*Indian Journal of Ecology* (2024) 51(1): 113-117
DOI: https://doi.org/10.55362/IJE/2024/4203

**Manuscript Number: 4203**
**NAAS Rating: 5.38**

# Geo-Statistical Analysis of Soil Pollutants of Tannery Industrial Area, Kanpur

**Nisha Sahu, Divya Pipalde, R.C. Jain[1] and S.C. Gupta**

*ICAR-Indian Institute of Soil Science, Bhopal-462 038, India*
*[1]R.A.K. College of Agriculture, Sehore-466 001, India*
*E-mail: nishasahu5@gmail.com*

***Abstract***: Earth's soil is a valuable natural resource that is easily damaged by excessive agriculture, industrialization, and other industrial or environmental activities. Accurate and rapid assessment of soil properties is key component of the agriculture. Heavy metals pollution and their toxic levels in soils is one of the major problems associated with the environment. The main objective is to measure the soil properties and their spatial variability. A combination of conventional analytical methods and geostatistical methods were used to analyze the data for spatial variability. A total of 50 Soil samples were collected grid wise at an interval of 500 m with the help of Global Positioning System (GPS) from the plough layer (0-25 cm) covering the study area which is highly contaminated. A classical ordinary kriging (OK) interpolation method with three types Spherical, Gaussian and Exponential were used in ArcGIS software. The findings showed that Cd varied from 3.6 to 24, Cr from 44.6 to 4596, EC from 0.41 to 1.230, OC from 0.2 to 1.54 and pH from 7 to 8.3. There was a strong spatial dependence for chromium, moderate for cadmium and weak for pH, EC and OC. Semivariograms showed spherical model best for cadmium and chromium, gaussian for OC and exponential model best fitted for pH and EC.

***Keywords***: Soil properties, Semivariogram, Spatial distribution, GIS, Ordinary Kriging

The presence of soil heavy metals is a severe threat to the environment, crop production, and human health. The entry of poisonous and toxic metals in human food chain by the vegetables crops is of great concern. Higher concentrations of heavy metals in field are often not caused by pedogenesis but by human activities, such as mining and smelting, combustion of fossil fuels, application of agrochemicals, and wastewater irrigation. It is crucial to ascertain and quantify the sources of heavy metals in farmland for designing pollution controls and remediation strategies. Spatial distribution and source identification of pollutants in soils are essential for risk evaluation and soil reclamation. With the development of geographic information system (GIS) and geostatistics, Kriging is a prevalent technique used in spatial interpolation on soil contamination. Multivariate analyses, including principal component analysis (PCA), cluster analysis (CA), enrichment factor (EF) and positive matrix factorization (PMF), are valuable tools for identifying sources of heavy metals in soil. Because of its environmental significance, studies to determine risk caused by metal levels in soil on human health and forest ecosystem have attracted attention in recent years.

Soil heavy metal pollution studies focus on the identification of high pollution risk areas. Samples from high pollution risk areas are usually local spatial outliers (Zhang et al 2009). Interpolation techniques all have a smoothing effect, which underestimates the local high values and overestimates the local low values (Journel et al 2000). This smoothing effect leads to bias in soil pollution assessment and has an effect on relevant environmental decision making (Goovaerts 2000). Consequently, there is need to monitor and evaluate the amount and distribution of soil heavy metals (Choe et al 2008). Mapping the spatial distribution of heavy metals in soils is critical for risk assessment of potential environmental pollution and for establishing protocols for pollution remediation. The use of advanced technology Geographical Information Systems (GIS) is one of the most efficient tool and valid alternative for studying soil heavy metals. This method allows repetitive coverage of large-scale areas in a relatively cost-effective way, have become attractive for identifying and assessing spatial patterns of the soil properties. There are a lot of studies of the performance of the spatial interpolation methods mentioned above, but the results are not clear-cut (Shi et al 2009). Some of them found that the kriging method performed better than IDW (Yasrebi et al 2009); while others showed that kriging was no better than alternative methods. Conventional statistical method and geographical information system (GIS) are powerful tools for predicting the spatial distribution of soil HMs at a regional scale (Yang et al 2020). Kriging is one of several methods that use a limited set of sampled data points to estimate the value of a variable over a continuous spatial field. It differs from simpler methods, such as Inverse Distance Weighted Interpolation, Linear Regression, or Gaussian decays in that it uses the spatial correlation between sampled points to interpolate the values in the spatial field. Kriging also generates estimates of the uncertainty surrounding each interpolated value. This helps to reduce bias in the predictions. Ordinary kriging, for which the assumption of stationary (that the mean and variance of the values is constant across the spatial field) must be assumed. The objective of the study was to prepare spatial variability distribution maps of soil properties of tannery industrial zone which is highly contaminated with long term irrigation of tannery effluents in the nearby zones.

