Study of Multiple Regression Modal between Soil Properties and Colony-forming Unit Collected from Diverse Regions of Rajasthan, India

Himanshi Mangla, Saurabh Dave[1], Smita Jain[2], Hardik Pathak

Department of Biotechnology, JECRC University, Jaipur, Rajasthan, India, [1]Department of Chemistry, JECRC University, Jaipur, Rajasthan, India, [2]Department of Mathematics, JECRC University, Jaipur, Rajasthan, India

Abstract

Aim: This study was done to predict the relation between soil properties and colony-forming unit (CFU) at different landscape positions from observed values using simplified regression models. **Methods:** Soil samples (normal and petroleum contaminated) were collected from diverse regions of Rajasthan, India. The coefficient of correlation was calculated between the CFU and the physicochemical properties of soils such as pH, conductivity, moisture content, organic matter, and heavy metals. Before applying the regression models, the dependency of several soil properties with CFU, the principal component (PC) analysis, was applied. **Results:** As a result, only three PCs have eigenvalues >1. In the regression analysis of variance, the P value is found to be >0.05, so the regression model is significant. **Conclusion:** This approach would promote soil sampling and variable rate application of agricultural chemicals and would serve as a fast, inexpensive, and reasonably accurate method to develop a soil database for fields that have similar soil characteristics.

Keywords: Colony-forming units, correlation, principal component analysis, regression

Introduction

Soil microbial communities are among the most abundant and diverse on earth and are responsible for many ecologically and economically important ecosystem functions.[1-5] For instance, soil microbes play critical roles in bioremediation process where these microorganisms used these highly carcinogenic compounds as an energy source and performed their metabolic activity.

Rajasthan is the most diversified state of India where India's largest saline water lake Sambhar is present and another side the Thar Desert in Jaisalmer is present.[6] In this investigation, we have collected normal soil samples along with the petroleum-contaminated soils from different region, namely Jaisalmer (Thar Desert), Sambhar lake, Sri Ganganagar (northern part of the state), and Banswara (southern part of the state).

According to the old soil taxonomy developed by the US Department of Agriculture in 1949, the soil of Rajasthan has been divided into 8 types based on difference in climate and mineralogy.[7] According to the classification, Sri Ganganagar comes under Desert soil (sandy to sandy loam), Jaisalmer comes into Dunes and Associate soil (loamy fine sands to coarse), Banswara comes in Red Loames (sandy loams to sandy), and Sambhar comes in Saline-Sodic soil (dark gray to pale brown).[7]

Different multivariate statistics such as generalized linear model, principal component analysis (PCA), structural equation models, generalized additive models, random forest, support vector machine, artificial neural network, and interpolation techniques such as ordinary kriging, regression kriging, and inverse distance weighting are used in soil property mapping.[8-14]

Access this article online	
Quick Response Code:	**Website:** https://journals.lww.com/IJEH
	DOI: 10.4103/ijehe.ijehe_47_22

Address for correspondence: Prof. Hardik Pathak, Department of Chemistry, JECRC University, Jaipur, Rajasthan, India. E-mail: saurabh.chem76@gmail.com

How to cite this article: Mangla H, Dave S, Jain S, Pathak H. Study of multiple regression modal between soil properties and colony-forming unit collected from diverse regions of Rajasthan, India. Int J Env Health Eng 2025;14:18.

Received: 21-09-2022, **Accepted:** 03-05-2023, **Published:** 25-06-2025

The interpretation of these data offers a proper understanding of soil quality and draws significant conclusions that interest soil scientists recently[15] and which can result in a more sustainable agricultural use of areas now and in future. Similarly, many international studies have stated in the literature that spatial variability is used as a method for evaluating soil data.[16-18]

The purpose of the present investigation was to carry out (1) physicochemical properties of normal soil as well as petroleum-contaminated soil, (2) find the spatial variability of some parameters (pH, electrical conductivity [EC], $CaCO_3$, and Organic Matter (OM)) creating the maps of spatial distribution using the geographic information systems, (3) evaluate soil quality status in surface for irrigation purpose using multivariate statistical approaches, and (4) isolate as many as microorganisms which are capable to degrade hydrocarbons and select the best degrader from these soil samples. Soil microorganisms were isolated from soil samples, and once colony-forming unit (CFU) has been determined, coefficient of correlation was calculated between the CFU and soil properties for the different sampling stations. After getting the coefficient, PCA was applied to make the regression modal.

