A Special Analytical Methodology for Variogram Modeling and Interpolation of Terrain Elevation Data by Kriging Method

Dr. Mohaammad Al-Abdallah[(1)]

Abstract

This paper presents a special scientific analytical methodology to conduct geostatistical spatial analysis, Variogram modeling and interpolation by Kriging method using terrain elevation data measured over geographical spatial unit, while accounting for anisotropic behavior of terrain within this unit. The methodology which includes the design of surface interpolation that gives weights to all data points, starts by performing geostatistical analysis and building the Variogram chart. The Variogram models that best representing the data is computed by using standard mathematical regression functions. The modeling process is achieved by using iterative methods and nonlinear least squares optimization process. The coherence between Variogram models constraint and the weights used in the kriging system ensures statistically the best unbiased estimators as well as minimum variances for the interpolated values. Kriging reduces the unrealistic smoothing surfaces inherited in other interpolation methods. It is also robust with respect to very small spatial differences in data points positions, where they are included in the process. There are a large number of semi-Variogram models that could be employed, although different models may lead to different interpolations. The study focuses on the ten most popular models (some of them recently discovered). The mean value of absolute variances provides valuable information help us to select which model is the best from several candidates. If anisotropy exists in variography according to different directions, then several Variogram models needs to be determined.

Special Matlab programs were written by the author for implementing all stages of the above methodology. The study has shown that the interpolation process by Kriging fails in some cases and inaccurate in other cases Thus we need easy and fast computational tools performing many experiments at the same time giving clear representation results and final error analysis, so that the best solution is reached at last. This was the main and most important achievement of this study.

Key Words: *Variogram Analysis and Estimation, Variogram Modeling, Optimization, Geospatial Interpolation, Ordinary Kriging, Terrain Elevation Data*

[(1)] Fculty of Civil Engineering, Damascus University, Syria.

تطبيق منهجية تحليلية خاصة من أجل نمذجة "الفاريوغرام" واستكمال بيانات ارتفاعية حقلية باستخدام طريقة "كريجينغ"

د. محمد صالح العبدالله[1]

ملخص

يقدم هذا المقال منهجية علمية خاصة في التحليل الإحصائي المكاني والنمذجة للفاريغرم والاستكمال بطريقة "كريجينغ" لبيانات طبوغرافية ارتفاعية مقيسة ضمن مساحة جغرافية، مع الأخذ بالحسبان السلوك غير المتجانس لتضاريس الأرض. تتضمن المنهجية تصميماً يوضّح عملية استكمال السطح بإعطاء أوزان لبيانات النقاط جميعها، إذْ يبدأ بإجراء تحليل جيوإحصائي لتغيرات الارتفاعات و إنشاء "مخطط التغيرية"، يتبع ذلك نمذجة "الفاريوغرام" والإلباس بمنحنيات رياضية معيارية، وأخيراً اختيار النموذج الأفضل. يجري الإلباس عن طريق تطبيق الحل اللاخطي بالتربيعات الصغرى والتقريب المتتالي حتى الوصول إلى لنهايات الحدية الصغرى لمربعات الرواسب. من خلال العلاقة المتينة بين شرط تبعية نموذج الفاريوغرام من جهة والأوزان المستخدمة في نظام كريجينغ من جهة أخرى، نضمن إحصائياً الحصول على أفضل تقديرات غير منحازة، وتباينات أصغرية للقيم المستكملة. أيضاً نقلل أو نتفادى بهذه الطريقة تأثير السطوح الملساء (الكاذبة) الملازمة لطرائق الاستكمال الأخرى. وأخيراً كريجينغ يسمح بوجود نقاط قريبة جداً من بعضها البعض ويستوعبها في العملية. هنالك عدد كبير من نماذج الفاريوغرم التي يمكن الاعتماد عليها، ومع الرغم من أنّ كلا منها قد يؤدي إلى نتيجة استكمال مختلفة. ركزت الدراسة على عشرة نماذج هي الأكثر استخداماً، وبعض منها حديث الاكتشاف. تقدم قيمة وسطي التباينات بالقيمة المطلقة معلومات مفيدة لاختيار النموذج الأمثل بين مجموعة من النماذج الصالحة. في حالة عدم وجود التجانس الذي نكتشفه من مخططات التغيرية باتجاهات مختلفة، عندها يجب الأخذ بالحسبان إدخال أكثر من نموذج واحد في الحل.

حضّر المؤلف مجموعة من البرامج الخاصة باستخدام برنامج ماتلاب بهدف اختبار مراحل المنهجية كلّها أعلاه. أثبتت الدراسة أن عملية الاستكمال بطريقة كريجينغ تخفق في بعض الحالات وغير دقيقة في حالات أخرى ولهذا نحتاج إلى أدوات سهلة وسريعة تقوم بعمل اختبارات عديدة في آن واحد وتعطي نتائج ومخططات واضحة وتحليل نهائي للأخطاء في كل تجربة، وذلك كله بقصد الوصول إلى أفضل حل. هذه كانت من أهم ما تم إنجازه من خلال هذه الدراسة.

[1] كلية الهندسة المدنية، جامعة دمشق، سورية.

Introduction

Spatial statistics and geostatistics have developed to describe and analyze the variation in both natural and man-made phenomena on, above or below the land surface. Spatial statistics includes any of the formal techniques that study entities that have a spatial index (Cressie 1993). Geostatistics is embraced by this general umbrella term, but originally it was more specifically concerned with processes that vary continuously, i.e. have a continuous spatial index. The term geostatistics applies essentially to a specific set of models and techniques developed largely by Matheron (1963) in the 1960s to evaluate recoverable reserves for the mining industry. These ideas had arisen previously in other fields; they have a long history stretching back to Mercer and Hall (1911), Youden and Mehlich (1937), Kolmogorov (1941), Matérn (1960) and Krige (1966).

Geostatistics have since been applied in many different fields, such as agriculture, fisheries, hydrology, geology, meteorology, petroleum, remote sensing, soil science, GIS and so on. In most of these fields the data are fragmentary and often sparse, therefore there is a need to predict from them as precisely aspossible at places where they have notbeen measured. This paper covers two of the principle techniques of geostatistics that solve this need for prediction; the *Variogram Estimation* and *Variogram Modeling*. The first one depends on geostatistics and spatial statistics while the second one depends on mathematics.

A brief summary only is given here of the theory that underpins geostatistics (formore detail see Journel and Huijbregts, 1978; Goovaerts, 1997; Webster and Oliver 2007). Most spatial properties vary in such a complex way that the variation cannot be defined deterministically. To deal with this spatial uncertainty a different approach from the traditional deterministic methods of spatial analysis was required that relies on a stochastic or probabilistic approach. The basis of modern geostatistics is to treat the variable of interest as a random variable. This implies that at each point *x* in space there is a series of values for a property, $Z(x)$, and the one observed, $z(x)$, is drawn at random according to some law, from some probability distribution. At *x*, a property $Z(x)$ is a random variable with a mean μ and variance σ^2. The set of random variables, $Z(x_1),\ldots,Z(x_N)$ is a random process, and the actual value of *Z* observed is just one of potentially any number of realizations of the random process. In classical statistics this set of observed values, there alization, is the population. The modeling and simulation of natural phenomena are based on the assumption that the a process $\{z(x), x \in D\}$ is a realization of a stochastic (or random) function $Z(x)$ where *D*is a fixed subset in R^d (a positive d-dimensional space). Matheron (1962) called the-quantity $Z(x)$ *a regionalized random variable*, allowing the presence of inhomogeneity in the physical process as well as emphasizing the natural continuity of space within the subset *D*. It has been established that fitting invalid covariance model to the Variogram can yield to a negative-definite variance $Var(Y)$, where *Y* represents any linear combination of $Z(x)$. The problem when using such models, is that it does not guarantee a unique solution of the ordinary kriging system and the same holds for any kind of simulation based on kriging, thus from this perspective we say that they are invalid. The idea is then to search for a valid Variogram model that, as a measure of correlation, is closest to the experimental Variogram. (Some authors call $\gamma(h)$ as the Variogram instead of Semi variogram). The space of valid variograms is a large set of parametric family or 'basic models', that are known to be positive-definite. We can also enlarge this family by combining those functions to form new ones that are also positive-definite and produce what is called a nested structures or nested models. Some preliminary assumptions have to be made in order to make the statistical inference about $Z(x)$ possible, thus we start from definition of those hypothesis that form the basis of most geostatistical theory.

Second-order Intrinsic Stationarity Assumptions

Stationarity exists when the probabilistic distribution of $Z(x)$ is invariant and does not depend on *x*. let us call *m* the trend (or drift), which can be expressed as the expectation of the random variable $Z(x)$

$$E(Z(x)) = m, \forall x \in D \quad (1)$$
$$F_x(z) = \Pr(Z(x) \le z), \forall x \in D \quad (2)$$

In order to estimate an optimal linear predictor (using Kriging for example), an additional assumption isneeded. Having sufficient number of sampled pairs $z(x_i), z(x_j), (i, j) \in 1,...,n$, where $(x_i, x_j) \in D$ refer to two different locations in D , and linked by a vector $h_{ij} = x_i - x_j$, let the function

$$Cov(Z(x_i), Z(x_j)) = C(x_i - x_j) = C(h_{ij}), \forall x_i, x_j \in D \quad (3)$$

defines the *Covariogram*, or the *stationary covariance function*. Any random function $Z(x)$ satisfying (1) and (2) and possess a stationary covariance function, i.e. the *Covariogram* (3) exists, is said to be a *Second-order Stationary Process*. Furthermore, if $Cov(x_i - x_j)$ is only a function of $\|x_j - x_i\| = \|h\|$ only, then $Z(\cdot)$ is called the isotropic Covariogram. On the other hand, if the random function $Z(x)$ satisfying (1) and (2) and possess a stationary variance function, i.e. the Variogram (4) exists, then it is said to be *Intrinsically Stationary Process*. It is clear that the *Second-order Stationarity* hypothesis implies the *Intrinsic Stationarity*, but the converse is not true.

