gamma(h) = \frac{1}{2}\{z(x) - z(x+h)\}^2 \quad (1)$$

where γ (h) is the semi-variance at lag h, h is the vector denoting the distance and direction between the two points x and x + h; z (x) and z (x + h) are the observed values of the soil property z at locations x, x + h, respectively. Here, x denotes the coordinates of the sample points in two dimensions.

Similarly, considering the soil property z being measured at numerous places in the region, and there are m pairs of observed values (considering all neighbouring pairs of observations), the mean semi-variance at any particular lag (h) was calculated following Equation (2) [14,57].

$$\gamma(h) = \frac{1}{2m}\sum_{i=1}^{m}\{z(x) - z(x+h)\}^2 \quad (2)$$

The functions relating the γ and h were expressed as sample semi-variograms (also known as experimental semi-variograms). Different mathematical models viz., circular, spherical, exponential, gaussian, and bessel were tested on each sample semi-variogram, to fit continuous smooth curves, i.e., model semi-variograms (also known as theoretical semi-variograms) to the respective sample semi-variograms. The best-fitted mathematical models for each soil property and each sampling density were chosen based on whichever provided the lowest mean standard error. Semi-variogram parameters, namely nugget (C_0; uncorrelated variance), partial sill (C; spatially correlated variance), sill (C_0 + C), and range (r; limit of spatial correlation) were calculated. Since skewness values of all the soil properties were close to zero, the datasets were normally distributed, and thereby, data transformation was not performed [39,58]. Nevertheless, preserving the original values was also useful to avoid complications of data back-transformation while estimating error components at the validation stage [15].

Predictions for the unsampled locations were carried out using the ordinary kriging interpolation method (Equation (3)) [59,60]. Respective prediction maps were prepared

using the 'Krige ()' function of the 'gstat' package of R (version 4.1.0) at a 100 m spatial resolution with a projected coordinate system of 'WGS 1984 UTM Zone 45N'.

$$\hat{z}\,(X_0) = \sum_{i=1}^{n} \lambda_i z(X_i) \quad (3)$$

where, $\hat{z}(X_0)$ is the predicted value of soil property 'z' at the unsampled location X_0, z (X_i) is the observed value of the same soil property 'z' at a neighbouring sampled location around the unsampled location, n is the number of neighbouring points, and λ_i are the respective weights assigned to each measured value at the sampled locations.

2.2.5. Evaluation of Spatial Prediction Accuracy

Goodness-of-fit for the semi-variogram models was measured using the coefficient of determination (R^2; Equation (4)). Subsequently, the prepared prediction maps were compared against the 224 independent validation points. Prediction accuracy of the maps were evaluated by determining CCC (Equation (5)) and RMSE (Equation (6)).

$$R^2 = 1 - \frac{SS_{error}}{SS_{total}} \quad (4)$$

$$CCC = \frac{2r\sigma_{obs}\sigma_{pred}}{\left(\overline{Z}_{obs} - \overline{Z}_{pred}\right)^2 + \sigma^2_{obs} + \sigma^2_{pred}} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[z(S_i) - \hat{z}(S_i)]^2} \quad (6)$$

where SS_{error} is the sum of the squares of the errors, SS_{total} is the total sum of the squares, r is the Pearson's correlation coefficient, $\overline{Z}$ is the mean of the variable, σ^2 is the variance, $z(S_i)$ is the observed value of variable z at the location S_i, $\hat{z}(S_i)$ is the predicted value of variable z at the location (S_i), and n is the number of sample points.

3. Results

In Section 3.1., goodness-of-fits for semi-variograms in terms of R^2 of model fit are presented. In Section 3.2., the prediction accuracies of the spatial soil maps in terms of independently validated CCC and RMSE are presented. In each section, trends are presented in three subsections viz., (a) general trends across all soil properties, (b) variations across categories of soil properties, and (c) variations across individual soil properties. Trends for RMSE values are presented across individual soil properties only since the units of RMSEs are specific to respective soil properties. Variations in the visual representation of the prediction maps are presented in Section 3.3.

3.1. Goodness-of-Fit for Semi-Variograms (R^2 of Model Fit)

3.1.1. General Trends across All Soil Properties

The sample semi-variograms obtained with smaller sampling densities viz., 3, 6, or 12 samples per 1000 km^2, were observed to be noisier and more erratic compared to those obtained with higher sampling densities, viz., 24, 48, and 96 samples per 1000 km^2 (Figure 3). Correspondingly, the goodness-of-fit for the model semi-variograms to the sample semi-variograms was also poorer for the smaller sampling densities than those obtained with the higher sampling densities. Semi-variograms and semi-variogram parameters, including the type of semi-variogram model, nugget, sill, and range for all the soil properties at different sampling densities, have been presented in the Supplementary Materials.

Irrespective of the soil property, the R^2 of the semi-variogram model fit initially increased with the increasing sampling density and then reached a plateau (Figure 4). The range for R^2 of model fit across soil properties gradually became narrower with the increasing sampling density, indicating that as the sampling density increased, the reliability of goodness-of-fit across all soil properties improved. In other words, the probability of

obtaining a higher R^2 of model fit, regardless of soil property, increased with increased sampling densities. The rate of gain in the mean R^2 of model fit decreased gradually from 20% for increasing the sampling density from 3 to 6 samples per 1000 km^2 to 0% for increasing the sampling density from 48 to 96 samples per 1000 km^2.

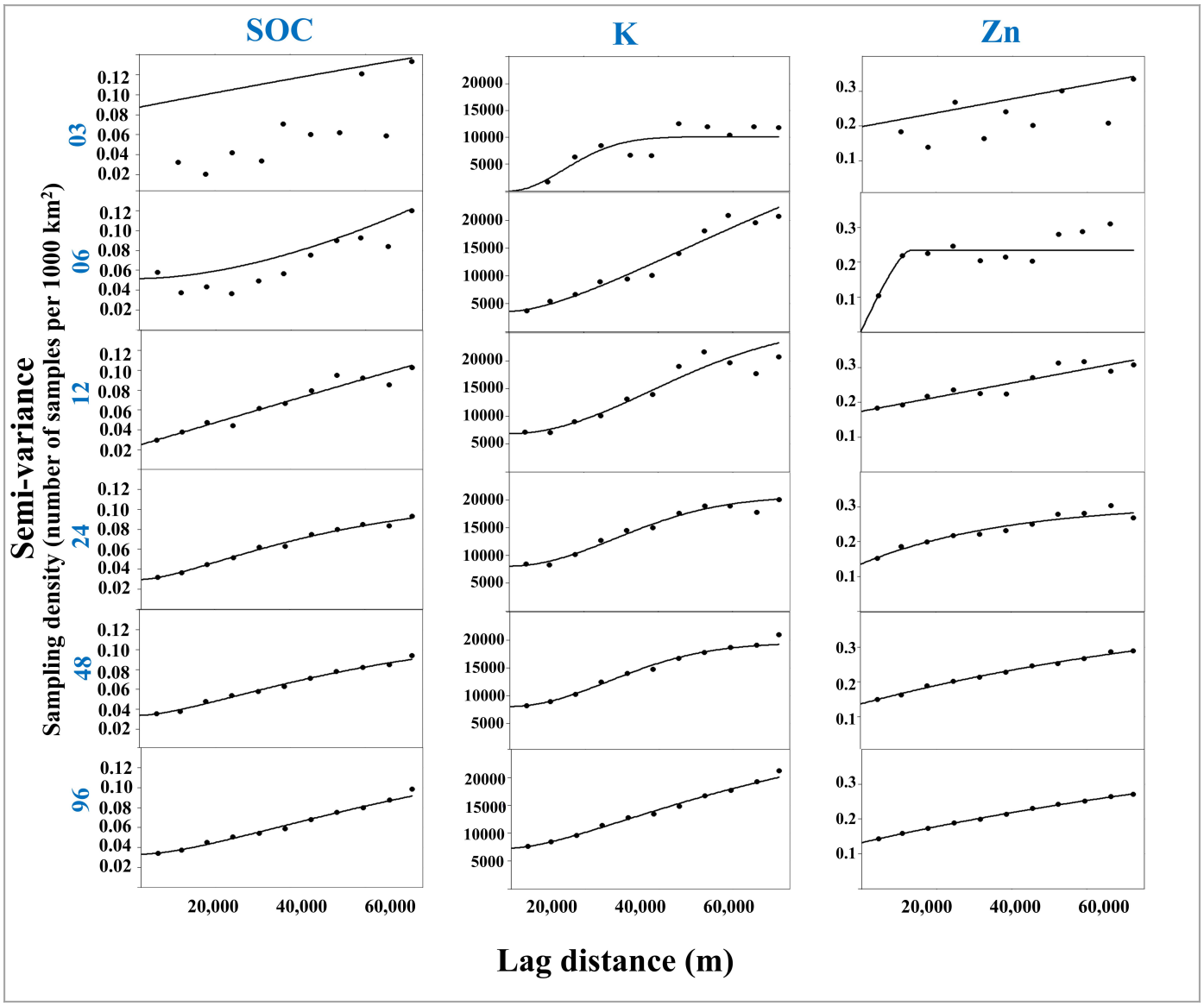


Figure 3. Semi-variograms of selected soil properties at different sampling densities. The dots represent the sample semi-variograms, and the solid lines represent the model semi-variograms. Increasing sampling density corresponds to a reduction in the noisiness of the sample semi-variograms and a coinciding improvement in the goodness-of-fit for modeled semi-variograms for each soil property.

