# Kriging in the Shadows: Geostatistical Interpolation for Remote Sensing

*Richard E. Rossi,* Jennifer L. Dungan,† and Louisa R. Beck†*

*It is often useful to estimate obscured or missing remotely sensed data. Traditional interpolation methods, such as nearest-neighbor or bilinear resampling, do not take full advantage of the spatial information in the image. An alternative method, a geostatistical technique known as indicator kriging, is described and demonstrated using a Landsat Thematic Mapper image in southern Chiapas, Mexico. The image was first classified into pasture and nonpasture land cover. For each pixel that was obscured by cloud or cloud shadow, the probability that it was pasture was assigned by the algorithm. An exponential omnidirectional variogram model was used to characterize the spatial continuity of the image for use in the kriging algorithm. Assuming a cutoff probability level of 50%, the error was shown to be 17% with no obvious spatial bias but with some tendency to categorize nonpasture as pasture (overestimation). While this is a promising result, the method's practical application in other missing data problems for remotely sensed images will depend on the amount and spatial pattern of the unobscured pixels and missing pixels and the success of the spatial continuity model used.*

## INTRODUCTION

Interesting ground features are often obscured in visible and near-infrared remotely sensed images. Clouds and cloud shadows are probably the most common culprits, but fire smoke, volcanic plumes, technological problems like line dropouts, and self-shadowing by mountains and buildings also may cause "gaps" in an otherwise complete image. If the ground feature of interest is more or less persistent over time, then a researcher may be able to use an unobscured image from a later date. When cloud or haze is thin enough to allow partial recovery of signal from the ground, the cloud or haze can be filtered from the data. When these solutions are not possible, some interpolation may be necessary to complete the image. One method that has often been used for classification of missing pixels is linear interpolation between adjacent pixels (Quarmby, 1992). In this article, we describe and illustrate an alternative method for interpolating gaps in remotely sensed images. The technique, known as "kriging," has theoretical roots in multiple linear regression, but was developed mainly in the fields of mining engineering and mathematical geology to model and map mineral deposits. The method is one tool of geostatistics, a branch of applied statistics that focuses on the modeling and estimation of spatial patterns.

Although developed for mineral exploration, some geostatistical tools have been used for analysis of remotely sensed images. Much of the remote sensing literature involving geostatistics has emphasized a tool called the "semivariogram" or "variogram" and how it may be used to describe an image's spatial structure (Woodcock et al., 1988a,b; Jupp et al., 1988; 1989; Curran, 1988; Smith et al., 1989; Cohen et al., 1990; Weiler and Stow, 1991). This statistic has appealing properties and is the traditional measure of spatial dependence in geostatistics. Others recognize these properties and go a step further by suggesting how the variogram can be used with kriging algorithms for image restoration. Atkinson et al. (1990) discuss the concept of subsampling an image and storing the subsample with the image's variogram(s) for efficient storage and reconstruction. Ramstein and Raffy (1989) illustrated this concept with an image degraded by preserving only every fifth pixel. An omnidirectional variogram calculated from the degraded data was used to assign weights to the preserved pixels to interpolate between

*Isaaks & Co., Redwood City, California.

†Johnson Controls World Services, NASA—Ames Research Center, Moffett Field, California.

Address correspondence to J. L. Dungan, MS 242-4, NASA Ames Research Center, Moffett Field, CA 94035-1000.

*Received 22 September 1992; revised 4 December 1993.*



0034-4257/94/$7.00

them, and thereby restore the image. Glass et al. (1988) discuss a different application; to create a high spatial resolution image from one of low spatial resolution, they used a variogram determined from a different high spatial resolution image of the same area. This has potential for the combination of data from sensors providing different spatial resolutions, such as the Système Probatoire d'Observation de la Terre (SPOT) high resolution visible (HRV) sensor and the Landsat Thematic Mapper (TM). Haining et al. (1989) show an example of estimating missing remotely sensed data using a conditional autoregressive model with maximum likelihood estimator, which is similar to kriging.

We demonstrate the usefulness and accuracy of kriging to interpolate values for missing pixels. Our example is from an ongoing National Aeronautics and Space Administration (NASA) program known as Global Monitoring and Human Health, in which pastures are to be mapped over large regions along the coastal plain in southern Chiapas, Mexico. These land covers are of interest because they provide potential larval habitat for the mosquito *Anopheles albimanus*, a major vector of human malaria in southern Chiapas. To identify and monitor pastures successfully, temporal data are required, once in the dry season and again in the wet season. Unfortunately, during the wet season cloud cover is an almost daily phenomenon. To obtain a comprehensive and timely accounting of mosquito habitats, a dependable interpolation procedure needs to be performed for ground areas obscured by clouds or cloud shadows. A geostatistical technique known as "indicator kriging" was used to interpolate under clouds and cloud shadow in a classified Landsat TM image of this area.

## METHODS

### Study Area

The study area is located in the southern Pacific coastal plain of Chiapas, Mexico, near the city of Tapachula. The coastal plain is 20–30 km wide, and consists of a series of coalescing alluvial fans created by the deposition of eroded volcanic material from the Sierra Madre. The plain increases gradually in elevation from sea level to approximately 150 m, where the foothills of the Sierra Madre begin. Soils are well drained, although clay content and soil compaction resulting from human activity create local variations in surface drainage. Areas with poor drainage tend to be used for livestock grazing, whereas the majority of the area is used for growing crops.

In 1991 research in the study area focused on developing a land cover map showing the location of pastures and subsequently determining if all pastures produce equal numbers of *Anopheles albimanus* larvae. Landsat TM scenes were acquired over the Chiapas study area on 23 March and 11 June 1990. These dates represent late dry season and early wet season, respectively. The March image contained clouds while the June image was essentially cloud-free. A subset of each scene covering approximately 1814 km$^2$ of the coastal plain was used for analysis.