Variogram and Semivariogram

Observations closer together tend to be more alike and statistically correlated than observations farther apart. In geostatistics, this idea of autocorrelation is quantifiedthrough a function called a s*emivariogram*. The quantity $2\gamma(h)$ that defined by,

$$Var(Z(x_i), Z(x_j)) = 2\gamma(x_i - x_j) = 2\gamma(h_{ij}), \forall x_i, x_j \in D \quad (4)$$

Which is a function of only the increment $h_{ij} = x_i - x_j$ is called the *Variogram* and $\gamma(h)$ or *Semivariogram* by Mather on (1962). The latter name is most popular, (and will be used frequently through this paper). Its estimation is achieved by takinghalf of the average square difference between two samples valuesapproximately separated by a predefined lags *h*:

$$\gamma_x(h) = \frac{1}{2N(h)} \sum_{k=1}^{N(h)} (Z(x_i) - Z(x_j))^2 \quad (5)$$

Where: $N(h)$ is the number of distinct pairs p_{ij} belonging to a separation vector *h*, and $Z(x_i), Z(x_j)$ are the sampled values at the beginning location and end location respectively. Thus, we can define the *Variogram* function as the variance of only the increment vector h.

Cross Variogram

Let
$Y(x) \equiv [Y(x_1),...,Y(x_n)], Z(x) \equiv [Z(x_1),...,Z(x_n)], \forall x \in D$

be two co-located spatial processes, where each is assumed to possess a Variogram thus

$$2\gamma_Y(x_i - x_j) = Var(Y(x_i), Y(x_j)) \quad (6)$$
$$2\gamma_Z(x_i - x_j) = Var(Z(x_i), Z(x_j)), \forall x_i, x_j \in D$$

There are two ways to generalize the previous notations to account for cross-dependence between the two processes $Y(\cdot)$ and $Z(\cdot)$. The most natural one for multivariate spatial prediction (*Cokriging*) is

$$2\gamma_{YZ}(x_i - x_j) = Cov(Y(x_i), Z(x_j)), \forall x_i, x_j \in D \quad (7)$$

In similar manner to the *Semivariogram*, another measure of spatial variability used by *Cokriging* under special conditions called Cross Semivariogram, due to Journel and Huijbregts (1978), and can be estimated by taking half of the average of cross product of all sampled pairs, having two different

attributes, and associated with two different locations separated by a predefined separation lag *h*,

$$\gamma_{YZ}(h)=\frac{1}{2N(h)}\sum_{k=1}^{N(h)}[Y(x_i)-Y(x_j)][Z(x_i)-Z(x_j)]. \qquad (8)$$

Covariogram and Correlogram

The function $Cov(Z(x_i),Z(x_j))=C(h_{ij}),\forall x_i,x_j\in D$, is given earlier by expression (8), defines the *Covariogram*. Notice that this statistics has another name like *Auto-Covariance function* known in *time series analysis*. The *Covariogram* can be estimated using the following formulae,

$$C(h)=\frac{1}{N(h)}\sum_{k=1}^{N(h)}\left[Z(x_i)\cdot Z(x_j)\right]-(x_{xi}\cdot x_{xj}) \qquad (9)$$

$$m_{xi}=\frac{1}{N(h)}\sum_{k=1}^{N(h)}Z(x_i) \quad (1), m_{xj}=\frac{1}{N(h)}\sum_{k=1}^{N(h)}Z(x_j) \quad (10)$$

On the other hand, the *Correlogram*is another spatial statistics denoted by $\rho(h)$ (in *time series* terms this is called *Auto-Correlation function*). This statistics can be estimated under the assumption that $C(h)>0$ as follows,

$$\rho(h)=\frac{C(h)}{C(0)} \qquad (11)$$

Checkthat $\rho(h)=1$ when $C(h)=C(-h)=C(0)$.

The quantity $C(0)$ is called the *sill of the Semivariogram*. In fact the sill $C(0)$ defines the upper bound of the Semivariogram model for $\|h\|\to\infty$ or practically for $\|h\|\geq\|h_0\|$ where $\|h_0\|$ defines the *range*. This quantity can be decomposed into a Variogram $\gamma(h)$ and Covariogram $C(h)$. First consider the relation

$$Var(Z(x_i),Z(x_j))=Var(Z(x_i))+Var(Z(x_j))...$$
$$-2\cdot Cov(Z(x_i),Z(x_j)),\forall x_i,x_j\in D \qquad (12)\ .$$

Recall form (3) and (4),

$$Cov(Z(x_i),Z(x_j))=C(x_i-x_j)=C(h_{ij})$$
$$Var(Z(x_i),Z(x_j))=2\gamma(h_{ij}),\forall x_i,x_j\in D\ .$$

In addition, under the second order stationarity assumption we can write

$$Var(Z(x_i))=Var(Z(x_j))=E[(Z(x)-m)^2]=C(0) \qquad (13)$$

$$2\gamma(h)=2C(0)-2C(h)\Rightarrow$$
$$\gamma(h)=C(0)-C(h) \qquad (14)$$
$$\rho(h)=\frac{C(h)}{C(0)}=1-\frac{\gamma(h)}{C(0)} \qquad (15)$$

A Variogram function can be deduced from a covariance function using the formula (14), but in general the reverse is not true because some *Semivariogram* models like the *linear models or power models* have no covariance function counterparts, as they grow without bounds. If the assumption that the mean of the tail values m_{xi}, is not the same as the mean of the head values m_{xj}, then the *Correlogram*, is slightly defined in different way,

$$\rho(h)=\frac{C(h)}{\sqrt{C(-h)}\cdot\sqrt{C(+h)}} \qquad (16)$$

$$C(-h)=\frac{1}{N(h)}\sum_{k=1}^{N(h)}x_i^2-m_{-h}^2 \qquad (17)$$

$$C(+h)=\frac{1}{N(h)}\sum_{k=1}^{N(h)}x_j^2-m_{+h}^2 \qquad (18)$$

Positive definite conditions

Let $Z(x)$ be a stationary random process with expectation m and covariance $C(h)>0$ or Semivariogram $\gamma(h)$. Let Y be any finite linear combination of $Z(x)$ as follows,

$$Y=\sum_{i=1}^{n}\lambda_i\cdot Z(x_i) \qquad (19)$$

for any set real numbers or weights vector $W=\{\lambda_i\},1\leq i\leq n$. This linear combination and its variance must be *positive-definite*, that is

$$Var(Y) = \sum_i \sum_j \lambda_i \cdot \lambda_j \cdot C(x_i - x_j) \geq 0 \qquad (20)$$

The last expression can be written in matrix form

$$Var(Y) = W^t \cdot C_h \cdot W \geq 0 \qquad (21)$$

Where C_h is the Covariance matrix that is defined by a covariance function $C(h) > 0$ and a set of points x_i, thus the function $C(h)$ is said to be positive-definite in order to ensure the positive-definiteness of the variance $Var(Y)$. On the other hand, the Semivariogram $\gamma(h)$ is said to be conditionally *negative-definite function* in order to guarantee the positive-definiteness of $Var(Y)$. If we rewrite (14) in matrix form corresponding to a set of points x_i,

$$\Gamma_h = C_0 - C_h \qquad (22)$$

where the matrix Γ_h represents all Semivariogram functions $\gamma(h)$, C_0 is a matrix of the same size as Γ_h whose all elements are equal to the sill $C(0)$ of Semivariogram. Therefore

$$Var(Y) = W^t \cdot C_h \cdot W \geq 0 \Rightarrow \quad W^t \cdot \Gamma_h \cdot W \leq 0 \qquad (23)$$

In the case when the sill does not exist and only the intrinsic hypothesis is assumed, then the variance of Y is defined on the condition that

$$\sum_i \lambda_i = 0 \qquad \Rightarrow \quad Var(Y) = -W^t \cdot \Gamma_h \cdot W \leq 0 \qquad (24).$$

Thus when handling linear combination of random variables, then the Semivariogram can only be used together with conditions on the weights guaranteeing its existence.