## MATERIAL AND METHODS

**Study area:** Jajmau Industrial Zone, Kanpur, India which is lies between 26.46$^0$ North latitude to 80.35$^0$ East longitude. It is situated on the banks of the Ganges River. The main industry is the leather industry. It is home to some of the biggest leather tanneries in Northern India. It is highly chronic polluted area and one of the biggest exporting centres of the tanned leather. The study area is situated in the zone of humid subtropical climate and the year is divided into three seasons with heavy rainfall during the monsoon season in the months of July, August and September about 70-80 % of the total rainfall.

**Soil sample collection:** Soil samples were collected grid wise at an interval of 500 m with the help of Global Positioning System (GPS). A total 50 soil samples are collected from the plough layer (0-25 cm) covering the study area which is highly contaminated.

**Soil analysis:** Soil sample collected are air dried under shade and then grinding is done using porcelain made mortar and pestle than the sample sieved through a 2.0 mm size sieve for pH, EC, Cd and Cr. For organic carbon 0.5 mm size sieve is used. The soil that passed through sieves is collected and stored in air-tight polythene bags for further use. Soil pH, EC, OC, Cd and Cr were determined using the standard analytical methods.

**Geostatistical analysis:** In general, geostatistical methods were used to estimate and map the soil properties. It is based on the theory of recognized variables, which was used to investigate the soil spatial variability. It is expressed by a Semivariogram, which measures the average dissimilarity between data separated by a vector h it is computed as half the average squared difference between the components of data pairs:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [\, Z(x) - Z(x_i + \mathrm{h})] \qquad (1)$$

Where, N(h) is the number of data pairs within a given class of distance and direction, z(xi) is the value of the variable at the location xi and z(xi+h)is the value of the variable at a lag of h from the location xi .

Experimental semivariogram value for each property was computed using ArcGIS 10.3. During pair calculation, maximum lag distance was taken half of the minimum extent of sampling area to minimize the border effect. Using the semivariogram model, basic spatial parameters such as nugget (C0), partial sill (C+ C0) and range (m) was calculated. Nugget is the variance at zero distance, partial sill is the lag distance between measurements at which one value for a variable does not influence neighbouring values and range is the distance at which values of one variable become spatially independent of another. Three commonly used semivariogram models were fitted for soil properties (pH, EC, OC, Cd and Cr). These are the Spherical, Exponential and Gaussian model. Expressions for different semivariogram models are below:

Spherical model

$$\gamma(h) = C_0 + C\left[1.5\tfrac{h}{a} - 0.5\left(\tfrac{h}{a}\right)^3\right] \quad \text{if } 0 \leq \mathrm{h} \leq \mathrm{a}, (2)$$

Exponential model

$$\gamma(h) = C_0 + C\left[1 - exp\left\{\tfrac{h}{a}\right\}\right] \qquad \text{for h} \geq 0 \ (3)$$

Gaussian model

$$\gamma(h) = C_0 + C\left[1 - exp\left\{\tfrac{-h}{a}\right\}^2\right] \qquad \text{for h} \geq 0 \ (4)$$

In all these models, nugget, sill and range were expressed by Co, (C+ Co) and m, respectively. From spatial data on soil properties corresponding point feature file was prepared in ArcGIS. ArcGIS geo-statistical analyst extension was used to carry out exploratory variogram analysis and then extend this exploratory approach to spatial interpolation by way of kriging. Geo-statistical analysis consisting of variogram calculation, kriging, cross-validation and mapping was performed using the geo-statistical analyst extension of ArcGIS 10.3. All the statistical calculations were performed using SPSS statistics 17.0.

