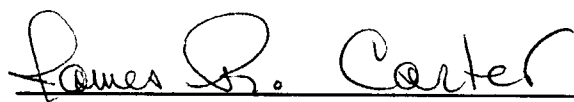


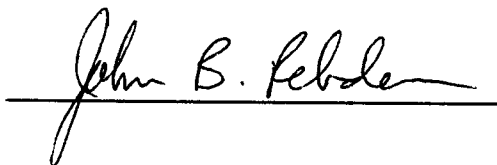
To the Graduate Council:

I am submitting herewith a thesis written by Clayton A. Crawford entitled "Evaluating the Quality of Arc/Info TINs Derived from USGS DEMs." I have examined the final copy of this thesis for form and content and recommend that it is accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Geography.


James R. Carter, Major Professor

We have read this thesis
and recommend its
acceptance:





Accepted for the Council:



Vice Provost
and Dean of the Graduate School

EVALUATING THE QUALITY OF ARC/INFO TINs
DERIVED FROM USGS DEMs

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Clayton A. Crawford

May 1990

ACKNOWLEDGEMENTS

I would like to thank Dr. John Rehder and Dr. Carol Harden for their helpful comments and assistance during the completion of this research. The support provided by Dr. Stephen Nodvin and the Southern Appalachian Man and the Biosphere Program is appreciated. In particular, I would like to thank Dr. James Carter who inspired this research, and who dedicated a great deal of time and effort to make it possible.

Dedicated to my wife,
Jane,
and my parents, David and Jane Crawford

ABSTRACT

There are a different forms of digital terrain models used to represent topography. The conversion of one model into another is often desirable and appropriate. However, the conversion process may affect the quality of the data in its newer format and thus the ability to represent topography. In this paper a methodology is developed to assess the effects of the transformation of gridded elevation data into triangulated irregular networks.

In particular, the quality of Arc/Info based triangulated irregular networks derived from USGS gridded digital elevation models is discussed. Quality is based on comparing a number of descriptive geomorphometric parameters between the original gridded models and the resulting triangulated surfaces created at different levels of generalization. Findings suggest that to maintain reliable definition of surface characteristics, samplings of at least five to fifteen percent of the critical points from the gridded elevation models will be required, depending on the nature of the land surface.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
On the Digital Representations of Terrain	2
Data Conversion	4
Purpose of Study	7
II. THE GRIDDED DIGITAL ELEVATION MODEL	9
Characteristics	10
Historical Developments	13
Sources of Data	15
Characteristics of USGS Gridded DEMs	17
Summary of Gridded DEMs	23
III. THE TRIANGULATED IRREGULAR NETWORK	24
Characteristics	25
Historical Developments	27
Sources of Data Used in Triangulation	31
Methods of Triangulating a Random Point Set	33
Summary of the TIN	37
IV. METHODOLOGY	38
Conversion of Gridded DEMs to TINS	38
Conversion of TINs to Raster Format	42
General Description of Evaluation Procedures ...	45
Specific Methodology	51

CHAPTER	PAGE
V. FINDINGS	55
Description of Original Datasets	56
Similarity Between Original and Interpolated DEMs	67
Departures Between Original and Interpolated DEMs	75
Comparison of Surface Structure	88
Summary of Findings	96
VI. CONCLUSION	101
BIBLIOGRAPHY	106
VITA	113

TABLE OF FIGURES

FIGURE		PAGE
1.1	Four Sampling Schemes for Digital Elevation Data	5
2.1	Structure of a 7.5-minute Digital Elevation Model	20
2.2	Structure of a 1-degree Digital Elevation Model	22
3.1	Delaunay Triangulation	30
3.2	Rotational Search of Points for Triangulation ..	36
4.1	Difference Between Actual and Predicted Elevations	41
4.2	Grid Superimposed on a TIN Surface	44
5.1	Hypsometric Curve - Site 250LR	58
5.2	Hypsometric Curve - Site 250HR	59
5.3	Hypsometric Curve - Site 24LR	60
5.4	Hypsometric Curve - Site 24HR	61
5.5	Contour Plot - Site 250LR (25 meter interval) ..	63
5.6	Contour Plot - Site 250HR (100 meter interval) .	64
5.7	Contour Plot - Site 24LR (25 meter interval) ...	65
5.8	Contour Plot - Site 24HR (100 meter interval) ..	66
5.9	Mean Elevation - 1:250,000 Scale Study Sites ...	68
5.10	Standard Deviation of Elevation - 1:250,000 Scale Study Sites	69
5.11	Mean Elevation - 1:24,000 Scale Study Sites	70

FIGURE	PAGE
5.12 Standard Deviation of Elevation - 1:24,000 Scale Study Sites	71
5.13 Correlation Between Original and Interpolated DEMs 1:250,000 Scale Study Sites	73
5.14 Correlation Between Original and Interpolated DEMs 1:24,000 Scale Study Sites	74
5.15 Maximum Absolute Departure - 1:250,000 Scale Study Sites	76
5.16 Maximum Absolute Departure - 1:24,000 Scale Study Sites	77
5.17 Mean Absolute Departure - 1:250,000 Scale Study Sites	78
5.18 Mean Absolute Departure - 1:24,000 Scale Study Sites	79
5.19 Standard Deviation of Departure - 1:250,000 Scale Study Sites	80
5.20 Standard Deviation of Departure - 1:24,000 Scale Study Sites	81
5.21 Distribution of Departure Between Original and Interpolated DEMs - Site 250LR	84
5.22 Distribution of Departure Between Original and Interpolated DEMs - Site 250HR	85
5.23 Distribution of Departure Between Original and Interpolated DEMs - Site 24LR	86
5.24 Distribution of Departure Between Original and Interpolated DEMs - Site 24HR	87
5.25 Ridge-Form of Original DEMs	90
5.26 TIN Derived Ridge-Form of Site 250LR	91
5.27 TIN Derived Ridge-Form of Site 250HR	92

FIGURE	PAGE
5.28 TIN Derived Ridge-Form of Site 24LR	93
5.29 TIN Derived Ridge-Form of Site 24HR	94

CHAPTER I

INTRODUCTION

Terrain mapping and analysis has evolved into a complex and dynamic field of study. Increased interest in this field has developed in response to the need for an understanding of terrain by a variety of disciplines concerned with its practical uses.¹ Geographers, hydrologists, ecologists, and other scientists in the academic community have found the need to become involved in terrain evaluation as part of their research. Other professionals in applied fields such as in civil engineering and the military have also experienced the need to develop new techniques and enhance their expertise in this area.

'Terrain' has been defined as the physical features of a tract of land.² This may include both man-made and natural features with emphasis on surface form. Typically, terrain has been portrayed with the use of topographic maps. Isolines, commonly referred to on these maps as contour lines, are used to represent the form of the surface.

¹Houston, J. M. (1973), Terrain Evaluation, London : Longman Group Limited, pp. 4.

²G. & C. Merriam Company (1973), Webster's New Collegiate Dictionary, Springfield, MA : G. & C. Merriam Company, pp 1195.

Analysts can gather information from topographic maps with the use of manual techniques but recent advances in computer technology have made the digital representation and analysis of terrain feasible. With the wide-spread adoption of geographic information systems there is an increasing use of the digital representation and analysis of terrain.

