MATEMATIKA, 2013, Volume 29, Number 1b, 33-41

A Comparison between the Performance of kriging and cokriging in Spatial Estimation with Application

[1]Ghanim Mhmood Dhaher and [2]Muhammad Hisyam Lee
Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia
81310 Skudai, Johor, Malaysia
e-mail: [1]ghanim_hassod@yahoo.com [2]mhl@utm.my

Abstract This paper deals with application of spatial prediction techniques (universal kriging and cokriging) to predict unmeasured locations in mining field such as mineral ores. However, the combination of the kriging and cokriging as many predictive techniques is still an active research area to obtain an adequate prediction model. The aim is to obtain solution of spatial prediction using multivariate. We experiment primary and secondary variables of the two techniques to create a prediction model that correlated covariance functions. Practically, we apply the model on real data samples of (120) of Copper and Nickel metals that is taken from the Korf property. We are able to minimize the error rates and satisfy the weights constraints comparing with Gaussian and Power models.

Keywords kriging; cokriging; Covariance Functions; Regionalized Variables.

2010 Mathematics Subject Classification 62H11

1 Introduction

Statistics study which relies on regionalized variables is called geostatiatics. According to this term, spatial data are collected at known sites in two dimensions in the plane, or three dimensions in space. The theory of spatial statistics is based on the theory of regionalized random spatial variables, and from the real spatial data which are obtained from areas of rainfall data, disease data, air pollution, ground water, natural vegetation, mineral ores, or mining engineering. These data form a model of varying trends under some predicted assumptions of geostatistics. Mining engineer, Krige, was the first to create the prediction of gold in mines, and that was in South Africa.

This field was later developed by the French mathematical scientist Georges Matheron, [1]; it was called 'kriging' of prediction method, and it was based on the master thesis of krige, and kriging is one of the best linear unbiased estimators used to estimate value in unobserved sites in this field of study. During the application of spatial statistics, various information of primary and secondary data going to be predicted. The difference between any two points of observations by the vector (h) is called the variogram function; this function is used as a term to describe the model of spatial continuity distribution. Where the variogram function, covariance, and correlogram are the most common model of spatial modeling; the variogram function are also used [2, 3, 4, 5, 6, 7]. Other researchers used kriging techniques [5, 8, 9, 10, 11, 12].

The methods of interpolation such as kriging and cokriging are used and provided good results [13, 14, 15]. These methods are applied with a real data in a soil, showing that co- kriging is used as secondary variables, many studies also deal with spatial prediction and parameters of forecasting of multivariate [13, 16, 17]. cokriging takes many benefits of spatial multivariate, to evaluate and estimate value unmeasured locations [6, 7, 18, 19]. Study of spatial relationships and theory of geology application give important to solve problems in spatial prediction in mining field, environmental sciences, meteorology, structural engineering and hydrology. In this work, real data with locations in south western British Columbia were applied [20].

2 Material and Methods

In this research, geostatistics techniques such as universal kriging and co-kriging have been applied in mining field using programming in Matlab language. The performance techniques are evaluated for four types of mineral ores in the study field.

2.1 Regionalized Variables

Any variable distributed in the place or in the space called a spatially variable. This variable describes certain phenomenon (such as degree of metal depth). Where spatial variable differs from normal variable because any spatial variable contains value of observation and coordinates represent the location on the surface or underground of earth or outside (in the space such as: Gas pollution). Let $z(s)$ Represent the spatial variable, and (s) is location in Euclid space where $s \in D \subseteq R^p$, p=1, 2, 3.

2.2 Variogram Function

Variogram function is used to describe the mathematical models, which are key factors in the process of spatial prediction. It is defined as follows:

$$2\gamma(h) = \frac{1}{n(h)} \sum_{i=1}^{n(h)} \left[z(s_i + h) - z(s_i) \right]^2 \tag{1}$$

where $2\gamma(h)$ is variogram function, $z(s_i + h), z(s_i)$ are values of observations separated by distance h with locations (s_i) , $(s_i + h)$ and $n(h)$ are the total number of pairs of observations. Whenever the lag h between the observations is increasing then the covariance became large until it is stabled i.e. when $(h = a)$ this called the range. After that the covariance fades which means that the phenomenon is very small.