### Image Classification

For land cover mapping, the March and June data were processed separately using an unsupervised clustering routine and maximum likelihood classification. Although the March classification provided the best opportunity to distinguish pastures from other land covers (e.g, fallow annual crops), the presence of clouds obscured some of the coastal plain. The final land cover classifications included the following classes: pasture / grassland, mangrove, transitional swamp, permanent / tree crop, annual crop, banana plantation, secondary forest, and riparian. Of these land cover types, pasture is associated with the highest *An. albimanus* larval production. Therefore, the classification was transformed into a binary (or indicator) coding, where 0 denoted the absence of pasture and 1 the presence of pasture.

To test the indicator kriging algorithm so that it could later be used for processing the classified March image, the cloud and cloud shadow patterns from the raw March image were used to replace the corresponding pixels in the classified June image. This allowed the assessment of the exact accuracy of the indicator kriging technique since the "true" presence or absence of a pasture in the June classification was known for all removed pixels.

### Modeling Spatial Dependence

A fundamental theme of geostatistics is the expectation that, on average, samples close together in time and / or space are more similar than those that are farther apart. Before using any geostatistical estimation methods, this spatial autocorrelation must be inferred using spatial continuity tools such as the variogram, covariance, and correlogram. These tools are used to gauge the strength of correlation among the samples or their similarity or dissimilarity with distance. The theoretical or population definitions of these tools are based on random function theory (Journel and Huijbregts, 1978; Isaaks and Srivastava, 1989), but we present here the experimental estimates of these statistics. Let $z(\mathbf{x})$ represent the value of a variable at location $\mathbf{x}$, where $\mathbf{x}$ is the vector $(x, y)$, and let $z(\mathbf{x}+\mathbf{h})$ represent the value of the same variable at some $\mathbf{h}$ distance (or lag) and direction away. One geostatistical tool known as a variogram summarizes the spatial continuity for all possible pairings of data for all significant $\mathbf{h}$:

$$\gamma^*(\mathbf{h}) = \frac{1}{2N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})}[z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})]^2, \tag{1}$$

where $\gamma^*$ ($\mathbf{h}$) is the estimated variogram value for lag $\mathbf{h}$ and $N(\mathbf{h})$ is the number of pairs of samples separated by $\mathbf{h}$. The variogram values can be computed either as averages over all directions (an omnidirectional variogram), or specific to a particular direction (a directional variogram). If there is a trend (i.e., local means and local variances change as a function of location within the sampling space), both the underlying lag-to-lag variability and the trend variability will be included in the variogram.

Two alternative spatial variability tools in geostatistics filter mean and variance trend effects so that the underlying lag-to-lag variability can be quantified. One tool, known as the "nonergodic" or spatial covariance, $C^*(\mathbf{h})$, is estimated:

$$C^*(\mathbf{h}) = \frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})}\{[z(\mathbf{x}_i) - m_{-\mathbf{h}}][z(\mathbf{x}_i + \mathbf{h}) - m_{+\mathbf{h}}]\} \tag{2}$$

(Isaaks and Srivastava, 1988). Datum $z(\mathbf{x}_i)$ is the tail and $z(\mathbf{x}_i + \mathbf{h})$ is the head of the vector, $N(\mathbf{h})$ is the total number of data pairs separated by lag $\mathbf{h}$, and $m_{-\mathbf{h}}$ and $m_{+\mathbf{h}}$ are the mean of the points that correspond to the tail values and head values of the vectors, respectively. These head and tail means are computed:

$$m_{+\mathbf{h}} = \frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})} z(\mathbf{x}_i + \mathbf{h}),$$

$$m_{-\mathbf{h}} = \frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})} z(\mathbf{x}_i). \tag{3}$$

The formal definition of "ergodicity" is rather involved and beyond the scope of this article (cf. Olea, 1990), but the main idea is that the traditional ergodic covariance considers $m_{+\mathbf{h}} = m_{-\mathbf{h}} = m$. Differences between the head and tail means are thus accounted for in the nonergodic covariance.

The nonergodic correlogram $\rho^*(\mathbf{h})$ filters both lag means and lag variances. It is related to the nonergodic covariance and is estimated similarly:

$$\rho^*(\mathbf{h}) = \frac{1}{N(\mathbf{h})}\,\frac{\sum_{i=1}^{N(\mathbf{h})}\{[z(\mathbf{x}_i) - m_{-\mathbf{h}}][z(\mathbf{x}_i + \mathbf{h}) - m_{+\mathbf{h}}]\}}{s_{-\mathbf{h}}s_{+\mathbf{h}}}$$

$$= \frac{C^*(\mathbf{h})}{s_{-\mathbf{h}}s_{+\mathbf{h}}} \tag{4}$$

(Srivastava and Parker, 1989). $s_{-\mathbf{h}}$ and $s_{+\mathbf{h}}$ are the standard deviations of the tails and heads of the vectors, respectively, and are computed:

$$s^2_{+\mathbf{h}} = \frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})} z(\mathbf{x}_i + \mathbf{h})^2 - m^2_{+\mathbf{h}},$$

$$s^2_{-\mathbf{h}} = \frac{1}{N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})} z(\mathbf{x}_i)^2 - m^2_{-\mathbf{h}}. \tag{5}$$

Theoretically, the correlogram can vary only from $+1$ to $-1$, depending upon whether the correlation between locations is positive or negative, but in practice values $>1$ and values $<-1$ are possible.

