

Short Communication

Determining the most optimal interpolation algorithm for groundwater data analysis

Authors:

Mohsen Hedarogle and Ebrahim Fataei

Institution:

Department of Environmental Sciences, Ardabil Branch, Islamic Azad University, Ardabil, Iran.

Corresponding author:

Ebrahim Fataei

ABSTRACT:

Contamination of heavy metals in ground water wells need to be reviewed and determination of the spatial distribution of the emissions are inevitable. For this purpose, the use of geostatistical techniques are very practical. In this study, to examine the spatial distribution of emissions of heavy metals, chromium, cadmium and lead in drinking water wells in the rural area of Meshkinshahr in high water and low water periods, different geostatistical methods were used. In this research, eight different interpolation algorithms were used in GIS and the most efficient algorithm to select the least square error function and maps related to the spatial distribution of heavy metals were drawn. The results showed that Kriging algorithm is the most optimal interpolation algorithm for groundwater data.

Keywords:

Potable water, Heavy metals, Kriging algorithm, Meshkinshahr.

Article Citation:

Mohsen Hedarogle and Ebrahim Fataei

Determining the most optimal Interpolation algorithm for groundwater data analysis

Journal of Research in Biology (2020) 10(3): 2817-2825

Dates:

Received: 16 March 2020 Accepted: 06 April 2020 Published: 18 April 2020

Web Address:

<http://jresearchbiology.com/documents/RA0717.pdf>

This article is governed by the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which gives permission for unrestricted use, non-commercial, distribution and reproduction in all medium, provided the original work is properly cited.

INTRODUCTION

Supplying the world's growing population and food security due to limited land resources so that they have the least effect on the environment is considered as one of the most important issues. Increasing industrial activities associated with the production of pollutants including heavy metals are one of the developing and serious problem faced by human beings now. There are two sources for contamination of heavy metals in the soil: natural resources and human resources. Important human sources for entering heavy metals in soil and water include mining, manufacturing, transportation, furnace burning waste and using fertilizer and agricultural chemical pesticides. The emissions from volcanoes, degassing processes in the earth's crust, forest fires or chemical composition of parent material can be of natural resources entering in the soil (Hansen *et al.*, 2002; Hutton and Meeus, 2001; Yalcin *et al.*, 2007; Lado *et al.*, 2008).

The use of fertilizers and chemicals in agriculture has led to an increase in the concentration of heavy metals such as chromium, cobalt, cadmium, lead, copper and zinc and increased the concentration of heavy metals in these regions (Lado *et al.*, 2008). Heavy metal contamination not only directly effects on soil physical and chemical properties, reduction of biological activity and biological availability of soil nutrients, but also is a serious risk to human health by entering the food chain and environmental security by penetration into groundwater (Boisson *et al.*, 1999). Most studies on heavy metal contamination such as heavy metals environmental impact assessment starts by determining their spatial distribution (Cattle *et al.*, 2002). Because of the high variance of the environmental data many researchers face issues related to the better evaluation of these data.

In this context, the need to use advanced statistical methods is necessary. Due to the location and extent of the problems associated with samples collected to determine contaminated areas and areas exposed to pol-

lution, the use of geostatistical methods are very useful (Jiachun *et al.*, 2007). Geostatistical methods are useful tools for analyzing the continuous data in terms of time and space. Interpolated methods using geostatistics have some advantages compared to the conventional interpolation techniques such as inverse distance weighting or levels of Legendre polynomials (Journel, 1986).

Geostatistical science is one of the most accurate methods that in addition to description of the spatial and temporal data, it can make contamination quantitative data map with minimum variance (Boisson *et al.*, 1999). These methods can provide some of the spatial concentration of elements (Boisson *et al.*, 1999; Bowen, 1979). The interpolation method is widely used in mapping and estimating the concentration of heavy metals in areas that have not been sampled (Xie *et al.*, 2011). Kriging methods have been proposed for determining the spatial distribution of heavy metals, as well as the high potential interpolation method for preparing maps of pollutants (Juang *et al.*, 2001).