2.3 Second Order Stationarity

Random function is called Second order stationarity if and only if

$$E\left(z\left(s\right)\right) = \mu, c\left(h\right) = \operatorname{cov}\left[z\left(s+h\right), z\left(h\right)\right] \tag{2}$$

defined for all values of s, h, where covariance function is:

$$c\left(h\right) = \operatorname{cov}\left[z\left(s+h\right), z\left(h\right)\right] \text{ and } \operatorname{var}\left[z\left(s\right)\right] = c\left(0\right) = \sigma^2$$

The relationship between the variogram function and the covariance function for second order stationary and isotropy variable can be written as [4]:

$$\gamma(h) = c(0) - c(h)$$

where $c(0)$ variance and covariance $c(h)$ is positive, while the cross-variogram is defined as:

$$2\gamma_{ab}(h)=\frac{1}{n(h)}\sum_{i=1}^{n(h)}\left[z_a(s_i)-z_a(s_i')\right]\left[z_b(s_i)-z_b(s_i')\right] \tag{3}$$

The cross-variogram can be negative, while the variogram function must always be positive [5].

2.4 kriging

Kriging technique is the supposed spatial association between the points known as the spatial variability in the Structural Created estimated values. Usually, universal kriging includes the equation of the trend surface in the prediction process.

$$m(s)=\sum_{p=1}^{L}\left[\alpha_P f_p(s)\right] \tag{4}$$

α_P is the P^{th} coefficient and f_p is the P^{th} basic function, that describes the trend of the model $z(s)=\mu(s)+e(s)$, then the predictor of Z is defined as:

$$z^*(s_o)=\sum_{i=1}^{n}\left[\alpha_i z(s_i)\right] \tag{5}$$

with the constraint;

$$\sum_{i=1}^{n}\alpha_i=1$$

where this constraint ensures the principle of unbiasedness;

$$E\left[z^*(s_o)\right]=\mu(s)$$

then the model Z(s) becomes:

$$z(s)=F'(S)\alpha+e(s)$$

kriging is found by minimizing the variance [2]:

$$\sigma_{UK}^2=\sum_{I=1}^{n}\alpha_i\gamma_{io}+\sum_{i=1}^{L}\phi_p f(s_o) \tag{6}$$

2.5 cokriging

Both techniques of universal kriging and cokriging required unbiasedness or minimum error. And by using Lagrange multipliers the system is solved. In different researches (e.g. [2], [10] and [19]), cokriging technique is used, which is based on the primary and secondary variables. For a primary variable P and a single secondary variable S, the following covariance matrix must be positive definite. cokriging technique requires two conditions which are:

$$\sum_{i=1}^{n} \alpha_i = 1, \sum_{j=1}^{n} \beta_j = 0 \text{ (unbiasedness)}$$

Estimated at any node Z^* is defined by:

$$\sigma_{UK}^2 = \sum_{I=1}^{n} \alpha_i \gamma_{io} + \sum_{i=1}^{L} \phi_p f\left(s_o\right) \tag{7}$$

where Z^* estimator is value at any weights; α_i, β_j are weighting factors of universal cokriging.

Drift $m_{cok}(s)$ can be written as:

$$m_{cok}\left(s\right) = \sum_{l=0}^{t} \eta_{k,l} f_{k,l}\left(s\right) \tag{8}$$

Generally, $m_{cok}(s)$ can be expressed as polynomial, with linear drift (l=2) is assumed, then the cokriging estimation variance σ_{cok}^2 can be written as [14]:

$$\sigma_{cok}^2 = \sum_{i=1}^{n} \alpha_i \gamma_{io} + \sum_{j=1}^{n} \beta_j \gamma_{jo} + \sum_{l=0}^{2} \phi_l f_l\left(s\right) \tag{9}$$