The nonergodic covariance and correlogram values can be plotted as a function of lag distance like a variogram. However, notice that, unlike the variogram which usually contains small values at short $\mathbf{h}$ and large values at larger $\mathbf{h}$, the nonergodic statistics' values are large for small $\mathbf{h}$ and small for large $\mathbf{h}$. The variogram, covariance, and correlogram of a stationary random function model, in which the population mean and variance are constant over the sampling space ($\mu_{-\mathbf{h}} = \mu_{+\mathbf{h}} = \mu$ and $\sigma^2_{-\mathbf{h}} = \sigma^2_{+\mathbf{h}} = \sigma^2$, i.e., there is no trend) are related:

$$\gamma(\mathbf{h}) = \sigma^2 - C(\mathbf{h}),$$

$$\rho(\mathbf{h}) = C(\mathbf{h}) / \sigma^2,$$

$$1 - \rho(\mathbf{h}) = \gamma(\mathbf{h}) / \sigma^2. \tag{6}$$

The experimental variogram, covariance, and correlogram defined above, which are estimates of the corresponding "true" measures, should therefore show similar relations. The covariance and correlograms can be reexpressed in variogram form to make this comparison easier. When the spatial covariance values are subtracted from the total sample variance, the resulting plot is in variogram form. When the correlogram values are subtracted from 1, then the resulting plot is in the form of a variogram. Differences between the variogram and covariance indicate changes in the local means. Plotting $m_{-\mathbf{h}}$ and $m_{+\mathbf{h}}$ as a function of $\mathbf{h}$ provides a description of the magnitude and direction of local mean changes. Differences between the covariance and correlogram indicate varying local variances. Plots of $s_{-\mathbf{h}}$ and $s_{+\mathbf{h}}$ versus $\mathbf{h}$ show the nature of the variance changes.

In addition to furnishing an appreciation for both lag-to-lag and local trend changes, the nonergodic statistics were computed along with the variogram because they provide alternative means of interpreting spatial structure. Isaaks and Srivastava (1988) provide an example of a gold deposit that has a noisy, pure nugget variogram [a nugget variogram is, by definition, $\gamma^*(\mathbf{h}) \simeq s^2$ for all $\mathbf{h}$] but displays a well-structured nonergodic covariance. By comparing these statistics calculated on a subsample with the "true" values calculated on an exhaustive data set, they showed that the nonergodic covariance was superior to the variogram for describing spatial structure in this case. By appraising only the variogram of these data, there would have been a false conclusion of no spatial dependence. In contrast, Rossi et al. (1992) show an example of a structured variogram with near-zero nugget (i.e., the apparent intercept at $\mathbf{h} = 0$), but a corresponding flat nonergodic covariance. In this case, interpreting spatial dependence using only the variogram would have led to the false conclusion that there is a small-scale, lag-to-lag spatial dependence. Knowing that a spatial pattern occurs on a small and/or large scale and having some idea of the relative strength of these patterns can be a valuable aid in understanding the distribution of the phenomenon it represents. Moreover, the scale of the spatial pattern will influence strongly the success of geostatistical tools that provide estimates for unsampled locations.

Variograms, nonergodic covariances, and nonergodic correlograms were computed omnidirectionally as well as for the four principal directions for both the March and June images. These statistics were computed using only pixels identified as either pasture or nonpasture land, and the June data had the March cloud and cloud shadow pixels removed prior to analysis. The three statistics for the omnidirectional and direction-specific models showed similar patterns. Therefore, the omnidirectional variogram in standardized form [i.e., $\gamma^*(\mathrm{h})/s^2$, so that $s^2 = 1$] was used for interpretation and kriging.

### Estimation for Unsampled Locations: Kriging

Geostatistics offers a wide and flexible variety of tools that provide estimates for unsampled locations. Known generally as "kriging" techniques, they estimate values by taking a weighted linear average of available samples, not unlike multiple linear regression. The term "kriging" was named by Matheron (1965) in honor of Danie Krige, who first formulated and implemented this form of interpolation in 1951. Kriging can be performed on nominal as well as continuous variables and is therefore suitable for the estimation of a binary variable such as the presence and absence of a pasture.

Like traditional point interpolation methods (e.g., inverse distance weighting, triangulation, and local sample means), kriging can provide an estimate for a specific location. And like traditional areal interpolation methods (e.g., polygonal weighting and cell declustering), kriging can estimate an average value over an area. Often the traditional methods are as accurate and less time-consuming than kriging (Agterberg, 1984; Isaaks and Srivastava, 1989; Weber and Englund, 1992). However, certain characteristics of kriging distinguish it from these other methods. First, kriging can provide an estimate that is either larger or smaller than any of the sample values. The traditional techniques are restricted to the range of sample values. Second, whereas the traditional methods use Euclidean distance to weight available samples, kriging takes advantage of both distance *and* geometry (i.e., the anisotropic relations) among samples. Third, unlike traditional methods, kriging attempts to minimize the variance of the expected error. The expected error is the difference between the estimate and the true value. Of course, the true value is never actually known, so kriging applies a conceptual, probabilistic random function model of the true values (cf. Isaaks and Srivastava, 1989).

The geostatistical interpolation procedure known as ordinary kriging (or OK) is essentially identical to multiple linear regression with a couple of important differences. In its general form, a multiple linear regression estimate $\hat{Y}$ may be expressed as

$$\hat{Y} = \beta_0 + \sum_{i=1}^{n} \beta_i X_i, \tag{7}$$

where the dependent variable ($\hat{Y}$) and independent variables ($X_i$) represent different variables usually measured at the same location in space or time. Thus, the matrices used to solve the system of simultaneous equations are inferred only once from the data. In geostatistics dependent and independent variables usually represent the same property, only now measured at different locations, and we wish to estimate (i.e., interpolate) values at unrecorded locations.

For example, if $z^*(\mathbf{x}_0)$ is the value to be estimated at location $\mathbf{x}_0$, $z(\mathbf{x}_i)$ are the sampled values at their respective locations, and $\lambda_i$ are the weights to be given to each sampled value, then an OK estimate may be expressed as

$$z^*(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i z(\mathbf{x}_i). \tag{8}$$

Notice that this expression is analogous to the formula in Eq. (7).