## RESULTS AND DISCUSSION

A descriptive statistic of Cd and Cr concentrations and soil properties (EC, OC and pH) value of plough layer (0-25 cm) soil of 50 soil samples is listed in Table 1. The Cd varied from 3.6 to 24, Cr from 44.6 to 4596, EC from 0.41 to 1.230, OC from 0.2 to 1.54 and pH from 7 to 8.3. The kurtosis values for Cr were low, however, that for Cd were high and skewness values for Cr is negative, however for EC is high. Figure 1 presents the semivariogram and fitted model for soil properties. The attributes of the semivariograms for each soil property is summarized in Table 2.

$C_0$ is the nugget variance; C is the structural variance, and $C_0$+C represents the degree of spatial variability, which is affected by both structural and stochastic factors. Nugget variance represents the experimental error and field variation within the minimum sampling spacing. The Nugget/Sill ratio can be regarded as a criterion to classify the spatial dependence of soil properties. The higher ratio

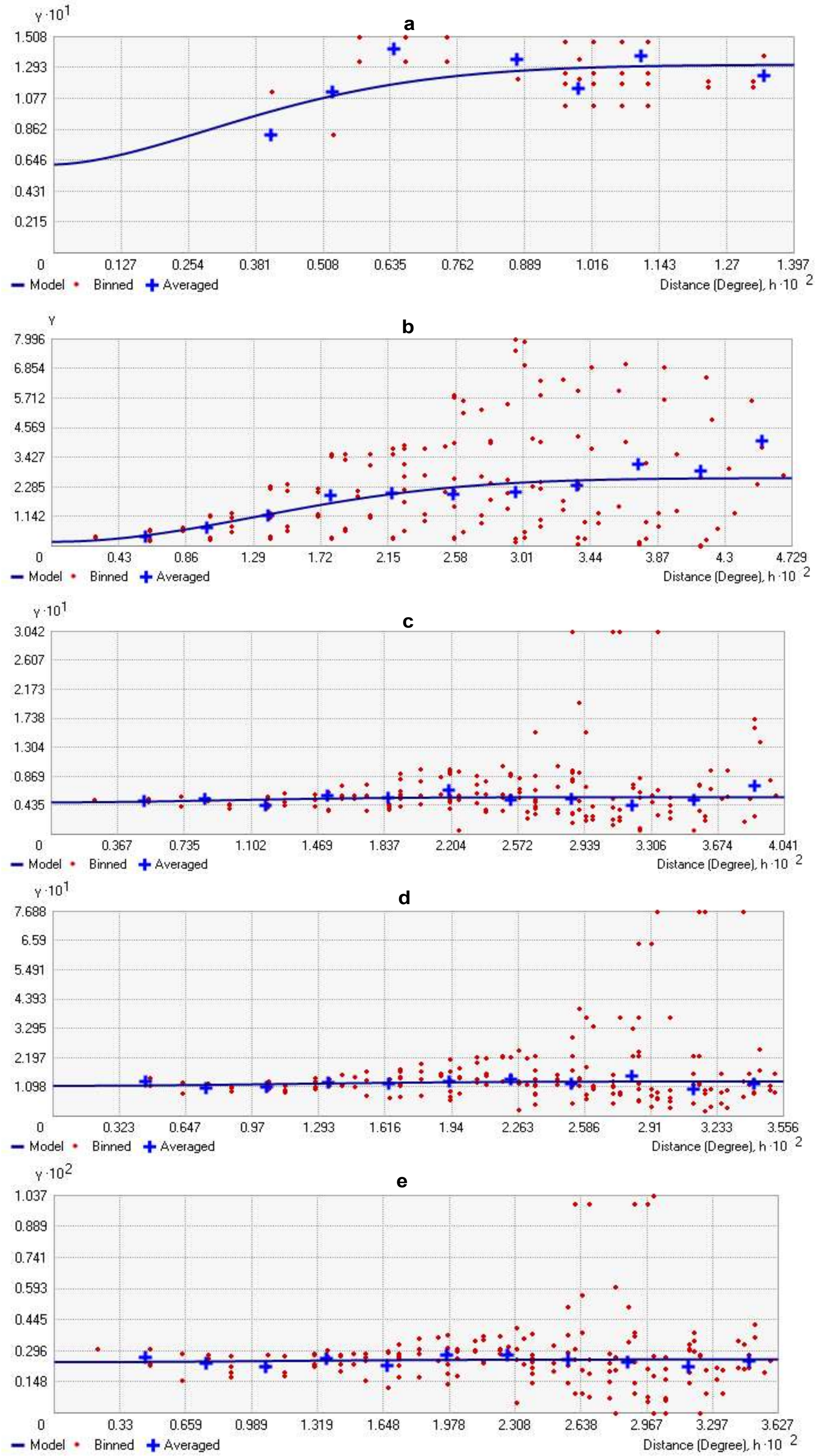


**Fig. 1.** Semivariogram parameters of best-fitted theoretical model to predict soil properties, a) Cd, b) Cr, c) EC, d) soil OC, and e) pH

indicates that the spatial variability is primarily caused by stochastic factors, such as fertilization, human activities etc.

The lower ratio suggests that structural factors, such as climate, parent material, topography and other natural factors, play a significant role in spatial variability. If the ratio is less than 25%, the variable has strong spatial dependence; between 25% and 75%, the variable has moderate spatial dependence; and greater than 75%, the variable shows only weak