Gregg Babish
GIS Specialist

Ecological Research Division
Environmental Conservation Branch
2365 Albert Street, Room 300
Regina, Saskatchewan S4P 4K1

Edition 2006.03

July 5, 2006

Environment Canada

Geostatistics Without Tears

A practical Guide to Surface Interpolation, Geostatistics, Variograms and Kriging

Gregg Babish
GIS Specialist
Environment Canada
Regina, Saskatchewan

GEOSTATISTICS WITHOUT TEARS

A Practical Guide to Surface Interpolation, Geostatistics, Variograms and Kriging

by Gregg Babish
GIS Supervisor / Data Management
Environment Canada
Illustrations by Mark Gilchrist
Environment Canada

PREFACE

Geostatistics (also known as kriging) was developed for the mining industry during the 1960s and 1970s to estimate changes in ore grade. The principles of geostatistics are now applied to many applications that require statistically based interpolation techniques. Geostatistics provides a data value estimate for locations that cannot be sampled directly by examining data taken at locations that can be sampled. If a model of spatial correlation can be established for the sampled data then this model is used to interpolate data values for those locations that cannot be sampled directly. The effectiveness of geostatistics depends upon how well the observation points represent the phenomenon under study as well as the appropriate selection of the model parameters.

Matheron introduced the term geostatistics in 1962 with this definition: "Geostatistics is the application of the formalism of random functions to the reconnaissance and estimation of natural phenomena" (Journel and Huijbregts, 1978). Oliver and Webster (1991) said, ". . . *there is to be no geostatistics without tears"*.

The intent of this paper is to provide the reader with an introduction and understanding of surface interpolation techniques and geostatistics (kriging). Hopefully no tears will be shed.

INTRODUCTION

Surface models are used for a variety of purposes including interpolating between actual data measurements, identifying data anomalies and establishing confidence intervals around predictions. The principles of surface analysis are applied in a number of applications such as climatic measurements, contamination plumes, and population densities. The interpolated surfaces are then used to examine weather change, contaminant distribution and population change. The accuracy of subsequent analyses depends directly on the accuracy of the surfaces created.

Data to create these surfaces are usually collected through field work. As a rule, data collection can only be conducted at a restricted number of point locations due to limited resources and high costs. In order to generate a continuous surface some type of interpolation method must be used to estimate data values for those locations where no samples or measurements were taken.

INTERPOLATION TECHNIQUES

Surface interpolation is the process of estimating a data value for unsampled sites within an area of existing data observations. The best surface interpolation method to use will depend on the data, the study objectives, and the ease of creating a surface versus the accuracy of the surface created.

In order to make spatial predictions there is a set of essential assumptions:

- the measurements taken are precise and reproducible
- the sample measurements are accurate and represent the true value at that location
- the samples are collected from a physically continuous, homogeneous population
- the values at unsampled locations are related to values at sampled locations

All interpolation methods depend on the similarity of nearby sample locations to derive the surface. Spatial independence means that the location of any data point is unrelated to the location of any other data point and therefore there is no spatial relationship between sampling locations. Spatial autocorrelation is the arrangement of data where the point locations are related to each other. The concept of spatial autocorrelation is one of the most important aspects of spatial statistics. Spatial autocorrelation is a normal result of physical and chemical processes in the environment and environmental parameters will exhibit spatial autocorrelation at some scale (Englund; Rigsby et al. 2001). Many social phenomena are spatially autocorrelated (Ned Levine & Associates, 1999).

Spatial Interpolation

Deterministic/Stochastic

Deterministic techniques are based on surrounding measurements (mathematical functions) to calculate the surface. These techniques are based on the measured values of a parameter at samples near the unmeasured location for which a prediction is made. For deterministic interpolation there is only one possible predicted value for a location. Thiessen polygons, IDW and spline interpolation are deterministic techniques

Stochastic techniques use both mathematical and statistical functions for prediction. The first step for such methods is to determine the statistical relationship between samples points. This information is then used to predict parameter values for unmeasured points. Stochastic interpolation incorporates the concept of randomness (through autocorrelation). Stochastic techniques are advantageous because their statistical basis allows for quantification of error. Polynomial regression and kriging are stochastic interpolation methods. Kriging attempts to statistically obtain the optimal prediction to provide the Best Linear Unbiased Estimation (BLUE), specifically when data are sparse.

Global/Local

A global interpolator derives a surface model using all of the available data to provide predictions for the whole area of interest by applying a single function across the entire region. The resulting surface gives a best fit for the entire sample data but may provide a very poor fit at some locations. There are a number of disadvantages to a global fit procedure. The most obvious is that global interpolation methods don't capture short-range variation. A change in one input value can affect the entire map. Global algorithms tend to produce smoother surfaces with less abrupt changes.

A local interpolator calculates predictions from the measured points within neighborhoods or smaller spatial areas within a larger study area to ensure that interpolated values are determined only by nearby points. Local interpolation applies an algorithm to a small portion at a time. If the interpolated plane can be bent in one place, it may be possible to obtain a better overall fit. To allow one bend is the basis for second-order global interpolation, two bends in the plane would be a third order and so forth (polynomial trend surface). If a number of planes are used for predicting many locations in a study area the final surface may fit more accurately. A change in one input value only affects results within a small area.

IDW and kriging are examples of local interpolation. Kriging interpolation depends on whether or not there is a trend in the data and probabilistic/stochastic variation or deviation from a trend (spatially autocorrelated random error) in addition to uncorrelated random noise.

Exact/Approximate

Exact interpolation honours the data points upon which the interpolation is based so that the interpolated surface passes through all points whose values are known. Kriging methods honor the sample data points, but kriging may incorporate a nugget effect (random noise) and in this case the concept of an exact interpolator ceases to be appropriate.

Approximate interpolation is used when there is uncertainty in the data values. In many data sets there are global trends and local variation that produces uncertainty (error) in the sample values. Approximate interpolation introduces smoothing so as to reduce the effects of error on the resulting surface.

Gradual/Abrupt

Gradual interpolation produces a surface with gradual changes by applying the same rules over the entire data source. Gradual interpolation is appropriate for interpolating data with low local variability.

Abrupt interpolation can involve the application of different rules at different points by including 'barriers' in the interpolation process. Semipermeable barriers, such as weather fronts, will produce quickly changing but continuous values. Impermeable barriers, such as a geologic fault, will produce abrupt changes. Abrupt interpolation produces surfaces with a stepped appearance and is appropriate for interpolating data of high local variability or data with discontinuities.

Climate Data

Collection of meteorological data requires specialized instrumentation collecting data over many years and is subject to siting, data collection errors and biases. Climate data is available only at a relatively small number of observed points. Given a set of meteorological data, researchers are confronted with a variety of interpolation methods to estimate meteorological variables at unsampled locations. Depending on the spatial attributes of the data the accuracy of the interpolated surface can vary widely among different methods (MacEachren and Davidson, 1987; and Rhind, 1975).

Climate interpolation models provide few estimates of error because these estimates rely on the same assumptions used in the interpolation process itself, and are therefore, not independent or reliable (Daly, 2006). Experts are the most useful source of verification because they can integrate information from disparate sources to give justifiable feedback on how well the model results reproduce their knowledge from a variety of perspectives. This feedback can be in the form of evaluation of the spatial patterns and magnitudes of the mapped climate values, as well as insight into station data quality issues (Daly and Johnson, 1999).

The interpolation of irregular meteorological-climatological point data onto a uniform grid has been the focus of research and a number of methods have been proposed ranging from simple Thiessen and distance weighting methods (Shepard, 1968; Willmott et al., 1985), to geostatistical methods such as kriging

(Phillips et al., 1992), splines (Hutchinson, 1995) and locally-varying regression techniques such as PRISM (Daly et al., 1994).

Kriging has been used for interpolation of precipitation data (Chua and Bras, 1982; Dingman et al., 1988; Phillips et al., 1992). Meteorologists sometimes consider that a potential drawback of kriging is that kriging implicitly relies on the data to directly represent the spatial variability of the actual precipitation field. If the data are not representative (as is often the case in complex terrain, such as mountainous regions), the accuracy of the resulting interpolated field will be in question (Daly et al. 1993).

The choice of spatial interpolator is especially important in mountainous regions where data collections are sparse and variables may change over short spatial scales (Collins and Bolstad, 2006). For many climate applications it is important that elevation is included as a covariate or independent variable because the climate variable is dependent on elevation in some manner (Willmott and Mastsuura, 1995; Briggs and Cogley, 1996; New and Hulme, 1997). Goovaerts (2000) confirmed the finding of Creutin and Obled (1982) that for low-density networks of rain gauges geostatistical interpolation outperforms techniques that ignore the pattern of spatial dependence which is usually observed for rainfall data. Prediction can be further improved if correlated secondary information, such as a DEM is taken into account.

Ordinary kriging has been applied for developing contour maps of design storm depth using intensity-duration-frequency (IDF) data (Cheng, et al., 2003). The design storm, a crucial element in urban design and hydrological modeling, is a hypothetical storm of specific storm duration, and recurrence interval. Variogram parameters, the sill and range are functions of the recurrence interval and the storm duration. The sill accounts for the time non-stationarity of the rainfall field, and design storms with higher total rainfall depths have higher sill values.

Some of the variations of meteorological interpolations are discussed below.

ANUSPLIN

ANUSPLIN software developed in the late 1980's by Michael Hutchinson (Hutchinson, 1991 and 1994), Australian National University (ANU), Centre for Resource and Environmental Studies (CRES), Canberra is based on the original thin-plate surface fitting technique (Wahba, 1979 and 1990) for mapping climate variables, especially for the interpolation of rainfall amounts (Hutchinson, 1995; 1998a; 1998b). ANUSPLIN fits thin-plate splines (usually second- or third-order polynomials) through station data in three dimensions: latitude, longitude, and elevation (Hutchinson, 1995). Booth and Jones (1996) conclude that ANUSPLIN provides a powerful set of programs for climatic analysis. ANUSPLIN has been used to develop globally consistent 30" climate surfaces covering most areas on the Earth's surface (Hijmans et al. 2004 a,b).

ANUSPLIN was specifically developed for interpolating climate data and is made up of nine programs that incorporate additional dependencies (the concept of "surface independent variables") such as elevation (Kesterven and Hutchinson, 1996). Despite having the capability of including many more dependencies ANUSPLIN routinely uses only three variables in practical applications. Being able to accommodate several covariates to aid interpolation is a significant advantage. A disadvantage of the method is ANUSPLIN's reliance on a dense network of stations covering all aspects of the topography being mapped. This is particularly problematic for the interpolation of rainfall, which is often underestimated at high elevations (Tait and Zheng, 2005). Since ANUSPLIN places a greater reliance a DEM than other methods ANUSPLIN is more likely to provide more robust surfaces, but that robustness is dependent on the accuracy of the DEM, and may results in problems where relief is subtle (Chapman et al., 2005). Because a spline is by definition smoothly varying, this approach has difficulty simulating sharply varying climate transitions, which are characteristic of temperature inversions, rain shadows and coastal effect (Daly, 2006).

ANUSPLIN produces results similar to kriging (Delfiner and Delhomme, 1975; Hutchinson, 1991b and 1993; Hutchinson and Gessler, 1994; Wahba and Wendelberger, 1980; Wahba, 1990; and Cressie, 2003). Unlike kriging the thin-plate spline method does not require development of a covariance function (variogram). Instead the degree of smoothing is optimized objectively by minimizing the predictive error of the fitted function as measured by generalized cross validation (Milewska et al., 2005; Tait and Zheng, 2005).

Centro Internacional de Agricultura Tropical

The Centro Internacional de Agricultura Tropical (the International Centre for Tropical Agriculture) (CIAT, http://www.ciat.cgiar.org/), Cali, Columbia, method uses a simple interpolation algorithm based on the inverse square of the distance between the station and the interpolated point of the nearest five stations (Jones et al. 1990, Jones, 1995).

The major difference between CIAT and ANUSPLIN is that the CIAT method uses a standard lapse rate applied over the whole dataset whereas ANUSPLIN uses a 3-dimensional spline algorithm to determine a local lapse rate from the data. The CIAT method introduces error when the local lapse rate deviates from the standard lapse rate function. ANUSPLIN suffers when there are erroneous data or insufficient data range in the local area, resulting in spurious correction for elevation (Booth and Jones, 1996). CIAT has the advantage of speed and ease of use for large data sets where computational capacity is limited. The influence of a bad data point can be significant and can cause significant circling in the resultant surface. Because CIAT uses only five data points, it relies less on the underlying DEM than other methods (Booth and Jones, 1996).

Climatologically Aided Interpolation

Climatologically Aided Interpolation (CAI) is a hybrid approach (incorporating elements of temporal substitution with distance weighted interpolation) that uses existing spatial climate data to improve the interpolation of another set of data (Willmott and Robeson, 1995; Robeson and Janis, 1998). CAI produces low validation errors, and its accuracy is attributed in part to the incorporation of terrain effects provided by the high resolution climatology (Thornton, et. al., 1997). CAI relies on the assumption that local spatial patterns of the element being interpolated closely resemble those of the existing climate grid (called the background or predictor grid). This method is useful for interpolating climate variables and time periods for which station data may be relatively sparse or intermittent.

Use of CAI fall into two broad categories (Daly, 2006):

1. using a long-term mean grid of a climate element to aid the interpolation of the same element over a different (usually shorter) averaging period; and
2. using a grid of a climate element to aid the interpolation of a different, but related, climate element (e.g. interpolating growing degree days using mean temperature as the predictor grid).

Daymet

Daymet (http://www.daymet.org/) is similar to PRISM in that both methods use local regression techniques, but Daymet and PRISM are different in their formulation. Daymet was developed at the University of Montana, Numerical Terradynamic Simulation Group (NTSG) to fulfill the need for fine resolution, daily meteorological and climatological data necessary for plant growth model inputs. Using a digital elevation model and daily observations of minimum/maximum temperatures and precipitation from ground-based meteorological stations, an 18 year daily data set (1980-1997) of temperature, precipitation, humidity and radiation have been produced as a continuous surface at a 1 km resolution for the United States (http://www.daymet.org/dataSelection.jsp).

Daymet develops local linear regressions between climate and elevation for each grid cell on a digital elevation model using data from surrounding stations. Each station is weighted in the regression function by its distance from the target grid cell. This method takes into account the elevation variation of climate and a simple station distance weighting algorithm. Daymet does not have the ability to simulate nonmonotonic relationships between climate and elevation, such as temperature inversions, and does not explicitly account for terrain-induced climatic transitions or coastal effects (Daly, 2006).

Gradient plus Inverse Distance Square

Gradient plus inverse distance square (GIDS) method has been used in various parts of the world to produce useable surfaces for meteorological parameters. GIDS combines multiple linear regression with distance weighting.

A comparison of GIDS and ANUSPLIN concluded that ANUSPLIN provides generally superior results. Both subjective assessment and statistical analysis showed that ANUSPLINE is generally more accurate in predicting climate variables. In addition ANUSPLIN produces better smoothing and better gradients at high elevations and in areas where climate station coverage was poor (Price et al., 2000; Chapman, 2003). GIDS is easy to implement and understand which can provide a useful baseline to compare with more sophisticated methods (Price et al., 2000).

Optimal Interpolation

Geostatistical techniques were originally developed by Soviet scientists for meteorological data predictions. The first book with complete explanations about simple and ordinary kriging and cokriging techniques was published in Leningrad (Gandin, 1963). According to this book, the original name of the technique is objective analysis. In atmospheric and oceanographic science this technique is known as optimal interpolation (OI). In spatial statistics this technique is known as simple kriging (Matheron, 1963). The multivariate analysis of meteorological data using an optimal interpolation method is presented in Gandin, 1963; Gandin and Kagan, 1974; Bretherton et al., 1976; Bretherton and Williams, 1980; Sarmiento et al., 1982; Hiller and Käse, 1983; Swanson, et. al., 2001; AWI, 2006.

The advantage of optimal interpolation is the simplicity of implementation and its relatively small cost if the right assumptions can be made on the selection of observation data. A drawback of optimal interpolation is that spurious noise is produced in the analysis fields because different sets of observations (and possibly different background error models) are used on different parts of the model state (Bouttier and Courtier, 1999). It is also impossible to guarantee the coherence between small and large scales of the analysis (Lorenc, 1981).

Parameter-evaluation Regressions on Independent Slopes Model

Parameter-evaluation Regressions on Independent Slopes Model (PRISM), (http://www.ocs.orst.edu/prism/) is an expert system developed at Oregon State University in the early 1990's for orographic precipitation estimates (Daly, 1996). PRISM uses point data and DEM data to generate gridded estimates of climate parameters (Daly et al., 1994; Taylor et al., 1997). PRISM is suited for use in mountainous regions because data relationships can be extrapolated beyond the lowest and highest station elevation. The use of PRISM has been extended to map temperature, snowfall, weather generator statistics, and more. Climate layers for the United States, western Canada and other a few other countries, have been developed using PRISM and are available on-line http://www.climatesource.com/ (Daly et al., 1994 and Gibson et al. 2004).

PRISM develops local climate-elevation regression functions for each DEM grid cell, but calculates station weights on the basis of an extensive spatial climate knowledge base that assesses each station's physiographic similarity to the target grid cell (Daly et al., 2002, 2003). The knowledge base and resulting station weighting functions currently account for spatial variations in climate caused by elevation, terrain orientation, effectiveness of terrain as a barrier to flow, coastal proximity, moisture availability, a two-layer atmosphere (to handle inversions), and topographic position (valley, midslope, ridge).

While PRISM accounts for more spatial climate factors than other methods, it also requires more effort, expertise and supporting data to take advantage of its full capability (Daly 2006). The effects of other variables, such as slope, aspect, coastal proximity, the influence of the boundary layer, etc. are controlled by the use of special algorithms, parameters and weights. Some of these values are set as defaults based on the general knowledge of physiographic and atmospheric processes; some are inferred by the model from station data; other are assigned manually by an expert climatologist through the user interface (Daly et al., 2002).

Although PRISM incorporates a number of spatial interpolation quality control measures in similar ways to ANUSPLIN, the PRISM and ANUSPLIN methods are fundamentally different (Simpson, et al., 2005). PRISM uses a two-layer atmosphere to model the effects of atmospheric inversions on surface temperature whereas ANUSPLIN cannot. PRISM analysis typically extends beyond the coastline to include near-shore areas, whereas ANUSPLINE analysis is largely restricted to land areas.

Thin-Plate Spline

The thin-plate spline (TPS) method refers to a physical analogy involving the bending of a thin sheet of metal constrained not to move at the sample points and is free from any external force. The resulting shape of the plate represents the natural bend of the metal to accommodate all of the fixed sample points. The use of thin-plate splines has been found useful for a number of applications including DEM construction, fingerprint modeling, image manipulation, geophysical applications, medical research, and rubber sheeting. Thin-plate splines have found application in meteorology (Hutchinson and Bischof, 1983; Hutchinson et al., 1984; Kalma et al., 1988; Hutchinson, 1990; Mitchell, 1991; Zheng and Basher, 1995; Chapman et al., 2005).

The local thin-plate spline method is an extension of the thin-plate Spline interpolation technique, and is recommended for use with a large number of grid points (>200). The only difference is that instead of using all the grid points for interpolation, the local version takes a maximum of 10 closest points to the sample point and fits a spline surface through them. The local spline surface is then used to determine the sample value (RocScience, 2006)

The key features of the thin-plate smoothing spline analysis are robustness, simplicity and the advantage of being able to map sparse and noisy data. One of the main shortcomings of the thin-plate spline method has been the limited range of practical options for user control of the fit. The thin-plate spline method can be problematic if sample points are sparse. Small changes in sample position could produce global changes in the warping transformation (Glasbey and Mardia, 1998). The thin-plate spline is equivalent to kriging with a specific spatial covariance structure (Kent and Mardia, 1994).

Digital Elevation Modeling

There are two forms of Digital Elevation Model (DEM): Grid and Triangular Irregular Network (TIN) with their own advantages and disadvantages (Cadell, 2002). A number of algorithms and computer programs have been developed to create a DEM.

ESRI’s ARCTIN method makes a Triangular Irregular Network (TIN) and then converts the TIN to a DEM. ANUDEM and ESRIs ArcInfo TOPOGRID are two common spline-based methods of obtaining a reasonable interpolation between spot elevations, and elevation contours while maintaining stream connectivity. ANUDEM was developed at the Centre for Resource and Environmental Studies (CRES) in Canberra (Hutchinson, 1988, 1989). A version of ANUDEM is included in ArcInfo as TOPOGRID.

ANUDEM preserves ridge and valley lines as well as maintaining drainage (Hutchinson, 2006). Input data to ANUDEM may include point elevations, elevation contours, streamlines, sink data points, cliff lines, boundary polygons, lake boundaries and data mask polygons. TOPOGRID is based on an older version of ANUDEM. ANUDEM/TOPOGRID use iterative interpolation starting from a very coarse model to end up at a user-specified grid spacing (Jaakkola and Oksanen, 2000). The strength of the ANUDEM method over most other methods is that ANUDEM imposes a global drainage condition through an approach known as drainage enforcement, to produce elevation models that represent more closely the actual terrain surface and which contain fewer artifacts than those produced with more general-purpose surface interpolation routines (USGS 2003).

Care must be taken with spline-based techniques, such as ANUDEM or TOPOGRID because certain topography (such as steep slope close to flat terrain) can cause undershoot and overshoot errors. Biasing towards contour elevations can occur and is sometimes quite noticeable.

When there is interest in maintaining an assumed network of hydrology then methods such as ANUDEM/TOPOGRID probably give a more acceptable result than the TIN method. This advantage tends to be more obvious in low-relief topography than in topography where the hydrology is constrained by strong slopes. If maintaining the hydrological networks isn't important then ANUDEM/TOPOGRID might not be the best answer. ANUDEM can give an inferior results compared to a DEM generated when the hydrological constraints are withdrawn and a linear interpolator is used. Then again, a simple linear terrain interpolation might not be good for some visualization work since it introduces visual artifacts. Analysis of DEMs suggests that the quality of the input data from which the DEM is generated has a more significant effect on DEM quality than do the algorithms employed by different methods (Barringer and Lilburne, 1997). Techniques have been developed that use the results derived from Radarsat imagery in conjunction with ANUDEM to provide a DEM of much greater accuracy. This approach has been valuable in areas with little terrain variability, and has been used to create a DEM of the Antarctic (Liu et al. 2001).

Triangular models have been used in terrain modeling since the 1970s, and are sometimes used to create Digital Terrain Models (DTMs). Commercial systems using TIN began to appear in the 1980's as contouring packages. Due to limitations of computers and the complexity of TIN data structures, gridded models have overshadowed triangular models. Certain types of terrain are very effectively divided into triangles with plane facets. This is particularly true with fluvially eroded landscapes. However, other landscapes, such as glaciated ones, are not well represented by flat triangles. Triangles work best in areas with sharp breaks in slope where TIN edges can be aligned with breaks such as mountainous terrain, but triangular models aren't as suitable for representing prairie landscapes (Goodchild, 1999).

Fourier Series

The Fourier series approximates a surface by overlaying a series of sine and cosine waves. The Fourier series is best suited for data sets that exhibit marked periodicity, such as ocean waves. Fourier series is a high-order surface representation with triangular mesh application (Bruno and Pohlman, 2002, 2003).

Geostatistics

Interpolation methods such as spline, IDW, triangulation, trend surfaces and Fourier series are based on mathematical models (deterministic). These methods assume that the sampled data has no errors, which is often an incorrect assumption since errors can be minimized but not eliminated. The best representative map created with these techniques may look like they model the spatial process but this model may not provide a correct representation of the spatial process (Krivoruchko, 1998).

Geostatistics is based on random processes with dependence. A variable that takes on values according to its spatial location (or time) is known as a regionalized variable. Regionalized variables describe phenomena with geographical distribution, such as disease incidence, geophysical measurements, insect counts, soil/water contamination, and surface elevation. Regionalized variables recognize the fact that properties measured in space (or time) follow a pattern that cannot be described by a mathematical function. In the spatial or temporal context such dependence is called autocorrelation. Statistical dependence provides valuable information for prediction. Geostatistical estimation is a two stage process:

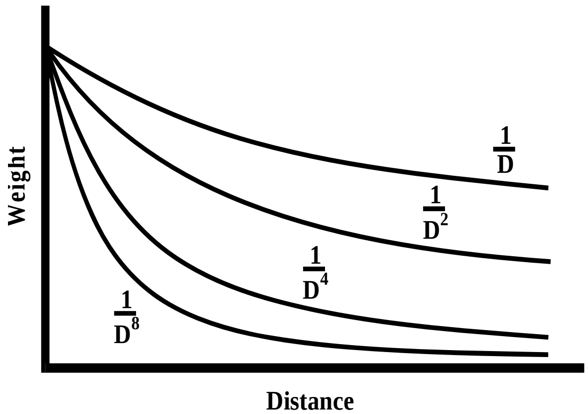
i. study the data to establish the predictability of values from place to place (uncover the dependency rules)

ii. interpolate values at those locations which have not been sampled based on the degree of spatial pattern (spatial autocorrelation). Kriging is the process of prediction of the unknown values

The results from kriging are generally of higher quality and have a more realistic look compared to techniques such as IDW and triangulation. The “bull’s eye” effect is avoided; a measure of error or uncertainty is provided; confidence placed in the estimates can be calculated; extreme weighting of values caused by irregular distribution of sample points is reduced; redundant data (clustering) and anisotropies is accounted for; and kriging can be used to determine where more data is needed if future sampling is planned.

The major disadvantage of the kriging process is that the construction of the model and evaluation of the surface interpolation results can be time consuming. Kriging is not a suitable method for data sets that have spikes, abrupt changes or break lines.

Inverse Distance Weighted



Inverse Distance Weighted (IDW) (also known as Inverse Distance to a Power) assumes that each coordinate (X,Y) and surface value (Z) has a local influence that diminishes with distance. Points closer to the processing cell are given more weight than those farther away. The power parameter in the IDW interpolation controls the significance of the surrounding points on the interpolated value. A higher power results in less influence from distant points.

Simplicity, speed of calculation, programming ease, and reasonable results for certain types of data are some of the advantages associated with IDW interpolation. The IDW method is based on the assumptions that the data values form a continuous surface across the whole area, and that the data is strongly correlated with distance. One potential advantage of IDW is the ability to control the influence of distance. IDW has been used to interpolate surfaces for consumer purchasing (distant locations have less influence because people are more likely to shop closer to home) and IDW is used for noise analysis since noise falls off very predictably with distance.

Although distance-weighted methods are one of the more commonly used interpolation approaches, they are far from ideal (Goodchild, 1999; Clark and Harper, 2000; Lembo, 2005). IDW doesn't work as well with phenomena that are dependent on complex variables because IDW can only account for the effects of distance. IDW works best with sample points that are dense and evenly distributed. IDW is easily affected by the uneven distribution of data points since an equal weight is assigned to each of the data points even if the data points are clustered. IDW is an averaging (smoothing) technique and cannot interpolate above or below the surrounding data which tends to generate flat areas. IDW has a tendency to generate patterns of concentric contours around the actual data points ("bull's eyes") and will average out trends and emphasize anomalies (outliers). For an elevation surface, this has the effect of flattening peaks and valleys (unless their high and low points are part of the sample). Smoother IDW surfaces can be created by decreasing the power, increasing the number of sample points used, or increasing the search radius – the opposite is done to create a more locally influenced surface (Zamkotowicz, 2005).

Shepard's Method (Shepard, 1968) is a variation on the inverse power with two different weighting functions (an inverse distance weighted least squares method, and the inverse distance to a power interpolator). Shepard's Method, while similar to IDW, eliminates or reduces "bull's eye" patterns, especially when a smoothing factor is used. The Shepard's Method can extrapolate values beyond the data's range.

Nearest Neighbour

The nearest neighbour technique is a very simple method in which grid point value are assigned the value at the station nearest them. The nearest neighbour method is similar to IDW, except that the grid values are not a weighted average of the values from nearby sampling locations. There are obvious problems with this method when there are no nearby stations. This method is the least accurate method for estimating values away from their sampling location.

For climate studies, both the nearest neighbor and IDW can be significantly improved by including topographical effects. For example, an environmental lapse rate (–6.5 °C/km) can be applied to temperature values at the stations before the interpolation is performed. Then, after the interpolation, the grid values can be transformed back by re-applying the lapse rate. This normalizing technique standardizes the temperature data with respect to elevation, resulting in a more accurate interpolation of 'like' data (Tait and Zheng, 2005).

Splines

Splines were originally flexible pieces of wood or metal used by draftsmen to fit curvilinearly smooth shapes when the mathematics and/or tools were not available to create the shapes directly (airplane wings/fuselages, boat hulls, railway tracks). Spline interpolation was later developed (early 1960's) to help in the computer aided manufacturing of car bodies, for computer graphics and to fit isopleths (contours) to continuous data.

Splines are polynomials (usually cubics) which are fitted to coordinate (X,Y) and surface values (Z) and forced to be smooth and continuous at the data joining points called 'knots' (Smith, et al., 2004). To visualize the spline in action, imagine a rubber sheet being used to interpolate the surface. The rubber sheet must be bent and stretched so that it passes through all of the measured values. The rubber sheet can either be forced to form nice curves (regularized spline interpolation) or the edges of the rubber sheet can be tightly controlled (tension spline). The spline curve goes through data points exactly and assumes a curved shape elsewhere. A spline interpolation demo can be seen at http://www.math.ucla.edu/~baker/java/hoefer/Spline.htm.

Splines are easy to compute mathematically and are useful in fitting a smooth line or surface to irregularly spaced data points while still retaining localized features. Spline functions produce good results for gently varying surfaces, such as elevation, water table heights or pollution concentrations.

The spline function is inappropriate if there are large changes in the surface within short horizontal distances and/or when the sample data are prone to error or uncertainty. With this type of data the spline method can wildly overshoot estimated values and closely spaced data points can develop embolisms. Extrapolating beyond the edges of the data domain often yields erratic results because there is no information beyond the data to constrain the extrapolation and splines will grow to large values (positive or negative). Splines tend to emphasize trend rather than anomalies (outliers).

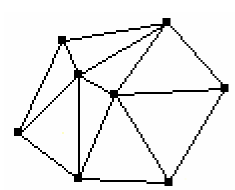
Trend Surface Analysis

Statistically based surface interpolation techniques assume that the sample data comes from a simple distributed in space distribution which means that the irregular spatial pattern cannot be describe by a mathematical function. Statistical techniques cannot handle the idea of a trend in the data. If there is a trend then the trend must be modeled so that the situation can be stabilized. With trend surface analysis drift can be analyzed for and subtracted out of the data much the same way an offset can be subtracted out of a data set (Glover, 1998, 2000).