In multiple linear regression, a least-squares variance is minimized with respect to each independent variable's coefficient or weight. Kriging has a parallel requirement that seeks to insure that over the estimation space the expected value of the estimates will equal the expected value of the true (i.e., the random function) values. In other words, we expect that our estimates will be unbiased. To be unbiased, on average, the difference between our estimates and the true, but unknown, values will be zero:

$$E[Z(\mathbf{x}) - Z^*(\mathbf{x})] = 0. \tag{9}$$

One consequence of this equation given the stationary random function model is that the sum of the weights must equal 1:

$$\sum_{i=1}^{n} \lambda_i = 1. \tag{10}$$

Just like multiple linear regression, ordinary kriging solves for the weights, $\lambda_i$, while making sure that they sum to unity and while simultaneously minimizing the quantity $[Z(\mathbf{x}) - Z^*(\mathbf{x})]$ for all estimated points. Like the minimization of least-squares variance in regression, kriging seeks to minimize the estimation variance $\sigma^2_{\mathrm{ev}}$, which is the variance of the error:

$$\sigma^2_{\mathrm{ev}} = \mathrm{Variance}[Z^*(\mathbf{x}) - Z(\mathbf{x})]. \tag{11}$$

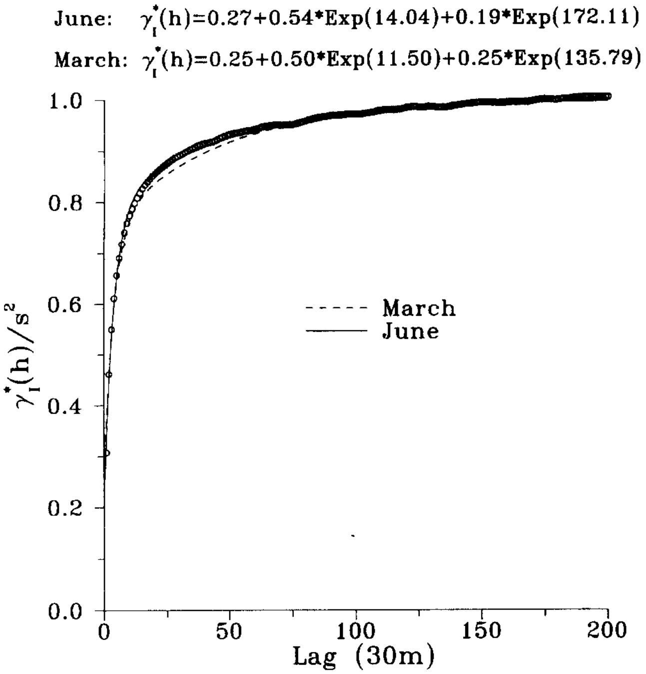


*Figure 1.* Plot of experimental variograms and fitted models. Circles represent the computed variogram values. The solid line is the fitted model for the June variogram, and the dashed line is the fitted model to the March variogram.

The geostatistical way to do all this is to infer $\sigma^2_{ev}$ from an empirical model of the existing spatial continuity or degree of spatial dependence with distance and direction. This empirical model is the variogram, covariance, or correlogram. We implicitly decide that the model is accurate over the whole estimation area. This so called "stationarity hypothesis" is perhaps the most consequential judgment in all of geostatistical estimation, for the more it is in error, the less accurate will be the kriging estimates.

Kriging uses matrix algebra to solve a set of simultaneous partial differential equations to minimize the error variance [Eq. (11)] with respect to each weight $\lambda_i$, while preventing bias [Eq. (10)]. To accomplish this task, positive definite matrices are built from the values of a model that is fit to the experimental variogram (or covariance or correlogram) points. Positive definiteness is a necessary condition in order to insure that there is a single solution to the set of simultaneous equations and that the estimation variance is positive. For convenience, the most commonly used models are those that have been demonstrated to produce positive definite matrices. A sufficiently large and flexible set of positive definite models exists so that nearly all experimental variograms can be modeled appropriately.

Kriging with binary or indicator variables is called indicator kriging or IK for short. IK is merely OK applied to indicator-coded data. As in our example, IK performed on binary data results in estimates that range from 0 to 1 corresponding to absence or presence, respectively. These IK values may be seen as a proportion or probability that a pasture exists at the given location. Journel (1983) provides a discussion of many other IK modifications and capabilities.

## RESULTS

### Variography and Kriging

Figure 1 is a plot of the variogram results and fitted models. Each fitted variogram model is composed of three structures: a nugget and two exponential models. An exponential model is a commonly used model that produces positive definite matrices in all dimensions. The exponential model's standardized equation is

$$\gamma(h) = 1 - \exp(-3h/a), \tag{12}$$

where $h$ is now the magnitude of vector $\mathbf{h}$ and $a$ is the "range." The practical range ($3a$) is the distance at which the variogram value is 95% of the sill. A shorthand notation for the fitted models is provided in Figure 1. In this shorthand, the first value is the nugget, "Exp" represents the exponential model, the coefficient before "Exp" is the sill of the model, and the value in parentheses is the practical range. Overall, there is slightly greater variability in the June image from about lag 10 to 50, but, for practical purposes, the variograms of both images are equivalent.

Figure 2 shows the IK results over the entire study area, with an inset showing a closer look at a subset of the area. Qualitatively, the results look realistic, since pasture boundaries appear to be followed underneath clouds and cloud shadows. These interpolated boundaries appear less sharp than those in the rest of the image because of the smoothing effect, a well-known side effect of kriging and most other interpolation methods (Lam, 1983; Carr and Myers, 1984).

### Assessing Kriging Accuracy

Since the truth is known from the intact June classification, the exact error, that is, the difference between the truth and the estimate, can be computed. Error that is positive represents underestimation while negative error depicts overestimation.

There are many ways to appraise the performance of kriging, but three common methods are: univariate summary statistics of the errors, a histogram of the errors, and a posting of errors. Perfect IK performance would result in zero mean, median, and all other quantiles, and zero variance. Deviations from these ideal results can be evaluated using a histogram of the errors. The overall shape of the histogram can reveal conditional bias. Conditional bias is a predisposition to either over- or underestimate given that the value is either a pasture or not.