Materials and Methods

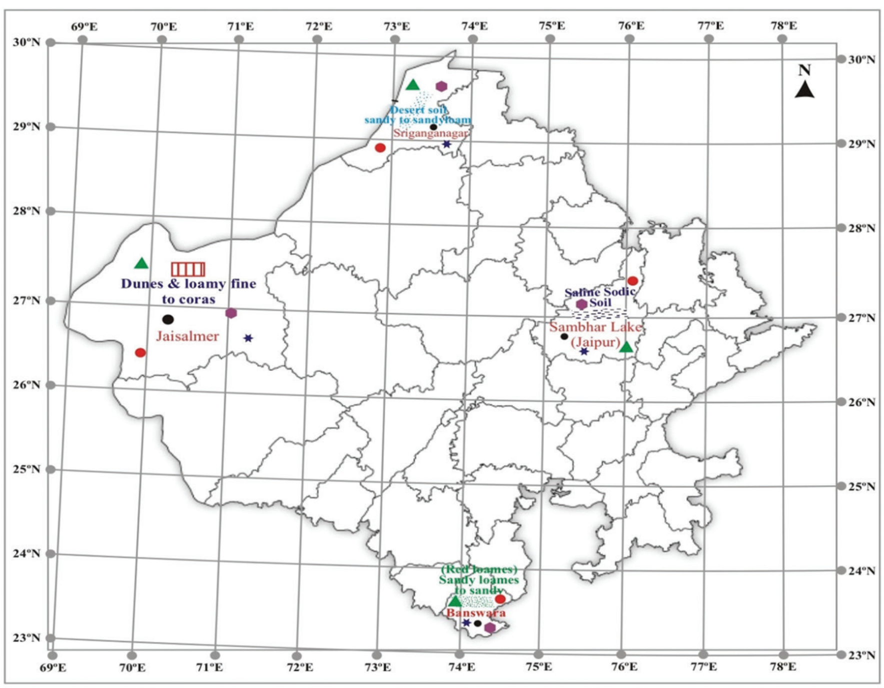
Sampling sites

Soil samples were collected from diverse regions of Rajasthan. Banswara district which is located at the southern part of Rajasthan is called Cherrapunji of Rajasthan. The average rainfall is 900–1200 mm.[19] In Jaisalmer district, which is situated in the Thar desert, the average temperature during the summer is 45 ºC- 48 ºC.[20] Another important part is Sambhar lake which comes in Jaipur district. At this place, cultivation is not possible due to the high degree of salinity. Sri Ganganagar is the northernmost district of Rajasthan. It is a plain region of Thar Desert where sandy soil in the west dotted with dunes is found.

Normal soil samples as well as petroleum-contaminated soils were collected from mentioned regions. Petroleum-contaminated soils were collected from Motor Garages of different stations. The samples were dug with a trowel from 0 to 20 cm below the topsoil surface. To remove plants or other waste residues, the soil is air-dried and sieved through 200 mesh sieves. The soil samples were then placed in sample storage bags and kept in the refrigerator until further analysis.

Physicochemical analysis

Soil properties were determined for each sample showing in Figure 1: (i) particle size analysis, (ii) percentage of organic matter content (OM), (iii) percentage of carbonate content ($CaCO_3$), (iv) EC (dS/m), and (v) pH. These properties were measured following standard techniques. The soil texture (sand, silt, and clay) was determined according to the international method of Robinson, standard AFNOR NF X31-107 using sodium hexametaphosphate for particle dispersion. The soil pH was measured in a suspension of 25 mL



S.No.	Symbol	Location	pH	Bulk density
1.		Ganganagar	7.5	1.30 gm/ml
2.		Banswara	7.65	1.63 gm/ml
3.		Jaisalmer	8.9	1.41 gm/ml
4.		Sambhar	9.4	1.10 gm/ml

S.No.	Location	Metal Symbol
1.	Ganganagar	● ▲ ⬢ *
2.	Banswara	* ▲ ● ⬢
3.	Jaisalmer	● ▲ ⬢ *
4.	Sambhar	⬢ ● ▲ *

S.No.	Metal Name	Symbol
1.	Iron	●
2.	Manganese	▲
3.	Zinc	⬢
4.	Coppper	*

Figure 1: Map of Rajasthan showing the physicochemical properties of four locations