1.1 On the Digital Representations of Terrain

The topographic surface can be considered to be a mathematically continuous surface. No matter how small an area, an infinite number of points may be selected to represent the topography. In practice, only a sample of observations are utilized to represent a surface. Observations are sampled in the field, calculated photogrammetrically, or taken from existing topographic maps. Digital elevation models (DEMs) may be created from samples of the topographic surface. A DEM is a numeric representation of a topographic surface where the primary data element is elevation. Most of the parameters needed in terrain analysis can be derived directly from such models of elevation.

A DEM should contain adequate elevation and planimetric measurements compatible in number and distribution with the terrain being modeled, so that the elevation of any location

can be interpolated accurately for any given application.³ This is needed to obtain a 'satisfactory' representation of the surface. Emphasis is therefore placed on data capture and storage techniques. Many sampling schemes are used to produce DEMs. Common forms of data capture, or storage, are with sampling at regularly spaced intervals, in parallel profiles, as digitized contours, or as triangulated irregular networks.

Data captured at regular intervals is referred to as gridded data. This is because points sampled at equal distances along profiles and equal distances between profiles produce a rectangular gridded matrix. Data captured or stored as parallel profiles is regularly spaced between profiles and irregularly spaced within them. Digitized contours are the digital counterpart of the isolines found on a topographic map where points are used to define the contour lines. An increasingly common form of data structure is the triangulated irregular network (TIN). This model represents a surface with a set of non-overlapping triangle facets formed by the connection of

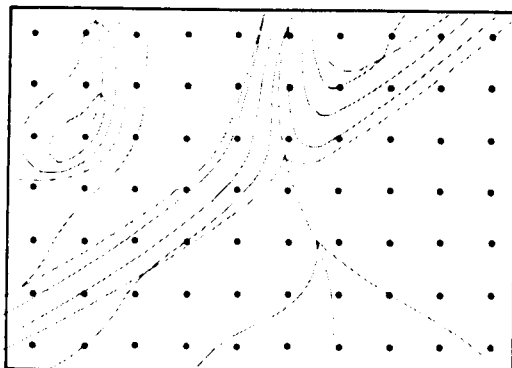
³Ayeni, O. O. (1982), "Optimum Sampling for Digital Terrain Models: A Trend Towards Automation", Photogrammetric Engineering and Remote Sensing, Vol. 48, No. 11, pp. 1687.

neighboring, irregularly distributed points. Figure 1.1 displays these sampling schemes.

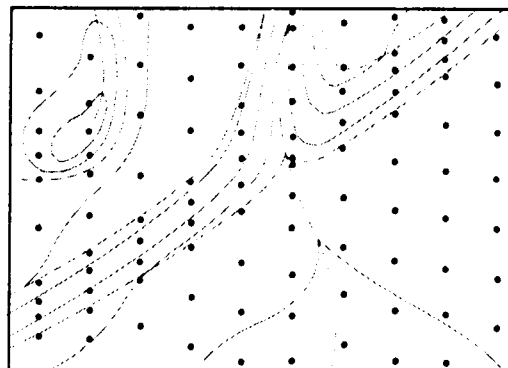
A variety of systems are available that can process one or more of these data types. Arc/Info along with its associated TIN package is one such system. Arc/Info is a collection of software tools which together provide the functionality of a Geographic Information System (GIS). A GIS is an automated system which has the ability to collect, store, retrieve, analyze, and display spatial data. A subset of tools available in the Arc/Info package, referred to as the TIN module, operates specifically with surface modeling. The TIN module permits the storage and analysis of both gridded DEMs and TINs. Procedures are also available to convert one data structure to the other.

1.2 Data Conversion

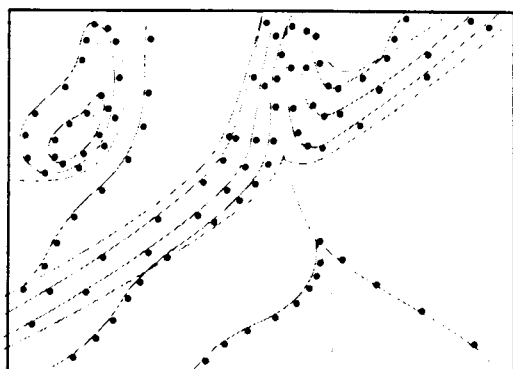
Several types of data structures represent topography. Each has its own set of advantages and disadvantages. In some cases, there may be a need to convert data from one structure into another. This can result from limitations of the system to model the topography, the desire to take advantage of some aspect of one data structure which is not available in another, or as a form of data reduction.



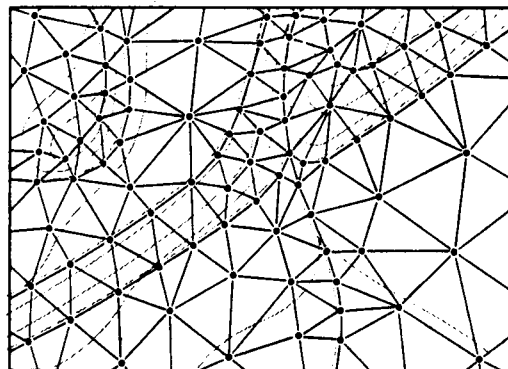
Gridded Matrix



Parallel Profiles



Digitized Contours



Triangulated Irregular Network

FIGURE 1.1

Four Sampling Schemes for Digital Elevation Data

Source: Carter, James R. (1988), "Digital Representations of Topographic Surfaces", Photogrammetric Engineering and Remote Sensing, Vol. 54, No. 11, pp. 1579.

Until recently, most systems capable of analyzing and displaying topographic surfaces had the limitation that the elevation data be in a gridded structure. If available surface points were irregularly distributed, a conversion to a rectangular matrix would have to be carried out before any other operation could be performed. Software systems which can process the TIN data structure have now become readily available. Depending on the algorithms employed, it is possible that a data conversion or transformation to a gridded format would be needed with some of these systems as well.

No single data structure best represents topography. Each data structure has advantages over its competitors if applied to certain tasks but can be inefficient when applied to others.⁴ While it may be better to convert data from one structure into another, it is important to note that with each transformation, information is lost and the definition of the surface can become more generalized.⁵

⁴Puecker, Thomas K. (1979), "Digital Terrain Models: An Overview", in Proceedings of the Fourth International Symposium on Computer Assisted Cartography, Vol. 1. Falls Church, VA: American Congress on Surveying and Mapping, pp. 98.

⁵Carter, James R. (1988), "Digital Representations of Topographic Surfaces", Photogrammetric Engineering and Remote Sensing, Vol. 54, No. 11, pp. 1577.

Generalization of surface definition may sometimes be desirable. Different levels of accuracy are required from one application to another. If data is only available from a relatively detailed, large scale source, but is to be used in a regional study, some form of data reduction would be in order. This can be necessary from a practical point of view, where computer storage space and processing time is prohibitive.

1.3 Purpose of Study

In the conversion from one data structure to another one should determine what aspects of surface definition have been lost, maintained, or altered. This evaluation is important as subsequent analyses may be affected. Results from these findings can help determine the nature and degree of generalization which has been introduced to the resulting dataset.

Work performed by the author has required the conversion of USGS DEMs, a common form of gridded elevation data, to TINs for use in an Arc/Info system. This was done because the Arc/Info system does not support the format of some of the DEMs being used.⁶ In addition, data reduction

⁶Support does not exist for the 1:250,000 scale USGS gridded DEMs because they are in the arc-second format.

was desired due to the large volume of data contained in the gridded models.