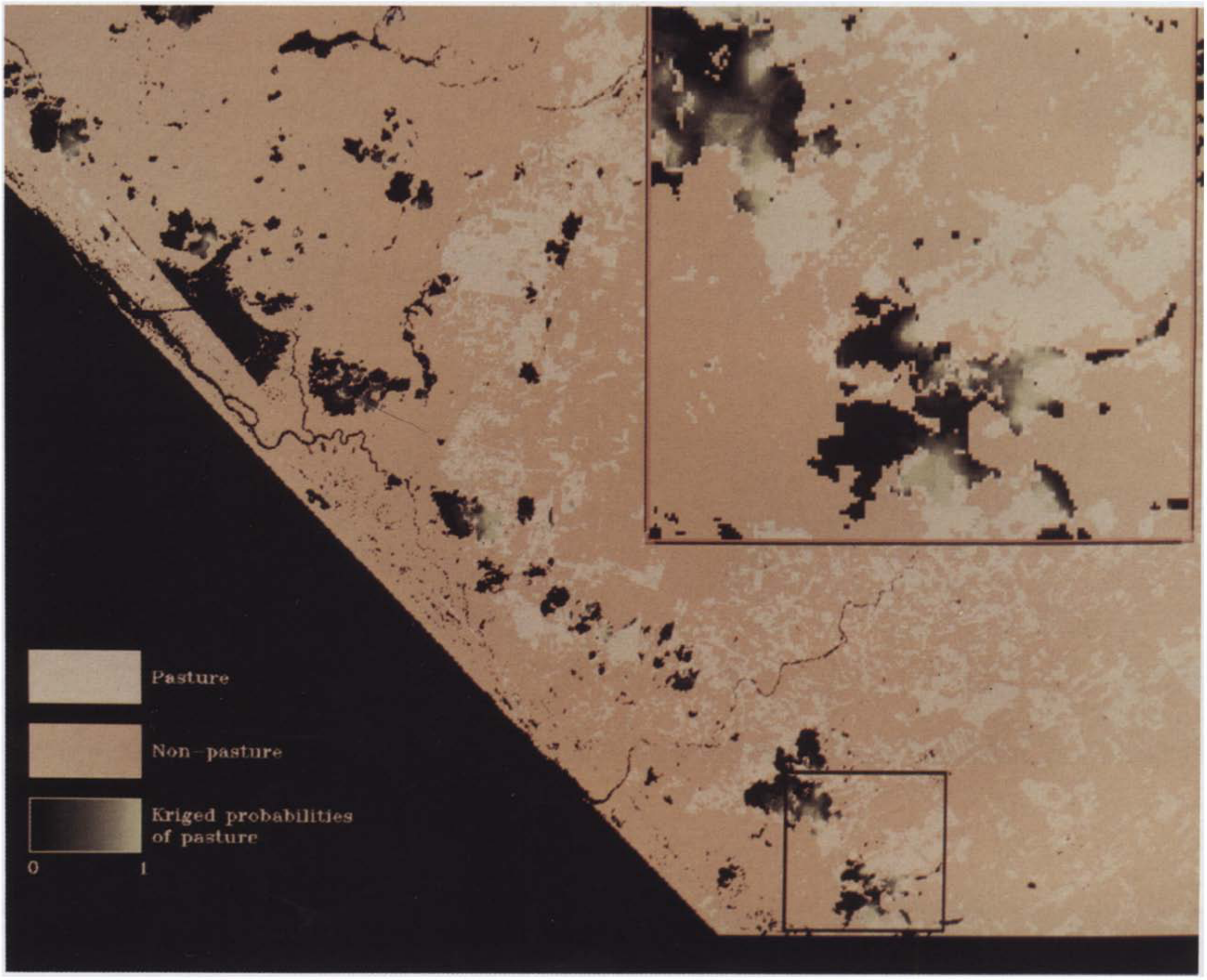


*Figure 2.* Results of kriging interpolation in clouds and cloud shadows over a portion of study area. The vertical of the image is oriented in the north-south direction, which represents approximately 27 km. Small inset box shows an area 4.5 km on a side, which is expanded to show detail in the large inset box.

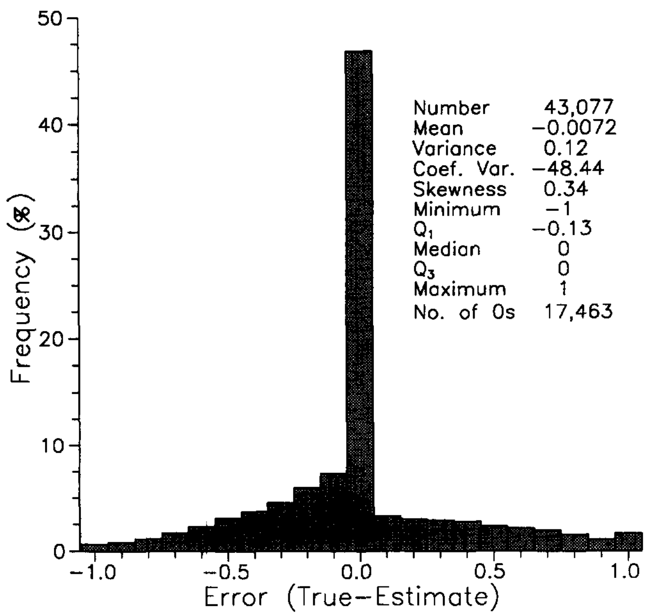


*Figure 3.* Histogram of errors with summary statistics. The true value is either the presence (= 1) or absence (= 0) of pasture and the estimate is the probability of pasture presence, [0,1].

Univariate summary statistics and a histogram of errors show the IK results to be very good (Fig. 3). Of the 43,077 kriged points, 17,463 or 40.54% were estimated perfectly. If we adopt a cutoff probability value of 0.5 for those pixels that are identified as pasture, the accuracy would be 83%. The mean of −0.0072 is nearly zero, and the variance of 0.12 is also quite small. More importantly, the first quartile value is only −0.13 while the median and the third quartile are both zero. The slight positive skew indicates that there is more overestimation than underestimation.