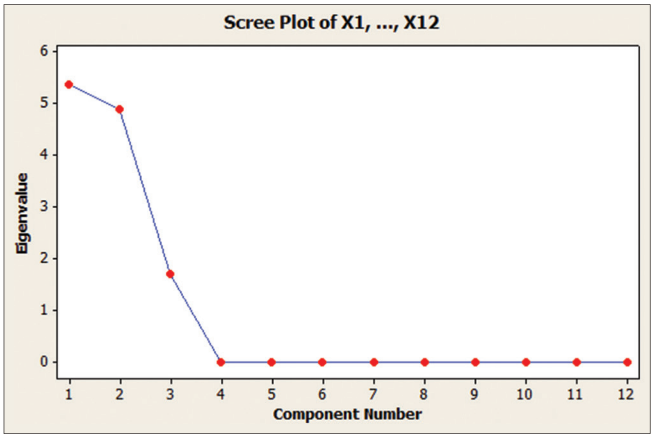


Figure 2: Scree plot of the parameters

deionized water to 10 g of soil (1/2.5 soil-to-water ratio [w/v]) after shaking for 1 h. The soil conductivity at a ratio of 1/5 (v/v) soil to distilled water was measured by preparing a slurry with 10 g of soil mixed with 50 ml of water. The slurry was vigorously shaken for 30 minutes prior to conducting the conductivity measurements. Samples were shaken for 1 h to obtain a balance between the solid phase and the liquid phase and allowed to settle for 2 h before analysis. The pH and Electric conductivity (EC) of the soil were determined using a Thermo Scientific Orion 4-Star Plus pH/conductivity meter. The total OM and carbonate content ($CaCO_3$) were determined using the Thermolyne furnace 6000°C for 4 h at 550°C and 2 h at 930°C, respectively.

Heavy metal detection

After the determination of pH, EC, and particle size distribution, samples were analyzed for heavy metals. The total contents of Fe, Mn Zn, and Co in the soil were determined by digesting about 5 g of oven-dried, grounded, sieved soil sample (taken into the flask) with 20 ml of extracting solution of 0.1N HCl. The flasks were then placed at mechanical shaker for 30 min. Subsequently, they were heated at 70°C on hot plate heating mantle for 30 min. After cooling, the mixture was filtered through Whatman Filter Paper No. 42. Then, the concentrate volume was made up to 50 ml with water. Quantification of heavy metals such as Fe, Mn, Zn, and Co was made using atomic absorption spectrophotometer (Thermo Fisher Scientific Model-AAS iCE 3000 Series).

Isolation of hydrocarbon utilizing bacterial strain

The enrichment culture technique was utilized to isolate petroleum hydrocarbon-degrading microbes. Bushnell-Haas media was used to isolate pure culture, and 1% (v/v) diesel was used as a carbon and energy source.[21] For screening, 1 g of contaminated soil samples was suspended and vortexed in 10 ml of sterile distilled water. Following that, 1 ml of solution was extracted and used as inoculum for the isolation of oil-degrading microorganisms. 100 ml of Bushnell-Haas minimal salts medium or Bushnell- Hass broth (BHM) broth medium was transferred to each flask and sterilized. Flask was kept in a rotary shaker at 150 rpm and 28°C for 7 days. After 1 week of incubation, 10 ml of sample from primary enrichment was transferred to a fresh BHM broth. Unless otherwise stated, after the second enrichment, 1 ml of medium was plated after appropriate dilution on BHM agar plate and incubated at 28°C. After a 48-h incubation period, pure colonies were isolated using the streak plate method. All isolated microorganisms were kept at 4°C for future use.

Statistical analysis

A variety of statistical assessment models, including correlation and regression model and PCA, have been used to study the physicochemical interrelationships and processes. Multivariate statistical analysis methods have the advantage of explaining complex data to gain a better understanding of the ecological status of the studied systems. A number of hydro-geochemical studies have successfully used multivariate statistical analysis. All the studies showed that multivariate statistical analysis can help to interpret the complex datasets and it is useful in verifying variations caused by natural and anthropogenic factors.

Multivariate statistics

The PCA allows similar variables to be grouped into dimensions or principal components (PCs) without distinguishing between independent and dependent variables.[22] A simple Pearson's correlation analysis was performed using Minitab software before implementing these multivariate statistics to check the correlations between soil properties and the selected soil-environmental covariates. Only PCs, or factors with eigenvalues greater than one, were considered to be contributing to the variability in soil properties in the PCA. The variability observed between soil properties and soil-environmental covariates reveals how important soil-environmental covariates are in estimating soil property distribution in a specified geographical position.[18,23] In PCA, variables with high factor loading (Eigen >1) can be used to suggest the relationship between variables under each factor.[13] The regression model between CFU and the scores of three PCs has been analyzed [Figure 2]. Hence, those parameters which are multicollinearity can be removed.