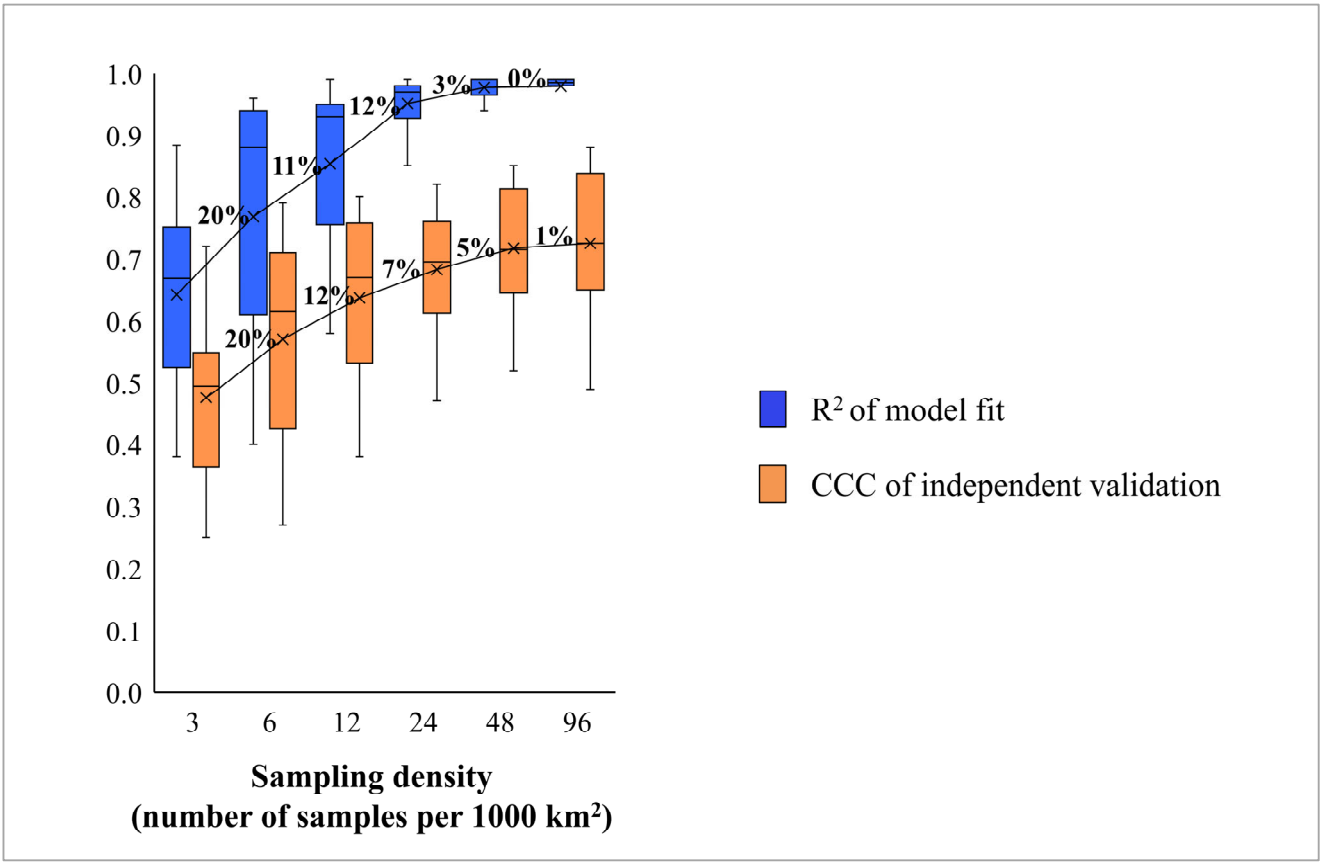


Figure 4. General trends for R^2 of model fit and CCC for independent validation at different sampling densities. The box plots represent the ranges for R^2 of model fit and independent validation CCC values for each of the 14 soil properties. The '×' marks represent the mean values; the horizontal bars inside the box plots represent the median values; the curves connect the mean values across different sampling densities. The values expressed in % represent the rate of gain in quality metrics at each step of sequential increases in the sampling density.

3.1.2. Variations across Categories of Soil Properties

All three categories of soil properties, viz., physico-chemical soil properties, macronutrients, and micronutrients, followed similar trends as the general trend for R^2 of model fit (Figure 5). For instance, for all three categories of soil properties, the minimum, maximum, and mean values of R^2 of model fit initially showed increasing trends, which later increased relatively slowly and finally remained constant with the increasing sampling density. For all three categories of soil properties, sampling densities of 48 and 96 samples per 1000 km^2 were able to generate smooth semi-variograms with R^2 of model fits varying between 0.92 and 0.99. However, at any sampling density, the mean R^2 values for model fit per category consistently decreased from physico-chemical soil properties (R^2 = 0.92) > macronutrients (R^2 = 0.90) > micronutrients (R^2 = 0.79).

The rate of gain in the mean R^2 of model fit values also followed similar trends as the general trend. Irrespective of the soil property category, the rate of gain in the mean R^2 of model fit reduced gradually and ultimately became 0% with the increasing sampling density. Nevertheless, the goodness-of-fit for micronutrients category responded more to the rise in sampling density compared to other categories of soil properties.

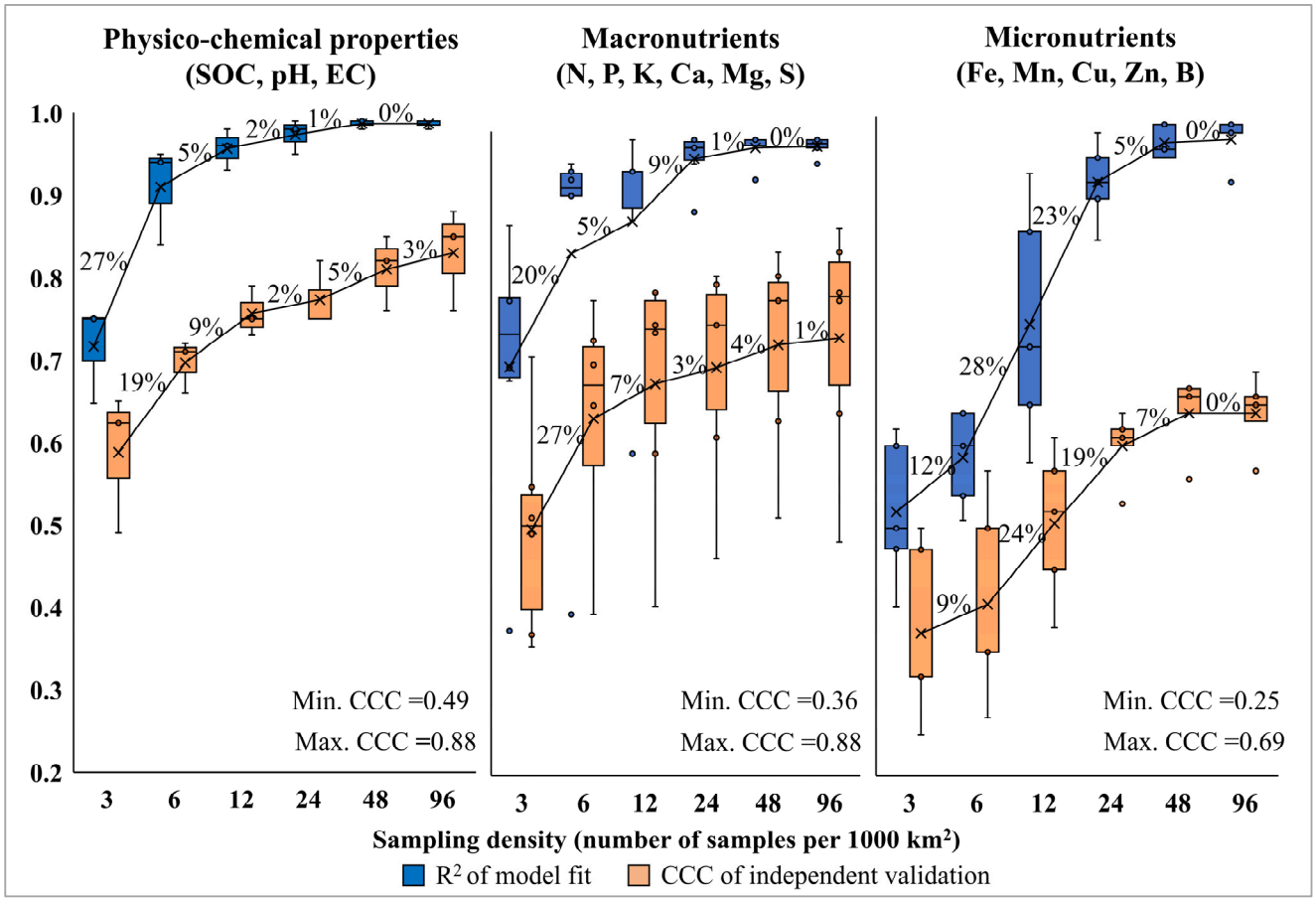


Figure 5. Trends for R^2 of model fit and CCC for independent validation at different sampling densities for different categories of soil properties. The box plots represent the ranges for R^2 of model fit and independent validation CCC values for the categories of soil properties. The '×' marks represent the mean values; the horizontal bars inside the box plots represent the median values; the curves connect the mean values across different sampling densities. The values expressed in % represent the rate of gain in quality metrics at each step of sequential increases in the sampling density.

3.1.3. Variations across Individual Soil Properties

All soil properties showed increasing trends for R^2 of model fit with the increasing sampling density (Figure 6). For each soil property, the lowest R^2 of model fit was observed at the sampling density of 3 samples per 1000 km^2. However, the highest R^2 of model fit was observed at varying sampling densities for each soil property. For instance, SOC, pH, K, Ca, Mg, Mn, Cu, and Zn reached their highest R^2 of model fits at the sampling density of 48 samples per 1000 km^2, whereas for P, Fe, and B, the highest R^2 of model fits were obtained at the sampling density of 96 samples per 1000 km^2. For a few soil properties viz., EC and N, the sampling density of 24 samples per 1000 km^2 was enough for generating the highest R^2 of model fit that could be obtained with even higher sampling densities. For the soil property S, even the sampling density of 12 samples per 1000 km^2 was enough to obtain the highest R^2 of model fit, which stayed constant thereafter, even with further increments in the sampling densities. Nevertheless, each soil property except P could achieve a R^2 of at least 0.95 or greater, usually at the higher side of sampling densities.