Conditional bias is readily apparent in the histogram of the errors. The overestimation (left side) portion of

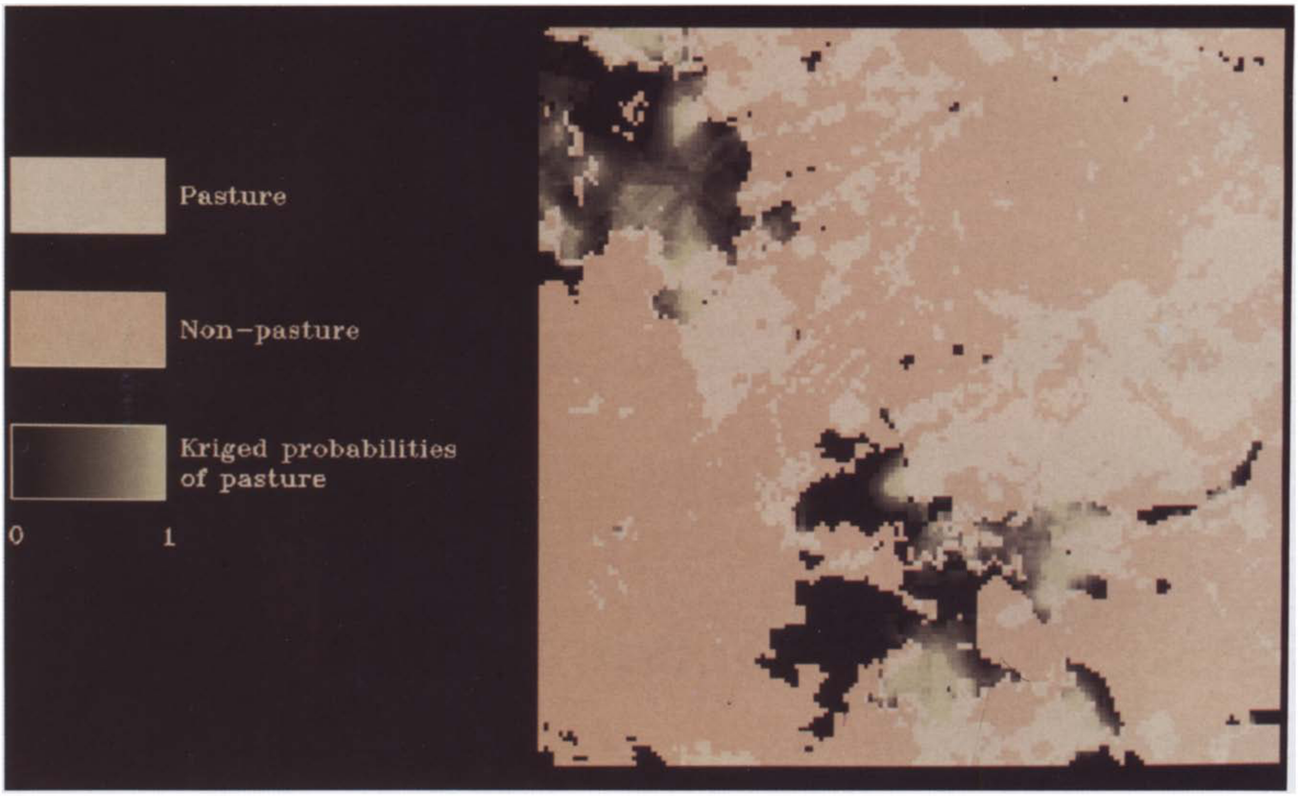


*Figure 4a.* Inset box of subarea from Figure 2.

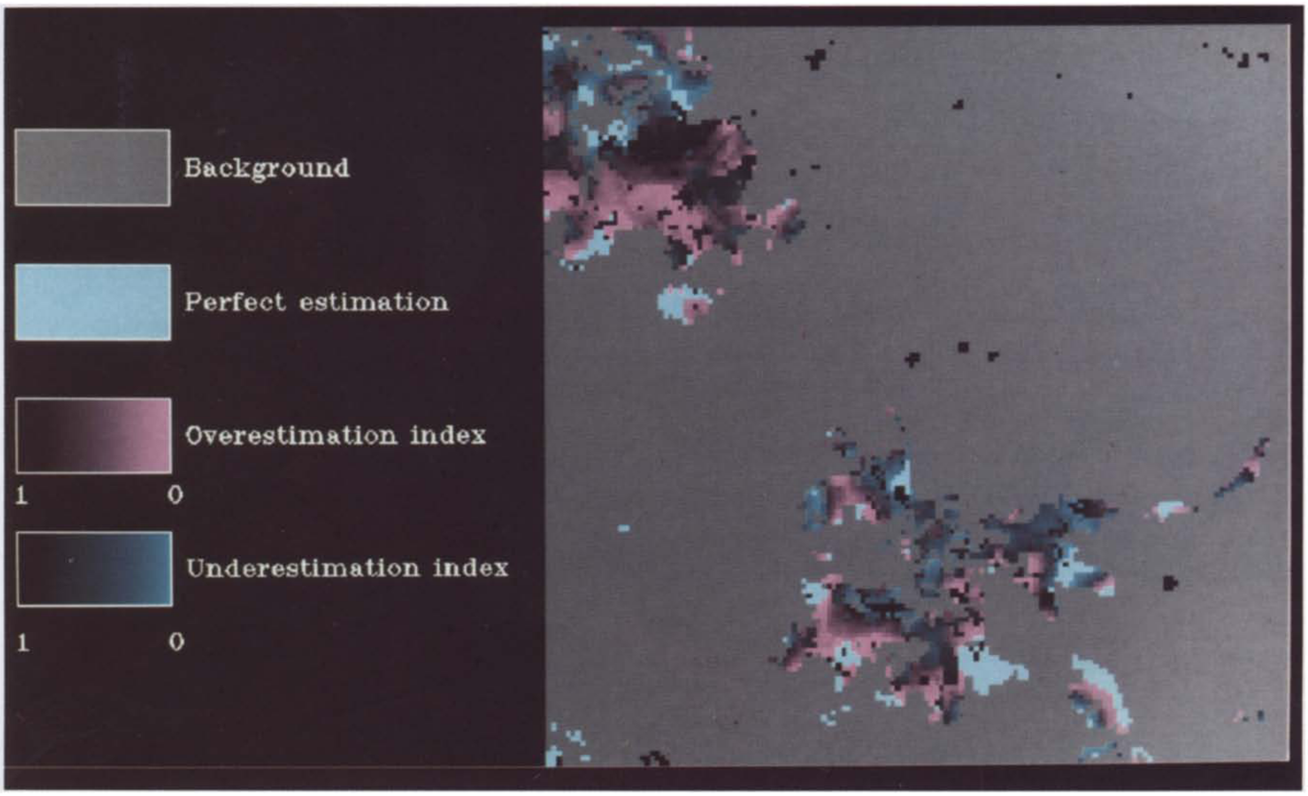


*Figure 4b.* Posting of errors in the area shown in Figure 4a.