A trend surface can be used to filter out large scale spatial trends in order to focus on the smaller scale variation (residuals) (Goodman, 1973, 1999, and Klinkenberg, 2002). Variability in spatial data is generally the product of two effects: the first the result of broad-scale regional changes in the value of the phenomena, the second the result of smaller-scale 'local' variations (Goodman, 1973, 1999; Davis, 1973). Trend surfaces typically fit one of three mathematically defined ideal surface models: linear (a constant dip in some single direction); quadratic (a bowl or dome shape, anticline or syncline); or cubic (saddle point, perhaps large scale folding).

Trend surface interpolation is highly sensitive to outliers (extreme values) and uneven distribution of the sample data points. The problem is further complicated by the fact that some data points are more informative than others. Polynomial trend surfaces have a tendency to estimate higher or lower values in areas where there are no data points, such as along the edges of maps. A polynomial model produces a rounded surface, which is rarely the case in many human and physical applications.

Triangulation

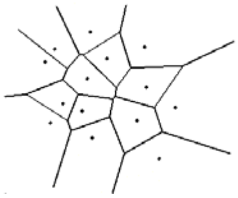


When a set of coordinate (X,Y) and surface value (Z) data points are connected in a triangular network of varying size the resulting lattice work is irregular and is called a Triangular Irregular Network (TIN) or irregular triangular mesh. The TIN partitions a surface into a set of contiguous, non-overlapping, triangles. The TIN is a vector data structure and is made from a set of points called mass points. Triangular structures were invented in the early 1970's because computers were limited and grids weren't as inefficient (Thurston, 2003).

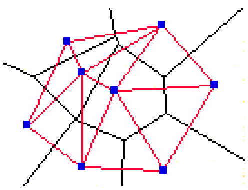
Triangulation is a fast, flexible and popular interpolation method that offers several advantages. When a surface has significant relief features, triangulation generates more accurate surface representation by maintaining break line features such as stream channels, shorelines, and slope features such as ridges. Triangulation can represent a surface with fewer data points than grid based methods. Triangulation works best when data points are evenly distributed over an area. The triangulation method honors the original data points exactly.

The main disadvantage of triangulation is that the generated surfaces are not smooth and may have a jagged appearance. The size and shape of the areas depend on the sample layout, which can lead to irregularly shaped polygons, especially on the edges. Sparse data sets result in distinct triangular facets.

Triangulation is not suitable for extrapolation beyond the observed data points. Analysis involving comparison with other layers of data is difficult. Estimation error can't be determined as the value assigned to each cell is based on only one value. Computation of a value at an unsampled point is a function of the polygon it lies within, rather than the values of the points closest to it.



The Voronoi diagram, sometimes named Dirichlet tessellation breaks an area into pieces or tiles. Lines that bisect the lines between a center point and its surrounding points define a Voronoi polygon. The bisecting lines and the connection lines are perpendicular to each other. Voronoi polygons are created so that every location within a polygon is closer to the sample point in that polygon than any other sample point. Interest in the Voronoi diagram originates from the climatologist A.H. Thiessen's use of Thiessen polygons to define regions that surround unevenly distributed weather stations.



Using the Voronoi diagram as a basis, Delaunay (named after B. Delaunay) triangulation is constructed by drawing lines between the points in adjacent polygons. Delaunay triangulation is a common and preferred technique since it provides nearly unique and optimal triangulation (Watson and Philip 1984; Tsai 1993).

HISTORY OF KRIGING

Given that concentrations of high-grade ore are easier and more profitable to mine, while regions of low-grade ore should be ignored, the estimation of recoverable ore is very important to the mining industry because local variability can make or break a mining venture. The first steps to resolve this problem were taken in the early 1950's in South Africa with the work of the mining engineer Danie Krige (Krige, 1951; Krige et al. 1989) and the statistician Herbert Sichel (1915-1995) working on the Witwatersrand goldfields.

The book "Historical Overview of the Witwatersrand Goldfields" (Handley, 2004) indicates that the discovery of the Witwatersrand goldfields in 1886 probably exerted a greater influence on the course of South African history than any other event. The Witwatersrand goldfield of South Africa has dominated world gold production for the last century and has been the source of almost one third of all the gold mined.

Krige and Sichel studied the best ways of predicting the value of a block to be mined using only samples from the faces already worked. Krige assumed that there was a relationship between neighboring blocks of ground in that blocks close to the stope (area of a mine from which ore is or has been extracted) would have a different and stronger relationship than those being estimated that are further away. Which is the basis for the assumption: "two blocks of ground will have a covariance which depends on the distance between them and on their relative locations." To simplify the calculations and cut down the amount of work necessary, Krige devised a template that could be laid over the mined areas.

Georges Matheron (1930-2000) an engineer with Ecoles des Mines, Centre de Morphologie Mathématique, Fontainebleaue, France, became aware of Krige's approach to ore reserves calculation while working with the French Geological Survey in Algeria and France from 1954 to 1963 (Rivoirard, 2000). Matheron adopted the pioneering work being done in South African and formalized the major concepts of the theory that he named geostatistics. In 1955 Matheron collaborated in a paper (Duval, et. al., 1955) that presented the work of Krige. Matheron's early work culminated in two books (Matheron, 1962 and Matheron, 1965). Out of Matheron's research came a spatial interpolation method that he called kriging in honor of Dr. Krige.

The word "krigeage" appears in Matheron's original work (in French) and the English translation was kriging. Matheron's first English paper appeared in 1963 (Matheron, 1963a). The two most common ways to pronounce kriging in English are "kree-ging" or "kree-jing". Danie Krige pronounced kriging with a hard 'g' (as in 'grand' or 'organize'), and not with the guttural Afrikaans 'g' (which is how he pronounced his name).

While Matheron was developing his theory of prediction in France, the meteorologist Lev S. Gandin (1921-1997) in the Soviet Union was doing remarkably similar work in meteorology and atmospheric sciences. Gandin first started publishing in 1959 (Krivoruchko, 2000, Armstrong and Galli, 2001), and

Gandin's first book complete with explanations about simple and ordinary kriging and cokriging techniques was published in Leningrad (Gandin, 1963) where the geostatistical technique was known as objective analysis or optimal interpolation. This work did not appear in English until much later when Gandin emigrated to Isreal (Myers, 1999).

B. Matern working in Sweden developed essentially a parallel theory to Matheron as a forestry application. Matern's work appeared in Swedish in 1960 and was not translated into English until 1986 (Myers, 1999). Matheron who is considered to have laid the foundation of geostatistics acknowledges that his work is similar to or duplicates the work of Matern and Gandin in some of his writings. Matheron knows Russian, was a communist supporter, studied statistics in the Soviet Union and was aware of Gandin's work.

From 1964 to 1968, Georges Matheron turned his attention to the mathematical characterization of geometric shapes and in collaboration with Jean Serra, created the discipline of "Mathematical Morphology" which has since become an essential branch of the field of image analysis (Rivoirard, 2000). In 1986 the Centre de Morphologie Mathématique became two programs, one on mathematical morphology and one on geostatistics (Centre de Geostatistique, 2000). Two of Matheron's first students (Journel and David) would start new centers of teaching and research in the USA and Canada.

In 1978 Shell Oil and the Bureau de Recherche Geologie Mathematique cooperated to develop a commercial software package called BLUEPACK. BLUE stands for Best Linear Unbiased Estimator. In 1980 the software MAGMA brought together BLUEPACK and the complete geostatistical library of the Centre de Géostatistique. Among the first users of MAGMA were Total, Exxon, Shell, British Petroleum, Agip, Amoco, and Gaz de France.

In 1993 ISATIS was released which offered in one package all the techniques previously available in MAGMA. ISATIS (http://www.geovariances.com/) is the result of 40 years of experience in industrial applications and applied research in geostatistics. Today, ISATIS is widely used the world over by more than 250 private oil & gas companies, consultant teams, mining corporations and environmental agencies.

In the mid 1980's the Environmental Protection Agency (EPA) commissioned a geostatistical software package, GEO-EAS, which was subsequently released in the public domain. GEO-EAS was a DOS program but included a menu system that made it fairly easy to use and was fairly inexpensive. The EPA did not continue to support the software and it has not been updated for a number of years. Go to http://www.epa.gov/ada/csmos/models/geoeas.html to find version 1.2.1 – April 1989.

In 1992 Andre Journel (Stanford University) and Clayton Deutsch published GSLIB (http://www.gslib.com/). GSLIB is an acronym for Geostatistical Software LIBrary. GSLIB is an extensive set of geostatistical programs and a user manual (second edition completed in 1997). The code is available on the website. GSLIB is compiled for a variety of platforms.

In 1996, Yvan Pannatier published VARIOWIN . Version 2.1 of the VARIOWIN (http://www-sst.unil.ch/research/variowin/) software was developed as part of a Ph.D. thesis that was presented on October 9, 1995 at the Institute of Mineralogy, University of Lausanne, Switzerland. VARIOWIN is a Microsoft Windows version of two of the components of GEO-EAS. VARIOWIN allows for much larger data sets than GEO-EAS and also allows for interactive variogram modeling. VARIOWIN 2.21 is still available.

Gstat is an open source computer code for multivariable geosatistical modeling, prediction and simulation. Gstat has been around from 1996 under the GNU General Public License (GPL) and is available from http://www.gstat.org/index.html. In the original form, gstat is a stand-alone executable, interfaced to various GIS. Gstat was not initially written for teaching purposes, but for research purposes, emphasizing flexibility, scalability and portability. As of 2003, the gstat functionality is also available as an S extension, either as R package or S-Plus library. Current development mainly focuses on the S extension. The gstat package provides multivariable geostatistical modeling, prediction and simulation, as well as several visualization functions.

GEMS (Geostatistical Earth Modeling Software) was designed at Stanford University (http://sgems.sourceforge.net/) with two aims in mind: to provide user friendly software that offers a large range of geostatistics tools complete with the ability to visualize data and results in an interactive 3D environment; and to design software whose functionalities could conveniently be augmented through a system of plug-ins.

Many applications of geostatistics continue to appear in: agroforestry, agronomy, air and water pollution, aerial distribution of acid rain and aerial contaminants, atmospheric sciences, disease outbreaks, entomology, environmental sciences, monitoring and assessment, epidemiology, fishery, forestry, geography, global change, meteorology, migratory bird population estimates, mining, oceanography, petroleum, plant pathology, surface hydrology, radioecology and more.

DATA ANALYSIS

Geographic Information Systems (GIS) includes tools to explore and visualize data to identify unusual data values or errors, detect patterns in data and to formulate hypothesis from data. Unfortunately this also inspires users to draw conclusions visually from maps and this often leads to faulty decision making (Krivoruchko, 2002b). Before starting any surface interpolation project it is important to do summary and exploratory data analysis to understand the classical statistics and spatial correlation of the data. Luc Anselin (Jacquez, 2005) defined exploratory spatial data analysis as "*techniques to describe and visualize spatial distributions, identify atypical locations (spatial outliers), discover patterns of spatial association (spatial clusters) and suggest different spatial regimes and other forms of spatial non-stationarity. . .* "

The first step is to verify three data features: dependency, stationarity and distribution (Krivoruchko, 2002a). If data are independent it makes little sense to analyze them geostatistically. Autocorrelation assumes stationarity, meaning that the spatial structure of the variable is consistent over the entire domain of the data set. If the data are not stationary, they need to be made so, usually by data detrending and data transformation. Geostatistics works best when input data are Gaussian (normal). If not, the data need to be made to be close to Gaussian distribution.

SUMMARY DATA ANALYSIS

Summary Statistics

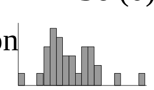
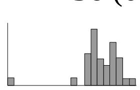
Summary statistics includes four groups:

- **measures of shape** –histogram (frequency and cumulative distribution), coefficient of skewness, kurtosis, normal probability plot, and quantile-quantile plot

- **measures of location** – number of samples, minimum and maximum values, mean, median, mode and quantiles

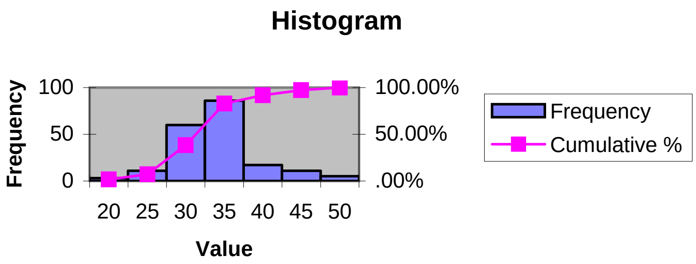
- **measures of spread** – variance and standard deviation and the coefficient of variation

- **measures of correlation** – coefficient of correlation, coefficient of determination, and standard error of estimate

	Untransformed	Transformed
mean	22.383	2.9600
std deviation	10.822	0.6650
sample variance	117.112	0.4423
minimum value	0.00	0.000
maximum value	55.10	4.009
n (n missing)	36 (0)	36 (0)
frequency distibution		
skewness (se)	0.73 (0.39)	-2.39 (0.39)
kurtosis (se)	0.89 (0.77)	9.03 (0.77)

Measures of Shape

Frequency Distribution



The histogram is the graphical version of a table which shows what proportion of cases fall into each of several or many specified categories. The frequency distribution summarizes discrete data by counting the number of observations falling into each category. The number associated with each category is called the frequency and the collection of frequencies over all categories gives the frequency distribution of that variable.

Cumulative Distribution

The cumulative distribution determines the number of observations that lie above (or below) a particular value in a data set. The normal distribution produces an s-shaped graph when the frequency value is plotted against cumulative proportion.

Skewness

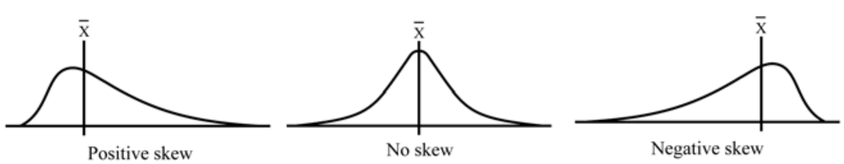
Skewness measures the degree to which a distribution is asymmetric. Normal distributions (symmetrical distribution) will have a skewness value near zero. If a distribution's mean and median do not coincide, the distribution is skewed.

Positive Skew

If the distribution's mean is greater than (to the right of) the median, the distribution is said to be skewed to the right. The tail of the distribution points to the right, and the mound of data are on the left. Data found in pollution studies or geological applications tend to have a lot of values clustered below the mean value and a long tail into the high values.

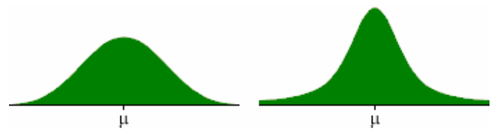
Negative Skew

The opposite is referred to as a negative skew. Data found in geological applications in limestone and sedimentary iron ores tend to have a lot of high values with a long tail into the lower values.



Right-skewed distributions will have a positive skewness value; left-skewed distributions will have a negative skewness value. Typically, the skewness value will range from negative 3 to positive 3.

Kurtosis



Kurtosis is a statistical measure used to describe the distribution of observed data around the mean. Kurtosis measures how flat or peaked the distribution is, and how thick or thin the tails of the distribution are. The graph on the right has a higher kurtosis than the graph on the left. It is more peaked at the centre and it has fatter tails.

Many classical statistical tests depend on normality assumptions. Significant skewness and kurtosis clearly indicate that data are not normal. If a data set exhibits significant skewness or kurtosis (as indicated by a histogram or the numerical measures) someway to deal with the problem must be attempted. One approach is to apply some type of transformation to try to make the data normal, or more nearly normal. The Box-Cox transformation is a useful technique for trying to normalize a data set. In particular, taking the log or square root of a data set is often useful for data that exhibit moderate right skewness.

Mesokurtic

A normal random variable has a kurtosis of 3 irrespective of its mean or standard deviation. Kurtosis should be near 3 for a normal distribution (normal tails).

Leptokurtic

Kurtosis is greater than 3 for distributions that have thicker than normal tails. Leptokurtosis is associated with distributions that are simultaneously peaked and have fat tails. The graph on the right is leptokurtic.

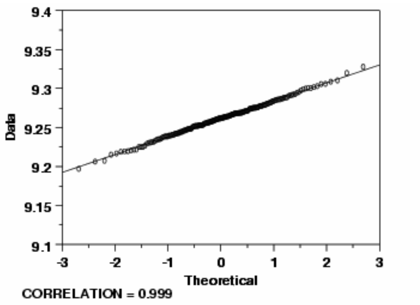
Platykurtic

Kurtosis is less than 3 for distributions with thinner than normal tails. Platykurtosis is associated with distributions that are simultaneously less peaked and have thinner tails. Platykurtic distributions are said to have shoulders. The graph on the left is platykurtic.

Normal Probability Plot

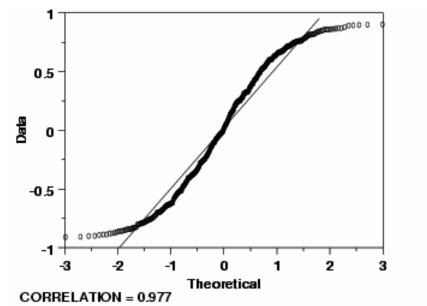
The following explanation of the normal probability plot and the explanation of the quantile-quantile plot were adapted from the “Engineering Statistics Handbook” A complete discussion on this topic can be found at the National Institute of Standards and Technology (NIST) – Information Technology Laboratory – Statistical Engineering Division website: http://www.itl.nist.gov/div898/handbook/eda/section3/eda33.htm. The normal probability plot is graphic technique for assessing whether or not a data set is approximately normally distributed. The data are plotted against a theoretical normal distribution in such a way that the points should form an approximate straight line. Departures from the straight line indicate departure from normality. It is not easy to judge how far the pattern in the plot can deviate from linearity before the assumption of normality should be judge implausible. Be careful about deciding against the plausibility of a normal distribution based on a normal probability plot when the sample size is small.

Normally Distributed Data



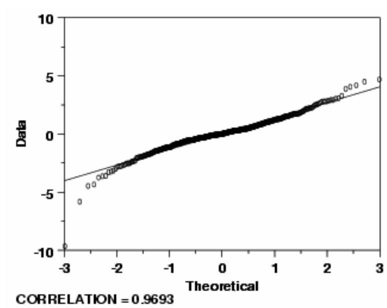
The normal probability plot based on normally distributed data shows a strong linear pattern. In this example, the normal distribution appears to be a good model for these data which is verified by the correlation coefficient of 0.999 of the line fit to the probability plot. The fact that the points in the lower and upper extremes of the plot do no deviate significantly from the straight line pattern indicates that there are no significant outliers.

Short Tails



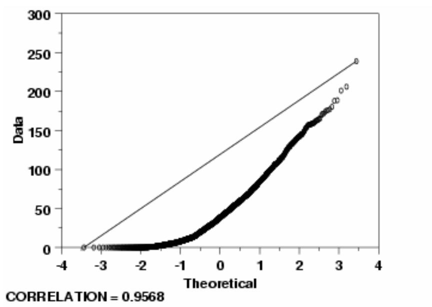
The normal probability plot based on data that has short tails shows a non-linear pattern. In this example, the normal distribution is not a good model for this data which shows up in two ways. First, the middle of the data shows an s-like pattern. This is common for both short and long tails. Second, the first few and the last few points show a marked departure from the reference fitted line. In comparing this plot to the long tail example the important difference is the direction of the departure from the fitted line for the first few and last few points. For short tails, the first few points show increasing departure from the fitted line above the line and the last few points show increasing departure from the fitted line below the line.

Long Tails



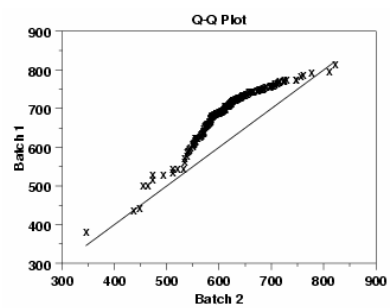
The normal probability plot based on data that has long tails shows a reasonably linear pattern in the centre of the data. However the tails, particularly the lower tail, show departures from the fitted line. For data with long tails relative to the normal distribution, the non-linearity of the normal probability plot can show up in two ways. First, the middle of the data may show an s-like pattern. This is common for both short and long tails. In this particular case, the s-pattern in the middle is fairly mild. Second, the first few and the last few points show marked departure from the reference fitted line. This is most noticeable for the first few data points. In comparing this plot to the short-tail example, the important difference is the direction of the departure from the fitted line for the first few and the last few points. For long tails, the first few points show increasing departure from the fitted line below the line and the last few points show increasing departure from the fitted line above the line.

Skewed



The normal probability plot based on data that are skewed right (example) shows a strongly non-linear quadratic pattern in which all the points are below a reference line drawn between the first and last points. The quadratic pattern in the normal probability plot is the signature of a significantly right-skewed data set. Similarly, if all the point on the normal probability plot fell above the reference line connecting the first and last points, that would be the signature pattern for a significantly left-skewed data set.

Quantile-Quantile Plot



A Quantile-Quantile (QQ) plot is a plot of the quantiles of one data set against the quantiles of a second data set. The QQ plot can determine if two data sets come from populations with a common distribution. The QQ plot has a number of advantages: sample sizes do not need to be equal; many distributional aspects can be simultaneously tested (shifts in location, shifts in scale, and changes in symmetry); the presence of outliers can be detected. For example, if the two data sets came from populations whose distributions differ only by a shift in location, the points should lie along a straight line that is displaced either up or down from the 45 degree reference line.

Measures of Location

Mean

The arithmetic mean (commonly called the average) is the sum of all data values divided by the number of data values. The mean is a good measure of central tendency for roughly symmetric distributions but can be misleading in skewed distributions since the mean can be greatly influence by extreme data values.

Median

The median is the middle of a distribution. To calculate the median the data must be ranked (sorted in ascending order). The median is the number in the middle. The median value corresponds to a cumulative percentage of 50% (i.e., 50% of the values are below the median and 50% of the values are above the median).

If the mean and the median are approximately the same value this is evidence that the data may be normally distributed. A frequency distribution is said to be skewed when its mean and median are different.

Mode

The mode is the most common (frequent value). The easiest way to look at modes is on the histogram. There may be no mode if no value appears more than any other. There may also be two modes (bimodal), three modes (trimodal), or four or more modes (multimodal). The mode will be close to the mean and median if the data have a normal or near-normal distribution.

Quantiles

A quantile is the variable value that corresponds to a fixed cumulative frequency. Any quantile can be read from the cumulative frequency plot:

- First quartile = 0.25 quantile
- Second quartile = median = 0.5 quantile
- Third quartile = 0.75 quantile

A quantile is the percent of points below a given value. That is, the 30% (0.3) quantile is the point at which 30 percent of the data fall below and 70 percent fall above that value. Probability intervals can also be read from the cumulative frequency plot, i.e. the 90% probability interval.

Measures of Spread

Measures of spread describe the variability of a data set. This group includes the range, variance, standard deviation, interquartile range and the coefficient of variation.

Range

The range is the largest data value minus the smallest data value. As a general rule of thumb, the range should be somewhere between 4 and 6 standard deviations depending on how many samples are present.

Variance

One of the most useful statistics to measure the variation of the sample values is the variance, which is found by considering the deviation of each sample from the average value. The variance is calculated by squaring the deviations, then adding them up and taking the average. Variance is a mean squared deviation and its units are squared rather than in the original sample units.

Standard Deviation

To get a descriptive statistic the square root of the variance is required to obtain the standard deviation (root mean squared deviation). The standard deviation can be interpreted as the difference between any data set member and the average of the data set. The mean value fixes the centre of the distribution, and the standard deviation scales the horizontal axis. In a normal distribution, 68% of the population values lie within one standard deviation of the mean value. Just over 95% of the values lie within two standard deviations. Approximately 99.7% of the observations are within 3 standard deviation of the mean.

Contour plots of the standard deviations of predicted values at non-sampled locations can be very useful. These contours show areas of higher uncertainty (higher standard deviations). Sampling from these locations can substantially improve the accuracy of predictions.

Geostatistics played a role in making the Channel Tunnel Project successful by assessing the geological risks and optimizing the alignment of the tunnel (RocNews, 2003). One of the most import criteria in optimizing the alignment was to ensure that the tunnel was bored within the Chalk Marl avoiding the Gault Clay material. Kriging was used to determine the boundary between the Chalk Marl and the Gault Clay, based on data available before construction. Contours of the standard deviations of predicted depths of this boundary were also generated.

The standard deviation contours helped engineers to realize that improved precision was required at certain tunnel sections. As more data became available from surveys and ongoing construction, geostatisticians enabled the tunnel engineers to readily improve the spatial model of the Chalk Marl/Gault Clay interface. Geostatistical analysis reduced the risk of penetrating the Gault Clay to acceptable levels. Penetration of the Gault Clay occurred only twice and in areas that already been predicted from the geostatistical model.

Interquartile Range

The interquartile range is used to describe the spread of a data set. The interquartile range is the difference between the first quartile (a number for which 25% of the data is less that that number) and third quartile (a number for which 75% of the data is less than that number) of a set of data. The interquartile range is the range of the central 50% of the sample values.

Coefficient of Variation

The coefficient of variation is calculated by dividing the mean by the standard deviation. The general rule of the coefficient of variation says, “if the standard deviation is relatively large compared to the arithmetic average, the average can’t be used to make decisions.” As a rule of thumb, when the standard deviation is smaller than the mean, the data are relatively closely clustered and mean is considered a reasonably good representation of the full data set. By contrast, if the standard deviation is greater than the mean, then the data are relatively widely dispersed and the mean is a rather poor representation of the full data set.

Correlation Analysis

Correlation analysis is used to determine the manner in which one (an independent) variable affects another (dependent variable). In correlation analysis, the coefficient of correlation, coefficient of determination, and standard error of estimate provide assessments of the reliability.

Coefficient of Correlation

Pearson's coefficient of correlation (r) is the most common measure of correlation or predictability between two variables. Pearson's r ranges in value from -1 to 1. The lager the r (ignoring the sign) the stronger the association and the more accurate it is to predict one variable from knowledge of the other variable. If r is 0 there is no correlation (no relationship) between the two variables.

The sign of the correlation implies the "direction" of the association. A positive correlation means that relatively high scores on one variable are paired with relatively high scores on the other variable, and low scores are paired with relatively low scores. A negative correlation means that relatively high scores on one variable are paired with relatively low score on the other variable.

A zero correlation does not necessarily mean there is no relationship between two variables – it means there is no linear relationship. For this reason, scatter plots are important supplements to statistical measures of association by providing a graphical picture of where variables may be associated with each other. Correlation measures the extent of association, but association does not imply causation. It can happen that two variables are highly correlated, not because one is related to the other, but because they are both strongly related to a third variable.

Deciding if r is "good enough" is an interpretive skill developed through experience. As a rule of thumb, if $r > 0.95$, there is good evidence for positive linear correlation, if $r < -0.95$, there is good evidence for negative linear correlation, and if $-0.95 < r < 0.95$, the data is either not linear, or "noisy". As an informal rule of thumb, call the relationship strong is r is greater than +0.8 or less than (more negative than) -0.8, weak if r is between -0.5 and 0.5, and moderate otherwise.

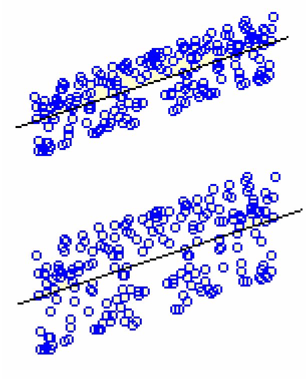
Coefficient of Determination

The coefficient of determination (r2, r^2 or R^2 – the square of the coefficient of correlation) is the ratio of the explained variation to the total variation and provides the proportion of the variation (fluctuation) of one variable that is predictable from the other variable. The coefficient of determination is a measure that provides the mechanism to determine how certain one can be in making predictions from a certain model or graph.

The coefficient of determination ranges from 0 to 1 and denotes the strength of the linear association between two variables. The coefficient of determination represents the percent of the data that is closest to the line of best fit. If the coefficient of correlation (r) is 0.922, then the coefficient of determination r2 = 0.850, which means that 85% of the total variation in the variable Y can be explained by the linear relationship between X and Y (described by the regression equation). The other 15% of the total variation in Y remains unexplained.

The better the regression line fits the data the larger the r2. If the regression line passes exactly through every point on the scatter plot, it would be able to explain all of the variation. The further the line is away from the points the less it is able to explain. A correlation coefficient of one (perfect correlation) is rarely, if ever, achieved. The coefficient of determination should be at least 0.6 and preferably between 0.8 and 1.0 for a reliable surface model.

Standard Error of Estimate



The standard error of estimate (se, SE, Se or SEE) is a measure of the accuracy of predictions. The top regression example show the points are closer to the regression line then they are in the bottom example. The predictions in the top graph are more accurate than the bottom graph. A regression line is the line that minimizes the sum of square deviations of prediction (also called the sum of squares error) . For example, if the standard error of the estimate is 775 then the prediction will be within +/- 775 units. The standard error of estimate is a measure of the variation (scatter) of the points about the regression line and is analogous to the standard deviation. However, the standard deviation measures the scatter about the arithmetical mean of a sample whereas the standard error of estimate measures the scatter of the dependent variable about the regression line. If the standard error of estimate is equal to the standard deviation of the dependent variable, then the correlation coefficient is equal to zero. The standard error of estimate provides a measure of the closeness of the point to the curve relation; it does not show the degree of correlation.

EXPLORATORY DATA ANALYSIS

Exploratory data analysis (EDA) is a set of graphical tools that help bridge the gap between data and modeling. The purpose of summary analysis is to arrive at a few key statistics such as mean and standard deviation, whereas the goal of exploratory data analysis is to gain insight into the process behind the data. Summary statistics are passive and historical, where as EDA is active and futuristic.

Censored Data

Censored data is data having a <,> or estimated qualifier, as is typical with water quality data. An accepted practice is to convert not detected (ND) values to one-half of the detection limit; to convert less than detection (<DL) values to the detection limit; and to convert estimated (E) values to the value estimated.

Co-located Samples

If the samples represent replicates of the same sample collected at the same time, an accepted practice is to average the samples and only include the average in the data set. If the sample represents a different element or a different time of collection an accepted practice is to change the location coordinates slightly. A problem arises with exact kriging when multiple values are measured at locations. A method of dealing with multiple samples was documented in Matheron's original works. With multiple samples at certain locations, the diagonal entry in the kriging system can be modified. Instead of $\gamma(0) = 0$ use $\gamma(0) = (n-1)/n$ times nugget effect. This tells the kriging system you have replicates and it will adjust weights and optimal estimator accordingly (Isobel Clark, 2001b)

Declustering

In many situations the spatial location of data collection sites are not randomly or regularly spaced. Data may have been sampled with a higher density in some places that in others. The mean value will be over-estimated if the samples are spatially clustered in the high value areas. A declustering method must be implemented to adjust for preferential sampling. One solution to preferential sampling is to weight the data, with data in densely sampled areas receiving less weight and data in sparsely sampled areas receiving greater weight. This can be accomplished in a couple of ways.