This work uses four case studies to assess what is lost in the transformation from a gridded matrix of elevations to TIN surfaces at varying levels of generalizations using the Arc/Info TIN procedures. Computer routines were developed by the author to manipulate the datasets, to analyze the surfaces, and to display the results. The methodologies developed here should be applicable to evaluating the quality of any data conversions between gridded DEMs and TINs.

CHAPTER II

THE GRIDDED DIGITAL ELEVATION MODEL

The gridded digital elevation model (DEM) is a numerical representation of terrain based on a regularly spaced array of surface elevations. Elevations are spaced at equal intervals along profiles, and at equal intervals between profiles.¹ This results in a rectangular grid, or matrix, of altitude data which can be used in conjunction with automated techniques for analysis. The gridded DEM facilitates efficient programming routines because the relationships between neighboring points are implicitly defined in the data structure.

The gridded DEM has historically been the most common form of a digital terrain model. This has logically resulted from two factors. Early computer technology enabled the input and analysis of elevation data, but with the constraint that the data structure could not be too complex. The gridded model is inherently simple and fits this criterion. In addition, the USGS began mass producing gridded models during the 1970's as part of its national

¹Evans, Ian S. (1980), "An Integrated System of Terrain Analysis and Slope Mapping", Zeitschrift fur Geomorphologie, N.F., Supp. Bd. 36, pp. 277.

mapping program. The production provided public access to thousands of models and prompted private industry to develop a substantial number of support programs.

2.1 Characteristics

The most important characteristic of a gridded DEM is the regular spacing of elevation points. This is significant both in terms of computer processing techniques and terrain representation. For computation, the spacing forms a simple matrix and thus permits uncomplicated data capture, storage, retrieval, and analysis. For terrain representation, the spacing and vertical resolution determines the model's ability to define the surface.

The gridded DEM attempts to represent a random surface with an ordered structure. The success of this is dependent on both surface variability and the resolution of the grid. One way of viewing this is in terms of spectral analysis.

Tobler² has stated:

"if a function has no spectral components of frequency higher than W , then the value of the function is completely determined by a knowledge of its values at points spaced $1/(2W)$ apart."

²Tobler, W. (1969). "Geographical Filters and Their Inverses", Geographical Analysis, Vol. 1, pp. 243.

Therefore, a regular grid of spacing D can only be expected to depict those feature of the surface having wavelengths of $2D$ or more.³ This can lead to problems, where the complexity of the terrain varies.

The fixed resolution of the grid does not adapt to surface variation which leads to problems of redundancy or over-simplification of the surface. Terrain complexity varies over horizontal distances whereas the gridded DEM does not. Therefore, if one produces a high resolution grid capable of defining the complex areas of a landscape, many points will be redundant in smooth areas.⁴ Conversely, if grid resolution is set so coarse as not to be redundant in smooth areas, it will be unable to capture the detail present in rough areas.

Vertical resolution is another factor which influences the gridded model's ability to define the surface. Vertical resolution refers to the precision of the units in which elevations are measured. These units are often measured in feet or meters with no fractional component. A model accurately recorded in whole feet would therefore be capable

³Mark, David M. (1975), "Computer Analysis of Topography: A Comparison of Terrain Storage Methods", Geografiska Annaler, Vol. 57A, No. 3-4, pp. 180.

⁴Ibid, pp. 180.

of defining a surface more precisely than a model recorded in whole meters.

The grid structure has the advantage of a simple implicit definition of topologic relations. Viewed as a matrix, the points are assigned their location based on row and column. Points in neighboring rows and columns can therefore be found without additional computation.⁵ This makes calculations of such parameters as slope and aspect relatively simple. Contours can be generated by linear interpolation between neighboring data points. Equations to calculate surface curvature and many other parameters have also been developed. Integration with satellite imagery, which has a similar gridded data structure, is another interesting possibility.⁶

The gridded DEM has an efficient storage structure. Since topologic relations are implicitly defined in the data structure, only z-values are needed for each point. In addition, the xy-coordinates of a point are determined

⁵Peucker, Thomas K. (1977), "Data Structures for Digital Terrain Models: Discussion and Comparison", in First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 5. Cambridge, MA: Harvard University, pp. 3.

⁶Franklin, Steven E. (1986), "An Approach to Processing Dense Grid Digital Elevation Models for Use in LANDSAT Terrain Classification", in 1986 ASPRS-ACSM Fall Convention Technical Papers. Falls Church, VA: ACSM, p. 219.

simply by identifying its row/column position in relation to a specified origin and grid spacing. Both of these aspects save storage space. This may be offset by the fact that some points are probably redundant.

2.2 Historical Developments

In the late 1950's, civil engineers in the Photogrammetry Laboratory at the Massachusetts Institute of Technology pioneered the use of digital terrain models. They were evaluating the utility of these models in several areas including transportation system location, volume measurements for mining operations, and surface change studies related to erosion.⁷ One of the data structures implemented was the rectangular grid. At the time, the rectangular grid permitted the highest degree of automation in the data procurement phase and the simplest processing of the data by the computer.

From that point on, the research community held great interest in digital terrain models. This was especially true in the fields of geodesy and photogrammetry.⁸ With

⁷Miller, C. L. and Laflamme, R. A. (1958), "The Digital Terrain Model: Theory and Application", Photogrammetric Engineering, Vol. 25, pp. 433-442.

⁸Peucker, Thomas K. (1979), "Digital Terrain Models: An Overview", in Proceedings of the Fourth International

possible exception of the military, little was achieved toward practical applications until the 1970's when datasets at reasonable sample densities and computers became readily available.

The gridded DEM continued to be the most common form of a digital terrain model. Peucker and Chrisman⁹ suggested this may have been a result of Boehm's¹⁰ research. Boehm described the advantages and disadvantages of different surface encoding schemes. He stated that the regular grid minimized computation time. He did this without comparison to a data structure with an irregular distribution of points. While Boehm may have influenced acceptance of the gridded model, the USGS probably played an even greater role.

The USGS was arguably the primary agent responsible for continuing the widespread use of the gridded DEM. In the mid-1970's, the USGS began mass producing and distributing

Symposium on Computer Assisted Cartography, Vol. 1. Falls Church, VA: American Congress on Surveying and Mapping, pp. 97.

⁹Peucker, Thomas K. and Chrisman, Nicholas (1975), "Cartographic Data Structures", The American Cartographer, Vol. 2, No. 1, pp. 61.

¹⁰Boehm, B. W. (1967), "Tabular Representation of Multivariate Functions With Application to Topographic Modeling", Association for Computing Machinery, 22nd. National Conference, pp. 403-415.

these models. Approximately twenty thousand models at various scales have been produced to date.¹¹

2.3 Sources of Data

There are three types of datasets used to create gridded DEMs. The first comprises sets of irregularly distributed points which are digitized with the understanding that each point is significant in defining the terrain. The second, more common type of dataset, contains regularly or irregularly distributed points where it is known that a number of points are redundant.¹² The third type of dataset is of digitized contour lines. For all three types of datasets, the information can be derived photogrammetrically, or from existing topographic maps. If the point sets are irregularly distributed, any one of a number of interpolation procedures can be used to construct the grids.