2.6 Model Evaluation

Mean absolute error denoted by the (ME) according the equation (10), which was a measure of the sum of absolute residue and the mean square error (MSE) refer to equation (11), which is the square of the total of square, and the root mean squares error is denoted by (RMSE) according the equation (12). Small mean absolute error values indicating a model with few errors have been measured accurately [17]. While the values of kriged reduced mean square error is denoted by KRMSE, refer to equation (13), showings more accurately the prediction on the point .These models are calculated by:

$$ME = \frac{\sum_{i=1}^{N}\left[z^*\left(s_i\right) - z\left(s_i\right)\right]}{N} \tag{10}$$

$$MSE = \frac{\sum_{i=1}^{N}\left[z^*\left(s_i\right) - z\left(s_i\right)\right]^2}{N} \tag{11}$$

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}\left[z^*\left(s_i\right) - z\left(s_i\right)\right]^2\right]^{0.5} \tag{12}$$

$$KRMSE = \frac{1}{N}\sum_{I=1}^{N}\left\{\frac{\left[z^*\left(s_i\right) - z\left(s_i\right)\right]^2}{\sigma_{UK,i}^2}\right\} \tag{13}$$

where N is the number of data points, and (KRMSE) are chosen as criteria for the final model geostatistical interpolation by using variance of kriging and cokriging [10].

3 Results and Discussion

The spatial data which are used in this paper contains (120) samples, for two sets of mineral ores (Copper and Nickel). Each metal has (60) values of metals with real (x, y) locations from the Korf Property in south western British Columbia, approximately 120 kilometres east of Vancouver [20]. The results in this paper, show that both of the proposed kriging and cokriging techniques as well as the parameters of variogram function used rely on comparison of errors. Table 1 shows the results of experimental variogram function was applied according equation (1) for all directions of (60) samples data and copper metal (Cu). Where h is lag or distance between the observations, $\gamma(h,\theta)$ is variogram function with angle (θ) and θ= (0, 90, 45, 135). We also show that (0, 90) have the same lag while (45,135) have another lags.

Table 1 The results of variogram function for Cu

Lag (h)	γ(h,0)	γ(h,90)	Lag (h)	γ(h,45)	γ(h,135)
1	0.3179	0.0211	1.414	0.5329	0.3119
2	0.4717	0.0273	2.828	0.9257	0.4629
3	0.8009	0.0388	4.243	1.6943	0.7533
4	1.4289	0.0488	5.657	3.3500	1.2478
5	2.7694	0.0602	7.071	0.5329	2.2546

Figure (1) below gives the results of the variogram function in all directions (0, 90, 45 and 135). We noted these curves between lag (h) on x-axis and vriogram function (gamma(h)) on y-axis have similar behaviors in all directions for copper metal (Cu) of (60) samples.

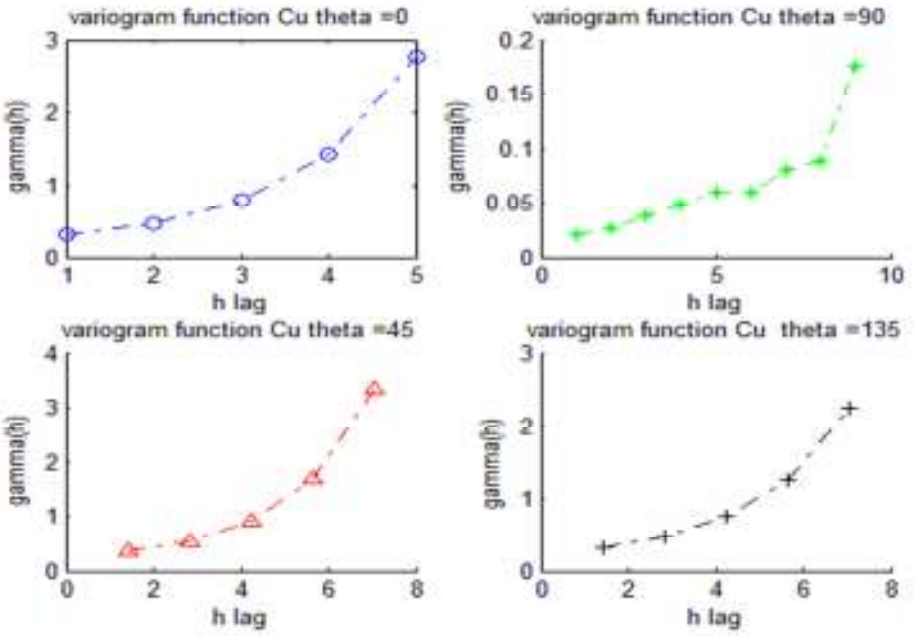

Figure 1 Variogram function for (Cu)