One method is cell declustering. A grid of cells is overlaid the data locations, and the weight attached to each data point is inversely proportional to the number of data points in its cell. Some researchers suggest that the cell size can be chosen corresponding to the minimum weighted mean if the data have been preferentially sampled in areas of high values. Conversely pick the cell size corresponding to the maximum weighted mean if the data have been preferentially sampled in areas of low value. This can lead to bias in other cases, and in general this method should not be used.

Another scheme uses a polygon method that defines a polygon around each spatial data location, such that all locations within that polygon are closer to the data location than any other data location. The idea is to weight each data point in proportion to the area that it represents. The problem with this method is that it is difficult to define weights toward the edge. The edge points can often receive large weights unless a border encloses the data.

Extreme Values

The first step in dealing with extreme values is to establish why the data contain extreme highs. Is there a high degree of imprecision in measuring values (the sample observations are actually inaccurate)? Is the distribution of data values skewed? Are there two (or more) populations, only one of which gives the high values? Are there other reasons?

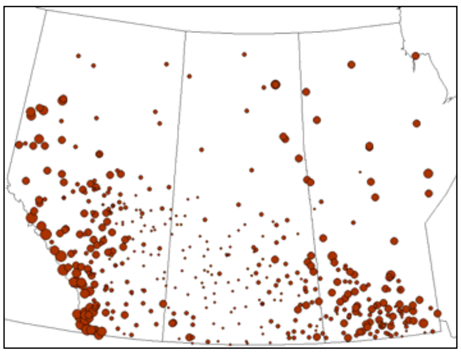
Once the reason for extreme values has been determined it will be possible to deal with this problem more objectively. You may be able to:

- use the error statistics in cross validation to assist in identifying erroneous sample measurements.
- use transformations or distribution-free approaches to geostatistics.
- use a mixed model together with indicator geostatistical approaches.
- identify extreme values as errors so they can be eliminated from the data set.
- stratify the data set by grouping extreme values together and treating them as a separate statistical population. Stratified kriging can be done by interpolating for different sub-regions separately and then the results of the interpolation can be combined as a single map.
- extreme values can be accepted and their effect on the distribution can be mathematically moderated. There are two sets of mathematical methods to do this. The first is to transform the data set before applying statistical measures. The other is to apply robust statistical measures that are better designed mathematically to reduce the effect of outliers or a small number of samples.

Skewed Data with Zeros

Isabel Clark (1987, 2002) suggests several ways of tackling skewed data with zeroes. Try a lognormal probability plot and see whether it is a straight line or if it drops off the line at low values. This is indicative of a three parameter lognormal distribution which needs an additive constant. Find the additive constant that makes the line straightest (Isabel’s criterion) or the skewness closest to zero (Sichel's recommendation). Treat the zeroes as a different population. Are the sampled locations zero because there are no fish there or because no fish were caught? If the later, use an indicator approach to separate the “no fish” population from the “some fish” population. Then do lognormal transformations on the “some fish” and recombine for a final result. The ‘not as nice’ approach is to use the lognormal probability plot to choose a threshold value to replace the zeroes. This assumes that all areas sampled are “some fish” areas and but no fish were caught.

Spatial Outliers



Special attention should be given to abnormal values (spatial outliers) since these values will seriously distort the interpolation process. A good way to detect the spatial outliers is to map or visually display the data set. Map display provides insights about the data set, how the samples are spatially distributed (uniform or clustered), their general pattern, trend and the extreme high or low values associated with the data set. The spatial location of extreme values is helpful in detecting erroneous data.

Care should be taken before deciding that a given outlier is actually a bad data point. An isolated extreme value may be suspicious, but this may not be sufficient to justify its removal. In some cases outlier values are obvious. There are a variety of methods that are used to treat outliers including inclusion, accommodation, replacement and rejection. The method to use depends on the objectives of the interpolation (depiction versus prediction; trend versus anomalies) and the nature of the data point. The decision to discard extreme values must be made with care, and data should be dismissed only if they are clearly wrong. It is possible to set a missing value indicator prior to surface generation in some software packages so that the missing values are ignored during analyses.

Stratification (Data set Subdivision)

If a phenomenon behaves differently in different locations the sample values collected will also differ significantly based on where they were taken. In this case, if a surface is created using all the measured points the basic principle of geography may be violated (things closer to one another are more alike). If this is the case it may be beneficial to divide or stratify the data into homogenous regions (or strata). Once the data is divided into parts the surface can be interpolated for each part separately and then the results can be combined.

The decision to split the data into more homogeneous subsets should be based on physical considerations and the density of the available data. Each subset should have enough data to allow the reliable inference of statistics for each population. Stratification should be based on information other than the data values themselves. Don’t stratify by putting high values into one stratum and low value into another. Secondary information (soil type, vegetation class) can be used when there is justification for stratification. Alternatively the data could be divided according to geographic location or elevation. Care should be taken because there may be no justification for stratification.

Time

According to Clark (2001c) there are two ways of approaching data that has a time element:

1. Treat time as a covariable and use cokriging.

2. Treat time as a dimension (an additional coordinate).

Transformation

Data transformations can be used to make data normally distributed. Kriging relies on the assumption of stationarity, and this assumption requires in part, that all data values come from distributions that have the same variability. Data transformations can improve predictions in many but not all cases. Transformations and trend removal are often applied to data to justify assumptions of normality and stationarity. The goal of transformations is to remove the relationship between the data variance and the trend. When data are composed of counts of events, such as crimes, the data variance is often related to the data mean. That is, if you have small counts in part of your study area, the variability in that region will be larger than the variability in a region where the counts are larger. The histogram and normal QQ plots can be used to see what transformations, if any, are needed to make a data set more normally distributed. The same transformation will likely equalize variances as well, helping to satisfy the stationarity assumption.

Arcsine Transform

For percentages or proportions (data between 0 and 1) the arcsine transform can be used. When data consists of proportions, the variance is smallest near 0 and 1 and largest near 0.5. The arcsine transformation often yields data that has constant variance throughout the study area and often makes the data appear normally distributed (Krivoruchko, 2005). Kriging prediction with a Box-Cox or arcsine transformation is known as transGaussian kriging (Krivoruchko, 2002a). To create an output map, the results of transGaussian kriging require back transformation to the original data, however, it should be noted that this can only be done approximately (Cressie, 1993).

Box-Cox Transform

Box-Cox (also known as power transformation) is a useful method to alleviate heteroscedasticity when the distribution of the dependent variable is not known. These transformations are defined only for positive data values. Heteroscedasticity is caused by nonnormality of one of the variables, an indirect relationship between variables, or the effect of a data transformation.

Lognormal Transform

Highly positively skewed data are encountered in many fields such as environmental, pollution and mining data. Kriging is a linear estimator and is sensitive to a few large samples, which may bias the results. To cope with highly positively skewed data geostatisticians developed lognormal kriging, which is kriging applied to lognormal transforms of data followed by a back-transform of the final estimates (Journel and Huijbregts, 1978). A back-transformation is not required if the original variable is not important, for example if the logarithm itself is a useful index. The log transformation requires that all data are positive.

Lognormal kriging must be used with caution because it is non-robust against departures from the lognormal model and the back-transform is very sensitive to semivariogram fluctuations. The back-transform tends to exaggerate any error associated with the interpolation and is most dramatic for extreme values. Although lognormal transforms attenuate the impact of high values it is not the most appropriate for censored data, which are better viewed as a different statistical population (Sito and Goovaert, 2000).

Sichel (1971, 1973) found that lower frequency distribution drop off can be brought back into line by adding a constant number to every one of the sample values before taking logarithms. If the probability plot shows a straight line, the population is said to be “three parameter lognormal.” According to Clark and Harper (2000) the additive constant has no physical meaning and is basically, just an artifact to normalize the logarithms. The size of the additive constant is a debatable issue with the most common rule of thumb being that it would be a problem if the additive constant starts to get to the same order of magnitude as the average value.

Sampling can reflect a mixture of normal or lognormal populations. For example if measurements are taken from both male and female populations the results may be comprise of two different normal populations. This structure will be reflected in the histogram and probability plots.

Normal Score Transform

The normal score transform ranks a data set, from lowest to highest values, and matches these ranks to equivalent ranks from a normal distribution. The transform is then defined by taking values from the normal distribution at that rank. Unlike the lognormal transform the normal score transform allows the symmetrization of the distribution of data regardless of the shape of the sample histogram (Goovaerts, 1997 and Deutsch and Journel, 1998).

A normal score transform allows the data distribution to be made symmetrical regardless of the shape of the sample histogram (Goovaerts, 1997 and Deutsch and Journel, 1998). The kriging of normal scores is referred to as multi-Gaussian kriging. Normal score transforms requires apriori specifications of the mean. An unfortunate feature of the normal score transform is the loss of the interval properties of the original data because the quantile definition depends on whether it occurs in the middle or at the tails of a standard normal distribution.

The normal score transformation cannot deal with any systematic trend in the data and should only be used in situations where the mean of the data is constant. Functional transformations such as lognormal or power work globally to transform data and fixed effects to a model with normal errors, where as the normal score transformation is working to make errors normally distributed. The normal score transform must occur after detrending since covariance and variograms are calculated on residuals after trend correction. Back transformation of a normal score transform is approximate and can be seriously biased when a trend exists in the data (Krivoruchko, 2002a). Although normal-score transforms attenuate the impact of high values it is not the most appropriate for censored data, which are better viewed as a different statistical population (Sito and Goovaert, 2000).

The goal of the normal score transform is to make all random errors for the whole population normally distributed. Thus it is important that the cumulative distribution from the sample reflect the true cumulative distribution of the whole population (Krivoruchko, 2005). An unfortunate property of the normal score transformation is that it destroys the interval property of the original data, because the quantile has a different meaning depending whether it occurs in the middle or at the tails of a standard normal distribution.

Scale Transform

In some cases it can be useful to scale data to a range of 0 to 1 if the values are extremely large.

Square Root Transform

The square-root transformation is a special case of the Box-Cox transformation. For count data the square root transforms can help make the variance more constant throughout the study area, and often makes the data appear normally distributed (Krivoruchko, 2005). Square-root transforms retain the measurement units. For square root, the back-transform through squaring tends to exaggerate any error associated with the interpolation. Such exaggeration of errors is most dramatic for extreme values (Sito and Goovaert, 2000).

Offsets and Back-transforms

Offset

In the case of a lognormal or square root transform if the Z value spans the range of less than one to greater than one (<1 to >1) all values should be made greater than one prior to transformation by adding an offset value.

Back-transformation

When a transformation is chosen, after analysis of the transformed data, the output data are customarily (but not necessarily) back-transformed to the original data domain. Three potential back-transformation choices are: none, standard or weighted.

Standard

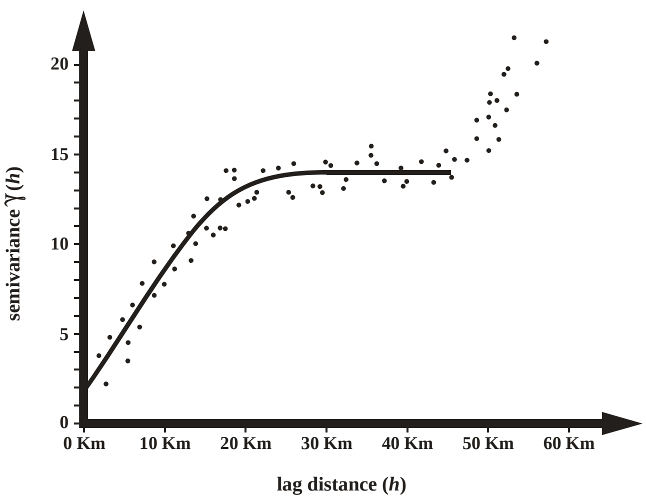
The standard back-transform is simply the converse of the transformation – scaled values are rescaled to the original range, log transformed values are raised to the natural exponent (*e)* and squared values are raised to the square root (0.5). Offset values are subtracted from the back-transformed values.

Weighted

The weighted back-transformation is a complex back-transformation that more closely approximates true population statistics than simple back transformations (Hann, 1977 and Krige, 1981).

Trends

When there is a trend in the data the first task is to identify the pattern and degree of the trend and then to decide whether the trend needs to accommodated. Trends can occur because of the measurement process, sampling design, or can be due to the physical qualities of the sampled object. Climatic data may follow a strong directional pattern. Increasing rainfall from west to east leads inevitably to constantly increasing local means in the same direction.



The semivariogram shows an example of data with definite trend. The fit appear to be fairly good for distances up to 45 kilometres, but beyond this the trend must be taken into account. The trend may or may not cause interpolation problems until after the range of influence is passed, however this is not always the case and the closer the parabolic behavior is to the origin the more attention will need to be paid to the trend.

There are several quantitative management tools for trends. One possibility is to remove the trend component. The trend component can be subtracted from the samples before calculation of variogram and identification of the spatial continuity. The trend surface can then be added back after kriging (Agterberg, 1974) or can be used as data input in universal kriging methods (Pebesma and Wesseling, 1998). Despite criticisms that the detrended data often have a different covariance function or variogram from the original data, they do have the same generalized covariance function (Kitanidis, 1993).

Secondary populations in the histogram, not removed by trend analysis, are strong suggestions to look for the existence of multiple populations, or perhaps an indicator method should be used. Geostatistical analysis is deeply dependent on the assumptions of single population with a consistent behavior.

APPLIED GEOSTATISTICS

The underlying principle behind geostatistics (and spatial interpolation in general) is that observation points closer together are more likely to have similar values than points further away. Tobler's Law of Geography (Tobler, 1979) is often cited as the First Law of Geography, where “everything is related to everything else, but near things are more related than distant things” (also known as autocorrelation). This means that samples collected close to one another are often more similar than samples collected further away, whether in space or in time.

Geostatistical interpolation is a two stage process:

1. Study the data to establish the predictability of values from place to place (determine dependency rules) and define the degree of autocorrelation.

The sample semivariance is used to estimate the shape of the variogram (the curve that represents the semivariance as a function of distance). The variogram describes the spatial relationship between the data points by showing the spatial structure and the data associations.

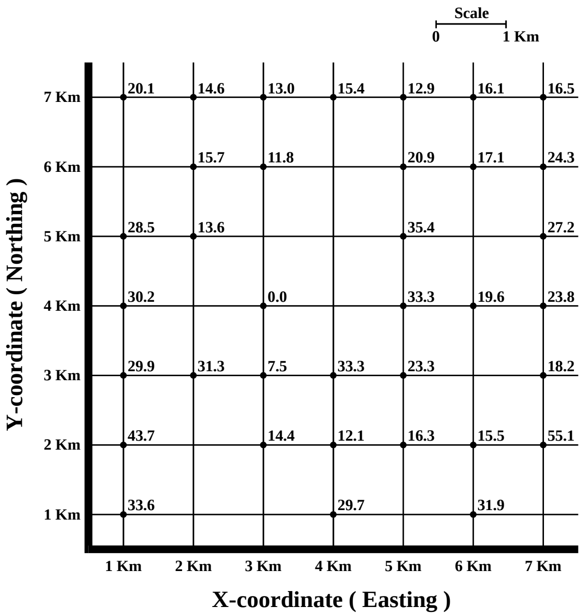
2. Once spatial or temporal dependency is established, the value at locations that have not been sampled are interpolated based on the degree of spatial autocorrelation. The process of prediction of the unknown values is called kriging.

The estimated semivariance function is used to determine the weights needed to define the contribution of each sampled point to the interpolation. Sample points close to the point for which an estimated value is to be generated contributes the most to the interpolation.

VARIOGRAMS

Geostatistics defines the correlation between any two values separated by a distance (known as the lag distance) then uses this information to make predictions at unsampled locations. Semivariance is a measure of the degree of spatial correlation among sample data points as a function of the distance and direction between the sample data points. The independent variable is the distance between pairs of points. The dependent variable is the semivariance of the differences in the data values for all samples a given distance apart. The semivariogram (or simply the variogram) controls how kriging weights are assigned to points during interpolation, and consequently controls the quality of the results. The semivariance increases with distance until the semivariance equals the variance around the average value and no longer increases, causing a flat region to occur on the semivariogram called the sill.

Calculating the variogram



It is not difficult to calculate and plot the semivariogram. The first step is to define a lag increment (*h*), which is the spacing between any two points in the data. For the example data a 1 kilometer lag increment is arbitrarily defined. For each pair separated by 1 kilometer, calculate the difference, and then square the difference. Sum up all the differences and divide by twice the number of pairs. This gives the measure of similarity $\gamma(h)$ for the variogram for a particular lag increment or distance. The same is done for other lag distances of 2, 3, 4, 5, and 6 kilometers. Grouping the pairs by location is referred to as binning. That is, all points that are within 0 to 1 kilometres apart are grouped into the first bin, those that are within 1 to 2 kilometres apart are grouped into the second bin, and so on.

$$\gamma(h) = \left(\sum_{i=1}^{N} [Z(x_i) - Z(x_i + h)]^2\right) / 2N$$

First Lag Distance[1]

Twenty (N = 20) pairs are found for lag distance h = 1. The pairs in the east-west direction and their calculated differences:

20.1 – 14.6 = 5.5	14.6 - 13 = 1.6	13 - 15.4 = -2.4	15.4 - 12.9 = 2.5
12.9 - 16.1 = -3.2	16.1 - 16.5 = -0.4	15.7 - 11.8 = 3.9	20.9 - 17.1 = 3.8
17.1 - 24.3 = -7.2	28.5 - 13.6 = 14.9	33.6 - 19.6 = 13.7	19.6 - 23.8 = -4.2
29.9 - 31.3 = -1.4	31.3 - 7.5 = 23.8	7.5 - 33.3 = -25.8	33.3 - 23.3 = 10
14.4 - 12.1 = 2.3	12.1 - 16.3 = -4.2	16.3 - 15.5 = 0.8	15.5 -55.1 = -39.6

The sum of the squared differences:

30.25 + 2.56 + 5.76 + 6.25 + 10.24 + 0.16 + 15.21 + 14.44 + 51.84 + 222.01 + 187.69 + 17.64 + 1.96 + 566.44 + 665.64 + 100.00 + 5.29 + 17.64 + .64 + 1568.2 = 3489.82

The gamma for the first lag distance is:

$\gamma(1) = 3489.82/(2*20) = 87.25$

Second Lag Distance

The lag distance h is incremented such that $h = h + h = 2h$. In this example the incremental value would be two kilometers. A search is conducted for pairs at a lag distance of two kilometers in the east-west direction. All pairs used for the h = 1 search and calculation are not included in the new lag interval.

[1] The example variogram calculations have been adapted from Carr (1995).

Twenty (N = 20) pairs are found for lag distance h = 2.

20.1 - 13.0 = 7.1	14.6 - 15.4 = -0.8	13.0 - 12.9 = 0.1	15.4 - 16.1 = -0.7
12.9 - 16.5 = -3.6	11.8 - 20.9 = -9.1	20.9 - 17.1 = -3.4	35.4 - 27.2 = 8.2
30.2 - 0 = 30.2	0 - 33.3 = -33.3	33.3 - 23.8 = 9.5	29.9 - 7.5 = 22.4
31.3 - 33.3 = -2.0	7.5 - 23.3 = -15.8	23.3 - 18.2 = 5.1	43.7 - 14.4 = 29.3
14.4 - 16.3 = -1.9	12.1 - 15.5 = -3.4	16.3 - 55.1 =-38.8	29.7 - 31.9 = -2.2

The sum of the squared differences:

50.41 + 0.64 + 0.1 + 0.49 + 12.96 + 82.81 + 11.56 + 67.24 + 912.04 + 1108.89 + 90.25 + 501.76 + 4.00 + 249.64 + 26.01 + 858.49 + 3.61 + 11.56 + 1505.40 + 4.84 = 5502.65

The gamma for the second lag distance is:

$\gamma(2) = 5502.65/(2*20) = 137.57$

Third Lag Distance

Pairs of observations in the east-west direction and with lag distances of three kilometers are now listed.

Fifteen (N = 15) pairs are found for lag distance h = 3.

20.1 - 15.4 = 4.7	14.6 - 12.9 = 1.7	13.0 - 16.1 = -3.1	15.4 - 16.5 = -1.1
15.7 - 20.9 = -5.2	11.8 - 17.1 = -5.3	13.6 - 35.4 = -21.8	0 - 19.6 = -19.6
29.9 - 33.3 = -3.4	31.3 - 23.3 = 8.0	33.3 - 18.2 = 15.1	43.7 - 12.1 = 31.6
14.4 - 15.5 = -1.1	12.1 - 55.1 = -43.0	33.6 - 29.7= 3.9	

The sum of the squared differences:

22.09 + 2.89 + 9.61 + 1.21 + 27.04 + 28.09 + 475.24 + 384.16 + 11.56 + 64.00 + 228.01 + 998.56 + 1.21 + 1849.00 + 15.21 = 4117.88

The gamma for the third lag distance is:

$\gamma(3) = 4117.88/(2*15) = 137.26$

Fourth Lag Distance

Pairs of observations in the east-west direction and with lag distances of four kilometers are now listed.

Twelve (N = 12) pairs are found for lag distance h = 4.

20.1 - 12.9 = 7.2	14.6 - 16.1 = -1.5	13.0 - 16.5 = -3.5	15.7 - 17.1 = -1.4
11.8 - 24.3 = -12.5	28.5 - 35.4 = -6.9	30.2 - 33.3 = -3.1	0 - 23.8 = -23.8
29.9 - 23.3 = 6.6	7.5 - 18.2 = -10.7	43.7 - 16.3 = 27.4	14.4 - 55.1 = -40.7

The sum of the squared differences:

51.84 + 2.25 + 12.25 + 1.96 + 156.25 + 47.61 + 9.61 + 566.44 + 43.56 + 114.49 + 750.76 + 1656.49 = 3413.51

The gamma for the fourth lag distance is:

$\gamma(4) = 3413.51/(2*12) = 142.23$

Fifth Lag Distance

Pairs of observations in the east-west direction and with lag distances of five kilometers are now listed.

Eight (*N* = 8) pairs are found for lag distance *h* = 5.

20.1 - 16.1 = 4.0	14.6 - 16.5 = -1.9	15.7 - 24.3 = -8.6	13.6 - 27.2 = -13.6
30.2 - 19.6 = 10.6	31.3 - 18.2 = 13.1	43.7 - 15.5 = 28.2	33.6 - 31.9 = 1.7

The sum of the squared differences:

16.00 + 3.61 + 73.96 + 184.96 + 112.36 + 171.61 + 795.24 + 2.89 = 1360.63

The gamma for the fifth lag distance is:

$\gamma(5) = 1360.63/(2*8) = 85.04$

Sixth Lag Distance

Pairs of observations in the east-west direction and with lag distances of six kilometers are now listed.

Five (*N* = 5) pairs are found for lag distance *h* = 6.

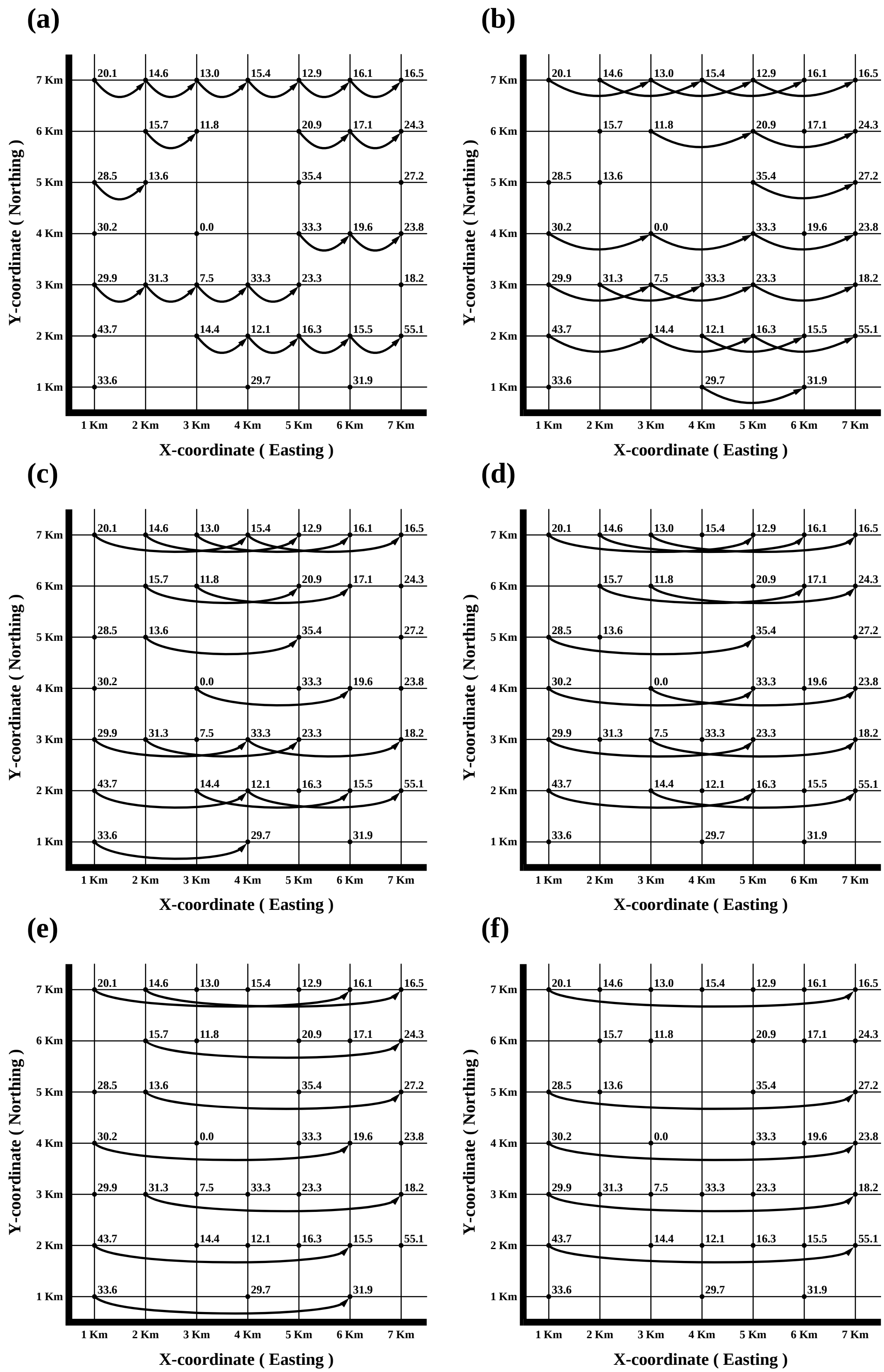
20.1 - 16.5 = 3.6	28.5 - 2.7.2 = 1.3	30.2 - 23.8 = 6.4	29.9 - 18.2 = 11.7
43.7 - 55.1 = -11.4			

The sum of the squared differences:

12.96 + 1.69 + 40.96 + 136.89 + 129.96 = 322.46

The gamma for the sixth lag distance is:

$\gamma(6)= 322.46/(2*5) = 32.25$



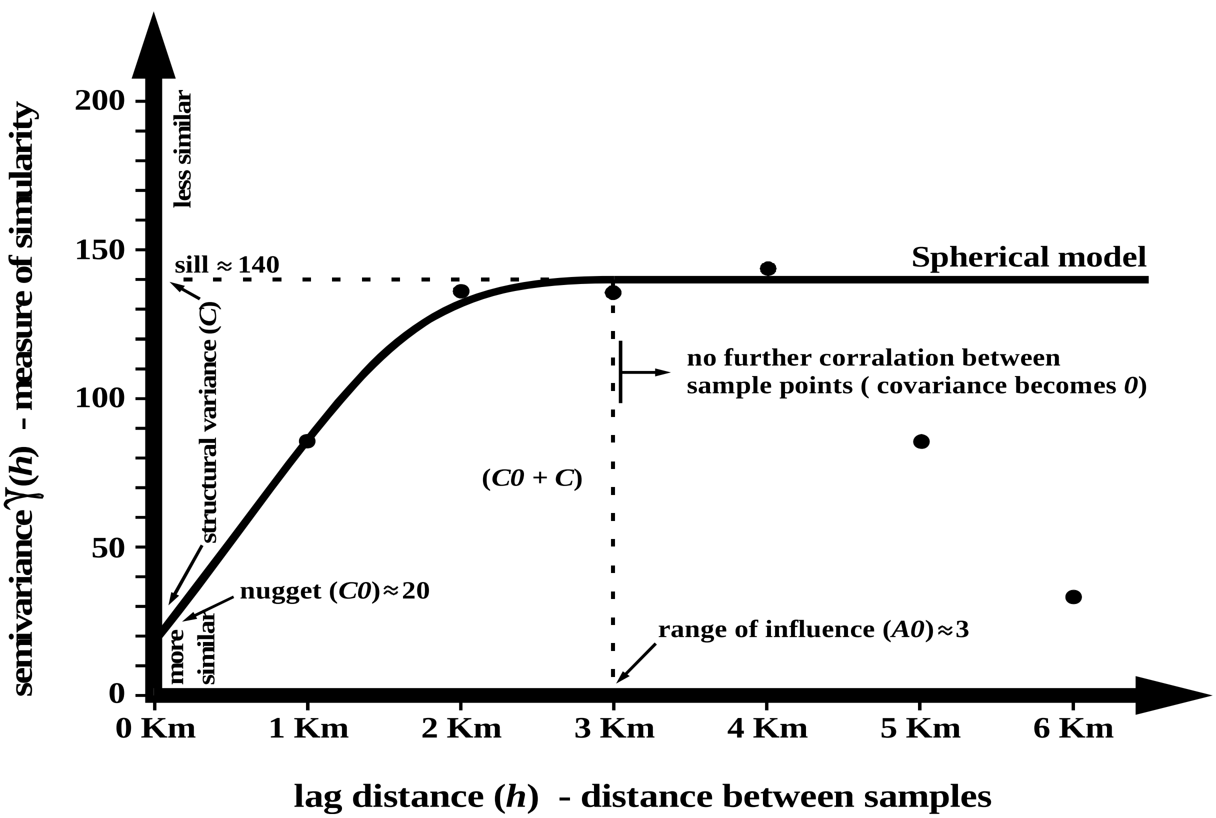
Transects used to estimate the semi-variances at varying lag distances (*h*)

Graphing the Semivariogram

A summary table of the east-west directional semivariogram calculations for the data represented is:

Interval of *h*	Number of Pairs	$\gamma(h)$
1	20	87.25
2	20	137.57
3	15	137.26
4	12	142.23
5	8	85.04
6	5	32.25

The semivariogram provides a graphical and numerical measure of the spatial continuity within a study area by plotting semivariance $\gamma(h)$ versus lag distance (h). The semivariogram values cannot be used directly instead a model must be fitted to the semivariogram. A line is fit through the plotted points to assess the spatial correlation. Once the model is fit, then this model is used to determine semivariogram values for various distances. The initial value of h is called the class size. Variogram modeling requires experience and practice.