As an exception, the R^2 of the model fits decreased in a few instances of increasing sampling density. For example, the R^2 of model fit for S and Mn decreased by 1–3.2% while increasing the sampling densities from 48 to 96 samples per 1000 km^2. Also, for K, the R^2 of the model fit decreased by 6.3% while increasing the sampling density from 6 to 12 samples per 1000 km^2. In terms of patterns for R^2 of model fit with increasing sampling density, SOC, pH, EC, N, Ca, Mg, and S had similar trends with the maximum rates of gain in R^2 of

model fits while increasing the sampling density from 3 to 6 samples per 1000 km^2, which thereafter showed an almost constant level with further increments in the sampling density. Soil property K also possessed a similar trend as above, except that there was also a decline in the R^2 of model fit while increasing the sampling density from 6 to 12 per 1000 km^2. Soil properties Cu and Zn possessed trends that showed exceptional improvements in the R^2 of the model fit while increasing the sampling densities from 6 to 12 samples per 1000 km^2. The trends of R^2 of model fits were distinguishably different for P, Fe, Mn, and B, which had their highest rates of gains in R^2 of model fit while increasing the sampling density from 12 to 24 samples per 1000 km^2.

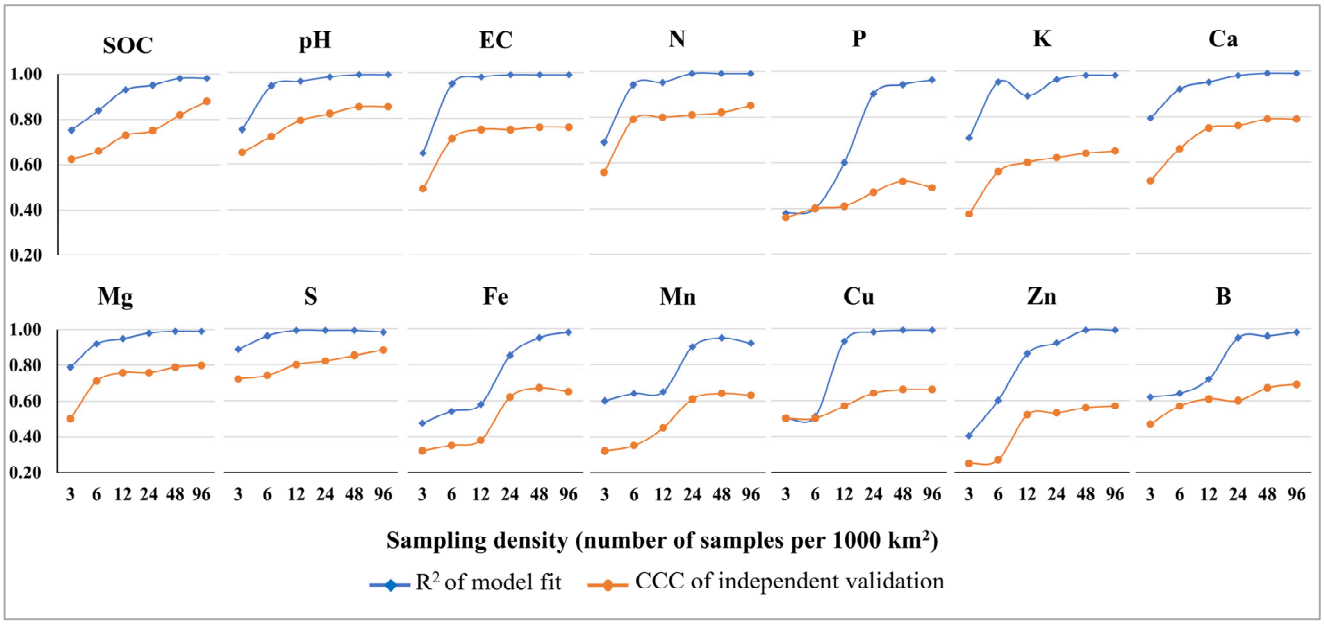


Figure 6. Trends for R^2 for model fit and CCC for independent validation at different sampling densities for individual soil properties. For each soil property, both R^2 of model fit and CCC of independent validation increased with increasing sampling density. Each soil property possessed its lowest R^2 of model fit and CCC of independent validation at the sampling density of 3 samples per 1000 km^2. The highest R^2 of model fit was observed at sampling densities varying between 12 and 96 samples per 1000 km^2. The highest CCCs were obtained at either 48 or 96 samples per 1000 km^2. However, for many of the soil properties, these performance metrics plateaued after 12 to 24 samples per 1000 km^2.

3.2. Prediction Accuracy (CCC and RMSE)

3.2.1. General Trends across All Soil Properties

The mean prediction accuracy, as measured by CCC, increased with increasing sampling density. The minimum and maximum CCCs obtained across soil properties increased with increasing sampling density. The rate of accuracy improvement with increasing sample density was initially higher but gradually slowed down and finally remained almost constant (Figure 4). The range of prediction accuracy across soil properties was widest at the sampling density of 6 samples per 1000 km^2. It was narrowest at the sampling density of 48 samples per 1000 km^2. Considering the range of CCC as an indicator of the reliability of prediction across soil properties, the sampling densities of 6 and 48 samples per 1000 km^2 were observed to be the most unreliable and most reliable, respectively.

The rate of gain in the mean CCC while increasing the sampling density from 3 to 6 samples per 1000 km^2 was observed to be 20%. However, this rate of gain decreased for each increment in the sampling density, and finally, the rate of gain in the mean CCC was only 1%, while the sampling density was increased from 48 to 96 samples per 1000 km^2. Thus, even though the mean CCC gradually increased with increasing sampling density, there was a simultaneous decline in the rate of gain in the mean CCC with each sequential

increase in the sampling density. Moreover, the R^2 of model fits, and CCCs were positively correlated (r = 0.80).

3.2.2. Variations across Categories of Soil Properties

Similar to the general trend, all three categories of soil properties showed improvements in the independent validation CCC with increasing sampling density (Figure 5). The mean CCCs at different sampling densities for different categories of soil properties indicated that mean CCCs initially increased rapidly, then slowly, and finally remained almost constant with increasing sampling density. In other words, the rate of improvement in the mean CCCs ultimately plateaued with the increasing sampling density for all three categories of soil properties.

The level of prediction accuracy decreased in the order of physico-chemical properties > macronutrients > micronutrients with mean CCCs of 0.74, 0.67, and 0.53 for physico-chemical soil properties, macronutrients, and micronutrients, respectively. Standard deviation values of CCCs for physico-chemical soil properties, macronutrients, and micronutrients were 0.09, 0.15, and 0.13, respectively, indicating that CCCs for the macronutrients and micronutrients were more variable across their constituent soil properties compared to that for the physico-chemical soil properties. However, the variability of independent validation results across micronutrients was much higher at lower sampling densities than at higher sampling densities.

For all three categories of soil properties, CCCs were observed to be positively correlated with the R^2 of model fits for semi-variograms. The respective correlation coefficients for the relationship between semi-variogram model fit and prediction performance for physico-chemical soil properties, macronutrients, and micronutrients were 0.89, 0.77, and 0.84, respectively. Although prediction accuracies were successively lower for macro and micronutrients, the generalization of sampling density being connected to the semi-variogram quality and prediction accuracy—with diminishing returns—remained true also for these soil property categories.

3.2.3. Variations across Individual Soil Properties

The CCCs of each of the soil properties increased with increasing sampling density (Figure 6). Each soil property possessed its lowest CCC at the sampling density of 3 samples per 1000 km^2. For SOC, N, K, Mg, S, Zn, and B, the highest CCCs were obtained at the highest sampling density of 96 samples per 1000 km^2, whereas the remaining soil properties reached their highest CCCs with 48 samples per 1000 km^2.

The low rate of improvement for CCC while increasing the sampling density from 48 to 96 samples per 1000 km^2 is noteworthy because of the relatively high effort required to produce that increase in sample density. This increase in sample density for the present study represents efforts of sampling an additional 400 sample points along with the associated costs for laboratory analysis. For SOC, N, K, Mg, S, Zn, and B, the rates of gain in CCC varied in the range from 1.3% to 7.3% while increasing the sampling density from 48 to 96 samples per 1000 km^2. However, in the same situation, pH, EC, Ca, and Cu did not show any improvement in CCC. In contrast, for P, Fe, and Mn, the CCCs decreased in the range of 1.6–5.8% while increasing the sampling density from 48 to 96 samples per 1000 km^2.

In terms of patterns of prediction accuracy, soil properties viz., EC, N, K, Ca, and Mg possessed similar trends having the maximum rates of gains in CCCs while the sampling density was increased from 3 to 6 samples per 1000 km^2, which thereafter decreased with the further increments in the sampling density. Boron possessed a similar trend, except that the rate of gain in CCC became negative (−1.6%) while increasing the sampling density from 12 to 24 samples per 1000 km^2, which again increased and later decreased with further increments in the sampling densities. Soil properties SOC, pH, and S possessed considerably higher rates of gains in CCCs at the initial two stages of increments in the sampling density, i.e., while increasing the sampling density from 3 to 6 and from 6 to

12 samples per 1000 km^2. Zinc possessed the highest rate of gain in CCC while increasing the sampling density from 6 to 12 per 1000 km^2, which further decreased, increased, and ultimately decreased with the further increments in the sampling density. Micronutrients Fe, Mn, and Cu possessed similar patterns of CCCs having higher rates of gains in CCCs at the second and third increments in sampling density viz., while increasing the sampling density from 6 to 12 and 12 to 24 samples per 1000 km^2, which then decreased thereafter with further increments in the sampling density. The trend of P suggests that the rate of gain in CCC was considerably higher while increasing the sampling densities from 12 to 24 and 24 to 48 samples per 1000 km^2, which then decreased with further increments in the sampling density.

In terms of levels of prediction accuracy achieved, maximum CCCs greater than 0.80 were obtained for soil properties SOC, pH, N, Mg, and S, whereas Ca and EC had maximum CCCs of 0.79 and 0.76, respectively. On the other hand, the maximum CCCs obtained for P, K, Fe, Mn, Cu, Zn, and B were observed to be varying between 0.52 and 0.69, which are poorer CCCs than that for other soil properties. In the context of the soil property categories, these results indicate that EC was the more difficult property to predict in the category of physico-chemical properties, while P was the more difficult property to predict in the macronutrient category. The high prediction accuracy for N was intriguing, considering how dynamic this soil property tends to be, both spatially and temporally.