Behavior of the phenomenon near the origin(Nugget Effect)

The Semivariogram expectation at a very small scale, which describes the behavior of phenomenon near the origin, is known as the nugget effect, after Matheron (1962). This is because it is believed that micro-scale variation is causing a discontinuity near the origin. In terms of Semivariogram prediction, nugget effect c_0 is defined by $\gamma(h)_{h \to 0} = c_0 > 0 \qquad (25)$

The behavior at a very small scale is very important as it indicates the type of discontinuity of the phenomenon near the origin, and we can distinguish three types of phenomena:

- *continuous and differentiable* near the origin;
$$\gamma(h) \to 0, \big| \quad \|h\| \to 0$$
- *discontinuous or non-differentiable* near the origin, then we have nugget effect;
$$\gamma(h) \to c_0 > 0, \big| \quad \|h\| \to 0$$
- *white noise process* with constant variance and zero-covariance (pure nugget);
$$\gamma(h) \to c_0 > 0, \big| \quad \forall h\,.$$

Statistically speaking, if the phenomenon is continuous (or expected to be continuous) at the micro-scale, then the only reason for $c_0 > 0$ is the measurement error. This means that if the *Variogram* is modeled with different sampling schemes or using different approaches, the value of c_0 would fluctuate around its true value, thus $c_0 = c_{ms} + c_{me}$, where c_{ms} represents the nugget effect at the micro-scale, while c_{me} represents measurements error. In practice, there is a problem to determine c_0 from data whose separations $\|h\|$ are too large to capture accurate micro-scale information. Typically, it is determined by extrapolation of Variogram estimates from lags closest to zero.

Parametric Isotropic Semivariogram Models

A review of the most frequently used isotropic *Semivariogram* models are given, as well as the general conditions that a model should satisfy in order to be valid. Those models can be classified into two categories:
models with a *sill*(or transition models) and Models without a *sill* (recall from a previous section that forthe second category a covariance function does not exist and only a Variogram model $\gamma(h)$ is defined).

To the first category goes: the *Spherical model*, the *Exponential model*, the *Gaussian model*, the *Rational quadratic model* and the *hole-effect model*, while To the second category goes: the *Linear model*, the *nugget effect model*, the *power model* and the *Logarithmic model*.

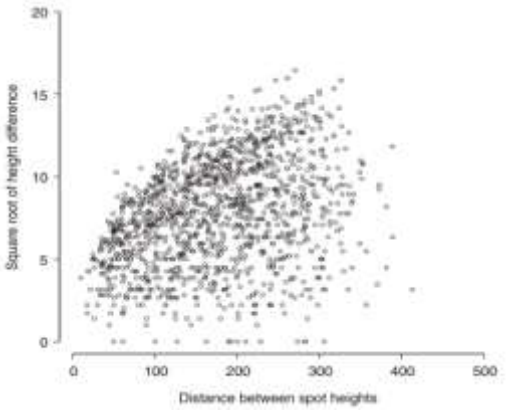


FIG.1 The square root differences cloud for elevation data

There are many parametric functions that satisfy the properties of the semivariogram (see, e.g., Journel and Huijbregts 1978; Chiles and Delfiner 1999). We say that a semivariogram model is valid in d dimensions (i.e., in R^d) if it satisfies the folloing conditions: (let's refer to $Z(x_i)$ by *Z(s)* and $Z(x_j)$ by *Z(u)* for simplicity):

- $\gamma(-h)=\gamma(+h),$ (26)
  the autocorrelation between *Z(s)* and *Z(u)* is the same as that between *Z(u)* and *Z(s)*].
- $\gamma(0)=0,$, since, *Var(Z(s) − Z(s)) = 0*. (27)
- $(\gamma(h)/\|h\|^2)\to 0,$ *as* $\|h\|\to\infty,$
- $\gamma(\cdot)$ must be conditionally negative definite, that is for any number of locations {*s(i), i=1,…,m*} and real numbers {a(i),….,a(m)} satisfying $\sum_{i=1}^{m}\sum_{j=1}^{m}a_i a_j\gamma(s_i-s_j)\le 0$ $\sum_{i=1}^{m}a_i=0$ this condition is analog of the positive-definite condition for variance-covariance matrices. Here below is given some of the ten most popular models:

- *Linear Model*

$$\gamma(h,\theta)=c_0+c\cdot\left(\frac{h}{a}\right),\qquad h>0, \tag{28}$$

- *Spherical Model*

$$\gamma(h,\theta)=\begin{cases}c_0+c_s\cdot\left(\frac{3}{2}\frac{h}{a_s}-\frac{1}{2}\left(\frac{h}{a_s}\right)^3\right),(0<h\le a_s)\\ c_0+c_s,(h>a_s)\end{cases} \tag{29}$$

- *Gaussian Model*

$$\gamma(h,\theta)=c_0+c_g\cdot\left(1-\exp(-\frac{h^2}{a_g^{\,2}})\right),\qquad h>0 \tag{30}$$

- *Exponential Model*

$$\gamma(h,\theta)=c_0+c_e\cdot\left(1-\exp(-\frac{h}{a_e})\right),\qquad h>0 \tag{31}$$

- *Circular Model* (32)

$$\gamma(h,\theta)=c_0+c_{cr}\cdot\left(1-\frac{2}{\pi}\cdot a\cos\left(\frac{h}{a_{cr}}\right)+\frac{2h}{\pi\cdot a_{cr}}\sqrt{1-\frac{h^2}{a_{cr}^2}}\right)$$

- *K-Bessel (Wittle) Model*

$$\gamma(h,\theta)=c_0+c_w\cdot\left(1-\frac{h}{a_w}\cdot Bessel^{K}\left(1,\frac{h}{a_w}\right)\right), \tag{33}$$

- *Sine Model* (*Hole Effect model*)

$$\gamma(h,\theta)=c_0+c_{he}\cdot\left(1-\exp\left(\frac{-a_{he}}{h}\right)\cdot\sin\left(\frac{h}{a_{he}}\right)\right) \tag{34}$$

- *Pentaspherical Model*

$$\gamma(h,\theta)=c_0+c_p\cdot\left(\frac{15}{8}\frac{h}{a_p}-\frac{5}{4}\left(\frac{h}{a_p}\right)^3+\frac{3}{8}\left(\frac{h}{a_p}\right)^5\right) \tag{35}$$

- *Rational Quadratic Model:*

$$\gamma(h,\theta)=c_0+c_r\cdot\|h\|^2/(1+,\|h\|^2/a_r)\quad h>0 \tag{36}$$

- *Power Model*

$$\gamma(h,\theta)=c_0+c\cdot\left(\frac{h}{a}\right)^p,\qquad p>1,h>0, \tag{37}$$

Remark:Parameters $\theta = (c_0, c_s, a_r)', c_0 \geq 0, c_s \geq 0, a_r \geq 0.$ refers to the three parameters : *Nugget effect, Sill and Range* respectively (FIG.3).Parameters $\{c_s, c_e, c_g, c_r, c_k, c_w, \ldots\}$ in all models refers to the *Sill*. $\{a_s, a_e, a_g, a_r, a_k, a_w, \ldots\}$ in all models refers to the *Range*. All models are valid in $R^d, d \geq 1$ except *Spherical, Sine* and *Pentaspherical* models are valid in $R^d, d = 1.$

There are many more parametric semivariogram models not described here [see, Armstrong (1999), Chiles and Delfiner (1999), and Olea (1999)].In addition, the sum of two semivariogram models that are both valid in R^d is also a valid semivariogram model in R^d, so more complex models can be generated by adding two or more of these basic semivariogram models (Christakos 1984)

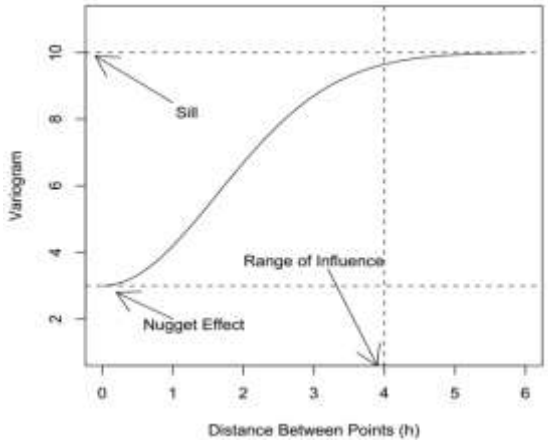


FIG.2Typical semivariogram with Sill, Range and Nugget Effect

Semivariogram models created this way are referred to as models of nested structures.