The Semivariogram

Kriging is a robust technique and minor errors in estimation of the semivariogram parameters make little difference to the reliability of interpolation (Trangmar, Yost and Uehara, 1985). Semivariograms are statistical measures that assume the input sample data are normally distributed and that there are no trends in the local neighborhood means and standard deviations. The semivariogram has a number of components to which the semivariogram model is fit.

Coordinate System

The geographic latitude/longitude coordinate reference system uses angular measurements to describe a position on the surface of the earth. Only at the equator is the distance represented by one degree of longitude equal to the distance represented by one degree of latitude. Moving towards the poles the distance between lines of longitude becomes progressively smaller until at the exact location of the pole all 360° of longitude are represented by a single point. Because lines of longitude are not uniform, a latitude/longitude is not a suitable coordinate system for lag distances.

Choosing a suitable projection is necessary to ensure that the value of x and y units are equivalent and constant across the study and that distance and area calculations are in real distance units. Coordinates for lag separation distances are presumed to be in Cartesian space. The coordinates start with a 0,0 origin that increases for X in an easterly direction and for Y in a northerly direction. Values can be less than or greater than 0.

Lag

Lag Distance

The lag distance is the range over which autocorrelation is calculated. The number of lags and the size of the lags must be specified based on knowledge of the phenomena being analyzed and the reason for modeling the variogram. The selection of a lag size has important effects on the semivariogram.

- If the lag distance is too large, short-range autocorrelation may be masked by forcing data pairs of widely varying separation distances into a single lag increment and therefore short-range autocorrelation may not be detected.
- If the lag distance is too small, there may be too many empty sample and class sizes to get representative averages for a particular lag distance because there are relatively few pairs contributing to any one lag increment resulting in a 'noisy' semivariogram plot. The number of samples in each bin may be too small to be representative.
- The general rule is that lag size times the number of lags is approximately equal to one-half of the largest distance among all points.

Lag Class

Lag class distance is selected by experimenting with increasing or decreasing the value and noting its influence on the semivariogram plot. The goal is to find lag distances that generate a relatively smooth continuous structure. If the range of the fitted variogram model is very small relative to the extent of the semivariogram then the lag class distance can be decreased. If the range of the fitted semivariogram model is large relative to the extent of the semivariogram then the lag size can be increased. Typically variograms should generally be limited to a maximum distance equal to half the sampled distance.

Optimally, a class size should be found where approximately the same number of pairs result for each lag increment. For some data sets such a class size cannot be found due to sampling geometry. For these data sets a class size should be sought yielding an interpretable semivariogram showing a clear sill and range. It is common that the sill will be reached in less than 20 lags (Myers, 1997).

Use the smallest lag class distance spacing that produce acceptable number of pairs. Ideally 30 to 50 data pairs are consider a practical minimum to obtain a reliable estimate of the variogram at any lag (Armstrong, 1984 and Russo, 1984). In statistics, this approximates the number of pairs needed to produce a normal distribution if drawing samples from a random distribution. Webster and Oliver (1992) reported that variograms computed on fewer than 50 sample points are of little value and that at least 100 data are needed. Their experiments suggest that for a normally distributed isotropic variable a variogram computed from a sample of 150 data might often be satisfactory, while one derived from 225 data will usually be reliable. The best lag spacing may differ for an isotopic versus anisotropic variograms of the same data (Welhan, 2005).

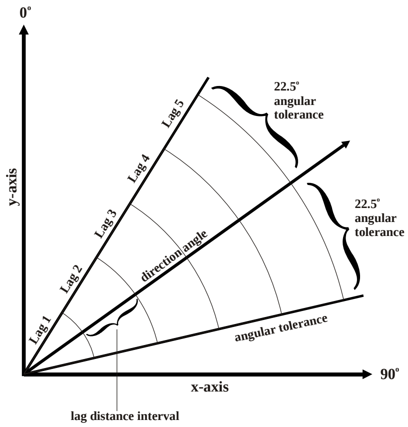
Variable Lag Class Distance

Specifying a variable non-uniform lag class is appropriate when locations are irregularly spaced or clustered across the interpolation area, or if the phenomenon being studied is known or suspected to be composed of several components. In this situation, a suggestion is to use lags with small separation distances at short distances and use lags with increasingly larger separation distances at longer distances. Changing the width of each lag may help improve detection of continuity.

Distance Tolerance

Lag distance tolerance is defined to deal with irregularly spaced data. Typically lag tolerance is defined to be half the lag increment used. For a lag increment of one kilometer the lag tolerance would be ± 500 meters. When the variogram is plotted the lag distance will be the average of all the distances between pairs falling within the tolerance. The larger the lag tolerance the more pairs that can be defined and the variogram will look smoother. Larger lag tolerances may also lead to unstable variograms that mask the continuity structure. The smaller the tolerance the fewer number of pairs will be defined and the variogram will not be as smooth.

Directional Tolerance



Choosing a lag distance and angular tolerance is similar to focusing a camera. There will be a certain lag (in each direction) at which the calculation will balance between detail on distance, detail on angles and number of pairs. Use the smallest directional tolerance angle that retains an acceptable number of data pairs per lag. Using too large a tolerance angle can over generalize by incorporating more pairs and will mask anisotropy that may be present. A tolerance of 90 degrees means an omni-directional variogram, regardless of azimuth. Directional tolerance must be chosen carefully and the process of examining the h -scatter plots at different lag widths or angular tolerances is essential.

Range

The range (A0) is the distance where the semivariance reaches the sill. The range of influence defines a neighborhood within which all locations are related to one another. Samples separated by distances in excess of this range are said to be spatially independent (not correlated). The range signifies the distance beyond which sample data should not be considered in the interpolation. A larger range means more continuous behavior, the variable is very well correlated and predictions result in fairly smooth maps. The range typically tends to increase as more and better data become available.

Sill

When the distance between sampling points is zero the value at each point is compared to itself. Therefore the difference is zero and the calculated semivariance for $\gamma(0)$ should be zero. If h is a small distance, the points being compared tend to be very similar, and the semivariance will be a small value. As the distance h increases, the points being compared are less and less closely related to each other and their differences become larger, resulting in larger values of $\gamma(h)$. Spatial variability in most environmental data between sample pairs increases as the separation distance increases.

At some distance the points being compared are so far apart that they are no longer related to each other, and their squared differences become equal in magnitude to the variance around the average value. The semivariance no longer increases and the semivariogram develops a flat region called the sill. The sill indicates that there is no further correlation between pairs of samples. The higher the sill value, the higher the prediction variances.

The semivariogram sill is theoretically equivalent to the variance of the spatial samples – however the sample variance is often an inappropriate measure of the population variance (Clark, 1979b, Rossi, et al. 1992; Isaaks and Srivastava 1989). The standard formula for estimating the variance assumes independent data, which is invalid in most environmental situations (Goovaerts, 2002, Barnes, 1991). In general if a sill is clearly present its value should be used as an estimate of the population variance, and the sample variance should not be used as an estimate of the variogram sill (Barnes, 1991). If the following conditions are met, then the sample variance is a reasonable approximation for the variogram sill (Barnes, 2005):

- the data are evenly distributed across the area of interest; there is no significant trend in the data across the area of interest; and
- the dimension of the area of interest is more than three times the effective variogram range.

Nugget

When the semivariogram is plotted back to the zero lag distance the nugget effect (nugget variance) is found. The nugget effect would be expected to be zero since two samples from the same point should have the same value. If the intercept is greater than zero then a random or unstructured component is present. No matter how close sampling is done the difference between neighboring samples will never be zero.

The term nugget effect was coined from gold mining. In most gold deposits the nugget effect tends to be quite large due to the nuggety nature of the mineralization, so that samples taken close together can potentially have very different grades. No matter how close the samples get, there will be large differences in value between the samples because of the nuggety occurrence of the gold. The value of the nugget effect will be close to zero in those deposits that have a very uniform grade distribution, such as porphyry coppers (Surpac, 2004).

The nugget effect represents the inherent variability of the data which can be interpreted as sampling errors, measurement errors, reproducibility issues, short scale variations or random variation at distances smaller than the sampling interval (white noise). A nugget structure increases the variability uniformly across the entire variogram because the variability is not related to distance or direction of separation. Increasing the nugget effect inflates the prediction error (kriging variance) and hence increases uncertainty of estimated values (Siska & Kuai Hung, 2001). Specifying a nugget effect causes kriging to become more of a smoothing interpolator, implying less confidence in individual data points versus the overall trend of the data. The higher the nugget affect the smoother the resulting grid.

Ideally the nugget effect should be as small as possible. A large nugget effect denotes unreliable data and more effort should be put into getting the data correct rather than using the unreliable data with large nugget effect (ACE, 2005). A high nugget variance indicates large point-to-point variation at short distances and increased sampling often reveals more detail in structure. The nugget effect tends to increase with lag tolerance and with data scarcity. Typically, the nugget effect decreases, as more and better data become available.

The way in which the sample variogram behaves at near zero separation distances is critical in describing the spatial continuity. A variogram that consists of pure nugget effect indicates complete lack of spatial dependence. It is not possible to get a nugget effect that is different in different directions. If there is an apparently higher nugget effect in one direction than another, there is most likely a short range component for which the sampling effort is not dense enough to pick up (Clark and Harper, 2000).

There are two possible ways in which geostatistical software will deal with the semivariogram model at zero distance:

1. force the model to go through 0 and a zero distance ($\gamma(0) = 0$)

2. allow the model to hit the vertical axis ($\gamma(0)$ = nugget effect.

The first option makes kriging an exact interpolator and the sample location will be honored and the kriging variance will be zero. This is what Matheron originally specified and is described in early test books. The second option means that kriging will not exactly honor the data but will put most of the weight on the sample and some weight on the other samples.

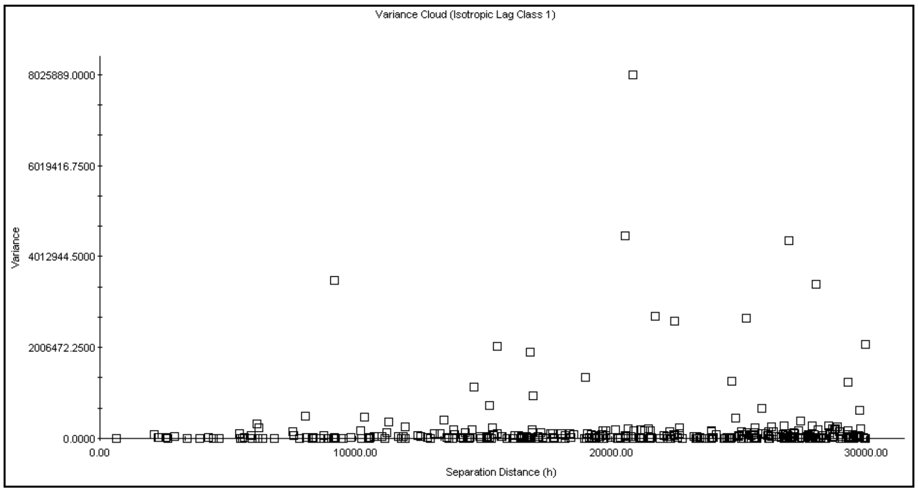
If the software only allows the second option the only way to honor the sample values is to have a zero nugget effect. The nugget effect does not have to be removed from the model, but another component (perhaps spherical) is added to the model where the sill equals the real nugget effect and whose range of influence is just below the closest sampling spacing. If it is not known which option is implemented in the software, run the interpolation with the nugget effect and with this alternative. If there is no difference in the results the software does the first option.

If the sampling errors can be quantified and the second method is implemented a combination can be used where short range spherical (for example) replaces the smaller scale variability and the nugget effect reflects the ‘true’ replication error. The choice to filter out the replication error by removing the nugget effect from the model must be made. It is important to keep in mind that when using the second method and/or removing the nugget effect when kriging, the calculated variances will be too low by a factor of 2*nugget effect. If the nugget effect is divided as suggested, the kriging variance will be too low by a factor of 2*replication error (Clark, 2006b).

Species of North American birds censused on the breeding bird survey vary considerably in their nugget variances (Villard and Maurer, 1996). Temperate migrants tend to have higher nugget variances than Neotropical migrants, and species that use a number of different habitat types have higher nugget variances than species using a single type of habitat. For estimates of abundances obtained from censuses like the breeding bird survey, part of the nugget variance could be due to the inherent variability among observers in counting birds at different breeding bird survey routes. The second component is due to the discontinuous nature of the process in space. For birds, this would be related to the continuity of the habitat in which they were being counted. Since the habitat of a species is generally discontinuous in space, this is a likely component of the nugget variance for data like the breeding bird survey.

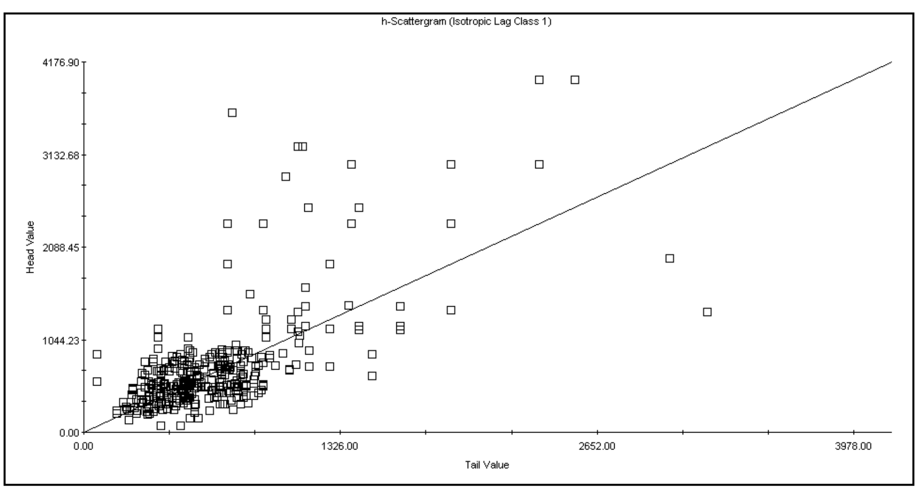
Variance Cloud

The variance cloud (variogram cloud) is a graph of the semivariances plotted against separation distance which is used to indicate how well the interpolated values fit the defined semivariogram model. The variance cloud can be used to refine the model. Variance clouds are specific to both direction and to a particular lag class. Spatial autocorrelation quantifies the basic principle of geography that things that are closer are more alike than those farther apart. Pairs of locations that are closer (far left on the x-axis of the variogram cloud) should have more similar values (low on the y-axis of the variogram cloud). As pairs of locations become farther apart (moving to the right on the x-axis of the variogram cloud), pairs of locations should become more dissimilar and have a higher squared difference (high on the y-axis of the semivariogram cloud). The variance cloud is useful for discovering outliers.



Scattergram

The scattergram (or scatter plot) is an XY representation of two variables used to analyze the spatial continuity of the data by displaying all the pairs of samples which are separated by a certain distance along a given direction. The coordinates correspond to the value of the first variable at the first sample location versus the value of the second variable (which can be identical to the first one) at the second sample location. The shape of the cloud of points spreads out as the spatial correlation between the two samples decreases or the relationship between the two variables weakens. The stattergram is specific to both direction and to its lag class. The stattergram can be used to look for outliers.



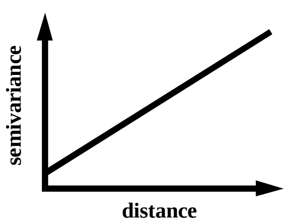
Semivariogram Models

The semivariogram provides a spatial picture of the data. Choosing which variogram model to be applied to the semivariogram is an art. Selecting the appropriate variogram model requires many correct decisions based on a solid understanding of the data and the underlying processes from which the data are drawn. The shape of the semivariogram function is typically one of four forms: linear, spherical, exponential, or Gaussian.

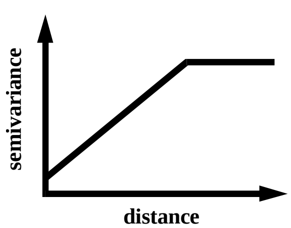
The major purpose of fitting a model is to give to provide a mathematical formula for the relationship between values at specified distances. Semivariances are estimated at discrete values of h, whereas the true variogram is continuous. The estimates are subject to error, and unless a large sample is taken the experimental variogram will appear erratic. The better correlated the variable the smoother the interpolated surface. Semivariograms can be a mix of two or more models.

The model selected influences the prediction of the unknown values, particularly when the shape of the curve near the origin changes significantly. The steeper the curve as it nears the origin, the more influence the closest neighbors will have on the prediction. As a result, the output surface will be less smooth. To consider the effect of the model on predictions, think in terms of the shape of the variogram for early lags

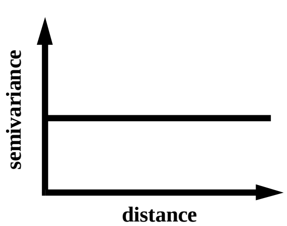
Linear



The linear model is the simplest model for a semivariogram and applies when the spatial variability increases linearly with distance and never levels off. The linear model indicates that the greater the separation of two samples, the greater the difference in the two samples. The linear model does not have a plateau and may be considered to be the beginning of the spherical or exponential model (Matheron, 1963a, 1965). The range is defined arbitrarily to be the distance interval for the last lag class in the variogram. Because the range is an arbitrary value it should not be compared directly with the ranges of other models. There is no sill, and the sill is the calculated semivariance for the arbitrarily defined range.

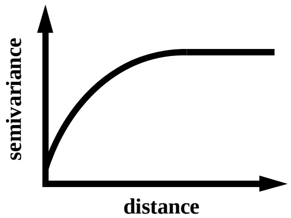


The linear/sill model is similar to the linear model except that at some distance (the range), pairs of points will no longer be autocorrelated and the variogram will reach a sill. This model should not be used unless the variation is limited to one dimension (Webster, 1985).



The semivariogram immediately takes its maximum value (represented by a flat semivariogram) if there is no correlation. In this case the phenomenon is completely random and there is no spatial dependence among the observations at the scale of sampling. Measurement errors or sampling sites too far apart (and as a result are spatially independent) or both may lead to noisy semivariograms that appear as pure nugget effect. Choosing a pure nugget effect model is an extreme modeling decision that precludes use of kriging.

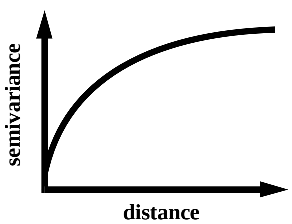
Spherical



The spherical model first proposed by Matheron (and sometimes referred to as the Matheron model) is based on the concept that a sample has a sphere of influence around it. Samples within this sphere have values that are related to the value at the central point. Imagine a second point with its own sphere of influence. If the spheres do not touch there is no relationship between the values at the two central points. If the spheres overlap, there will be a relationship, and the more the spheres overlap, the stronger the relationship. The spherical semivariogram is the simple geometric calculation for the volume of non-overlap of the two spheres, given the distance between their centres (Clark, 2006a)

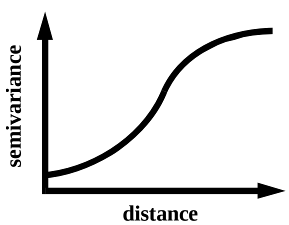
The spherical model shows a progressive decrease of spatial autocorrelation (increase of semivariance) of the variable with increasing distance. Dependence fades away altogether when the model reaches a sill. The spherical model is a modified quadratic function for which at some distance (the range) pairs of points will no longer be autocorrelated and the semivariogram reaches a sill (becomes asymptote). The spherical model is one of the most commonly used models.

Exponential



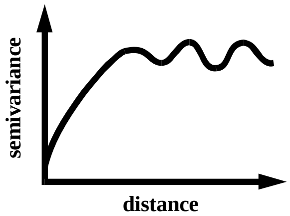
The exponential model was developed to represent the notion of exponential decay of "influence" between two samples. The exponential model applies when the spatial dependence decreases exponentially with increasing distance. This dependence disappears completely at infinite distance. The exponential model is similar to the spherical model in that it approaches the sill gradually, but different from the spherical model in the rate at which the sill is approached and in the fact that a definite sill is never actually achieved. The exponential model reaches 98% of its sill value at around three times the range of influence. The exponential model is also a commonly used model.

Gaussian



The Gaussian (hyperbolic) model is similar to the exponential model but assumes a gradual rise for the y-intercept. The Gaussian semivariogram model is named because the formula is basically identical to that for the normal (Gaussian) probability distribution. Spatial dependence vanishes only at an infinite distance. This model represents phenomena that are extremely continuous or similar at short distances. The main feature of this model is its parabolic shape at the origin. It expresses a rather smooth spatial variation of the variables. The Gaussian model does not have a true range but rather a parameter used in the model to provide range (called the effective range). For a Gaussian model the range of influence is square root of three times the distance scaling parameter. The Gaussian model is generally unstable in the presence of a nugget effect or of very dense data. This sort of model occurs in topographic applications or where samples are very large compared to the spatial continuity of the values being measured.

Hole Effect



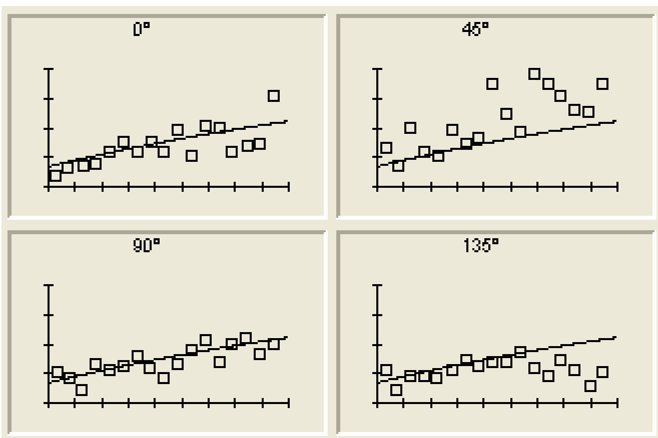
The hole effect (holesin) model suggests repetition in the variation that is neither wholly random nor periodic. This model was developed to represent a cyclic or periodic relationship between two samples. In practical circumstances, the relationship usually tends to trail off, but sometimes a dampening or decay parameter is included otherwise the cyclic effect would continue on to infinity. The hole effect model is not mathematically stable and can lead to some weird results if not used with care. Cyclicity may be linked to underlying periodicity or could be due to limited data.

Semivariograms from breeding bird survey data typically do not indicate the existence of an asymptotic maximum semivariance (Maurer 1994). Rather, they indicate that the maximum semivariance occurs at intermediate distances, and that points farther apart than this distance are actually more highly correlated. This increasing correlation with large distances separating sampling sites can be explained by the characteristic distribution of abundances across a geographic range. Generally a species is most abundant in one or a few places near the center of its geographic range, and its abundance decreases away from these points (Maurer and Villard, 1994). Abundances at points relatively large distances away from one another should be positively correlated because they represent peripheral sites in the geographic range.

Abundance may be determined by a number of factors, each operating at a different scale. For example, large-scale patterns in climate variation across a continent may determine large-scale patterns of vegetation productivity, and consequently variation in habitat for birds. Small-scale variation exists among sites located in the same biotic province, such as the physiographic regions used in the breeding bird survey. Sites that occur within the same physiographic regions, for example, might have abundances more similar than those that fall in different regions.

Anisotropic Variograms

Spatial autocorrelation can depend on more than the distance between two locations. Directional influence occurs when there is an effect that causes autocorrelation in one direction to be different from autocorrelation in another direction. Both isotropic (omnidirectional) and anisotropic (directional) variograms are examined to determine if autocorrelation is dependent on direction. Anisotropy alters the basic principle of geography to . . . things that are close together are more alike for distances in a certain direction than those in other directions. The isotropic variogram curve should be used to detect or confirm lag distance parameters to establish a clear continuity before examining anisotropic variograms.



Typically four anisotropic variograms with different orientations are constructed. These orientations are usually E/W (0°), NE/SW (45°), N/S (90°), and NW/SE (135° or -45°). The directional component is set at one of these directions, plus or minus a directional tolerance angle. Sometimes eight spatial directions analysis is used to yield more directional resolution. Larger data sets are usually required to yield enough information for the eight directional calculations.

If the directional variograms are reasonably similar then the spatial correlation can be assumed to be isotropic in which case the variability between samples would be a function of distance only. When the experimental variograms show different behaviors in different directions an anisotropic variogram model should be used since spatial correlation varies with direction and distance. For example, anisotropy associated with plant diseases can be interpreted as indicative of directional spread (Matheron, 1963a, Schotzko, et al., 1989, and Tangmar, et al., 1985.

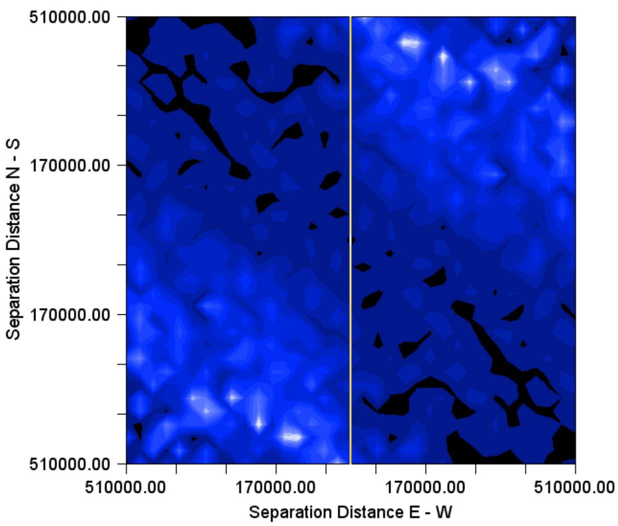
Nugget Variance

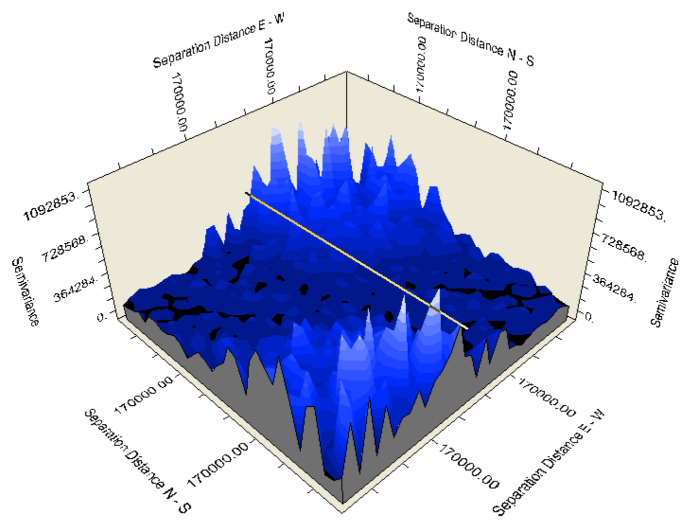
Nugget variance should be viewed as an isotropic parameter because it is not physically meaningful to say that uncorrelated (random) variability at a point is different depending on the direction away from that point. If the apparent nugget in two directional variograms is substantially different the data should be modeled as an additional anisotropic nested (non-nugget) structure with a maximum range so large as to add negligibly to the variogram at small lags (Welhan, 2005).

Anisotropic Variogram Surface

The anisotropic variogram (or variogram map) provides a visual picture of the semivariance in every direction by plotting the semivariances versus distance and direction. Each pair of samples corresponds to a distance and direction (converted to a grid cell), and to a variability (the cell value). Each location on the graph represents an approximate average of the pairs semivariance for the set of pair separation distances and directions, see Isaaks and Srivastava (1989; page 150) and Goovaerts (1997; page 98).

The anisotropic variogram is a good visual tool to highlight possible anisotropy in the data. The anisotropic variogram surface is used to find the axis of maximum spatial dependence (the lowest semivariances at a given distance) by viewing the anisotropic variogram surface to define and set the principal anisotropic axis to the direction aligned with the lowest semivariance values (the direction of maximum spatial continuity, or major axis of the anisotropic variogram model). The principal axis is the direction of maximum spatial continuity, or base axis from which the offset angles for anisotropic analyses are calculated. The axis orientation should correspond to the axis of maximum spatial continuity – the major anisotropic axis.





Geometric Anisotropy

Geometric (or affine) anisotropy is a situation where the variogram exhibits different ranges in different directions and the sill is constant in all directions. A longer range in one direction means that the values are more continuous in that direction than for a direction with a shorter range (Liebhold et al., 1993). For example, in an Aeolian deposit, permeability might have a larger range in the wind direction compared to the range perpendicular to the wind direction (Syed, 1997).

The anisotropy ratio is the ratio between the smallest range and biggest range (these directions are assumed to be approximately perpendicular to each other). A ratio of one denotes an isotropic variogram (Syed, 1997). The ratio between the lowest range and highest range, and the difference between their two angles can be used to transform the model to an isotropic model suitable for kriging (Ricci, A.K., 1998).

Zonal Anisotropy

Zonal anisotropy is a situation where the range remains constant in all directions but the sill varies with direction (see Isaaks and Srivastava, 1989, Figure 16.7 for a graphical representation). A variable sill means that the magnitude of spatial variation changes with direction (Liebhold et al., 1993). For example, a variogram in a vertical well bore typically shows a bigger sill than a variogram in the horizontal direction (Syed, 1997).

Since the sill of a semivariogram is equal to the population variance it is theoretically impossible to have different sills in different directions. A sill that varies with direction is likely a symptom of deeper problems such as non-stationarity, trend, discontinuities, etc. (Isobel Clark, 2001a). If different sills are seen in different direction look for one of the following:

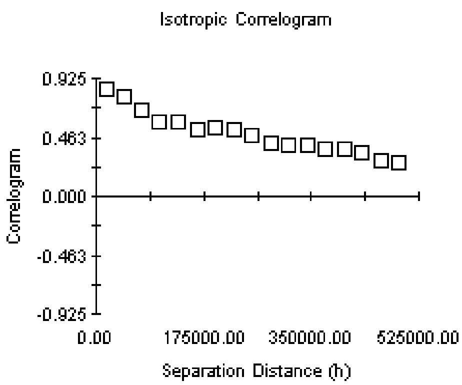
- the total sill has not been reached one direction;
- the data is skewed – a transformation should be tried;
- the assumption of homogeneity may have been violated – the sampling field may be crossing from one population group into another – the semivariogram in the direction of maximum variation will have the steepest slope, whereas the semivariogram in the direction of minimum variation will have the lowest slope (Issaks, et. al., 1989, Larkin, et. al., 1995, Trangmar, et. al., 1985, and Xiao, et. al., 1997);
- there may be a trend in the data

If an explanation cannot be found that fits one of these points or something similar there may be true zonal anisotropy, although true zonal anisotropy never occurs (Clark and Harper, 2000). Some variograms are a combination of both geometric and zonal anisotropies. For example geometric and zonal anisotropy is present in highly stratified conditions with strong horizontal layering (Onsoy, et al., 2005).