Results

Physicochemical properties of soil

A total of more than 10 soil samples were collected from each region of Rajasthan, encompassing both normal soil and petroleum-contaminated soil. Out of the 10 samples collected from each region, 9 samples were found to be petroleum-contaminated soil, while the remaining sample represented normal soil for that particular region.

Subsequently, the 9 petroleum-contaminated soil samples from each region were combined and mixed together to create a single composite sample representing petroleum-contaminated soil for that specific region of Rajasthan. This process was repeated for all regions, ensuring representative sampling.

As a result of the sampling and subsequent mixing of the soil samples, a total of 8 composite soil samples were obtained. Out of these, 4 samples were representative of petroleum-contaminated soil, and the other 4 samples represented normal soil from various regions of Rajasthan.

Comparative studies were made between soil samples taken from control site and petroleum-contaminated sites. The physicochemical properties of all the soil samples were summarized and are given in Table 1. Soil texture is divided into three major groups: sand (large particle), silt (intermediate), and clay (fine).

Bulk density

All the soil samples were reported to have a bulk density of around 1.00–1.70 g/ml. Porosity also governs the level of water retention and aeration in the soil, and the porosity in the polluted soil ranges from 60.32% to 65.15% and 68.03% in soil samples.[24]

pH

An appropriate range of pH and EC is responsible for the growth and activity of microorganisms.[25] The pH of soil indicates the concentration of hydrogen ions (H+) present in the soil. A lower pH value of the soil indicates a higher range of hydrogen ion concentration.[26] The pH ranges from 7.00 to 9.4 in all the normal as well as contaminated soil samples. The ideal pH range for optimum plant growth is suggested to be between 6.8 and 8.0. Within this range, the soil is neither too acidic nor too alkaline, approaching a neutral pH.

Heavy metal detection

Heavy metals such as Fe, Mn, Zn, and Co were evaluated in the present study. These metals have a tendency to decrease the crop yield due to biomagnifications and bioaccumulation in the food chain. Table 1 illustrates that petroleum-contaminated soil samples showed a higher load of these heavy metals in comparison to control site. Fe and Mn amounts were found high in all the soil samples (>2 g), respectively, and Pb, Cr, and Ni were not determined in any of the soil samples. The level of Zn was found higher in Sambhar-contaminated soil sample and the level of Fe was found high in Jaisalmer-contaminated soil samples. Heavy metals were always found to be a risk for public concerns.[27] They are not even required in small amounts for living organisms.[28]

Cell counts of microorganisms in soil samples

One of the objectives of the present investigation is to isolate microorganisms from petroleum-contaminated soil as well as normal soil samples and count the microbial populations in each sample. A total number of viable cells were estimated in each sample. Significant numbers of microbial population up to 10^{10} CFUs have been found. From the results in Table 2, a significant increase in hydrocarbon-degrading microorganisms was found in each soil sample. In Jaisalmer and Ganganagar samples, it was found highest after the second enrichment (3.74×10^9 and 4.6×10^9), and for the Banswara soil sample and Sambhar samples, it was after third enrichment (2.82×10^8 and 2.82×10^8). The normal does not show any significant increase in CFUs [Table 5]. Bacterial populations have the tendency to mineralize hydrocarbons present in petroleum-contaminated soil from various oil-contaminated sites.[29] Bushnell-Hass medium was proven to be a suitable medium for the isolation of hydrocarbon-degrading

Table 1: Physical properties of different soil samples

	Parameters	Unit	GPCS	GNS	SPCS	SNS	JPCS	JNS	BPCS	BNS
X1	Conductivity (dS/m)	dS/m	5.90±0.6	1.14±0.23	4.90±0.17	3.85±0.2	2.40±0.57	3.80±0.6	2.18±0.25	0.34±0.5
X2	Carbon estimation (%)	None	0.51±0.3	0.48±0.2	0.37±0.68	0.42±0.75	0.38±0.76	0.32±0.12	0.24±0.3	0.27±0.5
X3	Phosphate	kg/ha	41±0.2	24±0.41	25±0.55	28±1.0	28±0.8	32±0.76	50±0.23	36±0.45
X4	Potassium	kg/ha	258±0.56	245±0.66	297±1.0	310±1.5	320±1.2	260±1.3	298	246
X5	Zinc	ppm	0.41	0.42	5.75	4.32	0.48	0.52	0.62	0.48
X6	Copper	ppm	0.41	0.28	0.31	0.22	0.36	0.48	5.1	4.3
X7	Iron	ppm	2.88	2.72	2.99	3.22	4.16	4.2	0.32	0.24
X8	Manganese	ppm	2.95	2.9	2.91	2.82	2.88	2.8	2.98	2.92
X9	Moisture content	%	29.55±1.0	30.55±1.2	25.4±0.9	28.3±0.8	12.08±1.2	13.47±1.0	16.37±0.7	15.25±0.2
X10	Water-holding capacity	%	41.69±1.0	42.5±1.3	39.8±0.5	38.0±0.72	55.5±0.85	56.5±1.0	48.43±0.5	49.3±0.3
X11	Bulk density	g/mL	1.28±0.3	1.30±0.4	1.06±0.5	1.10±0.56	1.09±0.7	1.41±0.8	1.30±0.6	1.63±0.2
X12	pH	None	7.0±0.5	7.5±0.25	8.2±0.4	9.4±0.5	8.2±0.5	8.9±0.5	7.62±0.25	7.65±0.5