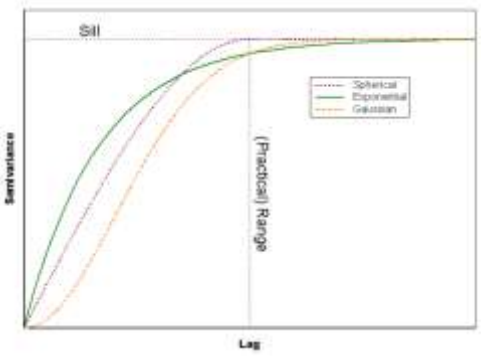


FIG.3Some theoretical semivariogram models showing Sill and Range Positions

Estimating the Semivariogram

The semivariogram can be estimated easily from data $\{Z(si) : i = 1, \ldots, N\}$ under the assumption of intrinsic stationary so that equations (6) and (11) hold. Using rules of expectation, we can write the Variogram as

$$2\gamma(h) = Var(Z(s+h) - Z(s)) \quad (38)$$

$$= E[(Z(s+h) - Z(s))^2] - [E(Z(s+h) - Z(s))]^2$$

From equation (6), $[Z(s_i)] = m \ \forall s_i \in D$, so the second term is zero. Thus, to estimate the Variogram we need only to estimate$E[(Z(s+h) - Z(s))^2]$ Since expectations are just statistical averages, one way to estimate this term is to average all observed squared differences$[Z(s_i) - Z(s_i)]^2$for pairs of observations taken thesame distance apart in the same direction. This is the rationale behind the method of moments estimator of the semivariogram, given by

$$\hat{\gamma}_x(h) = \frac{1}{2|N(h)|} \sum_{k=1}^{N(h)} \left(Z(s_i) - Z(s_j)\right)^2 \quad (39)$$

where *N(h)*is the set of distinct pairs separated by h [i.e., *N(h)* = $\{(s_i, s_j) : s_i - s_j = h, i, j = 1, \ldots, n\}$ and $|N(h)|$= the number of distinct pairs in *N(h)*]. Last Equation gives what is often referred to as the *classical semivariogram estimator*. It gives point estimates of γ (·) at observed values of h. If the process is isotropic, we need only consider pairs lag ||h|| apart. If the process is anisotropic, the semivariogram can be estimated in different directions by selectinga particular direction and

averaging pairs of data lag ||h|| apart in that particulardirection. With irregularly spaced data, there may be only one pair of locations that is *h* apart (two for||h||). Averages based on only one or two points are poor estimates with large uncertainties. We can reduce this variation and increase the accuracy of our point estimates by allowing a tolerance on the lags. Thus, we will define tolerance regions and group the sample pairs into these regions prior to averaging. This is analogous to the procedure used in making a histogram, adapted to two dimensions (FIG.4).

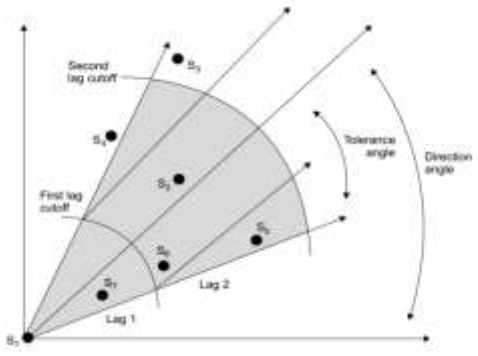


FIG.4 Tolerance regions for semivariogram estimation

Typically, one specifies tolerance regions through the choice of five parameters: the direction of interest; the angle tolerance, which defines a sector centered on the direction of interest; the lag spacing, which defines the distances at which the semivariogram is estimated; the lag tolerance, which defines a distance interval centeredat each lag; and the total number of lags at which we wish to estimate the semivariogram. Tolerance regions should include 20-30 pairs of points each to ensure that the empirical semivariogram at each point is well estimated (Journel and Huijbregts 1978).

Usually, a set of directions and associated angle tolerances are chosen together so that they completely cover two-dimensional space (Fig.5 and 6).

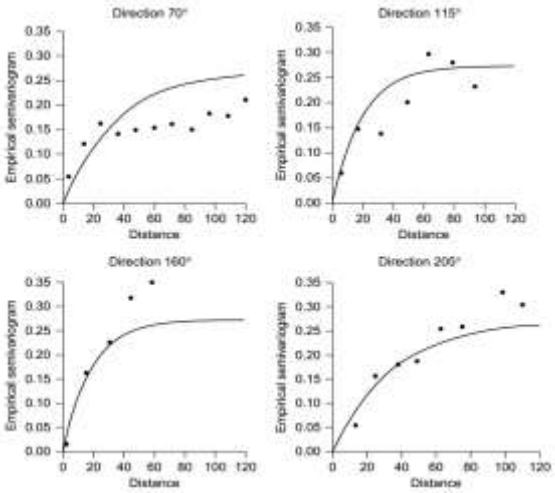


FIG.5 Empirical directional semivariograms and fitted models

One should construct lag in tervalsso that the total number of lags is between 10 and 25 in order to see the structure of the semivariogram. One should be careful of the use of very short maximum lag distances. The semivariogram is a picture of your data spatially: the sill and the range, if they exist, provide estimates of the processvariance and the zone of influence of the observations, and information at largerlags can indicate large-scale trends.

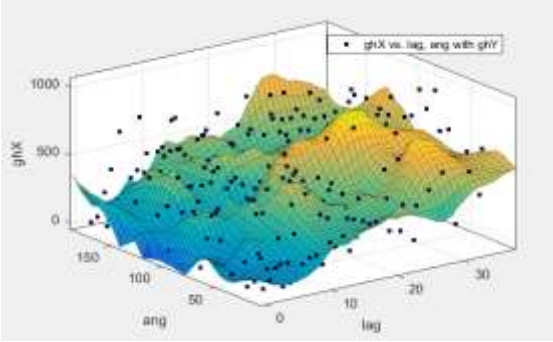


FIG.6Empirical directional semivariogram with 2D representation showing $\gamma(h)$ (z-axis), lags(x-axis) and directions(y-axis)

Fitting Semivariogram Models

The empirical semivariogram $\gamma(\cdot)$ is not guaranteed to be conditionally nonnegative definite. This is not a problem if we limit ourselves to inferences about

thespatial continuity of the process, but it can lead to problems when used for spatialprediction and mapping where we need reliable estimates of prediction uncertainty. Thus, we will need to find a valid theoretical semivariogram function that close lyre flects the features of our empirical semivariogram. We limit our choices to aparametric family of theoretical variograms (like those described in section *Parametric Isotropic Semivariogram Models* in this paper)and seek to find the parameter estimates that best fit the data.

Nonlinear Least Squares Regression Method

The idea here is to select a the oretical semivariogram family and find a vector of parameters $\hat{\theta}$ that makes this theoreticalmodel "*close enough*" to the empirical semivariogram. Let $\hat{\gamma}(\cdot)$ be the *empirical* semivariogram estimated at *K* lags, *h(1), . . . , h(K)* and let $\gamma(h;\theta)$be the *theoretical* semivariogram model whose form is known up to *θ*. Since the relationship between$\hat{\gamma}(h)$ and *h* is usually nonlinear, nonlinear least squares regression can be used toestimate *θ*. Nonlinear ordinary least squares (*OLS*) finds $\hat{\theta}$ minimizing the squared distance between the empirical and theoretical semivariograms, that is, minimizing

$$\sum_{j=1}^{K}\left[\hat{\gamma}(h(j)) - \gamma(h(j);\boldsymbol{\theta})\right]^2. \quad (49)$$

However, the estimates $\hat{\gamma}(h(j))$are correlated and have different variances, violating the general assumptions underlying OLS theory. The usual statistical adjustment to OLS when observations are correlated and heterois generalizedleast squares (*GLS*). Cressie (1985) applied nonlinear GLS to semivariogram estimation, finding $\hat{\theta}$ minimizing the objective function

$$[\hat{\gamma} - \gamma(\theta)]' \cdot V(\theta)^{-1} \cdot [\hat{\gamma} - \gamma(\theta)] \quad (50)$$

$V(\theta)$ the variance–covariance matrix that depends on $\theta$is unknown and θ is unknown, so the best estimator is computed iteratively from starting values that are improved at each iteration until the objective function is minimized. Taking $V(\theta) \equiv I$gives the *OLS* estimator, and taking

$$V(\theta) \equiv diag\{Var(\hat{\gamma}(h_1)\}, \ldots, Var(\hat{\gamma}(h_k)\} \quad (51)$$

gives a nonlinear weighted least squares (*WLS*) estimator. Determining the elements of$V(\theta)$ requires knowledge of the fourth-ordermoments of *Z*. Cressie (1985) showed that a nonlinear *WLS* estimatorbased on the expression:

$$Var[\gamma(h_j)] \approx 2\left[\gamma(h_j);\theta\right]^2/N(h_j) \quad (52)$$

yields an estimation procedure that often works well in practice. Thus, weighting the OLS objective function inversely proportional to the (approximate) variance of the empirical semivariogram estimator gives an estimator of θ that minimizes the weighted regression sum of squares:

$$WRSS(\theta) = \tfrac{1}{2}\textstyle\sum_{j=1}^{k}\frac{N(h_j)}{\left[\gamma(h_j);\theta\right]^2}[\hat{\gamma}(h_j) - \gamma(h_j);\theta]^2 \quad (53)$$

This approach is an approximation to *WLS* and offers a pragmatic compromise between *OLS* and *GLS*. It gives more weight where there is more "data" [large(h_j)] and near the origin [small$\gamma(h_j);\theta$], thus improving on *OLS*. Although it will not be as good as *GLS*, but ease of computation is a definite advantage. It can be used even when the data are not Gaussian, and empirical studies have shown (Zimmerman 1991) this approach to be fairly accurate in a variety of practical situations.

InverseDistance Interpolation

An inverse-distance interpolator is simply a weighted average of neighboring values. The weight given to each observation is a function of the distance between that observation's location and the grid point s_0 at which interpolation is desired.Mathematically, the general inverse-distance interpolator is written as

$$\hat{Z}_{\mathrm{ID}p} = \sum_{i=1}^{N} Z(\boldsymbol{s}_i)\; d_{0,i}^{-p} \Big/ \sum_{i=1}^{N} d_{0,i}^{-p}. \quad (54)$$

Here$d_{0,i}$ is the distance from the grid point location s_0 to the *i*th data locationss_i. The weighting power, *p*, is selected to control how fast the weights tend to zero as the distance from the grid node increases, based on assumed increasing similarity between

observations taken closer together. As the power increases, the contribution (to the interpolated value) from data points far from the grid nodedecreases. Distance powers between 1 and 3 are typically chosen, and taking $p = 2$gives the popular inverse-distance-squared interpolator. [Burrough (1986)].