There are several studies in different countries for interpolation, to determine the spatial distribution of heavy metals in the soil there. Lado *et al.* (2008) performed modeling for the distribution of eight heavy metals (arsenic, cadmium, chromium, copper, mercury, nickel, lead and zinc) in the surface soil of Europe with 1588 ground reference samples and kriging - regression method. Jiachun *et al.* (2007) determined the spatial variation of Hg, Cd, Pb, Cu, As, Cr in Zhyjiang province of China by analyzing 655 soil samples. They used geostatistical techniques and applied them to the log normal kriging and ordinary kriging GIS for mapping As, Cu, Pb and Hg (Juang *et al.*, 2001).

Also Rodriguez *et al.* (2009) in their study used geostatistics to map the distribution of heavy metal concentration. Faiz *et al.* (2009) investigated the spatial variability of soil Pb and Ni around industrial poles in the center of Isfahan using geostatistical methods. The mass balance modeling of cadmium and lead in the

arable field of Esfahan was also studied. Therefore, the use of geostatistics in addition to describing the spatial and temporal variation is to provide contamination quantitative data map with minimal variance.

These methods are able to quantify the spatial distribution of elements. Mei *et al.* (2005) investigated lead, zinc, cadmium, copper and chromium in the region of Hangzing in China and used ordinary kriging and kriging normal log to provide pollution map in the region. They reported that the most important cause of the pollution in the area is human interference and industrial activities. Chen *et al.* (2008) determined the origin of heavy elements in the Hong Zhou residential land using a combination of multivariate statistics and the geostatistics between the concentrations of heavy metals that were distinguished by industrial activities, natural factors and the soil characteristics.

Fetouani *et al.* (2008) evaluated soil pollution around the busy highways in Islamabad, Pakistan, on the basis of geostatistics studies. By determining the concentration of heavy metals such as cadmium, lead, zinc, nickel and copper, they concluded that soil pollu-

tion around the highway for the elements described above was higher than normal and nearer distance had more pollution than far distance (Faiz *et al.*, 2009). Lin *et al.* (2011) in geostatistics studied about contaminated elements in different parts of Taiwan agricultural soils and showed that the use of chemical fertilizers, industrial plants and mining activities influenced the addition of these compounds in soil surface horizons to these areas.

Maps indicating areas contaminated by heavy metals or at risk of contamination can have important information in the field of selecting and finding appropriate areas for the use of land or clean it for decision makers. The aim of this study was to determine the spatial distribution of heavy metals like chromium, lead and cadmium in groundwater wells of Meshkinshahr in Ardabil using geostatistical methods and techniques of GIS and to study geological structure and land use effects on the concentration of these metals in water. The study area is a major center of Ardabil agriculture and horticulture (Figure 1).