Alternate Variograms and Autocorrelation Measures

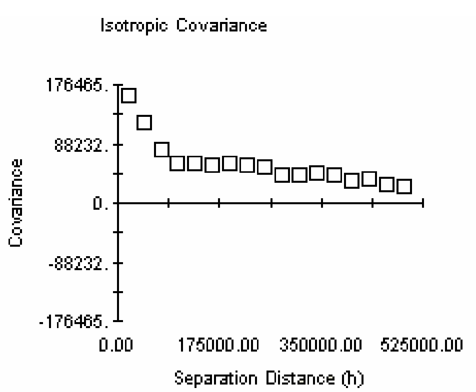
Alternative variograms and autocorrelation measures can provide additional estimates of spatial variability because these procedures are sometimes less sensitive to outliers, skewed distributions, or clustered data than ordinary variograms and these methods may aid in the recognition of data structure when the ordinary variogram is too noisy.

Correlogram



The correlogram shows the correlation among sample points by plotting autocorrelation against distance. Correlation usually decreases with distance until it reaches zero. A correlogram mirrors the variogram with the expectation that the correlogram assumes a stationary mean (Srivastava, 1996). As a variogram gradually rises to a plateau the correlogram gradually drops and levels out. Correlograms can be used to detect and describe regular spatial patterns between patches (Radeloff, et al. 2000). For a repeating pattern of patches separated by some distance, the patch size appears as a peak in autocorrelation for distances up to average patch size; the distance between patches appears as a second peak.

Covariance



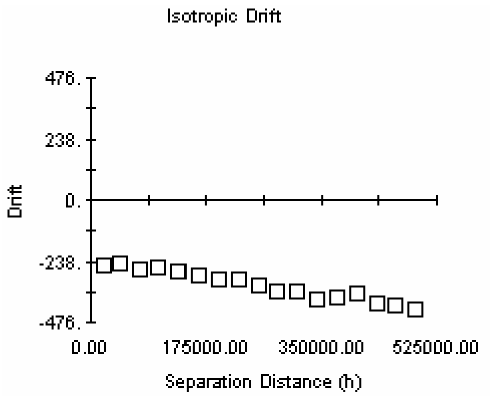
The covariance (like the semivariogram) measures the strength of statistical correlation as a function of distance – when two locations are close to each other they are expected to be similar and so their covariance will be large. As the two locations move further apart, they become less similar and their covariance eventually becomes zero. Since covariance decreases with distance, covariance can be thought of as a similarity function.

Cross-K

The Cross-K (Cressie, 1993) is similar to Ripley's K. A point from the first data set is chosen. A radius (*h*) is delineated and the points from the second data set that are within the circle are counted. A line above that designated by the interaction between two random point sets would imply that the data points are attracted to each other or tend to exist together in the same place. A line below shows that the points repel each other or tend to exist in opposing space.

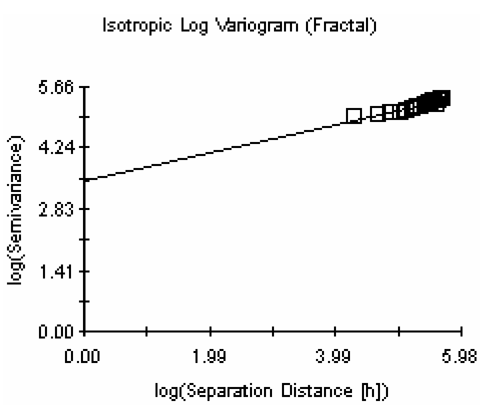
If Monte-Carlo envelopes were constructed around a K representation and a random point process was simulated and the K function for each was determined a plot of the maximum and minimum values at each distance (*h*) can be constructed. The envelope around the expected random K function represents a 99.9% confidence interval that the data are randomly distributed. If the K function from the observed data lies outside of this envelope there is very strong evidence that the data exhibit non-randomness.

Drift



Drift (also called trend) is calculated as the difference between values at different sample locations within a particular lag distance. Drift function is used to assess non-stationarity – does the mean value change over distance).

Fractal



Mandelbrot introduced the term fractal for temporal or spatial phenomena that are continuous but not differentiable, and that exhibit partial correlation over many scales. Since the mid-1980's fractal properties have been accepted widely by geologists and geophysicists as a quantitative characterization of complex geological phenomena and a fractal law has been accepted for many geological parameters. Many environmental variables suggest that not only are they fractals, but that they may have a wide range of fractal dimensions, including values that imply that interpolation mapping may not be appropriate in certain cases (Burrough, 1981).

In ecology, fractal theory has been used mostly as an analytical tool either to identify characteristic scales (Frontier, 1987) structural hierarchy (Burlando, 1990; Collins and Glenn, 1990) or to assess chaotic behavior (Sugihara and May, 1990) of phenomenon showing a spatial distribution or a temporal dynamics. One of the main appeals of fractals to ecologists lies in their ability to summarize the complexity and heterogeneity of a spatial or temporal distribution in a single value, the fractal dimension that is purportedly independent of scale (Leduc, et al., 1994).

The fractal dimension is directly related to the slope of the best-fitting line produced when the log of the distance between samples is regressed against the log of the mean-squared difference in the elevations for that distance. A graphical representation of variations of fractal dimension values as a function of scale has been called a fractogram (Palmer, 1988). According to Burroughs (1983), the slope of a variogram model plotted on log-log axes can be translated directly into a fractal (Hausdoff-Besicovitch) dimension. A linear variogram model on log-log axes would therefore indicate constant fractal dimension, and lead to simulations with similar textures at all scales (Englund, 1993). When a variogram is used as a means of determining the fractal dimension of the land surface, the normal practice of removing non-stationarity in the data should not be performed (Armstrong, 1986; Cressie and Hawkins, 1980). Fractal methods are not widely used in the mining industry (Vann, et. al., 2002).

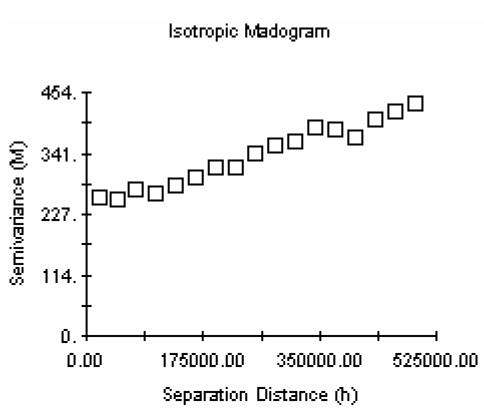
Geary's *C*

Geary's C (Geary, 1954) statistic is similar to Moran's I. Geary's C statistic is a squared difference statistic for assessing spatial autocorrelation. The values of C typically vary between 0 (similar) and 2 (dissimilar) although 2 is not a strict upper limit (Griffith, 1987). If values of any one location are spatially unrelated to any other location, then the expected value of C would be 1 (the theoretical value). Values between 0 and 1 indicate positive spatial autocorrelation while values between 1 and 2 indicate negative spatial autocorrelation. Gear's C is inversely related to Moran's I. Geary's C compares the squared differences in values between the center cell and its adjacent neighbors to the overall difference based on the mean of all the values. If the adjacent differences are less, then there is a positive correlation. If the differences are more, then there is a negative correlation. And if the adjacent differences are about the same, the variables are unrelated. See Comparison of Moran's *I* and Geary's *C* below.

Inverted Covariance

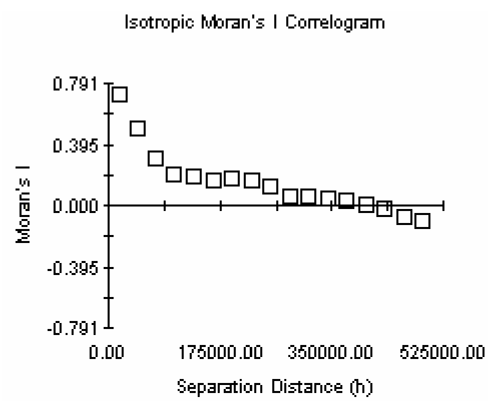
The inverted covariance (InvCov) is based on estimates of covariance rather than variance and is calculated by subtracting lag covariance from the sample variance (Srivastiva, 1988). This approach compensates for cases where the mean of the sample pairs in one direction is not the same as the mean of the sample pairs in another direction. The inverted covariance variograms have the same units (measurement units squared) as ordinary variograms and may be modeled and used for kriging in the same way.

Madogram



The madogram is a graph of mean absolute difference of sample measurements as a function of distance and direction. Madograms are less sensitive to extreme data values and can be useful for inferring range and anisotropy of data sets with outlier values that may make variogram features difficult to discern using traditional methods (Sobieraj, Elsenbeer and Cameron, 2003). The madogram should not be used for modeling the nugget of the semivariograms (Deutsch and Journel, 1992; Goovaerts, 1997).

Moran’s *I*



Moran's *I* (Moran, 1948 and 1950) is a weighted correlation coefficient used to describe autocorrelation by comparing the value of a variable at one location with values at all other locations. Moran’s *I* provides an indication of spatial patterns such as clusters and geographic trend. The semivariogram is a measure of dissimilarity while the Moran’s I is a measure of similarity. Moran’s *I* is one of the oldest indicators of spatial autocorrelation. Moran’s *I* varies between -1 and +1. Moran’s *I* shows a large positive value (close to 1) if a strong spatial correlation exists. Moran’s *I* becomes negative (close to -1) if the spatial autocorrelation is negative. Although it can be hard to detect spatial autocorrelation for shorter lag distances due to the lack of samples, a strong Moran’s *I* for smaller distances suggest a lower nugget effect. Moran’s analysis on a data set at different spacing (resolutions) provides insight on whether there is more than one controlling factor (Petrone et al., 2004).

No variography or kriging should be carried out for data with a zero Moran’s *I* for all lag distances since this does not respect Tobler’s Law. Deeper slopes in the Moran’s *I* correlogram indicate a shift from a spherical to an exponential model while flat profiles with a high spatial autocorrelation reveal a shift from a spherical to a Gaussian one (Garrott, 2003).

Comparison of Moran’s *I* and Geary’s *C*

Moran’s *I* is slightly more robust than the Geary’s *C* but Geary’s is often used as well. Moran’s *I* and Geary’s *C* statistics are essentially equivalent and interchangeable. Moran’s *I* gives a more global indicator whereas Geary’s *C* is more sensitive to differences in small neighborhoods.

The general interpretation of the Geary’s *C* and Moran’s *I* statistics can be summarized as:

autocorrelation	**Moran’s I**	**Geary’s C**
strong positive autocorrelation	I > 0	0 < C < 1
random distribution of values	I = 0	C = 1
strong negative autocorrelation	I < 0	1 < C < 2

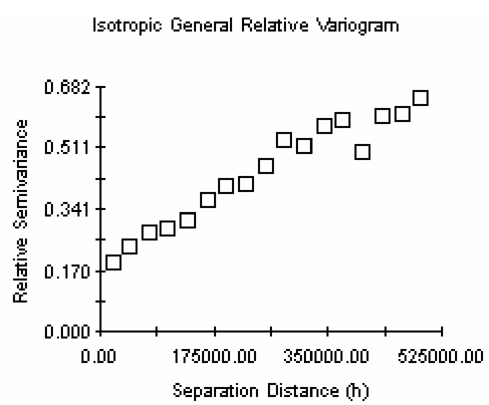
Relative Variograms

Often there is dependence between the semivariance and the mean of the data value for each lag distance. This leads to a proportional effect that complicates the interpretation of analysis. To detect the proportional effect a moving window statistic can be used. Relative variograms are analogous to the relative standard deviation often used to measure analytical variability. Relative variograms tend to be smoother than normal variograms and may aid in resolving the structure of the variogram. Relative variograms can be calculated only for data sets for which all Z value are positive. Relative variograms are unitless (decimal fraction squares). When modeled and used for kriging the relative kriging standard deviations must be multiplied by the estimated values to be comparable with kriging standard deviations produced with ordinary variogram models.

Local

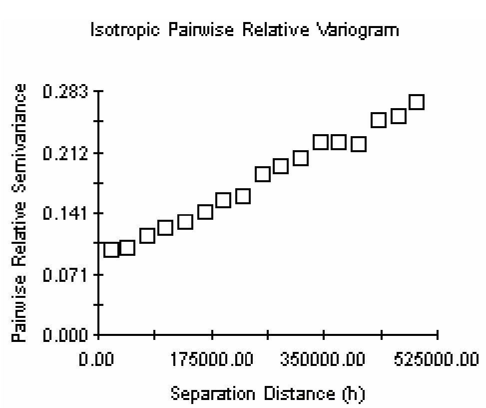
The local relative variogram considers the data within each region as a separate population and the values are proportional to their means. To describe the overall spatial continuity the local variogram values are scaled by their means. This method is rarely used because of the difficulty of determining the regions.

General



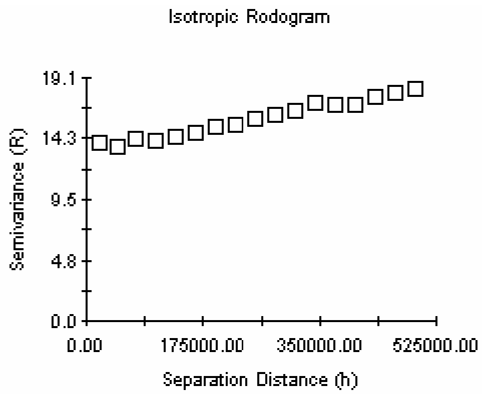
The general relative variogram weights the semivariance by the mean value of the data by dividing each semivariance by the squared mean of all samples used to estimate the semivariance. General relative variograms are less sensitive to clustering and data outliers and therefore may provide cleaner estimates of spatial correlation (Deutsch and Journel, 1997, p.45).

Pairwise



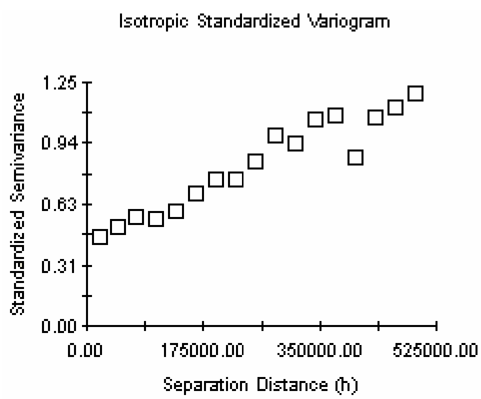
The pairwise relative variogram calculations are adjusted by a square mean performed for each pair of sample values by dividing each sample pair's difference by its mean. Experience has shown that the general relative and pairwise relative semivariograms are effective in revealing spatial structure and anisotropy when the scatter points are sparse (Deutsch & Journel, 1992). Pairwise relative variograms should only be used on positively skewed data sets. Pairwise relative variograms are less sensitive to clustering and data outliers and therefore may provide cleaner estimates of spatial correlation (Deutsch and Journel, 1997, p.45).

Rodogram



The rodogram is similar to the traditional variogram except that the square root of the absolute difference is used rather than the square difference. The Rodogram is a robust variogram that is useful for data with outliers.

Standardized



In standardized variogram analysis the variogram is computed on the natural logarithms of the variates.

Ripley's K

Ripley's K (Ripley, 1977; Cressie, 1993) is a spatial statistics function that is used to analyze spatial autocorrelation (usually clustering, rather than dispersion). Ripley's K is also called the reduced second moment measure – that is it measures second order trends (local clustering as opposed to a general pattern over the region. The function chooses a point within the data set and counts all the other points within a distance of that point. Like any statistic Ripley's K is prone to biases which include: edge biases, sample size, scale of interpretation, and the shape of the boundaries (Ned Levine & Associates, 2004).

KRIGING

Kriging can interpolate values for points at unsampled locations by using knowledge of the underlying spatial autocorrelation provided by the semivariogram to find an optimal set of weights to estimate the surface at unsampled locations. The kriging process involves the construction of a weighted moving average. Since the semivariogram is a function of distance, the weights change according to the geographic arrangements of the samples. Low weights are assigned to distant samples and vice versa. Kriging also take into account the relative position of the samples to each other.

The Kriging Algorithm

The understanding of the kriging algorithm is beyond what is needed for a pragmatic introduction to geostatistics, however the implementation of the kriging algorithm can be summarized (Ramirez-Beltrán, et al., 2006). Given a set of known points with each point having an X,Y coordinate value and an associated Z variable value, and a set of unknown points with X,Y coordinates and a Z variable value to be estimated.

The unknown values can be estimated as follow:

$$z_{ij} = \sum_{i=1}^{n} w_{ij} z_i \quad j = 1, \ldots, m$$

where the weight values W_{ij} can be calculated by solving the following system of equations:

$$\sum_{k=1}^{n} u_{ij} w_{kj} + \lambda_j = v_{ij} \qquad i=1, \ldots, m \qquad j=1, \ldots, m$$

$$\sum_{k=1}^{n} w_{ki} = 1 \qquad j = 1, \ldots, m$$

where u_{ij} is the distance from the known point (x_i,y_i) to the known point (x_j,y_j) and v_{ij} is the distance from the known point (x_i,y_i) to the unknown point (x_j,y_j) and λ_j is a weight value that can be estimated using the calculated variogram with the assumption that weights sum for each X should equal 1.

Kriging Variations

Four variations of kriging interpolation are ordinary kriging, simple kriging, universal kriging, and kriging with an external drift. Any one of these kriging methods can be applied as one of two forms: punctual or block.

Ordinary Kriging

Ordinary kriging is done when there is no underlying spatial trend (drift), the mean of the variable is unknown and the sum of the kriging weights is equal to one. This method assumes that the data set has a stationary variance but also a non-stationary mean within the search radius. Ordinary Kriging is highly reliable and is recommended for most data sets. One of the main issues concerning ordinary kriging is whether the assumption of a constant mean is reasonable. Sometimes there are good scientific reasons to reject this assumption.

Simple Kriging

Simple kriging assumes that the data set has no underlying spatial trend, a stationary variance and a known mean value. Because the mean is assumed known, it is slightly more powerful than ordinary kriging, but in many situations the selection of a mean value is not obvious. Simple kriging also assigns a weight to the population mean, and is in effect making a strong assumption that the mean value is constant over the site. Simple kriging also requires that the available data be adequate to provide a good estimate of the mean. Simple kriging uses the average of the entire data set (a constant known mean) while ordinary kriging uses a local average (the average of the scatter points in the kriging subset for a particular interpolation point). Simple kriging can be less accurate than ordinary kriging, but it generally produces a result that is "smoother" and more aesthetically pleasing.

Universal Kriging

Universal kriging combines trend surface analysis (drift) with ordinary kriging (Bonham-Carter, 1994). Universal kriging is more difficult to apply. Once trends have been accounted for (or assumed not to exist), all other variation is assumed to be a function of distance. This method represents a true geostatistical approach to interpolating a trend surface of an area.

Universal kriging is a two-stage process.

1. Fit a trend (local or global) and calculate the residuals. From these residuals obtain a de-trended (driftless) semivariogram.

2. Using the semivariogram derived in step one together with the established trend do the kriging with the trend from the original sample values -- not from the residuals. The only time the residuals are used is to develop the semivariogram model.

The recommended setting for the trend is a first degree polynomial which will avoid unpredictable behavior at the outer margins of the data set (Loránd, 2005).

The problem with this method is that the semivariogram itself is sensitive to the form of the deterministic surface. It is possible to over fit the trend surface, which does not leave enough variation in the random errors to properly reflect uncertainty in the model. When used properly universal kriging is more powerful than ordinary kriging because it explains much of the variation in the data through the nonrandom trend surface.

Kriging with External Drift

Kriging with an external drift (KED) is an extension of universal kriging in that the external drift is intended to be defined by a secondary variable that may incorporate some information of relevance to the primary variable. Universal kriging assumes that the shape of the trend is known, but its magnitude (or coefficients) is unknown. It is also possible to specify a trend mathematically but the original intent of kriging with external drift was to incorporate a secondary variable relevant to estimation of the first.

Punctual and Block Kriging

Any one of the four kriging methods can be applied as either a punctual (point) or block method. In punctual kriging the goal is to predict the value of a variable over a specified region. In block kriging the goal is to predict the average value of a variable over a specified region (a block).

Punctual Kriging

Punctual (point) kriging provides an estimate for a given point by estimating the value of that point from a set of nearby sample values. Punctual kriging can be used if sampling was done to represent point values in a field or in time.

Punctual kriging is an exact interpolator – estimated values are identical to measured values where interpolated points coincide with sample locations. In the presence of a nugget variance there will be local discontinuities at the sampling points (Dorsel & La Breche, 1998). If the nugget variance is large, undesirably large estimation variances may be produced. These discontinuities depend on the particular locations where the samples were collected. Changing the sampling pattern could result in a different map of kriged values and their estimation variances. Therefore, results obtained by punctual kriging depend strongly on the sampling methodology.

Despite being computationally simple, punctual kriging punctual has some drawbacks: it does not work unless the variable being mapped is stationary; it is not unbiased in the presence of a trend; the estimates shift up or down from the true values depending upon the arrangements of points and the trend direction. Block kriging eliminates some of these short comings (Castrignanò and Lopez, 2003).

Block Kriging

Block kriging (block averages), can be used to estimate an attribute value over an area instead of at one point location. Due to an averaging out effect, block kriging variances are generally much smaller than corresponding punctual kriging variances (Castrignanò and Lopez, 2003). Block kriging provides better variance estimation and has the effect of smoothing interpolated results. Block kriging is an approximate interpolator (based on a block of points) so that estimated values are not necessarily identical to measured values where the interpolated points coincide with sample locations.

A fine spatial resolution (small blocks) will require denser sampling than a coarser resolution (large blocks) to predict block means at a given level of precision. An optimal sampling scheme will attempt to find the combination of block size and sampling spacing in order to achieve a given level of precision at the lowest total cost (Castrignanò and Lopez, 2003).

Block kriging can be used if sampling was done to represent the area around the actual sample point. For example if soil samples from an area around the sampling location were composited before analysis then block kriging may be more appropriate. Block kriging produces smoother maps than punctual kriging by interpolating average values over blocks with the effect of smoothing local discontinuities. These effects are particularly desirable when the main interest is to study regional patterns of variation rather than local details (Castrignanò and Lopez, 2003). Consequently block estimates appear more reliable than those for points. In environmental work block kriging is usually more appropriate than punctual kriging.

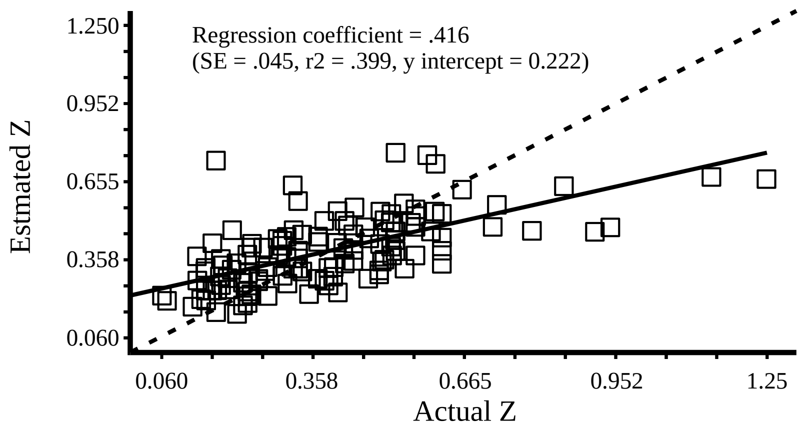
By computing block estimates, the nugget effect, which may be due either to measurement error or very short-range variation, can be avoided. Block interpolation may be more appropriate than punctual interpolation where average values of properties are more meaningful than exact single-point values, especially where spatial or temporal dependence is weak (Burgess and Webster 1980).

Kriging Search Neighborhood

The search neighborhood specifies which data points are included in the weighted moving spatial average. The use of a search neighborhood limits the estimates to just the data within some predefined radius of estimation called the search radius. The search radius does not need to be a circle or a sphere. To account for possible anisotropies in the data the search radius can be elliptical or ellipsoidal (in 3D). Kriging weights become small for sample points at a distance away from the location estimated and have a negligible influence especially if there is a strong spatial correlation. Points beyond the sill contribute in no way to the kriged estimate and so can be ignored. The use of the search neighborhood is the assumption of stationarity within the search window (quasi-stationarity). This concept can be useful if there is a trend in the data by assuming a small enough search neighborhood that the trend component can be ignored.

Cross-Validation

Cross-validation is used to check kriging model assumptions. After specifying the variogram model, the search neighborhood and kriging the surface the resulting kriged values and the true values are compared. The difference between these two values is called the cross-validated residual. A graph is constructed of the estimated versus the actual values for each sample location. Each point on the graph represents a location in the input data set for which an actual and estimated value are available.



If all estimated values were the same as the true value the scatter plot would be a 45-degree straight line. The distance each point is from the 45-degree line is the estimation error for that sample value. The closer the cloud of points is to this line, the better the estimation. The slope of the cross-validation scatter plot of predicted versus true values is usually less than one. It is a property of kriging that it tends to under-predict large values and over-predict small values. Systematic deviations from the line indicate lack-of-fit with the model. Individual points deviating from the line may be considered outliers.

Cross-validation is used to assess performance of the model assumptions used in kriging the surface. Is the variogram model appropriate, is the search neighborhood too small, are some regions being over or under estimated, should a trend model be incorporated, are the prediction errors comparable? Cross-validation does not prove that the variogram model is correct but rather that it is not grossly incorrect [Cressie, 1990].

Kriging Predictors

There are a wide variety of predictors used in geostatistics.

Cokriging

Cokriging is a multivariate extension of kriging to allow the incorporation of two or more variables that are correlated with each other by using one set of data to help explain and improve the description of variability in another (Goovaerts, 1997, pp 203-248, Wackernagel, 1995). Cokriging assumes that the second data set is highly correlated with the primary data set to be interpolated. For example altitude is considered as an important additional variable when estimating air temperature and hence cokriging should provide a better estimation than kriging because cokriging would take altitude as an additional variable. As another example, the quantity of vegetation (the biomass) may be related to elevation and soil moisture, and if this relationship can be determined, it would be possible to use cokriging to predict biomass based on both elevation and soil moisture.

Disjunctive Kriging

Disjunctive kriging (full indicator cokriging) is a specific nonlinear transformation of the original data that does not have simple (Gaussian) distributions (Rivoirard, 1994). The first step is to perform a normal score transformation of the data. The next step is to express the normalized function as an expansion of hermitian polynomials. Disjunctive kriging assumes all data pairs come from a bivariate standard normal distribution. The bivariate normal distribution is weaker than the multi-Gaussian condition. If bivariate normality exists, disjunctive kriging can produce better results than other geostatistical methods. If not, another type of kriging should be used. Disjunctive kriging tries to do more than ordinary kriging and indicator kriging by considering functions of the data and requires making stronger assumptions. The technique is computational daunting even with the use of variogram modeling software.

Indicator Kriging

Indicator kriging (Journel, 1993, and Journel and Arik, 1988) is a method developed for use on data populations with any distribution. The method is easy to understand. A cutoff value is defined, and all data below the cutoff are assigned a value of 0, all data above the cutoff are assigned 1. Variogram modeling and kriging is done on binary (0 or 1) values. An advantage of indicator kriging is to minimize the influence of outliers, to contend with skewed distributions, to deal with populations of both high value and values below the detection limit. Indicator kriging is not recommended for data having a trend.

Kriged indicator values can be viewed either as probabilities (the probability of exceedence), with a higher value denoting a higher probability or vice versa; or as proportions (the proportion of the attribute above the specified cut-off). The final result of indicator kriging is a set of probabilities that a grid cell exceeds a specific set of cut-off values with the values ranging between 0 and 1. This method is especially useful when a decision will be made based on a specific cutoff, or if the probability of exceeding some threshold is the final use of the interpolated data. Although geostatistical analysis is not specifically designed for binary data, it is tolerant of such data (Chellemi et. al, 1988, Lecaust et. al., 1989, Maison, et. al. 1976, Materson, 1963a, Tangmar, et. al., 1985).

Although sampling error is not taken into account the indicator transform can improve estimation of spatial continuity parameters in two ways. By reducing the scale of the variogram the behavior of the sample variograms frequently improves. Being an arithmetic average kriging can be drastically affected by even a single outlier. Deleting outliers is one approach to dealing with unusual values, but this may not be desirable because data are thrown away. Alternatively an indicator transform may be applied to the data, because no matter how unusual the data values are the data are transformed to either 0 or 1. Since the indicators are 0 or 1 the indicator variogram is very well behaved and resistant to outliers. Indicator variograms have a linear relationship (Lemmar, 1985).

When the threshold level does not correspond well with the continuity structure the indicator variogram will not produce good result, for example, when a skewed proportion of the data fall into one level relative to the other. Experience shows transforming the data into indicators using the median of the distribution as the threshold can provide a variogram revealing the approximate description of variability which at times is a reasonable approximation (Myers, 1997).

Among the shortcomings of indicator kriging is the loss of information after coding data through indicator functions. For example, if the data values are in the range of 1 to 100 and the indicator value (cutoff) selected is 60, then the data points with values 1 and 59 will be interpreted as being equivalent. The practice of using what was developed as a linear statistical model for continuous data to operate instead on discrete data (ones and zeroes) could be considered an “engineer’s method” – it works well as long as you don’t think too hard about its theoretical underpinning (Henley, 2001).

Lognormal Kriging

To cope with the presence of a few high grades South African mining geostatisticians developed lognormal kriging, which is kriging applied to lognormal transforms of data followed by a back-transform of final estimates (Journe, Huijbregts, 1978). The lognormal hypothesis is very strict in that any departure can result in completely biased estimates. Lognormal kriging is rarely used in mining applications today (Vann and Guibal, 2000). Some environmental parameters are lognormally distributed, and in such cases, lognormal kriging may be useful. The lognormal transform is referred to as a “sledgehammer” approach to data manipulation and can significantly reduce or even eliminate true distribution characteristics of the data (Myers, 1997). It is also believed that log transforming data often goes too far, especially if data are not originally lognormally distributed (Cressie and Hawkins, 1980).

Median Indicator Kriging

Median indicator kriging is an approximation of multiple indicator kriging which assumes that the spatial continuity of indicators at various cut-offs can be approximated by a single function.

Multiple Indicator Kriging

Multiple indicator kriging can be used if the data has a strange distribution or is a mixture of different populations. Multiple indicator kriging involves kriging of indicators at several cut-offs. Multiple indicator kriging is an approach to recoverable resource estimation which is robust to extreme values and practical to implement (Vann and Guibal, 2000).

The distribution histogram, probability plot, and other characteristics of the data should be examined before resorting to the complicated multiple indicator approach. If a decent semivariogram is obtained from ordinary values, indicator kriging is not needed. Theoretically multiple indicator kriging gives a worse approximation of the conditional expectation than disjunctive kriging, which can be shown to approximate a full cokriging of the indicators at all cut-offs, but does not have the strict stationarity restriction of disjunctive kriging (Vann and Guibal, 2000).