In the first situation, sample points are selected based on their relative importance in defining the terrain. These should be from critical locations on the surface where

¹¹Carter, James R. (1989), "Indices of the Planar Nature of Gridded DEMs", unpublished manuscript.

¹²Peucker, Thomas K. and Chrisman, Nicholas (1975), "Cartographic Data Structures", The American Cartographer, Vol. 2, No. 1, pp. 68.

major breaks in slope occur such as peaks, pits, ridges, and stream beds. This is considered surface specific, location random, sampling. After a sufficient number of points are assembled to define the surface, the irregularly distributed points are passed through an interpolation procedure to produce a regular grid. The accuracy of this method relies heavily on the operator's knowledge of the landscape.

The second type of dataset used to create gridded DEMs is derived from points randomly sampled from the surface. This approach is considered random in terms of defining the surface but points can be spaced at regular intervals along and/or between profiles and be location specific. In other words, points are not chosen by their relative importance in capturing significant features of the terrain, but rather by their position in reference to an origin and distance from other points. Thus, point selection is considered arbitrary.¹³ This approach is most conducive to automation. The success of this technique requires sufficient sampling intensity as determined by the spatial variability of the surface being modeled. Research has produced evidence that

¹³Mark, David M. (1975), "Computer Analysis of Topography: A Comparison of Terrain Storage Methods", Geografiska Annaler, Vol. 57A, No. 3-4, pp. 179.

random sampling may provide equal levels of accuracy to surface specific sampling.¹⁴

The third type of dataset used to create gridded DEMs is from a matrix of elevations interpolated from digitized contours. This technique contains characteristics of both surface specific and surface random sampling. Contours adapt to the variation of the surface being modeled. Therefore, high information content exists between contours. A drawback is redundancy along contours. Procedures have been developed to successfully interpolate a gridded surface from digitized contour lines. This often involves processing the contours into profile lines and interpolating them to a grid.¹⁵

2.4 Characteristics of USGS Gridded DEMs

The USGS now distributes gridded DEMs at three scales: 1:24,000, 1:100,000, and 1:250,000. The 1:24,000 and 1:100,000 scale series are produced by the USGS. The 1:250,000 scale series has been produced by the Defense Mapping Agency (DMA) but is distributed by the USGS.

¹⁴Davidson, John V. and MacEachren, Alan M. (1987), "Sampling and Isometric Mapping of Continuous Geographic Surfaces", American Cartographer, Vol. 14, No. 4, pp. 304.

¹⁵USGS (1987), Digital Elevation Models: Data Users Guide. Reston, VA: Department of the Interior, pp. 4.

Partial coverage of the United States is available in the two larger scales, with complete coverage being available at the 1:250,000 scale.

Several techniques have been used by the USGS to produce 1:24,000 scale DEMs. The Gestalt Photomapper II (GPM2) image correlation system, used in orthophoto production, samples points which can also be used to create the 1:24,000 scale DEMs.¹⁶ It correlates approximately 500,000 points from National High Altitude Photography (NHAP) for each 7.5-minute quadrangle.¹⁷ These points are regridded to form a standard USGS DEM. Other methods used to generate 1:24,000 scale DEMs are manual profiling from photogrammetric stereomodels, stereomodel digitizing of contours, and derivation from USGS Digital Line Graph (DLG)¹⁸ contour, hypsography, and hydrology categories.

¹⁶Allder, W. R., et al. (1983), "An Overview of Digital Elevation Model Production at the United States Geological Survey", in Proceedings of the Fifth International Symposium on Computer Assisted Cartography, Falls Church, VA: American Congress on Surveying and Mapping, pp. 24.

¹⁷USGS (1987), Digital Elevation Models: Data Users Guide. Reston, VA: Department of the Interior, pp. 4.

¹⁸Digital Line Graphs contain map information stored as vector data defining points, lines, and polygons. The USGS distributes these at various scales.

A standard 7.5-minute DEM has the following characteristics:

- 1) Consists of a regular array of elevations cast on the UTM coordinate system.
- 2) Data are stored as profiles in which the spacing along and between each profile is thirty meters. Data are stored from south to north along profiles and profiles are ordered from west to east.
- 3) The unit of coverage is the 7.5-minute quadrangle; no overedge coverage is provided.
- 4) Not all profiles have the same number of elevations.

The structure of the 7.5-minute DEM can be found in figure 2.1. Significant errors have been detected in the models derived from the GPM2.¹⁹ These usually occur over water bodies where image correlation is difficult to maintain. Blunders of this nature have supposedly been edited out of the newer models before they are released for distribution.

The USGS distributes 1:250,000 scale DEMs which were produced by the DMA. These are byproducts from the creation of three dimensional plastic relief models. To produce the plastic relief models, the 1:250,000 scale topographic maps were digitized. This included the digitization of contours, stream beds, ridges, and spot elevations. It was discovered that the information obtained from this procedure could also

¹⁹Carter, James R. (1989), "Relative Errors Identified in USGS Gridded DEMs", in Proceedings of the Ninth International Symposium on Computer Assisted Cartography. Falls Church, VA: American Congress on Surveying and Mapping, pp. 255-265.

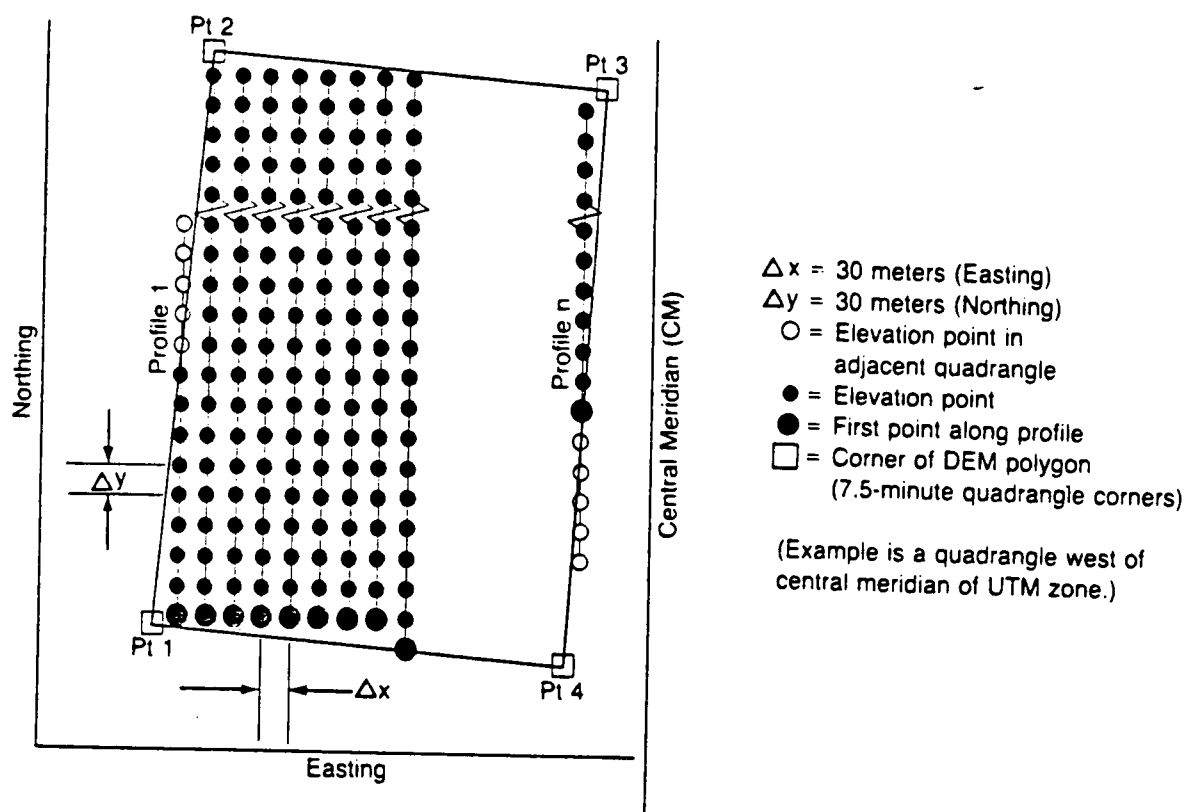


FIGURE 2.1

Structure of a 7.5-minute Digital Elevation Model

Source: USGS (1987), Digital Elevation Models: Data Users Guide. Reston, VA: Department of the Interior, pp. 4.