Consequently, Table 2 shows the results of experimental variogram function according equation (1.1) for all directions of angles (0, 90, 45, and 135) of (60) samples for Nickel metal where lag h is the distance between the observations, the second and the third columns gave the variogram function with theta 0,and 90 (have the same lag). Similarly, we obtained the variogram functions in fifth and the sixth columns

Table 2 The results of variogram function for Ni

Lag (h)	ɣ(h,0)	ɣ(h,90)	Lag (h)	ɣ(h,45)	ɣ(h,135)
1	0.0524	0.0231	1.414	0.0594	0.0356
2	0.1192	0.0293	2.828	0.1381	0.1023
3	0.2289	0.0311	4.243	0.2957	0.1888
4	0.3333	0.0375	5.657	0.4281	0.2561
5	0.4260	0.0502	7.071	0.5642	0.2298

Figure (2) shows four curves of all directions (0, 90, 45, and 135) for the second metal (Ni) in all directions of variogram functions where the x-axis represent the lag h and the y-axis represent the variogram functions according the results that are obtained in Table 2.

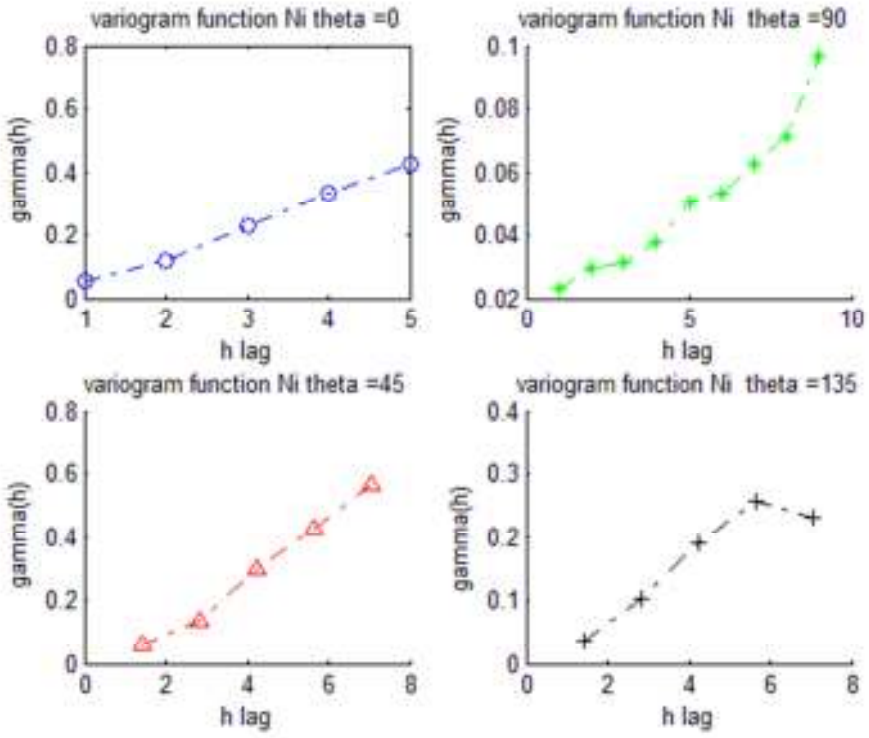

Figure 2 The variogram function for (Ni)

Figure (3a) illustrates the average of variogram function for copper metal (Cu) with two directions (0, 90, blue curve) and (45, 135, red curve). While Figure (3b) demonstrates the average of variogram function for Nickel metal (Ni) with two directions (0, 90 blue curves) and (45, 135 red curve).