spatial dependence (Chien et al 1997, Chang et al 1998). Therefore, Spatial dependence of Cr is strong, Cd is moderate and for EC, OC and pH are weak. These results indicate that the theoretical model was an adequate representation of the spatial structural properties of soil.

Semivariograms showed that soil Cd and Cr were best fitted for spherical model, organic carbon for Gaussian and other properties of soil which are electrical conductivity and pH were best fitted for an exponential model. Figure 2 shows the spatial distribution maps of soil properties using ordinary kriging. Cadmium availability range varies with the distance from the industrial zone. Near the industrial zone Cd has concentration 16.63 ppm and away from the industrial zone its concentration decreases to 4.54 ppm. In tannery industries, the bi-products generated is rich in chromium therefore, a very high concentration is noted near industrial zone which decreases with distance.

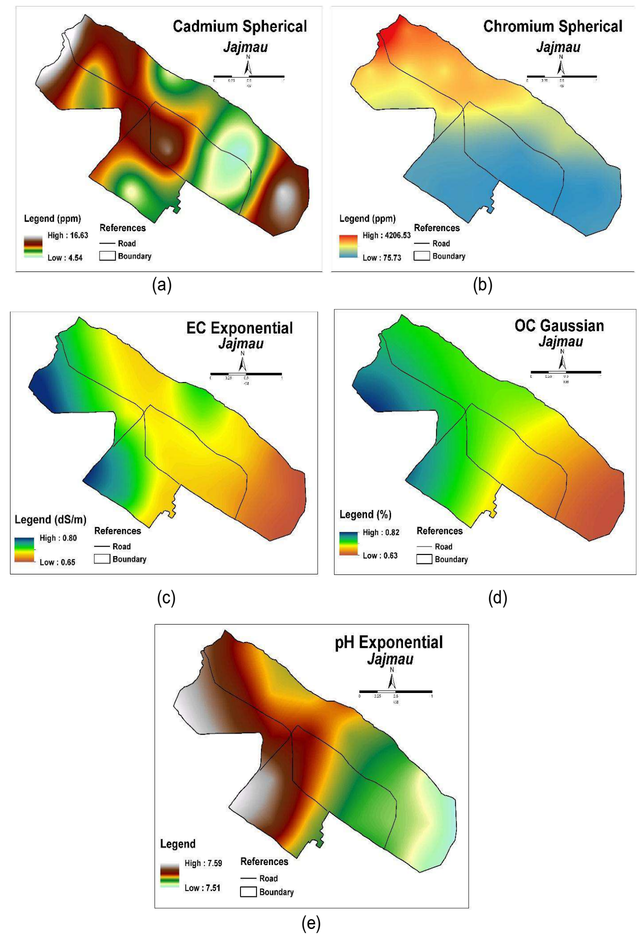


**Fig. 2.** Spatial distribution maps of soil properties, a) Cd, b) Cr, c) EC, d) soil OC, and e) pH

**Table 1.** Descriptive statistics of soil properties of industrial area

| Parameters | Distribution | Minimum | Maximum | Mean | Median | Standard deviation | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|
| Cadmium (mg/kg) | Log | 3.6 | 24 | 9.25 | 9 | 0.364 | 0.144 | 3.377 |
| Chromium (mg/kg) | Log | 44.6 | 4596 | 1066.8 | 942.2 | 1.242 | -0.336 | 1.874 |
| EC (dS/m) | None | 0.41 | 1.230 | 0.719 | 0.620 | 0.231 | 0.738 | 2.446 |
| OC (%) | None | 0.2 | 1.54 | 0.722 | 0.73 | 0.352 | 0.663 | 2.725 |
| pH (1:2.5) | Log | 7.0 | 8.3 | 7.56 | 7.5 | 0.050 | 0.353 | 1.902 |