Isofactorial Disjunctive Kriging

There are several versions of isofactorial disjunctive kriging.

Gaussian Disjunctive Kriging

By far the most common is Gaussian disjunctive kriging which is based on an underlying diffusion model where the variable tends to move from lower to higher values and vice versa in a relatively continuous way. The initial data are transformed into values with a Gaussian distribution, which can easily be factorized into independent factors called Hermite polynomials (Rivoirard, 1994). Gaussian disjunctive kriging has proved to be relatively sensitive to stationarity decisions (Vann and Guibal, 2000).

Uniform Conditioning

Uniform condition is a variation of Gaussian disjunctive kriging more adapted to situations where stationarity is not very good. In order to ensure that the estimation is locally well constrained, a preliminary ordinary kriging of relatively large panels is made, and the proportions per panel are conditional to that kriging value. Uniform conditioning is a robust technique but does depend heavily on the quality of the kriging of the panels.

Probability Kriging

Probability kriging (Sullivan, 1984) represents an attempt to alleviate the order relationship problems associated with multiple indicator kriging by considering discretized (0 and 1) standardized rank transforms (the variables are sorted by increasing value and assigned their cumulative frequency which is referred to as the rank order transform) in addition to the indicator values. Probability kriging is thus cokriging between the indicator and the rank transform of the data.

Since cokriging must be done for a discrete number of cutoffs, there are a large number of variograms required to implement probability kriging. If ten cutoff values are used, then twenty-one variograms are necessary. The hybrid nature of this estimate as well as the time-consuming complexity of the structural analysis makes probability kriging rather unpractical (Vann and Guibal, 2000). Since probability kriging uses additional information it should theoretically produce better estimates than indicator kriging (Knudsen, 1987). Probability kriging is not recommended for data having a trend.

Regression Kriging

Regression kriging (RK) is a spatial interpolation technique that combines a regression of the dependent variable on auxiliary variable (such as terrain parameters, remote sensing imagery and thematic maps) with kriging of the regression residuals. Regression kriging is mathematically equivalent to interpolations methods such as universal kriging and kriging with external drift (KED) where auxiliary predictors are used directly to solve the kriging weights (Hengl, et. al., 2003). Hengl (2006) has a web page that presents step-by-step instruction for running regression kriging in a statistically sound manner – called the generic framework for regression-kriging (http://spatial-analyst.net/regkriging.php).

Residual Indicator Kriging

Residual indicators (Rivoirard, 1994, chapter 4 and 13) are one way to cokrige indicators by separately kriging independent combinations of the indicators and then recombining these to form the cokriged estimate. The residuals are defined from indicator functions. In practice the residuals are calculated at each data point, their variograms are then evaluated and independent kriging is performed.

Kriging and Remotely Sensed Images

Kriging has been found to be useful for assessing spatial patterns in remotely sensed images (Curran, 1988; McBratney and Webster, 1981; Ten Berge et al., 1983; Webster et al., 1989; Webster and Oliver, 1992; Woodcock et al., 1988a; and Woodcock et al., 1988b). Variograms of remotely sensed images should be interpreted with care because these variograms may differ from variograms resulting from ordinary samples. In remote sensing, the support size (support is a geostatistical term used to describe the size, geometry and orientation of the space on which an observation is defined) is equivalent to the spatial resolution. For example, reflection values are average over the field of view or pixel size of the measuring device. The shape and distribution of elements in the image influence variograms of data collected by remotely sensed devices.

According to the authors cited above some of the major points for variogram interpretation are:

- the range presents information on spatial dependence or reflectance – the range is related to sizes of objects in the terrain (groupings of shrubs)
- the variogram model or shape of the variogram reveals information on the spatial behavior of the data – the shape of the variogram is related to variability of the size of objects in the terrain
- the sill gives information on the total variability – the height of the variogram is influenced by the density of coverage of the objects and the spectral differences between the objects
- the nugget reveals information on variability between adjacent pixels
- as the spatial resolution become coarser the overall variance of the data is reduced and the fine scale variation becomes blurred – consequently, the sill height will reduce, the range will increase, and the nugget will increase
- anisotropy in the image is expressed by the variation of variogram parameters with the direction of the transect

Although the variogram seems to be a robust tool, a number of disadvantages of variograms can be identified. A fractal approach to assess spatial patterns from images provides an easier and more rapid method to assess spatial patterns from remotely sensed images (Ardini et al., 1991, Burrough, 1993, De Cola, 1989, de Jong, 1993, de Jong, et at. 1995, Jones et al., 1989, LaGro, 1991, Lovejoy, 1982, Vasil'yev and Tyuflin, 1992, Walsh et al. 1991).

GEOSTATISTICAL CONDITIONAL SIMULATION

Spatial uncertainty modeling has been a growing area of research in geostatistics for the last couple of decades (Goovaerts, 1997, 2001 and 2006; Chiles and Delfiner, 1999). Geostatistical conditional simulation (stochastic simulation) is more computationally demanding than geostatistical estimation. Increasing computer processor speed, memory and data storage capacity have provided an opportunity to apply these techniques. Geostatistical conditional simulation is a spatial extension of the concept of Monte Carlo simulation and provides a method to quantify uncertainty and to minimize risk. Conditional simulation is an interpolation technique that builds many realizations in which known data values are honored at their locations (hence 'conditional'). Each realization is different to others because there is always uncertainty away from known data points, hence individually such realizations are simulations not estimates. A conditional simulation passes through, or honors, the known data (hence it is said to be conditioned by the data).

A geostatistical simulation model is different than geostatistical estimation (kriging) in that a family of model realizations is generated. A series of realizations are produced that presents a range of plausible possibilities. The plausibility of these realizations is dependent on the assumptions and methodology of the simulation process. There is a clear distinction between methods where conditioning is built in (sequential methods) and where conditioning takes place as a separate kriging step (turning bands).

Each realization is different to others because there is always uncertainty away from known data points and so individually realizations are simulations not estimates. Simulation has different objectives to estimation. The point of simulation is to reproduce the variance of the input data, both in a univariate sense via the histogram and spatially via the variogram. Thus simulations provide an opportunity to study any problem relating to variability, for example risk analysis, in a way that estimates cannot (Vann, et. al., 2002)

Simulation provides a method to arrive at a theoretically infinite number of realizations (renditions) of the surface, each of which has approximately the same variogram and variance of the original data. These are referred to as equiprobable or equally likely images. Any individual simulation is a poorer estimate than kriging; however averaging a very large set of simulations can yield a good estimate. Simulation strives for realism whereas estimation strives for accuracy.

Simulations can be used to deal with a range of problems that involve variability, and for such problems they are far superior than any estimate (including kriging). One example of an application is to generate confidence intervals for an estimated block. Other applications include modeling local spatial variability for blending, stockpiling or mining selectivity studies (Bertoli, 2005).

Gaussian Simulation and Sequential Gaussian Simulation are the most common conditional simulation methods. Non-parametric simulation techniques such as Sequential Indicator Simulations and Probability Field Simulations are becoming more and more popular.

Conditional Cosimulation

Conditional Cosimulation is a multivariate conditional simulation of more than one variable.

Gaussian Simulation

Variogram modeling is performed on normally distributed data. The variogram model is used to krig the data at all locations providing a base map. In simulation, multiple realization maps are created that honor the actual data value (a conditional map), and approximately the same variogram and distribution. To generate each realization, an unconditional map is simulated which honors the variogram model but not the data at the data locations. Then the same unconditional map is used to create another map by kriging the values at the same locations as the actual data. At each grid node a simulated error is obtained – the difference between the kriged and simulated value. This error is added to the base map at each grid location. This gives the first realization which can then be back transformed to its original distribution. This is repeated for other realization using a different random number sequence to generate multiple realizations of the map (Syed, 1997).

Plurigaussian simulation

Plurigaussian simulation tries to simulate categorical variables by the intermediate simulation of continuous Gaussian variables (Armstrong et al., 2003). A Plurigaussian simulation is way of simulating the geometry of reservoirs and deposits. Fluid flow in oil reservoirs depends on their internal architecture - on the spatial layout of the lithofacies. A similar problem arises in mining when different facies with different grades or metallurgical recoveries are intermingled in a complicated way. Plurigaussian simulations make it possible to simulate complex types of geometry such as double (or multiple) anisotropies (e.g. gold bearing veins in one dominant direction with conjugate veins in another direction) or ore types that cannot touch each other. Plurigaussian simulation is an extension of Truncated Gaussian simulation. Plurigaussian simulation has proved to be very efficient in providing images reproducing the main features of the geology encountered in kimberlite crater deposits (Deraisme et al., 2004).

Probability Field Simulation

Probability Field (p-field) simulation is performed once to estimate the histogram at each grid location. Then using a uniform random number generator, a simulated value is drawn at each node using the estimated probability density function (or histogram). The only other constraint to honor is that the distribution of uniform random numbers should honor the variogram of the data, or simply that they should be correlated. This makes for fast conditional simulations compared to Sequential Indicator Simulation (Syed, 1997). Probability Field Simulation is very efficient and conceptually simple. Probability Field Simulation has two main drawbacks which can be detrimental. Nodes close to the conditioning data commonly appear as local minima or maxima of the simulated realizations. The simulated values usually show greater continuity than the original data. Pyrcz and Deutsch (2001) detail these problems and recommend that Probability Field Simulation not be used.

Sequential Gaussian Simulation

Sequential Gaussian simulation is an efficient method widely used in the mining industry (Vann et al., 2002). Variogram modeling is performed on normally distributed data. One grid node is selected at random then the value at that location is kriged which also gives the kriged variance. A random number is drawn from a normal (Gaussian) distribution that has a variance equivalent to the kriged variance and a mean equivalent to the kriged value. This number is then the simulated number for that grid node. Another grid node is selected at random and the process is repeated. This procedure defines a random path through all the gird nodes. For the kriging all previously simulated nodes are included to preserve the spatial variability as modeled in the variogram. When all the nodes have been simulated a back transform to the original distribution is performed. This gives the first realization. This process is repeated for all the other realizations using a different random number sequence to generate multiple realizations of the map (Syed, 1997). For Sequential Gaussian Simulation the biggest problem is the search neighborhood selection. The selection of small neighborhoods can lead to poor condition and poor replication of the variogram.

Sequential Indicator Simulation

Sequential Indicator simulation was developed for application in petroleum reservoir modeling where extreme values (high permeabilities) are well connected in space. The high connectivity of extreme values is difficult to model in the multigaussian paradigm. Another major advantage of the Sequential Indicator Simulation is that hard data and soft data can be easily mixed (Vann, et. al., 2002)

Variables are divided into several thresholds (or cut-offs). The data indicators are then coded (1 or 0) depending on exceedence or not of the thresholds. Sequential indicator simulation proceeds as follows: a random location is chosen and indicator kriging is performed for the threshold at that location. After correcting for possible order relations violations the indicator kriging will define a full cumulative probability density Function at that location. A simulated value from this Cumulative Probability Density Function is drawn based on a uniform random number generator. The sequence simulations then moves on to another location with this simulated value assumed as “hard” data until all nodes are simulated. This gives one realization (Syed, 1997).

A main difficulty with Sequential Indicator Simulation is the same as for multiple indicator kriging which is order relation problems (Vann and Guibal, 2000; Vann et al., 2000). Because indicator variogram models may be inconsistent from one cut-off to another a higher prediction about a certain cut-off have be predicted than for a lower cut-off. A further drawback is that the quality of the simulation is sensitive to the kriging neighborhood employed which is often too small (Vann et al., 2002).

Truncated Gaussian Simulation

Truncated Gaussian simulation was first designed to provide stochastic images of sedimentary geology, mostly in fluvio-deltaic environments. The basic principle is to replace the handling of the geological description by the handling of a Random Function with multigaussian distribution for which geostatistical simulations are used routinely (Carrasco et al., 2006).

Truncated Plurigaussian Simulation

The truncated plurigaussian model was introduced by Le Loc'h (1994), Gallie (1994), Le Loc'h and Gallie (1997) and others as a more flexible alternative to the standard truncated Gaussian model for describing facies distributions with complicated facies arrangements (Liu and Oliver, 2003)

Turning Bands Simulation

Turning Bands was the first large-scale 3D Gaussian simulation algorithm implemented (Journel, 1974, Mantoglou and Wilson, 1982). The Turning Bands method works by simulating a one-dimensional process on lines regularly spaced in 3D. The one-dimensional simulations are then projected onto the spatial coordinates and average to give the required 3D simulated value.

The Turning Bands method is very efficient for generating non-conditional simulations and reproduces the variogram better than other methods (Vann, et al., 2002). The Turning Bands method used to suffer mechanical limitations. Only certain specific variogram models (including spherical and exponential models) could be simulated and Turning Bands is considerably slower to process than sequential methods.

APPLYING GEOSTATISTICS

Rather than simply seeking to produce a visually-pleasing map geostatistics provides tools to explore and understand the nature of a data set. Do geostatistical tools produce useful results? “Useful” has to be decided by the user. Across a wide spectrum of applications the answer is yes but in some cases the answer may be not positive due to: a lack of data, difficulty in estimating or modeling the variogram, the sensitivity of the estimator to unusual data values, etc. It has been stated that “the interpretation of variograms has been found in practice to be almost as difficult as Freud’s interpretation of dreams”(Henley, 2001). Geostatisticians have developed a number of “rules of thumb” to help. There are no “correct answers in geostatistics, only the opportunity to gain more knowledge about the data and the interpolated surface, and to improve on the surface model.

Geostatistics is concerned with spatial data and each data value is associated with a location in space and there is at least an implied connection between the location and the data value. The semivariogram provides a spatial picture of the data and choosing the correct variogram model to use is an art. The development of an appropriate variogram model requires numerous correct decisions based on a solid understanding of the data and the underlying processes from which the data are drawn. The following is provided as a summary guide to understanding the kriging process.

Data Collection

- State the objective of the modeling exercise in clear quantitative term.
- Select a sample size and a sample collection strategy that are consistent with the objectives.
- Try to obtain a sample size of 100 data points or more – any calculations using less than 100 samples may give poor results.
- Errors made during sampling add a random component of variance to the variogram and increases the nugget effect. If the data are really bad, the variogram may show a total nugget effect.
- If necessary, re-project data from geographic to a projected coordinate system.
- Mixed data sets – if a number of sampling methods were used, each data set should be statistically analyzed separately to see if there are any significant differences.

Data, Summary Statistics and Histogram Visualization

- Examine a graduated symbols map looking for data trends, anomalies and errors.
- Is the standard deviation larger than the mean?
- Are the data skewed?
- If the data are multi-modal, the data may need to be split into different sets.
- Are there data anomalies?
- Possibly transform the data of skewed distributions (i.e., logarithmic transforms; indicator).

Semivariogram Visualization

- Check for enough number of data pairs at each lag distance (a minimum of 30 to 50 pairs).
- Truncate at half the maximum lag distance to ensure enough pairs.
- Use a large lag tolerance to get more pairs and a smoother variogram.
- Sometimes the variogram is sensitive to the choice of class size, lag distance and tolerance. A good starting point for the class size is the average spacing between the samples.
- Is the variogram nice, parabolic, discontinuous, erratic or completely random.
- If the spatial structure in the semivariogram doesn't look good, don't revert to IDW.
- Start with an omnidirectional variogram before proceeding with directional variograms.
- Remove trends.

Evaluation of Model

- Correct data errors.
- Manage extreme values or remove outliers – the estimator of the variance is very sensitive to outliers. A chaotic looking variogram can often be improved by limiting the maximum value used in the calculation.
- Select an appropriate model for the data and the objectives.

- Use other variogram measures to take into account lag means and variances (inverted covariance, correlogram or relative variograms).
- Interpret the model to obtain the best reliable fit possible.
- Understand the uncertainty and the precision of the predictions.
- Make decisions that are sensible.
- Reevaluate the model with new data when environmental conditions change.

CONCLUSION

Geostatistics (kriging) is an advanced technique for interpolation and mapping of data that relies on statistics (theory of regionalized variables) and computer processing. While there are individuals who would identify themselves as "geostatisticians" it is more likely that individuals using geostatistics would call themselves: ecologists, environmental scientists, geographers, geologists, hydrologists, mathematicians, meteorologists, mining engineers, petroleum engineers, plant pathologists, soil scientists, statisticians, etc.

Geostatistics is not a cure-all nor is it useful for all problems. It is not a black box scheme and must be used with some care. In some cases, and for certain data sets given some objectives, a user need only have access to a geostatistical software package. In other cases it may be necessary to seek the advice and assistance of someone more experienced and with a stronger understanding of the mathematics/statistics in order to adequately apply geostatistics to a data set.

The information presented in this paper was developed to provide a good base understanding to the field of geostatistics to either design or to analyze the data from environmental monitoring surveys. Hopefully no tears were shed during this learning process. An extensive bibliography is given which goes beyond the references cited in this paper and is provided to give readers an entry to the primary literature on geostatistics and surface interpolation.

REFERENCES

ACE, 2005. Options for Variogram. Applied Computer Engineering (ACE), Inc. Integrated Reservoir Studies. Houston Texax 77036. http://www.gridstat.com/ACEweb/example8.htm

Acevedo. M.F., 1997. Geog 5190 Advanced Stats and Quantitative Methods, Lecture Outlines, Lab Guides and Assignments. http://www.geog.unt.edu/~acevedo/courses/590/week10/lectur10.htm

Agterberg, F.P., 1974. Geomathematics, Elsevier Science Inc.

Annales des Mines (Accueil), 2000. Georges François Paul Marie Matheron (1930-2000). http://www.annales.org/archives/x/matheron.html

Australian National University. 2004. Centre for Resource and Environmental Studies (CRES)-ANUSPLIN Version 4.3. http://cres.anu.edu.au/outputs/anusplin.php

ANUDEM Version 5.2. The Australian National University. http://cres.anu.edu.au/outputs/anudem.php

Arad, N., Dyn, N., Reisfeld, D. and Yeshurun, Y. 1994. Image warping by radial basis functions: applications to facial expressions. CVGIP: Graphical Models and Image Processing, 56, 161-172.

Ardini, F., Fioavanti, S. and Giusto, D.D., 1991. Multifractals towards Remote Sensing Surface Texture Characterization. Proc. Int. Geoscience and Remote Sensing Symp. (IGARSS'91), 3-6 June, Espoo, Finland, pages 317-320.

Armstrong A., Galli A., Le Loc'h G., Geffroy G. and Eschard R., 2003. Plurigaussian Simulations in Geosciences, Springer.

Armstrong, M., 1984. Common problems with universal Kriging, Mathematical Geology 16(1): 101-108.

Armstrong, M., 1986. On the Fractal Dimensions of some Transient Soil Properties, Journal of Soil Science, 37, 641-652.

Armstrong, M., and Alain G., 2001. Matheron's Contributribution to Geostatistics, Centre de Géostatistique, Fontainebleau, France, 2001 Annual Conference of the International Association for Mathematical Geology, Cancún, Mexico, September 6-12, 2001.

Ashraf, M., Loftis, J.C., and Hubbard, K.G., 1997. Application of Geostatistics to Evaluate Partial Weather Station Networks. Agricultural & Forest Meteorology. 84(3-4): 225-271.

Atkinson, P.M. and Lloyd, C.D., 1998. Mapping Precipitation in Switzerland with Ordinary and Indicator Kriging. Journal of Geographic Information and Decision Analysis, Vol. 2, No. 2.: 72-86. ftp://ftp.geog.uwo.ca/SIC97/Atkinson/Atkinson.html

AWI. 2006. Objective Interpolation Method. Stiftung Alfred-Wegener-Istitut für Polar-und Meeresforschung in der Helmholtz-Gemeinschaft. http://www.awi-bremerhaven.de/Atlas/SO/A4.html

Barnes, R.J. 1991. The variogram sill and the sample variance. Mathematical Geology, Vol. 23, No. 4: 673-678.

Barnes, R. 2005. Variogram Tutorial. Golden Software, Inc. Golden, Colorado 80401-1866, U.S.A. http://www.goldensoftware.com/variogramTutorial.pdf

Barringer, J.R.F. and Lilburne, L., 1997. An Evaluation of Digital Elevation Models for Upgrading New Zealand Land Resource Inventory Slope Data. Presented at the Second Annual Conference of GeoComputation '97 & SIRC '97, University of Otago, New Zealand, 26-29 August 1997.
http://divcom.otago.ac.nz/sirc/webpages/Conferences/SIRC97/97Proceedings/97Barringer.pdf

Bastin, G., Lorent, B., Doque, C., and Gevers, M., 1984 Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations, Water Resources Res. 20(4): 463-470.

Bates, D., Lindstrom, M., Wahba, G. and Yandell, B. 1987 GCVPACK - Routines for Generalized Cross Validation. Communications in Statistics B - Simulation and Computation, 16: 263-297.

Berry, J.K, 1999. Extending Spatial Dependency to Maps. Beyond Mapping, Geoworld, Vol 12 (1): 26-27.

Bertoli, O., 2005. Conditional Simulation. Quantitative Geoscience.
http://www.quantitativegeoscience.com/references_simulation.asp

Billings, S.D., Newsman, G.N., and Beatson, R.K. 2006. Fourier Transformation of Geophysical Data Using Continuous Global Surfaces. University of British Columbia.
http://www.geop.ubc.ca/~sbilling/papers/fourier_transformation_using_cgs.PDF

Binkley, M.R., Carlson, R. and Lee, J., 1997. Spatial Interpolation of Lake Water Quality Data. Kent State University. http://humboldt.kent.edu/~binkley/poster.html

Bogaert, P., Mahau, P. and Beckers, F., 1995. The Spatial Interpolation of Agro-Climatic Data – Cokriging Software and Source Code, Use' Manual, Version 1.0b. Unité de Biométrie, Faculté des Sciences Agronomiaques, Université Catholique de Louvain, Louvain-la-Neuve, Belgium. Agrometeorology Series Working Paper, Number 12, FAO Rome, Italy

Bonham-Carter, G.F., 1994. Geographic Information Systems for Geoscientists: Modelling with GIS, Pergamon, Tarrytown, New York.

Boomer, K.M. and Brazier, R.A., 1997. Geostatistics, Seismic Risk Analysis of Northern Arizona from Acceleration Contour Maps, Enhancing the Sampling Procedure through a Geostatistical Analysis. http://www.geo.arizona.edu/geophysics/students/brazier/geo.html

Booth, T.H. and Jones, P.G. 1996. Climatic Databases for use in Agricultural Management and Research Proceedings of the Arendal II Workshop on UNEP/GRID and CGIAR cooperation to meet requirements for the use of digital data in agricultural management and research. Arundal, Norway 9-11 May 1995.
http://www.grida.no/cgiar/htmls/climpres.htm#climpres

Boutteir F. and Courtier, P. 1999. Data Assimilation Concepts and Methods. European Centre for Medium-Range Weather Forecasts (ECMWF).
http://www.ecmwf.int/newsevents/training/rcourse_notes/DATA_ASSIMILATION/ASSIM_CONCEPTS/Assim_concepts8.html

Bretherton F.P., R.E. Davis and C.B. Fandry. 1976. A Technique for Objective Analysis and Design of Oceanic Experiments applied to MODE-73. Deep Sea Research, 23, p. 559-582.

Bretherton F.P. and J.C. Williams. 1980. Estimations from Irregular Arrays. Reviews of Geophysics and Space Physics, Vol.18, 4, p.789-812.

Briggs, P.R., and Cogley, J.G. 1996. Topographic Bias in Mesoscale Precipitation Networks. Journal of Climate, 9, 205-218.

Brill, D.J., 1997. Skewness and Kurtosis - A Discussion http://www.uark.edu/plscinfo/pub/methods/help/skkrdis.htm

Bruce, J.P. and Clark, R.H., 1980. Introduction to Hydrometeorology. Pergamon Press, Ontario. 324 pages.

Bruno, O. and Pohlman, M. 2002. High-order Surface Interpolation. Applied and Computational Mathematics, California Institute of Technology. http://www.acm.caltech.edu/~bruno/matt_lockheed_12_2_2002.pdf

Bruno, O.P, and Pohlman, M.M., 2003. High-order Surface Representation. California Institute of Technology. http://www.acm.caltech.edu/~bruno/afosrposter03.pdf

Burgess, T.M. and Webster, R., 1980. Optimal Interpolation and Isarithmic Mapping of Soil Properties I. The semivariogram and punctual kriging. Journal of Soil Science 31: 315-331.

Burgess, T.M. and Webster, R., 1980. Optimal Interpolation and Isarithmic Mapping of Soil Properties II. Block kriging. Journal of Soil Science 31: 333-341.

Burrough, P.A., 1981. Fractal dimensions of landscapes and other environmental data. Nature 294:240-242.

Burrough, P.A., 1993a. Fractals and Geostatistical Methods in Landscape Studies. Fractals in Geography (N.S.N. Lam and L. De Cola, editors), Prentice Hall.

Burrough, P.A., 1993b. Soil Variability: A Lake 20th Century View. Soils and Fertilizers, pages 529-562.

Cadell, W., 2002. Report on the Generation and Analysis of DEMs for Spatial Modelling. The Macaulay Institute. http://www.macaulay.ac.uk/ladss/documents/DEMs-for-spatial-modelling.pdf

Calter, P., 1990. Technical Mathematics with Calculus. Prentice Hall, Engelwoods Cliffs, New Jersey.

Carr, James R., 1995. Numerical Analysis for the Geological Sciences. Prentice Hall, Englewoods Cliffs, New Jersey.

Carrasco, P., Ibarra, F., Le Loc'h, G., Rojas, R., Seguret, S., 2006. Application of the Truncated Gaussian Simulation Method to the MM deposit at Codelco Norte, Chile. http://www.geovariances.com/IMG/pdf/EAGE_Madrid_Leloch.pdf

Carrat, F. and Valleron, A., 1992. Epidemiological Mapping Using the "Kriging" Method: Application to an Influenza-like Illness Epidemic in France. American Journal of Epidemiology, 135(11): 1293-1300.

Castrignanò, A. and Lopez, R., 2003. GIS and Geostatistics: An Essential Coupling For Spatial Analysis. Istituto Sperimentale Agronomico. Italian Interest Group on Spatial Statistics. Text: http://194.119.196.14/istituti/ieif/spatialstatistics/res/gg.pdf, PPT: http://194.119.196.14/istituti/ieif/spatialstatistics/res/gg_ppt.pdf

Centre de Geostatistique de l'Ecole des Mines de Paris, 2005. Professor Georges Matheron (1930-2000), A life in probabilistic modelling. 35 Rue St Honore, 77300 Fontainebleau, FRANCE. http://cg.ensmp.fr/Presentation/Matheron/Matheron_en.shtml

Chandila, P.K., 2004. Strategy for Global Optimization and Post-Optimality Using Local Kriging Approximations. A Thesis submitted to the Graduate School of the University of Notre Dame in Partial Fulfillment of the Requirements for the Degree of Master of Science. Graduate Program in Aerospace and Mechanical Engineering, Notre Dame, Indiana. http://etd.nd.edu/ETD-db/theses/available/etd-04142004-144454/unrestricted/ChandilaP042004.pdf

Chapman, A.D., 2003. The Case for a 3-minute Climate Surface for South America. SpeciesLink, Appendix_e.pdf. http://cres.anu.edu.au/outputs/anudem.php

Chapman, A.D., Muňoz, M.E.S., and Koch, I. 2005. Environmental Information: Placing Biodiversity Phenomena in an Ecological and Environmental Context. Biodiversity Informatics, 2, 24-41. http://jbi.nhm.ku.edu/index.php/jbi/article/viewFile/5/3

Chang, T.J. and Teoh, C.B., 1995. Use of the Kriging Method for Studying Characteristics of Ground Water Droughts. Water Resources Bulletin, American Water Resources Association, Vol. 31, No. 6: 1001- 1007.

Cheng, K-S, Wei, C., Cheng, Y-B, and Yeh, H-C. 2003. Effect of Spatial Variation Characteristics on Contouring of Design Storm Depth. Hydrological Processes, 17, 1755-1769. Published online 28 January 2003 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hyp.1209. http://www.rslabntu.net/KSC/Publications/IDF_Contouring.pdf

Chellemi, D.O., Rohrbach, K.G., Yost, R.S. Yost, and Sonoda, R.M., 1988. Analysis of the Spatial Pattern of Plant Pathogens and Diseased Plants Using Geostatistics. The American Phytopathological Society. Vol. 78, No. 2: 221-226.

Chilès, J.-P., and Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty, John Wiley & Sons, New York and Toronto, 695 pages.

Chua, S. H. and Bras, R. L., 1982. Optimal Estimators of Mean Areal Precipitation in Regions of Orographic Influence. J. Hydrology, 57: 23-48.

Clark, I., 1979a. Practical Geostatistics. http://uk.geocities.com/drisobelclark/practica.html)

Clark, I., 1979b. Does Geostatistics Work? 16th International APCOM Symposium, T.J. O'Neil (Ed), McGraw-Hill, New York, 213-225.

Clark, I., 1983. Regression Revisited. Mathematical Geology, Vol. 15, No. 4, 1983.

Clark, I., 1986. The Art of Cross Validation in Geostatistical Applications. 19th Application of Computers and Operations Research in the Mineral Industry, R.V. Ramani , Editor. Sponsored by The Pennsylvania State University, April 14-16, 1986. Published by Society of Mining Engineers, Inc., Littleton, Colorado 1986.

Clark, I., 1987. Turning the tables an interactive approach to the traditional estimation of reserves", J. S. Afr. Inst. Min. Metall., Vol 87, No 10, pp.293-306. Full copy at http://uk.geocities.com/drisobelclark/resume/Publications.html

Clark, I., 1998. Geostatistical Estimation Applied to Highly Skewed Data. Joint Statistical Meetings, Dallas, Texas, August, 1998.

Clark, I., 2000. Erratic highs -- a perennial problem in resource estimation SME 2000, Salt Lake City, Utah.

Clark I., 2001. Practical Geostatistics. Geostokos Limited, Alloa, Scotland.

Clark, I., 2001a. AI-GEOSTATS listserver email: Anisotropy with varying Range and Sill.

Clark, I., 2001b. AI-GEOSTATS listserver email: Kriging of multiple samples.

Clark, I., 2001c. AI-GEOSTATS listserver email: interannual spatial "stability" of variable.

Clark, I., 2002. AI-GEOSTATS listserver email: Log transformation and zeros.

Clark, I., 2006a. AI-GEOSTATS listserver email: spherical model.

Clark, I., 2006b. AI-GEOSTATS listserver email: kriging without a nugget.

Clark I. and Harper,W.V., 2000. Practical Geostatistics 2000. Ecosse North America Llc, Columbus, Ohio, U.S.A.

Clark I. and Harper,W.V., 2000. Practical Geostatistics 2000, Answers to the Exercises. Ecosse North America Llc, Columbus, Ohio, U.S.A.

Cox D.D., Cox L.H., and Ensor, K.B., 1997. Spatial Sampling and the Environment - Some Issues and Directions. Environmental & Ecological Statistics. 4(3): 219-233.