GPCS: Ganganagar petroleum contaminated soil, GNS: Ganganagar normal soil, SPCS: Sambhar petroleum contaminated soil, SNS: Sambhar normal soil, JPCS: Jaisalmer petroleum contaminated soil, JNS: Jaisalmer normal soil, BPCS: Banswara petroleum contaminated soil, BNS: Banswara normal soil

Table 2: Cell counts in each soil sample

	1st enrichment	2nd enrichment	3rd enrichment	4th enrichment	Average CFU
Sample 1 (Jaisalmer)	2.51×10^8	2.75×10^{10}	3.74×10^9	2.8×10^7	2.95×10^9
Sample 2 (Ganganagar)	2.1×10^9	3.76×10^9	4.6×10^9	3.73×10^6	3.45×10^9
Sample 3 (Sambhar)	1.7×10^8	1.36×10^{10}	4.39×10^8	3.79×10^8	2.81×10^9
Sample 4 (Banswara)	1.395×10^{10}	3.35×10^7	4.52×10^7	2.82×10^8	3.02×10^8

CFU: Colony-forming unit

Table 3: Descriptive statistics for different physiochemical parameters of soil

Variables	Mean	Median	Skewness	Variance	CV
X1 conductivity (dS/m)	3.845	3.65	0.15	2.719	42.89
X2 carbon estimation (%)	0.375	0.375	0	0.00973	26.31
X3 phosphate	36	34.5	0.28	108.27	28.9
X4 potassium	293	297.5	−0.61	531.93	7.86
X5 zinc	1.815	0.55	1.27	5.512	129.35
X6 copper	1.545	0.385	1.28	4.495	137.22
X7 iron	2.588	2.935	−0.77	2.097	55.96
X8 manganese	2.93	2.93	0	0.00155	1.34
X9 moisture content	20.85	20.89	−0.01	34.44	34.44
X10 water-holding capacity	47.11	46.56	0.12	14.37	14.37
X11 bulk density	1.1825	1.185	−0.02	0.0125	9.45
X12 pH	7.755	7.91	−0.58	0.262	6.61

CV: Coefficient of variation

Table 4: Correlation matrix between different parameters of soil

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1											
X2	0.80	1										
X3	−0.20	−0.31	1									
X4	−0.80	−0.61	−0.40	1								
X5	0.35	−0.062	−0.61	0.11	1							
X6	−0.59	−0.80	0.81	0.11	−0.31	1						
X7	0.28	0.65	−0.83	0.17	0.13	−0.93	1					
X8	0.03700899	−0.2886615	0.9253613	−0.55	−0.28	0.76	−0.90	1				
X9	0.96	0.64	0.003	−0.87	0.35	−0.36	0.019	0.280	1			
X10	−0.91	−0.49	0.17	0.69	−0.62	0.38	−0.030	−0.16	−0.95	1		
X11	0.013	−0.06	0.96	−0.60	−0.64	0.63	−0.71	0.90	0.19	0.034	1	
X12	−0.47	−0.44	−0.70	0.89	0.52	−0.17	0.34	−0.70	−0.56	0.30	−0.86	1

Table 5: Correlation matrix between the colony-forming unit and the physical properties of soil for all the four sample locations

	X (G)	X (S)	X (J)	X (B)
Y (G)	0.145			
Y (S)		0.8		
Y (J)			0.953	
Y (B)				0.03

microorganisms.[30] In the present investigation, a high number of bacterial colonies were observed control plate where 4T engine oil was not added as a carbon source.