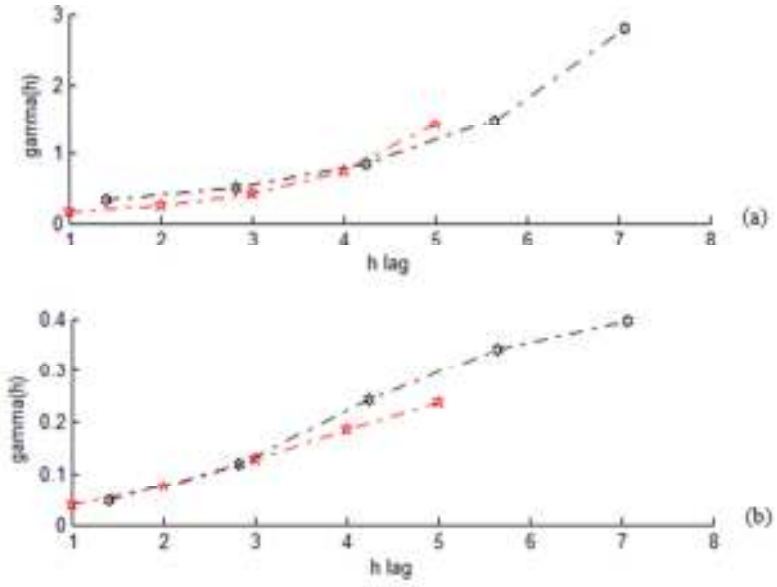

Figure 3 (a) The average of variogram function for Cu and **(b)** The average of variogram function for Ni

The curves of the first metal (Cu) shown in Figure (3a) are the best fitted of power model that is defined as:

$$\gamma(h) = c_o h^a \text{ , } 0 < h < 2 \tag{14}$$

While the curves of the second metal (Ni) refer to Figure (3b) are the best fitted of Gaussian model that is defined as:

$$\gamma(h) = c_o \left[1 - \exp\left(\frac{-h^2}{a} \right) \right], h \succ 0 \tag{15}$$

In Table (3) we note the parameters of variogram function for the data of copper metal, these data has no sill or range. In other words, we cannot clearly identify of sill or range and nugget effect is =0.1695 for copper in angle (0,90) and nugget effect is =0.3363for copper in angle (45,135) the curves of this metal nearest to power model refer to equation (1.14), while the Gaussian model equation (1.15) was chosen of nickel metal because the curves of all directions for variogram function are similar of Gaussian model with all parameters of sill (0.2381)l, nugget effect (0.03775) and range (4) in angle (0,90), while in angle (45,135) sill =0.3970 , nugget effect = 0.04752 and the range =5.657.

Table 3 The parameters of variogram function for the fundamental directions

Metal	Direction	Nugget effect	Sill	Range
Copper	$[\theta=0^{\circ}, \theta=90^{\circ}]$	0.1695	---	---
	$[\theta=45^{\circ}, \theta=135^{\circ}]$	0.3363	---	---
Nickel	$[\theta=0^{\circ}, \theta=90^{\circ}]$	0.03775	0.2381	4
	$[\theta=45^{\circ}, \theta=135^{\circ}]$	0.04752	0.3970	5.657

The estimated values of each metal with the models of covariance function are shown in Table 4 below, and the error in the prediction insignificant, and we used to variance estimator kriging. In addition that, the minimum error in prediction, also the weights are satisfied the conditions and computed under minimum variance which makes the mean prediction error minimum.

Table 4 The comparison between both techniques kriging and cokriging for four mineral ores

		kriging		**cokriging.**	
Metal	Observed of metal	Estimated of metal	Absolute bias	Estimated of metal	Absolute bias
Copper	35.000	35.023	0.023	35.002	0.002
	10.000	9.768	0.232	10.001	0.001
Nickel	234.000	233.754	0.246	233.896	0.104
	392.000	392.469	0.469	391.999	0.001

The cross validation results in Table 5 below, contain mean error, mean square error, root of mean square error, and kriged reduced mean square error refer to equations (10, 11, 12, and 13) respectively. In this table, we show there is a small difference between the results of both kriging and cokriging techniques. Although the results of both techniques close to 1, but still, the second technique (cokriging) is better than the first technique (kriging).