Cressie, N., 1985. Fitting variogram models by weighted least squares. Mathematical Geology 17: 563 - 586.

Cressie N., 1990. The Origins of Kriging. Mathematical Geology, Vol. 22, No. 3, 1990: 239 - 252.

Cressie, N.A.C. 1991. Statistics for Spatial Data. John Wiley, New York.

Cressie, N., 1993. Statistics for Spatial Data. 2nd edition. John Wiley and Sons, Inc, New York.

Cressie, N., 2003. Statistics for Spatial Data. Revised edition. John Wiley and Sons, Inc, New York.

Cressie, N.A.C. and Hawkins D.M., 1980. Robust Estimation of the Variogram. Journal International Association Mathematical Geology 12:115-125.

Cressman, G., 1959. An Operation Objective Analysis System. Monthly Weather Review, 87, 367-374.

Creutin, J.D., and Obled. C., 1982. Objective Analysis and Mapping Techniques for Rainfall Fields: an objective comparison. Water Resource Research 18(2):413-431.

Daly, C. 1996. Overview of the PRSIM Model. http://www.ocs.orst.edu/prism/docs/overview.html

Daly, C. 2006. Guidelines for Assessing the Suitability of Spatial Climate Data Sets. International Journal of Climatology. 26: 707-721. Published online 27 March 2006 in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/joc.1322.

Daly, C., Gibson, W.P., Taylor, G.H. Johnson, G.L., and Pasteris P. 2002. A Knowledge-based Approach to the Statistical Mapping of Climate. Climate Research 22: 99-113.

Daly, C., Gibson, W.P., Doggett, M., Smith, J and Taylor, G. 2004. A Probabilistic-spatial Approach to the Quality Control of Climate Observations. In Proceedings of the 14th AMS Conference on Applied Climatology, American Meteorological Society, Seattle, Washington, January 13-16, 2004. http://www.ocs.orst.edu/pub/prism/docs/appclim04-probabilistic_spatial_approach-daly.pdf

Daly, C., Hlemer, E.H., Quinones, M. 2003. Mapping the Climate of Puerto Rico, Vieques, and Culebra. International Jounal of Climatology 23: 1359-1381.

Daly, C, and Johnson, G.L. 1999. PRISM Spatial Climate Layers: Their development and use. Short course on topics in applied climatology. 79th Annual Meeting of the American Meteorological Society, American Meteorological Society: Dallas, Texax, 10-15 January. http://www.ocs.orst.edu/pub/prism/docs/prisguid.pdf

Daly, C., Nielson, R.P., and Phillips, D.L. 1994. A Statistical-topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. Journal of Applied Meteorology, 33:140-158. http://www.ocs.orst.edu/prism/docs/overview.html

Dandurand, L.M., Knudsen, G.R., and Schotzko, D.J., 1995. Quantification of *Phythium Ultimum* var. *sporangiiferum* Zoospore Encystement Patterns Using Geostatistics. Phytopathology, Vol. 85, No. 2: 184 - 190.

David, M., 1977. Geostatistical Ore Reserve Estimation, Elsevier, 364 pages.

Davis, J.C., 1973. Statistics and Data Analysis in Geology. John Wiley and Sons, New York.

Davis, J.C., 1986. Statistics and Data Analysis in Geology, 2nd edition, Wiley, New York.

Davis, B.M., and Borgman L.E., 1979. A Test of Hypothesis Concerning a Proposed Model for the Underlying Variogram, Proceedings 16th APCOM, 163-181.

De Cola, L. 1989. Fractal Analysis of a Classified Landsat Scene. Photogrammetric Engineering & Remote Sensing, 55:601-610.

De Jong, S.M. 1993. An Application of Spatial Filtering Techniques for Land Cover Mapping Using Tm Images, Geocarto International, 8(1):43-49.

Delfiner, P. and Delhomme, J.P. 1975. Optimum Interpolation y Kriging. Display and Analysis of Spatial Data, Davis, J.C. and McCullagh, M.J. (Eds.) Wiley and Sons, New York. pp. 96-114.

Delhomme, J.P., 1979. Kriging in the Hydrosciences, Advance in Water Res. 1(5): 251-266.

Deraisme. J., 1997. Interpolation and Reality. Geovariances News Letter #7, page 3. http://www.geovariances.fr/publications/n1/7/page3.html

Deraisme, J. and Farrow, D. 2004. Geostatistical Simulation Techniques Applied to Kimberlite Orebodies and Risk Assessment of Sampling Strategies. Banff. http://www.geovariances.com/IMG/pdf/Banff2004.pdf

Deutsch, C.V., 1996. Correcting for Negative Weights in Ordinary Kriging. Computers and Geosciences Vol. 22, No. 7: 765-773.

Deutsch, C.V. and A.G. Journel 1992. GSLIB Geostatistical Software Library. Oxford University Press, New York. www.gslib.com

Deutsch, C.V. and A.G. Journel 1997/1998. GSLIB Geostatistical Software Library and Users Guide, 2nd Edition. Oxford University Press, New York.

Dingman, S.L., Seely-Reynolds, D.M., and Reynolds III R.C. 1988. Application of Kriging to Estimate Mean Annual Precipitation in a Region of Orographic Influence. Water Resource Bulletin, 24,329-339.

Dorsel, D. and La Breche, T., 1998. Environmental Sampling & Monitoring Primer, Kriging. http://www.ce.vt.edu/enviro2/smprimer/kriging/kriging.html

Duchon, J. 1976. "Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces." RAIRO Analyse Numérique 10, 5-12.

Duval, R., Lévy, R., Matheron, G., 1955. Travaux de D. G. Krige sur l'évaluation des gisements dans les mines d'or sud-africaines. *Annales des Mines*, Paris, vol. 12, p. 3-49.

Eastman, J., 1999. IDRISI32, Guide to GIS and Image Processing – Volume 2, p 139 and p 155-161. Clark Labs, Clark University.

Eastman, J., 1999. IDRISI32, Tutorial. 179-201. Clark Labs, Clark University.

Ehlschlaeger, C.R., 2001. Representing Multiple Spatial Statistics in Generalized Elevation Uncertainty Models: Moving Beyond the Variogram. Department of Geography, Hunter College, New York, New York.
http://www.geo.hunter.cuny.edu/~chuck/RGSIBG/beyondVariogramSmallMaps.pdf

Englund, E.J. Undated. Spatial Autocorrelation: Implications for Sampling and Estimation. U.S. Environmental Protection Agency, P.O. Box 93478, Las Vegas, Nevada 89193-3478.
http://www.epa.gov/nerlesd1/cmb/research/papers/ee107.pdf

Englund, E.J., 1990. A Variance of Geostatisticians. J. Math. Geology. vol. 22, no. 4, pp. 417-455.

Englund, E.J., 1993. Spatial Simulation: Environmental Applications. A version of this manuscript has been published in Environrnental Modeling with GIS, M.F. Goodchild et al., eds., Oxford University Press (1993) 432-437.
http://www.epa/gov/nerlesd1/cmb/research/papers/ee104.pdf

Franke, R. 1985. Thin-plate Splines with Tension. Computer Aided Geometric Design 2, 87-95, North-Holland.

Freek, D., 1997. Classification of High Spectral Resolution Imagery Using an Indicator Kriging Based Technique. International Institute for Aerospace Survey and Earth Sciences (ITC), Department of Earth Resources Surveys. 350 Boulevard 1945, P.0. Box 6, 7500 AA, Enschede, The Netherlands.

Galli, A., Beucher H., Le Loc'h, G., Doligez, B. and Group H., 1994. The pros and cons of the truncated Gaussian method, in Geostatistical Simulations, pp. 217-233, Kluwer Academic, Dordrecht.

Gamma Design. 2004. GS+ for Windows is a Windows 95/98/NT version of a geostatistical mapping program for the PC. Gamma Design Software, LLC, Plainwell, Michigan 49080 USA. http://www.gammadesign.com

Gambolatic, G. and Volpi, G., 1979a. A Conceptual Deterministic Analysis of the Kriging Technique in Hydrology. Water Resources Res. 15(3): 625-629.

Gandin, L.S., 1963. Objective Analysis of Meteorological Fields: Gidrometeorolgicheskoe Izdatel'stvo (GIMIZ), Leningrad (translated by Israel Program for Scientific Translations, Jerusalem, 1965, 238 pages.

Gandin, L.S., and Kagan, R.L. 1974. Construction of a System for Objective Analysis of Heterogeneous Data Based on the Method of Optimum Interpolation and Optimum Agreement. Meteorological I Gidrol., No. 5 (Moscow), 1-11. [Translated by Joint Publications Research Service.]

Garrott, J., 2003. SAKWeb: (Spatial Autocorrelation and Kriging Web) A W3 Geocomputation Perspective. Universidade Nova De Lisboa.
http://www.isegi.unl.pt/labnt/presentations/2003/031215Garrot.pdf

Geary, R., 1954. The Contiguity Ratio and Statistical Mapping. The Incorporated Statistician, Volume 5, 115-145.

Genoton,, M.G. and Furrer, R., 1998. Analysis of Rainfall Data by Simple Good Sense: is Spatial Statistics Worth the Trouble? Journal of Geographic Information and Decision Analysis, Vol. 2, No. 2: 11-17. ftp://ftp.geog.uwo.ca/SIC97/Genton_1/Genton_1.html

Genton, M.G. and Furrer, R., 1998. Analysis of Rainfall Data by Robust Spatial Statistics using S+Spatialstats. Journal of Geographic Information and Decision Analysis, Vol. 2, No. 2: 12 -136. ftp://ftp.geog.uwo.ca/SIC97/Genton_2/Genton_2.html

Gibson, W.P., Daly, C., Kittel, T., Nychka, D, Johns, C., Rosenbloom, N., McNab, A. and Taylor, G. 2002. Development of a 103-year High-resolution Climate Data Set for the Conterminous United States in Proceedings 13th AMS Conference on Applied Climatology, American Meteorological Society, Portland, Oregon, may 13-16, 2002. pp 181-183. http://www.ocs.orst.edu/pub/prism/docs/appclim02-103yr_hires_dataset-gibson.pdf

Glasbey, C.A., and Mardia, K.V. 1998. A Review of Image Warping Methods. Journal of Applied Statistics, 25, 155-177. http://www.bioss.sari.ac.uk/staff/chris/warp.pdf

Glover, D.M., 1998,2000. 12.747 Lecture 5:Section 4: Objective Mapping and Kriging. Woods Hole Oceanographic Institution. http://w3eos.whoi.edu/12.747/notes/lect05/l05s04.html

Goodchild, M., 1999. Spatial Analysis and GIS. 1999 ESRI User Conference, Pre-Conference Seminar.

Goodman, A., 1973,1999. Trend Surface Analysis in the Comparison of Spatial Distributions of Hillslope Parameters. http://www.deakin.edu.au/~agoodman/masters/

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York. 483 pages.

Goovaerts, P., 1999. Performance Comparison of Geostatistical Algorithms for Incorporating Elevation into the Mapping of Precipitation. Proceedings of, the IV International Conference on GeoComputation, Mary Washington College in Fredericksburg, VA, USA, 25-28 July 1999. http://www.geovista.psu.edu/sites/geocomp99/Gc99/023/gc_023.htm

Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228 (2000) 113-129. http://www.terraseer.com/training/geostats/jhydrol00.pdf

Goovaerts, P., 2001. Geostatistical Modelling of Uncertainty in Soil Science. Geoderma 103:3-26.

Goovaerts, P. 2002. AI-GEOSTATS: Tuning a semivariogram.

Goovaerts, P. 2005. Geostatistical Analysis of Disease Data: Estimation of cancer mortality risk from empirical frequencies using Poisson kriging. International Journal of Health Geographics 4:31. http://www.ij-healthgeographics.com/content/pdf/1476-072X-4-31.pdf

Goovaerts, P. 2006. Geostatistical Analysis of Disease Data: Visualization and propagation of Spatial Uncertainty in Caner Mortality Risk using Poisson kriging and p-field simulation. International Journal of Health Geographics 2006, 5:7. http://www.ij-healthgeographics.com/content/pdf/1476-072X-5-7.pdf

Gottwald, T.R., Avinent, L, Llacer, G., de Mendoza, H. and Cambra, M., 1995. Analysis of the Spatial Spread of Sharka (Plum Pox Virus) in Apricot and Peach Orchards in Eastern Spain. Plant Disease. Vol. 79 No. 3: 266-278.

Gottwald, T.R., Cambra, M., Moreno, P, Camarasa, E. and Piquer, J., 1996. Spatial and Temporal Analyses of Citrus Tristeza Virus in Eastern Spain. The American Phytopatholgical Society. Vol. 86, No. 1: 45 - 55.

Griffith, D.A., 1987. Spatial Autocorrelation. A Primer. Resource Publication in Geography, Association of American Geographers, Washington D.C.

Griffith, D.A., Amrhein, C.G. and Desloges, J.R., 1991. Statistical Analysis for Geographers. Prentice Hall, Englewoods Cliffs, New Jersey.

Guocheng, P. Garrd, D. Moss, K. and Heiner, T., 1993. A Comparison Between Cokriging and Ordinary Kriging: Case Study with a Polymetallic Deposit. Mathematical Geology, Vol. 25, No. 3: 377-398.

Haan, C.T., 1977. Statistical Methods in Hydrology. Iowa State University Press, Ames, Iowa.

Handley, J.R.F., 2004. Historic Overview of the Witwatersrand Goldfields. Available from the Geological Society of South Africa. ISBN 0-620-32127-X. 224 pages.

Hansen, R.O., 1993. Interpretive Gridding by Anisotropic Kriging. Geophysics, Vol. 58, No. 10: 1491-1497.

Hengl, T. 2006. Regression-Kriging. Spatial-analyst.net. http://spatial-analyst.net/regkriging.php

Hengl, T., Heuvelink, G.B.M., and Stein, A. 2003. Comparison of Kriging with External Drift and Regression-kriging. Technical note, ITC, Available on-line at http://www.itc.nl/library/Academic_output/
http://spatial-analyst.net/GRK/Hengl_et_al_Comparison_RK_KED.pdf

Henley, S. 2001. Geostatistics – Cracks in the Foundations? Stephen Henley, Resources Computing International Ltd. Derbyshire, U.K.
http://www.infomine.com/softwaremine/articles/GeostatCracks.asp and
http://www.earthsciswinfo.com/swmine/swarticles/GeostatCracks.html

Herzfeld, U.C, Eriksson, M.G., and Homlund P., 1993. On the Influence of Kriging Parameters on the Cartographic Output – A study in Mapping Subglacial Topography. Mathematical Geology, Vol. 25, No. 7: 881-900.

Hevesi, J.A., and Istok, J.D. 1992. Precipitation Estimation of Mountainous Terrain Using Multivariate Geostatistics. Part I: Structure Analysis. Journal of Applied Meteorology, 31, 661-667.

Hijmans, R.J., Cameron, S., and Para J. 2004a. WorldClim, version 1.2. A Square Kilometre resolution database of global terrestrial surface climate.

Hijmans, R.J., Cameron, S., and Para J. 2004b. WorldClim, a new high-resolution global climate database. Abstract: Inter-American Workshop on Environmental Data Access. Campinas, Brazil.

Hiller W. and R. Käse. 1983. Objective Analysis of Hydrographic Data Sets from Mesoscale Surveys. Berichte aus dem Institut für Meereskunde an der Christian-Albrechts-Universität Kiel, 116**,** 78 pages.

Holawe, F. and Dutter, R., 1999. Geostatistical Study of Precipitation Series in Austria: Time and Space. Vienna University of Technology.

Hu, J., 1998. Methods of Generating Surfaces in Environmental GIS Applications. Bechtel Environmental Inc., 151 Lafayette Drive, Oak Ridge, Tennessee 37922.
http://pasture.ecn.purdue.edu/~edu/~aggrass/esri95/to100/p089.html

Hutchinson, M.F., 1988. Calculation of Hydrologically Sound Digital Elevation Models. Proceedings of the Third International symposium on Spatial Data Handling, August 17-19, Sydney. International Geographic Union, Columbus, Ohio.

Hutchinson, M.F., 1989. A New Method for Gridding Elevation and Stream Line Data with Automatic Removal of Pits. Journal Hydrology 106:211-232.

Hutchinson, M.F., 1990. Robust Calibration of Seasonally Varying Stochastic Weather Models Using Smoothing Splines. Math. Compt. Simulation, 32, 125-130.

Hutchinson, M.F. 1991a. The Application of Thin-plate Smoothing Splines to Continent-Wide Data Assimilation. In BRMC Research Report No. 27, Data Assimilation Systems, Bureau of Meterology, Melbourne, Jasper, J.D. (Ed.) pp 104-113. In Data Assimilation Systems: papers presented at the 2nd BMRC modeling workshop, September 1990. BMRC Research Report No. 27, 104-113.

Hutchinson, M.F. 1991b. Climatic Analysis in Data Sparse Regions. Climatic Risk in Crop Production, Muchow, R.C. and Bellamy, J.A. (Eds), CAB International, pp. 55-71.

Hutchinson, M. F. 1993. On Thin-plate Splines and Kriging. Computing Science and Statistics, *Vol. 25*. Edited by M.E.Tarter and M.D.Lock, pp. 55-62. Berkeley: Interface Foundation of North America.

Hutchinson, M.F. 1994. ANUSPINE (Revisions 1 July 1991). Centre for Research and Environmental Studies. Available from CRES, Australian National University, G.P.O. Box 4, Canberra, ACT 2601, Australia.

Hutchinson, M. 1995. Interpolating Mean Rainfall Using Thin-plate Smoothing Splines. International Journal of Geographic Information Systems, 9, 385-403.

Hutchinson, M. 1998a. Interpolation of Rainfall Data with Thin-plate Smoothing Splines – Part I: Two Dimensional Smoothing of Data with Short Range Correlation. Journal of Geographic Information and Decision Analysis, vol. 2, no. 2, pp. 139-151, 1998. http://www.ai-geostats.org/events/sic97/Hutchinson_1.pdf

Hutchinson, M. 1998b. Interpolation of Rainfall Data with Thin-plate Smoothing Splines – Part II: Analysis of Topographic Dependence. Journal of Geographic Information and Decision Analysis, vol. 2, no. 2, pp. 152-167, 1998. http://www.ai-geostats.org/events/sic97/Hutchinson_2.pdf

Hutchinson, M.F. 2006. ANUDEM Version 5.2. Centre for Resource and Environmental Studies, The Australian National University, Caneberra. http://cres.anu.edu.au/outputs/anudem.php

Hutchinson, M. F. and Bischof, R. J. 1983. A New Method for Estimating Mean Seasonal and Annual Rainfall for the Hunter Valley, New South Wales. Australian Meteorological Magazine, 31; 179-184.

Hutchinson, M. and P. Gessler. 1994. Splines – more than just a smooth interpolator. *Geoderma*, 62, 45-67.

Hutchinson, M.F., Kalma, J.D., and Johnson, M.E. 1984. Monthly Estimates of Wind Speed and Wind Run for Australia. Journal of Climatology, 4, 311-324

Inggs, M.R. and Lord R.T., 1995. Interpolating Satellite Derived Wind Field Data Using Ordinary Kriging, with Application to the Nadir Gap, Proc. IEEE Geoscience Remote Sensing Symp., IGARSS'95, Firenze, Italy, vol. 1, pp. 141-143, July 1995.

Ingram, P., 1998. An Introduction to Geostatistics. School of Earth Sciences, and Kriging. Macquarie University, Sydney, NSW, 2109, Australia. http://137.111.98.10/users/pingram/geostat.html and http://atlas.es.mq.edu.au/users/pringram/krig.html#hrgprcss

Isaaks, E.H., 1989. Applied Geostatistics. Oxford University Press, New York and Toronto. 561 pages.

Isaaks, E.H. and Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.

Jaakkola, O.A. and Oksanen, J., 2000. Creating DEMs from Contour Lines --- Interpolation Techniques Which Save Terrain Morphology. GIM International, September.

Jacquez, G., 2005. Exploratory Spatial Data Analysis. University of Michigan, Department of Epidemiology. http://zappa.nku.edu/~longa/geomed/modules/esda/index.html

Jones, J.G., Thomas, R.W., and Earwicker, P.G. 1989. Fractal Properties of Computer-Generated and Natural Geophysical Data. Computer & Geosciences, 15(2):227-235.

Jones, T.A., Hamilton, D.E., and Johnson, C.R., 1986. Contouring Geological Surfaces With the Computer. Van Nostrand Reinhold, London, England.

Jones, P.G. 1995. Centro Internacional de Agricultura (CIAT) climate database version 3.41, Digital data tape, Cali, Columbia: CIAT.

Jones, P.G., Robison, D.M. and Carter, S.E. 1990. A geographical information approach for stratifying tropical Latin America to identify research problems and opportunities in natural resource management for sustainable agriculture in Centro Internacional de Agricultura Tropical (CIAT). Cali, Colombia: Agroecological Studies Unit., CIAT.

Journel, A.G., 1974**.** Geostatistics for Conditional Simulation of Ore Bodies. Economic Geology: Volume 69 pp.673-687.

Journel, A.G., 1983. Mathematical Geology, 445-468.

Journel, A.G. and Arik, A., 1988. In Computer Applications in the Mineral Industry; Fytas, et. Al., Editors; Balkema: Rotterdam; 161-171.

Journel, A.G. and Hujibregts, C.J., 1978. Mining Geostatistics. Academic Press, New York.

Kalma, J.D., Speight, J.D. and Wasson, R.J. 1988. Potential Wind Erosion in Australia: A continental perspective. Journal of Climatology, 8, 411-428.

Kent, J.T. and Mardia, K.V. 1994. The link between kriging and thin-plate splines. In Probability, Statistics and Optimization (ed. F.P. Kelly). Wiley, New York. 324-339.

Kesterven J., and Hutchinson, M. 1996. Spatial Modeling of Climate Variables on a Continental Scale. Proceedings of the Third International Conference Integrating GIS and Environmental Modeling. University of California, Santa Barbara, National Center for Geographic Information and Analysis: CD-ROM.

Kitanidis, P., 1985. 'Minimum-Variance Quadratic Estimation of Covariances of Regionalized Variables.' Mathematical Geology, 17(2):195-208.

Kitanidis, P.K., 1997. Introduction to Geostatistics: Applications to Hydrogeology. Cambridge University Press, New York. 249 pages.

Kitterod NO. Langsholt, E. Wong, W.K. and Gottschalk, L., 1997. Geostatistic Interpolation of Soil Moisture. Nordic Hydrology. 28(4-5): 307-328.

Klinkenbery, B., 2002. Geography 516, Graduate GIS Seminar: Linking Statistical Models with Spatial Data. Department of Geography, University of British Columbia, Vancouver, B.C. http://www.geog.ubc.ca/courses/geog516/notes/statistics_spatial.htm

Klinkenberg, B. and Goodchild, M.F. 1992. The Fractal Properties of Topography: A Comparison of Methods. Earth Surface Process and Landforms, Vol. 17, 217-234.

Knudsen, H.P. and Baafi, E.Y., 1987. Indicator Kriging and Other New Geostatistical Tools. Pacific Rim Congress 87.

Kohl, M. and Gertner, G., 1997. Geostatistics in Evaluating Forest Damage Surveys - Considerations on Methods for Describing Spatial Distributions. Forest Ecology & Management. 95(20): 131-140.

Krige, D.G., 1951. A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand. Journal of the Chemical and Metallurgical Mining Society of South Africa 52: 119-139.

Krige, D.G., 1959. A Study of the relationship between development values and recovery grades on the South African goldfields". Journal of the South African Institution of Mining and Metallurgy, No. 61, 317-331.

Krige, D.G., 1966. Two Dimensional Weighted Moving Average Trend Surfaces for Ore-Evaluation. Journal of the South African Institute of Mining and Metallurgy 66:13-38.

Krige, D.G., 1981. Lognormal-de Wijsian geostatistics for ore evaluation. South African Institute of Mining and Metallurgy Monograph Series. Geostatistics I. South Africa Institute of Mining and Metallurgy, Johannesburg, South Africa.

Krige DG, Guarascio M, and Camisani-Calzolari FA., 1989. Early South African Geostatistical Techniques in Today's Perspective. In: Armstrong M, editor. Geostatistics. Amsterdam: Kluwer Academic Publisher. p 1-19.

Krivoruchko, K., 1998. GIS and Geostatistics: Spatial Analysis of Chernobyl's Consequences in Belarus. Workshop on Status and Trends in Spatial Analysis. Santa Barbara, California., December 10-12, 1998. 20 pages. http://www.ncgia.ucsb.edu/conf/sa_workshop/papers/ncgia.html

Krivoruchko, K., 2000. Applied Reservoir Characterization Using Geostatistics: The Value of Spatial Modeling. AAPG Hedberg Conference, Houston, Texas, December 3-6.

Krivoruchko, K., 2002a. Introduction to Modeling Spatial Processes Using Geostatistical Analyst. ESRI Australia. http://zappa.nku.edu/~longa/geomed/modules/esda/index.html

Krivoruchko, K. 2002b. Bridging the Gap Between GIS and Solid Spatial Statistics. New Tools for Spatial Data Analsysis: Procceding of the CSISS Specialist Meeting, Santa Barbara, California, May 10-11, 2002. University of California, Santa Barbara, Center for Spatially Integrated Social Science. http://www.dpi.inpe.br/gilberto/csiss/papers/krivoruchko.pdf

Krivoruchko, K., 2002. Working on Nonstationarity Problems in Geostatistics Using Detrending and Transformation Techniques: An Agricultural Case Study. paper was presented at the Joint Statistical Meetings, New York City, August 2002. Available at http://www.esri.com/software/arcgis/arcgisxtensions/geostatistical/research_papers.html and http://campus.esri.com/campus/library/ConfProc/GeostatisticsTeam/Krivoruchko_2002_Working.pdf

Krivoruchko, K. 2005. Introduction to Modelling Spatial Processes Using Geostatistical Analyst. ESRI White Papers. http://www.esri.com/library/whitepapers/pdfs/intro-modeling.pdf

Kufs, C., 1987. Trend-Surface Modeling of Groundwater Data. Management of Uncontrolled Hazardous Waste Sites. Silver Springs, MD. Hazardous Materials Control Research Institute. http://members.home.net/terrabyte/t_surf.pdf

Kufs, C., 1998. Interpolation for Fun and Profit. http://members.home.net/terrabyte/didyou.htm

LaGro, Jr., J. 1991. Assessing Patch Shape in Landscape Mosaics. Photogrammetric Engineering & Remote Sensing, 57(3):285-293.

Larkin, R.P., Gumpertz, M.L., and Ristaino, J.B., 1995. Geostatistical analysis of *Phytophthora* epidemic development in commercial bell pepper fields. Phytopathology 85:191-203.

Le Loc'h, Beucher G., Galli, A., Doligz, B., and Group, H., 1994. Improvement in the truncated Gaussian method: Combining several Gaussian Functions, in Proceedings of ECMORE IV, Fourth European Conference on the Mathematics of Oil Recovery.

Le Loc'h, G. and Galli, A. 1997 Truncated plurigaussian method: Theoretical and practical points of view, in Geostatistical Simulations, pp. 211-222, Kluwer Academic.

Lecoustre, R. Fargette, D., Fauquet, C. and de Reffye, P., 1989. Analysis and Mapping of the Spatial Spread of African Cassava Mosaic Virus Using Geostatistics and the kriging technique. The American Phytopatholgical Society. Vol 79, No. 9: 913-920.

Leduc, A., Prairie, Y.T., Bergeron, Y., 1994. Fractal dimension estimates of a fragmented landscape: sources of variability. Landscape Ecology Volume 9 Number 4 pp 279-286. SPB Academic Publishing bv, The Hague. http://landscape.forest.wisc.edu/landscapeecology/articles/v09i04p279.pdf

Liebhold, A.M., Rossi, R.E., and Kemp, W.P., 1993. Geostatistics and Geographic Information Systems in Applied Insect Ecology. Gypsy Moth Slow the Spread Foundation, Inc. http://www.gmsts.org/pubs/liebhold_1993.pdf

Lembo, A.J., Jr., 2005. Spatial Interpolation, Lecture Power Point Presentation, Cornell University. http://www.css.cornell.edu/courses/420/lecture18.ppt

Lemmar, I.C., 1985. The Mononodal Cutoff-Application to Robust Variography and Spatial Distribution Estimation, Ph.D. Thesis, Stanford University.

Levine, N. & Associates. 1999. CrimeStat (Version 1.0) – A Spatial Statistics Program for the Analysis of Crime Incident Locations. The National Institute of Justice, Washington, D.C. http://www.ojp.usdoj.gov/cmrc/

Little, L.S., Edwards, D. and Porter, D.E., 1997. Kriging in Estuaries - As the Crow Flies, Or As the Fish Swims. Journal of Experimental Marine Biology & Ecology. 213(1):1-11.

Liu, H.K., Jezek, B. L., and Zhao, Z. 2001. Radarsat Antarctic Mapping Project Digtial Elevation Model Version 2. NDSIC, Boulder, Colorado.

Liu, N., and Oliver, D., 2003. Conditional Simulation of Truncated Random Fields Using Gradient Methods. International Association for Mathematical Geology (IAMG). IMAG 2003 Proceedings. http://www.iamg.org/meetings/Proceedings_2003/papers/Liu.pdf

Loránd, 2005. Kriging Interpolation. Eötvös Loránd University, Faculty of Informatics, Department of Cartography and Geoinformatics. http://lazarus.elte.hu/hun/digkonyv/havas/mellekl/vm25/vma07.pdf

Lorenc, A., 1981: A Global Three-Dimensional Multivariate Statistical Interpolation Scheme. Monthly Weather Review, 109, 701-721.

Lovejoy, S., 1982. Area-Perimeter Relation for Rain and Cloud Areas. Science, 216:185-187.

Lynch, S.D. and Schulze, R.E., 1997. Techniques for Estimating Areal Daily Rainfall. Department of Agricultural Engineering, University of Natal, Private Bag X01, Scottsville, 3209, South Africa. http://www.ccwr.ac.za/~lynch2/p241.html

Maclean, A.L., Cleland, D.T., The Use of Geostatistics to Determine the Spatial Extent of Historical Fires as an Aid in Understanding Fire Regimes for Northern Lower Michigan. USDA Forest Service, Great Lakes Ecological Assessment Reports. In press: Proceedings of a Conference on Fire, Fuel Treatments and Ecological Restoration. April 16-18, 2002. Ft.Collins, CO. Rocky Mountain Research Station General Technical Report. http://www.ncrs.fs.fed.us/gla/reports/geostatisticsinfire.pdf

Mandelbrot, B.B., 1975. Stochastic Models for the Earth's Relief, the Shape and the Fractal Dimension of Coastlines, and the Number-Area Rule for Islands. Proceedings of the National Academy of Sciences, U.S.A. 72, 3825-3828.