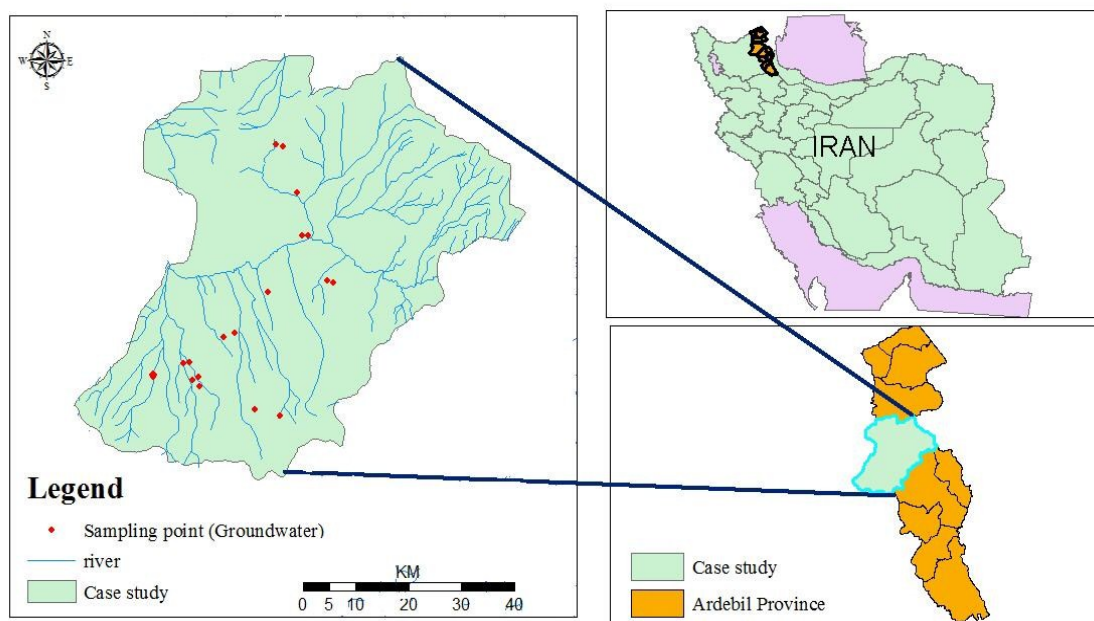


Figure 1. The study region

Table 1. The concentration of heavy metals measured during high and low water season in the drinking water network of Meshkinshahr

S. No	Lead (ppm)	Chromium (ppm)	Cadmium (ppm)	Season	Station
1	0.089	0.013	0.005	Summer	Doshanlou
2	0.369	0.52	0.15	Summer	
3	0.231	0.011	0.003	Spring	Moradlou
4	0.263	0.052	0.018	Summer	
5	0.271	0.026	0.011	Spring	Ahle Iman
6	0.186	0.034	0.023	Summer	
7	0.025	0.098	Nil	Spring	Barzil
8	0.174	0.039	0.029	Summer	
9	0.257	Nil	0.001	Spring	Sorkhanlou
10	0.137	0.055	0.014	Summer	
11	0.028	0.047	Nil	Spring	Tazekand
12	0.095	0.058	0.015	Summer	
13	0.019	0.005	0.012	Spring	Alouj
14	0.141	0.032	0.024	Summer	
15	Nil	Nil	Nil	Spring	Mouil
16	0.177	0.044	0.014	Summer	
17	0.221	Nil	0.025	Spring	Nasirabad
18	0.257	0.063	0.017	Summer	
19	0.842	0.007	Nil	Spring	Mazagha
20	0.215	0.019	0.019	Summer	
21	0.071	0.005	0.008	Spring	Ghasabeh
22	0.066	0.022	0.017	Summer	
23	0.583	0.051	Nil	Spring	Goudeh Kahriz
24	0.031	0.029	0.017	Summer	
25	0.014	Nil	0.001	Spring	Kourbalagh
26	0.056	0.022	0.021	Summer	
27	0.032	0.029	0.027	Spring	High
28	0.069	0.044	0.022	Summer	
29	0.014	Nil	Nil	Spring	Khourshidabad
30	0.143	0.048	0.024	Summer	
31	0.008	0.008	Nil	Spring	Ounar
32	0.124	0.066	0.006	Summer	
33	0.052	0.015	0.002	Spring	Kangarlou
34	0.281	0.069	0.019	Summer	
35	0.044	0.023	0.004	Spring	Meshkinshahr
36	0.077	0.033	0.007	Summer	
37	0.012	Nil	0.004	Spring	Arjeh
38	0.039	0.049	0.017	Summer	
39	0.014	0.005	Nil	Spring	Doost Beiglou
40	0.123	0.036	0.022	Summer	

Table 2. The results of the validation of different interpolation algorithms

S. No	Kriging–universal	Kriging–simple	MSE				IDW	Season	Heavy metal
			Kriging–ordinary	LPI	RBF	GPI			
1	0.0301	0.0237	0.0301	0.0281	0.0286	0.026	0.03169	Low water	Chromium
2	0.1287	0.1045	0.1287	0.1323	0.1198	0.1261	0.1437	High water	
3	0.0086	0.0078	0.0086	0.0094	0.0088	0.0089	0.009	Low water	Cadmium
4	0.0373	0.02931	0.0373	0.0376	0.0345	0.0348	0.0422	High water	
5	0.2378	0.2135	0.2378	0.2564	0.257	0.2457	0.262	Low water	Lead
6	0.1007	0.0877	0.1007	0.1072	0.1022	0.1059	0.1178	High water	