Independent validation RMSEs of individual soil properties have been presented in Figure 7. All soil properties showed declining trends of RMSEs with increasing sampling density. Soil properties N, K, Ca, Zn, and B had sharp declines in RMSEs between 3 and 6 samples per 1000 km^2. SOC, pH, Mg, S, and Cu, in contrast, had gradual declining trends of RMSEs at each increment of sampling density. For EC and P, the RMSEs initially decreased with the increments in sampling density up to 12 samples per 1000 km^2 and thereafter remained constant with further increases in the sampling density. Although still following the general trend, RMSEs for Fe and Mn were more erratic.

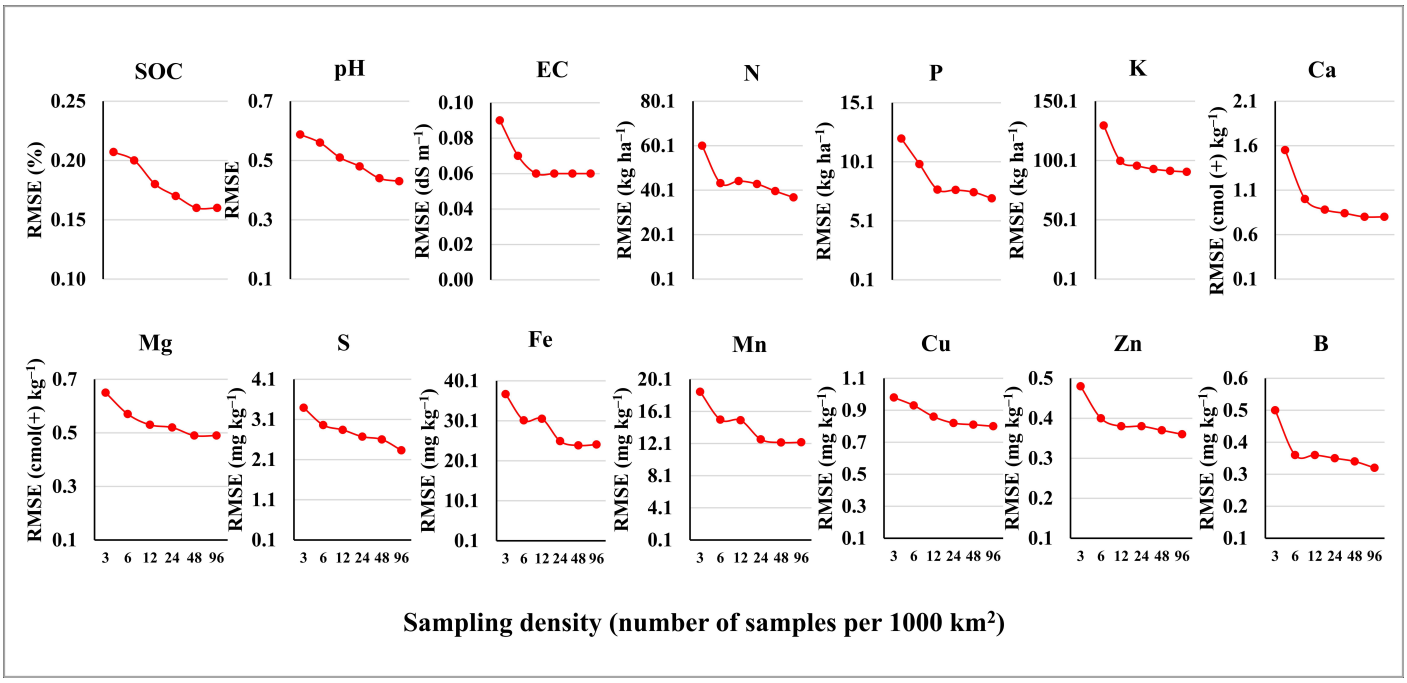


Figure 7. Trends of RMSEs with increasing sampling densities for individual soil properties. While gradual declines in RMSE values are observed for the soil properties SOC, pH, Mg, S, and Cu, other soil properties exhibited sharp declines in RMSE initially and then slow declines thereafter with increasing sampling density.

3.3. Prediction Maps

Soil properties SOC, EC, N, P, S, Fe, Mn, Cu, Zn, and B were observed to have the highest concentrations in the North-West part of the study area, whereas the remaining soil properties were observed to have higher concentrations towards the South-East part of the study area (Figure 8a,b). These patterns suggest that many of the soil properties covary with each other, while the other soil properties (i.e., pH, K, Ca, Mg) appear to have an inverse relationship with those properties.