be used to create gridded DEMs. 1:250,000 scale DEMs have the following characteristics:

- 1) The product consists of a regular array of elevations referenced horizontally on the geographic (latitude/longitude) coordinate system of the World Geodetic System 1972 Datum.
- 2) The spacing of elevations along and between each profile is three arc-seconds. Therefore, the actual spacing between profiles varies with latitude.
- 3) The unit of coverage is a one degree by one degree block with data ordered from south to north in profiles ordered from west to east. The first and last data points are at the integer degrees of latitude. Two models are needed to cover the area of a standard DMA 1:250,000 scale topographic map.
- 4) Each model, south of fifty degrees north latitude, has 1,201 profiles containing 1,201 elevations.

The structure of a 1:250,000 scale, 1-degree, USGS DEM can be found in figure 2.2.

1:100,000 scale models are now being produced from DLG contour data at a spacing of two arc-seconds. Therefore, the grid system of these models have the same characteristics as the three arc-second data. They are usually divided into several sub-units which, when combined, cover the same area as a standard USGS 1:100,000 scale topographic map.

All gridded DEMs released by the USGS are classified into one of three levels of quality. Level-1 DEMs are considered to be elevation datasets in a standard format

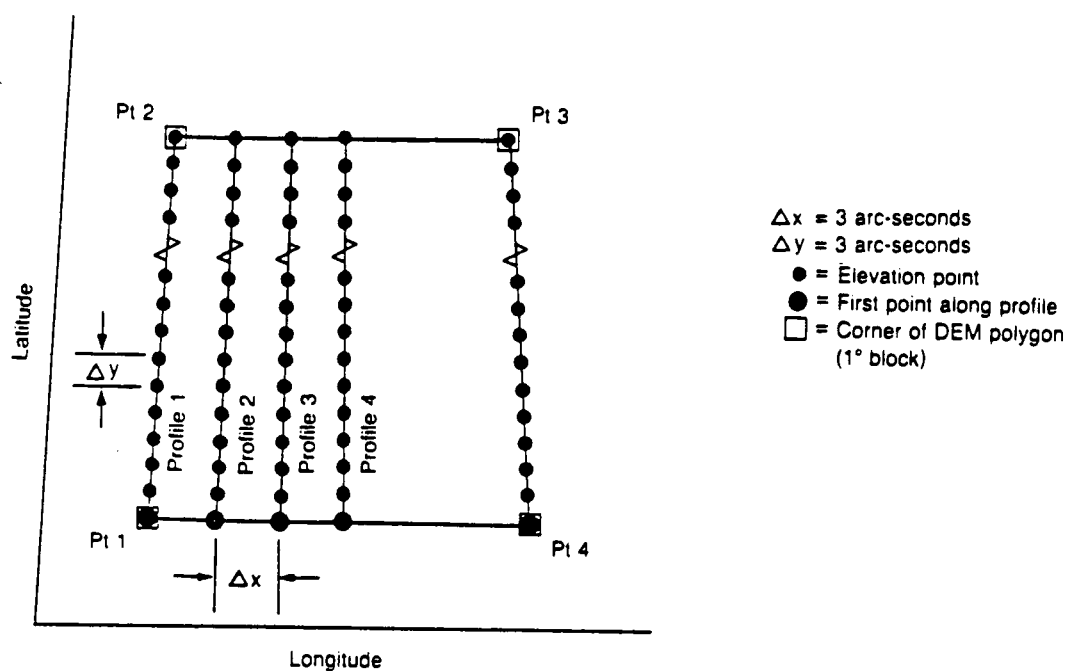


FIGURE 2.2

Structure of a 1-degree Digital Elevation Model

Source: USGS (1987), Digital Elevation Models: Data Users Guide. Reston, VA: Department of the Interior, pp. 6.

where only gross blunders have been removed. Virtually all 1:24,000 scale DEMs are classified in this level. Level-2 DEMs have been processed or smoothed for consistency and edited to remove identifiable systematic errors. Level-3 DEMs are datasets which have been vertically integrated to insure positional consistency with planimetric data such as hydrography and transportation. All 1:250,000 scale DEMs are classified as level-3.

2.5 Summary of Gridded DEMs

Gridded DEMs are the most common form of a digital terrain model. This is the result of data acquisition and computing simplicity. Although a gridded definition of a topographic surface has its shortcomings, such as fixed resolution, portrayals of the terrain adequate for most applications can still be achieved. Long term use of these models can be expected, particularly with agencies such as the USGS continuing to mass produce them.

CHAPTER III

THE TRIANGULATED IRREGULAR NETWORK

A Triangulated Irregular Network (TIN) is a digital terrain model which is based on an irregular array of points which form the nodes of a sheet of non-overlapping contiguous triangle facets.¹ The points used to construct a TIN are a set of critical locations on the surface which define breaks in slope such as peaks, pits, ridges, valleys, and passes. Their locations and height values are preserved in the model. No transformation into an intermediate structure takes place. Analysis is performed on the original data points with the necessary neighborhood relationships, the topology, being explicitly maintained.

The TIN model has gained acceptance during the past twenty years because of its efficiency and the manner in which it represents terrain. It adapts to the variation in non-stationary surfaces, maintains surface structure, and permits rapid computations. The increased availability of technologies which can create, store, and manipulate the TIN has also promoted its acceptance.

¹Peucker, Thomas K. et al. (1978), "The Triangulated Irregular Network", in Proceedings of the Digital Terrain Models (DTM) Symposium. Falls Church, VA: American Society of Photogrammetry, pp. 517.

3.1 Characteristics

A well designed TIN maintains consistency with the degree of variation in surface heights found in the terrain. As the terrain becomes more complex the resolution of the TIN should increase accordingly. This occurs because more points are sampled and included in the TIN model in areas of high complexity. In computations, such as spot height estimation, elevation values are interpolated for a given location based on the triangle in which it falls. Because triangles in the TIN are more densely packed in rough terrain, and the original data values are used as triangle nodes, relatively precise interpolations can be performed.

Surface structure is preserved in the TIN. The model is based on the principle that a flat plane can be fit to any 3 non-collinear points.² Thus, areas of consistent slope and aspect are represented in the TIN by individual triangles. Areas can be considered uniform if changes in slope or aspect are within a given tolerance level. A tolerance allowing for high levels of error will result in a generalized model of the surface. Slope and aspect can be

²Carter, James R. (1988), "Digital Representations of Topographic Surfaces", Photogrammetric Engineering and Remote Sensing, Vol. 54, No. 11, pp. 1578.

computed rapidly as only individual triangles need to be examined, rather than the relationship between them.

