

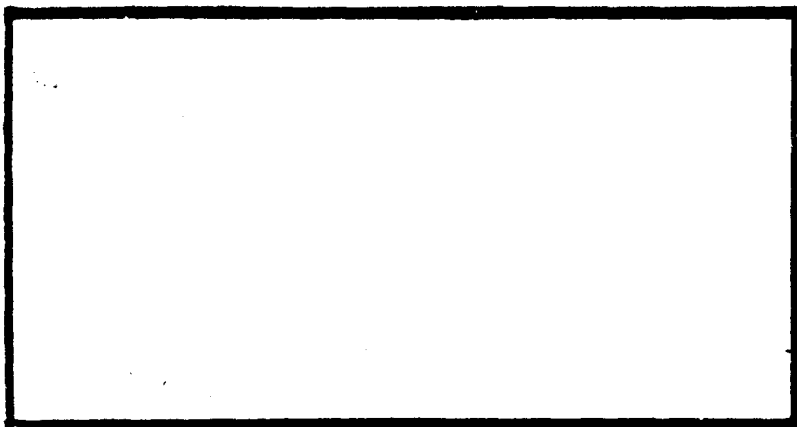
AD-A244 014



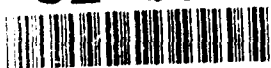
①



DTIC
ELECTE
JAN 6 1992
S B D



92-00185



DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

PARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

92 1 2 124

AFIT/GSO/ENS/ENY/91D-13

THE APPLICATION OF STATISTICAL KRIGING
TO IMPROVE
SATELLITE IMAGERY RESOLUTION

THESIS

Donald W. McGee, Jr.
Captain, USAF

AFIT/GSO/ENS/ENY/91D-13

Approved for public release; distribution unlimited

Thesis Approval

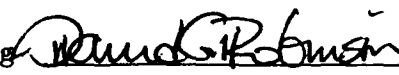
STUDENT: Capt. Donald W. McGee, Jr.

CLASS: GSO-91D

THESIS TITLE: The Application of Statistical Kriging
to Improve Satellite Imagery Resolution

DEFENSE DATE: November 22, 1991

COMMITTEE: NAME/DEPARTMENT SIGNATURE

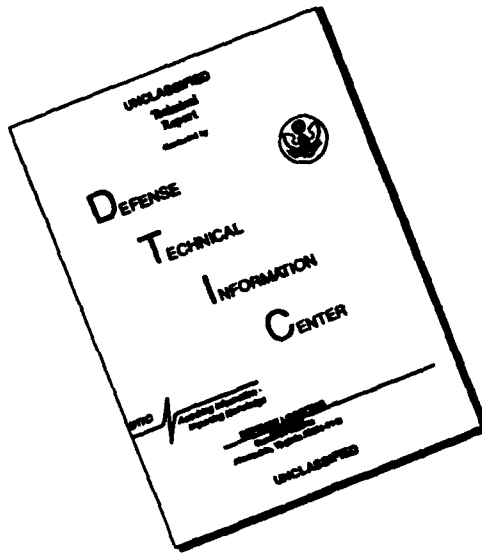
Co-Advisor David G. Robinson, Maj, USAF
Associate Professor of System Engineering 

Co-Advisor Thomas S. Kelso, Maj, USAF
Assistant Professor of Space Operations 



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

DISCLAIMER NOTICE



**THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE COPY
FURNISHED TO DTIC CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.**

AFIT/GSO/ENS/ENY/91D-13

THE APPLICATION OF STATISTICAL KRIGING
TO IMPROVE
SATELLITE IMAGERY RESOLUTION

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Space Operations

Donald W. McGee, Jr., B.S.
Captain, USAF

December, 1991

Approved for public release; distribution unlimited

Acknowledgments

Kriging digital satellite images has been an interesting and rewarding challenge. Studying kriging has greatly enhanced my knowledge and appreciation of stochastic processes and digital image enhancement. I would like to thank Dr. T. S. Kelso for suggesting the application of kriging to digital images. He personally downlinked the TIROS-N APT satellite imagery used in this thesis, and taught me everything I know about remote sensing and image processing. His guidance, support, and enthusiasm for this effort has constantly motivated me. I also owe a great debt of gratitude to Dr. David G. Robinson for his informative instruction on kriging, his advice, his understanding, and especially for his many hours of guidance and direction. He has been the cornerstone of this thesis. I would like to thank Captains Don Duckett, Chris Brodtkin, and Rob Parrott for their comraderie and for their expertise in developing kriging programs. I am grateful for the support provided by John Querns, Bill Resimont, Dennis Crnich, and John Labadan, all of the Air Force's Foreign Technology Division, for their assistance in image processing and photographic productions. The two remaining constant driving forces in my life are my wife, Margaret, and our daughter, Megan. I would like to thank them for their strength, support, and love. Without all of the above friends, this work would not have been possible. Any remaining inaccuracies that may appear in this thesis are my own.

Donald W. McGee, Jr.

Table of Contents

	Page
Acknowledgments	ii
Table of Contents	iii
List of Figures	vi
List of Tables	ix
Abstract	x
I. Introduction	1-1
1.1 Background	1-1
1.2 Research Objective	1-3
1.3 Scope	1-3
1.4 Sub-Objectives	1-3
II. Literature Review	2-1
2.1 Introduction	2-1
2.2 Digital TIROS-N Imagery	2-1
2.3 Current Enlargement Techniques	2-2
2.3.1 Nearest-Neighbor	2-2
2.3.2 Bilinear Interpolation	2-3
2.3.3 Cubic Convolution	2-3
2.4 A Statistics Review	2-3
2.5 Definition of Kriging	2-5
2.6 Kriging Theory	2-6

	Page
2.6.1 Regionalized Variables.	2-6
2.6.2 Stationarity Assumptions.	2-7
2.6.3 Structural Analysis.	2-9
2.6.4 Kriging Categories.	2-16
2.6.5 Exactness.	2-26
2.7 Summary	2-27
 III. Methodology	 3-1
3.1 Data Collection	3-1
3.2 Structural Analysis	3-4
3.2.1 Trend Analysis.	3-5
3.2.2 Variogram Determination.	3-6
3.3 Universal Kriging	3-9
3.3.1 Kriging Assumption.	3-9
3.3.2 Kriging Program.	3-10
3.3.3 Error Map.	3-14
3.4 Current Enlargement Techniques	3-14
3.5 Summary	3-14
 IV. Results	 4-1
4.1 Initial Results	4-1
4.2 Improved Structural Analysis	4-9
4.3 Numerical Analysis Comparison	4-18
4.4 Error Map	4-28
4.5 Summary	4-30
 V. Recommendations	 5-1
5.1 Improving the Kriged Sub-Images	5-1
5.2 Kriging Program Enhancements	5-1

	Page
5.3 Applicability of Kriging	5-2
5.4 Dedicated Kriging Convolution Program	5-2
5.5 Summary	5-3
Appendix A. Conversions	A-1
A.1 The Kriging Weights and the a Parameter	A-1
A.2 Estimation Error Variance	A-2
Appendix B. Original Images	B-1
Appendix C. Variograms of Residuals	C-1
Appendix D. Variogram Models	D-1
Appendix E. Typical Variogram Output	E-1
Bibliography	BIB-1
Vita	VITA-1

List of Figures

Figure	Page
2.1. Increase in $Z(\vec{x})$ with coordinate x	2-7
2.2. Second-Order (weak, wide-sense) Stationary Process	2-8
2.3. Variogram for a 4×4 of pixel values	2-12
2.4. Spherical variogram model showing sill and range	2-13
2.5. De Wijsian variogram model	2-15
2.6. Linear variogram model	2-16
3.1. Kriging Program Calculated Kernel	3-11
4.1. Image 04, Convolution Comparison	4-2
4.2. Image 07, Convolution Comparison	4-3
4.3. Image 10, Convolution Comparison	4-4
4.4. Image 13, Convolution Comparison	4-5
4.5. Image 14, Convolution Comparison	4-6
4.6. Image 19, Convolution Comparison	4-7
4.7. Image 22, Convolution Comparison	4-8
4.8. Improved Image 04, Convolution Comparison	4-11
4.9. Improved Image 07, Convolution Comparison	4-12
4.10. Improved Image 10, Convolution Comparison	4-13
4.11. Improved Image 13, Convolution Comparison	4-14
4.12. Improved Image 14, Convolution Comparison	4-15
4.13. Improved Image 19, Convolution Comparison	4-16
4.14. Improved Image 22, Convolution Comparison	4-17
4.15. Original Images, 201-by-201 pixels	4-19
4.16. Mississippi Delta, Variograms of Residuals	4-20

Figure	Page
4.17. Aerial Photograph, Variograms of Residuals	4-20
4.18. Mississippi Delta, 0° Variogram Models	4-21
4.19. Mississippi Delta, 90° Variogram Models	4-22
4.20. Aerial Photograph, 0° Variogram Models	4-23
4.21. Aerial Photograph, 90° Variogram Models	4-24
4.22. Comparison of Convolutions with Full-Sized Aerial Photograph	4-26
4.23. Comparison of Convolutions with Full-Sized Mississippi Delta	4-27
4.24. Aerial Photograph, Prediction Error Map	4-29
B.1. The Original 100-by-100 Satellite Sub-Images.	B-1
C.1. Image 04, Variograms of Residuals	C-1
C.2. Image 07, Variograms of Residuals	C-2
C.3. Image 10, Variograms of Residuals	C-2
C.4. Image 13, Variograms of Residuals	C-3
C.5. Image 14, Variograms of Residuals	C-3
C.6. Image 19, Variograms of Residuals	C-4
C.7. Image 22, Variograms of Residuals	C-4
D.1. Image 04, 0° Variogram Models	D-1
D.2. Image 04, 90° Variogram Models	D-2
D.3. Image 07, 0° Variogram Models	D-3
D.4. Image 07, 90° Variogram Models	D-4
D.5. Image 10, 0° Variogram Models	D-5
D.6. Image 10, 90° Variogram Models	D-6
D.7. Image 13, 0° Variogram Models	D-7
D.8. Image 13, 90° Variogram Models	D-8
D.9. Image 14, 0° Variogram Models	D-9
D.10. Image 14, 90° Variogram Models	D-10

Figure	Page
D.11. Image 19, 0° Variogram Models	D-11
D.12. Image 19, 90° Variogram Models	D-12
D.13. Image 22, 0° Variogram Models	D-13
D.14. Image 22, 90° Variogram Models	D-14

List of Tables

Table	Page
2.1. Water Table Elevation Data	2-23
3.1. Simple Correlations of 0° Variogram Models	3-8
3.2. Simple Correlations of 90° Variogram Models	3-9
3.3. Global Polynomial Trend Coefficients	3-10
4.1. Mississippi Delta Differenced Data Distribution	4-28
4.2. Aerial Photograph Differenced Data Distribution	4-28

Abstract

Imagery analysts are always looking for improved methods of analyzing digital satellite imagery. The resolution of satellite imagery can be improved by enlarging the images since the result will be a higher degree of discernable detail. Currently, nearest-neighbor, bilinear interpolation, and cubic convolution techniques are used for this purpose. The nearest-neighbor technique produces "block-like" images. The latter two methods produce sharp imagery, but the original information contained in the pixel values is changed in the process of convolving the image. These techniques cannot, therefore, be considered true representations of the original image.

Kriging is a statistical technique which can be applied to enlarging satellite imagery. Specifically, it is a method of best linear unbiased prediction of spatial data. One of the benefits of kriging is that it is an exact interpolator: the original pixel values will not be modified in the resulting kriged image.

This thesis develops the application of universal punctual kriging to the analysis of digital satellite imagery. Current convolution techniques and kriging are used to produce enlarged images and comparisons are made. Images are also sub-sampled and enlarged back to the original size using convolution and statistical methods. This allows the products of cubic convolution and kriging to be subtracted from the original image. This procedure provides an additional quantitative comparison of kriging and cubic convolution.

Results indicate that kriging performs as well as or better than cubic convolution when used to enlarge images. When sub-sampled data was enlarged back to the original density, the kriging methods were clearly superior to the convolution methods in recreating the original image. Kriging produced a mean difference that was 6.4 times smaller than the mean difference produced from cubic convolution. The standard deviation of the kriging difference was 4.73 times smaller than that produced from the cubic convolution technique.

THE APPLICATION OF STATISTICAL KRIGING TO IMPROVE SATELLITE IMAGERY RESOLUTION

I. Introduction

This research effort investigated the application of *kriging* to digital image enhancement. Kriging was the name given to the technique of spatial optimal linear prediction by Matheron (7:251) who chose to honor Danie G. Krige's contributions in the field of geostatistics. While kriging originated in geostatistics, its nature as an optimal predictor has made it attractive to other disciplines. For instance, according to Cressie (7:249), it has been used in physics to characterize turbulence structure; it was used in World War II to temporally predict enemy aircraft movements; it has been used in plant and animal breeding to predict genotypes based on measured characteristics; it has been used in geodesy to draw spatial maps; it has been used in the analysis of anthropometric data (15); and it has been applied here to digital image enhancement.

The imagery used in this research is the digital data reconstructed from a TIROS-N weather satellite in either an Automatic Picture Transmission (APT) or High-Resolution Picture Transmission (HRPT) resolution mode. Polar-orbiting TIROS-N satellites are designed to collect meteorological, oceanographic, and space environment data at high latitudes (1:1-1). This chapter provides the background, research objectives, scope, and methodology of the study.

1.1 Background

Enhancing photographs which are in digital form is an ongoing area of interest, especially to users of satellite imagery. Tactical battlefield commanders, for example, would find it useful to obtain and enhance satellite intelligence imagery as quickly as possible. If access to a satellite transmission was available, directly receiving and processing the images on-site would be the most expedient method. One source of these

images is the polar-orbiting TIROS-N (Television and Infrared Observation Satellite) series satellites, operated by the National Oceanographic and Atmospheric Administration (NOAA) (16:11), (1:1-1). The TIROS-N imaging function is performed by the onboard Advanced Very High Resolution Radiometer (AVHRR). The AVHRR is a cross-track multispectral scanning radiometer which continuously transmits unencrypted meteorological images to earth (1:2-1). However, it scans a 3000-km swath width and, in the highest resolution mode, still has only a relatively low resolution of 1.1 km. Both the military and the civilian communities would benefit from an enhancement technique which could produce images with better resolution than that which was recorded by the TIROS-N or any other type of satellite.

“The goal of image enhancement is to improve the visual interpretability of an image by increasing the apparent distinction between features in the scene” (21:626). A few digital image enhancement techniques currently in use are histogram equalization, various types of neighborhood averaging, and median filtering (14:146-151,161-162). Histogram equalization and neighborhood averaging techniques produce a resultant image that does not retain the original picture element (pixel) gray values. The information contained in the original image is altered. Median filtering is suitable for removing noise since it is resistant to single outlier data points, but when applied as a filter over the entire data set, it, too, suffers the disadvantage of altering the original data. These techniques, therefore, alter the presentation of the existing data and its resolution.

The current image enlargement methods are *nearest-neighbor*, *bilinear interpolation*, and *cubic convolution* (21:615). The images produced using these methods will be compared to images produced by kriging. Kriging is a statistical technique that is an exact method of spatial determination and was initially applied in the field of geostatistics to determine the grade of ore deposits (22:1259). The exactness property indicates that kriging retains all of the originally sampled pixel values: they are exactly predicted. Kriging should, therefore, be considered a faithful representation of the data in the original scene. An understanding of the exactness property is essential to the acceptance of kriging as a method of image enhancement, and is developed further in Chapter II.

Kriging predicts the gray value of all pixels which are spatially distributed on an image of gray values, whether or not the data previously existed at that point. For example, spacing the original gray values onto a finer grid will leave unknown pixel

holes that can be predicted using kriging. The original pixels are predicted exactly—with zero error. After kriging enhancement, an enlarged image would not exhibit the usually characteristic “graininess” or “block-like” appearance of an image produced by the nearest-neighbor method of replicating pixels into larger blocks. Kriging will predict better than the other spatial determination techniques partly because it uses a best linear unbiased predictor (BLUP). On the topic of spatial determination, Cressie states that “In all comparisons, on real and simulated data, universal kriging generally did as well as or better than the other methods” (6:202). *Universal kriging* was the method used in this research and is developed in Chapter II.

1.2 *Research Objective*

It is the purpose of this research to determine how the statistical technique of kriging can be applied as a method of digital image enhancement to improve satellite imagery resolution. Specifically, images enlarged by kriging will be compared to those produced by other currently accepted methods. Kriging algorithms written in the C++ computer language are used in an attempt to enhance the resolution of TIROS-N APT digital satellite imagery. One program analyzes the image and produces a *variogram* function (22:1250) to represent the amount of gray value correlation between pixel elements. The variogram function and the kriging program are then used to obtain optimum values for the pixels which were not sampled.

1.3 *Scope*

There is one clarification to make on the scope of this research:

- Kriging is applied to estimating *regionalized variables* which are spatially related in two dimensions. For this reason, only the two-dimensional version of the kriging equations will be presented when defining certain aspects of kriging. While kriging can also be applied to three dimensions, it is not done here.

1.4 *Sub-Objectives*

The following sub-objectives were accomplished to meet the research objective:

1. Produce enlarged satellite imagery to investigate the application of nearest-neighbor, bilinear interpolation, and cubic convolution to image enhancement.

2. Produce enlarged satellite imagery to investigate the application of kriging to image enhancement. Compare the results to the current enhancement techniques.
3. Analyze the resulting images and improve the kriging application, as necessary, to produce visually interpretable results.
4. Perform a quantitative comparison of kriging and cubic convolution using kriging and cubic convolution to enlarge sub-sampled imagery back to the original size.

II. Literature Review

2.1 Introduction

Remote sensing data, merged into a Geographical Information System (GIS), has many applications including soil conservation, city zoning, assessment, forecasting, and intelligence gathering (21:612). To get the most out of the information contained in the image, it may be beneficial to enhance or enlarge sections of the images prior to merging. Currently, nearest-neighbor, bilinear interpolation, and cubic convolution techniques are used to zoom or increase the resolution of digital imagery. Kriging is a statistical method of spatial determination and will be compared to the methods currently in use. A current literature review revealed no previous application of kriging to the improvement of satellite imagery resolution.

This chapter is a review of the literature pertaining to digital satellite imagery from NOAA TIROS-N weather satellites, current image enlargement techniques, the statistics applicable to kriging, and the theory of kriging.

2.2 Digital TIROS-N Imagery

Starting in 1972, radiometers were used on TIROS-N satellites as the primary source of data for the picture transmission service (1:1-2). In particular, the Advanced Very High Resolution Radiometer (AVHRR) records visible and infrared data which can be transmitted in two primary formats: Automatic Picture Transmission which is a reduced-resolution mode (APT) with 4-km resolution, and High Resolution Picture Transmission (HRPT) with 1.1-km resolution (1:2-1, 4-1). There are advantages to using AVHRR data. One advantage is its availability. AVHRR images are available two to four times each day from each satellite. Another advantage is that the data is easily accessible. TIROS-N satellite transmissions are unencrypted, and the telemetry wavetrain can, therefore, be inexpensively received.

Larcomb (20) discusses the AVHRR and the exact format of the APT and HRPT data in detail. The following paragraphs summarize this information. Visible wavelength imagery in APT format which was recorded during daylight hours will be used for this research.

The TIROS-N AVHRR records data in four or five spectral regions depending on the AVHRR type—AVHRR/1 or AVHRR/2 (1:2-2). Regardless of the type of

AVHRR in use by a particular satellite, it records visible wavelength imagery on its Channel 1 (1:2-3). In HRPT mode, all of the channels (four or five) are transmitted at full resolution. In APT mode, only two of the channels are transmitted to maintain bandwidth restrictions (1:4-1). Usually, a channel containing visible wavelength imagery and a channel containing infrared imagery will be transmitted during the daytime portion of the satellite's orbit. During nighttime, both APT channels may contain infrared imagery (1:4-1).

The APT data is transmitted as an analog signal. Upon reaching the ground receiving station, the signal is reconverted to a digital format using an analog-to-digital converter. In this format, each digital word is eight binary bits long and represents a gray value for a pixel in the image. This gray value is an integer which can range from 0 to 255. The gray values are then displayed on a computer monitor or are converted to a format that is compatible with other types of display software.

2.3 Current Enlargement Techniques

Once the pixel values are available in a two-dimensional image format, resolution enhancement techniques can be applied. The goal is to make features become noticeable that were not interpretable in the original image. Resolution is defined as the degree of discernable detail in the image, and is strongly dependent on the image dimensions and the number of gray values which can be assigned to each pixel (14:23). This research applies convolution and statistical techniques to improve the resolution of satellite imagery by increasing the image dimensions—enlarging the image. Enlargement techniques which are in common use today are nearest-neighbor, bilinear interpolation, and cubic convolution (21:615).

2.3.1 Nearest-Neighbor is by far the easiest enlargement method. At an unsampled location, the gray value is assigned to be the same as the closest neighboring pixel. It is simply a method of pixel replication. For a two times enlargement, additional rows and columns will be inserted between the existing ones. Pixel values in the inserted rows and columns will be determined by assigning them to the same value as the nearest surrounding pixel (14:249). Assignments are made at unsampled locations only, and the original image data is retained. However, the resulting "block-like" appearance makes interpretation difficult.

2.3.2 *Bilinear Interpolation* is a distance-weighted spatial prediction technique which determines the value at an unsampled point by considering a linear combination of the four nearest pixels. The procedure is to use an equation of the form:

$$Z_p^* = ax + by + cxy + d \quad (2.1)$$

where Z_p^* is the predicted pixel value, and x and y are the coordinates of the known pixels (14:250). This equation produces four simultaneous equations corresponding to the four known pixel gray values. Solving for the four coefficients, inserting them back into the predictor, and using the coordinates of the prediction point as the x and y value will produce the interpolated pixel value. Interpolation is performed at all locations, altering the original pixel gray values as well as filling-in the unsampled locations.

2.3.3 *Cubic Convolution* performs a distance-weighted prediction similar to bilinear interpolation, except that the prediction is based on a cubic combination of the surrounding 16 pixels. There are various forms of the cubic function that can be used. For example, Gonzales and Wintz suggest that cubic convolution "fits a surface of the $(\sin x)/x$ type" through the 16 nearest neighbors to produce the gray value prediction (14:249). A cubic spline convolution is another possible method. Also, in the field of computer graphics, various parametric cubic curves are used (13:482-488). The derivation of cubic convolution in this case requires identifying geometric constraints, incorporating them into a geometry vector, choosing a type of curve (e.g. Hermitian), and producing a blending function (13:482-485). However the cubic convolution is established, it will produce a slightly sharper image than bilinear interpolation. However, it too, alters the original pixel gray values.

2.4 *A Statistics Review*

A review of the statistics involved in kriging will be presented before proceeding.

This research considers the possible outcomes for Z , the pixel gray value, to be a random variable. The expected value of Z , the mean, and is defined as:

$$E[Z] = \mu \quad (2.2)$$

The variance is the expected squared deviation of Z about its mean:

$$\text{VAR}[Z] = \sigma^2 = E[(Z - \mu)^2] = E[Z^2] - \mu^2 \quad (2.3)$$

The covariance, represented by σ_{XY} , where X and Y are random variables is:

$$\begin{aligned} \sigma_{XY} &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - 2E[X]E[Y] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y] \end{aligned} \quad (2.4)$$

The following quote explains the covariance relationship between two random variables, Y_1 and Y_2 :

The larger the absolute value of the covariance of Y_1 and Y_2 , the greater the linear dependence between Y_1 and Y_2 . Positive values indicate that Y_1 increases as Y_2 increases; negative values indicate that Y_1 decreases as Y_2 increases. A zero value of the covariance would indicate no linear dependence between Y_1 and Y_2 (23:237).

Two distinct random variables can reference the same attribute measured at different locations (17:8), for example, a pixel value at location \vec{x} and $\vec{x} + \vec{h}$. The covariance between two such locations can then be written as:

$$\sigma_{Z_1 Z_2} = \sigma(Z[\vec{x}], Z[\vec{x} + \vec{h}]) \quad (2.5)$$

and if $Z(\vec{x})$ is a stationary random function:

$$E[Z(\vec{x})] = \mu = E[Z(\vec{x} + \vec{h})] \quad (2.6)$$

then:

$$\sigma_{Z_1 Z_2} = E[Z(\vec{x})Z(\vec{x} + \vec{h})] - \mu^2 \quad (2.7)$$

At a distance of $\|\vec{h}\| = 0$ between the two points, $Z(\vec{x})$ and $Z(\vec{x} + \vec{h})$ are the same, and the covariance simply becomes the variance. From the definition of the covariance

(and using a capital letter C as is sometimes used to represent the covariance):

$$C(0) = C(Z, Z) = \sigma_{ZZ} = E[(Z - E[Z])(Z - E[Z])] = E[(Z - \mu)^2] \quad (2.8)$$

which is the definition of the variance by Equation 2.3.

$$C(0) = \text{VAR}[Z(\vec{x})] \quad (2.9)$$

2.5 Definition of Kriging

Kriging is a method of interpolation which uses a "best linear unbiased predictor" to optimally predict an attribute of a variable by minimizing the variance associated with its prediction (8:237). The first step is to perform a structural analysis of the spatial data. This will determine the second-order correlation structure of the data and will determine the kind of kriging which should be employed. Then, minimizing the estimation variance will be shown to produce a series of equations which will be solved simultaneously for the weighting factors. These equations are referred to as the kriging equations. Their general form, in matrix notation, is:

$$\begin{bmatrix} \gamma(h_{11}) & \gamma(h_{12}) & \cdots & \gamma(h_{1n}) & 1 \\ \gamma(h_{21}) & \gamma(h_{22}) & \cdots & \gamma(h_{2n}) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(h_{n1}) & \gamma(h_{n2}) & \cdots & \gamma(h_{nn}) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma(h_{1p}) \\ \gamma(h_{2p}) \\ \vdots \\ \gamma(h_{np}) \\ 1 \end{bmatrix} \quad (2.10)$$

where $\gamma(h_{ij})$ is related to the covariance between spatially-distributed, known samples. w_j are the weights associated with those samples, λ is a Lagrange multiplier, and $\gamma(h_{ip})$ is related to the covariance between known samples and the unknown point being predicted. γ is actually the *variogram* value which is discussed below, and the weights are calculated by matrix operations. Finally, the weights are inserted into the linear predictor. This is how the kriging predictor is used to calculate the attribute value of a regionalized variable based on the amount of correlation that exists between known samples in the area.

2.6 Kriging Theory

The following paragraphs will present the theory of kriging as it was developed by Matheron and others in the field of geostatistics. The reviewed topics include regionalized variables, stationarity assumptions, variogram estimation, structural analysis, kriging methodology, types of kriging, and exactness.

2.6.1 Regionalized Variables. A regionalized variable is a random function which varies continuously between points, but cannot be represented exactly by a deterministic function (9:239). Often in statistics, observations of a random variable are assumed to be independent and identically distributed. This indicates that the observations are uncorrelated to one another. However, the correlation between data points is the foundation of a statistical analysis of spatial data (6:197). Therefore, Matheron speaks of using a regionalized variable to “stress the spatial aspect of the phenomena” (22:1248). He goes on to say that a regionalized variable is a function that is localized, continuous, and may display *anisotropies*—be more correlated in one direction than another. To model a regionalized variable at a location, \vec{x} , assume that $E[Z(\vec{x})]$ and $VAR[Z(\vec{x})]$ exist. Then, the variable is usually modeled as an expression of the mean and a residual at each location (2:263).

$$Z(\vec{x}) = \mu(\vec{x}) + e(\vec{x}) \quad (2.11)$$

where the mean function is:

$$E[Z(\vec{x})] = \mu(\vec{x}) \quad (2.12)$$

For example, Figure 2.1 shows an increasing mean, $\mu(\vec{x})$, in the measured spatial attribute $Z(\vec{x})$ with coordinate \vec{x} (18).