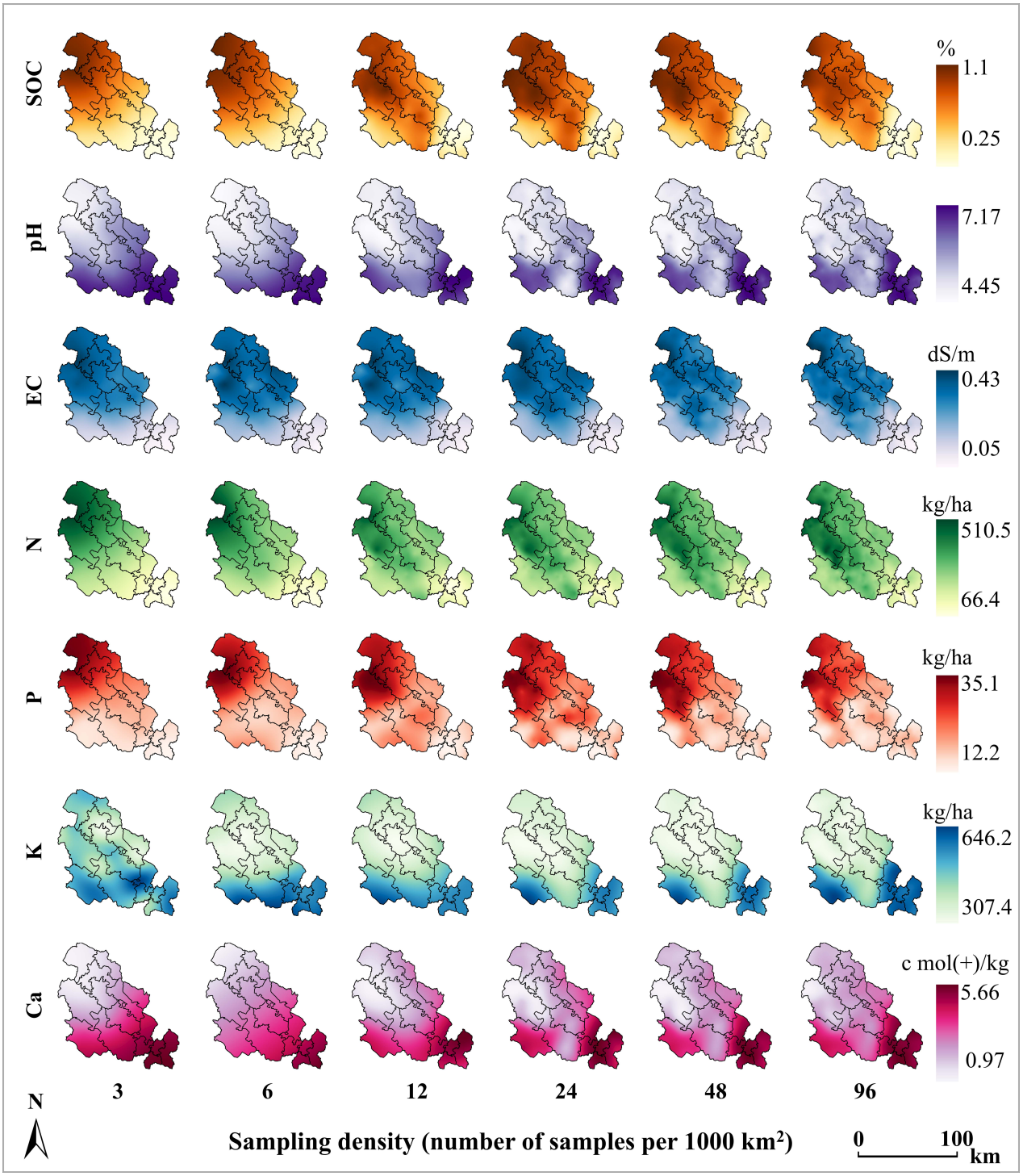


Figure 8. *Cont.*

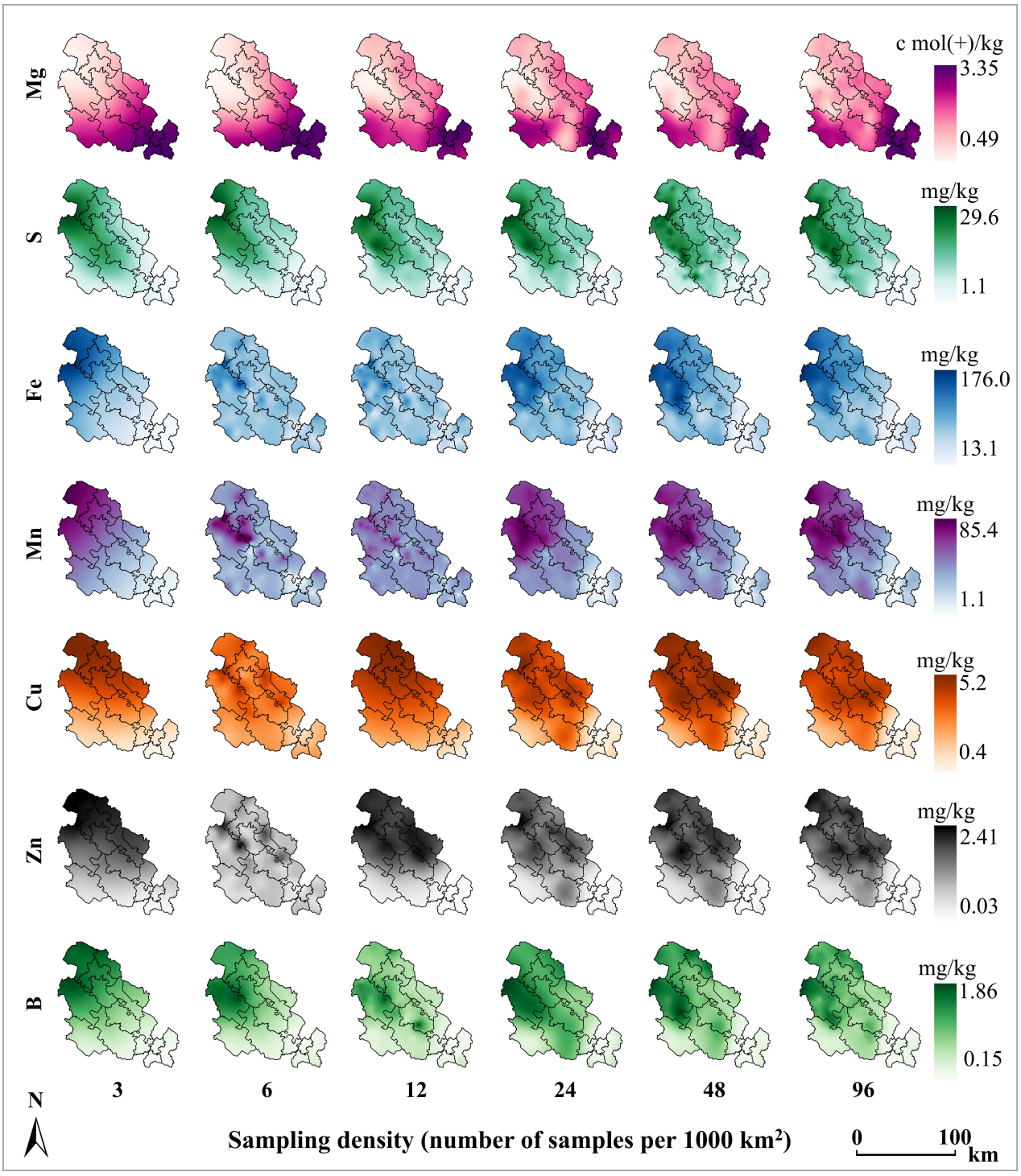


Figure 8. (**a**) Spatial prediction maps for SOC, pH, EC, N, P, K, and Ca at different sampling densities. For any soil property, the overall spatial pattern of the resulting maps remained nearly the same across all the sampling densities. Increasing sampling density corresponded to a decline in the smoothing effect and a subsequent increase in spatial heterogeneity of the resulting maps. (**b**) Spatial prediction maps for Mg, S, Fe, Mn, Cu, Zn, and B at different sampling densities. Similar to (**a**), the representation of spatial structures became more detailed with increasing sampling density.

For any soil property, the prediction maps generated similar patterns of spatial distribution at all the sampling densities. However, the prediction maps obtained from lower sampling densities were comparatively smoother across the study area compared to the ones produced with larger sampling densities. The map smoothness could be caused by

either a lower proportion of the map being outside the range of spatial autocorrelation identified by the respective semi-variograms or the semi-variograms identifying longer ranges of spatial autocorrelation with increased sampling densities. An intriguing pattern in the semi-variogram ranges with different sampling densities was that the shortest ranges were usually around 12 to 24 samples per 1000 km^2, with lower and higher sampling densities tending to have semi-variograms with longer ranges (Supplementary Materials). This pattern suggests that a consistent increase in map detail with increasing sampling density is more related to the proportion of the map area observed rather than differences in the semi-variograms generated at different sampling densities. In other words, once a plateau for improving the quality of the semi-variogram is reached, additional samples distributed across the mapping area can still improve the level of detail represented in the map.

4. Discussion

4.1. Effect of Sampling Density on Goodness-of-Fit for Semi-Variograms

For an effective kriging interpolation, the goodness-of-fit between the sample (experimental) and model (theoretical) semi-variograms should be satisfactory [12,61]. A reliable semi-variogram requires sufficient sample points for measuring and describing the characteristics of spatial autocorrelation [62]. Previous recommendations for sampling in support of spatial interpolation by kriging have tended to focus on minimum sample quantities. Past studies have recommended minimum quantities of samples ranging from 50 to 225 [13,29,63–65]. The improved model fits in this experimental dataset at 200 sample points, suggests a minimum quantity of points towards the higher end of this range for generating reliable semi-variograms at the regional scale.

Recommendations for minimum sample size have sometimes been linked to methods of semi-variogram estimation but likely also correspond with the sample set's ability to capture the pattern of spatial autocorrelation in the soil landscape of interest. This spatial pattern could be produced by physical processes in the landscape or moving average processes [66]. In any case, the chances of capturing that full pattern with the sample set decrease as sample locations become further apart. As an extreme example, a sample set of 100 data points distributed over a 100 km^2 study area would not be expected to capture the same patterns as they would if they were distributed over a 100,000 km^2 area. For this reason, each of the sample sets in the present study has been expressed in terms of sampling density. However, this concept does not exclude the possibility that there is also a minimum quantity of samples needed to construct a reliable semi-variogram, regardless of the study area size.

In this experiment, the sample semi-variograms obtained with sampling densities below 24 samples per 1000 km^2 (sample size below 200 points) were observed to be noisy and erratic, whereas those obtained at the higher sampling densities were smooth and stable (Figure 3). Reflecting this pattern, the goodness-of-fit between the sample and model semi-variograms (R^2 of model fit) was also observed to improve with increasing sampling density (Figure 4). Thus, higher sampling densities or higher sample sizes favoured generating semi-variograms with less noise and improved the goodness-of-fit. As an example of an analysis to identify sufficient sampling for supporting a reliable semi-variogram, Oliver and Webster reanalyzed exchangeable potassium data from the Broom's Barn farm [57,67]. They observed that the sample semi-variogram obtained from a smaller sample size (87) was much more erratic and noisier compared to the one obtained from a larger sample size (434). Accordingly, the goodness-of-fit for semi-variograms was also poorer in the case where a smaller sample size was used. Ye et al. while predicting SOC over a spatial extent of 16,400 km^2, also observed that the goodness-of-fit for semi-variograms improved with an increasing sampling density [32]. The reason behind the poor semi-variogram fittings at lower sampling densities could be that too few sample points might not sufficiently capture the complex spatial structures of soil properties. In such situations, semi-variogram estimates generated from very few sample points are usually unreliable [57,68].

4.2. Effect of Sampling Density on Spatial Prediction Accuracy

Trends of increasing CCCs and decreasing RMSEs for independent validation with the increasing sampling density indicate that irrespective of the soil property, prediction accuracy improved with sampling density. These results are in line with many other studies, both at local [28,41] and regional scales [16,31,33,35]. Improvements in prediction accuracy with the increase in sampling density can be attributed, at least in part, to the increased amount of representativeness of the spatial variation of the soil properties. In other words, more of the map area had been directly observed. The mean distance between samples decreased with the increased sampling density, which resulted in a reduced overall standard error of kriging interpolation. Additionally, increased sample density could have better captured the spatial autocorrelation structures occurring at smaller phenomenon scales [39]. Decreased distances between sample locations associated with the increased sampling density might have also added useful spatial information, leading to improved prediction accuracies at the higher sampling densities.

The decrease in the rate of gain in prediction accuracy with the increasing sampling density could be because the sample points became increasingly spatially clustered, causing the information to reach a saturation level [39]. This is in line with McBratney and Webster's observation that spatial autocorrelation reduces the quantity of samples needed to achieve lower standard errors [69]. The relatively consistent trend observed in the present study of increasing prediction accuracy with increasing sample density up to the point of diminishing improvement largely agrees with trends reported in the literature. Nonetheless, some deviations from that trend have been observed. For example, similar to the results in the present study, Long et al. observed that the rate of improvement in the prediction accuracy of SOM was gradual in the first stage of increasing sampling density, which was then followed by the rate of improvement being stagnated in the second stage [39]. However, in the third stage of increasing sampling density, the rate of improvement in prediction accuracy increased again and at a faster rate. In their case, the mapping extent was sufficiently larger (geographical area of 124,000 km^2), and the calibration sample sizes were also sufficiently larger (varied between 9365 and 188,247 soil samples). The findings of Li are also interestingly different as the validation RMSE of ordinary kriging interpolation initially decreased and then again started increasing linearly with the increasing sampling density while predicting SOM content over a geographical area of 400 km^2 in China [70]. Recently, Sun et al. observed that even though prediction accuracy in general improved with the sampling density, the increasing trend was abrupted at certain levels of increasing sampling density [33]. These exceptions suggest that plateaus of prediction accuracy may occur in certain situations.

4.3. Effect of Sampling Density on Spatial Patterns Observed in the Prediction Maps

For any soil property, the prediction maps prepared from different sampling densities followed similar patterns of spatial distribution. Long et al. also observed that the spatially interpolated prediction maps of SOM obtained at 20 different sampling densities were visually similar [39]. Similarly, Domenech et al. observed that the prediction maps for soil depth obtained with six different sample sizes possessed similar spatial patterns [41].

Although the prediction maps across the sampling densities possessed similar spatial patterns, the ones obtained from smaller sampling densities were smoother. The more generalized maps could roughly represent the spatial patterns but lacked details about the local variations. Another problem with the smoothed maps is that they might increase the low values and reduce the high values of soil properties [71]. Even though the smoother maps might be useful for certain broad-level applications, they may be less suitable for the more precise management recommendations for soil properties at finer scales [72]. In contrast, the maps obtained with higher sampling densities were more fragmented, which could represent more details about the local variations. Similarly, Sun et al. observed that spatial heterogeneity increased with increasing sampling density, which could capture the spatial variations of soil properties caused by anthropogenic activities at smaller distances [37].

Thus, even though the lower sampling densities can produce the overall pattern of spatial distributions, the local variations are less likely to be represented [32]. Overall, results from these and the present study agree with Loiseau et al., who observed that with an increasing sampling density, the stabilization of spatial prediction increased, the smoothing effect decreased, and spatial structures were described in finer detail with the kriged prediction maps [16].