Table 4 The cross validation results

Medel	**ME**	**MSE**	**RMSE**	**KRMSE**	
				kriging	Cokriging
Power	-0.3143	1.2204	1.1047	0.7999	0.8976
Gaussian	-0.1857	2.5578	1.5993	0.8984	1.0146

5 Conclusion

This study shows that kriging techniques can be used for the best prediction of mining field. There are many benefits over universal kriging and cokriging techniques with the small variations in forecasting process. Each combination of both techniques is added to the spatial interpolation as auxiliary variables. The behaviour of each metal has different curve than the others. The models of covariance function are shown with their features by using the variogram and cross-variogram functions where the power model is clearly for the copper metal, while the Guassian model represents the data of nickel metal. The differences between both techniques were little. The results of this paper show that cokriging technique has better performance than universal kriging as the variance of the prediction error is minimized and satisfies the weights.

Acknowledgements

The authors would like to sincerely thank ISM-1 committee for giving them the opportunity to present their work.

References

[1] Matheron, G. Principles of geostatistics. *Economic Geology*. 1963. 58: 1246–1266.
[2] Cherry, S., Banfield, J. and Quimby, W. F. An evaluation of a non-parametric method of estimating semi-variograms of isotropic spatial processes. *USA Journal of Applied Statistics*. 1996. 23(4): 435–449.

[3] Cressie, N. A. *Statistics for Spatial Data*. New York: Wiley. 1991.
[4] Cressie, N.A. *Statistics for Spatial Data*. 2nd Edition. New York: John Wiley and Sons. 1993.
[5] Journel, A. G. and Huijbregts, C. J. *Mining Geostatistics London*. Academic Press. 1978.
[6] Philip, R.D. and Kitanidis, P.K. Geostatistical estimation of hydraulic head gradients ground water. 1989. 27(6): 855–865.
[7] Robertson, G. P. Geostatistics in ecology: Interpolating with known variance. *Ecology*. 1987. 68(3): 744–748.
[8] Eldeiry, A. and Garcia, L. Comparison of regression kriging and cokriging techniques to estimate soil salinity using landsat image. *Hydrology*. 2009. CO 80523-1372.
[9] Johansson, B., and Chen, D. The influence of wind and topography on precipitation distribution in Sweden: Statistical analysis and modeling. *International Journal of Climatology*. 2005. 23: 1523-1535.
[10] Kumar, V. and Ramadevi J. kriging of groundwater levels: a case study. *Spatial Hydrol*. 2006. 6(1): 81–94.
[11] Rudolf, D. *Introduction to Geostatistics and Variogram Analysis*. Vienna University of Technology. 2003.
[12] Webster, R. and Oliver, M.A. *Reliability of the Experimental Variogram*. In: Geostatistics for Environmental Scientists (Eds.). Barnett V. Chichester: John Wiley and Sons. pp. 78-92. 2001.
[13] Chiles, J. P. and Delfiner, P. *Geostatististics, Modeling Spatial Uncertainty*. Wiley-Interscience. 1999.
[14] Odeh, I.O. A., McBratney, A. B. and Chittleborough, D. J. Further results on prediction of soil properties from terrain attributes: Heterotopic cokriging and regression-kriging. *Geoderma.* 1995. 67(3-4): 215–226.
[15] Zhang, R. Myers, D.E. and Warrick A.W. Estimation of the spatial distribution of the soil chemicals using pseudo cross-variograms, *Soil Sci. Soc. Am. J.* 1992. 56: 1444–1452.
[16] Diggle, P. J. and Ribeiro, P. *Model-based Geostatistics.* Springer Series in Statistics, Springer Science, Business Media, LLC. pp. 228. 2007.
[17] Schabenberger, O. and Gotway, C.A. *Statistical Methods for Spatial Data Analysis Texts in Statistical Science*. Chapman and Hall. 2005.
[18] Hoeksema, R.J. and Kitanidis, P.K. Prediction of transmissivities, heads, and seepage velocities using mathematical modeling and geostatistics. *Adv Water Resources*. 1989. 12: 90-102.
[19] Miller, J., Franklin, J. and Aspinall, R. Incorporating spatial dependence in predictive vegetation models, *Ecol. Modelling*. 2007. 202(3-4): 225-242.
[20] John, A. and William, A. *Diamond Drilling and Geochemical Assessment Report on Krof Property*. (Revision No. 1) Soil Geochemical Survey and Core Drilling. 2010.