MATERIALS AND METHODS

Sampling methods

Interpolation algorithm

Recent advances in spatial statistics and replacing variable region instead of random variables, has attracted scientists to change quantitative methods in classical statistics. The development of geographic information systems in relation to spatial data has caused many scientists to use new and different interpolation methods. There are various procedures for survey and mapping changes in groundwater characteristics that each of them depend on the area/position, statistics and data. Early detection and preparing maps of quality changes and identifying its factors by traditional and observing methods, is working hard, time-consuming and costly. One of the best systems is technology, Geographic Information System (GIS) that was developed in 1960 in America and Canada.

GIS aspects include groundwater management, planning, optimized design and analysis of water distribution networks etc. In interpolation methods to produce maps of changes of nitrate in the study area, we can point to kriging geostatistical methods, inverse distance weighting, yield general, radial function and yield focal. Interpolation methods are the perfect solutions that are used in estimating non sampled locations.

Interpolation is a process to gain unknown points by the known. In this method, a surface raster (continuous) is made using the discrete sample points that the third dimension of it is sampled, and if neces-

sary can be transformed into the lines. Estimates anywhere by any interpolation methods give different results. Because according to the functions that each of these methods use, interpolation results are also different. On the other hand, all the planning is done on the basis of the estimates and if the estimate is wrong and inappropriate to design, it leads to weak and confused conclusions.

MSE cross validation

To assess the accuracy, the estimation of geostatistical model and interpolation error and determining the parameters of the study, cross validation technique was used. In this technique, for each of the measured points that usually are the only means of comparison, it could be estimated, then compared with the observed and estimated payments. Thus, in eliminated validation method, one point removed by use of other points and applying the interpolation method, estimation in this point occurs. Then, this point is returned to its point and the next point is removed. Similarly assessment is done to all parts. Estimated values are compared and estimation error is calculated using statistical and standardized root mean square error models. The lower the value, the more accurate the model.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n [z^*(x_i) - z(x_j)]^2}{n}}$$