the histogram corresponds to those locations that are not pastures while the underestimation (right side) represents locations that were pastures. Although there is a tendency to overestimate the probability of pasture, the frequency of overestimation is greatest when the actual error is smallest. This tendency is also evident in the underestimation portion of the histogram.

Summary statistics and a histogram of the errors tell us something about how well IK performed overall, but performance may be location-dependent. Plotting or posting the errors helps to assess spatially IK's effectiveness. That is, some regions of the entire estimation space may contain concentrations of systematic over- or underestimation. One expected pattern is that the error will be greatest in those areas that have the least amount of local data, that is, the centers of the larger clouds.

The kriged results and posting of errors for a subset of the data are displayed in Figures 4a and 4b. No one region of the whole estimation space contained an unusual share of either large or small errors. As expected, the largest errors occur in those regions where there is a minimum amount of information, that is, in the centers of the clouds and cloud shadow patches. Kriging's best information is the closest information, so the kriging error will be greatest ordinarily when local data are scarce.

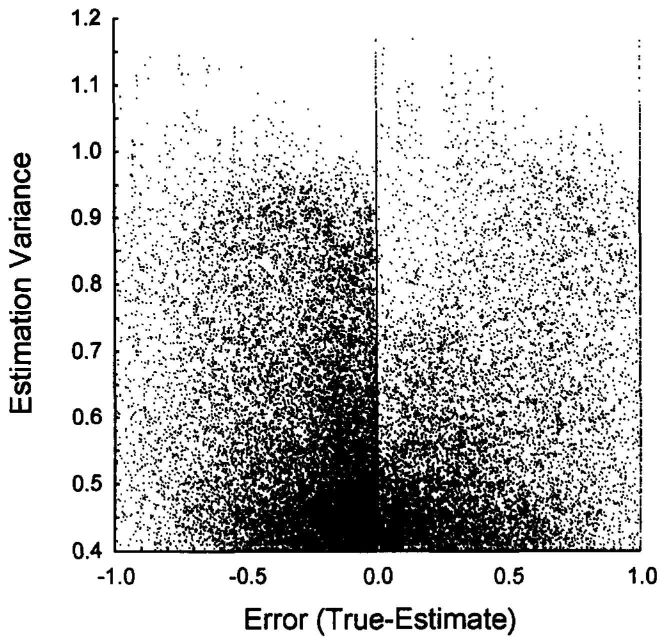


*Figure 5.* Scatterplot of estimation variance (probability$^2$) from kriging versus actual error measured by the difference between true and estimated values.

## DISCUSSION AND CONCLUSIONS

Geostatistical estimation procedures like indicator kriging and ordinary kriging can be useful in remote sensing by estimating values for missing pixels. The particular phenomenon of interest may be represented by either a nominal or a continuous variable. In addition, remote sensing researchers may gain new insights into the spatial variability inherent in their images by calculating nonergodic covariances and nonergodic correlograms along with the traditional variogram.

Currently, there are many public domain and commercial computer packages available to implement geostatistics (e.g., Deutsch and Journel, 1992; Englund and Sparks, 1988). One implementation problem encountered in this study concerned the efficient selection of neighboring samples. Typically, millions of samples are not available to the geostatistician, but they may be routine in remote sensing. Given the large number of data (3.6 million pixels), searching all data to select each kriging neighborhood was prohibitive. Therefore, a spiral search algorithm (Deutsch and Journel, 1992) was created to select only the closest data.

The IK results presented above were especially accurate due to the abundant data in the TM image. The large amount of data for pasture / nonpasture made possible the well-structured experimental variograms, covariances, and correlograms. Since these three statistics all displayed essentially identical behaviors, the decision of stationarity (i.e., the decision that the variogram was indeed representative of the presence / absence of pastures over the whole estimation space) was reasonable. Moreover, the relatively small size of clouds in relation to the size of the whole image provided many local known points for the most accurate estimation.

Although the errors demonstrated some conditional bias due to overestimation, this result is not necessarily unfavorable. Since the objective of the kriging exercise was to predict pasture presence so that a complete pool of available pastures could be identified and later sampled, overestimation represents an error of commission rather than an error of omission. Identifying an area as pasture, when in fact it is not, would not encumber the objective because once the error was detected another potential site could be chosen easily. Not identifying a pasture, when in fact it exists, might limit unrealistically the available pool of sites.

Although the estimation error variance [Eq. (11)] is minimized and can be estimated using the kriging algorithm, only when the random variable is considered multivariate normal can it be used legitimately as a measure of confidence or reliability in the resulting estimate. Journel and Rossi (1989) clearly demonstrate this point. Raw data are rarely univariate normal and multivariate normality can only be checked, not tested. The fact that kriging estimation variances are not a good local measure of confidence in the estimate is demonstrated in Figure 5, a plot of the IK estimation variances for the 43,077 kriged locations as a function of the actual error. If $\sigma^2_{ev}$ is a worthwhile measure of estimate confidence, then $\sigma^2_{ev}$ will be large when the error is large and it will be small when the error is also small. This would correspond to a "U"-shaped cloud centered on 0 error. Clearly, Figure 5 shows no such distinguishing pattern. Estimation error is, however, a function of the type (shape) of variogram model and the data configuration (i.e., areas represented by a lot of data will generally have lower estimation variance).