The TIN can also incorporate break lines. These are linear or areal features which represent important natural or artificial breaks in slope contained in the landscape. In many instances it is desirable to maintain the full representation of these in the terrain model. Examples of natural features would be ridges, channels, geological faults, and lakes. Artificial features could be tunnels, reservoirs, and buildings. The incorporation of break lines into a TIN is achieved either by not permitting a triangle edge to cross them during triangulation or by adding an additional localized triangulation into an existing model.³ This feature is not inherent in the raster based DEM.

A topological structure is another important characteristic of the TIN. Once the topology of the TIN has been defined, the neighborhood of a sample point can be reconstructed without resorting to searching.⁴ A triangle record contained in a TIN with complete topology has

³Chen, Zi-Tan (1988), "Break Lines on Terrain Surface", in Proceedings, GIS/LIS '88. Falls Church, VA: American Society for Photogrammetry and Remote Sensing, pp. 781.

⁴Peucker, Thomas K. et al. (1978), "The Triangulated Irregular Network", in Proceedings of the Digital Terrain Models (DTM) Symposium. Falls Church, VA: American Society of Photogrammetry, pp. 519.

explicit reference to the three points which define it and the three associated triangles which share its borders. Thus there are six references maintained for each triangle. This greatly reduces processing time but increases storage space. Storage requirements can be reduced by referencing only the points defining each triangle. This method's primary disadvantage is that it increases computing time for applications which require topology because neighborhood relationships between triangles must be determined during processing.

3.2 Historical Developments

Originally, manual techniques were used to represent surfaces. Contour maps would be based on information provided at certain control points gathered from field surveys or aerial photography. A triangular framework was provided by the surveyor to serve as a guide for hand contouring. The cartographer would visually interpolate elevation values with the use of this frame.

For more than twenty years after the development of computers, automated methods for triangulating random data sets were not well developed. It was more common to interpolate regular grids from random point sets. The implementations to generate TINs which did exist were

extremely slow.⁵ This was because the TIN structure, requiring the relation of triangle to triangle, was complex. The human brain can easily distinguish relationships through visual inspection, but the computer has a much harder time making these connections. The mathematics of topology and its implementation in computer programming was still in its infancy.

One of the first automated triangulation procedures was developed in the mid 1960's. This method was published by IBM in 1966 as part of the documentation for a contouring system.⁶ It had difficulty performing the triangulation unless the point set contained a regular spatial distribution. In 1973, Rhynsburger⁷ published a solution for automated triangulation through use of Thiessen polygons. Thiessen polygons are important not only in triangulation but in geographical thought as well.

⁵McCullagh, Michael J. (1979), "Triangular Systems in Surface Representations", in Proceedings of the Fourth International Symposium on Computer Assisted Cartography, Vol. 1. Falls Church, VA: American Congress on Surveying and Mapping, pp. 146.

⁶Elfic, M. H. (1979), "Contouring by Use of a Triangular Mesh, Cartographic Journal, Vol. 16, No. 1, pp. 24.

⁷Rhynsburger, D. (1973), "Analytic Delineation of Thiessen Polygons", Geographical Analysis, Vol. 5.

In 1911, the climatologist A. H. Thiessen suggested a method of representing precipitation data collected from unevenly distributed weather stations. He defined regions based on a set of data points such that regions be enclosed by a line midway between the station under consideration and the nearest surrounding stations.⁸ The important concept is it defines the region of influence of any point in an areal context. The dual of the Thiessen diagram is a triangulation based on the same proximity criteria.

Delaunay first recognized the relationship between the Thiessen polygon and triangulation. Instead of showing regions around points, it is also possible to connect the neighboring points. When this is accomplished, a diagram of non-overlapping triangle facets has been created (see figure 3.1). A Delaunay triangle forms a circumscribing circle which includes no other points in the data distribution.⁹ To date, this technique is the most frequently used in TIN construction.

⁸Thiessen, A. H. and Alter, J. C. (1911), "Precipitation Averages for Large Areas", Monthly Weather Review, Vol. 39, pp. 1082-1084.

⁹McCullagh, Michael J. and Ross, Charles G. (1980), "Delaunay Triangulation of a Random Data Set for Isarithmic Mapping", Cartographic Journal, Vol. 17, No. 2, pp. 94.