This is a global trend over the entire sample area and can be approximated by fitting a linear or a polynomial function to the data. For example, a step function could be fitted to the data to create neighborhoods where the localized trend, or drift, is, on the average, constant. Removing the global trend from the data will leave the residual component, which is a zero-mean stochastic process:

$$E[e(\vec{x})] = 0 \quad (2.13)$$

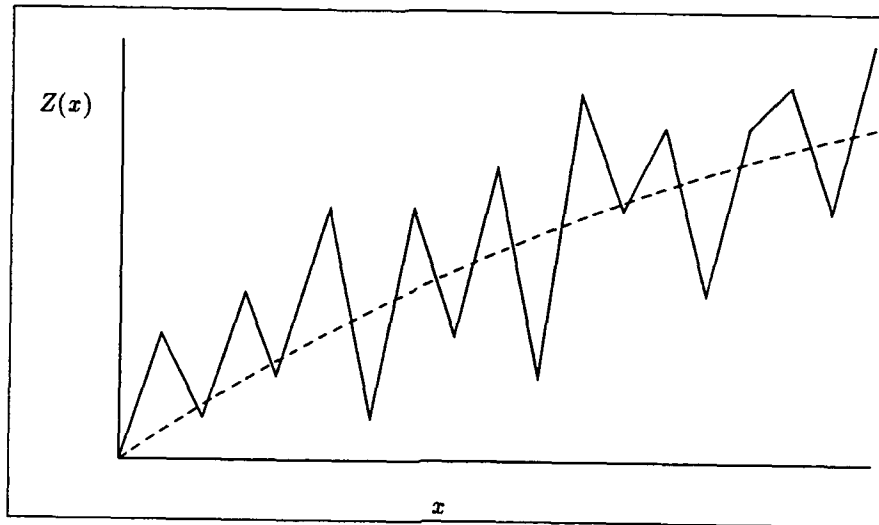


Figure 2.1. Increase in $Z(\vec{x})$ with coordinate x (Adapted from Journal (18:716))

and can be modeled by a regionalized variable. The type of kriging model employed will depend on the structure of the mean function and the stationarity assumption as defined below.

2.6.2 Stationarity Assumptions. Stationarity is a model decision, and is not a property of the $Z(\vec{x})$ values themselves (17:8). The type of kriging model employed will depend on the stationarity assumption.

2.6.2.1 Strict Stationarity means that the entire cumulative distribution function for the regionalized variable must be the same at all points. This is impractical for kriging since an infinite number of samples would be required (24).

2.6.2.2 Weak, Wide-Sense, or Second-Order Stationarity is the next least restrictive case, where the mean and covariance, $\mu(\vec{x})$ and $\sigma(\vec{x})$, of the regionalized variable are known to exist, and are both constant over the area (8:92). This is seldom found in nature, but is easily shown for the covariance as in Figure 2.2. The origin is arbitrarily placed at the center of the surface being considered.

The two samples under consideration, Z_1 and Z_2 , are the same attribute of the surface at locations \vec{x}_1 and \vec{x}_2 , respectively. Under second-order stationarity,

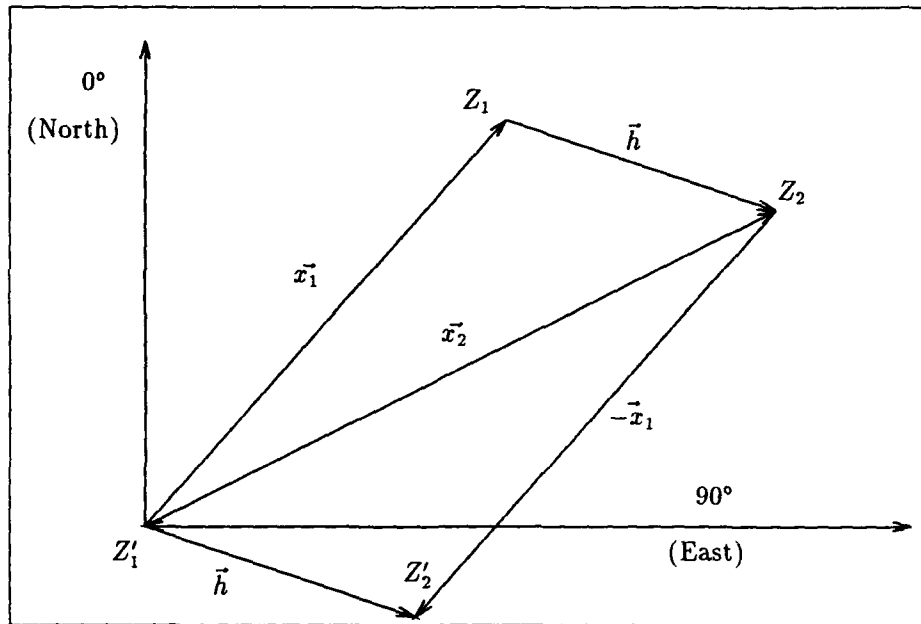


Figure 2.2. Second-Order (weak, wide-sense) Stationary Process

the covariance of the two attributes is independent of their location and dependent only on their relative distance apart. Therefore, moving both points by \vec{x}_1 to the prime coordinates will not change the value of the covariance. Subtracting \vec{x}_1 and substituting $(\vec{x}_1 + \vec{h})$ for \vec{x}_2 gives the following relationship:

$$C(Z_1, Z_2) = \sigma(\vec{x}_1, \vec{x}_2) \quad (2.14)$$

$$= \sigma(\vec{x}_1 - \vec{x}_1, \vec{x}_2 - \vec{x}_1) = C(Z'_1, Z'_2) = \sigma'(\vec{h}) \quad (2.15)$$

$$= \sigma(0, \vec{x}_1 + \vec{h} - \vec{x}_1) = \sigma(\vec{h}) \quad (2.16)$$

This shows that if second-order stationarity is assumed, the covariance between any two points, $\sigma(\vec{h})$, depends only on the vector distance between them. It does not depend on the position of the samples or of the sample values themselves. The variance is equal to the covariance at zero: $\sigma(0)$ (8:93).

2.6.2.3 Intrinsic Stationarity indicates that the relationship between any two points depends only on the distance between them, and not on the direction between samples, or on the position of the samples, or on the values at the locations.

This is called the stationarity of increments and is a further relaxation of the stationarity constraint. Under intrinsic stationarity, the covariance may not even exist. A random function is intrinsically stationary if:

1. $E\{Z(x)\} = \mu$. That is, the mean is known and constant in a particular region.
2. For all vectors in the area, the difference, $[Z(\vec{x} + \vec{h}) - Z(\vec{x})]$, has finite variance that does not depend on \vec{x} (24).

To restate the second condition, the covariance may be infinite, but the variogram must exist. The relationship between the variogram and the covariance will be derived in Equation 2.24.

2.6.3 Structural Analysis. Structural analysis involves determining the covariance, or variogram, and the degree of trend, or drift, in the regionalized variable (11:327). Analysis must also be performed to compensate for any *anisotropic* behavior of the sampled data (8:134).

2.6.3.1 Variograms. The kriging equations require knowledge of the correlation between the known data samples in order to determine the optimum weights for the predictor. The variogram is a function which fulfills this requirement by representing the correlation of the sampled data in a particular direction at any given distance h apart. However, the data being sampled must at least be assumed to be intrinsically stationary. That is, it must have a constant mean, at least locally. Therefore, the residual samples are used instead of the original data, since it has already been shown that the residual component is a zero-mean stochastic process. Next, the residual distribution can be represented by defining the mean and variance. The mean is known to be zero, and the variogram is defined to represent the variance. The variogram refers to the discrete representation of the true variogram since a finite number of known samples are used in the calculation.

Definition. Like the covariance, the variogram is a measure of similarity between two values, and it is logical that a similarity index between values at two points be based on a difference between them. Also, it is desirable that the magnitude of this difference always be positive (8:74). Therefore, the variogram is defined as the mean square of the difference between two samples, and is represented by $2\gamma(\vec{h})$.

Kerbs calls the variogram the “variance of the differences” (19:5).

$$2\gamma(\vec{h}) = \text{VAR}[Z(\vec{x}) - Z(\vec{x} + \vec{h})] \quad (2.17)$$

$$= E[Z(\vec{x}) - Z(\vec{x} + \vec{h})]^2 \quad (2.18)$$

$$\approx \frac{1}{N(\vec{h})} \sum_{i=1}^{N(\vec{h})} [Z(\vec{x}_i) - Z(\vec{x}_i + \vec{h})]^2 \quad (2.19)$$

$N(\vec{h})$ is the number of pairs of points being compared. The intrinsic assumption that the increment, h , is stationary led to the computation of the variogram (8:113).

Semivariogram. $\gamma(\vec{h})$ is half of the variogram, and is named the semivariogram. However, it has been suggested (8:94) (and it will be the convention in this thesis) that for the sake of simplicity, $\gamma(\vec{h})$ be called the variogram.

Example. The following example (adapted from Kerbs (19:11-14)) will show how a variogram is calculated on a four-sample-square area. Assume the samples are pixel gray values in an image.

91	28	43	72
55	86	32	41
72	71	59	81
38	19	51	44

Since this data is already on a regular grid, variogram calculations can be started immediately. Let 0° be defined as the north direction, and 90° be defined as the east direction. Choose the direction to be examined as 90° . Next, $\gamma(h)$, as defined in Equation 2.19, can be calculated for $\gamma(0)$ through $\gamma(3)$. Arbitrarily choose the top left corner as the place to begin variogram calculations. Assume that the samples are two meters apart, $h = 2$ meters. This is equivalent to a hypothetical image resolution of two meters. It should be apparent from Equation 2.19, that $\gamma(0) = 0$, since $Z(\vec{x}_i + \vec{h})$ reduces to $Z(\vec{x}_i)$ for all i . The others are calculated as follows:

$$\begin{aligned}
\gamma(2 \text{ m}) &= [(91 - 28)^2 + (28 - 43)^2 + (43 - 72)^2 + (55 - 86)^2 + \\
&\quad + (86 - 32)^2 + (32 - 41)^2 + (72 - 71)^2 + (71 - 59)^2 + \\
&\quad + (59 - 81)^2 + (38 - 19)^2 + (19 - 51)^2 + (51 - 44)^2] \div [(2)(12)] \\
&= 460.67 \text{ (gray values)}^2
\end{aligned}$$

$$\begin{aligned}
\gamma(4 \text{ m}) &= [(91 - 43)^2 + (28 - 72)^2 + (55 - 32)^2 + (86 - 41)^2 + \\
&\quad + (72 - 59)^2 + (71 - 81)^2 + (38 - 51)^2 + (19 - 44)^2] \div [(2)(8)] \\
&= 491.06 \text{ (gray values)}^2
\end{aligned}$$

$$\begin{aligned}
\gamma(6 \text{ m}) &= [(91 - 72)^2 + (55 - 41)^2 + (72 - 81)^2 + (38 - 44)^2] \div [(2)(4)] \\
&= 84.25 \text{ (gray values)}^2
\end{aligned}$$

Since, by definition, all variogram values are positive, and all lags are positive, a single quadrant of a cartesian coordinate system is all that is necessary to plot the variogram. The vertical axis coordinate is the variogram, and the horizontal axis is plotted as the lag, h . The variogram values at lags of zero through three are then plotted and are displayed in Figure 2.3.

Covariance Relationship. The relationship between the variogram and the covariance function can now be derived as follows (2:266), (6:298):

$$\gamma(\vec{h}) = \frac{1}{2} \text{VAR}[Z(\vec{x}) - Z(\vec{x} + \vec{h})] \quad (2.20)$$

and from basic statistics:

$$\gamma(\vec{h}) = \frac{1}{2} [\text{VAR}(Z(\vec{x})) + \text{VAR}(Z(\vec{x} + \vec{h})) - 2\sigma(Z(\vec{x}), Z(\vec{x} + \vec{h}))] \quad (2.21)$$

Assuming at least second-order stationarity:

$$\sigma(0) = \text{VAR}[Z(\vec{x})] = \text{VAR}[Z(\vec{x} + \vec{h})] \quad (2.22)$$

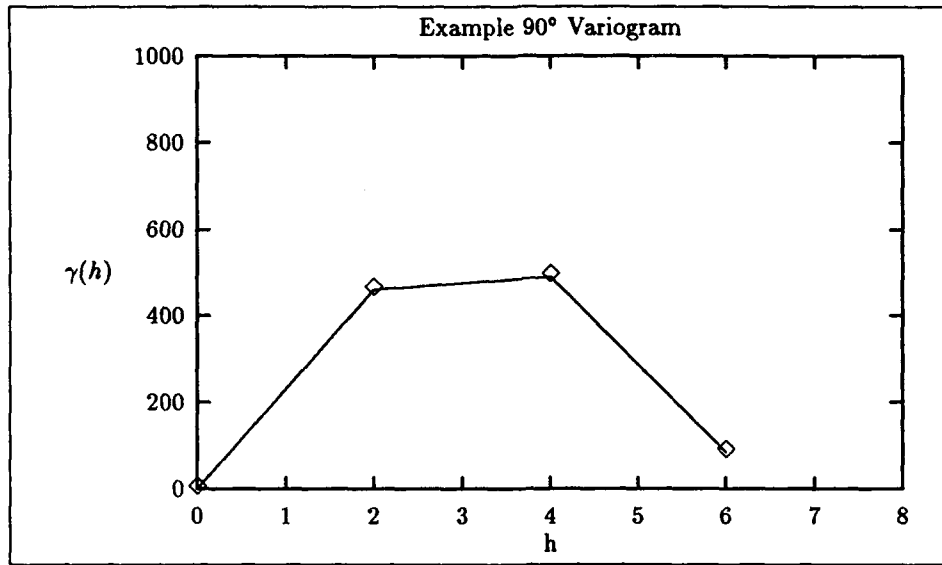


Figure 2.3. Variogram for a 4×4 of pixel values

and substituting $\sigma(0)$ gives:

$$\gamma(\vec{h}) = \frac{1}{2}[\sigma(0) + \sigma(0) - 2\sigma(\vec{h})] \quad (2.23)$$

$$\gamma(\vec{h}) = \sigma(0) - \sigma(\vec{h}) \quad (2.24)$$

By this relationship, the variogram will react oppositely from the covariance. As the correlation increases, the variogram value decreases. Since closely spaced data should be more alike than data that are spaced far apart, the variogram will indicate a small value for short separations h , and large values at larger separations. This relationship also shows how the variogram can exist in cases where the covariance may not.

Sill and Range. The *sill* and the *range* are parameters of some forms of the variogram. Their locations are shown on a spherical variogram curve in Figure 2.4 (the spherical variogram model and its parameters, C and C_0 , are defined later).

For small values of h , closely positioned points have a high correlation, and the variogram value is small. At greater distances between sampled points, the correlation is less, and the variogram value becomes large until the variogram is

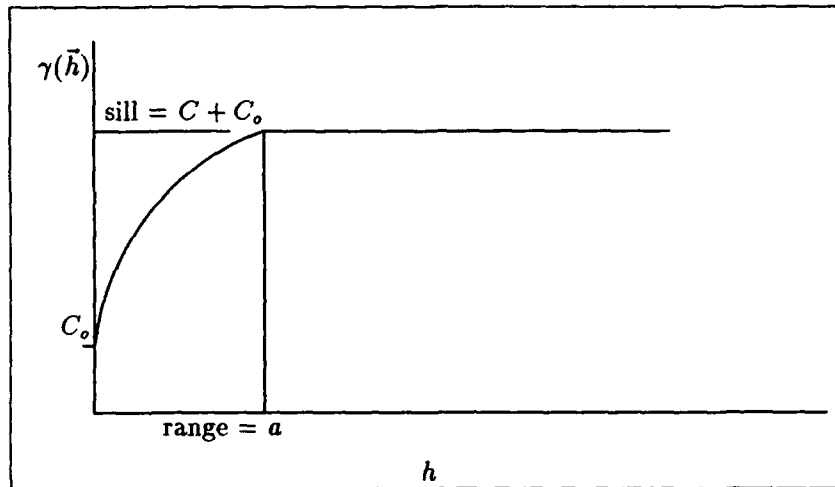


Figure 2.4. Spherical variogram model showing sill and range

equal to the variance around the mean of the data set. If the sill exists, it is the ordinary variance of the samples (4:42) (5:408). That is, as:

$$\|\vec{h}\| \rightarrow \infty, \gamma(\vec{h}) \rightarrow \sigma(0) = \text{VAR}[Z(\vec{x})] \quad (2.25)$$

The value of h at this point represents the range at which the data remain correlated (19:54). The range, a , defines a neighborhood of related points. At $h \geq a$, the samples become independent of one another (4:6).

Nugget Effect. Measurement error and microscale variation comprise the nugget effect (5:410-411). Ideally, resampling a point should give the same value as it did in the original sample. However, this is not always the case. Variations in the sampling technique or the sample itself may cause subsequent samples to be different. This is known as measurement error. If the samples are pixel values on a single image, there is only one value at each location, and measurement error is zero. The other source of nugget effect is microscale variations. These are high frequency changes in the data relative to the lag. If either measurement error or microscale variation exist, the resulting variogram curve will not intersect the origin as $\|\vec{h}\| \rightarrow 0$. (However, $\gamma(0) \equiv 0$ since a point is, by definition, not different from itself.) The result is that the variogram curve is shifted upward. The sill, and therefore the variance between the samples, will be increased. It is significant to note, however, that the

covariance remains unaffected (8:99). The nugget effect is shown in Figure 2.4 and is represented by C_o .

Variogram Modeling. In the kriging system of equations, the variogram values between various locations are required. The spherical model, the De Wijsian model, the linear model, the exponential model, the hole-effect model, and linear combinations of these are typical continuous functions assumed and may be appropriate depending on the shape of the variogram (8:102-112).

Spherical. The shape of the spherical model is shown in Figure 2.4. The three parameters necessary to define the model are a , C , and C_o . The C parameter is used with the nugget effect, C_o , to define the sill, $C + C_o$. The range of influence is defined as a , and the actual model is as follows:

$$\gamma(h) = \begin{cases} C\left(\frac{3h}{2a} - \frac{1}{2}\frac{h^3}{a^3}\right) + C_o & \text{for } h \leq a \\ C + C_o & \text{for } h > a \\ 0 & \text{for } h = 0 \end{cases}$$

De Wijsian. The De Wijsian model can be used if there is no range of dependence evident in the discrete representation of the variogram (8:106). The general shape of this model is a natural log curve as shown in Figure 2.5. The model is as follows:

$$\gamma(h) = A \ln(h) + B \quad (2.26)$$

The De Wijsian model is named after Prof. H. J. de Wijs and can be used if the variogram is linear when plotted against $\ln(h)$. The model shown was fitted to actual variogram data and, as such, does not exactly coincide with the data and intersect the origin. The consequence of this type of modeling behavior will be discussed in Chapter IV.

Linear. The linear model is simple and can be desirable in cases where computational efficiency is required. It is represented as:

$$\gamma(h) = Ah + B \quad (2.27)$$

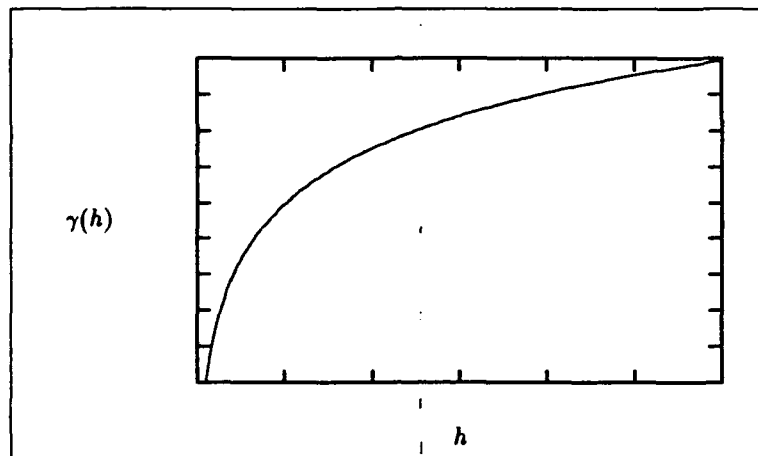


Figure 2.5. De Wijsian variogram model

Within a limited neighborhood, the linear model can be a good representation of the discrete variogram (8:108). The shape of the linear model is shown in Figure 2.6.

2.6.3.2 Drift. A regionalized variable can be considered to be composed of a *drift* and a *residual* (9:243). If there is no trend in the data, the drift will not exist. However, if there is a trend, the data and the variogram will exhibit a drift. That is, the variance of the data will continue to increase as the distance between samples, h , is increased (19:55). As a result, no sill can be determined, and kriging will produce a biased result if it is performed in the presence of a strong global trend (drift) (4:120). The solution to the problem is to model the drift, possibly as a linear or quadratic function, subtract it from the original variable, and perform a structural analysis on the residuals. However, a localized drift is allowable if considered during kriging and is explained in detail below.

2.6.3.3 Anisotropy. To represent the sampled data, variograms are calculated over the area in different directions. Regionalized variables are said to be anisotropic if they are not the same in all the directions or vary from location to location over the surface. In geostatistics, anisotropy explains runs of a mineral in a certain direction (22:1249). There are two types of anisotropies: geometric and zonal (8:134). Geometric anisotropy exists when the variation in one direction at a distance h away is the same as the variation in another direction at a distance of

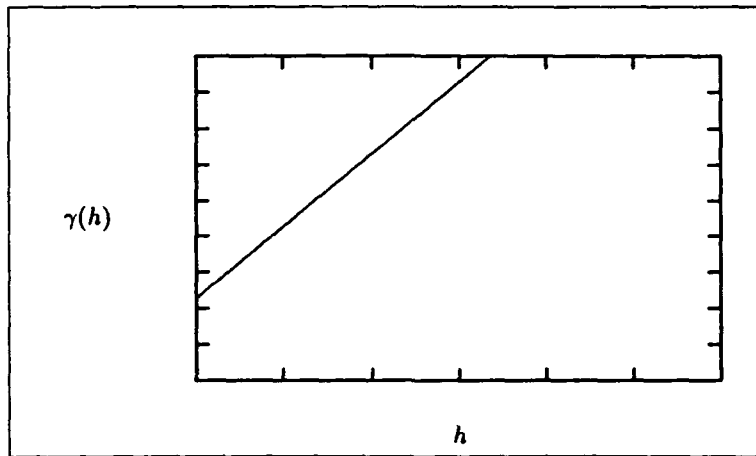


Figure 2.6. Linear variogram model

some multiple of h . They have different ranges a , but the same sill. This is easily corrected by scaling the second distance by a factor of k to make the ranges equal. Then, one variogram can be used to represent the entire data set correlation in either direction. Isotropy is simulated by substituting kh for h in the variogram model being used. For the spherical model:

$$\gamma_1(h) = C \left(\frac{3h}{2a} - \frac{h^3}{2a^3} \right) + C_0 \quad (2.28)$$

$$\gamma_2(h) = C \left(\frac{3kh}{2a} - \frac{k^3 h^3}{2a^3} \right) + C_0 \quad (2.29)$$

Zonal anisotropy is exhibited when variograms are different at different locations (24). Therefore, a single variogram will not be able to represent the entire sampled data. Zonal anisotropy is not as easily corrected as geometric anisotropy. However, the problem can be approached by partitioning the data into isotropic regions.

2.6.4 Kriging Categories. There are two broad categories of kriging. Block kriging is concerned with both volume samples and point samples. It is used extensively in geostatistics. Punctual kriging is concerned with predicting the values of a

stationary regionalized variable based on point samples. This is the simplest form of kriging (9:383). Also, there are different types of kriging that are used depending on the assumption of stationarity. This section will develop the three main types of kriging which are used, based on the stationarity assumption. These three types are simple, ordinary, and universal kriging.

2.6.4.1 Simple Kriging. The following derivation is adapted from Journel (17:10-11). First, a linear predictor, $Z^*(\vec{x})$, is defined to predict the unknown value $Z(\vec{x})$. The prediction at an unsampled location, p , is established to be a linear combination of the n known sample values Z_i :

$$Z_p^* = w_o + w_1 Z_1 + w_2 Z_2 + \dots + w_n Z_n \quad (2.30)$$

$$= \sum_{i=1}^n w_i Z_i + w_o \quad (2.31)$$

w_o is a shift parameter (constant). The goal in kriging is to calculate the weights w_i by ensuring that the best linear unbiased predictor is used. To be unbiased, the average prediction error should be zero:

$$E[Z_p - Z_p^*] = 0 \quad (2.32)$$

and from properties of the expected value:

$$E[Z_p] - E[Z_p^*] = 0 \quad (2.33)$$

since $E[Z_i] = \mu_i$, simple substitutions for Z_p and Z_p^* gives:

$$\mu_p - E\left[\sum_{i=1}^n w_i Z_i\right] - E[w_o] = 0 \quad (2.34)$$

and $E[w_o] = w_o$. Then,

$$w_o = \mu_p - \sum_{i=1}^n w_i \mu_i \quad (2.35)$$

Substituting this result into the predictor gives:

$$Z_p^* = \mu_p + \sum_{i=1}^n w_i(Z_i - \mu_i) \quad (2.36)$$

$$Z_p^* - \mu_p = \sum_{i=1}^n w_i(Z_i - \mu_i) \quad (2.37)$$

but, if Z_i is a regionalized variable, then the quantity $(Z_i - \mu_i)$ is just the residual from Equation 2.11.

$$e_p^* = \sum_{i=1}^n w_i e_i \quad (2.38)$$

In other words, a predictor of the residual term is automatically unbiased if the mean is constant. The predictor has been shown to be linear and unbiased, but to show that e_p^* is the "best" predictor, the predictions must be shown to produce a minimum estimation error variance. The estimation error at the unknown point is defined to be ϵ_p :

$$\epsilon_p = Z_p - Z_p^* \quad (2.39)$$

and its variance is:

$$VAR[\epsilon_p] = \sigma_\epsilon^2 = VAR[Z_p - Z_p^*] \quad (2.40)$$

Substituting for Z_p , and Z_p^* :

$$VAR[\epsilon_p] = \sum_{i=0}^n \sum_{j=0}^n a_i a_j \sigma_{ij} \quad (2.41)$$

where i and j are locations of known sample points, σ_{ij} is the covariance of the increment between them, and a_i is defined in terms of the w_i weights (see Appendix A for details).

Having defined the criteria for the BLUP, the necessary values for w_j are found by minimizing the error variance in Equation 2.41. Taking the partial derivatives of the error variance and setting them equal to zero gives:

$$\frac{\partial}{\partial a_i} \left(\sum_{i=0}^n \sum_{j=0}^n a_i a_j \sigma_{ij} \right) = 0 \quad (2.42)$$

and, since there are only 1 to n samples, the unsampled point p is defined to be the $i = 0, j = 0$ subscripts. The result of the derivative is:

$$\begin{aligned} \sum_{j=0}^n a_j \sigma_{ij} &= 0 \\ a_0 \sigma_{i0} + \sum_{j=1}^n a_j \sigma_{ij} &= 0 \end{aligned} \quad (2.43)$$

The weights w_j from Appendix A are substituted for the a_j coefficients, and the p subscript replaces the 0 subscript to maintain consistent notation. The p subscript indicates the unsampled point value being predicted: $a_0 = a_p = 1$, and $a_j = -w_j$ for $j = 1, \dots, n$. The result is:

$$\begin{aligned} \sigma_{ip} - \sum_{j=1}^n w_j \sigma_{ij} &= 0 \quad \forall j = 1, \dots, n \\ \sum_{j=1}^n w_j \sigma_{ij} &= \sigma_{ip} \quad \forall j = 1, \dots, n \end{aligned} \quad (2.44)$$

Equation 2.44 represents n equations which will be solved simultaneously. An additional equation must be added to ensure that the weights sum to one (as a result of the unbiased condition). Since there are now $n+1$ equations and n unknown weights, a Lagrange multiplier, λ , will be added to the system. This procedure will be shown in more detail under ordinary kriging. Note since σ_{ij} and σ_{ip} are the covariance of the increment between points, they can more appropriately be written as $\sigma(h_{ij})$ and $\sigma(h_{ip})$, respectively. In matrix notation Equation 2.44 becomes:

$$\begin{bmatrix} \sigma(h_{11}) & \sigma(h_{12}) & \cdots & \sigma(h_{1n}) & 1 \\ \sigma(h_{21}) & \sigma(h_{22}) & \cdots & \sigma(h_{2n}) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma(h_{n1}) & \sigma(h_{n2}) & \cdots & \sigma(h_{nn}) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \lambda \end{bmatrix} = \begin{bmatrix} \sigma(h_{1p}) \\ \sigma(h_{2p}) \\ \vdots \\ \sigma(h_{np}) \\ 1 \end{bmatrix} \quad (2.45)$$

These are the kriging equations and are analogous to Equation 2.10 through the variogram and covariance relationship of Equation 2.24. The predicted value of

the regionalized variable Z_p^* can then be calculated by substituting the weights into the residual predictor and adding the constant mean value.