Interpolation by Kriging

Kriging is a geostatistical technique for optimal spatial prediction. We emphasize the distinction between prediction, which is inference on random quantities, and estimation, which is inference on fixed but unknown parameters. Georges Mather on, the founding father of geostatistics, introduced this term in one of his early works developing geostatistical theory (Matheron 1963).There are many different types of kriging, differing by under lying assumptions and analytical goals. One can consult some references to learn more about *Kriging methods*, (e.g., Journel and Huijbregts 1978; Isaaks and Srivastava 1989; Cressie 1993; Wackernagel 1995; Chiles and Delfiner 1999; Olea 1999; Stein 1999).The basic and most popular method is *Ordinary Kriging* which will be briefly discussed here.

Ordinary Kriging (OK):Assume that $Z(\cdot)$is intrinsically stationary process [i.e., having unknown mean, μ, and known semivariogram, $\gamma(h)$, where $\{Z(si): i = 1,\dots,N\}$represent the data and we want to predict the value of the $Z(\cdot)$process at an unobserved location,$Z(s_0), s_0 \in D$. As with the inverse distance methods described in previous section, the ordinary kriging (OK) predictor is a weightedaverage,

$$\hat{Z}_{\mathrm{OK}}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i), \quad \sum_{i=1}^{N} \lambda_i = 1. \tag{55}$$

However, instead of specifying an arbitrary function of distance, we determine the weights based on the data using the semivariogram and two statistical optimality criteria: unbiasedness and minimum mean-squared prediction error. For unbiasedness, the predicted value should, on average, coincide with the value of the unknown random variable, $Z(s_0)$ In statistical terms. Unbiasedness constraint requires$E\left[\hat{Z}_{ok}(s_0)\right] = E[Z(s_0)] = \mu$, which means that$\sum_{i=1}^{N} \lambda_i = 1$. To ensure the second optimality criterion, we need to minimize *mean-squared prediction error* (***MSPE***), defined as$E\left[\hat{Z}_{ok}(s_0) - Z(s_0)\right]^2$, subject to the unbiasedness constraint. One method for solving constrained optimization problems is the method of *Lagrange multipliers*. With this method, we need to find$\lambda_1, \dots, \lambda_N$and a Lagrange multiplier, *m*, that minimize the objective

$$E\left[\left(\sum_{i=1}^{N} \lambda_i Z(s_i) - Z(s_0)\right)^2\right] - 2m\left(\sum_{i=1}^{N} \lambda_i - 1\right). \tag{56}$$

Function. The second term is essentially a penalty, minimized when$\sum_{i=1}^{N} \lambda_i = 1$, thus ensuring that our overall minimization constraint. Now this implies that

$$\left[\sum_{i=1}^{N} \lambda_i Z(s_i) - Z(s_0)\right]^2 = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j \left[Z(s_i) - Z(s_j)\right]^2 + \sum_{i=1}^{N} \lambda_i \left[Z(s_0) - Z(s_i)\right]^2. \tag{57}$$

gives

$$E\left[\left(\sum_{i=1}^{N} \lambda_i Z(s_i) - Z(s_0)\right)^2\right] = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j E\left[\left(Z(s_i) - Z(s_j)\right)^2\right] + \sum_{i=1}^{N} \lambda_i E\left[\left(Z(s_0) - Z(s_i)\right)^2\right].$$

Thus the equation (56) becomes

$$-\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j \gamma(s_i - s_j) + 2\sum_{i=1}^{N} \lambda_i \gamma(s_0 - s_i) - 2m\left(\sum_{i=1}^{N} \lambda_i - 1\right) \tag{59}$$

To minimize (59), we differentiate with respect to $\lambda_1, \dots, \lambda_N$ and *m* in turn and set the partial derivatives equal to zero. This gives a system of equations, referred to as the *ordinary kriging equations*,

$$\sum_{j=1}^{N} \lambda_j \gamma(s_i - s_j) + m = \gamma(s_0 - s_i), \qquad i = 1, \dots, N \;))$$

We solve these equations for $\lambda_1, \dots, \lambda_N$(and *m*), and use the resulting optimal weights in equation (55) to give the ordinary kriging predictor. Note that $\hat{Z}(s_0)$ has weights that depend on both the spatial correlations between, $Z(s_0)$ and each data point $Z(s_i): i = 1,\dots,N$, and the spatial correlations between all pairs of data points $Z(s_i)$and $Z(s_i): i = 1,\dots,N$ and $Z(s_j): j = 1,\dots,N$. usually we write the ***Kriging System of Equations***(60) in matrix form as

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ m \end{bmatrix} = \begin{bmatrix} \gamma(s_1 - s_1) & \cdots & \gamma(s_1 - s_N) & 1 \\ \gamma(s_2 - s_1) & \cdots & \gamma(s_2 - s_N) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(s_N - s_1) & \cdots & \gamma(s_N - s_N) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(s_0 - s_1 \\ \gamma(s_0 - s_2 \\ \vdots \\ \gamma(s_0 - s_N \\ 1 \end{bmatrix} \quad (61)$$

Note that we must calculate $\lambda_0$for each prediction location s_0. However, only the right-hand side of equation (61) changes with the prediction locations through γ_0. Since the coefficient matrix depends only on the data locations and not on the prediction locations, we need only invert this w matrixonce and then multiply by $\gamma_0$vector to obtain a prediction for any $s_0 \in D$. The minimized MSPE, also known as the ***kriging variance***, which is a measure of the uncertainty in the prediction of$Z(s_0)$, is obtained as follow

$$\begin{aligned} \sigma_k^2(s_0) &= \boldsymbol{\lambda}_O' \boldsymbol{\gamma}_O \\ &= \sum_{i=1}^{N} \lambda_i \gamma(s_0 - s_i) + m \\ &= 2\sum_{i=1}^{N} \lambda_i \gamma(s_0 - s_i) - \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j \gamma(s_i - s_j). \end{aligned}$$

Implementation in Matlab(3 stages):

(1)Variogram Estimation,
(2)Semivariogram Modeling,
(3) Kriging Interpolation.

- Dataset that has been used for testing programs performance (which cover area of 1000×1000m)is given in the form of 3-column matrix (x,y,z) . It is a terrain elevation data consist of 266 points distributed as shown in (Fig.8).

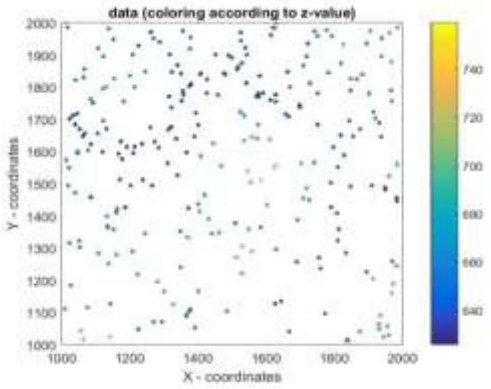


Fig.8 Dataset Locations and their distribution (color indicate the z-value)

Data in the study has been downloaded from internet which was related to a small forested area in Wisconsin, USA, provided by Department of Forest Resources, University of Minnesota.

- By constructing histogram of the dataset values (z values) (Fig.9), we see that some z-values have very large frequencies (like z=630 and z=650) and this means that lower areas is represented by large number of points.
-

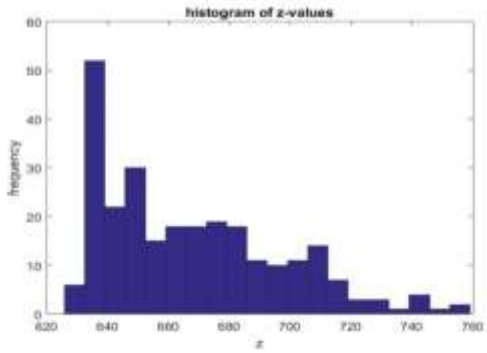


Fig.9 Data histogram shows the frequency of Z-Values

- ***EstiamteVariogram.m*** is a Matlab functions were written by the author for computation and visualization the variogram. It calculates the experimental variogram for concrete number of lags-distances and directions (anisotropic variogram).

Program Input:

x - array with coordinates. Each row is a location in a size(x,2)-dimensional space (e.g. [x y elevation]). y - column vector with values of the locations in x.

nrbins - number bins the distance should be grouped into(default = 20).

maxdist - maximum distance for variogram calculation (default = maximum distance in the dataset / 2).

type - 'gamma' returns the variogram value (default) '*cloud1*' returns the binned variogram cloud , '*cloud2*' returns the variogram cloud .

plotit - true -> plot variogram - false -> don't plot (default). ***subsample*** - number of randomly drawn points if large datasets are used. scalar (positive integer, e.g. 3000) inf (default) = no subsampling

anisotropy - false (default), true (only in 2D)

thetastep - if anisotropy is set to true, specifying thetastep allows you to change the angle width (default 30 degrees)

Program Output:

d - structure array with distance and gamma – vector

plot gamma results as shown below (Fig.10).