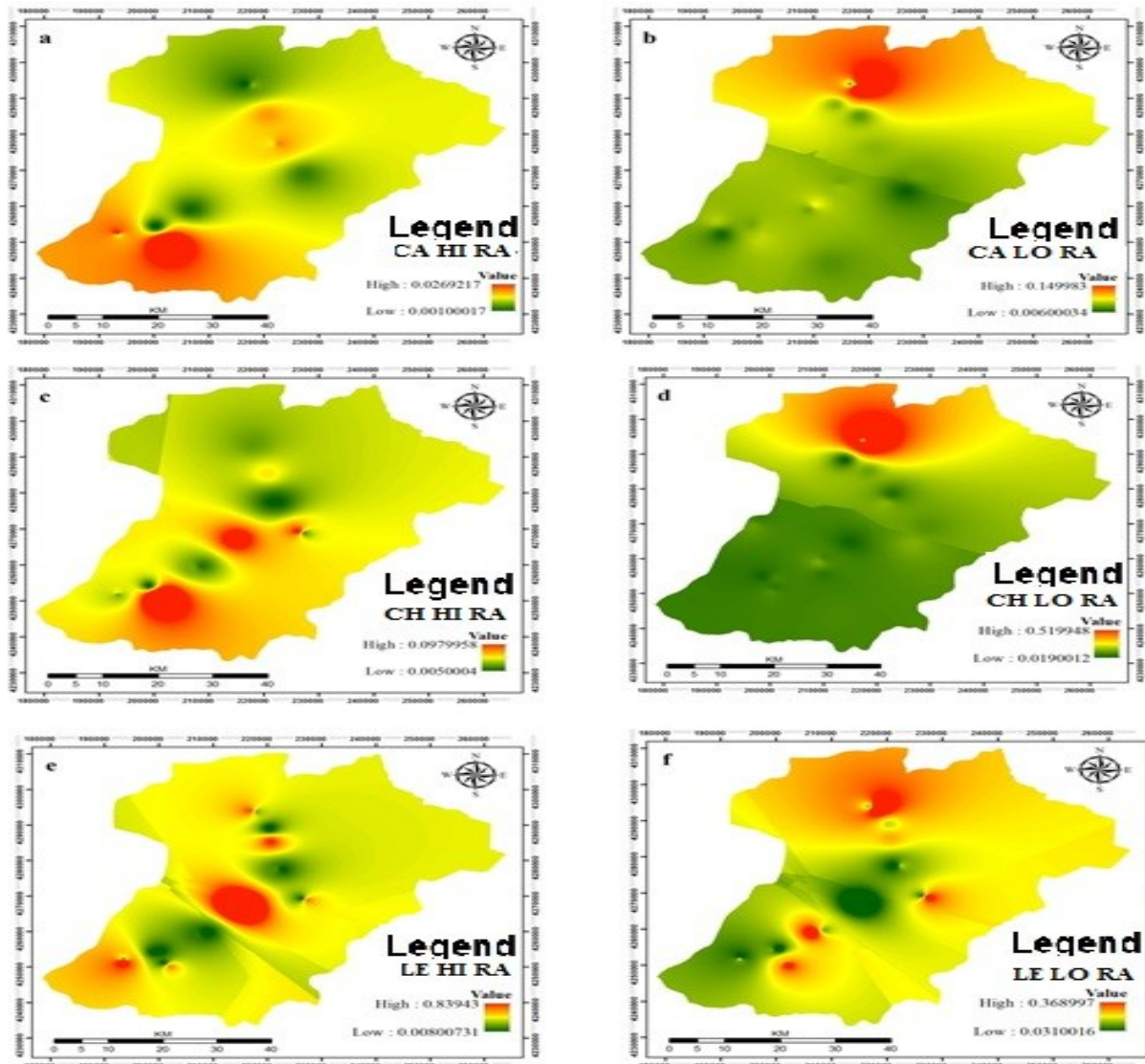


Figure 2. Map of the most optimal algorithm for the heavy metals like cadmium, lead and chromium

RESULTS AND DISCUSSION

Summarized Data

According to the aim of this study, heavy metals, including cadmium, chromium, and lead were sampled in underground water resources of Meshkinshahr in 2014 in both high and low water season. The number of samples were 20 that had a good relative distribution in the city. Table 1 shows the initial values for heavy metals.

The results of the statistical analysis

In order to determine the most optimal interpolation algorithm shown in Table 2, method seven was used. Table 2 also shows the RMSE for each algorithm.

Mapping of Results

Based on the results obtained from the analysis of optimum model interpolation values for heavy metals, like chromium, cadmium and lead in high and low water seasons, for each criterion the most optimal algorithm was interpolated and mapped in GIS system, which is shown in Figure 2.

Optimal management of water resources and improvement in their quality requires data on the location, amount and distribution of chemical agents in a given geographical area. The aim of this study was to determine the most appropriate interpolation method to analyze spatial changes in the size of the heavy metals