Kriging also has the advantage of calculating the estimation error variance, or kriging variance, of its predictions. This is a measure of how well each unsampled point was predicted. The error variance, σ_e^2 , is:

$$\sigma_e^2 = \sigma(0) - [w]^T [b] \quad (2.46)$$

where $[w]$ is the $[w_j]$ weights matrix and

$$b = \begin{bmatrix} \sigma(h_{1p}) \\ \sigma(h_{2p}) \\ \vdots \\ \sigma(h_{np}) \\ 1 \end{bmatrix} \quad (2.47)$$

The error variance is different if the variogram form of the kriging equations is used.

$$\sigma_e^2 = [w]^T [b] \quad (2.48)$$

where

$$b = \begin{bmatrix} \gamma(h_{1p}) \\ \gamma(h_{2p}) \\ \vdots \\ \gamma(h_{np}) \\ 1 \end{bmatrix} \quad (2.49)$$

The relationship between the two σ_e^2 calculations is in Appendix A. The error variance calculation can then be immediately performed once the weights are known. Calculating the error variance at each predicted point will enable one to calculate an error map of the predictions.

2.6.4.2 Ordinary Kriging. It is unlikely that the mean of a set of samples will be known, and that it will be constant as was assumed under simple kriging.

If the mean was known, that would imply an infinite number of samples, and there would be no predictions to be made. Ordinary kriging allows the mean to be unknown but still requires it to be constant. In this case, it is not sufficient to define the best predictor as that which has the minimum estimation error variance; it should also be free of systematic error (8:238). To remain unbiased, an additional condition should be added to the kriging equations such that the predictor is still linear, and the mean is still constant, albeit unknown, over the area:

$$Z^* = \sum_{i=1}^n w_i Z_i \quad (2.50)$$

$$E[Z^*] = \mu \quad (2.51)$$

A simple substitution for Z^* gives:

$$\begin{aligned} E \left[\sum_{i=1}^n w_i Z_i \right] &= \mu \\ \sum_{i=1}^n w_i E[Z_i] &= \mu \\ \sum_{i=1}^n w_i \mu &= \mu \\ \sum_{i=1}^n w_i &= 1 \end{aligned} \quad (2.52)$$

The problem, now, is to minimize the error variance subject to the constraint that the kriging weights must sum to one. The solution utilizes the method of Lagrange to incorporate the constraint into the objective function (17:15) as follows:

$$F = \sigma_e^2 + \lambda \left[\sum_{i=1}^n w_i - 1 \right] \quad (2.53)$$

where σ_e^2 is the estimation error variance, and λ is the Lagrange multiplier. Another partial derivative is then taken with respect to the Lagrange multiplier, and is set

equal to zero. The resulting kriging equations are shown below.

$$\begin{array}{rcccccccc}
 0 & + & w_1 & + & w_2 & + & w_3 & + \dots + & w_n & = & 1 \\
 \lambda & + & w_1\gamma(h_{11}) & + & w_2\gamma(h_{12}) & + & w_3\gamma(h_{13}) & + \dots + & w_n\gamma(h_{1n}) & = & \gamma(h_{1p}) \\
 \lambda & + & w_1\gamma(h_{21}) & + & w_2\gamma(h_{22}) & + & w_3\gamma(h_{23}) & + \dots + & w_n\gamma(h_{2n}) & = & \gamma(h_{2p}) \\
 \lambda & + & w_1\gamma(h_{31}) & + & w_2\gamma(h_{32}) & + & w_3\gamma(h_{33}) & + \dots + & w_n\gamma(h_{3n}) & = & \gamma(h_{3p}) \\
 \vdots & + & \vdots & + & \vdots & + & \vdots & + \dots + & \vdots & = & \vdots \\
 \lambda & + & w_1\gamma(h_{n1}) & + & w_2\gamma(h_{n2}) & + & w_3\gamma(h_{n3}) & + \dots + & w_n\gamma(h_{nn}) & = & \gamma(h_{np})
 \end{array}$$

Again, the covariance can also be used in lieu of the variogram, $\gamma(\vec{h})$, and the result is the following system of equations:

$$\begin{array}{rcccccccc}
 0 & + & w_1 & + & w_2 & + & w_3 & + \dots + & w_n & = & 1 \\
 \lambda & + & w_1\sigma(h_{11}) & + & w_2\sigma(h_{12}) & + & w_3\sigma(h_{13}) & + \dots + & w_n\sigma(h_{1n}) & = & \sigma(h_{1p}) \\
 \lambda & + & w_1\sigma(h_{21}) & + & w_2\sigma(h_{22}) & + & w_3\sigma(h_{23}) & + \dots + & w_n\sigma(h_{2n}) & = & \sigma(h_{2p}) \\
 \lambda & + & w_1\sigma(h_{31}) & + & w_2\sigma(h_{32}) & + & w_3\sigma(h_{33}) & + \dots + & w_n\sigma(h_{3n}) & = & \sigma(h_{3p}) \\
 \vdots & + & \vdots & + & \vdots & + & \vdots & + \dots + & \vdots & = & \vdots \\
 \lambda & + & w_1\sigma(h_{n1}) & + & w_2\sigma(h_{n2}) & + & w_3\sigma(h_{n3}) & + \dots + & w_n\sigma(h_{nn}) & = & \sigma(h_{np})
 \end{array}$$

In either case, the BLUP calculation remains the same. Furthermore, a confidence interval can be established around the prediction. If the estimation error is assumed to be normally distributed about the true (but unknown) value at the unsampled point, σ_e can be used to establish the confidence interval.

2.6.4.3 Ordinary Kriging Example. Davis shows an example of ordinary kriging by considering water table elevations based on three well samples (9:386-392). The water table depth at a new well, p , at a known coordinate is to be predicted. Table 2.1 shows the well locations. The variogram for this example is determined (given) to be linear, have no nugget effect, and have a slope of $4.0 \text{ m}^2/\text{km}$. The distances between wells, h_{ij} , are calculated from the coordinates. The variogram slope then maps the distances into $\gamma(h_{ij})$ by simply multiplying, $\gamma(h_{ij}) = 4.0h$. In a similar manner, the new well coordinate is used to calculate h_{ip}

Location	X Coordinate	Y Coordinate	Water Table Elevation
Well 1	3.0	4.0	120.0
Well 2	6.3	3.4	103.0
Well 3	2.0	1.3	142.0
Point p	3.0	3.0	

Table 2.1. Water Table Elevation Data

and then $\gamma(h_{ip})$. The kriging equations are written as:

$$\begin{bmatrix} 0.0 & 1 & 1 & 1 \\ 1 & \gamma(h_{11}) & \gamma(h_{12}) & \gamma(h_{13}) \\ 1 & \gamma(h_{21}) & \gamma(h_{22}) & \gamma(h_{23}) \\ 1 & \gamma(h_{31}) & \gamma(h_{32}) & \gamma(h_{33}) \end{bmatrix} \cdot \begin{bmatrix} \lambda \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ \gamma(h_{1p}) \\ \gamma(h_{2p}) \\ \gamma(h_{3p}) \end{bmatrix}$$

and substituting the calculations for $\gamma(h_{ij})$ and $\gamma(h_{ip})$:

$$\begin{bmatrix} 0.0 & 1 & 1 & 1 \\ 1 & 0.0 & 13.4 & 11.5 \\ 1 & 13.4 & 0.0 & 19.1 \\ 1 & 11.5 & 19.1 & 0.0 \end{bmatrix} \cdot \begin{bmatrix} \lambda \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4.0 \\ 13.3 \\ 7.9 \end{bmatrix}$$

Notice that the diagonal values are all zero since $\gamma(h = 0) \equiv 0$, and the matrix is symmetric since $\gamma(h_{ij}) = \gamma(h_{ji})$. The zero diagonal can cause a problem during matrix inversion and is one reason why the covariance could be used instead of the variogram. Solving the system for the weights gives:

$$\begin{bmatrix} \lambda \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -0.7267 \\ 0.6039 \\ 0.0868 \\ 0.3093 \end{bmatrix}$$

Solution to the prediction is calculated by substituting the weights into the predictor:

$$\begin{aligned} Z_p^* &= \sum_{i=1}^3 w_i Z_i \\ &= 0.6039(120.0) + 0.0868(103.0) + 0.3093(142.0) \\ &= 125.3 \text{ meters} \end{aligned}$$

and the error variance is:

$$\begin{aligned} \sigma_e^2 &= [w]^T [b] \\ &= [w_j]^T [\gamma(h_{ip})] \\ &= -0.7267 + 0.6039(4.0) + 0.0868(13.3) + 0.3093(7.9) \\ &= 5.3 \text{ m}^2 \end{aligned}$$

The confidence interval of the prediction can now be calculated.

$$Z_p^* \pm 1.96\sigma_e \quad (2.54)$$

is the prediction with 95 percent confidence. This result implicitly assumes a Gaussian model (22:1259). The confidence interval for the water table example under ordinary kriging can be calculated as follows:

$$\begin{aligned} \sigma_e^2 &= 5.3 \text{ m}^2 \\ \sigma_e &= 2.3 \text{ m} \end{aligned}$$

and

$$Z_p^* = 125.3 \pm 4.5 \text{ m} \quad (2.55)$$

with 95 percent confidence.

2.6.4.4 Universal Kriging. If there is a trend, or a drift, in the sampled data, the regionalized variable will not be stationary. As a consequence, the predictor will no longer be unbiased (9:393). In this case, universal kriging must be used and is described by Davis (9:393) as consisting of three operations:

“First, the drift must be estimated and removed. Then, the stationary residuals are kriged to obtain the needed estimates. Finally, the estimated residuals are combined with the drift to obtain estimates of the actual surface.” (9:393)

The trend can further be characterized as having both a local and global component. Performing a least-squares fit to the data and removing the polynomial will produce stationary residuals. The variogram of the residuals can then be calculated. However, universal kriging only assumes the data is intrinsically stationary. That is, it only requires the data in the local neighborhood, as defined by the range a , to be stationary. Then, only the neighborhood sample points will be used in the predictor. Local stationarity is accounted for by incorporating a local, first-order drift polynomial into the kriging equations. The first-order polynomial terms are added to the kriging matrix, and the following system of equations is solved to determine the weights:

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 0 & X_1 & X_2 & \cdots & X_n \\ 0 & 0 & 0 & Y_1 & Y_2 & \cdots & Y_n \\ 1 & X_1 & Y_1 & \gamma(h_{11}) & \gamma(h_{12}) & \cdots & \gamma(h_{1n}) \\ 1 & X_2 & Y_2 & \gamma(h_{21}) & \gamma(h_{22}) & \cdots & \gamma(h_{2n}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_n & Y_n & \gamma(h_{n1}) & \gamma(h_{n2}) & \cdots & \gamma(h_{nn}) \end{bmatrix} \cdot \begin{bmatrix} \lambda \\ \alpha_1 \\ \alpha_2 \\ w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} 1 \\ X_p \\ Y_p \\ \gamma(h_{1p}) \\ \gamma(h_{2p}) \\ \vdots \\ \gamma(h_{np}) \end{bmatrix}$$

where X_i and Y_i are the coordinates of the i th sample point, X_p and Y_p are the coordinates of the unsampled point, and the α s are unknown local drift coefficients which are to be found. Once again, the alternate covariance form of the kriging

equations can be written as:

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 0 & X_1 & X_2 & \cdots & X_n \\ 0 & 0 & 0 & Y_1 & Y_2 & \cdots & Y_n \\ 1 & X_1 & Y_1 & \sigma(h_{11}) & \sigma(h_{12}) & \cdots & \sigma(h_{1n}) \\ 1 & X_2 & Y_2 & \sigma(h_{21}) & \sigma(h_{22}) & \cdots & \sigma(h_{2n}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_n & Y_n & \sigma(h_{n1}) & \sigma(h_{n2}) & \cdots & \sigma(h_{nn}) \end{bmatrix} \cdot \begin{bmatrix} \lambda \\ \alpha_1 \\ \alpha_2 \\ w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} 1 \\ X_p \\ Y_p \\ \sigma(h_{1p}) \\ \sigma(h_{2p}) \\ \vdots \\ \sigma(h_{np}) \end{bmatrix}$$

and the error variance is calculated as in Equation 2.48 or Equation 2.46. The prediction and confidence interval are subsequently calculated as in the previous example.

2.6.5 Exactness. The exactness property of kriging has been described as its ability to predict a known value with zero error (9:392). For instance, in the water table prediction example, if the elevation for Well Two was to be predicted, $\gamma(h_{ip})$ would be equal to $\gamma(h_{i2})$ since Well Two and the prediction point coincide. Using ordinary kriging for simplicity, the resulting equations are as follows:

$$\begin{bmatrix} 0.0 & 1 & 1 & 1 \\ 1 & 0.0 & 13.4 & 11.5 \\ 1 & 13.4 & 0.0 & 19.1 \\ 1 & 11.5 & 19.1 & 0.0 \end{bmatrix} \cdot \begin{bmatrix} \lambda \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 13.4 \\ 0.0 \\ 19.1 \end{bmatrix}$$

In $[A][w] = [b]$ notation, row three of $[A]$ is identical to $[b]$. The solution, as would be expected, is:

$$\begin{bmatrix} \lambda \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

and the prediction is:

$$\begin{aligned} Z_p^* &= 0(120) + 1(103) + 0(142) \\ &= 103 \text{ meters} \end{aligned}$$

and the error variance is:

$$\begin{aligned} \sigma_e^2 &= [w]^T [b] \\ &= [w_j]^T [\gamma(h_{ip})] \\ &= 0(1) + 0(13.4) + 1(0) + 0(19.1) \\ &= 0.0 \text{ m}^2 \end{aligned}$$

Thus, the known point is exactly predicted, and kriging is often referred to as being an "exact interpolator" (9:392).

2.7 Summary

This chapter reviewed digital TIROS-N imagery, current enlargement techniques, and kriging. The images are visible wavelength weather imagery at 4.0-km resolution and are stored in digital format. Currently, nearest-neighbor, bilinear interpolation, and cubic convolution are commonly used to enlarge digital images. Kriging, however, can also be applied to enlarge images since the pixel gray values are spatially distributed in two dimensions, and can be modeled by a regionalized variable. Punctual kriging was identified as the method to use. It uses point samples and produces point predictions. Universal kriging was shown to be applicable in situations where there was a drift in the regionalized variable. Simple and ordinary kriging were defined in the process and help complete the kriging types. Additionally, several variogram types, and the more common models, were presented. Each of the topics were discussed to the degree that they were used in this research.

III. Methodology

This chapter presents the methodology used in meeting the objectives outlined in Chapter I. The kriging methodology will be discussed by chronologically progressing through the data collection procedure and the application of each of the kriging programs. The included sections discuss data collection, structural analysis, universal kriging, and current enlargement techniques.

3.1 Data Collection

The visible wavelength imagery used in this research was collected by Dr. T. S. Kelso of the Air Force Institute of Technology. The initial imagery collected was 4.0-km resolution AVHRR data as received from the NOAA operated, TIROS-N weather satellites. The data was reconstructed by ground processing into a digital format. The resulting eight-bit data are termed to be in "raw" format. These eight digital bits allow the pixel gray values to range between 0 and 255. The resultant images were 800-by-1440 pixels in size.

All programs used in this research were written in the *C* or *C++* programming language by various authors who will be acknowledged when their program is introduced. A matrix object was written by Brodtkin (3) to facilitate programming efficiency and memory conservation. This research effort also involved defining simple interface programs (such as "krigeinterface.c" written by Duckett (12), Brodtkin (3), and McGee) which allow users to customize the operation of each program to their needs. Among other things, the specific header lines in the image data file were accommodated. Once the raw files were available and the appropriate interface files were established, the data was prepared for kriging using a three step process.

Step 1. The first step in preparing the data was to cut a tractable size sub-image from the original image. Image processing software provided by Kelso was used to cut 100-by-100 raw sub-images. This particular size was chosen to be compatible with current program capabilities while allowing reasonable processing times, and to allow sufficient room for enlargements. The original 100-by-100 sub-images used in this research are shown in Appendix B. The numbering scheme on the images is arbitrary, and is not intended to indicate that any images were discarded. The image contents are identified below.

- Image 0004 was chosen for its anticipated isotropic characteristics. It is a section of plains in the north-central United States located west of the Great Lakes. The image was recorded by the NOAA-11 satellite on Day 231 of 1991.
- Image 0007 is the western coast of Lake Michigan with significant cloud cover over the land. This sub-image was cut from the same image as Image 0004.
- Image 0010 is Hurricane Bob. It was recorded by the NOAA-11 satellite on Day 227 of 1991. The hurricane was then located off of the eastern coast of Florida.
- Image 0013 is a section of Cuba. It was recorded by NOAA-11 on Day 230 of 1991.
- Image 0014 is the Mississippi Delta region opening into the Gulf of Mexico. It was cut from the same image as Image 0013.
- Image 0019 is a section of southern Florida. It was also taken from the same image as Image 0013.
- Image 0022 is a section of the Rocky Mountains. It was recorded by NOAA-11 on Day 227 of 1991.

Image 0004 was selected as a sample which would display a high degree of correlation throughout the image. The remaining six images were chosen to contain varying degrees of clouds, land, and water. This mixture of different features was necessary to test kriging over a large variance in pixel values over the image areas. The test cases were chosen to span a wide range of possible variations in the input images.

Step 2. The second step was to convert the raw data into a form that could be processed. The first output of the conversion program is a file of data values which is compatible with the kriging programs. The data file lists the x-coordinate, y-coordinate, and the pixel value, z , in three columns. The second file is created in the Run-Length Encoded (RLE) format for display. This file was necessary to further view the 100-by-100 sub-images. The RLE functions were compiled from an existing RLE database by Parrott and McGee and belong to the Utah RLE 3.0 Toolkit. Some modifications were made to incorporate the image size and to write the data output. The usage of the program is as follows:

```
makedata infile outfile x-dimension y-dimension
```

The x-dimension and y-dimension are those of the original sub-image. In this case they were both entered as 100.

Step 3. The third step was to prepare the data file for future processing. Three header lines listing the desired processing parameters were added to the top of the data file. The first two header lines include parameters which are used by the variogram program, and the third line is a place-holder for the coefficients of the global trend polynomial calculated by the trend removal program. The header format that must be added to the data file is as follows:

```
NS      idir  step
phi     psi   k
0 0 0 0 0 0
```

Each parameter is described in detail below.

- *NS* is the number of samples in the image. *NS* is 10000 for the 100-by-100 pixel sub-images. No commas are allowed.
- *idir* is the number of variograms that will be calculated over the image. Initially, this value was set to one, and each directional variogram was calculated by separate runs of the program. The variogram calculation program calculates the 0° and 90° variograms regardless of this parameter setting. Subsequent variogram calculations used in this research used this updated version of the program. Other directional variograms can be calculated as specified below in the *Variogram Determination* section.
- The *step* value allows the variogram to be calculated in increments set by the user. The variogram does not necessarily need to be calculated at every incremental lag. As long as enough points are used, the continuous variogram model will still be representative of the discrete variogram. For example, if a 45° variogram was desired, the step should be at least $\sqrt{2}$, which is the diagonal length of a pixel. If the step is allowed to remain at one for calculation of a 45°

variogram, the variogram value will be zero, indicating the correlation between a point and itself. Since only 0° and 90° variograms were calculated for this research, the step size was always set to one.

- *phi* is the direction for which the variogram is to be estimated. The convention is for 0° to be in the north direction with angles increasing clockwise. Unless a variogram angle is specified on the command line of the variogram program, and a variogram angle specification flag is set in the interface file, only 0° and 90° variograms will be calculated, and the ϕ angle setting will be ignored.
- *psi* is the regularization angle. When the variogram is calculated in a desired direction, the variogram program must first determine which pairs of points to use in the calculation. It does so by looking at the variogram angle, ϕ , and including all of the sample pairs within a semi-inclusion angle, ψ on either side of ϕ . To decide which pairs of points are to be included in the variogram calculation in a particular direction, the angle between pairs of points must be within the regularization angle. Effectively, the angle being scanned ranges from $\phi - \psi$ to $\phi + \psi$, where ψ is the regularization, or semi-inclusion, angle. The variogram program can also be configured to skip the regularization angle and perform a quicker procedure of simply searching the rows and columns; however, to use this method, the input data must already be assigned to a regular grid. This search replaces the trigonometric calculations associated with the regularization angle and non-gridded data.
- *k* is the anisotropic correction factor for geometric anisotropy. The structural analysis of the sub-images did not include a manual determination of the anisotropic factor. Instead, geometric anisotropy was incorporated in the kriging program and is presented at the end of this chapter. *k* is only displayed in the variogram model. Therefore, it was set to one.

Upon completion of data collection and preparation, the structural analysis can begin.

3.2 Structural Analysis

A structural analysis includes a trend analysis, or determination of the drift, and estimation of the variogram. Also necessary is a treatment of any anisotropic effects. Since it is reasonable to believe that data sets obtained from nature are

nonstationary, it is assumed that stationarity does not initially exist in the images processed in this study; however, all of the kriging types assume that the mean of the regionalized variable is stationary (at least within a local neighborhood), whether or not it is known. The first step in any structural analysis must, therefore, be a trend analysis and subsequent removal of the global trend, followed by determination of the variogram.

3.2.1 Trend Analysis. It is assumed that some form of a trend existed in the mean of the regionalized variable. Therefore, stationarity was induced in a data set by fitting a function to the global trend in the regionalized variable and subtracting this function from the variable. Removing the global trend left the zero-mean residuals, as indicated by Equation 2.11, which were then used to calculate the variogram. A sample trend as shown in Figure 2.1 can be represented by a step function, a linear function, or a polynomial function; the choice is usually arbitrary (18:716). The global trend in the images was represented by a second-order polynomial of the form (18:718-719) (17:18):

$$z_i = \beta_0 + \beta_1 x_i + \beta_2 y_i + \beta_3 x_i y_i + \beta_4 x_i^2 + \beta_5 y_i^2 \quad (3.1)$$

where $i = 1, \dots, n$, and n is the number of known samples. In matrix notation this becomes:

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 \\ 1 & x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 \\ 1 & x_3 & y_3 & x_3 y_3 & x_3^2 & y_3^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & x_n y_n & x_n^2 & y_n^2 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{bmatrix}$$

This is also represented as $[A][w] = [b]$, and the goal is to estimate the β coefficients in the $[w]$ matrix. Matrix manipulation gives the following equations:

$$A^T(Aw) = A^T(b) \quad (3.2)$$

and

$$w = (A^T A)^{-1} A^T b \quad (3.3)$$

This is the solution to the least-squares regression equations (23:505). The global trend removal program solves for the β coefficients using this method of least-squares. The polynomial trend, \hat{Z} , was then calculated from $\hat{Z} = [A][w]$ and subtracted from the original data set, creating an output file of residuals. To save computer memory, only every other sample point was actually included in the A matrix. The first two header lines in the output file are the same as the first two input header lines. The zeros in the third header line are now replaced by the β coefficients, and the remaining columns in the output file are the x-component, y-component, and the residual pixel value, z , at the coordinate. The usage of the program is as follows:

```
resid [-i infile] [-o outfile] [-h]
```

Residual files were created using this program for each of the images used in this research. The [-h] option will print the usage line shown above.

An advantage of the least-squares polynomial regression technique is that it is capable of very closely modeling the trend of the data. Another advantage is that the trend is retained by the polynomial, and can be recombined with the kriged residuals. Also, since the trend model is continuous, a trend value can be calculated for positions that were not originally sampled. In kriging, for example, unsampled points can be predicted based on residual samples, and a trend value will exist for the unsampled location even though no trend was calculated at that point.

3.2.2 Variogram Determination. The trend analysis program outputs a residual data set where $E[Z(\vec{x})]$ is constant and nearly zero over the entire sub-image. The next step is to estimate the variograms of the residuals.

The variogram program identifies sample pairs lagged h apart and estimates the variogram in the chosen directions. As previously mentioned, the program will run faster if it is configured to take advantage of data which is already arranged on a regular grid. When the variograms were estimated for the initial sub-images, they were calculated by taking advantage of the grid structure. The residual pixel values were used, and 0° and 90° variograms were calculated. The curves in Appendix C

show the resulting variograms; both variograms are plotted in the same figure to test for anisotropy. All subsequent variograms were also calculated by taking advantage of the regularly spaced grid structure. This precluded the need for the program to search through the semi-inclusion angle for the sample pairs, and it reduced computation time. The usage of the variogram program is as follows:

```
varfit [-i infile] [-o outfile] [-p plotfile]
      [-v variogramfile] [-a variogram-angle] [-h]
```

The variogram program was written by Robinson and modified by Duckett. The variogram program creates three output files. The first file is a replication of the input data file. The variogram file is the third output file. It contains the results of the program. A sample output of this file is included in Appendix E. It shows that the mean drift (shown as elevation) is nearly zero. Also displayed in columns are the lag for each variogram calculation, the drift value at each lag, the variogram calculation $\gamma(h)$, and the number of samples which were included in each variogram calculation. The second output file is a subset of the variogram file. It contains the lag, h , and variogram values, $\gamma(h)$, necessary to plot the discrete variogram curve.

One observation is apparent from the variograms in Appendix C. For all of the sub-images, the sills (based on the spherical model) for the 0° and 90° variograms are roughly equal. Since zonal anisotropy is characterized by differing sills, the sub-images were deemed zonally isotropic. If this were not the case, it would be necessary to partition the sub-image into zonally isotropic regions in order for a single variogram to be used by the kriging program. Duckett has written a C++ partitioning program which divides the data set based on row and column sum comparisons to their respective median (12). Partitioning can occur before or after global trend removal. New variograms would then have to be calculated for each zone. After kriging is performed, the data set would have to be reassembled.

Further observation of the variograms reveals that geometric anisotropy is present in some of the sub-images. The indication of geometric anisotropy is that the range of correlation, as defined in the spherical variogram model, for each directional variogram is different. While it is difficult to detect any geometric anisotropy in Image 04, it is much more clearly defined in Image 07 and Image 13. Image 14 may

also be geometrically anisotropic. Images 10, 19, and 22 appear to be geometrically, as well as zonally, isotropic. While geometric anisotropy is present in the sub-images indicated, no change in the continuous variogram model is made. However, geometric anisotropy is accounted for in the kriging program and will be discussed below. No nugget effect is detectable in any of the discrete variograms.

If it is desired to calculate variograms at other than the standard 0° and 90°, two parameters should be set in the interface file. The first parameter is `var_angle_specified`. This is a flag which is set to one and then compiled with the variogram program. Then the `var_angle` parameter can be set to the desired direction. Alternatively, the [-a] option on the command line of the program allows the variogram angle to be defined.

The last items shown in the variogram output file are the variogram models calculated by the program and the simple correlation showing how well each of the models were fitted to the discrete curve. The results are presented next.