In contrast, John et al. observed that SOM maps prepared with ordinary kriging of 150 samples differed from those prepared with 250, 450, 600, and 1200 sample points [17]. However, those differences were related to semi-variogram ranges that were relatively short to sample spacing, leading to different representations of 'hot spots' surrounding point locations. Lai et al. also observed that the spatial distribution pattern of prediction maps derived from ordinary kriging interpolation with lower sampling densities differed from those obtained with larger sampling densities [40]. In this case, the map of SOM became smoother with the use of additional samples. The patterns in the resulting map suggest that the additional samples modified the modeling of the semi-variogram such that the range was likely increasing with the additional samples, which converted the local 'hot spots' around a few points to more regional groupings with the added samples.

Irrespective of the similarity of the overall spatial pattern of the prediction maps across all the sampling densities, the spatial patterns obtained with higher sampling densities (viz., 24, 48, and 96 samples per 1000 km^2) were observed to be visually more identical to each other. Visual similarities among prediction maps prepared with different sampling densities could be related to the reuse of sample locations across the sampling densities in the present study's experimental design [40]. However, the experimental sampling densities were increased exponentially to keep the proportion of new sample locations at 50% for each increment. Visual comparison of the whole study area will tend to emphasize the major trend. The increase in CCC is likely a better indicator of how well the map represents more local variations.

4.4. Which Sampling Density Should Be Considered for Optimal Interpolation of Soil Properties?

The utility of soil maps and the requirements of the stakeholders are the ultimate driving forces in deciding the accuracy requirements. For example, if the stakeholders are policy makers or decision makers, the coarse resolution maps prepared from a small sampling density with a relatively lower accuracy might be sufficient, but not for farmers for their field management. While lower levels of accuracy may be acceptable for lowering costs, this study focused on the inflection point for the rate of gain in improved model performance. At the sampling density of 3 samples per 1000 km^2, all the soil properties possessed their lowest R^2 of model fits and CCCs for independent validation. By these metrics, model quality improved—with diminishing returns—for increasing sample density until the improvement was negligible, moving from 48 to 96 samples per 1000 km^2. Across the soil properties examined in this study, a sampling density of 48 per 1000 km^2 was identified as that inflection point where there was no substantial improvement in model performance for the next increment of increased sampling density. A possible exception to that was the lack of a plateau for independent validation of the physico-chemical category of soil properties. Although the goodness-of-fit for semi-variogram models did not change, the prediction accuracy of the resulting maps may have still benefited from additional increases in sampling density.

A brief literature review was conducted to compare the results of this study to those of previous studies. For this purpose, the results of the previous studies, either reported an optimal sampling density or sample size, were converted into sampling density for quantitatively identifying the relationship between sampling density and mapping extent. When considered in the context of previous research along with the size of the mapping extents, the results of this study fit a pattern of optimal sampling density decreasing as a function of increasing mapping extent (Figure 9). A decrease in optimal sampling density for larger mapping extents suggests potential explanations. One possibility is that more

samples are providing less additional information for spatial modeling due to spatial autocorrelation. McBratney and Webster demonstrated how spatial autocorrelation reduces the number of samples needed to achieve a desired level of standard error, which may translate to lower densities of samples needed to produce maps with optimized prediction performance with kriging [69]. Another possibility is that the statistics focus on larger-scale trends with larger mapping extents, with more local variation being treated as noise. However, the maps produced in this study with lower sampling densities were the most generalized and became more spatially detailed up to optimal sample density. The degree to which spatial detail was added to the resulting maps corresponded to the decrease in the rates of gain for semi-variogram goodness-of-fit and independent validation CCC.

In the literature review of optimal sampling density for kriging, a few studies did not fit the trend as well as the others (Figure 9). The one study that mapped soil depth was an outlier with the identification of around 323 samples per km^2 for fields ranging from 0.2 to 0.7 square km [41], which was a much higher sample density than would be expected for those mapping extents. Nanni et al. identified different optimal sampling densities for mapping clay content, SOM, P, K, and base saturation for an area of 1.84 square km [38]. The optimal sampling densities for clay content and SOM aligned with the observed trend with 35 and 48 samples per km^2, respectively. However, the authors concluded that a sampling density greater than 100 samples per km^2 was needed to map P, K, and base saturation in the area of highly managed sugar cane.

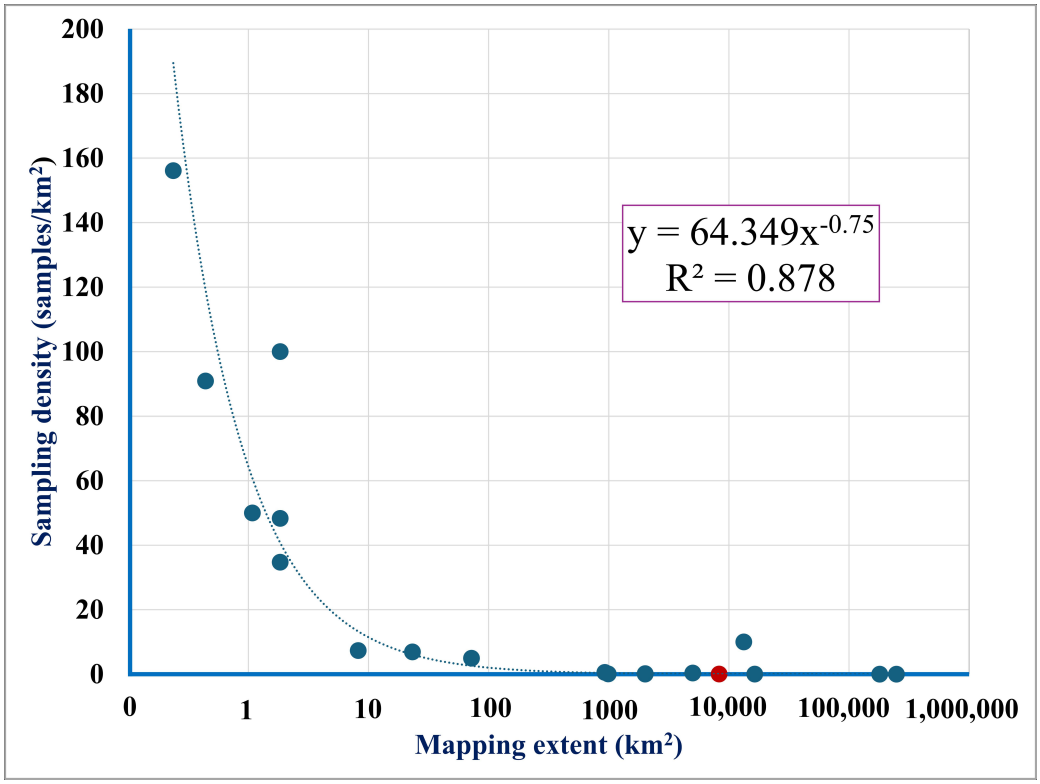


Figure 9. The relationship between mapping extent and optimal sampling density is based on a literature review of 21 studies (blue dots) and the present study (red dot). These studies included SOC, SOC stock, SOM, particle size fraction, electrical conductivity, pH, total nitrogen, phosphorus, potassium, iron, manganese, copper, zinc, boron, aluminium, and base saturation. A study mapping soil depth that identified the equivalent of 323 samples per km^2 for fields ranging from 0.2 to 0.7 km^2 [41] has been excluded as an outlier.

The optimal sampling densities identified in the present study for macro and micronutrients were consistent with the general trends, but the area being mapped has likely been less intensely managed than the fields in [38]. This difference suggests that management

may induce an additional level of spatial variability that requires a higher sampling density. Moreover, landform complexity and land use patterns are also supposed to affect the sampling density requirement and prediction accuracy of soil mapping. For example, the findings of Sun et al. demonstrate that for the same set of sample quantities, accuracy metrics were better in plain than in the hill mountainous landform [33]. Similarly, Safaee et al. observed that prediction performance was more consistent and accurate in uniformly managed lands than in intensively managed ones [26]. Therefore, the sampling density-prediction accuracy relationships could be rather more complex and unique (site-specific) to an area depending on the local factors such as existing spatial autocorrelations, climate, topography, landform complexity, and management practices. Thus, future studies may include these additional local factors in addition to the experience gained in this study to explore the impact of sapling density on spatial prediction accuracy further.

5. Conclusions

This study highlights the critical role of sampling density in generating reliable predictions of soil properties through semi-variogram models. Sampling densities below 24 samples per 1000 km² produced noisy and unstable results, but as densities increased, the semi-variograms stabilized, and model fit improved, with higher R^2 values and better prediction accuracy (increased CCCs and decreased RMSEs). However, the rate of improvement diminished as sampling density increased, indicating diminishing returns beyond a certain point. For instance, increasing sampling density from 48 to 96 samples per 1000 km² had minimal effect on the model's goodness-of-fit and prediction accuracy, suggesting a saturation threshold.

A strong correlation was observed between model fit (R^2) and prediction performance (CCC), especially for physico-chemical properties, which showed higher accuracy compared to macronutrients and micronutrients. The reliability of prediction accuracy or probability of obtaining a higher R^2 of model fit or CCC, regardless of soil property, increased with increased sampling densities. Prediction accuracy across all the soil properties was most reliable at the sampling density of 48 samples per 1000 km^2. The mean CCC values decreased in the order of physico-chemical properties (0.74) > macronutrients (0.67) > micronutrients (0.53). This indicates that different soil properties respond differently to increases in sampling density, with physico-chemical properties being the most predictable. Nevertheless, considering results in the context of previous work on mapping with kriging, this study identified a general trend of optimal sampling density decreasing with exponential increases in mapping extent. The study emphasizes that regional soil maps, despite lower prediction accuracy for nutrients, provide valuable spatial information that can guide policymakers in soil management. These maps can inform planners where nutrient management is less intensive, offering a practical tool for broad-scale agricultural recommendations when site-specific data are unavailable or too costly to obtain.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/land13101615/s1, Figure S1: Semi-variograms for induvial soil properties at different sampling densities; Table S1: Semi-variogram parameters of soil properties at different sampling densities.