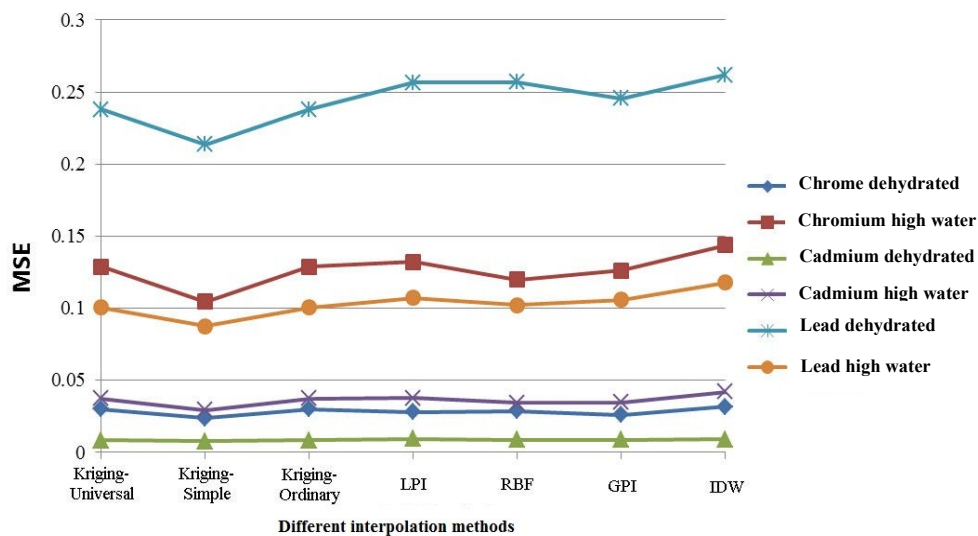


Figure 3. The values of different MSE variables in interpolation methods

like chromium, cadmium and lead in high and low water season in and Meshkinshahr ground water resources. For this purpose, methods such as ordinary kriging, simple kriging, universal kriging, inverse distance weighting, yield general, radial function, yield were used. The results showed that among the method used, simple kriging method at all levels of the heavy metals like chromium, cadmium and lead in high and low water had less subject error than other methods and detected to be the most appropriate algorithm.

The results of this study are consistent with the results of researchers such as Kresic (1997), Ahmed (2002), Gaus *et al.* (2003), Taghizadeh Mehrjardi *et al.* (2008), Barca and Passarella, (2008) and Fetouani *et al.* (2008). Researchers with geostatistical methods, such as simple and ordinary kriging, had suggested a useful tool in the study of ground water and mapping in various regions of the world that the results of this study also indicate their opinion. According to the Table 2 error values, the inverse distance weighting method, is above all methods. Among these methods, kriging methods have the fewest errors. Figure 3 shows MSE index in different variables in the interpolation methods.

As observed carefully, all methods are acceptable but simple kriging method is more accurate

that this area is a plain area, slope is from south to the north, on this basis. Figure 2 indicates that different doses of heavy metals in the region and its distribution is not a function of the slope and topography of the region and other factors such as the type of substrate, the characteristics of the power supply, ground cover and soil texture had the greatest impact. Among the human factors, human settlements also has a role by building municipal sewage plants and pollution of surface and subsurface as well as the use of chemical fertilizers and various pesticides had more decisive impact on the structure of the study area's water. Of course, this can also be stated that the increase in distance traveled by water on the way south to the northern plains, can cause changes in the amounts of heavy metals in different areas.

REFERENCES

- Ahmed S. 2002.** Groundwater monitoring network design: Application of geostatistics with a few case studies from a granitic aquifer in a semi-arid region. *Groundwater Hydrology*, 2: 37-57.
- Journel AG. 1986.** Geostatistics models and tools for the earth sciences. *Mathematical Geology*, 18: 119-140.