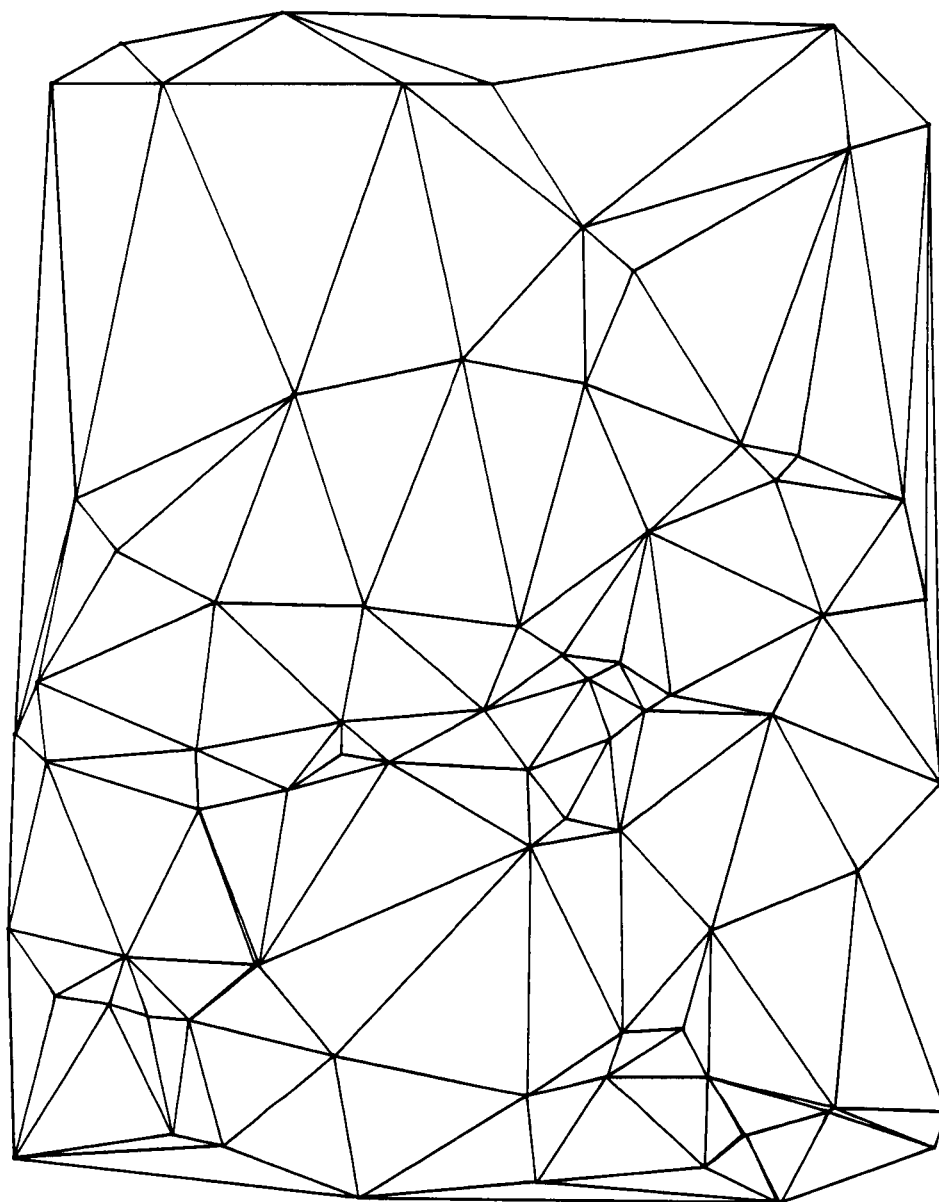


FIGURE 3.1

Delaunay Triangulation

3.3 Sources of Data Used in Triangulation

Currently, the most difficult process in generating TINs is obtaining points. A poor selection in either point quantity or pattern can lead to substandard results. The most common ways of acquiring points are by sampling critical spot heights with use of aerial photography or topographic maps, using points defining digitized contours, and selecting points from existing gridded DEMs. It is important to note that the distribution of points is random in a geometric sense. Geomorphologically, they should be characteristic points of the terrain.¹⁰

Sampling critical spot heights is one of the better means of obtaining points for TINs. This approach allows one to visually pick the most appropriate points. These points should represent major breaks in slope such as ridge and valley lines, peaks, pits, and passes. The operator must choose the points to represent the surface. As in many processes, the skill of the operator in picking points is a significant factor in the quality of the TIN. Some computer programs work interactively with the operator to determine

¹⁰Yoeli, Pinhas (1977), "Computer Executed Interpolation of Contours into Arrays of Randomly Distributed Height-Points", Cartographic Journal, Vol. 14, No. 2, pp. 103.

optimal sample size and pattern for any given terrain.¹¹

This may help lead to more dependable results.

The use of points defining contour lines is also a fairly common method of obtaining points for TINs. Contours are frequently output from photogrammetric digitizing and some USGS DEMs are now available as digitized contours - those in the DLG format. Contours have the advantage that the density of the contours changes with the dynamics of the relief. This keeps the volume of redundant data low between contours, although along contours there tends to be a dense collection of points. Problems result if highly detailed contour lines are widely spaced because triangles can form on three vertices from the same contour line.¹² These triangles would then be considered flat when the area they represent truly is not.

Points can be sampled from gridded DEMs. This is an efficient process because computer programs can automatically select critical points contained within the gridded datasets. A force promoting the use of this

¹¹Ayeni, O. O. (1982), "Optimum Sampling for Digital Terrain Models: A Trend Towards Automation", Photogrammetric Engineering and Remote Sensing, Vol. 48, No. 11, pp. 1692.

¹²Environmental Systems Research Institute (1988), TIN Training Course Class Materials. Redlands, CA: ESRI, pp. 3(11).

procedure is that a tremendous amount of gridded DEMs already exist. The USGS mass produces them at varying scales and they are relatively inexpensive. A drawback is that any TIN created by this process cannot be better than the original model and could be of lesser quality.

3.4 Methods of Triangulating A Random Point Set

Recent methods of triangulation are based on automated procedures stressing the use of computational geometry. This approach attempts to understand the interaction between geometry and computing. The task is to isolate fundamental problems and develop optimal algorithms to solve them. This creates a set of generic tools which can be used in diverse applications.¹³ Work in the field of computational geometry has not only helped triangulation procedures but a multitude of other problems implemented in GISs as well.

In triangulation, the identification of neighboring points is critical. Much of the complexity lies in finding potential neighbors and verifying whether or not they should be combined to form triangles. If the number of data points equals N the number of possible lines connecting them is

¹³Shamos, Michael I. and Bentley, J. L. (1978), "Optimal Algorithms for Structuring Geographic Data", in Harvard Papers on Geographical Information Systems, Vol. 6. pp. 1.

$N * (N - 1) / 2$.¹⁴ While many algorithms have been developed to locate closest points, some perform the operations more efficiently than others. The best utilize the 'divide-and-conquer'¹⁵ method where points are sorted and split into groupings based on their spatial location. Therefore, when a search is performed it need only be performed on points that are potential participants rather than the entire dataset.

Brassel and Reif¹⁶ described a Thiessen polygon approach that sorted points in one dimension based on the x-coordinate. For triangulation a search is performed to determine the closest neighbor of the arbitrarily chosen starting point. From here the appropriate points to form triangles around the starting point, or rotation point, are located. This is achieved by searching in a clockwise direction from the base line formed by the first two points

¹⁴Yoeli, Pinhas (1977), "Computer Executed Interpolation of Contours into Arrays of Randomly Distributed Height-Points", Cartographic Journal, Vol. 14, No. 2, pp. 103.

¹⁵Bentley, Jon L. and Shamos, Michael I. (1976), "Divide-and-Conquer in Multidimensional Space", in Proceedings, Eighth Annual ACM Symposium on Theory of Computing. New York: Association for Computing Machinery, pp. 220-230.

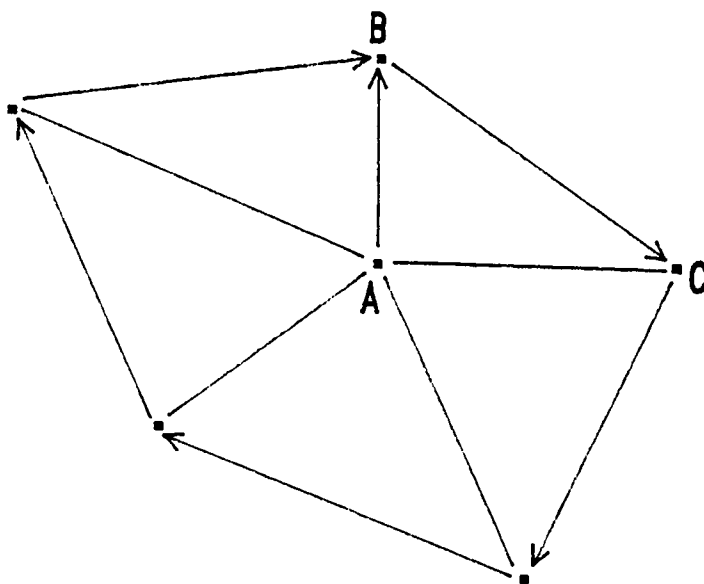
¹⁶Brassel, Kurt E. and Reif, Douglas (1979), "A Procedure to Generate Thiessen Polygons", Geographical Analysis, Vol. 11, No. 3, pp. 289-303.

(see figure 3.2). The first point encountered which lies in the proper plane is considered a potential neighbor. A circumscribing circle is computed around the three points and a search is done limited to the spatial extent of the circle. If a new point is found to be inside the circle and in the correct plane, it becomes a new potential neighbor. This can be referred to as a limited outward-inward search.

When a third point is found which forms a circumscribing circle which contains no other points in the distribution, a triangle can be formed. A new base line is created between the rotation point and third point; a new search is then performed. This continues around the rotation point until all possible triangles about it are defined. The process is then repeated using each of the found points as rotation points themselves. Search time is reduced as progression continues because an index is set up, tracking neighborhood relationships. By the time any point is visited as a rotation point, a partial listing of its neighbors already exists.