All three variogram models were plotted, and the results are included in Appendix D, along with the mathematical model representations. Fitting the spherical model to the 0° variogram of Image 13 produced a negative nugget effect value. Since the variogram is defined as the mean squared difference of the sample pairs, a negative nugget is actually impossible. Therefore, prior to kriging Image 13, the nugget was reset to zero. This is in agreement with the actual value displayed by the discrete variogram calculation. The simple correlations for each of the three models show how well the fits performed. The results are shown in Table 3.1 and Table 3.2. Both tables indicate that the spherical variogram model is generally

Direction	Image Number						
	04	07	10	13	14	19	22
0°							
Linear	0.992	0.924	0.670	0.754	0.430	0.618	0.508
De Wijsian	0.841	0.951	0.898	0.873	0.804	0.911	0.864
Spherical	0.993	0.937	0.980	0.997	0.787	0.942	0.780

Table 3.1. Simple Correlations of 0° Variogram Models

the best correlated of the three. The spherical variogram model was incorporated into the kriging program by Grant and Robinson because it was the most flexible

Direction	Image Number						
	04	07	10	13	14	19	22
90°							
Linear	0.505	0.507	0.637	0.963	0.316	0.318	0.605
De Wijsian	0.864	0.778	0.956	0.927	0.727	0.726	0.894
Spherical	0.872	0.982	0.910	0.990	0.694	0.763	0.970

Table 3.2. Simple Correlations of 90° Variogram Models

of the three models. Another reason it was incorporated is that it offered the most parameters of the three models. Image 14 would have benefitted from use of the De Wijsian model. However, at the time it was kriged, the De Wijsian model was not yet incorporated into the kriging program.

3.3 Universal Kriging

Armed with the variogram, the kriging process can continue with performing the predictions and calculating the error variances. Universal kriging was chosen to perform the calculations based on its ability to handle any remaining local drift. To address the implementation of universal kriging, the topics discussed are the kriging assumption, the kriging program, and the error variance calculation.

3.3.1 Kriging Assumption. Kriging must make an assumption as to the stationarity of the data. Global stationarity over the sub-image was required for estimation of the variogram. However, universal kriging allows assumption of the least restrictive case of stationarity—the intrinsic hypothesis (intrinsic assumption). Under the intrinsic hypothesis, stationarity must be attained at least within the local neighborhood. Therefore, it is not necessary to krig the residual data as it may contain as much local drift as does the original sub-image data. The procedure, then, is to determine the order of the local polynomial to include in the universal kriging equations. The method used was to examine the overall global trend in the data and to make an inference as to the local drift present. The coefficients of the global trend for the seven original sub-images are listed in Table 3.3.

Table 3.3 indicates that the global trend is dominated by the first-order terms in all of the sub-images. A local linear trend was, therefore, incorporated into the universal kriging equations, as was demonstrated in Chapter II, and kriging was

Image Number	β_0	$\beta_1 X$	$\beta_2 Y$	$\beta_3 XY$	$\beta_4 X^2$	$\beta_5 Y^2$
04	56.4247	.1263	.0973	.0007	.0006	-.0015
07	116.6112	-1.8049	.4621	-.0086	.0230	-.0054
10	40.9010	3.2095	1.4883	.0076	-.0365	-.0189
13	-9.3885	.9023	2.3841	-.0180	.0021	-.0159
14	26.3550	.8139	.8031	.0035	-.0096	-.0026
19	86.3372	.4613	.2158	.0111	-.0108	-.0085
22	83.1519	.2721	.1914	-.0061	-.0048	-.0048

Table 3.3. Global Polynomial Trend Coefficients

performed on the original data. The kriging program was written in C++. Grant, Brodtkin, and Robinson are the authors of the program (15) (3).

The capability to perform kriging on residual data has been retained. After kriging residual data, a simple program is necessary to recombine the global trend polynomial with the kriged residuals. Since the polynomial is a continuous function, trend values will be available at grid locations that did not contribute to the polynomial's determination. The usage of the rebuild program which performs this function is described here:

```
rebuild [-i infile] [-o outfile] [-h]
```

The rebuild program was written by Duckett. The usage of the program is self-explanatory, except that the [-h] option will print the usage line shown above.

3.3.2 Kriging Program. Discussion of the kriging program involves its implementation, the treatment of anisotropy, and the resulting image size.

3.3.2.1 Implementation. The kriging program is implemented by setting the necessary parameters in a control file and executing the program on the control file. Then, the known samples within a neighborhood are used in the prediction of the unsampled pixels in the enlarged image. The neighborhood can be calculated by the kriging program or set by the user.

The kriging program will calculate a square neighborhood (or rectangle), or kernel as it is known in image processing, with a diagonal radius equal to the range.

This is shown in Figure 3.1 where IAX and IAY are termed the semi-inclusion dis-

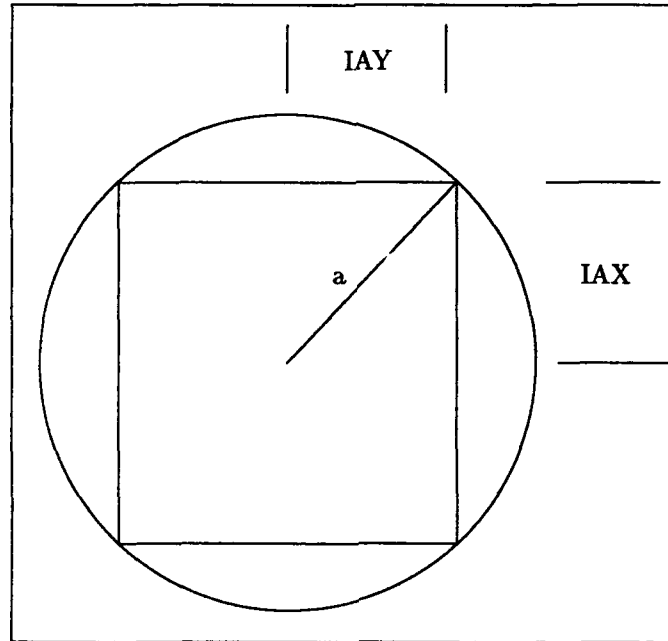


Figure 3.1. Kriging Program Calculated Kernel

tances (3) and are half the length of the side of the kernel. The pattern below shows four neighboring pixels of the original sub-image displayed on the four times expanded grid,

```

X . . . X
. . . . .
. . . . .
. . . . .
X . . . X

```

where "X" is a known point, and "." represents a grid location to be predicted. This structure would be included many times within the kernel defined by the kriging program. No points will be included in the kernel if they are beyond the range. For this research, the kernel size was set to a smaller moving square. The smaller kernel

includes fewer known samples and will produce less than optimal predictions. However, Davis indicates that neighborhood reduction is a common practical application in kriging and that it does not significantly affect the predictions (10:599-600).

To set the kernel, the semi-inclusion distances are set in the control file. Since the kernel is applied to an enlarged grid, the semi-inclusion distances are multiplied by the expansion factor. At the time that the sub-images were processed, the kernel was slightly enlarged to include one extra row and column on each edge. The concept of a kernel as applied in kriging is different from that which is applied in the other convolution techniques. Since the kriging kernel is applied to an expanded grid, the number of known pixel values within the kernel is dependent on where the kernel is centered. Therefore, the "effective kernel size" only includes known pixels on the expanded grid; it varies as the kernel moves over the image.

In this research, the semi-inclusion distance was set to two, and the effective kernel size included 16, 20, or 25 pixels. This is comparable to the 16 pixels included in the cubic convolution technique.

The initial sub-images shown in Appendix B were expanded four times by using a minimization path in the kriging program. The impetus for using this method was the amount of computer processing time that was saved. The minimization path through the kriging program takes advantage of the grid structure of the data.

If a 100-pixel-square sub-image is kriged into a four times enlargement using the expansion method, 150,000 matrices are formed containing the neighborhood samples, and 150,000 matrix inversions are calculated. Most of these calculations become unnecessary by designating the minimization path through the kriging program and only calculating the weights once for the central region.

One important assumption is made when applying this technique; the local drift in the central blocks are identical. The local drift coefficients are included in the universal kriging equations and affect the solution for the weights. The weights are thus not only determined by the distance between samples, but also on the local drift. To replicate the weights between central blocks implies that the local trend within each block is equal. While this may not be true, kriging Image 13 by using the minimization path and by simply expanding it (with the kriging program) produced no noticeable visual difference. However, using the same computer for two test runs, minimization took approximately twenty-five minutes for a job which would have taken days using the expansion method.

One of the parameters included in the control file is the nugget effect, C_0 , of the directional variograms. Since it is possible for the nugget effect to be modeled as a negative number, a criteria was established to offset this error. If C_0 was negative and $C_0 \leq 10$ percent of the sill, then the nugget was entered in the control file as being zero. This was performed on the 0° variogram entry for kriging Image 13.

3.3.2.2 Anisotropy. The known sample points inside the defined kernel are used to calculate the second-order moments used in the universal kriging equations. However, the moments are based on h —the distance between known samples, and, unless the sub-image is isotropic, h is a vector and depends on the relative direction between the samples. Hence, the parameters for both of the variograms are entered into the kriging control file and, before the second-order moment can be calculated, geometric anisotropy must be addressed and a single variogram model produced.

The kriging program treats zonal anisotropy by simply producing an average value for the sills of the two variograms. For geometric anisotropy, the \vec{h} is decomposed into two components, h_1 and h_2 , corresponding to the image axes. The anisotropic correction factor, k , is then used to scale the range of one variogram to match the range of the other. For example, if the k factor is the ratio of the range of the 0° variogram to the range of the 90° variogram, then the distance between two points located at (x_1, y_1) and (x_2, y_2) is represented as $h' = \sqrt{k^2(x_1 - x_2)^2 + (y_1 - y_2)^2}$, and the variogram calculated in the 0° direction is used (8:134-135). h' is then used as the distance between two points to calculate the second-order moment.

3.3.2.3 Kriged Image Size. The kriging program does not calculate predictions outside of the last row and column of samples on the expanded grid. The 100-by-100 pixel sub-images are placed on a four-times enlarged grid at locations 0 through 396. Since no predictions are made outside the last data points, the resultant kriged image is 397-by-397. These are the resultant image sizes produced in this thesis.

Beginning with a 51-by-51 initial sub-image size and expanding four times will produce a 201-by-201 kriged image since the 51st pixel is placed on grid location 201. In Chapter IV, four-times expansions were performed on two 51-by-51 sub-images.

3.3.3 Error Map. The estimation error variance calculations are written to an output file for each of the points which were predicted by universal kriging. These variance values can be used to create a visual map of how well the predictions were made. For irregularly spaced data, the error map would provide a useful indication of the amount of error associated with the predictions.

Exactly predicted pixels will have a zero error variance. Therefore, at locations with known pixels, a black pixel will be present on the error map. At locations which were predicted, the variance can be used to establish an error map pixel value based on the confidence interval of the prediction.

If a Gaussian normal distribution is assumed for the prediction errors (8:56), then the standard error is simply the square-root of the error variance, and the pixel value for a 95 percent confidence interval can be established as in Equation 2.54. This pixel value can be displayed at each location over the image to produce an error map. A typical error map from a kriged image is presented in Chapter IV.

3.4 Current Enlargement Techniques

The nearest-neighbor, bilinear interpolation, and cubic convolution image processing techniques were also applied to each of the sub-images. Four-times enlargements were produced for each of the sub-images using each of these three currently accepted methods of enlargement. As mentioned in Chapter I, the only method which retains the original pixel information is the pixel replication used by the nearest-neighbor method, but it has the disadvantage of producing "block-like" images which are difficult to interpret. Bilinear interpolation and cubic convolution alter the original pixel values over the entire image and are not faithful representations of the original image data. The results of each technique are compared to those produced by kriging and are presented in Chapter IV.

3.5 Summary

The data collection was performed by cutting seven 100-by-100 sub-images of raw NOAA TIROS-N AVHRR data and converting them into data files compatible with the kriging software. Stationary residuals were created by fitting a second-order polynomial to the global trend and removing it from the data. The structural analysis of 0° and 90° variograms of residuals indicated zonal isotropy. Analysis of the

global trend polynomial indicated a local linear drift. Models were fitted to the variograms and the spherical model parameters were entered into the universal kriging program. Geometric anisotropy was handled automatically by the kriging program by calculating the anisotropy ratio, k , as a ratio of the range of the 0° variogram to the range of the 90° variogram, and a scaled version of the separation vector, h , was produced and used in the calculation of the covariance. The covariance function replaced the variogram for numerical stability reasons. The resulting predictions using universal punctual kriging produced kriged images and error maps. Images were also produced using the nearest-neighbor, bilinear interpolation, and cubic convolution techniques.

IV. Results

Nearest neighbor, bilinear interpolation, cubic convolution, and universal punctual kriging were applied to seven TIROS-N AVHRR, 4-km resolution satellite sub-images. The original sub-images are presented in Appendix B. Cubic convolution and kriging were also applied to one aerial photograph. This chapter includes the initial results, the results obtained from an improved structural analysis, a numerical analysis comparison of kriging and cubic convolution, and a typical estimation error variance (error map) presentation.

4.1. Initial Results

The first kriged sub-images were produced using the spherical variogram model, universal kriging, and a kernel size defined by a semi-inclusion distance of two. There were 16, 20, or 25 pixels used in each predicted point. This is a maximum effective kernel size of 25 pixels. Geometric anisotropy was corrected, and no zonal anisotropy was present. This initial trial did not produce results of the desired quality. These images are shown in Figure 4.1 through Figure 4.7.