- The Program (***EstiamteVariogram.m***) which is Matlab function, computes also the Variogram in all possible directions defined by separation distances is called lag (number of lag distances = 22).

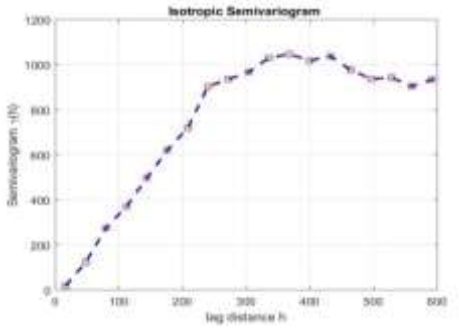


Fig.10 the Isotropic Variogram of our data

- One can visualize *Variogram anisotropy*, which is computed and stored in matrix form. This feature is very useful to detect certain directions where variogram shows distinct characteristics. The*quadratic 'lowess' interpolant*has been used to create a smooth surface representation. Interpolation here gives better understanding of anisotropic variogram than plotting the original one. As we see below the behavior of the data for each direction is different (directions are given in degrees). (Fig.11).
- The programs ***VarFitModelm*** is written by the author and used for *Semivariogram Modeling* and visualization of results using the dataset (shown in Fig.8) as 2D irregularly spaced data. VarFitModel.m a Matlab function that performs a least squares fit of various theoretical variograms to an experimental, isotropic variogram. The user can choose between various ***bounded*** (e.g. spherical) and ***unbounded*** (e.g. exponential and power) models.

- A nugget variance can be modeled as well, but higher nested models are not supported. ***VarFitModel*** uses Matlab ***fminsearch*** function, but it should be used carefully, because the problem is, that it might return negative variances or ranges. *The variogram fitting algorithm is in particular sensitive to initial values below the optimal solution. Hence, visually inspecting the data and estimating a theoretical variogram by hand should always be your first choice*. Note that for unbounded models, the supplied parameter a0 (range) is the distance where gamma equals 95% of the sill variance. The returned parameter a0, however, is the parameter r in the model. The range at 95% of the sill variance is then approximately 3*r.

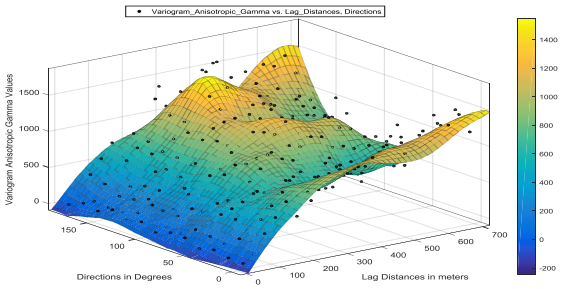
- 

Fig.11 Anisotropic Variogram (γ values vs. lag Distances and Directions in degrees). As we see that Variability behavior in the East-West Direction is completely different from that behavior in North-South Direction

- Ten most popular Fitting models types that have been analyzed and fitted to the empirical VariogramData. Theyare:***Spherical,Gaussian, Exponential,Circular, Hole Effect (Wave), Pentaspherical, Rational Quadratic,, Power,K-Bessel (Wittle) and.BLinear*** model gives strange output for ***Kriging*** so it is avoided. Below 10 Models (Figures 12,…,20) have been fitted with the empirical variogram data using program ***EstiamteVariogram.m*** **.** Table.1 shows the summary of semivariogram results**.**

Table.1 Summary of the Semivariogram Analysis Results

Model	*Sill*	*Range*	*RMS*	*Type*
Spheical	961.4	365.2	64.37	*Bounded*
Gaussian	963.7	306.5	70.61	*Unbounded*
Exponential	998.6	307.3	103.58	*Unbounded*
Circular	954.8	316.8	60.57	*Bounded*
Hole Effect	965.3	448.0	71.53	*Unbounded*
Pentaspherical	969.1	421.4	70.32	*Bounded*
Rational Quadratic	982.1	528.0	85.33	*Unbounded*
Power	1159.0	376.1	135.19	*Unbounded*
Wittle	975.6	480.0	81.47	*Unbounded*
Blinear	950.3	268.2	56.26	*Bounded*

Model parameters (shownon each of the figures)are: model *Name*, model *Type*, *Sill*, *Range*, *RMS* which reflects the goodness of fit, and model function.