Author Contributions: Conceptualization, P.K.D. and B.A.M.; Formal analysis, P.K.D., B.A.M. and N.P; Software, P.K.D. and B.A.M.; Supervision, A.M.; Writing—original draft, P.K.D.; Writing—review and editing, B.A.M., N.P. and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Soil data generated through laboratory analysis can be made available upon request.

Acknowledgments: The authors are thankful to the Department of Science and Technology, Government of India, for providing an Inspire Fellowship to the first author (IF180424). The authors are also thankful

to the US Department of State and the United States-India Educational Foundation (USIEF), New Delhi, India, for providing a Fulbright-Nehru Doctoral Research Fellowship to the first author (Award No. 2598/DR/2020-2021). Laboratory facility by ICAR-AICRP on STCR is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Franzen, D.W.; Peck, T.R. Field Soil Sampling Density for Variable Rate Fertilization. *J. Prod. Agric.* **1995**, *8*, 568–574. [CrossRef]
2. Agriculture Census Division. *All India Report on Number and Area of Operational Holdings, Agricultural Census (2015–2016)*; Government of India: New Delhi, India, 2019.
3. Ministry of Finance. *Department of Economic Affairs Economic Survey 2023–2024*; Government of India: New Delhi, India, 2024.
4. Ministry of Finance. *Budget Division Expenditure Budget 2024–2025*; Government of India: New Delhi, India, 2024.
5. Dasgupta, S.; Debnath, S.; Das, A.; Biswas, A.; Weindorf, D.C.; Li, B.; Shukla, A.K.; Das, S.; Saha, S.; Chakraborty, S. Developing Regional Soil Micronutrient Management Strategies through Ensemble Learning Based Digital Soil Mapping. *Geoderma* **2023**, *433*, 116457. [CrossRef]
6. Hinsinger, P.; Gobran, G.R.; Gregory, P.J.; Wenzel, W.W. Rhizosphere Geometry and Heterogeneity Arising from Root-mediated Physical and Chemical Processes. *New Phytol.* **2005**, *168*, 293–303. [CrossRef] [PubMed]
7. Liu, W.; Du, P.; Zhao, Z.; Zhang, L. An Adaptive Weighting Algorithm for Interpolating the Soil Potassium Content. *Sci. Rep.* **2016**, *6*, 23889. [CrossRef]
8. Krige, D. Two-Dimensional Weighted Moving Average Trend Surfaces for Ore Valuation. In Proceedings of the Symposium on Mathematical Statistics and Computer Applications in Ore Valuation, Johannesburg, South Africa, 7–8 March 1966; pp. 13–38.
9. Miller, B.A. Chapter 5–Geographic Information Systems and Spatial Statistics Applied for Soil Mapping: A Contribution to Land Use Management. In *Soil Mapping and Process Modeling for Sustainable Land Use Management*; Pereira, P., Brevik, E.C., Muñoz-Rojas, M., Miller, B.A., Eds.; Elsevier: Amsterdam, The Netherlands, 2017; pp. 127–149. ISBN 978-0-12-805200-6.
10. Erdogan Erten, G.; Yavuz, M.; Deutsch, C.V. Combination of Machine Learning and Kriging for Spatial Estimation of Geological Attributes. *Nat. Resour. Res.* **2022**, *31*, 191–213. [CrossRef]
11. Zhu, C.; Wei, Y.; Zhu, F.; Lu, W.; Fang, Z.; Li, Z.; Pan, J. Digital Mapping of Soil Organic Carbon Based on Machine Learning and Regression Kriging. *Sensors* **2022**, *22*, 8997. [CrossRef]
12. Oliver, M.A. Geostatistics and Its Application to Soil Science. *Soil. Use Manag.* **1987**, *3*, 8–20. [CrossRef]
13. Oliver, M.A.; Webster, R. *Basic Steps in Geostatistics: The Variogram and Kriging*; Briefs in Agriculture; Springer International Publishing: Cham, Switzerland, 2015; ISBN 978-3-319-15864-8.
14. Goovaerts, P. *Geostatistics for Natural Resources Evaluation*; Oxford University Press on Demand: Cary, NC, USA, 1997; ISBN 0-19-511538-4.
15. Miller, B.A.; Koszinski, S.; Hierold, W.; Rogasik, H.; Schröder, B.; Van Oost, K.; Wehrhan, M.; Sommer, M. Towards Mapping Soil Carbon Landscapes: Issues of Sampling Scale and Transferability. *Soil. Tillage Res.* **2016**, *156*, 194–208. [CrossRef]
16. Loiseau, T.; Arrouays, D.; Richer-de-Forges, A.C.; Lagacherie, P.; Ducommun, C.; Minasny, B. Density of Soil Observations in Digital Soil Mapping: A Study in the Mayenne Region, France. *Geoderma Reg.* **2021**, *24*, e00358. [CrossRef]
17. John, K.; Bouslihim, Y.; Ofem, K.I.; Hssaini, L.; Razouk, R.; Okon, P.B.; Isong, I.A.; Agyeman, P.C.; Kebonye, N.M.; Qin, C. Do Model Choice and Sample Ratios Separately or Simultaneously Influence Soil Organic Matter Prediction? *Int. Soil. Water Conserv. Res.* **2022**, *10*, 470–486. [CrossRef]
18. McBratney, A.B.; Webster, R.; Burgess, T.M. The Design of Optimal Sampling Schemes for Local Estimation and Mapping of Regionalized Variables—I. *Comput. Geosci.* **1981**, *7*, 331–334. [CrossRef]
19. Hu, M.-G.; Wang, J.-F. A Spatial Sampling Optimization Package Using MSN Theory. *Environ. Model. Softw.* **2011**, *26*, 546–548. [CrossRef]
20. Ferreyra, R.A.; Apezteguía, H.P.; Sereno, R.; Jones, J.W. Reduction of Soil Water Spatial Sampling Density Using Scaled Semivariograms and Simulated Annealing. *Geoderma* **2002**, *110*, 265–289. [CrossRef]
21. Qiao, P.; Dong, N.; Lei, M.; Yang, S.; Gou, Y. An Effective Method for Determining the Optimal Sampling Scale Based on the Purposes of Soil Pollution Investigations and the Factors Influencing the Pollutants. *J. Hazard. Mater.* **2021**, *418*, 126296. [CrossRef] [PubMed]
22. Saurette, D.D.; Heck, R.J.; Gillespie, A.W.; Berg, A.A.; Biswas, A. Sample Size Optimization for Digital Soil Mapping: An Empirical Example. *Land* **2024**, *13*, 365. [CrossRef]
23. Schmidinger, J.; Schröter, I.; Bönecke, E.; Gebbers, R.; Ruehlmann, J.; Kramer, E.; Mulder, V.L.; Heuvelink, G.B.M.; Vogel, S. Effect of Training Sample Size, Sampling Design and Prediction Model on Soil Mapping with Proximal Sensing Data for Precision Liming. *Precis. Agric.* **2024**, *25*, 1529–1555. [CrossRef]
24. De Gruijter, J.; Brus, D.J.; Bierkens, M.F.; Knotters, M. *Sampling for Natural Resource Monitoring*; Springer: Berlin/Heidelberg, Germany, 2006; Volume 665.
25. Brungard, C.W.; Boettinger, J.L. Conditioned Latin Hypercube Sampling: Optimal Sample Size for Digital Soil Mapping of Arid Rangelands in Utah, USA. In *Digital Soil Mapping*; Boettinger, J.L., Howell, D.W., Moore, A.C., Hartemink, A.E., Kienast-Brown, S., Eds.; Springer: Dordrecht, The Netherlands, 2010; pp. 67–75. ISBN 978-90-481-8862-8.