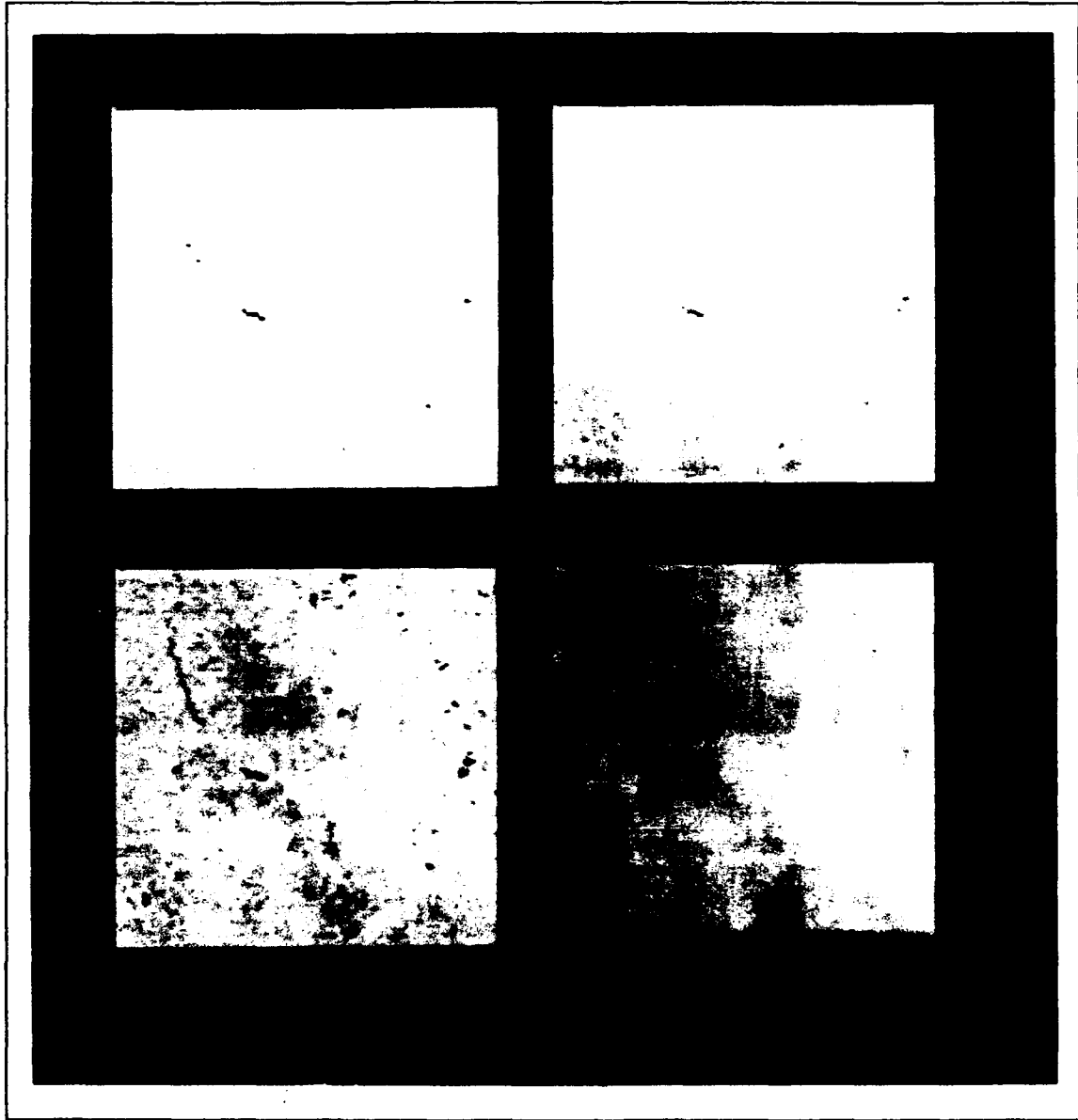


Figure 4.1. Image 04, Convolution Comparison

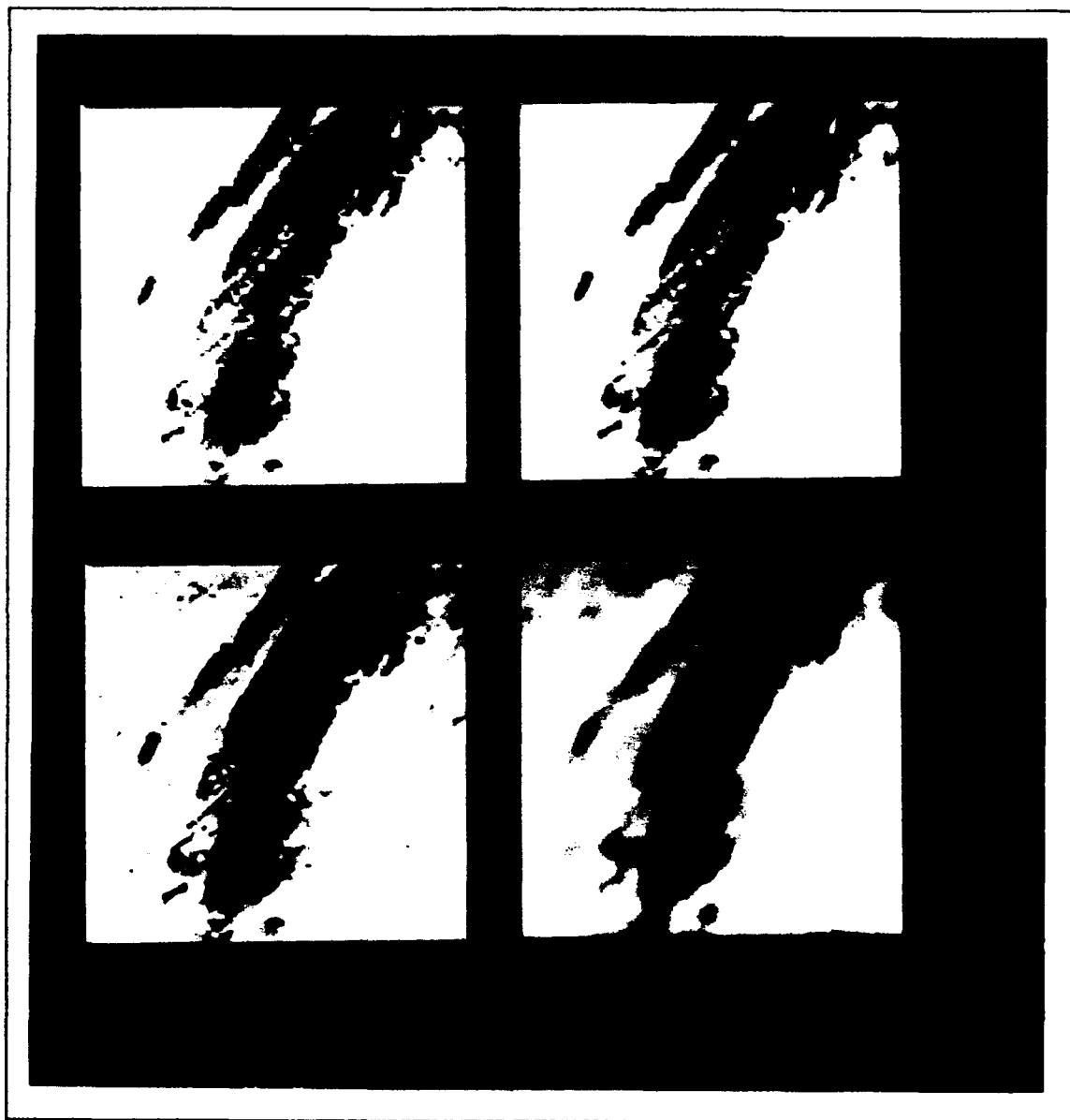


Figure 4.2. Image 07, Convolution Comparison

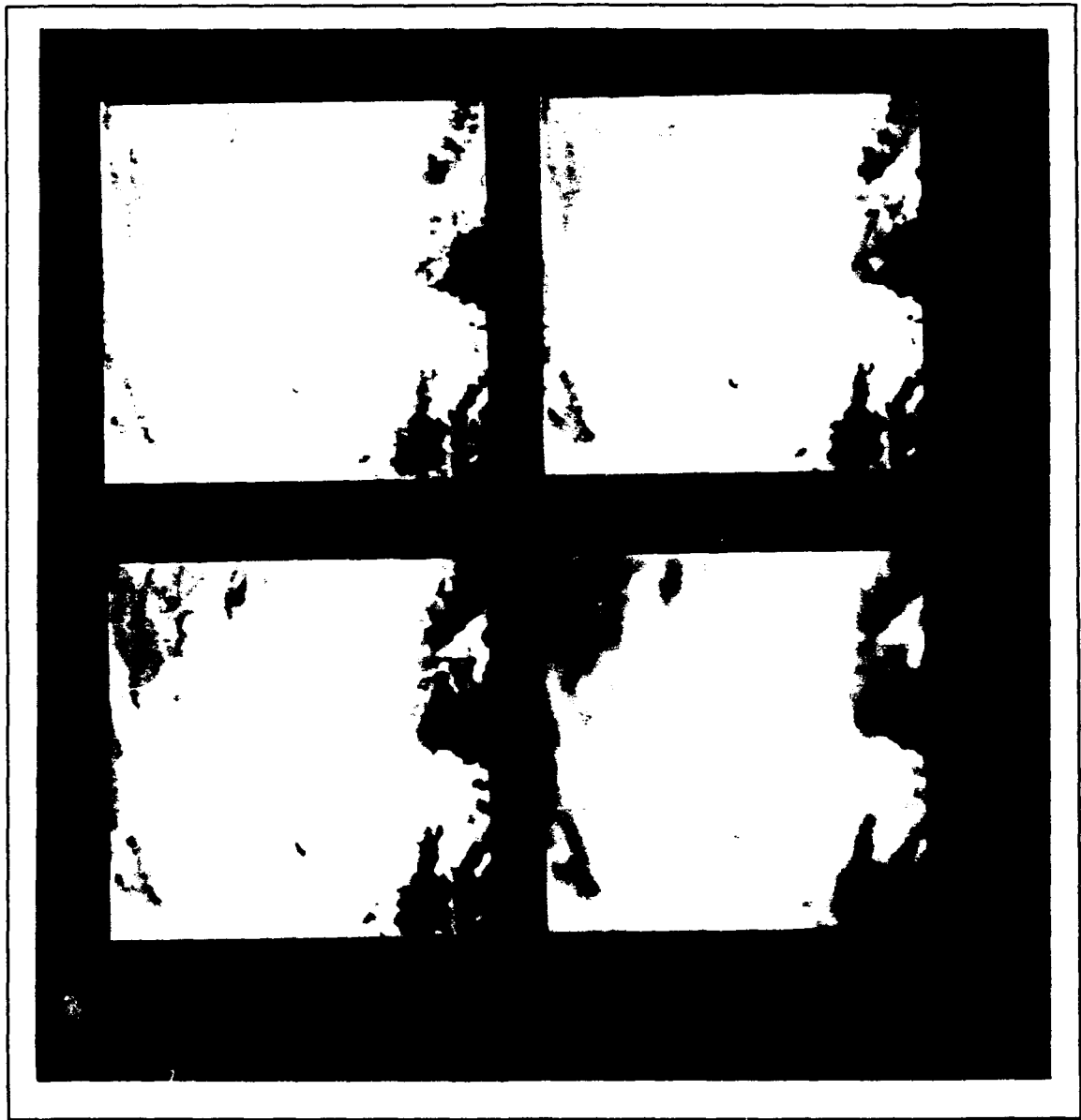


Figure 4.3. Image 10, Convolution Comparison

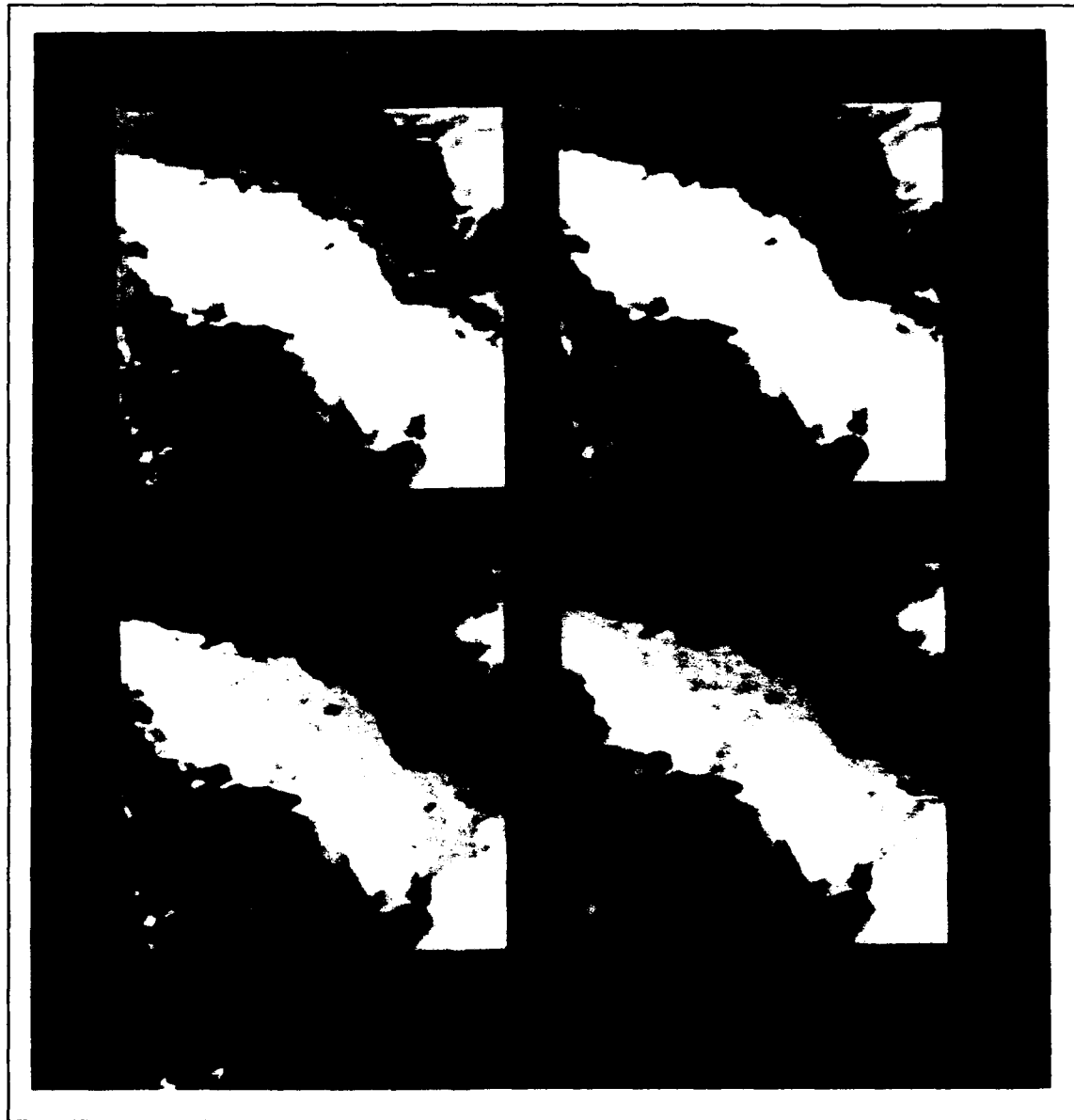


Figure 4.4. Image 13, Convolution Comparison



Figure 4.5. Image 14, Convolution Comparison

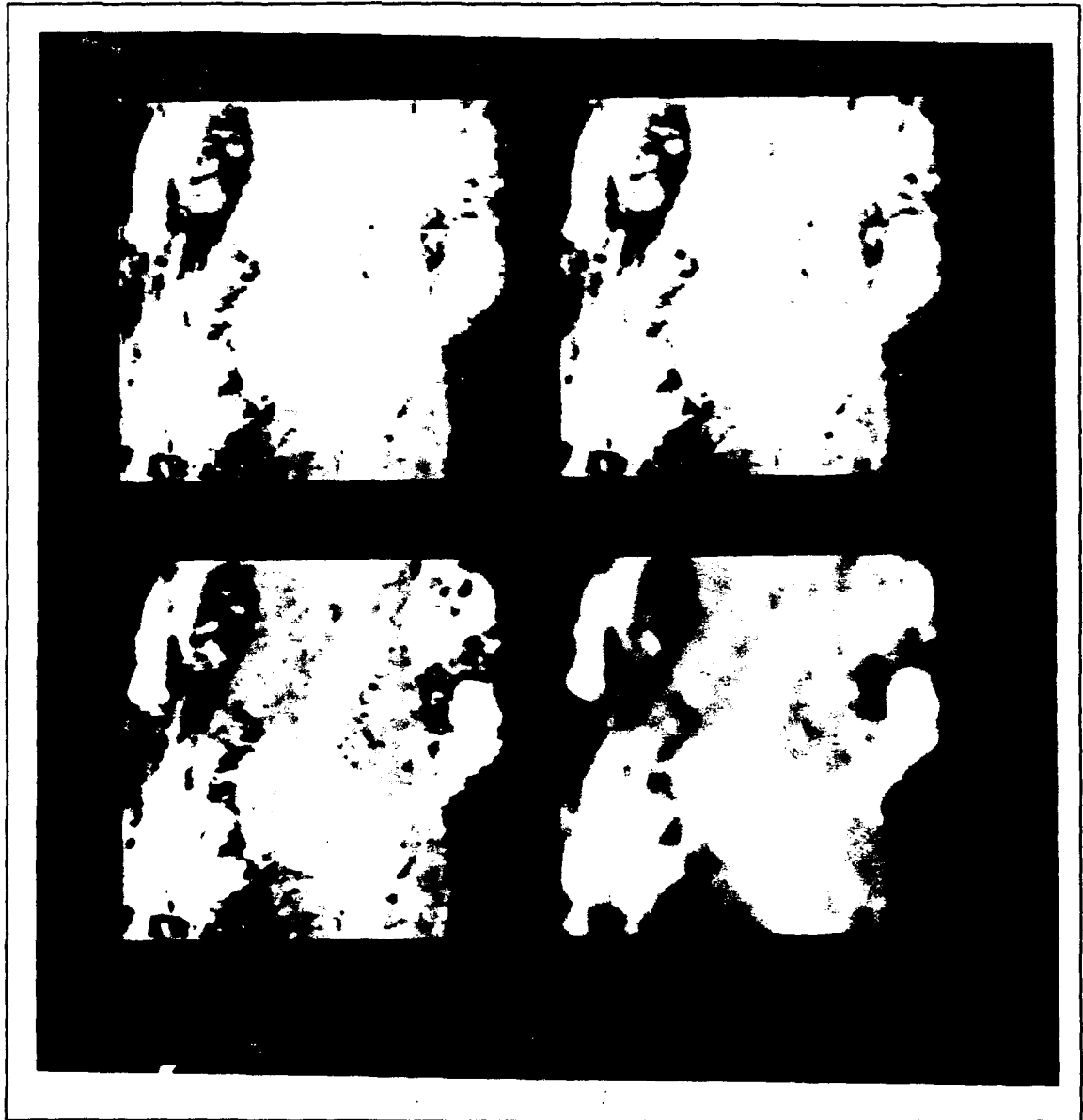


Figure 4.6. Image 19, Convolution Comparison



Figure 4.7. image 22, Convolution Comparison

To determine whether or not the kernel size was a significant factor in the blurring, Image 04 was re-kriged with a semi-inclusion distance of one. Four to nine pixels would be included in each prediction using this kernel size. The resulting image was still blurry. An improved structural analysis was, therefore, indicated if a sharper image was to be produced.

4.2 Improved Structural Analysis

The source of the blur in the initial images was determined to be the model representation of the variogram. Therefore, an improved structural analysis is necessary in order to produce sharper images.

The problem is that the second-order moment was being calculated from a model which did not adequately represent the variogram. David mentions that it is not very important for the C and a parameters of the spherical model to be misinterpreted, but more care should be taken in interpreting C_0 (8:122). While the spherical variogram model generally produced the best fit, it did not usually fit well at small lags, and the small lags are exactly what are being used within the kernel to make predictions.

The largest lag that is being used can be determined by multiplying $\sqrt{2}$ by the expansion factor and the semi-inclusion distance. This distance is half of the kernel diagonal and is the largest lag that can exist between the predicted pixel and a known pixel within the kernel. The result is that lags of 11.3 or less are being used to make predictions.

The poor fit of the spherical model is clearly evident in the variograms of the sub-images. For instance, the 0° spherical model fit of Image 13 displayed $C_0 = -111.081$, but the discrete variogram is actually zero at $h = 0$. It can be seen from the variograms in Appendix D that there was a significant nugget effect in each of the spherical curve fits. The nugget effect in the models is artificial in that it is not representative of the true variogram.

The result was that the kriging weights were calculated from an inadequate representation of the variogram, and the known pixel values were improperly weighted in the predictions. Thus, the kriged sub-images were blurry.

The solution was to use a model which fit well near the origin of the variogram. The nugget effect of the spherical model was, therefore, set to zero within the control

file before kriging the sub-images. This forced the spherical model to begin at the origin. While this caused the model to fit poorly at other areas of the curve, it was irrelevant since no large lags were being used.

The sub-images were re-kriged using zero for the nugget of both of the 0° and 90° variograms. The semi-inclusion distance was again set to two. With a better model fit at short lags, the kriged images became comparable to those which were produced by cubic convolution. The resulting comparisons are shown in Figure 4.8 through Figure 4.14.

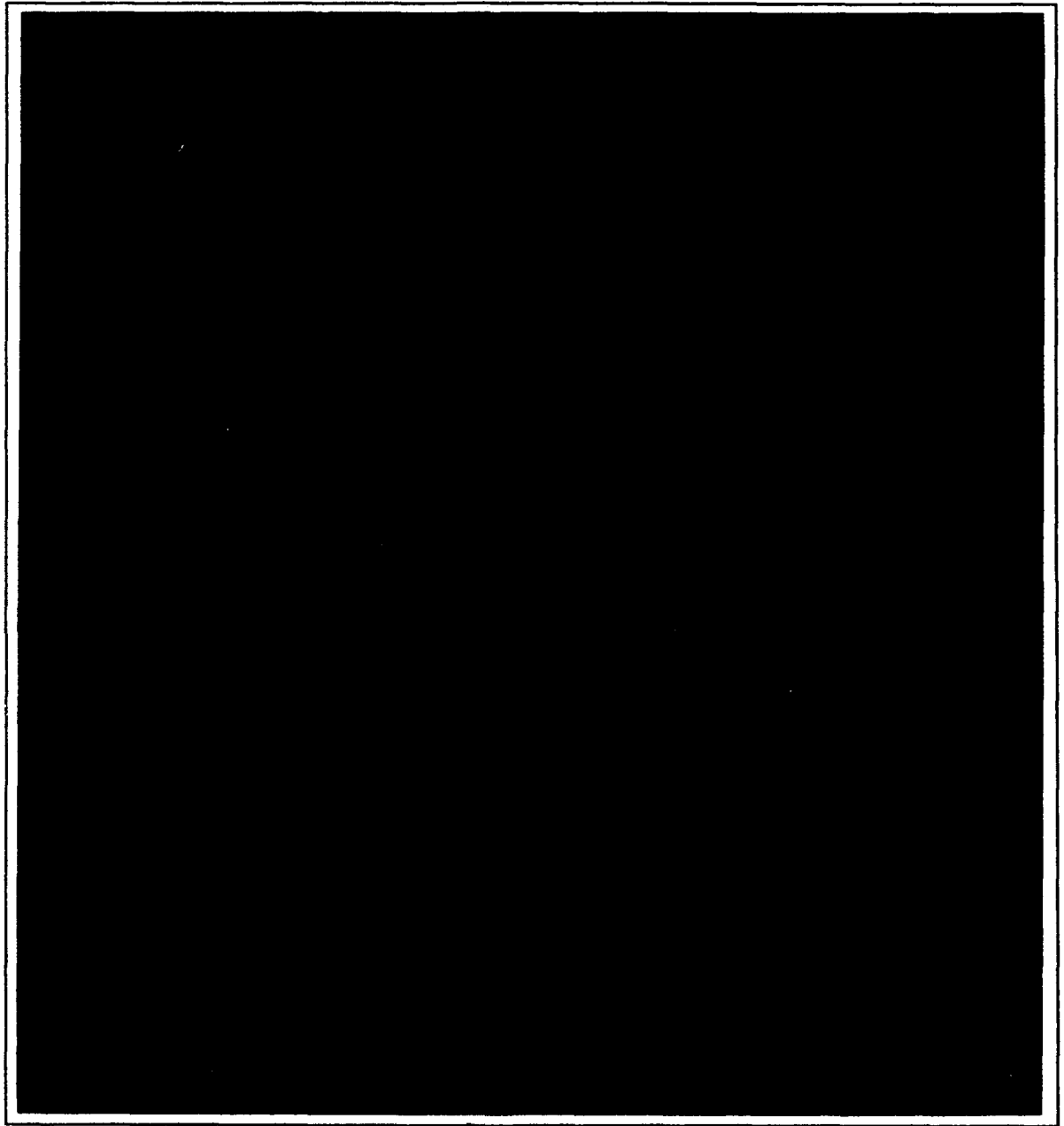


Figure 4.8. Improved Image 04, Convolution Comparison



Figure 4.9. Improved Image 07, Convolution Comparison

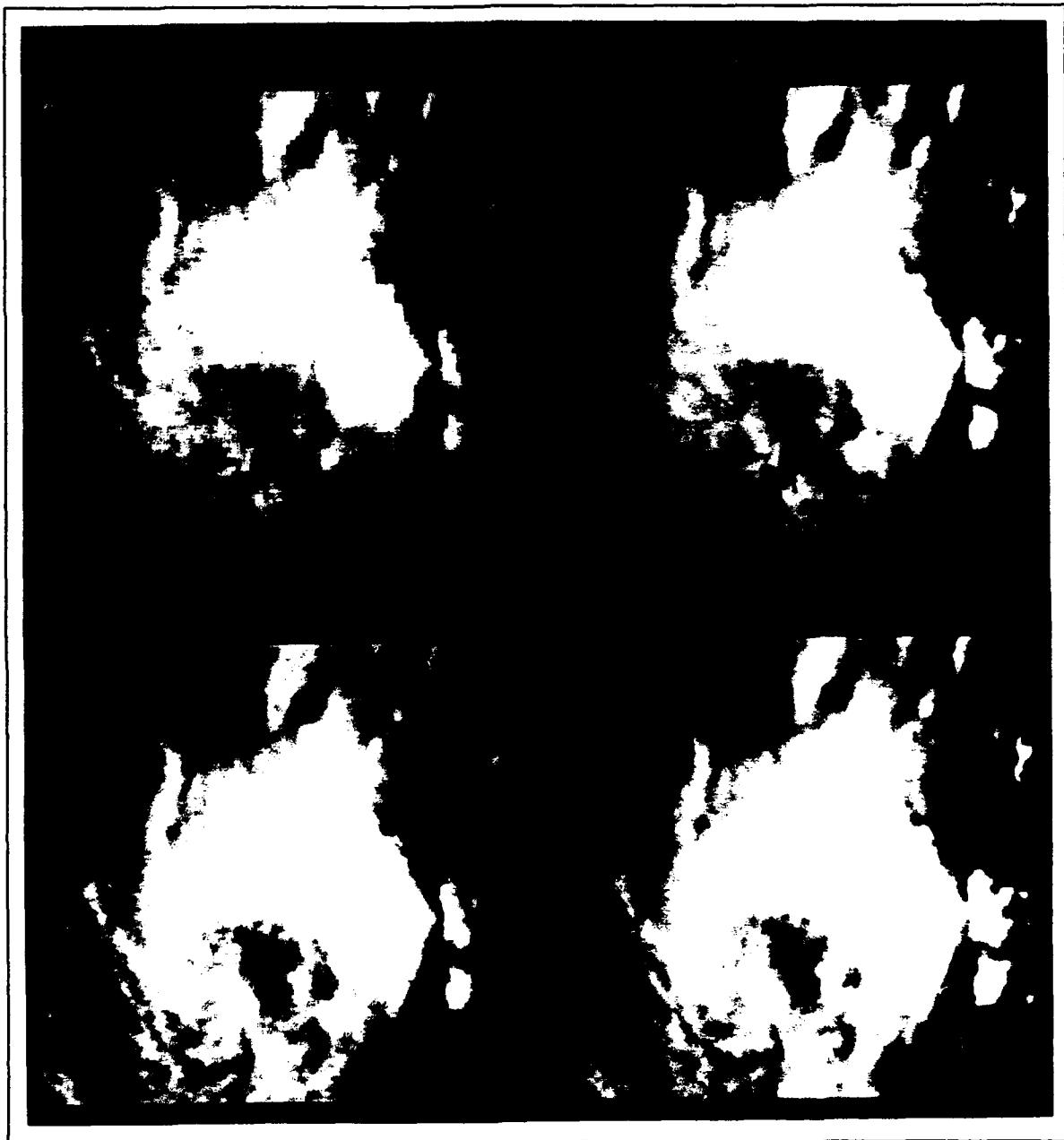


Figure 4.10. Improved Image 10, Convolution Comparison



Figure 4.11. Improved Image 13, Convolution Comparison



Figure 4.12. Improved Image 14, Convolution Comparison

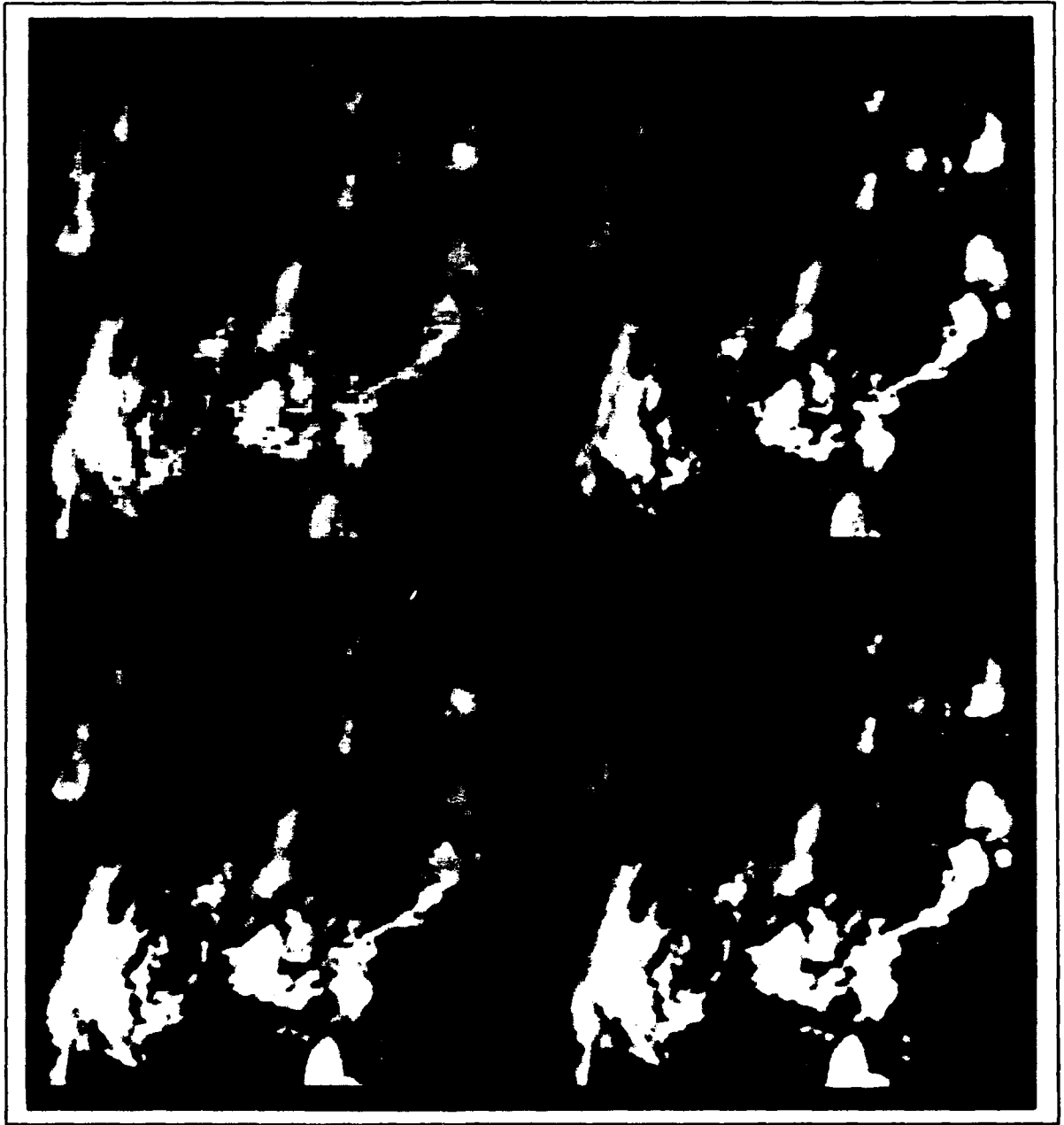


Figure 4.13. Improved Image 19, Convolution Comparison



Figure 4.14. Improved Image 22, Convolution Comparison

4.3 Numerical Analysis Comparison

To determine a quantitative comparison of kriging and cubic convolution, two new sub-images were cut to a size of 201-by-201 pixels. The first sub-image is just a larger cut of the Mississippi Delta region of the original Image 14. The second sub-image is an aerial photograph of a residential area. This photograph was used to test the results of kriging higher resolution imagery. The original images are shown in Figure 4.15.

The methodology used to compare kriging and cubic convolution was to sub-sample the images down to a size of 51-by-51 and use both techniques to enlarge the images back to their original size. Sub-sampling simply consisted of retaining every fourth column and row of the image. Then, the convolved images were subtracted from the original. The difference between a kriged or cubic convolved image and the original produced a differenced image for each of the two techniques. Next, statistics were collected on the differenced images to determine how well each technique could reproduce the original image.

To implement this methodology, a structural analysis had to be performed. First, the global trend was removed from each of the images, and the variogram of residuals was produced in the 0° and 90° directions. The variograms are shown in Figure 4.16 and Figure 4.17.



Figure 4.15. Original Images, 201-by-201 pixels

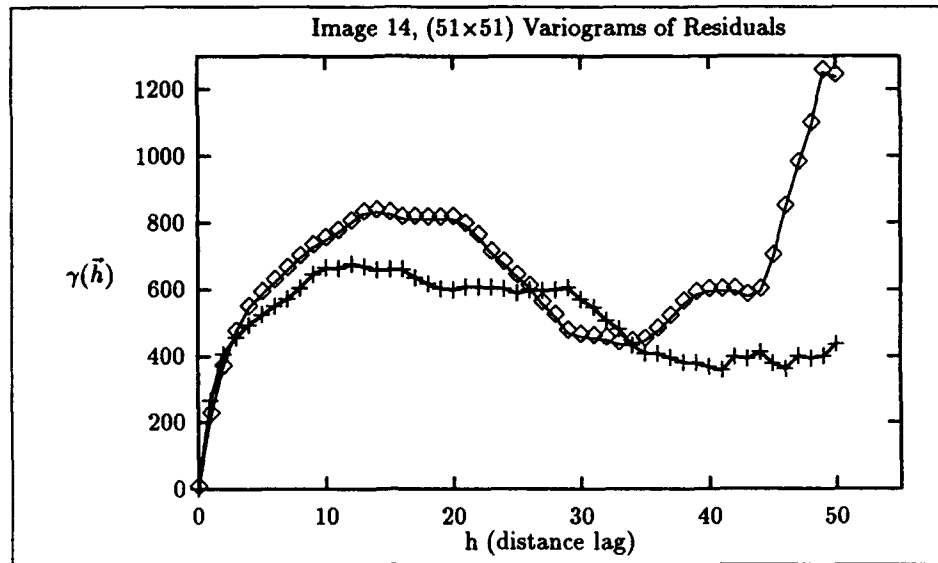


Figure 4.16. Mississippi Delta, Variograms of Residuals

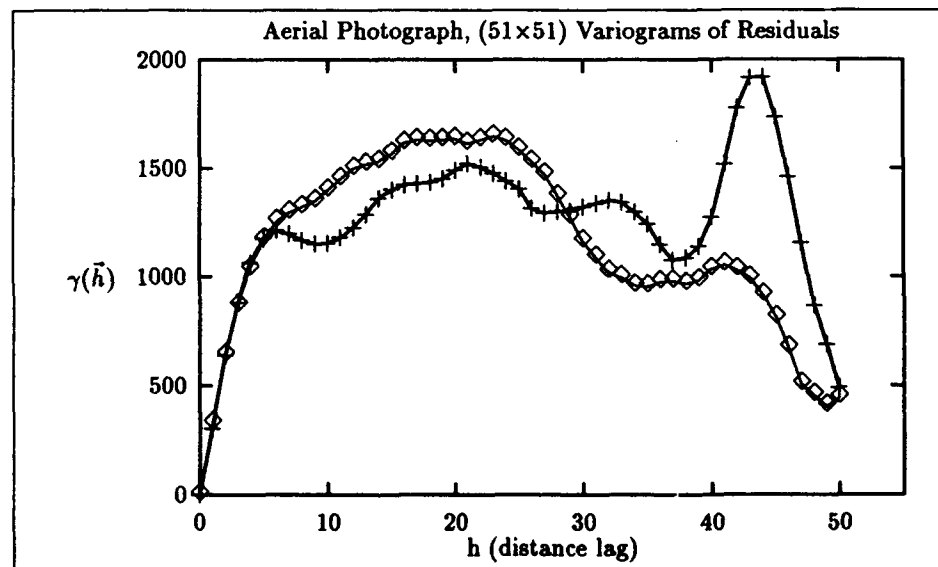


Figure 4.17. Aerial Photograph, Variograms of Residuals

Then, the variograms were fitted with the linear, De Wijsian, and spherical models. The 0° and 90° continuous variograms for Image 14 and the aerial photograph are shown in Figure 4.18 through Figure 4.21.

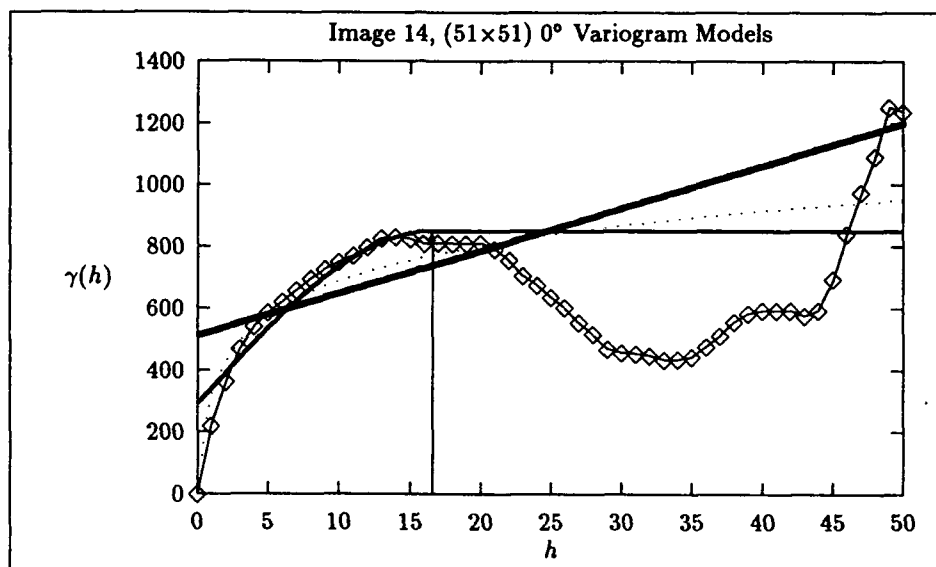


Figure 4.18. Mississippi Delta, 0° Variogram Models

$$\text{Spherical model} = 554.555 \left(\frac{3h}{(2)16.611} - \frac{1}{2} \left(\frac{h}{16.611} \right)^3 \right) + 296.769$$

$$\text{Linear model} = 513.370 + 13.754h$$

$$\text{DeWijsian model} = 314.188 + 162.915 \ln h$$

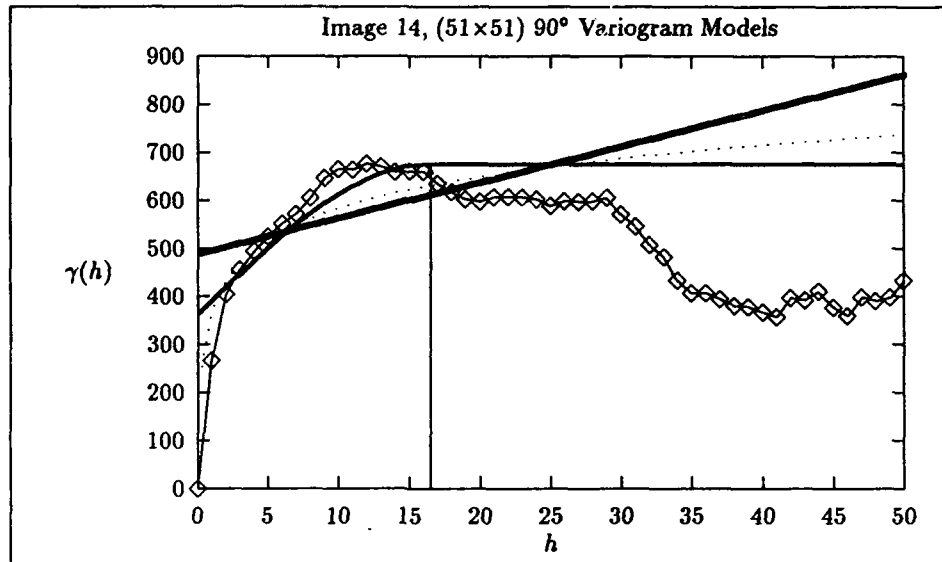


Figure 4.19. Mississippi Delta, 90° Variogram Models

$$\text{Spherical model} = 311.030 \left(\frac{3h}{(2)16.497} - \frac{1}{2} \left(\frac{h}{16.497} \right)^3 \right) + 363.913$$