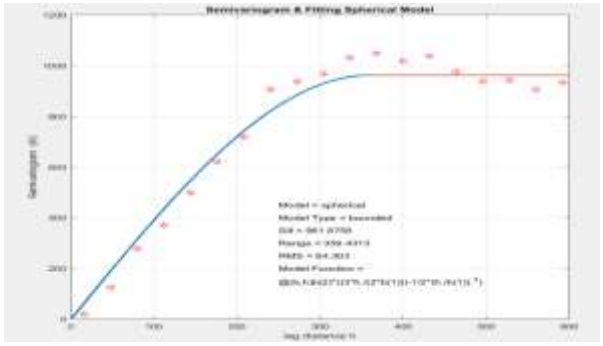


FIG.12 Fitting Spherical Model

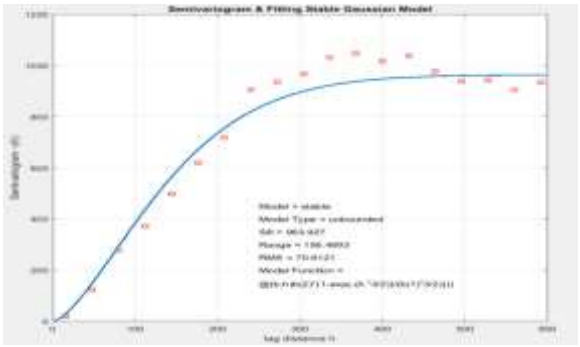


Fig.13 Fitting Gaussian Model

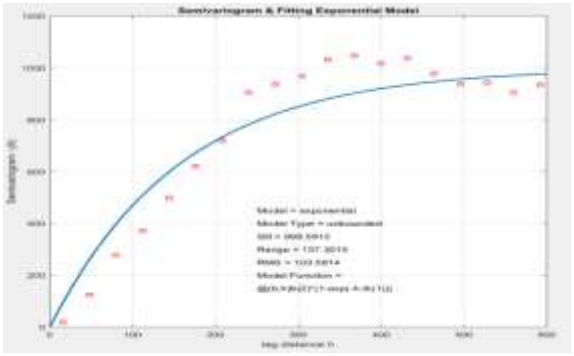


Fig.14 Fitting Exponential Model

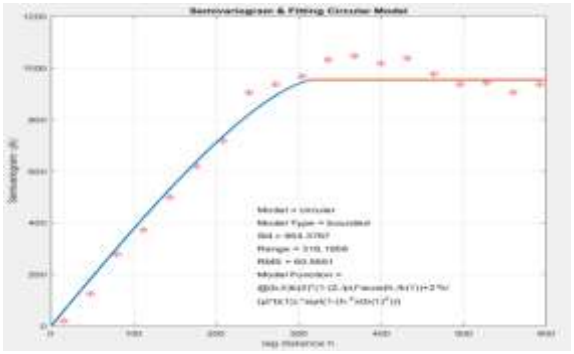


Fig.15 Fitting Circular Model

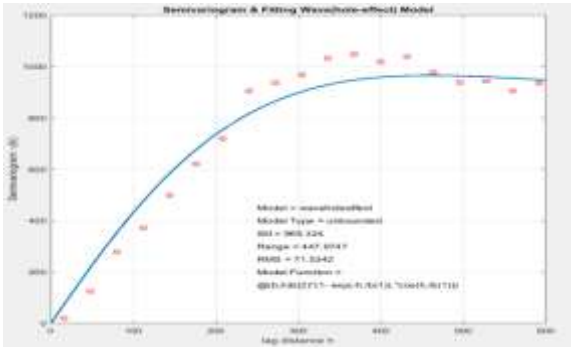


Fig.16Fitting Hole Effect (wave) Model

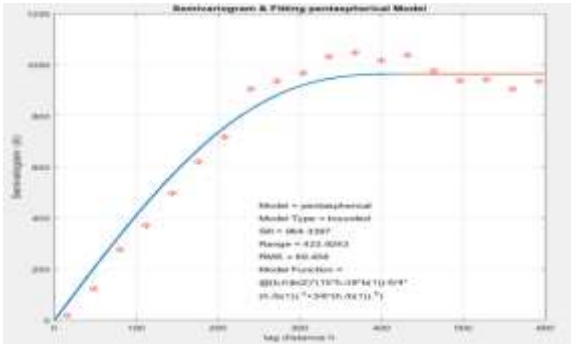


Fig.17 Fitting Pentaspherical Model

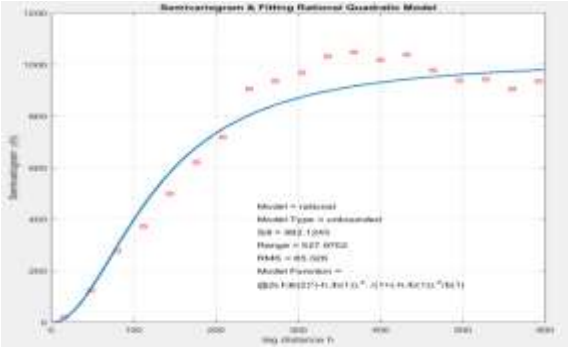


Fig.18 Fitting Rational Quadratic Model

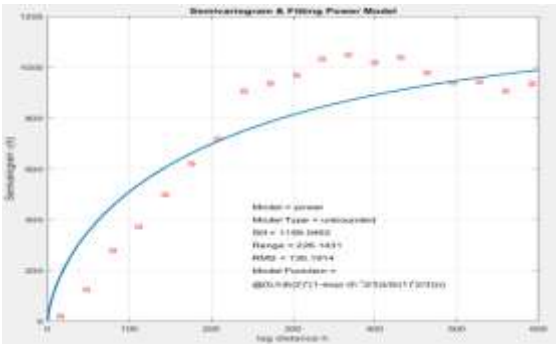


Fig.19 Fitting Power Model

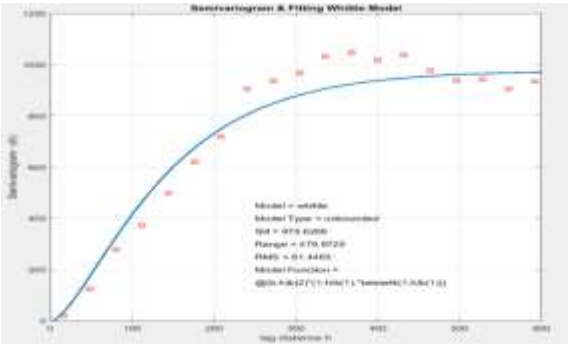


Fig.20 Fitting Whittle Model

Implementing Kriging Interpolation

Krig Interpolate.mis a Matlab function written by the author, which uses ordinary kriging to interpolate a variable ***z*** measured at a locations (*x, y*) at unsampled locations (*xi, yi*).The function requires the variable ***vstruct*** that contains allnecessary information on the variogram. vstruct is the main output argument of the previous function ***VarFitModel.m***. The function always includes all observations to estimate values at unsampled locations. This may not be necessary when sample locations are not within the autocorrelation range but would require something like a ***k nearest neighbor*** search algorithm. Thus, the algorithms works best for relatively small numbers of observations (100-500).Note that kriging fails if there are two or more observations at the same location.

Input arguments:

vstruct structure array with variogram information as returned from***VarFitModel.m***function.

x,y coordinates of observations

z values of observations

xi,yi coordinates of locations for predictions

chunksize number of elements in **zi** that are processed at one time. The default is 100, but this depends largely on the available main memory and numel(x).

Output arguments:

zi kriging predictions, ***zivar*** kriging variance

Summary of Resultsand Conclusions

Table.1 shows the summary of semivariogram results. From the table we see that the average ***Sill*** is 988, the minimum is 950.3 (*Blinear* model) and the maximum is 998.6(Power model is excluded). The average ***Range***is 381.7 the minimum is 268.2 (Blinear model) and the maximum is 528.0 (*Rational Quadratic* model). The average ***RMS***is 79.9 (which indicates goodness of fit) fluctuates between 56.3 and 135.2, thus *Exponential* and *Power* models are not among the best fitting models.

The Final digital terrain model (DTM)is generated (as contour lines) by above programmed *Kriging Function* (***krigInterpolate.m***) and is visualized by a *Matlab Function* (***contourf.m***) **(Figure No.21,.., No.26)**.

As we see from the figure that the performance of seven models (namely: *Spherical, Gaussian, Exponential, Circular, Hole Effect, Pentaspherical*) was very good in spite of some differences. *Rational Quadratic model* produced bad result and

unrealisticcontoursshape. *Wittle and Blinear* models produces very unrealistic results and sometimes the Kriging process fails without results. Matlab gives an error due to the pseudo-inverse of the kriging matrix cannot be executed. Finally *Circular, Exponential and Power* models generated some artifacts.

Figures (No.27,..,No. 30) represent ***Kriging Variance***for several models (namely, *Spherical, Gaussian, Exponential, Circular, Hole Effect, Pentaspherical and Rational Q.*).

We notice that *Gaussian, Circular and Ratinal Q*. models have been produced small Variances.

Finally, one thing has to be considered that the semivariogram is estimated from the data available (in our case the terrain data), it is describing the variability of a spatial process. So even though aparticular model is deemed best for a particular data set by a statistical comparison,it may not be the best choice. For example, the Gaussian model is often selected asbest with automatic fitting criterion, but it also corresponds to a process that is of tenunrealistically smooth. Ultimately, the final choice of model should reflect both the results of the statistical model fitting procedure and an interpretation consistent with the scientific understanding of the process being studied.

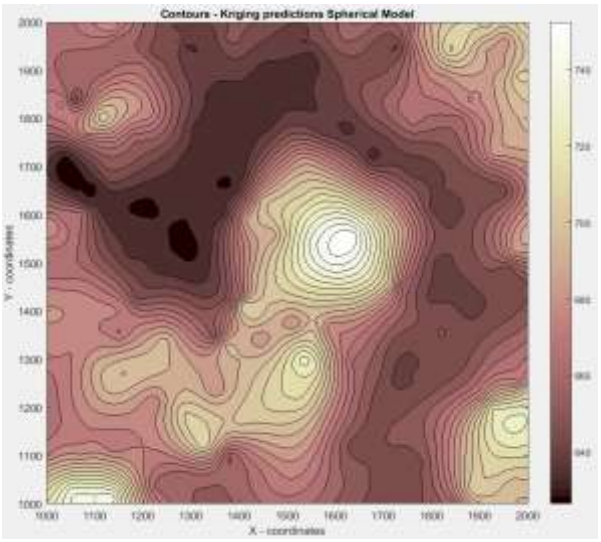


Fig.21 Contours representation of Kriging (*Spherical model*)

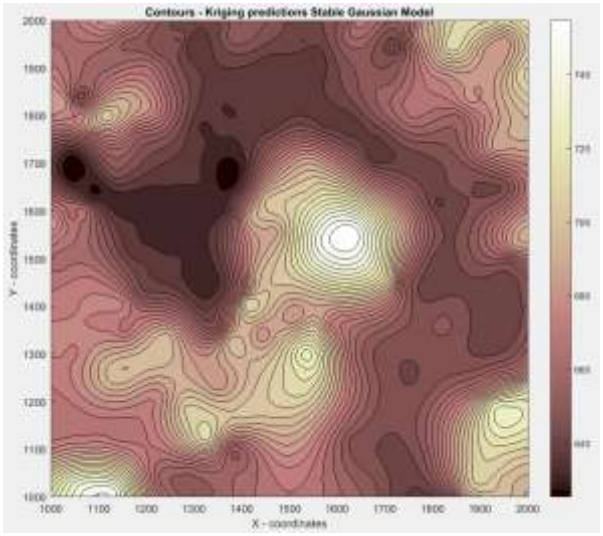


Fig.22 Contours representation ofKriging (*Gaussian Model*)

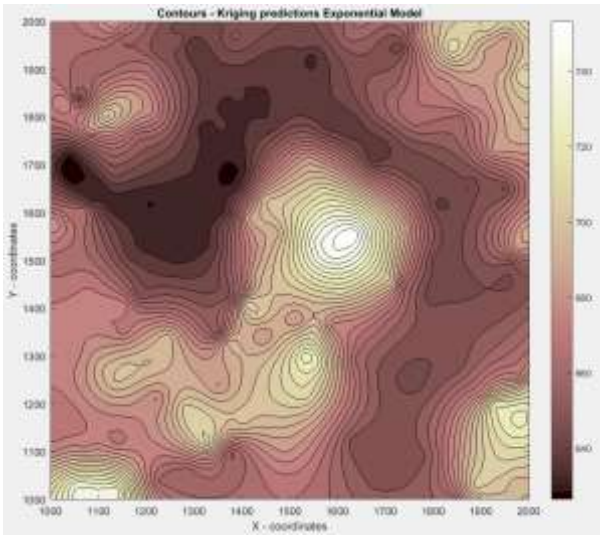


Fig.23Contours representation of Kriging (*Exponential Model*)

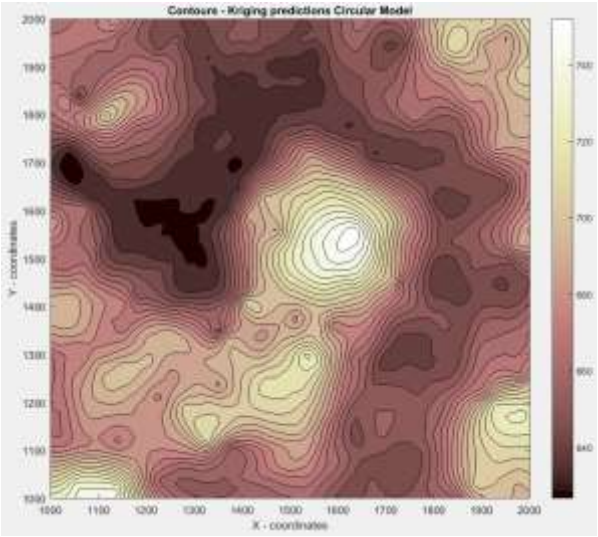


Fig.24Kriging Contours (*Circular Model*)