26. Safaee, S.; Libohova, Z.; Kladivko, E.J.; Brown, A.; Winzeler, E.; Read, Q.; Rahmani, S.; Adhikari, K. Influence of Sample Size, Model Selection, and Land Use on Prediction Accuracy of Soil Properties. *Geoderma Reg.* **2024**, *36*, e00766. [CrossRef]
27. Vašát, R.; Heuvelink, G.B.M.; Borůvka, L. Sampling Design Optimization for Multivariate Soil Mapping. *Geoderma* **2010**, *155*, 147–153. [CrossRef]
28. Wang, J.; Yang, R.; Bai, Z. Spatial Variability and Sampling Optimization of Soil Organic Carbon and Total Nitrogen for Minesoils of the Loess Plateau Using Geostatistics. *Ecol. Eng.* **2015**, *82*, 159–164. [CrossRef]
29. Sherpa, S.R.; Wolfe, D.W.; Van Es, H.M. Sampling and Data Analysis Optimization for Estimating Soil Organic Carbon Stocks in Agroecosystems. *Soil Sci. Soc. Amer. J.* **2016**, *80*, 1377–1392. [CrossRef]
30. Zhang, Z.; Huang, X.; Zhou, Y. Spatial Heterogeneity of Soil Organic Carbon in a Karst Region under Different Land Use Patterns. *Ecosphere* **2020**, *11*, e03077. [CrossRef]
31. Yu, D.-S.; Zhang, Z.-Q.; Yang, H.; Shi, X.-Z.; Tan, M.-Z.; Sun, W.-X.; Wang, H.-J. Effect of Soil Sampling Density on Detected Spatial Variability of Soil Organic Carbon in a Red Soil Region of China. *Pedosphere* **2011**, *21*, 207–213. [CrossRef]
32. Ye, H.; Huang, W.; Huang, S.; Huang, Y.; Zhang, S.; Dong, Y.; Chen, P. Effects of Different Sampling Densities on Geographically Weighted Regression Kriging for Predicting Soil Organic Carbon. *Spat. Stat.* **2017**, *20*, 76–91. [CrossRef]
33. Sun, X.-L.; Lai, Y.-Q.; Ding, X.; Wu, Y.-J.; Wang, H.-L.; Wu, C. Variability of Soil Mapping Accuracy with Sample Sizes, Modelling Methods and Landform Types in a Regional Case Study. *Catena* **2022**, *213*, 106217. [CrossRef]
34. Lark, R.M. Multi-Objective Optimization of Spatial Sampling. *Spat. Stat.* **2016**, *18*, 412–430. [CrossRef]
35. Wang, J.; Wang, X.; Zhou, S.; Wu, S.; Zhu, Y.; Lu, C. Optimization of Sample Points for Monitoring Arable Land Quality by Simulated Annealing While Considering Spatial Variations. *Int. J. Environ. Res. Public Health* **2016**, *13*, 980. [CrossRef]
36. Li, X.; Gao, B.; Bai, Z.; Pan, Y.; Gao, Y. An Improved Parallelized Multi-Objective Optimization Method for Complex Geographical Spatial Sampling: AMOSA-II. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 236. [CrossRef]
37. Sun, W.; Zhao, Y.; Huang, B.; Shi, X.; Landon Darilek, J.; Yang, J.; Wang, Z.; Zhang, B. Effect of Sampling Density on Regional Soil Organic Carbon Estimation for Cultivated Soils. *J. Plant Nutr. Soil. Sci.* **2012**, *175*, 671–680. [CrossRef]
38. Nanni, M.R.; Povh, F.P.; Demattê, J.A.M.; Oliveira, R.B.D.; Chicati, M.L.; Cezar, E. Optimum Size in Grid Soil Sampling for Variable Rate Application in Site-Specific Management. *Sci. Agric.* **2011**, *68*, 386–392. [CrossRef]
39. Long, J.; Liu, Y.; Xing, S.; Qiu, L.; Huang, Q.; Zhou, B.; Shen, J.; Zhang, L. Effects of Sampling Density on Interpolation Accuracy for Farmland Soil Organic Matter Concentration in a Large Region of Complex Topography. *Ecol. Indic.* **2018**, *93*, 562–571. [CrossRef]
40. Lai, Y.-Q.; Wang, H.-L.; Sun, X.-L. A Comparison of Importance of Modelling Method and Sample Size for Mapping Soil Organic Matter in Guangdong, China. *Ecol. Indic.* **2021**, *126*, 107618. [CrossRef]
41. Domenech, M.B.; Castro-Franco, M.; Costa, J.L.; Amiotti, N.M. Sampling Scheme Optimization to Map Soil Depth to Petrocalcic Horizon at Field Scale. *Geoderma* **2017**, *290*, 75–82. [CrossRef]
42. Radočaj, D.; Jug, I.; Vukadinović, V.; Jurišić, M.; Gašparović, M. The Effect of Soil Sampling Density and Spatial Autocorrelation on Interpolation Accuracy of Chemical Soil Properties in Arable Cropland. *Agronomy* **2021**, *11*, 2430. [CrossRef]
43. Marques, J., Jr.; Alleoni, L.R.F.; Teixeira, D.D.B.; Siqueira, D.S.; Pereira, G.T. Sampling Planning of Micronutrients and Aluminium of the Soils of São Paulo, Brazil. *Geoderma Reg.* **2015**, *4*, 91–99. [CrossRef]
44. Five decades of Odisha Agriculture Statistics, Technical Report, Directorate of Agriculture and Food Production, Government of Odisha. 2020. Available online: https://agri.odisha.gov.in/sites/default/files/2023-03/Five%20Decades%20of%20Odisha%20Agriculture%20Statistics_compressed.pdf (accessed on 10 July 2021).
45. Mishra, A.; Saren, S.; Dey, P. *Soils of Keonjhar District and Improvement of Crop Productivity in Tribal Villages through Adaption of STCR Technology*; AICRP on STCR, Department of Soil Science and Agricultural Chemistry, OUAT: Bhubaneswar, India, 2016.
46. Sarkar, D.; Mandal, D.; Haldar, A. Soil Maturity Assessment along a Toposequence in Chotanagpur Plateau, West Bengal Using Inorganic Soil Phosphorus Based Weathering Index, Soil Taxonomy and Other Chemical Indices: A Comparative Study. *Agropedology* **2014**, *24*, 82–94.
47. Walkley, A.; Black, I.A. An Examination of the Degtjareff Method for Determining Soil Organic Matter, and a Proposed Modification of the Chromic Acid Titration Method. *Soil. Sci.* **1934**, *37*, 29–38. [CrossRef]
48. Jackson, M.L. *Soil Chemical Analysis*; Prentice Hall of India Pvt. Ltd.: New Delhi, India, 1973; Volume 498.
49. Subbiah, B.V.; Asija, G.L. A Rapid Method for the Estimation of Nitrogen in Soil. *Curr. Sci.* **1956**, *26*, 259–260.
50. Bray, R.H.; Kurtz, L.T. Determination of Total, Organic, and Available Forms of Phosphorus in Soils. *Soil. Sci.* **1945**, *59*, 39–46. [CrossRef]
51. Hanway, J.; Heidel, H. Soil Analysis Methods as Used in Iowa State College Soil Testing Laboratory. *Iowa Agric.* **1952**, *57*, 1–31.
52. Page, A.; Miller, R.; Keeney, D. *Methods of Soil Analysis. Part 2. Chemical and Microbiological Properties*; Agronomy, No. 9; Soil Science Society of America: Madison, WI, USA, 1982; p. 1159.
53. Chesnin, L.; Yien, C. Turbidimetric Determination of Available Sulfates. *Soil. Sci. Soc. Am. J.* **1951**, *15*, 149–151. [CrossRef]
54. Lindsay, W.L.; Norvell, W. Development of a DTPA Soil Test for Zinc, Iron, Manganese, and Copper. *Soil. Sci. Soc. Am. J.* **1978**, *42*, 421–428. [CrossRef]
55. John, M.K.; Chuah, H.H.; Neufeld, J.H. Application of Improved Azomethine-H Method to the Determination of Boron in Soils and Plants. *Anal. Lett.* **1975**, *8*, 559–568. [CrossRef]

56. Pebesma, E.J. *Gstat User's Manual*; Department of Physical Geography, Utrecht University: Utrecht, The Netherlands, 2001; Volume 1, pp. 499–505.
57. Oliver, M.A.; Webster, R. A Tutorial Guide to Geostatistics: Computing and Modelling Variograms and Kriging. *Catena* **2014**, *113*, 56–69. [CrossRef]
58. Silva, R.A.; Siqueira, G.M.; Costa, M.K.L.; Guedes Filho, O.; Silva, Ê.F.D.F.E. Spatial Variability of Soil Fauna under Different Land Use and Managements. *Rev. Bras. De Ciência Do Solo* **2018**, *42*, e0170121. [CrossRef]
59. Matheron, G. *Les Variables Régionalisées et Leur Estimation: Une Application de La Théorie de Fonctions Aléatoires Aux Sciences de La Nature*; Masson et Cie: Echandens, Switzerland, 1965.
60. Webster, R.; Oliver, M.A. *Geostatistics for Environmental Scientists*; John Wiley & Sons: Hoboken, NJ, USA, 2007.
61. Webster, R. Quantitative Spatial Analysis of Soil in the Field. In *Soil Restoration*; Lal, R., Stewart, B.A., Eds.; Advances in Soil Science; Springer: New York, NY, USA, 1985; Volume 17, pp. 1–70. ISBN 978-1-4612-7684-5.
62. Lark, R.M. Optimized Spatial Sampling of Soil for Estimation of the Variogram by Maximum Likelihood. *Geoderma* **2002**, *105*, 49–80. [CrossRef]
63. Webster, R.; Oliver, M.A. Sample Adequately to Estimate Variograms of Soil Properties. *J. Soil. Sci.* **1992**, *43*, 177–192. [CrossRef]
64. Gascuel-Odoux, C.; Boivin, P. Variability of Variograms and Spatial Estimates Due to Soil Sampling: A Case Study. *Geoderma* **1994**, *62*, 165–182. [CrossRef]
65. Kerry, R.; Oliver, M.A. Comparing Sampling Needs for Variograms of Soil Properties Computed by the Method of Moments and Residual Maximum Likelihood. *Geoderma* **2007**, *140*, 383–396. [CrossRef]
66. Lark, R.M. Two Contrasting Spatial Processes with a Common Variogram: Inference about Spatial Models from Higher-order Statistics. *Eur. J. Soil. Sci.* **2010**, *61*, 479–492. [CrossRef]
67. Webster, R.; McBratney, A.B. Mapping Soil Fertility at Broom's Barn by Simple Kriging. *J. Sci. Food Agric.* **1987**, *38*, 97–115. [CrossRef]
68. Roberton, S.D.; Bennett, J.M.L.; Lobsey, C.R.; Bishop, T.F.A. Assessing the Sensitivity of Site-Specific Lime and Gypsum Recommendations to Soil Sampling Techniques and Spatial Density of Data Collection in Australian Agriculture: A Pedometric Approach. *Agronomy* **2020**, *10*, 1676. [CrossRef]
69. McBratney, A.B.; Webster, R. How Many Observations Are Needed for Regional Estimation of Soil Properties? *Soil. Sci.* **1983**, *135*, 177–183. [CrossRef]
70. Li, Y. Can the Spatial Prediction of Soil Organic Matter Contents at Various Sampling Scales Be Improved by Using Regression Kriging with Auxiliary Information? *Geoderma* **2010**, *159*, 63–75. [CrossRef]
71. Zhang, Z.; Sun, Y.; Yu, D.; Mao, P.; Xu, L. Influence of Sampling Point Discretization on the Regional Variability of Soil Organic Carbon in the Red Soil Region, China. *Sustainability* **2018**, *10*, 3603. [CrossRef]
72. Dash, P.K.; Panigrahi, N.; Mishra, A. Identifying Opportunities to Improve Digital Soil Mapping in India: A Systematic Review. *Geoderma Reg.* **2022**, *28*, e00478. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



https://dspace.org/

Exploring the effect of sampling density on spatial prediction with spatial interpolation of multiple soil nutrients at a regional scale

Dash, Prava Kiran; Miller, Bradley A.; Panigrahi, Niranjan; Mishra, Antaryami

2024-10-01
MDPI
https://doi.org/10.3390/land13101615
Dash PK, Miller BA, Panigrahi N, Mishra A. (2024) Exploring the effect of sampling density on spatial prediction with spatial interpolation of multiple soil nutrients at a regional scale. Land, Volume 13, Issue 10, October 2024, Article number 1615

Downloaded from a DSpace repository. https://dspace.org/