$$\text{Linear model} = 487.932 + 7.483h$$

$$\text{DeWijsian model} = 364.519 + 95.120 \ln h$$

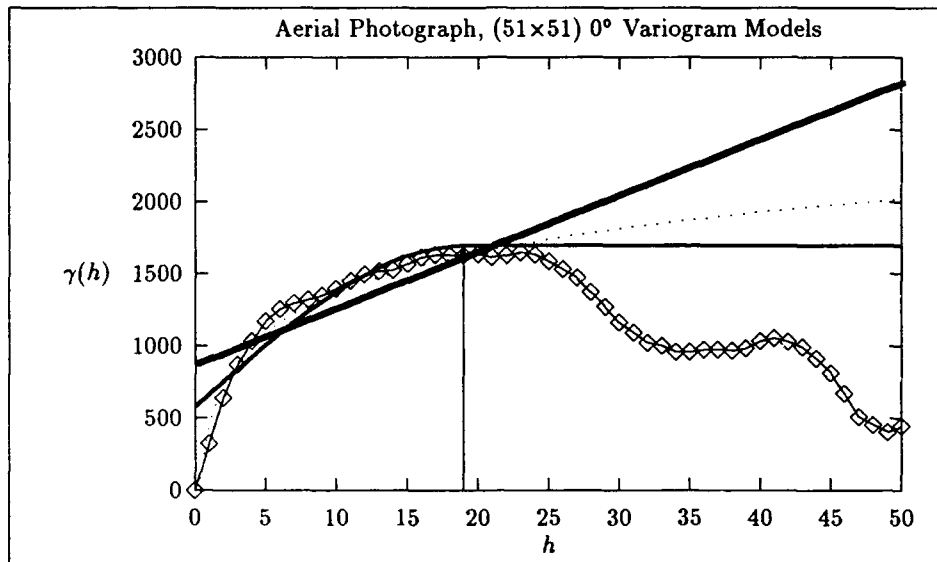


Figure 4.20. Aerial Photograph, 0° Variogram Models

$$\text{Spherical model} = 1111.239 \left(\frac{3h}{(2)18.997} - \frac{1}{2} \left(\frac{h}{18.997} \right)^3 \right) + 581.048$$

$$\text{Linear model} = 871.215 + 39.083h$$

$$\text{DeWijian model} = 447.315 + 401.693 \ln h$$

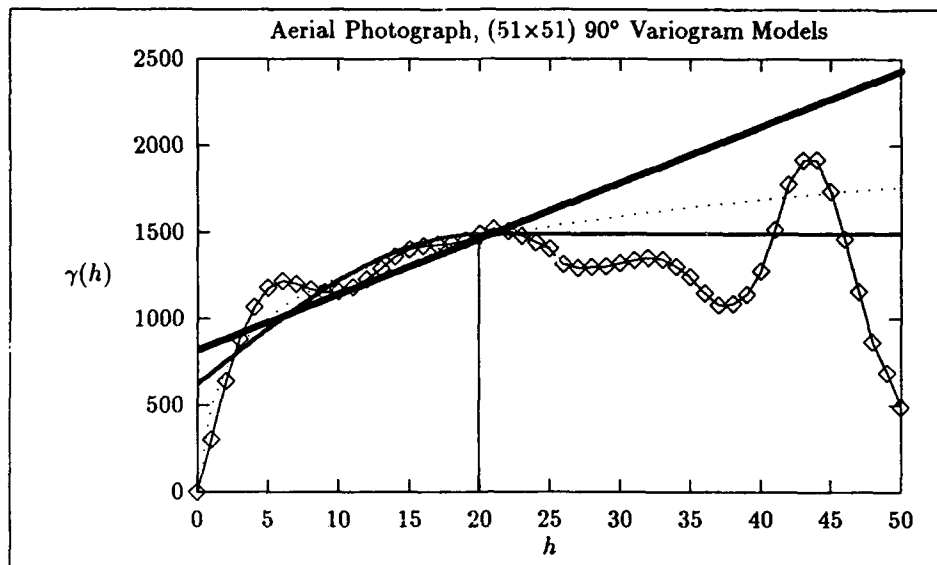


Figure 4.21. Aerial Photograph, 90° Variogram Models

$$\text{Spherical model} = 864.284 \left(\frac{3h}{(2)19.911} - \frac{1}{2} \left(\frac{h}{19.911} \right)^3 \right) + 624.832$$

$$\text{Linear model} = 820.841 + 32.241h$$

$$\text{DeWijian model} = 480.006 + 327.554 \ln h$$

The spherical variogram model was chosen and added to the kriging control file. The nugget was reset to zero in all of the models. The semi-inclusion distance of two was used, and a four-times expansion was performed to create kriged images of 201-by-201 pixels. The cubic convolution image and kriged images are shown in Figure 4.22 and in Figure 4.23 for the aerial photograph and the Mississippi Delta image, respectively. Also shown are the differenced images, which were produced by subtracting the convolved images from the original.

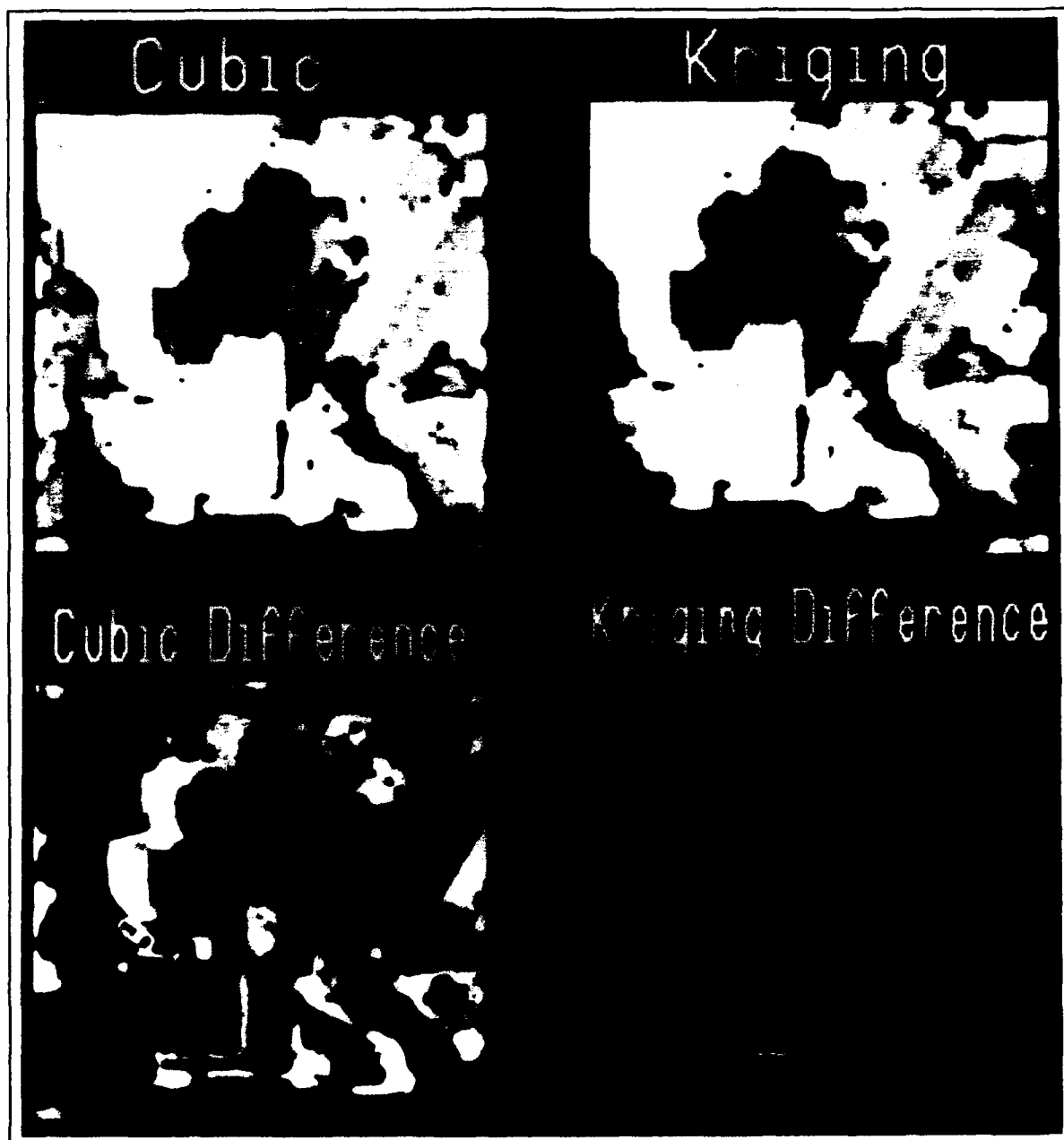


Figure 4.22. Comparison of Convolutions with Full-Sized Aerial Photograph

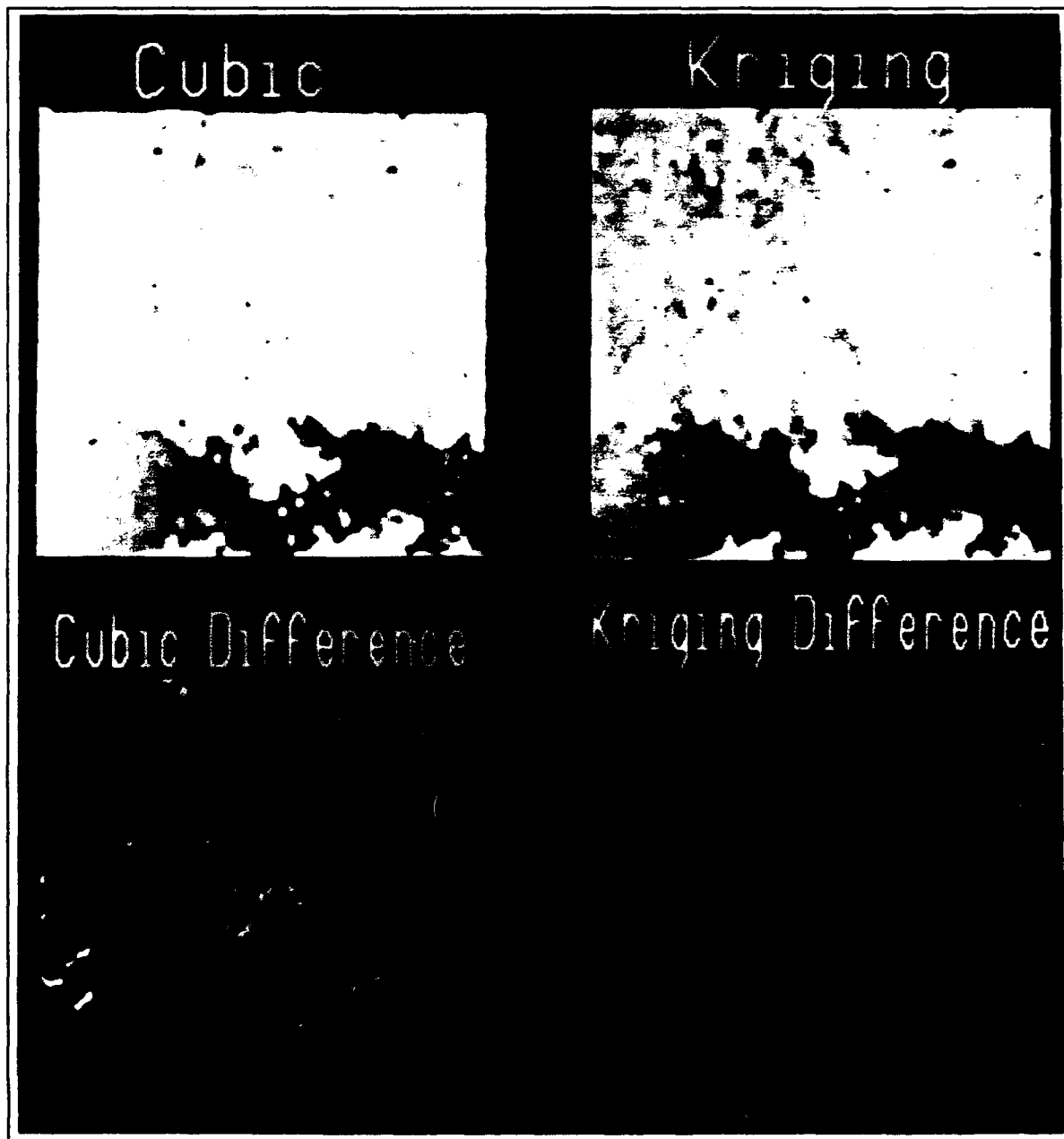


Figure 4.23. Comparison of Convolutions with Full-Sized Mississippi Delta

The mean and standard deviation of the differences were then calculated, and are presented in Table 4.1 and Table 4.2.

Mississippi Delta	Cubic	Kriging
μ	4.929	4.752
σ	10.068	10.076

Table 4.1. Mississippi Delta Differenced Data Distribution

Aerial Photograph	Cubic	Kriging
μ	14.780	2.309
σ	27.321	5.782

Table 4.2. Aerial Photograph Differenced Data Distribution

In both of the differenced images, the mean difference using kriging was less than the mean difference using cubic convolution. The standard deviation of the difference was slightly greater using kriging than using cubic convolution for the Mississippi Delta image, but the standard deviation of the difference was much less using kriging than using cubic convolution for the aerial photograph. The results do not show a statistically significant difference between the kriging and cubic convolution techniques when applied to the Mississippi Delta image. However, the improved results using kriging are prominent on the high-resolution aerial photograph. These results demonstrate that an image produced using kriging is a more faithful representation of the true image than an image that was produced using the cubic convolution technique.

4.4 Error Map

Kriging produces the estimation error variance which can be used to establish a confidence interval around each of the predictions. This was demonstrated in Chapter II. The variance can also be scaled to represent a pixel value, and then displayed as an error map image. This was performed using the error variance of the aerial photograph and is shown in Figure 4.24.

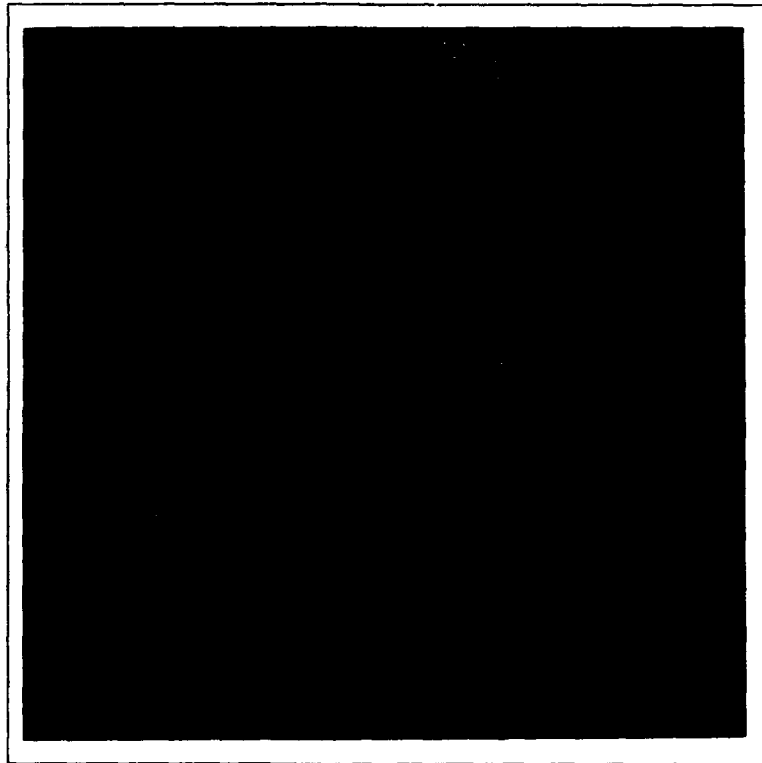


Figure 4.24. Aerial Photograph, Prediction Error Map

The aerial photograph was enlarged four times. Every fourth pixel is reproduced exactly—with zero error, and is represented by a black pixel (zero value) on the error map. The remaining error values are based on the distance from known pixels which were used in the prediction for that location.

4.5 Summary

The kriged images produced using the spherical variogram and a zero nugget prove that kriging can perform well in image enhancement by improving the resolution of satellite imagery. A quantitative comparison of the kriging and cubic convolution techniques indicates that image enhancement by universal punctual kriging can produce images which are more faithful than cubic convolution to the original image. This is indicated by a lower mean difference from the original image than the mean difference produced by subtracting a cubic convolution image from the original. This result was expected since kriging is a method of exact interpolation. That is, it retains the original pixel values.

V. Recommendations

The recommendations made in this chapter deal with further improvements of the kriged sub-images, enhancements to the kriging program, a further examination of kriging's applicability to image processing, and the production of a dedicated kriging convolution program to be designed and used strictly for image processing.

5.1 *Improving the Kriged Sub-Images*

In Chapter IV when the kriged images needed to be improved, an improved structural analysis was performed. The spherical variogram was the only model incorporated into the kriging program, so it was modified to produce the desired results. The improvement in image quality was obtained by setting the nugget effect to zero which, in effect, forced the model to begin at the origin. However, this was a first attempt at improved images, and it did not ensure that the model would fit well as the lags increased. In reality, the fit was then determined by how near the initial slope of the variogram curve was to the reduced spherical model. The optimal solution would have been to fit a better curve to the variogram at lags within the size of the kernel.

Since, for this thesis, the kernel size only included lags of approximately 11.3, there is no need to fit a curve beyond this point. Also, Davis writes that there is no significant difference between the spherical and linear models near the origin of the variogram as long as there is a sufficient sample density (9:248). Therefore, the kriged sub-images produced in this thesis can be improved by fitting a linear model to the variogram up to the first 12 (approximately) lags, and forcing the nugget to be zero. Not only will the resultant images be better, there will also be a slight savings in processing time by not curve fitting the remaining portion of the variogram curve.

5.2 *Kriging Program Enhancements*

Another possible way to improve the kriged sub-images produced in this thesis would have been to use one of the other commonly used variogram models. Therefore, in the future, the linear and De Wijsian models should be incorporated into the kriging program. This will allow the users to choose the best model based on their structural analysis and their particular application.

Lognormally distributed data should also be accommodated in the kriging program. If the data is indicated to be lognormally distributed, the program should transform it to normal, krig it, and transform it back to lognormal. This procedure may produce slight improvements in the mean difference between a kriged and a true image as calculated in Chapter IV.

5.3 Applicability of Kriging

To correctly apply kriging to image processing, an analysis of the parameters involved and their effect on the quality of the resultant image is necessary. This can be done by performing a controlled experiment and sensitivity analysis on the parameters involved in kriging. For example, the variogram of Image 04 indicated that it was extremely continuous and virtually isotropic. However, kriging it with the semi-inclusion distance of two, which was chosen for this thesis, there were some discontinuities in the resultant image. Examining the variogram for variance and continuity may suggest what an appropriate kernel size should be for existing conditions within a particular image. This implies that it may be significant to treat different images with different kernel sizes based on the type of image that it is, instead of kriging them all with a semi-inclusion distance of two.

Another application of kriging to image enhancement can be studied in the area of noise reduction. A recommendation is to eliminate noise in satellite imagery using kriging, apply image classification and forecasting to kriged images, and evaluate the results. Noise elimination using kriging should produce better results than other methods and should improve image classification and the forecasts made from the enhanced imagery.

Finally, the kriged, differenced imagery produced in Chapter IV showed the kriging error to be concentrated at the borders of the features in the image. This strongly suggests that this technique should be tested as a method of edge-detection.

5.4 Dedicated Kriging Convolution Program

The application of universal punctual kriging has been successful in improving the resolution of satellite imagery while maintaining a more faithful representation than cubic convolution of the true image. This success was performed using kriging programs and procedures which were designed to be generally applicable to a wide

variety of data types and structures. The setup and execution of the various programs associated with kriging is often intricate. To make the process more streamlined and efficient, a dedicated kriging program can be developed to operate strictly on regularly gridded digital imagery, and take advantage of the recommendations made in this chapter. In particular,

- Read in the raw data. Save time by not performing any unnecessary conversions.
- Calculate residuals.
- Calculate the variograms only for the lags that are within the kernel size to be used. Calculate the variogram based on the regular grid pattern. No searches for paired points is necessary.
- Curve fit a linear model to the truncated variograms. Ensure the curve intersects the origin. The linear model is a perfectly good approximation of the variogram at distances much less than the range, as shown by the spherical model (9:247).
- The minimization approach used in this thesis and the accompanying kriging program should be incorporated if it is desirable to keep kriging processing times low (minutes instead of hours). As presented in Chapter III, this method assumes the local drift is the same within the central region.
- The option to krig residual data sets and reincorporate the trend should be incorporated.

5.5 Summary

The success of applying universal punctual kriging to image processing indicates that the current procedures can be made available to image analysts. However, incorporating program enhancements and developing a streamlined program strictly for use on digital imagery will shorten processing times and produce further improved imagery.

Universal punctual kriging can also be applied to noise removal in digital satellite imagery. The results should be improvements in image interpretation and classification, as well as improved forecasting based on the kriged images. Finally, the methodology used in Chapter IV to quantitatively compare kriging to cubic convolution showed promise as a method of edge detection and should be explored.

Appendix A. Conversions

A.1 The Kriging Weights and the a Parameter

The kriging weights w_i are defined in Chapter II in terms of a_i for notational simplicity. Their relationship is:

$$\begin{aligned} a_p &= 1 \\ a_i &= -w_i \quad \text{for } i = 1, \dots, n \end{aligned}$$

where n is the number of known samples, and the p subscript represents the point being kriged. The relationship was derived as follows:

$$\begin{aligned} \epsilon_p &= Z_p - Z_p^* \\ &= e_p - e_p^* \quad \text{for a known constant mean} \\ &= Z_p - \mu_p - \sum_{i=1}^n w_i (Z_i - \mu_i) \\ &= Z_p - \mu_p - \sum_{i=1}^n w_i (Z_i - \mu_i) + [w_o (Z_p - \mu_p) - w_o (Z_p - \mu_p)] \\ &= Z_p - \mu_p + w_o (Z_p - \mu_p) - \sum_{i=0}^n w_i (Z_i - \mu_i) \\ &= Z_p (1 + w_o) - \mu_p (1 + w_o) - \sum_{i=0}^n w_i (Z_i - \mu_i) \\ &= (Z_p - \mu_p) (1 - a_p) + \sum_{i=0}^n a_i (Z_i - \mu_i) \end{aligned}$$

where the coefficient switch from w to a has been made such that $a = -w$. And, if a_p is defined to be one, then the equation further simplifies to:

$$\epsilon_p = \sum_{i=0}^n a_i (Z_i - \mu_i)$$

Again the substitution for Chapter II is:

$$\begin{aligned} a_p &= 1 \\ a_i &= -w_i \quad \text{for } i = 1, \dots, n \end{aligned}$$

A.2 Estimation Error Variance

When $[b]$ is written in terms of the variogram, $\gamma(h)$, the estimation error variance is calculated as:

$$\sigma_e^2 = [w]^T [b]$$

The covariance and the variogram are second-order moments, and their relationship was established in Chapter II as:

$$\gamma(\vec{h}) = \sigma(0) - \sigma(\vec{h})$$

Using this relationship, when $[b]$ is written in terms of the covariance, $\sigma(h)$, the estimation error variance is calculated as:

$$\sigma_e^2 = \sigma(0) - [w]^T [b]$$

For simplicity, consider only two sample points. The transformation can be made as follows:

$$\sigma_e^2 = [w]^T [b]$$

$$\sigma_e^2 = [w_1 \ w_2] \cdot \begin{bmatrix} \gamma_1(h) \\ \gamma_2(h) \end{bmatrix}$$

$$\sigma_e^2 = [w_1 \ w_2] \cdot \begin{bmatrix} \sigma(0) - \sigma_1(h) \\ \sigma(0) - \sigma_2(h) \end{bmatrix}$$

$$\sigma_e^2 = w_1[\sigma(0) - \sigma_1(h)] + w_2[\sigma(0) - \sigma_2(h)]$$

$$\sigma_e^2 = w_1\sigma(0) + w_2\sigma(0) - w_1\sigma_1(h) - w_2\sigma_2(h)$$

$$\sigma_e^2 = (w_1 + w_2)\sigma(0) - (w_1\sigma_1(h) + w_2\sigma_2(h))$$

and since the weights sum to one, the result becomes:

$$\sigma_e^2 = \sigma(0) - [w_1 \ w_2] \cdot \begin{bmatrix} \sigma_1(h) \\ \sigma_2(h) \end{bmatrix}$$

$$\sigma_{\epsilon}^2 = \sigma(0) - [w]^T [b]$$

when $[b]$ is in terms of the covariance, $\sigma(h)$.

Appendix B. *Original Images*

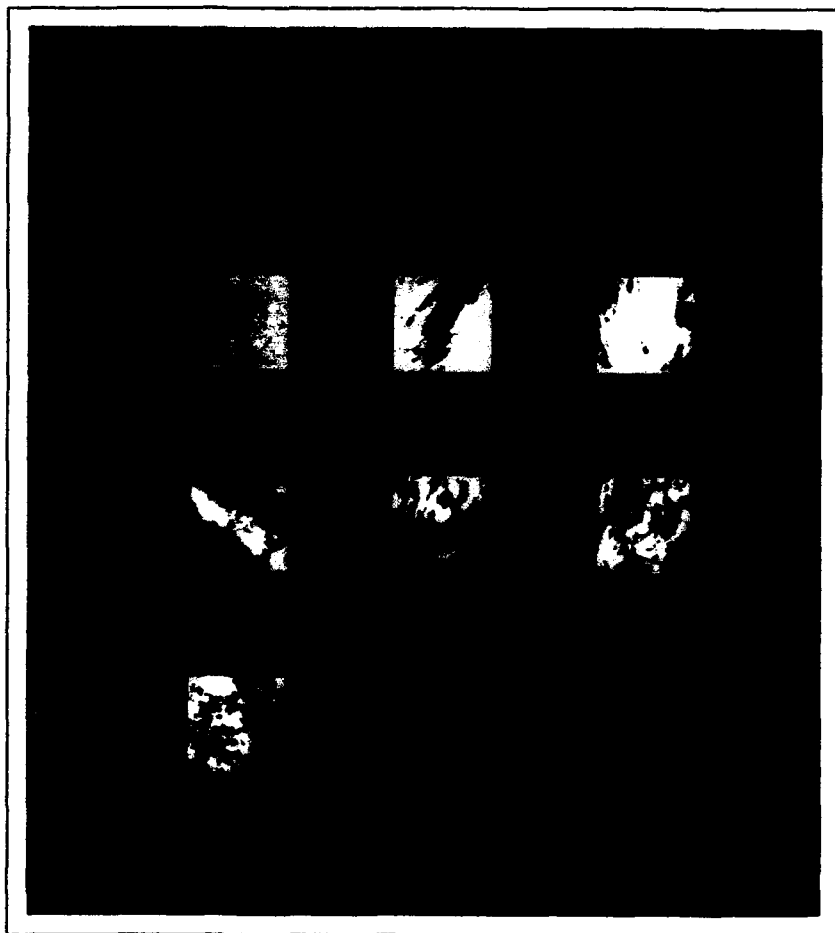


Figure B.1. The Original 100-by-100 Satellite Sub-Images.

Appendix C. Variograms of Residuals

The following variograms are calculated in the 0° and 90° directions. The diamond points represent values calculated in the 0° direction; the plus points represent values calculated at 90°. The variograms of the initial seven sub-images are shown below.