Mandelbrot, B.B., 1977. Fractals, Form, Chance and Dimension. Freeman, San Francisco.

Mandelbrot, B.B., 1982. The Fractal Geometry of Nature. W.H. Freeman, London, New York.

Mandelbrot, B.B. and van Ness, J.W. 1968. Fractional Brownian Motions, Fractional Noises and Applications, Siam Review, 10(4), 422-437.

Mantoglou, A. and Wilson, J.L., 1982. The turning bands method for simulation of random fields using line generation by a spectral method: Water Resources Research Vol. 18, No.5, pp.1379-1394.

Mark, D.M., 1997. The History of Geographic Information Systems: Invention and Reinvention of Triangulated Irregular Networks (TINs). Proceedings, GIS/LIS'97. http://www.geog.buffalo.edu/ncgia/gishist/GISLIS97.html

Matheron, G., 1962. Traité de geostatisque appliqueée, vol. I: Memories du Bureau de Géologiques et Miniéres, no 14, Editions Techniq, Paris. 333 pages.

Matheron, G., 1963a. Principles of Geostatistics. Economic Geology 58: 1246-1266.

Matheron, G., 1963b. Traité de geostatistique appliqée, vol. II, Le krigeage: Memories du Bureau de Recherches Géologiques et Miniéres, no 24. Éditions Bureau de Recherche Géologiques et Miniéres. Paris. 171 pages.

Matheron, G., 1965. Les Variables Régionalisées et Leur Estimation. Une Application de la Thérie des Fonctions Aléatoires aux Sciences de la Nature, Masson, Paris. 305 pages.

Matheron, G., 1971. The Theory of Regionalized Variables and its Applications. Cahiers du Center de Morphologie Mathematique, Fontainbleau, No. 5.

Maurer, B.A., 1994. Geographical Population Analysis. Blackwell Scientific, Oxford, UK.

Maurer, B.A. and Villard, M.-A., 1994. Population Density: Geographic Variation in Abundance of North American birds. National Geographic Research and Exploration 10: 306-317.

Meinguet, J. 1979. Multivariate Interpolation at Arbitrary Points Made Simple, Journal of Applied Mathematics and Physics, 30, pp. 292–304

McBratney, A.B. and Webster, R., 1981. Spatial Dependence and Classification of the Soil along a Transect in Northeast Scotland. Geoderma, 26:63:82.

McBratney, A.B. and Webster, R., 1986. Choosing Functions for Semi-variograms of Soil Properties and Fitting them to Sampling Estimates, Journal of Soil Science 37: 617-639.

Milewska, E.J., Hopkinson, R.F. and Niitsoo, A. 2005. Evaluation of Geo-Referenced Grids of 1961-1990 Canadian Temperature and Precipitation Normals. Atmosphere-Ocean 43 (1) 2005, 49-75. Canadian Meteorological and Oceanographic Society. http://www.cmos.ca/Ao/Abstracts/v430104.pdf

Mitchel, N.D. 1991. The Derivation of Climate Surface for New Zealand, and Their Application to the Bioclimatic analysis of the Distribution of Kauri (Agathis australis). Journal Royal Society of New Zealand. 21, 13-24.

Moran, P.A.P., 1948. The interpretation of statistical maps. Journal of the Royal Statistical Society, Series, B, 10, 243-250.

Moran, P.A.P., 1950. Notes on Continuous Stochastic Phenomena. Biometrika, 37, 1950; 17-23.

Mitasova, H. and Mitas, L., 1993. Interpolation by Regularized Spline with Tension: I. Theory and Implementation. Mathematical Geology, Vol. 25, No. 6: 641-655.

Mitasova, H. and Hofierka, J., 1993. Interpolation by Regularized Spline with Tension: II. Application to Terrain Modeling and Surface Geometry Analysis. Mathematical Geology, Vol. 25, No. 6: 657-669.

Myers, J., 1997. Geostatistical Error Management: Quantifying Uncertainty for Environmental Sampling and Mapping, Van Nostrand Reinhold, New York.

Myers, D.E., 1999. What is Geostatistics? http://www.u.arizona.edu/~donaldm/whatis.html

Nalder, I.A. and Wein, R.W. 1998. Spatial interpolation of climatic Normals: test of a new method in the Canadian boreal forest. Agric. For. Meteorol. 92: 211-225.

Ned Levine & Associates, 2004. Chapter 5 – Distance Analysis I and II. CrimeStat III, The national Institute of Justice, Washington, DC. http://www.icpsr.umich.edu/NACJD/crimestat.html/ and http://www.icpsr.umich.edu/CRIMESTAT/files/CrimeStatChapter.5.pdf

New, M., and Hulme. M. 1997. Monthly Rainfall and Temperature Surfaces for Africa: 1951-1995 (CD-ROM) information: Production of Climate Surfaces. Climatic Research Unit, University of East Anglia Norwich NR4 7TJ, UK. http://www.mara.org.za/climatecd/info.htm

NIST, 2005. Engineering Statistics Handbook. NIST/SEMATECH e-Handbook of Statistical Methods, Home: http://www.itl.nist.gov/div898/handbook/index.htm, Normal Probability Plot: http://www.itl.nist.gov/div898/handbook/eda/section3/normprpl.htm, Quanitile-Quantile Plot: http://www.itl.nist.gov/div898/handbook/eda/section3/qqplot.htm

Olea, R., Ed., 1991. Geostatistical Glossary and Multilingual Dictionary, New York: Oxford University Press.

Oliver, M.A. and Webster, R., 1991. How Geostatistics Can Help You. Soil Use and Management. Volume 7, Number 4, December 1991: 206-217.

Oliver, M., Webster, R. and Gerrard, J., 1989. Geostatistics in Physical Geography. Part I: theory. Trans. Inst. Br. Geogr. N.S. 14: 259-269 (1989) ISSN:0020-2754.

Oliver, M. Webster, R. and Gerrard, J., 1989. Geostatistics in Physical Geography. Part II: applications. Trans. Inst. Br. Geogr. N.S. 14: 270 -286 (1989) ISSN:0020-2754.

Omre, H., 1984. The Variogram and its Estimation, in Verly, G., David, M., Journel, A. G. and Marechal, A. (eds) Geostatistics for Natural Resources Characterization (Reidel, Dordrecht): 107-125.

Onsoy, Y.S., Harter, T., Ginn, T.R., and Horwath, W.R., 2005. Spatial Variability and Transport of Nitrate in a Deep Alluvial Vadose Zone. Published in Vadose Zone Journal 4:41-54 (2005), Soil Science Society of America 677 S. Segoe Rd., Madison, WI 53711 USA. http://groundwater.ucdavis.edu/Publications/Onsoy%20et%20al%20-%20Spatial%20variability%20of%20nitrate%20-%20Vad%20Zone%20J%202005.pdf

Oregon State University. 2006. Spatial Climate Analysis Service. http://www.ocs.orst.edu/prism/

Palmer, M.W., 1988. Fractal geometry: a tool for describing spatial patterns of plant communities. Vegetatio 75: 91-102.

Parker, H.M. Journal, A.G. and Dixon, W.C., 1979. The Use of the Conditional Lognormal Probability Distribution for the Estimation of Open-pit Ore Reserves in Stratabound Uranium Deposits – A Case Study", Proceedings 16th APCOM, 133-148.

Pebesma, E.J. and Wesseling, C.G., 1998. Gstat: a program for geostatistical modelling, prediction and simulation. Computers & Geosciences, 24(1): 17-31.

Petrone, R.M., Price, J.S., Carey, S.K., and Waddington, J.M., 2004. Statistical characterization of the spatial variability of soil moisture in a cutover peatland. Hydrological Processes 18, 41-52. Published online 2003 in Wiley InterScience. DOI: 10.1002/hyp.1309. http://www.gret-perg.ulaval.ca/Petrone_et_al_HP18_2004.pdf

Phillips, D.L., Dolph, J. and Marks, D. 1992. A Comparison of Geostatistical Procedures for Spatial Analysis of Precipitation in Mountainous Terrain. Agric. Forest. Meteor., 58, 199-141.

Pincock. 2001. Geostatistics – What Exactly Does All this Mumbo-Jumbo Mean? Pincock Solutions newsletter, Issue No. 20 – July 2001. Lakewood, Colorado 80228. http://www.pincock.com/Perspectives/Issue20-Geostatistics.pdf

Poiker, T. and Strobl, J., 2000. UniGIS: Course 7, Spatial Operations, Class Notes. Simon Fraser University and the University of Salzburg, 2nd Edition. 99 pages.

Price, D.T., McKenney, D.W., Nalder, I.W., Hutchinson, M.F. and Kesteven, J.L. 2000. A Comparison of Two Statistical Methods for Spatial Interpolation of Canadian Monthly Mean Climate Data. Agriculter For. Meteorology. 101: 81-94 (2000).

Pyrcz, M.J., and Deutsch, C. V., 2001. Two artifacts of probability field simulation. Mathematical Geology: Volume 33 No.7: pp.775-800.

Radeloff, V.C., Miller, T.F., He, H.S. and Mladenoff, D.J., 2000. Periodicity in spatial data and geostatistical models: autocorrelation between patches. Ecography 23: 81-91. Copenhagen. http://landscape.forest.wisc.edu/pdf/Radeloff_etal2000_Ecog.pdf

Ramierz-Beltrán, N.D., Winter, A., Veneros, A., Rengifo, F.C., Escalante, N.R., 2006. Neural Networks Estimate Atmospheric Variables in the Caribbean Basin. American Meteorological Society. http://ams.confex.com/ams.confex.com/ams/pdfpapers/55135.pdfm

Rendu, J-M.M., 1979. Kriging, Logarithmic Kriging, and Conditional Expectation: Comparison of Theory with Actual Results, Proceedings 16th APCOM, 199-212.

Ricci, A.K., 1998. Geostatistics Using SAS. NorthEast SAS Users Group (NESUG) 1998 Conference. October, 4-7, Pittsburgh, PA. http://www.nesug.org/html/Proceedings/nesug98/stat/p003.pdf

Ringsby, H., B-E Saether, J. Tufto, H. Jensen, and E.J. Solberg. 2001. Asynchronous Spatiotemporal Demography of a House Sparrow Metapopulation in a Correlated Environment. Ecology: Vol. 83, No. 2, 561-569.

Ripley, B.D., 1977. Modelling spatial point patterns (with discussion). Journal of the Royal Statistical Society. B39, 1972-212.

Rivoirard J. 1994 Introduction to Disjunctive Kriging and Non-Linear Geostatistics. Claredond Press, Oxford, New York. 180 pages.

Rivoirard J., 2000. Concepts and Methods of Geostatistics. Centre de Géostatistique, Ecole des Mines de Paris. http://www.fast.u-psud.fr/~cargese/papers/chiles/chiles1.pdf

Robeson, S.M. and Janis, M.J., 1998. Comparison of temporal and unresolved spatial variability in multiyear time-average of air temperature. Climate Research, Vol. 10: 15-26. http://www.int-res.com/articles/cr/10/c010p015.pdf

Robertson, G.P., 1987. Geostatistics in Ecology: Interpolating with Known Variance. Ecology 68:744 -748.

RocNews, 2003. Geostatistics. Article prepared for RocNews Spring 2003. http://www.rocscience.com/library/rocnews/Spring2003/GeostatisticsArticle.pdf

RocScience. 2006. Interpolation Method. http://www.rocscience.com/downloads/phase2/webhelp/phase2_model/Interpolation_Method.htm

Rouhani, S., Srivastava, R., Desbarats, A, Cromer, M, and Johnson, A. editors. 1996. Geostatistics for Environmental and Geotechnical Applications, ASTM STP 1283. Special Technical Publication of the ASTM.

Rosenbaum, M., and Söderström, M., 1997. Geostatistics as an Aid to Mapping. http://www.esri.com/base/common/userconf/europroc96/PAPERS/PN11/PN11F.HTM

Rossi, R.E., Mulla, D.J., Journel, A.G. and Franz, E.H., 1992. Geostatistical Tools for Modeling and Interpreting Ecological Spatial Dependence. Ecological Monographs 62: 277-314.

Royle, A.G., 1979. A Practical Introduction To Geostatistics. Department of Mining and Mineral Sciences, University of Leeds. 103 pages.

Russo, D., 1984. Design of an Optimal Sampling Network for Estimating the Variogram. Soil Science Society of America J. 48: 708 - 716.

Saito, H., and Goovaerts, P. 2000. Geostatistical Interpolation of Positively Skewed and Censored Data in a Dioxin-Contaminated Site. Environmental Science & Technology, volume 34, number 19, 4228-4235.

Sarmiento, J.L., J. Willebrand and S. Hellerman. 1982. Objective Analysis of Tritium Observations in the Atlantic Ocean During 1971-74. Ocean Tracers Laboratory Technical Report #1.

Schotzko, D.J., and O'Keeffe, L.E., 1989. Geostatistical Description of the Spatial Distribution of *Lygus Hesperus* (*Heteroptera: Miridae*) in Lentils. Journal Economic Entomology. 82:1277-1288.

Sharov, A., 1996. Quantitative Population Ecology. http://www.gypsymoth.ento.vt.edu/~sharov/PopEcol/popecol.html

Shepard, D., 1968. A Two-Dimensional Interpolation Function for Irregularly Spaced Data. Proceedings, 23rd National Conference, Association for Computing Machinery, Brabdon Syst. Press, 517-524.

Shibli, S.A.R., 1997. Geostatistics FAQ. http://curie.ei.jrc.it/faq/index.html

Sichel, H.S., 1971. On a Family of Discrete Distributions Particularly Sited to Represent Long Tailed Frequency Data, in Laubscher, N.F., (editor). Proceedings of the third symposium on Mathematical Statistics, Pretoria, 51-97.

Sichel, H.S., 1973. Statistical Valuation of Diamondiferous Deposits. Journal of the South African Institute of Mining and Metallurgy, February, 2335-243.

Simpson, J.J., Hufford, G.L., Daly, C., Berg, J.S., and Fleming, M.D. 2005. Comparing Maps of Mean Monthly Surface Temperature and Precipitation for Alaska and Adjacent Areas of Canada Produced by Two Different Methods. Arctic. Vol. 58, No. 2, pp 137-161 http://www.ocs.orst.edu/pub/prism/docs/ComparingMapsFinalProof.pdf

Siska, P.P. and I-Kuai Hung, 2001. Propagation of Errors in Spatial Analysis. Presented in: the 24th Applied Geographic Conferences Vol. 24, Fort Worth, Texas. College of Forestry, Stephen F. Austin University, Nacogdoches, Texas, 75692-6901. http://www.faculty.sfasu.edu/ihung/pdf/research/agc24.pdf

Smith, L., R.J. Hyndman, S.N. Wood. 2004. Spline interpolation for demographic variables: the monotonicity problem. Journal of Population Research, May, 2004. http://www.findarticles.com/p/articles/mi_m0PCG/is_1_21/ai_n6155266#continue

Sobieraj, J.A., Elsenbeer, H., Cameron, G., 2003. Scale dependency in spatial patterns of saturated hydraulic conductivity. www.elevier.com/locate/catena and http://www.uni-potsdam.de/u/Geooekologie/download/elsenbeer/publikationen/RG5.pdf

Sokal, R.R. and Oden, N.L., 1978. Spatial Autocorrelation in Biology 1. Methodology 2. Some Biological Implications and Four Applications of Evolutionary and Ecological Interest. Biological Journal of the Linnean Society 10:199-228.

Solow, A.R., 1990. Geostatistical Cross-Validation: A Cautionary Note. Mathematical Geology, Vol. 22, No. 6: 637-639.

Sotter, A.P.R, Ipia, A.H.S., Pulido,J.W.C., Vaca,W.L.S., 2003. Applied Geostatistics to Studies of Environmental Contamination, Alternative Analysis Viewpoints for Air Pollutants in Bogotá, D.C. ESRI 2003 User Conference Proceedings. http://gis.esri.com/library/userconf/proc03/p0442.pdf

Srivastava, R.M., 1988. A Non-ergodic Framework for Variograms and Covariance Functions. SIMS Technical Report No. 114, Department of Applied Earth Sciences, Stanford University.

Stein, Michael, L., 1999. Interpolation of Spatial Data. Springer Series in Statistics. 256 pages.

Sterk, G. and Stein, A., 1997. Mapping Wind-Blown Mass Transport by Modeling Variability in Space and Time. Soil Science Society of American Journal. 61(1): 232-239.

Sullivan, J., 1984. Conditional recovery estimation through probability kriging – theory and practice. In: Geostatistics for natural resources characterization, Part 1. Ver, G. et al. (Eds), Reidel (Dordrecht), pp. 365-384.

Surfer, 1995. Surfer for Windows, Version 6 User's Guide – Contouring and 3D Surface Mapping. Golden Software, Inc. Golden, Colorado, 80401-1866, U.S.A.

Surfer, 1999. Surfer for Windows, Version 7 User's Guide – Contouring and 3D Surface Mapping for Scientists and Engineers. Golden Software, Inc. Golden, Colorado, 80401-1866, U.S.A. www.goldensoftware.com

Surpac Minex, 2004. Polygon Kriging. Surpac Minex Group, Perth Western Australia 6850. http://www.surpac.com/refman/default/stats/pkrig.htm

Swanson, R.T., Ritz, R. L. , McAtee, M.D. 2001. Operational Mesoscale Data Assimilation at Air Force Weather Agency Using a Parallelized Multivariate Optimal Interpolation Scheme. PSU/NCAR Mesoscale Modeling System Users' Workshop 2001. http://www.mmm.ucar.edu/mm5/workshop/ws01/swanson.doc

Syed Abdul Rahman Shibli. 1997. Spatial Correlation: General. AI-GEOSTATS. http://www.ai-geostats.org/Geostats_Faq/Syed/sc_general.html

Syed Abdul Rahman Shibli., 1997. Conditional simulation: implementation. AI-GEOSTATS. http://www.ai-geostats.org/Geostats_Faq/Syed/cs_implement.html

Syed Abdul Rahman Shibli., 1998. Geostatistics FAQ. AI-GEOSTATS. Desa Pandan 55100 Kuala Lumpur, Malaysia. http://curie.ei.jrc.it/faq/index.html

Taghvakish, S. and Amini, J., 2004. Optimum Weight in Thin-plate Spline for Digital Surface Model Generation. TS26 Positioning and Measurement Technologies and Practices II. FIG Working Week, 2004. Athens, Greece, May 22-27, 2004. http://www.fig.net/pub/athens/papers/ts26/TS26_6_Amini_Taghvakish.pdf

Tait, A., and Zheng, X. 2005. Final Report: Optimal Mapping and Interpretation of Fire Weather Information, prepared for New Zealand Fire Service. NWA Client Report: WLG2005-1, June 39, 2005. National Institute of Water and Atmospheric Research (NIWA) Project: NZF03301 / Year 2. http://www.fire.org.nz/research/reports/reports/Report_49.pdf

Tangmar, B.B., Yost, R.S., and Uehara, G., 1985. Application of Geostatistics to Spatial Studies of Soil Properties. Advanced Agronomics, 38:45-94.

Taylor, G.H., Daly, C., W.P. Gibson, and Sibul-Weisberg, J. 1997. Digital and Map Products Produces using PRISM. Proceedings 10th AMS Conference on Applied Climatology, American Meteorological Society 20-24 October 1997, Reno, Nevada pp. 217-218

Ten Berge, H.F.M., Stroosnijder, L., Burrough, P.A., Bregt, A.K., and de Haes, M.J. 1983. Spatial Variability of Soil Properties Influencing the Temperature of the Soil Surface, Agricultural Water Management, 6:213-226.

Thomas, G.S., 1997. Interactive Analysis and Modeling Semi-Variograms. Snowden Associates Pty Ltd. Australia. http://www.metal.ntua.gr/msslab/MineIT97/papers/GT67/GT67.html.

Thornton, P.E., Running, S.W., and White, M.A. 1997. Generating Surfaces of Daily Meterological Variable over Large Regions of Complex Terrain. Numerical Terradynamics Simulation Group, School of Forestry, University of Montana, Missoula. Journal of Hyrdology 190 (1997) 214-251. http://www.cgd.ucar.edu/tss/aboutus/staff/thornton/pubs/thornton_jh_1997.pdf

Thurston, J. (editor), 2003. Looking Back and Ahead: The Triangulated Irregular Network (TIN). GeoInformatics, October. http://www.vectorone.info/Publish/tin.pdf

Tobler, W., 1979: A Transformational View of Cartography. The American Cartographer, 6, 101-106.

Todini, E., and Ferrarcsi, M., 1996. Influence of Parameter Estimation Uncertainty in Kriging. Journal of Hydrology 175: 555-566.

Tolosana-Delgado, R., and Pawlowsky-Glahn. 2006. Kriging Coordinates: What Does that Mean? Departament d'Informàtica i Matemàtica Aplicada Universitat de Girona, E-17071, Girona, Spain. http://ima.udg.es/Activitats/CoDaWork03/paper_Tolosana_Pawlowsky.pdf

Tomczk, M., 1998. Spatial Interpolation and it Uncertainty Using Automated Anisotropic Inverse Distance Weighting (IDW) - Cross-Validation/Jackknife Approach. Journal of Geographic Information and Decision Analysis, Vol. 2, No. 2: 18-33. ftp://ftp.geog.uwo.ca/SIC97/Tomczak/Tomczak.html

Tsai, V.J.D., 1993. Delaunay triangulations in TIN creation: An overview and a linear-time algorithm. Int. J. Geographic Information. Science, 7~6!, 501-524.

Turcotte, D.L. 1992. Fractals and Chaos in Geology and Geophysics, Cambridge University Press, Cambridge.

University of Queensland, Mathematics. 1999. MN309 -- Advanced Numerical Analysis: 22 pages. http://www.maths.uq.oz.au/~gac/mn309/bspl.html

USGS. 2003. GTOP30 Documentation. USGS-NASA.

Vann, J., and Guibal, D., 2000. Beyond ordinary kriging – An overview of non-linear estimation. In: Mineral Resource and Ore Reserve Estimation – The AusIMM guide to good practice (Monograph 23): pp. 249-256. The Australasian Institute of Mining and Metallurgy: Parkville. [originally published as: Vann, J., and Guibal, D., 1998. Beyond ordinary kriging – An overview of non-linear estimation. In: Vann, J., (Ed.), Beyond Ordinary Kriging Seminar, October 30th, 1998 Perth, Western Australia. Geostatistical Association of Australasia Monograph 1: Perth.]. http://www.quantitativegeoscience.com/images/pdf/vann_guibal_beyond_ok.pdf

Vann, J., Guibal, D., and Harley, M., 2000**.** Multiple indicator kriging – Is it suited to my deposit? Proceedings of the 4th International Mining geology Conference, Coolum, Queensland: pp.171-179.The Australasian Institute of Mining and Metallurgy: Parkville.

Vann, J., Bertoli, O. and Jackson, S. 2002. An Overview of Geostatistical Simulation for Quantifying Risk. Paper originally published at a Geostatistical Association of Australasia symposium "Quantifying Risk and Error" March 2002. http://www.quantitativegeoscience.com/images/pdf/vann_bertoli_jackson_simulation_for_risk_distribution.pdf

Vanouplines, P., 1995. Rescaled Range Analysis and the Fractal Dimension of pi. University Library, Free University Brussels, Brussels Belgium. http://homepages.vub.ac.be/~pvouplin/pi/fracwhat.htm

Vasil'yev, L.N., and Tyuflin, A.S. 1992. Fractal Characteristics of Geostystem Spatial Structure from Space Imagery, Mapping Sciences and Remote Sensing, 29:93-102.

Verly, G., and Sullivan, J., 1985. Multigaussian and probability kriging – an application to the Jerrit Canyon deposit. Mining Engineering, June, 1985, pp. 568-574.

Villard, Marc-André and Maurer, B.A., 1996. Geostatistics as a Tool for Examining Hypothesized Declines in Migratory Songbirds. Ecology, 77(1), 1996: 59 - 68.

Vivoni, E.R., Ivanov V.Y., Bras R.L., and Entekhabi D., 2004. Generation of Triangulated Irregular Networks Based on Hydrologic Similarity. http://hydrology.mit.edu/viva/research/publications/Vivoni2004TINprint.pdf

Wackernagel, H., 1994. Cokriging versus Kriging in Regionalized Multivariate Data Analysis. Geoderma, 62 (1994): 83-92.

Wackernagel, H., 1995. Multivariate Geostatistics: An Introduction with Applications. Springer, Berlin, New York. 256 pages.

Wahba, G. 1979. How to Smooth Curves and Surfaces with Splines and Cross-validation. Proceedings 24th Conference on the Design of Experiments. US Army Research Office 79-2, Research Triangle Park, NC: 167-192.

Wahba, G. 1990. Spline Models for Observational Data. CBMS-NSF Regional Conference Series in Applied Mathematics 59, SIAM, Philadelphia, Pennsylvania.

Wahba, G. and Wendelberger, J. 1980 Some New Mathematical Methods for Variational Objective Analysis Using Splines and Cross Validation. Monthly Weather Review, 108, 1122-1143.

Walsh, S.J., Bian, L., and Brown, D.G. 1991. Issues of Spatial Dependency for Surface Representation through Remote Sensing and GIS. The Integration of Remote Sensing and Geographic Information System (J.L. Star, editor), ASPRS, Bethesda, Maryland.

Warnes, J.J., 1986. A Sensitivity Analysis for Universal Kriging. Math. Geol, Vol. 18, No. 7: 653-676.

Watson, D.F., and Philip, G.M., 1984. Systematic triangulations, Computer Vision Graphic Image Processing, 26, 217–223.

Webster, R., 1985. Quantitative Spatial Analysis of Soil in the Field. Advances in Soil Science Volume 3. Springer-Verlag, New York: 1-70.

Webster, R., Curran, P.J., and Munden, J.W. 1989. Spatial Correlation in Reflected Radiation from the Ground and Its Implication for Sampling and Mapping by Ground-Based Radiometry, Remote Sensing of Environment, 29:67-78.

Webster, R. and Oliver, M.A., 1992. Sample Adequately to Estimate Variograms of Soil Properties. Journal of Soil Science, 43, 177-192.

Webster, R. Oliver, M.A., Muir, K.R. and Mann, J.R., 1994. Kriging the Local Risk of a Rare Disease from a Register of Diagnoses. Geographic Analysis, 26(2) 168 - 185.

Weber, D. And Englund D., 1992. Evaluation and Comparison of Spatial Interpolators. Mathematical Geology, Vol. 24, No. 4: 381-391.

Wen, R., and Sinding-Larsen, R. 1997. Uncertainty in Fractal Dimension Estimated from Power Spectra and Variograms. Mathematical Geology, 29(6):727-753.

Welhan, J. 2005. Geology 606: Geostatistics, Course notes: Week 6-10 – Week 6.a Lecture Material, Experimental Variography. Idaho State University. http://giscenter.isu.edu/training/

Wiener, N., 1949. Extrapolation, Interpolation and Smoothing of Stationary Time Series: MIT Press Cambridge, Massachusetts, 158 pages.

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R. O'Donnell L., and Rowe, C.M. 1985. Statistics for the Evaluation and comparison of Models. Journal of Geophysical Research 90: 8995-9005.

Willmott, C.J., Matsuura, K. 1995. Smart Interpolation of Annually Averaged Air Temperature in the United States. Journal of Applied Meteorology, 34, 2577-2586.

Willmott, C.J., and Robeson S.M. 1995. Climatologically Aided Interpolation (CAI) of Terrestrial Air Temperature. International Journal of Climatology 15:221-229.

Willmott, C.J., Rowe, C.M. and Philpot, W.D. 1985. Small-scale Climate Maps: A sensitivity analysis of some common assumptions associated with grid point interpolation and contouring. American Cartogr. 12, 5-16.

Wingle, W.L., 1992. Examining Common Problems Associated with Various Contouring Methods, Particularly Inverse-Distance Methods, Using Shaded Relief Surfaces. Geotech '92 Conference Proceedings, Lakewood, Colorado. http://magma.mines.edu/fs_home/wwingle/pub/contour/

Wingle, W.L., and Poeter, E.P., 1998. Classes Versus Thresholds: A Modification to Traditional Indicator Simulation. American Association of Petroleum Geologists (AAPG) '98 Annual Convention, Advances in Geostatistics, Salt Lake City, Utah, May 17-20, 1998. http://www.mines.edu/fs_home/wwingle/pubs/aapg_98/class.html

Woodcock, C.E., Strahler, A.H., and Jupp, D.L.B. 1988a. The Use of Variograms in Remote Sensing: I. Scene Models and Simulated Images, Remote Sensing of Environment, 25:323-348.

Woodcock, C.E., Strahler, A.H., and Jupp, D.L.B. 1988b. The Use of Variograms in Remote Sensing: II. Real Digital Images, Remote Sensing of Environment, 25:349-379.

Wu, B.M., vanBruggen, A.H.C., Subbarao, K.V., and Pennings, G.G.H., 2000. Spatial Analysis of Lettuce Downy Mildew Using Geostatistics and Geographic Information Systems. Phytopathology, The American Phytopathological Society. Publication no. P-2000-1206-01R. http://www.apsnet.org/phyto/pdfs/2000/1206-01R.pdf

Zamkotowicz, M. 2005. Advanced Applications in GIS, Week 8, Interpolating Surfaces Power Point Presentation. Burlington County College, New Jersey. http://staff.bcc.edu/mzamkoto/advancedgis/Powerpoint/Spatial%20Analyst%20Lesson%204.ppt

Zeng, X. and Basher, R. 1995. Thin-Plate Smoothing Spline Modeling of Spatial Climate Data and Its Application to Mapping South Pacific Rainfalls. National Institute of Water and Atmospheric Research Limited, Wellington, New Zealand. American Meteorological Society. Monthly Weather Review, 3086, Volume 123. http://ams.allenpress.com/pdfserv/10.1175%2F1520-0493(1995)123%3C3086:TPSSMO%3E2.0.CO%3B2

Zimmerman, D., Pavlik, C., Ruggles, A., Armstrong, M.P. An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting. Mathematical Geology, Vol. 31. No. 4, 1999.

Zirschky, J., 1985. Geostatistics for Environmental Monitoring and Survey Design. Environmental International, Vol. 11: 515 -524.

Zrinji, Z., and Burn, D.H., 1992. Application of Kriging to Surface Water Level Estimation. Canadian Journal of Civil Engineering, 19: 181 - 185.

Zuring, Hans, 2005. F503 GIS: Methods and Applications I Notes – Part 7. University of Montana, College of Forestry and Conservation. http://www.forestry.umt.edu/academics/courses/for503/SPATIAL_Stats.htm

Xiao, C.L., Hao, J.J., and Subbarao, K.V., 1997. Spatial Patterns of Microsclerotia of *Verticillum dahliae* in Soil and Verticillium Wilt of Cauliflower. The American Phytopathological Society Vol. 87, No. 3: 325 - 331.

INDEX

E

F

G

H

I

K

T