Since the skewness coefficient is almost Zero for all that mention parameters so there distribution is symmetrical and since the value of Coefficient of variation is indicating the scaredness in the observed data set which is highest for Zinc and Copper.

From Table 4, we can observe a positive and high correlation between conductivity (X1) and moisture content (X9) and a negative and high correlation between water-holding capacity (X10) and moisture content (X9).

Principal component analysis

Principal component analysis: X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X11, and X12 (Table 8)

In these results, the first three PCs have eigenvalues >1 as shown in Tables 6 and 7. These three components explain 100% of the variation in the data. The screen plot shows that the eigenvalues start to form a straight line after the third PC. 100% is an adequate amount of variation explained in the data, so we should use the first three PCs.

Scores are linear combinations of the data that are determined by the coefficients for each PC. To obtain the score of a principal component, substitute observed values of all the parameters (X1, X2,……….X12.) in the following linear equation.

$$PC1 = 0.064X1 + 0.134X2 - 0.429X3 + 0.192X4 + 0.234X5 - 0.35X6 + 0.373X7 - 0.412X8 - 0.03X9 - 0.041X10 - 0.416X11 + 0.308X12.$$

Regression analysis: Y (colony-forming unit) versus PC1, PC2, and PC3

The regression equation is as per in Table 9:

$$Y (CFU) = -4.92 \times 10^9 + 2.03 \times 10^8 PC1 + 18196658 PC2 + 3.15 \times 10^8 PC3.$$

Table 6: Eigenanalysis of the correlation matrix

Principal Components	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Eigenvalue	5.3866	4.8983	1.7151	0.0000	0.0000	0.0000	0.0000	−0.0000
Proportion	0.449	0.408	0.143	0.0000	0.0000	0.0000	0.0000	−0.0000
Cumulative	0.449	0.857	1.000	1.000	1.000	1.000	1.000	1.000

Table 7: Eigenanalysis of the correlation matrix

Principal Components	PC9	PC10	PC11	PC12
Eigenvalue	−0.0000	−0.0000	−0.0000	−0.0000
Proportion	−0.000	−0.000	−0.000	−0.000
Cumulative	1.000	1.000	1.000	1.000

Table 8: Principal component analysis of different parameters

Variable	PC1	PC2	PC3	PC4	PC5	PC6
X1	0.064	−0.445	−0.062	0.382	0.223	0.012
X2	0.134	−0.365	0.381	−0.301	0.211	−0.171
X3	−0.429	0.023	0.059	0.256	−0.170	−0.264
X4	0.192	0.403	−0.057	−0.094	0.232	−0.570
X5	0.234	−0.097	−0.620	−0.261	−0.419	−0.254
X6	−0.350	0.230	−0.219	0.253	−0.284	0.047
X7	0.373	−0.088	0.351	0.007	−0.462	−0.314
X8	−0.412	−0.073	−0.186	−0.614	0.275	0.005
X9	−0.030	−0.438	−0.180	0.265	0.034	−0.054
X10	−0.041	0.394	0.368	0.009	−0.077	0.184
X11	−0.416	−0.079	0.149	0.119	0.085	−0.603
X12	0.308	0.278	−0.254	0.311	0.513	−0.100

Table 9: Regression table of Y (colony-forming unit) versus PC1, PC2, and PC3

Predictor	Coefficient	SE	Coefficient *T*	*P*
Constant	−4919320453	24794376975	−0.20	0.846
PC1	203185539	385495633	0.53	0.608
PC2	18196658	249685776	0.07	0.943
PC3	315139996	776136743	0.41	0.692

SE: Standard error

Discussion

Sandy soil is unable to hold water due to its heavy texture and broader pore size whereas clay supports water-holding capacity due to its narrower pore size.[31] The percentage of sand was highest and clay was lowest in Control soil or Normal Soil (CS) soil which makes it loam sandy, while petroleum-contaminated soils were found loam clayey and rich in mineral content. No significant variations were observed in silt percentage. Bulk density of the different soil samples was measured and it is obvious that an increase in OM will increase the bulk density. The reason behind this is that contaminating petrochemicals will occupy the pores inside the soil without any change in volume. It can be easily observed that contaminated soil has lower porosity in comparison to normal soil samples. This may be due to hydrocarbons occupying the interstitial space and decreasing the water-holding capacity of the soil.