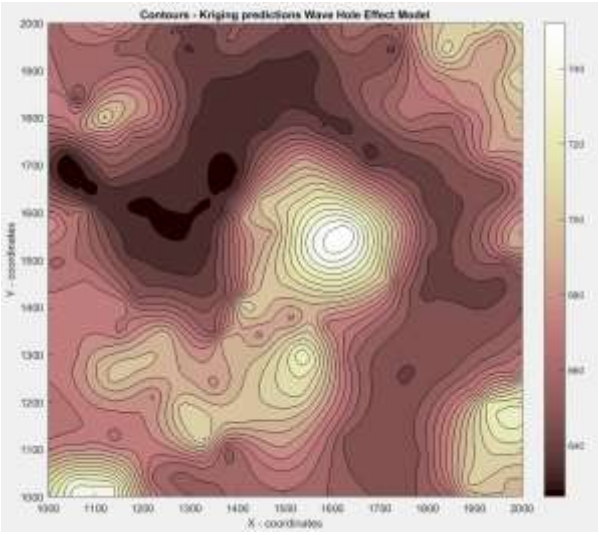


Fig.26 Kriging Contours (*Hole Effect Model*)

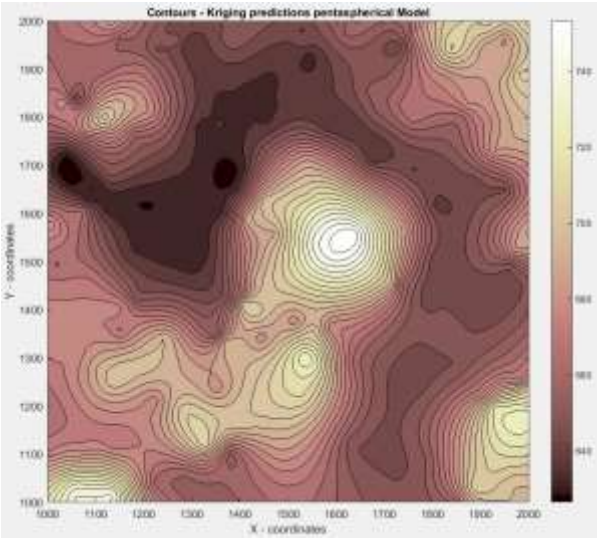


Fig.25Kriging Contours (*Pentaspherical Model*)

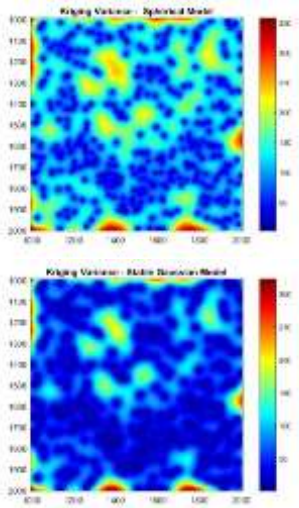


Fig.27 Kriging Variance -Spherical Model (left), Gauss Model (right)

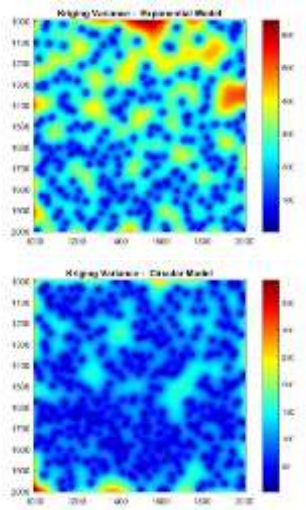


Fig.28 Kriging Variance-Exponential Model (left) Circular model (right)

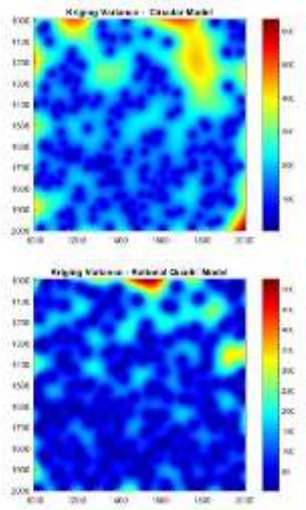


Fig.29 Kriging Variance – Hole Effect model (left), Pentaspherical model (right)

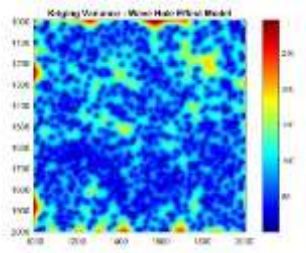


Fig.30 Kriging Variance - RationalQuadratic model

REFERENCES

Armstrong M. (1998) Basic linear geostatistics. Springer.

Burrough, P. A. and McDonnell, R. (1998) Principles of Geographical Information Systems, 2nd ed. (Oxford: Clarendon Press).

Deutsch CV, Journel AG (1998) GSLIB: Geostatistical software library and user's guide, 2nd edn. Oxford University Press, New York, 369 p

Chiles, J-P. and Delfiner, P. (1999) Geostatistics: Modeling Spatial Uncertainty(New York: Wiley). 720 p.

Cressie N (1993) Statistics for spatial data. Wiley, New York.

Davis, B. M. and L. E. Borgman (1979). Some exact sampling distributions for variogramestimators. Journal of the International Association for Mathematical Geology 11.

Davis, G. and M. D. Morris (1997). Six factors which affect the condition number of matrices associated with kriging. Mathematical Geology 29, 91–98.

Davis, J. C. (2002) Statistics and Data Analysis in Geology (Hoboken, NJ: Wiley).

De Floriani, L. and Magillo, P. (1994) Visibility algorithms on triangulatedterrain models. International Journal of Geographical Information Systems,8 (1): 13–41.

DeMers, M. (2001) GIS Modeling in Raster (New York: Wiley).

Deutsch, C. V. and A. G. Journel (1992). GSLIB: Geostatistical Software Library and User's Guide. New York: Oxford University Press.

Diggle, P. J., J. A. Tawn, and R. A. Moyeed (1998). Model-based geostatistics. Applied Statistics 47, 229–350.

Dykes, J. (1996). Dynamic maps for spatial science: A unified approach to cartographic visualization. In D. Parker (Ed.), Innovations in GIS 3, pp. 177-188. Taylor& Francis.

Ecker, M. D. and A. E. Gelf and (1999). Bayesian modeling and inference for geometrical lyanisotropic spatia data. Mathematical Geology 31, 67–83.

Goovaerts P. (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York, 483 p

Goovaerts P (2005) Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging.

Gotway CA, Young LJ (2002) Combining incompatible spatial data. J Am Stat Assoc 97(459):632–648.

Harvey, F. (2008) A Primer of GIS: Fundamental Geographic and Cartographic Concepts (New York).

Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York.

Journel AG, Huijbregts CJ (1978) Mining geostatistics. Academic Press, London

Journel AG (1983) Non-parametric estimation of spatial distributions. J Int Ass Math Geol. 5(3):445-468

Kolmogorov AN (1941) The local structure of turbulence in an incompressible fluid at verylarge Reynolds numbers.

Kyriakidis P (2004) A geostatistical framework for area-to-point spatial interpolation. Geogr Anal 36(2):259–289.

Li, X. and Hodgson, M. E. (2004) Vector field data model and operations. GIScience and Remote Sensing, 41 (1).

Matérn B (1966) Spatial variation: Stochastic models and their applications to problems inforest surveys and other sampling investigations.

Matheron, G. (1963) Principles of geostatistics. Economic Geology, 58: 1246–1266.

Mercer WB, Hall AD (1991) The experimental error of field trials. J A gricult Sci 4:107-132

Oliver M (2009) The variogram and kriging. In Fischer MM, Getis A (eds) Handbook ofapplied spatial analysis. Springer, Berlin, Heidelberg and New York, pp.319-352

Oliver MA, Webster R, Slocum K (2000) Filtering SPOT imagery by kriging analysis. Int. JRemote Sensing 21(4)

Ripley BD (1977) Modeling spatial patterns. J Roy Stat 39(2)

Rhind, D. W. (1971) Automated contouring: an empirical evaluation of some differing techniques. Cartographic Journal, 8: 145–158.

Stein ML (1999) Interpolation of spatial data: some theory for kriging. Springer, Berlin, Heidelberg and New York.

Tomlin, D. (1990) Geographical Information Systems and Cartographic Modeling (Englewood Cliffs, NJ: Prentice H)

Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading [MA]

Wackernagel H (2003) Multivariate geostatistics (3rd edition). Springer, Berlin, Heidelberg and New York.

Walvoort, D. J. J. (2004) E(Z)-Kriging: Exploring the World of Ordinary Kriging (Wageningen: Wageningen University & Research Centre).

Webster R, Oliver MA (1992) Sample adequately to estimate variograms of soil properties. J Soil Sci 43(1)

Webster R, Oliver MA (2007) Geostatistics for environmental scientists (2nd edition). Wiley, New York.

Wheeler D, Paéz A (2009) Geographically weighted regression. In Fischer MM, Getis A (eds) Handbook of applied spatial analysis. Springer, Berlin.

Received	**2017/01/04**	**إيداع البحث**
Accepted for Publ.	**2017/02/23**	**قبول البحث للنشر**