- Barca E and Passarella G. 2008.** Spatial evaluation of the risk of groundwater quality degradation: A comparison between disjunctive kriging and geostatistical simulation. *Journal of Environmental Monitoring and Assessment*, 137(1-3): 261-273.
- Boisson J, Ruttens A, Mench M and Vangronsveld J. 1999.** Evaluation of hydroxyapatite as a metal immobilizing soil additive for the remediation of polluted soils. Part 1. Influence of hydroxyapatite on metal exchange ability in soil, plant growth and plant metal accumulation. *Environmental Pollution*, 104(2): 225-233.
- Bowen HJM. 1979.** The environmental chemistry of elements, Academic press, London, New York. 333 P.
- Cattle JA, McBratney AB and Minasny B. 2002.** Kriging method evaluation for assessing the spatial distribution of urban soil lead contamination. *Journal of Environmental Quality*, 31(5): 1576-1588.
- Chen T, Xingmei L, Muzhi Z, Keli Z, Jianguo W, Jianming X and Panming H. 2008.** Identification of trace element source and associated risk assessment in vegetable soils of the urban-rural transitional of Hangzhou, China. *Environmental Pollution*, 151(1): 67-78.
- Faiz Y, Tufail M, Tayyeb Javed M, Chaudhry MM and Siddique N. 2009.** Road dust pollution of Cd, Cu, Ni, Pb and Zn along Islamabad Expressway, Pakistan. *Microchemical Journal*, 92(2): 186-192.
- Fetouani S, Sbaa M, Vanclooster M and Bendra B. 2008.** Assessing groundwater quality in the irrigated plain of Triffa (north-east Morocco). *Journal of Agricultural Water Management*, 95(2): 133-142.
- Gaus I, Kinniburgh DG, Talbot JC and Webster R. 2003.** Geostatistical analysis of arsenic concentration in groundwater in Bangladesh using disjunctive kriging. *Environmental Geology*, 44(8): 939-948.
- Hansen E, Lassen C, Stuer-Lauridsen F and Kjølholt J. 2002.** Heavy metals in waste. European Commission DG ENV. E3, Project ENV.E.3, ETU/200/0058, Final Report.
- Hutton M and C De Meeûs. 2001.** Analysis and conclusions from member states' assessment of the risk to health and the environment from cadmium in fertilisers. Final report European commission - Enterprise DG, Environmental Resource Management, Europe.
- Jiachun S, Haizhen W, Jianming X, Jianjun W, Xingmei L, Haiping Z and Chunlan Yu. 2007.** Spatial distribution of heavy metal in soil: A case study of changing, China. *Environmental Geology*, 52(1): 1-10.
- Juang KW, Lee DY and Ellsworth TR. 2001.** Using rank-order geostatistics for spatial interpolation of highly skewed data in heavy metal contaminated site. *Journal of Environmental Quality*, 30(3): 894-903.
- Kresic N. 1997.** Quantitative solutions in hydrogeology and groundwater modeling, CRC Press, Lewis Publishers, New York. 461 P.
- Lado LR, Hengl T and Reuter HI. 2008.** Heavy metals in European soils: A geostatistical analysis of the FOREGS geochemical database. *Geoderma*, 148(2): 189-199.
- Lin YP, Cheng BY, Chu HJ, Chang TK and Yu HL. 2011.** Assessing how heavy metal pollution and human activity are related by using logistic regression and kriging methods. *Geoderma*, 163(3-4): 275-282.
- Rodriguez L, Ruiz E, Azcárate JA and Rincon J. 2009.** Heavy metal distribution and chemical speciation in tailing and soils around a Pb-Zn mine in Spain. *Journal of Environmental Management*, 90(2) :1106-1116.
- Taghizadeh Mehrjardi R, Zareian MJ, Mahmodi Sh and Heidari A. 2008.** Spatial distribution of groundwater quality with geostatistics (Case study: Yazd-Ardakan plain). *World Applied Science Journal*, 4

(1): 9-17.

Xie Y, Chen TB, Lei M, Jun Yang, Guo Q, Song B and Zhou XY. 2011. Spatial distribution of soil heavy metal pollution estimated by different interpolation method: Accuracy and uncertainty analysis. *Chemosphere*, 82(3): 468-476.

Xing Mei L, Jianjun WU and Jiangming XU. 2005. Characterizing the risk assessment of heavy metal and sampling uncertainty analysis in paddy fields by geostatistics and GIS. *Environmental Pollution*, 141(2): 257-264.

Yalcin MG, Battaloglu R and Ilhan S. 2007. Heavy metal sources in Sultan Marsh and its neighborhood, Kayseri, Turkey. *Environmental Geology*, 53(2): 399-415.

Submit your articles online at www.jresearchbiology.com

Advantages

- **Easy online submission**
- **Complete Peer review**
- **Affordable Charges**
- **Quick processing**
- **Extensive indexing**
- **You retain your copyright**

submit@jresearchbiology.com

www.jresearchbiology.com/Submit.php