The pH 9.4 ± 0.5 of Sambhar lake and 8.9 ± 0.5 of Jaisalmer sample are more basic in nature and not suitable for crop production. EC is the measure of the conductivity of soil due to soluble salt concentration present in the soil suspension.[31] The source of organic carbon is manure obtained from plants and animals, crop residue, and organic fertilizers. The values of carbon estimation were also estimated in all the soil samples and as predicted more carbon percent were observed in petroleum-contaminated soil. Furthermore, the presence of carbonates and bicarbonates depends on the climatic conditions.[31] The role of potassium is to help in maintaining photosynthesis rate by involving in plants' various metabolic reactions.[32,33] As the values are in ppm level and change in value is not significant, it was assumed that contamination has a minimum role on potassium concentration. The presence of potassium in nominal amount in all the samples signifies sustained osmotic pressure in the plant[34] which also indicates better water absorptive capacity. Multiple regression modal constitutes an accurate tool to evaluate soil quality, because it provides a set of minimum data of indicators[34] that seem to be able to reflect equilibrium among its physical, chemical, and biochemical properties. In consequence, different models were needed for each group. Moreover, the equations obtained with Multiple Linear Regression (MLR) are sensitive to changes in soil quality, because they have been able to reflect deviations respecting the relationships established in the reference soils when disturbed soils have been applied. From the present investigation, we conclude that zinc (X5) and copper (X6) have the maximum variability in the soil for the studied sample locations and almost all the parameters have correlated with each other. By the Tables 3 and 4 as the value of correlation coefficient between conductivity (X1) and moisture content (X9) is 0.96 and between water-holding capacity (X10) and moisture content (X9) is -0.95 therefore we can observed in the variables conductivity (X1) and moisture content (X9) is highly positively associated and water-holding capacity (X10) and moisture content (X9) is highly negatively associated.

Conclusion

It has been concluded that a positive correlation between the CFU and the physical properties of soil for all the four sample locations is also observed. PCA based on the correlation matrix is an effective approach in discriminating the variables. In our study, the results of PCA show that the first three components are able to describe 100% of the total variability in soil properties. This approach would promote soil sampling and variable rate application of agricultural chemicals and would serve as a fast, inexpensive, and reasonably accurate method to develop a soil database for fields that have similar soil characteristics.

Financial support and sponsorship

Nil.

Ethics code

There is no need for ethical clearance. There is no such permission required for this study. JECRC University, Jaipur, Rajasthan.

Conflicts of interest

There are no conflicts of interest.