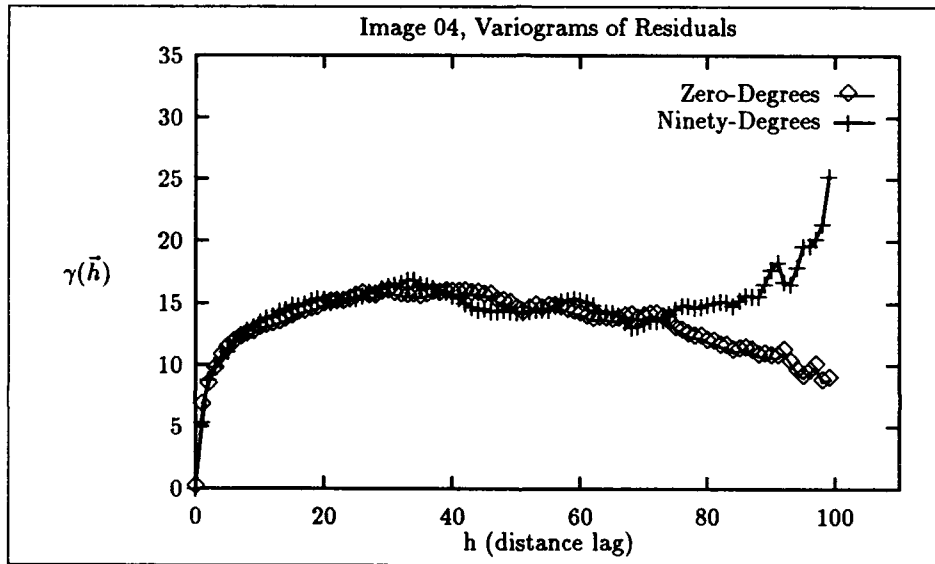


Figure C.1. Image 04, Variograms of Residuals

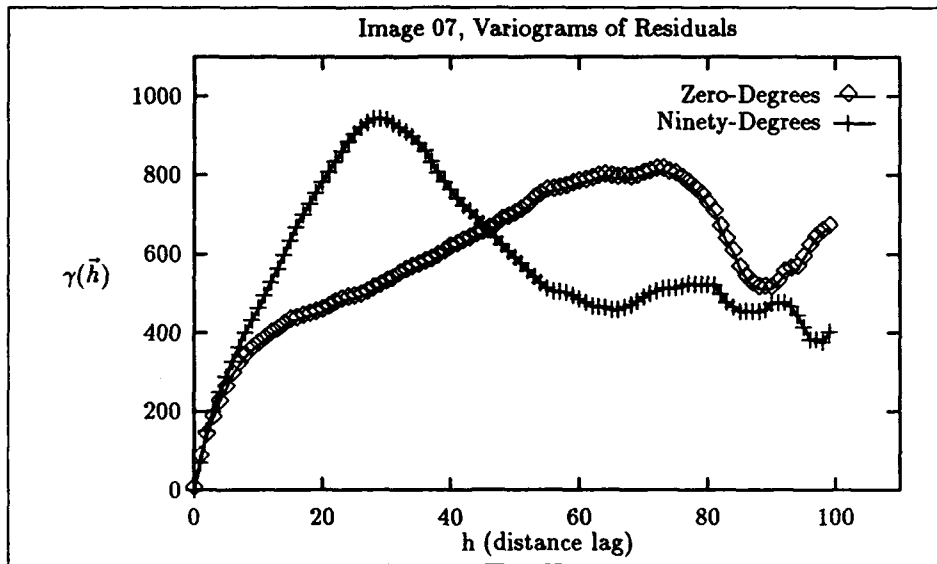


Figure C.2. Image 07, Variograms of Residuals

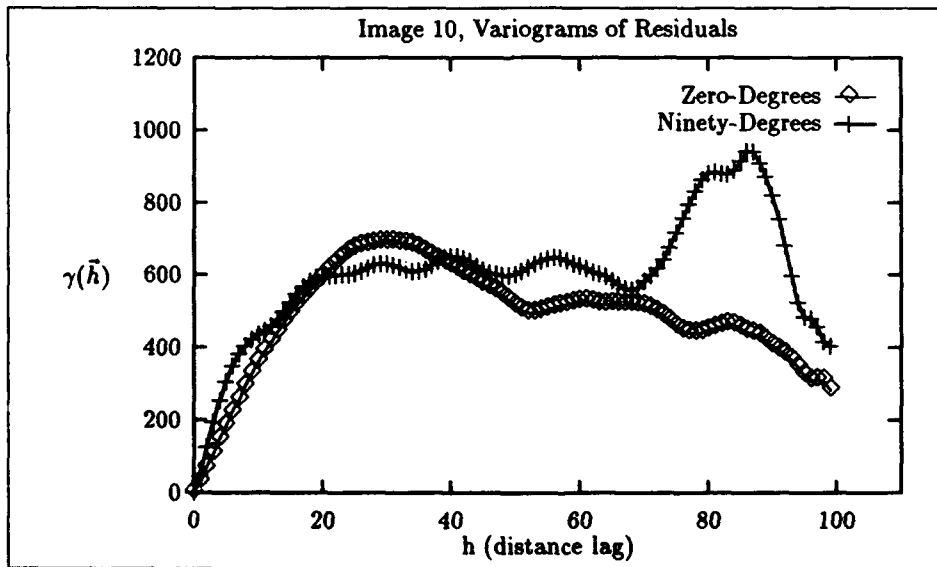


Figure C.3. Image 10, Variograms of Residuals

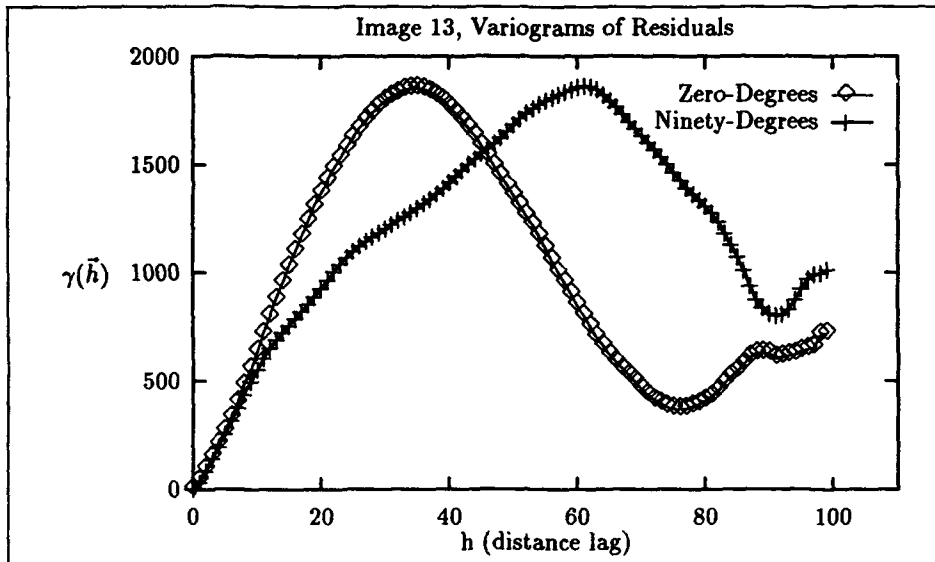


Figure C.4. Image 13, Variograms of Residuals

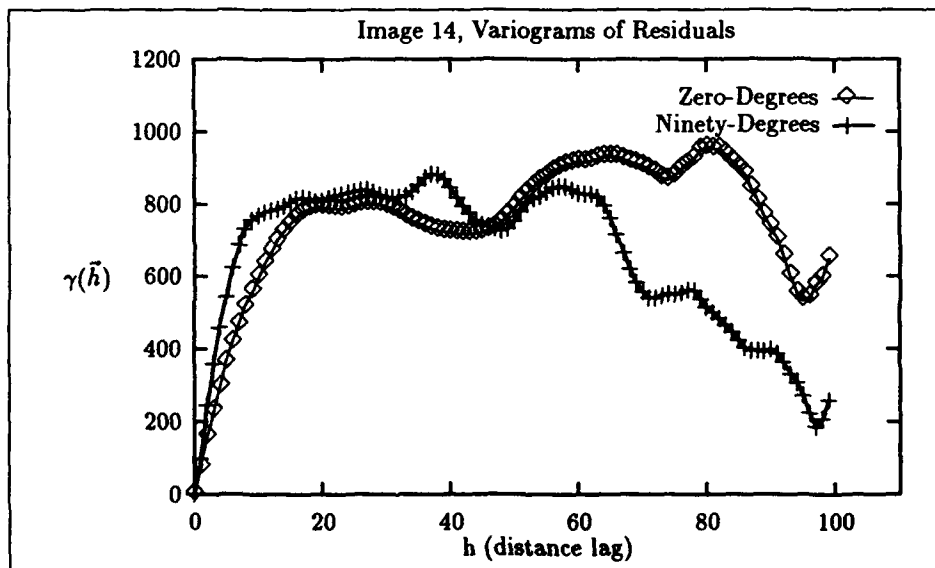


Figure C.5. Image 14, Variograms of Residuals

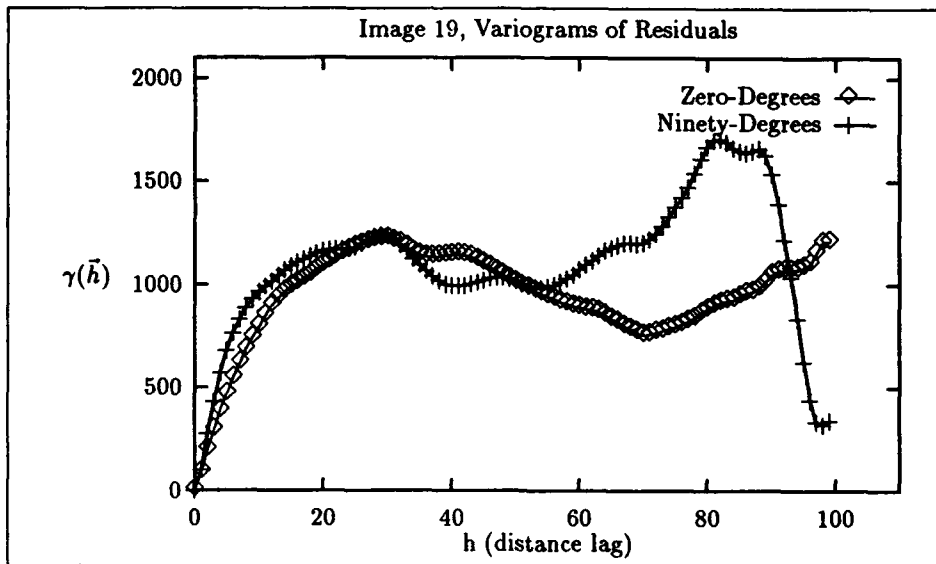


Figure C.6. Image 19, Variograms of Residuals

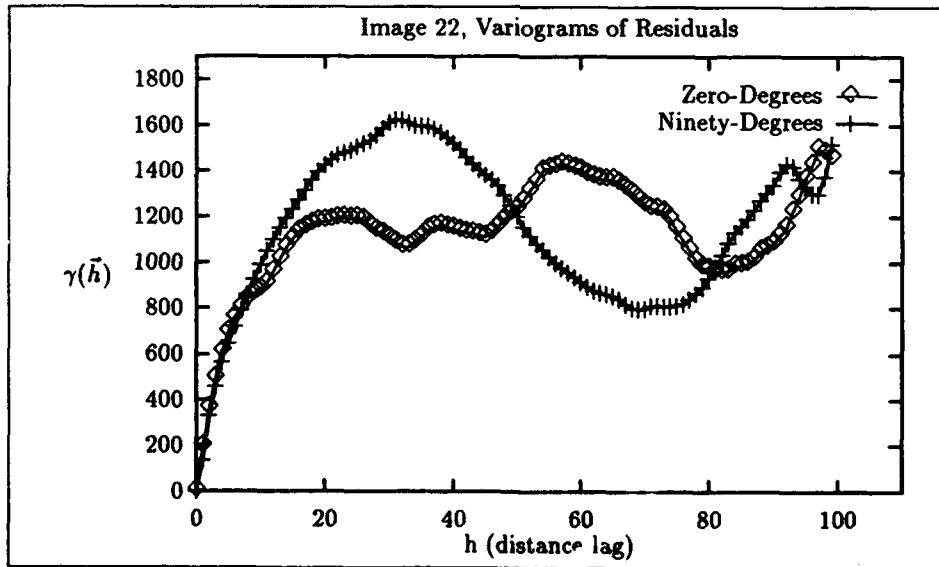


Figure C.7. Image 22, Variograms of Residuals

Appendix D. *Variogram Models*

The graphs in this appendix show the 0° and 90° variograms for each of the subject images, along with the three variogram models. The thick solid line is the linear model; the thin solid line is the spherical model; the dotted line is the De Wijsian model. Each model is displayed below its respective plot.

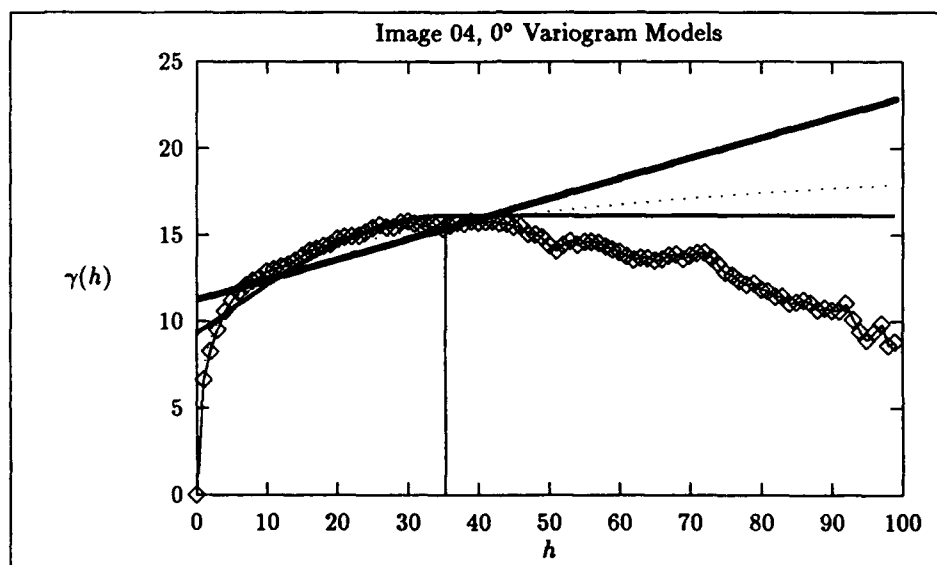


Figure D.1. Image 04, 0° Variogram Models

$$\text{Spherical model} = 6.81 \left(\frac{3h}{(2)35.334} - \frac{1}{2} \left(\frac{h}{35.334} \right)^3 \right) + 9.304$$

$$\text{Linear model} = 11.252 + 0.117h$$

$$\text{DeWijsian model} = 7.467 + 2.271 \ln h$$

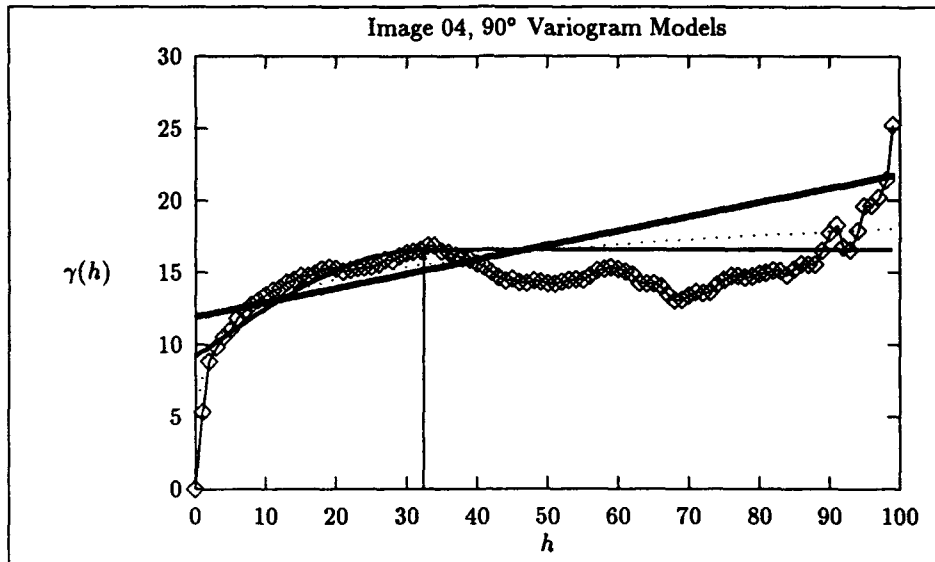


Figure D.2. Image 04, 90° Variogram Models

$$\text{Spherical model} = 7.379 \left(\frac{3h}{(2)32.407} - \frac{1}{2} \left(\frac{h}{32.407} \right)^3 \right) + 9.204$$

$$\text{Linear model} = 11.941 + 0.099h$$

$$\text{DeWijian model} = 7.906 + 2.209 \ln h$$

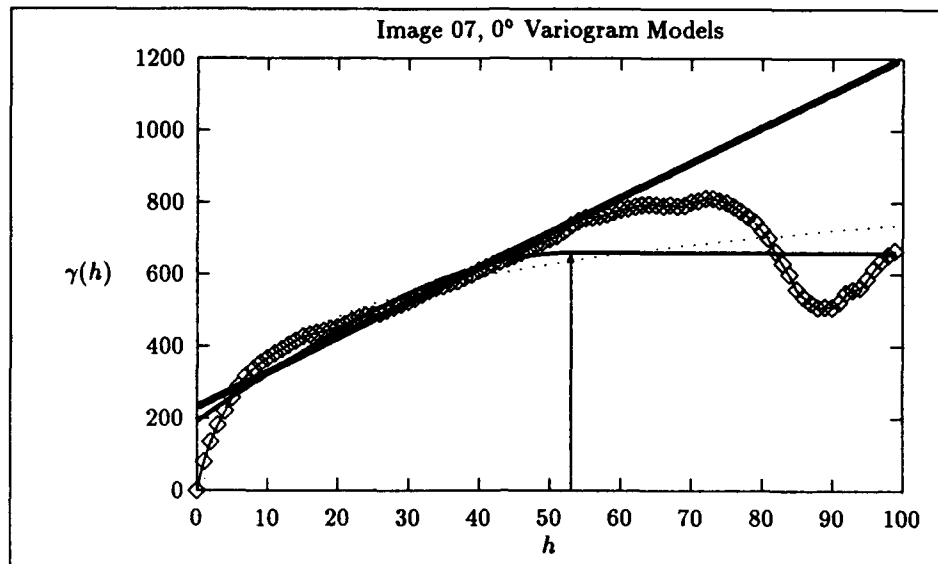


Figure D.3. Image 07, 0° Variogram Models

$$\text{Spherical model} = 466.880 \left(\frac{3h}{(2)52.958} - \frac{1}{2} \left(\frac{h}{52.958} \right)^3 \right) + 192.628$$

$$\text{Linear model} = 232.305 + 9.712h$$

$$\text{DeWijsian model} = 4.124 + 159.633 \ln h$$

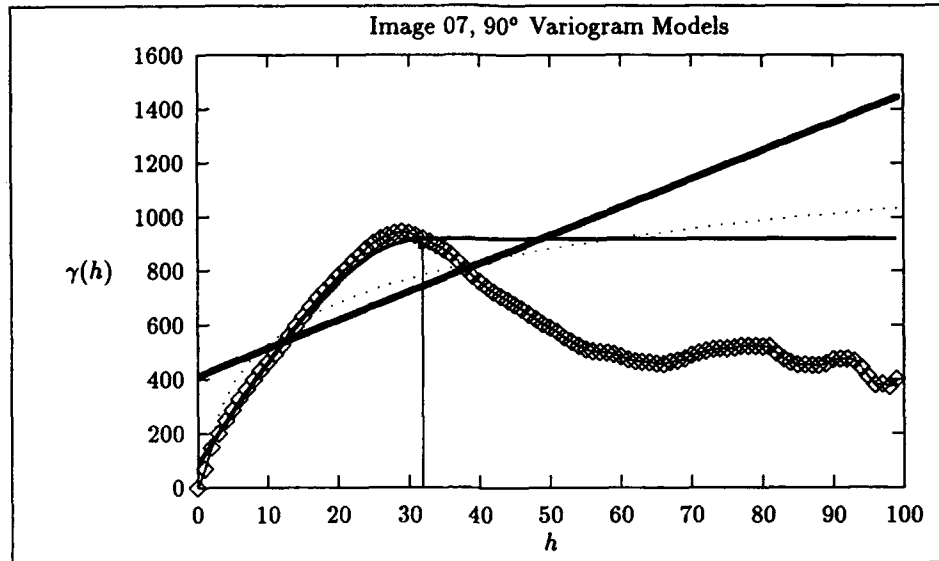


Figure D.4. Image 07, 90° Variogram Models

$$\text{Spherical model} = 826.993 \left(\frac{3h}{(2)31.914} - \frac{1}{2} \left(\frac{h}{31.914} \right)^3 \right) + 91.509$$

$$\text{Linear model} = 412.647 + 10.441h$$

$$\text{DeWijsian model} = 26.390 + 219.393 \ln h$$

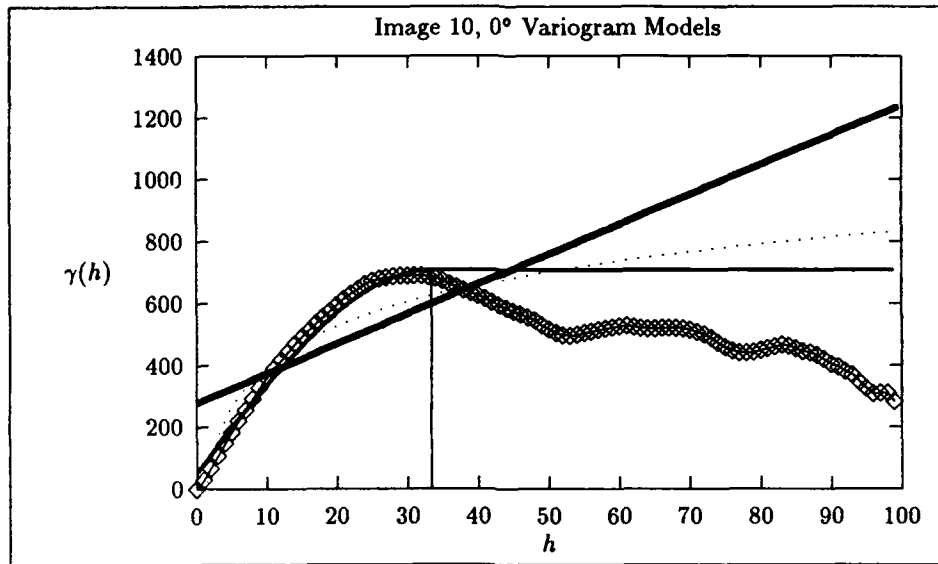


Figure D.5. Image 10, 0° Variogram Models

$$\text{Spherical model} = 651.108 \left(\frac{3h}{(2)33.332} - \frac{1}{2} \left(\frac{h}{33.332} \right)^3 \right) + 55.762$$

$$\text{Linear model} = 277.694 + 9.655h$$

$$\text{DeWijsian model} = -43.489 + 190.675 \ln h$$

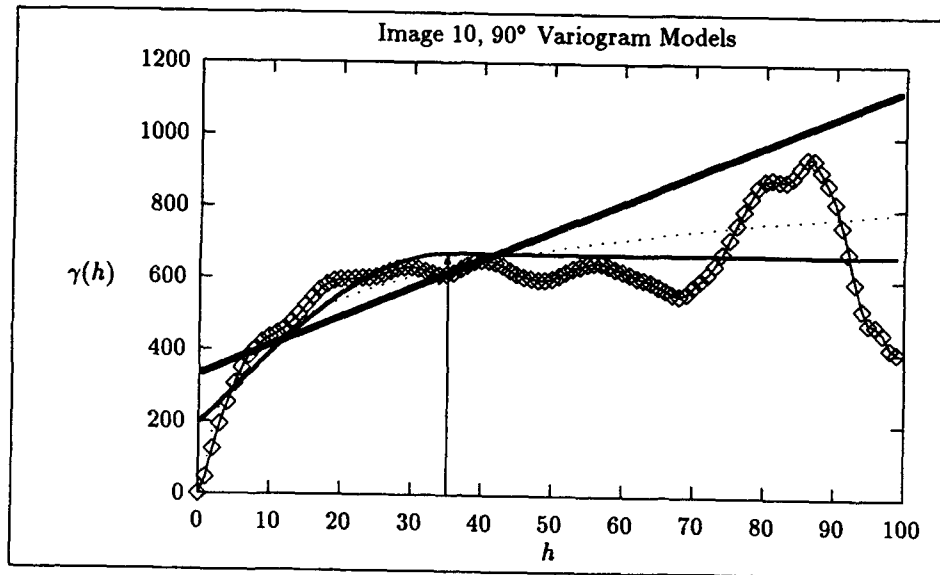


Figure D.6. Image 10, 90° Variogram Models

$$\text{Spherical model} = 473.103 \left(\frac{3h}{(2)35.144} - \frac{1}{2} \left(\frac{h}{35.144} \right)^3 \right) + 196.214$$

$$\text{Linear model} = 333.794 + 8.014h$$

$$\text{DeWijsian model} = 70.756 + 157.064 \ln h$$

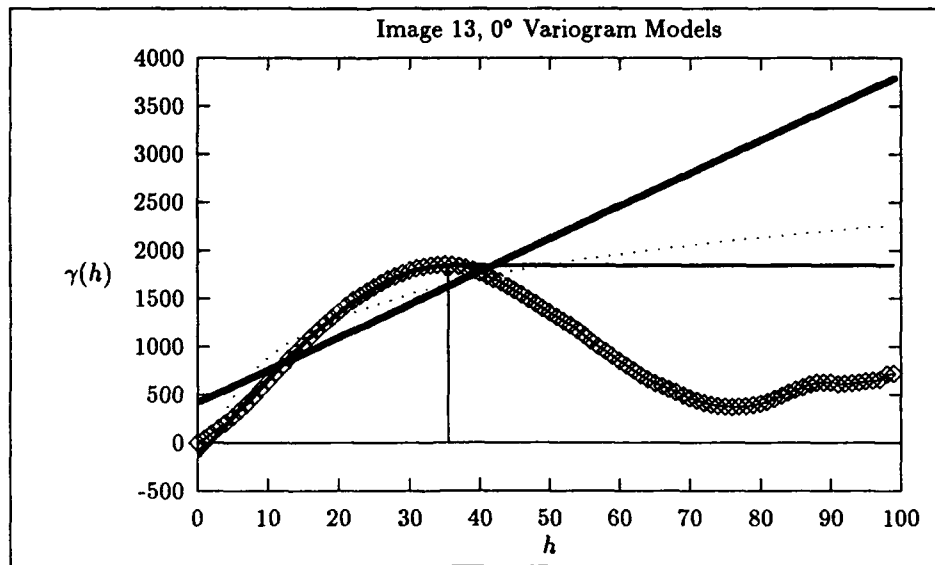


Figure D.7. Image 13, 0° Variogram Models

$$\text{Spherical model} = 1974.232 \left(\frac{3h}{(2)35.459} - \frac{1}{2} \left(\frac{h}{35.459} \right)^3 \right) + -135.012$$