**Table 2.** Geo-Statistical parameters of the fitted semivariogram models for soil properties

| Soil properties | Semivariogram model | Range (m) | Nugget ($C_0$) | Partial sill (C) | $C_0$+C | NS ratio | Spatial dependence |
|---|---|---|---|---|---|---|---|
| Cadmium (mg/kg) | Spherical | 0.0091 | 0.0378 | 0.0922 | 0.13 | 0.290 | Moderate |
| Chromium (mg/kg) | Spherical | 0.0472 | 0.00 | 2.8431 | 2.8431 | 0.00 | Strong |
| EC (dS/m) | Exponential | 0.3459 | 0.0457 | 0.0108 | 0.0565 | 0.808 | Weak |
| OC (%) | Gaussian | 0.0264 | 0.1144 | 0.0165 | 0.1309 | 0.87 | Weak |
| pH (1:2.5) | Exponential | 0.0362 | 0.0024 | 0.0001 | 0.0025 | 0.96 | Weak |

## CONCLUSION

The spatial variability of soil properties in tannery industrial zone is very important for environmental monitoring and planning remediation measures in this region. The spherical model was best fitted for cadmium and chromium, exponential for EC and pH and gaussian for soil organic carbon for prediction. Spatial variability maps of soil properties were prepared with best models using ArcGIS software. Future studies is required to analysis the environmental and human risk due to presence of high concentration of soil heavy metals.

## REFERENCES

Chang YH, Scrimshaw MD, Emmerson RHC and Lester JN 1998. Geostatistical analysis of sampling uncertainty at the Tollesbury Managed Retreat site in Blackwater Estuary, Essex, UK: Kriging and cokriging approach to minimise sampling density. *Science of Total Environment* **221**: 43-57.

Choe E, vander Meer F, van Ruitenbeek F, vanderWerff H, de Smeth B and Kim KW 2008. Mapping of heavy metal pollution in stream sediments using combined geochemistry, field spectroscopy, and hyperspectral remote sensing: A case study of the Rodalquilar mining area, SE Spain. *Remote Sensing of Environment* **112**: 3222-3233.

Chien YJ, Lee DY, Guo HY and Houang KH 1997. Geostatistical analysis of soil properties of Mid-West Taiwan soils. *Soil Science* **162**: 291-298.

Goovaerts P 2000. Estimation or simulation of soil properties? An optimization problem with conflicting criteria. *Geoderma* **97**(3-4): 165-186.

Journel A, Kyriakidis P and Mao S 2000. Correcting the smoothing effect of estimators: A spectral postprocessor. *Mathematical Geosciences* **32**: 787-813.

Shi W, Liu J, Du Z, Song Y, Chen C and Yue T 2009. Surface modelling of soil pH. *Geoderma* **150**(1-2): 113-119.

Yang S, Qu Y, Ma J, Liu L, Wu H and Liu Q 2020. Comparison of the concentrations, sources, and distributions of heavy metal(loid)s in agricultural soils of two provinces in the Yangtze River Delta, China. *Environmental Pollution* **264**: 114688.

Yasrebi J, Saffari M, Fathi H, Karimian N, Moazallahi M and Gazni R 2009. Evaluation and comparison of ordinary kriging and inverse distance weighting methods for prediction of spatial variability of some soil chemical parameters. *Research Journal of Biological Sciences* **4**: 93-102.

Zhang C, Tang Y, Luo L and Xu W 2009. Outlier identification and visualization for Pb concentrations in urban soils and its implications for identification of potential contaminated land. *Environmental Pollution* **157**(11): 3083-3090.

*Received 02 September, 2023; Accepted 30 November, 2023*