References

1. Singh KP, Malik A, Mohan D, Sinha S. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) – A case study. Water Res 2004;38:3980-92.
2. Strickland MS, Osburn E, Lauber C, Fierer N, Bradford MA. Litter quality is in the eye of the beholder: initial decomposition rates as a function of inoculum characteristics. Funct Ecol 2009;23:627-36.
3. Bodelier PL. Toward understanding, managing, and protecting microbial ecosystems. Front Microbiol 2011;2:80.
4. Philippot L, Spor A, Hénault C, Bru D, Bizouard F, Jones CM, *et al.* Loss in microbial diversity affects nitrogen cycling in soil. ISME J 2013;7:1609-19.
5. Bardgett RD, van der Putten WH. Belowground biodiversity and ecosystem functioning. Nature 2014;515:505-11.
6. Scanlon BR, Mukherjee A, Gates J, Reedy RC, Sinha AK. Recarga de água subterrânea em sistemas de dunas naturais e em ecossistemas agrícolas da região do Deserto de Thar, Rajastão, Índia. Hydrogeol J 2010;18:959-72.
7. Kumar KA, Lalitha M, Shivanand, Sujatha K, Nair KM, Hegde R, Singh SK, Mishra BB. Major soil types and classification. The Soils of India. 2020:81-108.
8. Bishop T, McBratney A. A comparison of prediction methods for the creation of field-extent soil property maps. Geoderma 2001;103:149-60.
9. Afu S, Isong I, Awaogu C. Agricultural potentials of floodplain soils with contrasting parent material in Cross River State, Nigeria. Glob J Pure Appl Sci 2019;25:13-22.
10. Odeh IO, Crawford M, McBratney AB. Digital mapping of soil attributes for regional and catchment modelling, using ancillary covariates, statistical and geostatistical techniques. Dev Soil Sci 2006;31:437-622.
11. Beguería S, Spanu V, Navas A, Machín J, Angulo-Martínez M. Modeling the spatial distribution of soil properties by generalized least squares regression: Toward a general theory of spatial variates. J Soil Water Conserv 2013;68:172-84.
12. Carvalho Junior WD, Chagas CD, Lagacherie P, Calderano Filho B, Bhering SB. Evaluation of statistical and geostatistical models of digital soil properties mapping in tropical mountain regions. Rev Bras Ciênc Solo 2014;38:706-17.
13. Zeraatpisheh M, Ayoubi S, Sulieman M, Rodrigo-Comino J. Determining the spatial distribution of soil properties using the environmental covariates and multivariate statistical analysis: a case study in semi-arid regions of Iran. J Arid Land 2019;11:551-66.
14. Gribov A, Krivoruchko K. Empirical Bayesian Kriging implementation and usage. Sci Total Environ 2020;722:137290.
15. Nourzadeh M, Mahdian MH, Malakouti MJ, Khavazi K. Investigation and prediction spatial variability in chemical properties of agricultural soil using geostatistics. Arch Agron Soil Sci 2012;58:461-75.
16. Erşahin S, Karaman R. Use of factor analysis in the assessment of soil variability for site specific management and soil fertility studies. Tarım Bilimleri Derg 2000;6:76-81.
17. Shange RS, Ankumah RO, Githinji L, Zabawa R. Spatial assessment of selected soil properties within an industrial poultry production site. Air Soil Water Res 2012;5:ASWR-S9268.
18. Shukla M, Lal R, Ebinger M. Determining soil quality indicators by factor analysis. Soil Tillage Res 2006;87:194-204.
19. Khan A. A study on mapping out the best practices under different irrigation source for sustainable agriculture in Sajjangarh block of Banswra, Rajasthan. Int J All Res Writ 2020;1:75-83.
20. Sangamesh MB, Jambagi S, Vasanthakumari MM, Shetty NJ, Kolte H, Ravikanth G, *et al*. Thermotolerance of fungal endophytes isolated from plants adapted to the Thar Desert, India. Symbiosis. 2018;75:135-47.
21. Bushnell LD, Haas HF. The utilization of certain hydrocarbons by microorganisms. J Bacteriol 1941;41:653-73.
22. Borůvka L, Vacek O, Jehlička J. Principal component analysis as a tool to indicate the origin of potentially toxic elements in soils. Geoderma 2005;128:289-300.
23. Malinowski ER, Howery DG. Factor Analysis in Chemistry. Vol. 3. New York: Wiley; 1980.
24. Machender G, Dhakate R, Prasanna L, Govil PK. Assessment of heavy metal contamination in soils around Balanagar industrial area, Hyderabad, India. Environ Earth Sci 2011;63:945-53.
25. Raman N, Sathiyanarayanan D. Physico-Chemical characteristics of soil and influence of cation exchange capacity of soil in and around Chennai. Rasayan J Chem 2009;2:875-85.
26. Agehara S, Warncke D. Soil moisture and temperature effects on nitrogen release from organic nitrogen sources. Soil Sci Soc Am J 2005;69:1844-55.
27. Adeniyi AA, Afolabi JA. Determination of total petroleum hydrocarbons and heavy metals in soils within the vicinity of facilities handling refined petroleum products in Lagos metropolis. Environ Int 2002;28:79-82.
28. Martinez-Tabche L, Mora BR, Faz CG, Castelan IG, Ortiz MM, Gonzalez VU, *et al*. Toxic effect of sodium dodecylbenzenesulfonate, lead, petroleum, and their mixtures on the activity of acetylcholinesterase of Moina macrocopa *in vitro*. Environmental Toxicology and Water Quality: An International Journal 1997;12:211-5.
29. Akhavan SA, Dezhban GP, Emami M, Nakhoda AM. Isolation and characterization of Crude oil degrading *Bacillus* spp; 2008.
30. Gómez-Ullate E, Bayon JR, Castro D, Coupe SJ. Efficiency of MPN method to indicate hydrocarbon biodegradation processes within permeable pavements. InProceeding of 11th International Conference on Urban Drainage, Edinburgh, Scotland, UK 2008.
31. Marbaniang I, David AA, Thomas T, Swaroop N, Hassan A. Assessment of different soil properties of Mawkynrew block, Meghalaya, India. The Pharma Innovation Journal. 2021;10:88-92.
32. Solanki HA, Chavda N. Physicochemical analysis with reference to seasonal changes in soils of victoria park reserve forest, Bhavnagar (Gujarat). Life Sci Leaf 2012;30:62, 68-2, 68.
33. Kekane SS, Chavan RP, Shinde DN, Patil CL, Sagar SS. A review on physico-chemical properties of soil. International Journal of Chemical Studies 2015;3:29-32.
34. Wilson N. Soil water and ground water sampling. CRC Press; 2020.