$$\text{Linear model} = 423.931 + 34.035h$$

$$\text{DeWijsian model} = -495.287 + 599.964 \ln h$$

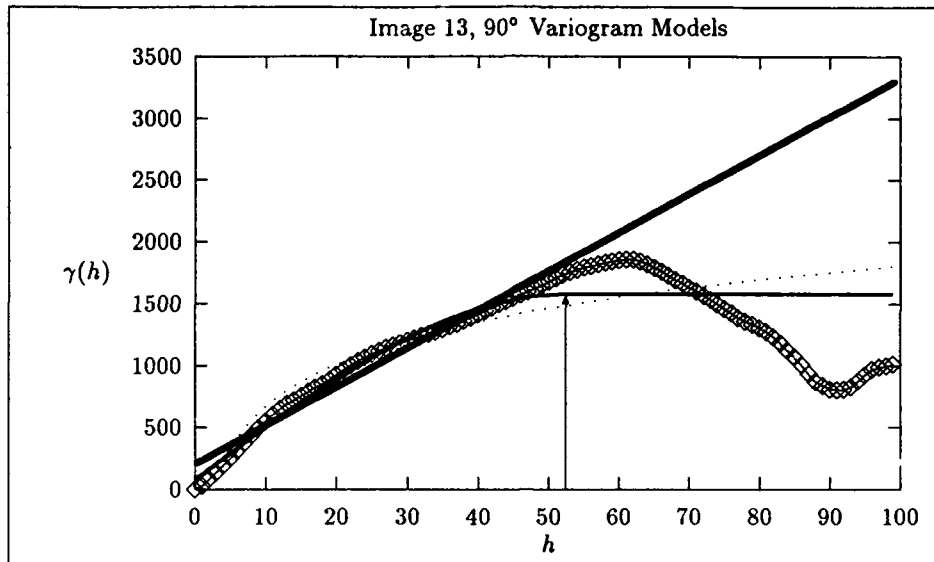


Figure D.8. Image 13, 90° Variogram Models

$$\text{Spherical model} = 1495.693 \left(\frac{3h}{(2)52.357} - \frac{1}{2} \left(\frac{h}{52.357} \right)^3 \right) + 79.885$$

$$\text{Linear model} = 211.426 + 31.206h$$

$$\text{DeWijsian model} = -463.640 + 493.241 \ln h$$

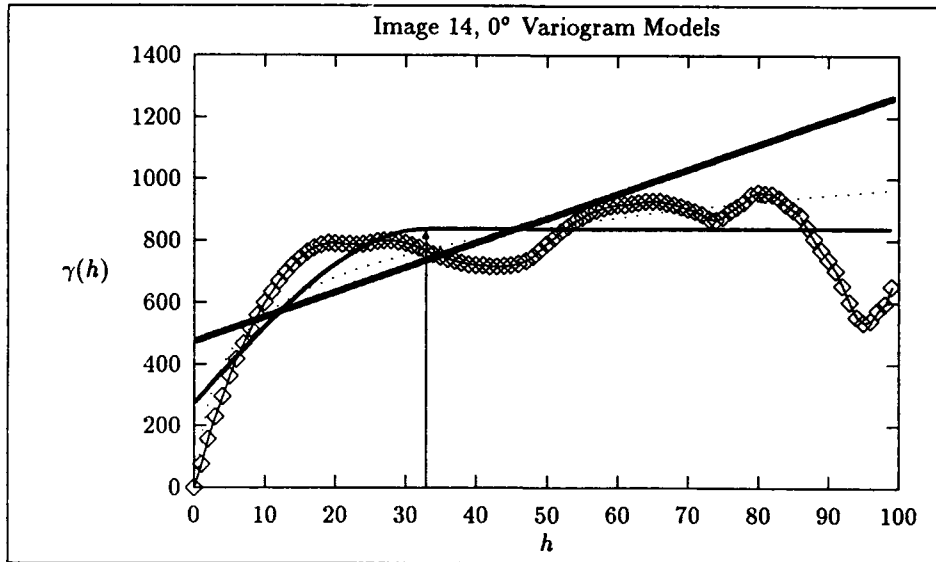


Figure D.9. Image 14, 0° Variogram Models

$$\text{Spherical model} = 559.396 \left(\frac{3h}{(2)32.911} - \frac{1}{2} \left(\frac{h}{32.911} \right)^3 \right) + 276.949$$

$$\text{Linear model} = 475.026 + 7.961h$$

$$\text{DeWijsian model} = 151.181 + 177.225 \ln h$$

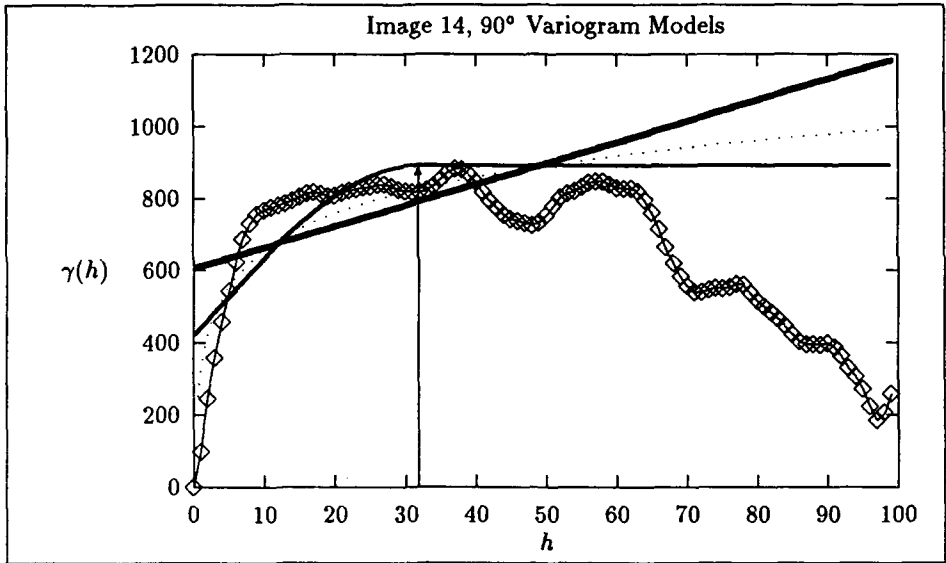


Figure D.10. Image 14, 90° Variogram Models

$$\text{Spherical model} = 470.171 \left(\frac{3h}{(2)31.793} - \frac{1}{2} \left(\frac{h}{31.793} \right)^3 \right) + 421.863$$

$$\text{Linear model} = 606.541 + 5.834h$$

$$\text{DeWijsian model} = 321.578 + 146.026 \ln h$$

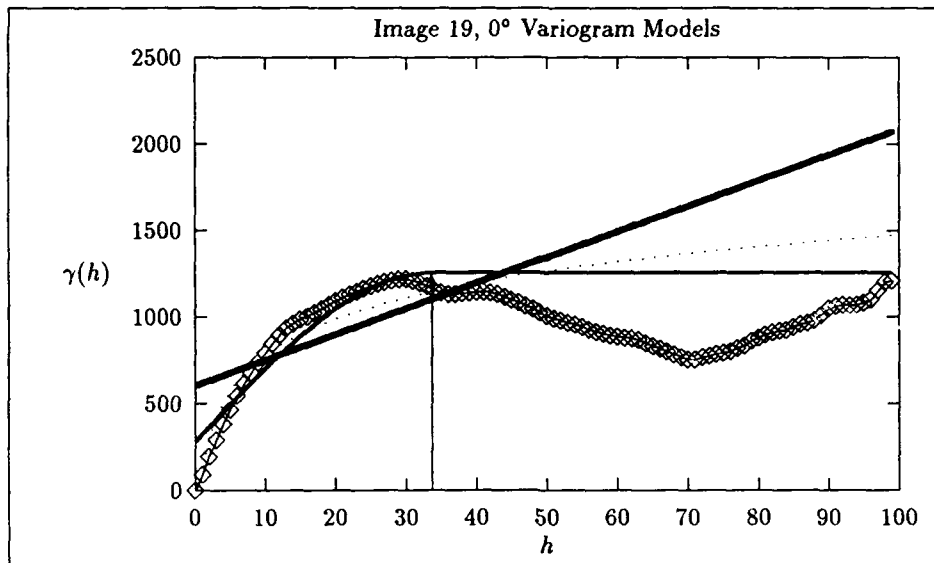


Figure D.11. Image 19, 0° Variogram Models

$$\text{Spherical model} = 969.047 \left(\frac{3h}{(2)33.674} - \frac{1}{2} \left(\frac{h}{33.674} \right)^3 \right) + 286.381$$

$$\text{Linear model} = 606.714 + 14.809h$$

$$\text{DeWijian model} = 88.910 + 300.990 \ln h$$

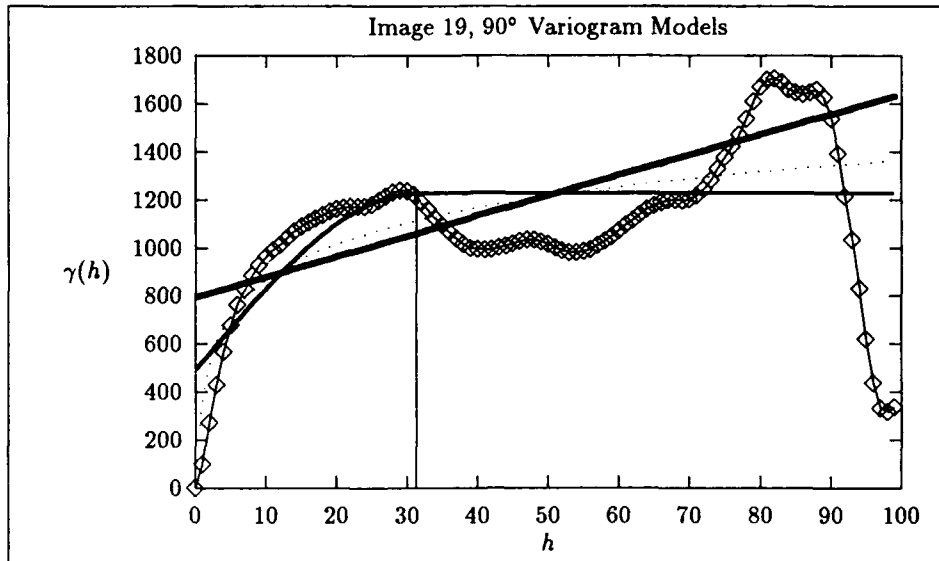


Figure D.12. Image 19, 90° Variogram Models

$$\text{Spherical model} = 730.959 \left(\frac{3h}{(2)31.351} - \frac{1}{2} \left(\frac{h}{31.351} \right)^3 \right) + 495.731$$

$$\text{Linear model} = 795.157 + 8.467h$$

$$\text{DeWijian model} = 367.829 + 216.585 \ln h$$

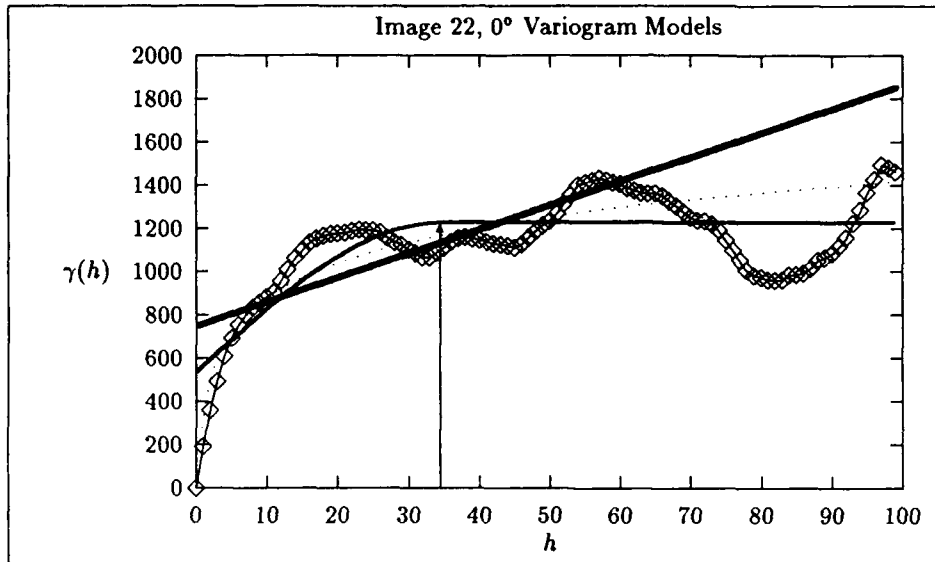


Figure D.13. Image 22, 0° Variogram Models

$$\text{Spherical model} = 687.392 \left(\frac{3h}{(2)34.452} - \frac{1}{2} \left(\frac{h}{34.452} \right)^3 \right) + 538.185$$

$$\text{Linear model} = 750.364 + 11.145h$$

$$\text{DeWijsian model} = 335.710 + 234.985 \ln h$$

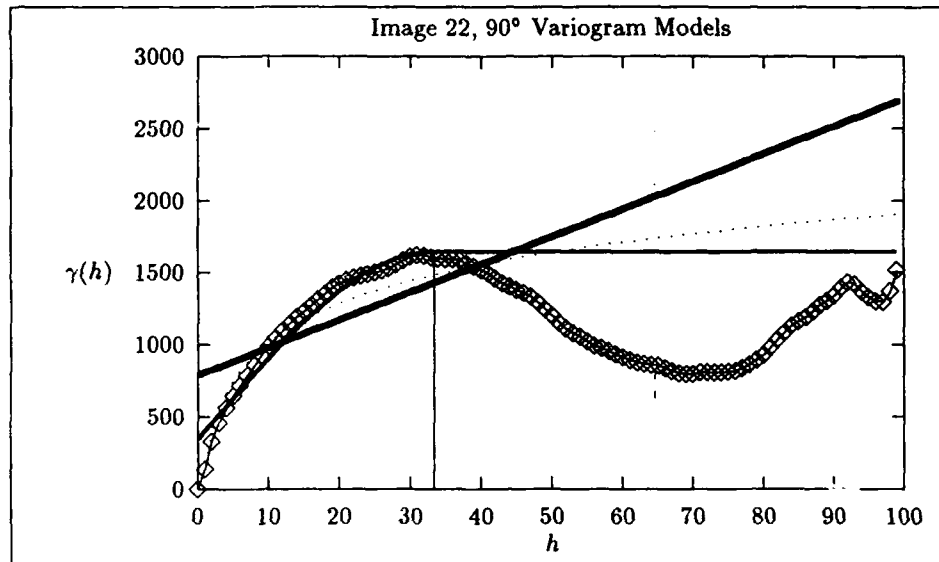


Figure D.14. Image 22, 90° Variogram Models

$$\text{Spherical model} = 1292.612 \left(\frac{3h}{(2)33.342} - \frac{1}{2} \left(\frac{h}{33.342} \right)^3 \right) + 352.069$$

$$\text{Linear model} = 792.268 + 19.185h$$

$$\text{DeWijsian model} = 135.067 + 385.319 \ln h$$

Appendix E. *Typical Variogram Output*

This is a typical output from the variogram program. The mean elevation indicates the effectiveness of the global trend removal program. The variogram models are included at the end of the output.

sum: -163.304535 sum2: 11252960.000000

mean: -0.016330 var: 1125.295776

Direction number: 1

Direction angle (degrees): 90

Step size: 1.00

Overall mean elevation: -0.02

variance: 1125.30

Field of 1.00 degrees in each direction (regularization factor)

Anisotropic correction factor: 1.00

Distance		Num of Samples	Drift	Average gamma(h)	Distance(h)
0.0 -	1.0	0	0.00	0.00	0.00
1.0 -	2.0	9900	0.14	30.96	1.00
2.0 -	3.0	9800	0.29	84.05	2.00
3.0 -	4.0	9700	0.44	137.02	3.00
4.0 -	5.0	9600	0.60	193.03	4.00
5.0 -	6.0	9500	0.75	251.64	5.00
6.0 -	7.0	9400	0.90	312.83	6.00
7.0 -	8.0	9300	1.03	375.02	7.00
8.0 -	9.0	9200	1.14	436.31	8.00
9.0 -	10.0	9100	1.18	495.70	9.00
10.0 -	11.0	9000	1.19	552.05	10.00
11.0 -	12.0	8900	1.18	603.48	11.00
12.0 -	13.0	8800	1.16	650.34	12.00
13.0 -	14.0	8700	1.10	689.21	13.00

14.0 -	15.0	8600	1.00	722.37	14.00
15.0 -	16.0	8500	0.86	754.98	15.00
16.0 -	17.0	8400	0.71	787.65	16.00
17.0 -	18.0	8300	0.52	821.05	17.00
18.0 -	19.0	8200	0.32	854.07	18.00
19.0 -	20.0	8100	0.11	888.48	19.00
20.0 -	21.0	8000	-0.09	925.78	20.00
21.0 -	22.0	7900	-0.32	964.69	21.00
22.0 -	23.0	7800	-0.56	1002.81	22.00
23.0 -	24.0	7700	-0.76	1037.39	23.00
24.0 -	25.0	7600	-0.93	1067.45	24.00
25.0 -	26.0	7500	-1.10	1095.51	25.00
26.0 -	27.0	7400	-1.24	1120.07	26.00
27.0 -	28.0	7300	-1.33	1141.77	27.00
28.0 -	29.0	7200	-1.41	1160.96	28.00
29.0 -	30.0	7100	-1.49	1180.48	29.00
30.0 -	31.0	7000	-1.54	1201.30	30.00
31.0 -	32.0	6900	-1.57	1220.47	31.00
32.0 -	33.0	6800	-1.56	1239.77	32.00
33.0 -	34.0	6700	-1.53	1256.96	33.00
34.0 -	35.0	6600	-1.47	1276.07	34.00
35.0 -	36.0	6500	-1.38	1296.92	35.00
36.0 -	37.0	6400	-1.26	1317.77	36.00
37.0 -	38.0	6300	-1.17	1338.46	37.00
38.0 -	39.0	6200	-1.06	1362.26	38.00
39.0 -	40.0	6100	-0.87	1388.69	39.00
40.0 -	41.0	6000	-0.64	1417.23	40.00
41.0 -	42.0	5900	-0.45	1446.41	41.00
42.0 -	43.0	5800	-0.28	1473.41	42.00
43.0 -	44.0	5700	-0.10	1500.55	43.00
44.0 -	45.0	5600	0.08	1526.87	44.00
45.0 -	46.0	5500	0.24	1550.55	45.00
46.0 -	47.0	5400	0.37	1574.59	46.00
47.0 -	48.0	5300	0.47	1599.08	47.00
48.0 -	49.0	5200	0.56	1625.86	48.00

49.0 -	50.0	5100	0.61	1655.38	49.00
50.0 -	51.0	5000	0.63	1685.54	50.00
51.0 -	52.0	4900	0.63	1714.13	51.00
52.0 -	53.0	4800	0.61	1740.25	52.00
53.0 -	54.0	4700	0.54	1762.59	53.00
54.0 -	55.0	4600	0.43	1781.13	54.00
55.0 -	56.0	4500	0.30	1797.77	55.00
56.0 -	57.0	4400	0.10	1811.55	56.00
57.0 -	58.0	4300	-0.14	1824.84	57.00
58.0 -	59.0	4200	-0.39	1836.95	58.00
59.0 -	60.0	4100	-0.65	1847.94	59.00
60.0 -	61.0	4000	-0.96	1858.92	60.00
61.0 -	62.0	3900	-1.37	1863.19	61.00
62.0 -	63.0	3800	-1.74	1859.64	62.00
63.0 -	64.0	3700	-2.00	1845.82	63.00
64.0 -	65.0	3600	-2.24	1819.52	64.00
65.0 -	66.0	3500	-2.56	1790.91	65.00
66.0 -	67.0	3400	-2.86	1761.94	66.00
67.0 -	68.0	3300	-3.11	1728.89	67.00
68.0 -	69.0	3200	-3.31	1693.11	68.00
69.0 -	70.0	3100	-3.50	1661.53	69.00
70.0 -	71.0	3000	-3.60	1628.16	70.00
71.0 -	72.0	2900	-3.64	1598.05	71.00
72.0 -	73.0	2800	-3.62	1564.38	72.00
73.0 -	74.0	2700	-3.60	1530.19	73.00
74.0 -	75.0	2600	-3.53	1496.11	74.00
75.0 -	76.0	2500	-3.31	1458.96	75.00
76.0 -	77.0	2400	-2.95	1422.90	76.00
77.0 -	78.0	2300	-2.55	1389.08	77.00
78.0 -	79.0	2200	-1.98	1361.14	78.00
79.0 -	80.0	2100	-1.22	1335.96	79.00
80.0 -	81.0	2000	-0.37	1309.81	80.00
81.0 -	82.0	1900	0.49	1279.86	81.00
82.0 -	83.0	1800	1.45	1239.73	82.00
83.0 -	84.0	1700	2.56	1190.46	83.00

84.0 - 85.0	1600	3.72	1137.66	84.00
85.0 - 86.0	1500	4.88	1083.83	85.00
86.0 - 87.0	1400	6.14	1025.55	86.00
87.0 - 88.0	1300	7.39	955.25	87.00
88.0 - 89.0	1200	8.48	893.91	88.00
89.0 - 90.0	1100	9.52	859.20	89.00
90.0 - 91.0	1000	10.72	834.34	90.00
91.0 - 92.0	900	11.95	828.79	91.00
92.0 - 93.0	800	13.10	833.69	92.00
93.0 - 94.0	700	13.75	856.19	93.00
94.0 - 95.0	600	14.11	900.57	94.00
95.0 - 96.0	500	14.27	948.17	95.00
96.0 - 97.0	400	14.31	992.90	96.00
97.0 - 98.0	300	14.19	1016.76	97.00
98.0 - 99.0	200	14.32	1024.25	98.00
99.0 - 100.0	100	13.70	1041.81	99.00

You may have to increase the value of MaxLag

Number of points used in linear model fit: 49

Linear Model:

$$\text{gamma}(h) = 210.553 + 31.256 h$$

$$\text{Simple Correlation: } 0.963$$

Number of points used in DeWijsian model fit: 49

DeWijsian Model:

$$\text{gamma}(h) = -464.823 + 493.769 \ln(h)$$

$$\text{Simple Correlation: } 0.927$$

Number of points used in spherical model fit: 49

Spherical Model:

$$C0 = 80.457$$

$$C = 1499.873$$

$$\text{(range) } A = 52.599$$

$$\text{(sill) } C0+C = 1580.330$$

$$\gamma(h) = 1499.873 \left[\frac{3h}{2 \cdot 52.599} - \frac{1}{2} \left(\frac{h}{52.599} \right)^3 \right] + 80.457$$

Remember to incorporate k if appropriate!

Simple Correlation: 0.990

NOTE:

Simple correlation should be close to +/- 1 for good model fit but ---
please be cautious of negative slopes for linear and ln models and
spherical cases where $C < C_0$!

Bibliography

1. Barnes, James C. *TIROS-N Series Direct Readout Services User's Guide*. Washington: U.S. Department of Commerce, NOAA/NESDIS, March 1982.
2. Brockwell, Peter J. *Time Series: Theory and Methods* (Second Edition). New York: Springer-Verlag, 1991.
3. Brodtkin, Chris, "The Application of Kriging for Controlled Minimization of Large Data Sets," December 1991. Draft MS Thesis, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH.
4. Clark, Isobel. *Practical Geostatistics*. London: Elsevier Applied Science, 1987.
5. Cressie, Noel. "Spatial Prediction and Ordinary Kriging," *Mathematical Geology*, 20(4):405-421 (1988).
6. Cressie, Noel. "Geostatistics," *The American Statistician*, 43(4):197-202 (November 1989).
7. Cressie, Noel. "The Origins of Kriging," *Mathematical Geology*, 22(3):239-252 (1990).
8. David, Michel. *Geostatistical Ore Reserve Estimation*. New York: Elsevier Scientific Publishing Company, 1977.
9. Davis, John C. *Statistics and Data Analysis in Geology* (Second Edition). New York: John Wiley and Sons, 1986.
10. Davis, M. W. and P. G. Culhane. "Contouring Very Large Datasets Using Kriging." *Geostatistics for Natural Resources Characterization, Part 2* edited by G. et al. Verly, 599-619, D. Reidel Publishing Company, 1984.
11. Dubrule, Olivier. "Comparing Splines and Kriging," *Computers and Geosciences*, 10(2-3):327-338 (1984).
12. Duckett, Donald P., "The Application of Statistical Estimation Techniques to Terrain Modeling," December 1991. Draft MS Thesis, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH.
13. Foley, James D., et. al. *Computer Graphics: Principles and Practice* (Second Edition). Addison-Wesley Publishing Company, 1990.
14. Gonzalez, Rafael C. and Paul Wintz. *Digital Image Processing* (Second Edition). Addison-Wesley Publishing Company Inc., November 1987.
15. Grant, Michael. *The Application of Kriging in the Statistical Analysis of Anthropometric Data, Vol 1*. MS thesis, AFIT/GOR/ENY/ENS/90M-8. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 1990, (AD-A220 613).

16. Hubert, Lester F. and Paul E. Lehr. *Weather Satellites*. Waltham MA: Blaisdell Publishing Company, 1967.
17. Journel, Andre G., "Fundamentals of Geostatistics in Five Lessons." Short Course Presented at the 28th International Geological Congress, Washington, D.C., 1989.
18. Journel, Andre G. and M. E. Rossi. "When Do We Need a Trend Model in Kriging," *Mathematical Geology*, 21(7):715-739 (1989).
19. Kerbs, Lynda. "GEO-Statistics: The Variogram," *COGS Computer Contributions*, 12(2):54-59 (August 1966).
20. Larcomb, Charles H. *Spatial Registration of TIROS-N Weather Satellite Data*. MS thesis, AFIT/GSO/ENS/89D-10. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, December 1989, (AD-A216 041).
21. Lillesand, Thomas M. and Ralph W. Kiefer. *Remote Sensing and Image Interpretation* (Second Edition). New York: John Wiley and Sons, 1987.
22. Matheron, G. "Principles of Geostatistics," *Economic Geology*, 58:1246-1266 (1963).
23. Mendenhall, William, et al. *Mathematical Statistics with Applications* (Fourth Edition). Boston: PWS-KENT Publishing Company, 1990.
24. Robinson, David G., "Kriging Lectures," 1991.

Vita

Captain Donald Wayne McGee, Jr. was born in Edinburgh, Scotland on February 14, 1963. He graduated with honors from Cabot High School in Cabot, Arkansas in 1981. In 1985, he graduated from Valdosta State College in Valdosta, Georgia with a Bachelor of Science degree in Physics. He was an ROTC Distinguished Graduate, received a regular commission upon graduation, and entered the Air Force Space Command, where he trained for 15 months to become a Satellite Operations Officer. In 1987 Captain McGee was received into the Outstanding Young Men of America. Later that year, he married the former Margaret Lynne Ryan of Moultrie, Georgia. After five years of operating Defense Satellite Communications System III (DSCS III) satellites in the Air Force's Mission Control Center-2, Captain McGee attended Squadron Officers' School in residence where he was a Distinguished Graduate. In May 1990, Captain McGee entered the School of Engineering, Air Force Institute of Technology. Megan Elizabeth McGee was born to Wayne and Margie on July 15, 1991 at Wright-Patterson AFB Medical Center.

Permanent address: 1416 4th Avenue N.E.
Moultrie, Georgia 31768

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December 1991	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE THE APPLICATION OF STATISTICAL KRIGING TO IMPROVE SATELLITE IMAGERY RESOLUTION			5. FUNDING NUMBERS	
6. AUTHOR(S) Donald W. McGee, Jr., Captain, USAF				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology, WPAFB OH 45433-6583			8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GSO/ENS/ENY/91D-13	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This thesis applies the statistical technique of universal punctual kriging to improving the resolution of satellite imagery. Kriging is used to optimally predict pixel values between the sampled pixels. Nearest-neighbor, bilinear interpolation, and cubic convolution are the currently accepted interpolation methods which were applied and compared to kriging. Kriging produced visually interpretable images which were as good as those produced by bilinear interpolation and were often as good as those produced by cubic convolution. Kriging has the advantage of exactly reproducing the sampled pixels. Bilinear interpolation and cubic convolution lack this ability, and nearest-neighbor cannot produce interpretable images of the kriging quality. A quantitative comparison of kriging and cubic convolution indicates that resolution improvement using kriging can produce images which are more faithful than cubic convolution to the original image. The application of kriging to edge-detection and noise reduction should be examined.				
14. SUBJECT TERMS anisotropy, convolution, correlation, covariance, image processing, interpolation, kriging, resolution, stationary processes, statistical functions, statistics, structural analysis, variogram			15. NUMBER OF PAGES 121	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	