In summary, indicator kriging worked very well for interpolation of pixels obscured by clouds and cloud shadows in our example. Its practical application in other missing data problems for remotely sensed images will depend on the amount and spatial pattern of both the unobscured pixels and missing pixels and the success of the spatial continuity model used.

*The principal author would like to thank Walter Dettman for many illuminating late-night conversations and to Austin Nichols for his 101 reasons. The authors wish to acknowledge Dr. Mario Rodríguez of the Centro de Investigación de Paludismo for suggesting the possibility of "kriging in the shadows." This research was funded by the National Aeronautics and Space Administration's Biospheric Monitoring and Disease Prediction Program, which is a multidisciplinary effort supported by the following collaborating organizations: NASA's Ames Research Center Earth Systems Science Division; Uniformed Services University of the Health Sciences; University of California, Davis; California State University, Fresno; University of Texas, Houston and El Paso; Stanford University; and the Centro de Investigación de Paludismo, Mexico.*

## REFERENCES


Agterberg, F. P. (1984), Trend surface analysis, in *Spatial Statistics and Models* (G. L. Gaile and C. J. Willmott, Eds.), D. Reidel, Dordrecht, The Netherlands, pp. 147–171.

Atkinson, P. M., Curran, P. J., and Webster, R. (1990), Sampling remotely sensed imagery for storage, retrieval and reconstruction, *Prof. Geogr.* 37:345–353.

Carr, J. R., and Myers, D. E. (1984), Application of the theory of regionalized variables to the spatial analysis of Landsat data, in *Proceedings of the Ninth Pecora Symposium*, 2–4 October, Sioux Falls, SD, IEEE, Silver Spring, MD, pp. 55–61.

Cohen, W. B., Spies, T. A., and Bradshaw, G. A. (1990), Semivariograms of digital imagery for analysis of conifer canopy structure, *Remote Sens. Environ.* 34:167–178.

Curran, P. J. (1988), The semivariogram in remote sensing: an introduction, *Remote Sens. Environ.* 24:493–507.

Deutsch, C. V., and Journel, A. G. (1992), *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 340 pp.

Englund, E., and Sparks, A. (1988), Geo-EAS 1.2.1 User's Guide, EPA Report No. 60018-91 / 008. EPA-EMSL, Las Vegas, NV.

Glass, C. E., Carr, J. R., Yang, H. M., and Myers, D. E. (1988), Application of spatial statistics to analyzing multiple remote sensing data sets, in *Geotechnical Applications of Remote Sensing and Remote Data Transmission*, (C. B. Petterson, Ed.), American Society for Testing and Materials, Philadelphia, pp. 138–150.

Haining, R. P., Griffith, D. A., and Bennett, R. J. (1989), Maximum likelihood estimation with missing spatial data and with an application to remotely sensed data, *Commun. Stat. Theory Methods* 18:1875–1894.

Isaaks, E. H., and Srivastava, R. M. (1988), Spatial continuity measures for probabilistic and deterministic geostatistics, *Math. Geol.* 20:313–341.

Isaaks, E. H., and Srivastava, R. M. (1989), *An Introduction to Applied Geostatistics*, Oxford University Press, New York, 561 pp.

Journel, A. G. (1983), Non-parametric estimation of spatial distributions, *Math. Geol.* 15:445–468.

Journel, A. G., and Huijbregts, Ch. J. (1978), *Mining Geostatistics*, Academic, London, 600 pp.

Journel, A. G., and Rossi, M. (1989), When do we need a trend model in kriging?, *Math. Geol.* 21:715–739.

Jupp, D. L. B., Strahler, A. H., and Woodcock, C. E. (1988), Autocorrelation and regularization in digital images I. Basic theory, *IEEE Trans. Geosci. Remote Sens.* 26:463–473.

Jupp, D. L. B., Strahler, A. H., and Woodcock, C. E. (1989), Autocorrelation and regularization in digital images II. Simple image models, *IEEE Trans. Geosci. Remote Sens.* 27:247–258.

Lam, N. S.-N. (1983), Spatial interpolation methods: a review, *Am. Cartographer* 10:129–149.

Matheron, G. (1965), *Les Variables Régionalisées et Leur Estimation*, Masson, Paris.

Olea, R. A., Ed. (1990), *Geostatistical Glossary and Multilingual Dictionary*, Oxford University Press, New York.

Quarmby, N. A. (1992), Towards continental scale crop area estimation, *Int. J. Remote Sens.* 13:981–989.

Ramstein, G., and Raffy, M. (1989), Analysis of the structure of radiometric remotely-sensed images, *Int. J. Remote Sens.* 10:1049–1073.

Rossi, R. E., Mulla, D. J., Journel, A. G., and Franz, E. H. (1992), Geostatistical tools for modeling and interpreting ecological spatial dependence, *Ecol. Monogr.* 62:277–314.

Smith, R. C. G., Pratehrapar, S. A., Barrs, H. D., and Slavich, P. (1989), Use of a thermal scanner image of a water stressed crop to study soil spatial variability, *Remote Sens. Environ.* 29:111–120.

Srivastava, R. M., and Parker, H. M. (1989), Robust measures of spatial continuity, in *Geostatistics, Volume 1* (M. Armstrong, Ed.), Kluwer Academic, Dordrecht, The Netherlands, pp. 295–308.

Weber, D., and Englund, E. (1992), Evaluation and comparison of spatial interpolators, *Math. Geol.* 24:381–391.

Weiler, R. A., and Stow, D. A. (1991), Spatial analysis of land cover patterns and corresponding remotely-sensed image brightness, *Int. J. Remote Sens.* 11:2237–2257.

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 Sens. Environ.* 25:324–348.

Woodcock, C. E., Strahler, A. H., and Jupp, D. L. B. (1988b), The use of variograms in remote sensing: II. Real digital images, *Remote Sens. Environ.* 25:349–379.