McCullagh and Ross¹⁷ presented a similar technique but improved the efficiency of the neighborhood search. Rather

¹⁷McCullagh, Michael J. and Ross, Charles G. (1980), "Delaunay Triangulation of a Random Data Set for Isarithmic Mapping", Cartographic Journal, Vol. 17, No. 2, pp. 93-99.



A = rotation point B = nearest neighbor
C = third point forming first triangle

FIGURE 3.2

Rotational Search of Points for Triangulation

than using an outward-inward approach, they suggested one which was inward-outward. Points are sorted in two dimensions dividing space into an imaginary box structure. Points are labeled by the box they are contained in. As searches are needed it is determined what boxes to limit the searches to and thus which points. The circle test is used but instead of it proceeding from potentially large to small, it starts small and if no points are contained within the circle, it is enlarged. If multiple points are found in the circle the one which subtends the largest angle to the base line is used to form a triangle.

3.5 Summary of the TIN

The TIN is rapidly becoming one of the most common forms of a terrain model. It arguably represents surfaces more accurately and efficiently than other models because it conforms to surface structure and maintains explicit topological relations. The majority of GISs which offer surface processing capabilities now support the TIN.

CHAPTER IV

METHODOLOGY

The set of procedures needed to access gridded DEMs, generate various TIN representations from the DEMs, interpolate from the TIN surface back to a grid, and compare the reconstituted gridded surface with the original is complex. The USGS DEMs must be read from a tape and imported into a system which is designed to handle them. Critical points have to be extracted from the DEMs for use in TIN construction. A conversion of each TIN surface back to a gridded format is necessary for comparison to the original data. Statistics then are performed with the results examined. These have all been accomplished with the use of existing software, Arc/Info, and programs developed by the author.

4.1 Conversion of Gridded DEMs to TINs

One procedure selectively extracts points from the gridded DEMs to provide a surface composed of triangle facets. It would be undesirable to select all points for several reasons. First, a full USGS gridded DEM contains too many points to construct a TIN. At present only 50,000 points can be used for any single model in Arc/Info. If one

could use all points, some of the advantages of the TIN are lost. For example, many points in a raster based DEM are redundant. Numerous points define the surface in areas of relatively homogeneous characteristics. Since not all points are needed to represent the surface in a TIN, a method should be used to filter out the less significant points and extract the ones with more information content.

Arc/Info provides a routine in its TIN module to select critical points from a lattice. A lattice represents a grid structure, like the USGS DEMs, and is simply a proprietary format. These files are created from DEMs when an import routine is run. The procedure to extract points from a lattice is referred to as VIP.¹ An acronym for "Very Important Points", the algorithm examines the significance of every point in the lattice. The user is allowed to specify the number of points to be extracted by stating a percentage; ie. the ratio of desired points to the total number of points in the lattice. For any given percentage, VIP returns the best set of points which can be used to construct a TIN.

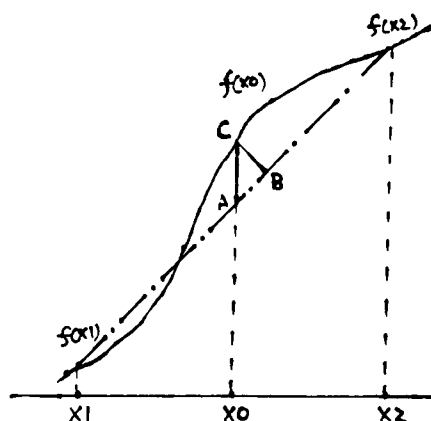
The VIP algorithm uses a roving 3x3 window which is passed through the entire dataset. Within the window, the

¹ESRI (1988), "TIN Users Guide", Redlands CA: Environmental Systems Research Institute.

difference in the predicted to the actual elevation of the center pixel is calculated in four directions: up-down, left-right, upper left-lower right, and lower left-upper right. To see how this measurement is made, refer to figure 4.1. The absolute values of the four differences are added together for each pixel with the result representing its relative significance.

A histogram depicting the distribution of significance for the pixels is then constructed. This forms a distribution with many points having less significance and fewer points having more. Given the desired percentage of points a threshold is determined which provides the significance level at which to start selecting pixels. When a point is encountered with a significance level above the threshold, its x, y, and z coordinates are placed into a vector point file. The number of pixels above the threshold is not precisely the number of points desired, but it is within a fraction of a percent. It is possible to obtain a printout of the histogram without actually selecting the points. This enables the user to see the distribution of differences with different percentages.

After points have been selected from the lattice files, they are used to form TINs with Arc/Info's ARCTIN procedure. An option with the procedure allows for a point dataset to



$f(x_0)$ = elevation of center pixel
 $f(x_1)$ = elevation of first neighboring pixel
 $f(x_2)$ = elevation of second neighboring pixel

In a given direction, the difference between the actual and predicted elevations for the center pixel is calculated by measuring the distance BC.

FIGURE 4.1

Difference Between Actual and Predicted Elevations

Source: Chen, Zi-Tan and Guevara, Armando (1987), "Systematic Selection of Very Important Points (VIP) from Digital Terrain Model for Constructing Triangular Irregular Networks", in Proceedings of the Eighth International Symposium on Computer Assisted Cartography. Falls Church, VA: American Congress on Surveying and Mapping, pp. 52.

be used as the source file.² Distance tolerances can be specified which eliminate points that are considered too close together in xy-coordinate space. The tolerances used in this study were such that no points from any of the gridded DEM datasets would be excluded from consideration in the construction of TINs of various percentages.

4.2 Conversion of TINs to Raster Format

In order to quantitatively compare the TIN surfaces to the gridded DEMs it is desirable to have both in a similar format. For this study the TINs, generated from a lattice, were converted from their vector representation back to a lattice of the same resolution as the original gridded datasets. With two lattices of the same resolution a comparison can be made. Arc/Info provides an efficient means to create a new lattice from a TIN.

The TINLATTICE procedure generates a regularly spaced grid of height values by interpolating from a TIN model. It allows for both linear and smooth interpolations. The number of rows and columns as well as the spacing between points can be specified. To understand this procedure,

²Polygon and line files can be input into the ARCTIN procedure as well with points defining these features being used in triangulation.

think of a grid superimposed on a TIN (see figure 4.2). For each intersection in the grid an elevation is interpolated.

For these studies a linear interpolation from the TINs to lattices was performed. In many mapping situations grid point elevation values are predicted by the use of a distance weighted function against the three vertices of the triangles in which they fall, or in other cases, elevation values are based on four or more nearest known points. Although such weighted algorithms are used frequently in producing 'smooth' contour maps, it was felt that a weighted interpolation would not be appropriate here. Peucker³ has stated that a linear interpolation is the most likely approximation between a few known points. In addition, the documentation to the TINLATTICE procedure states that unpredictable results can occur with smooth interpolations.

After a lattice is created in Arc/Info, the data matrix is converted to a generic binary integer format for evaluation on a microcomputer. The lattice is stored in real numbers, but since the original data was integer and the precision of real numbers is unwarranted, the switch to

³Peucker, Thomas K. (1979), "Digital Terrain Models: An Overview", in Proceedings of the Fourth International Symposium on Computer Assisted Cartography, Vol. 1. Falls Church, VA: American Congress on Surveying and Mapping, pp. 102.

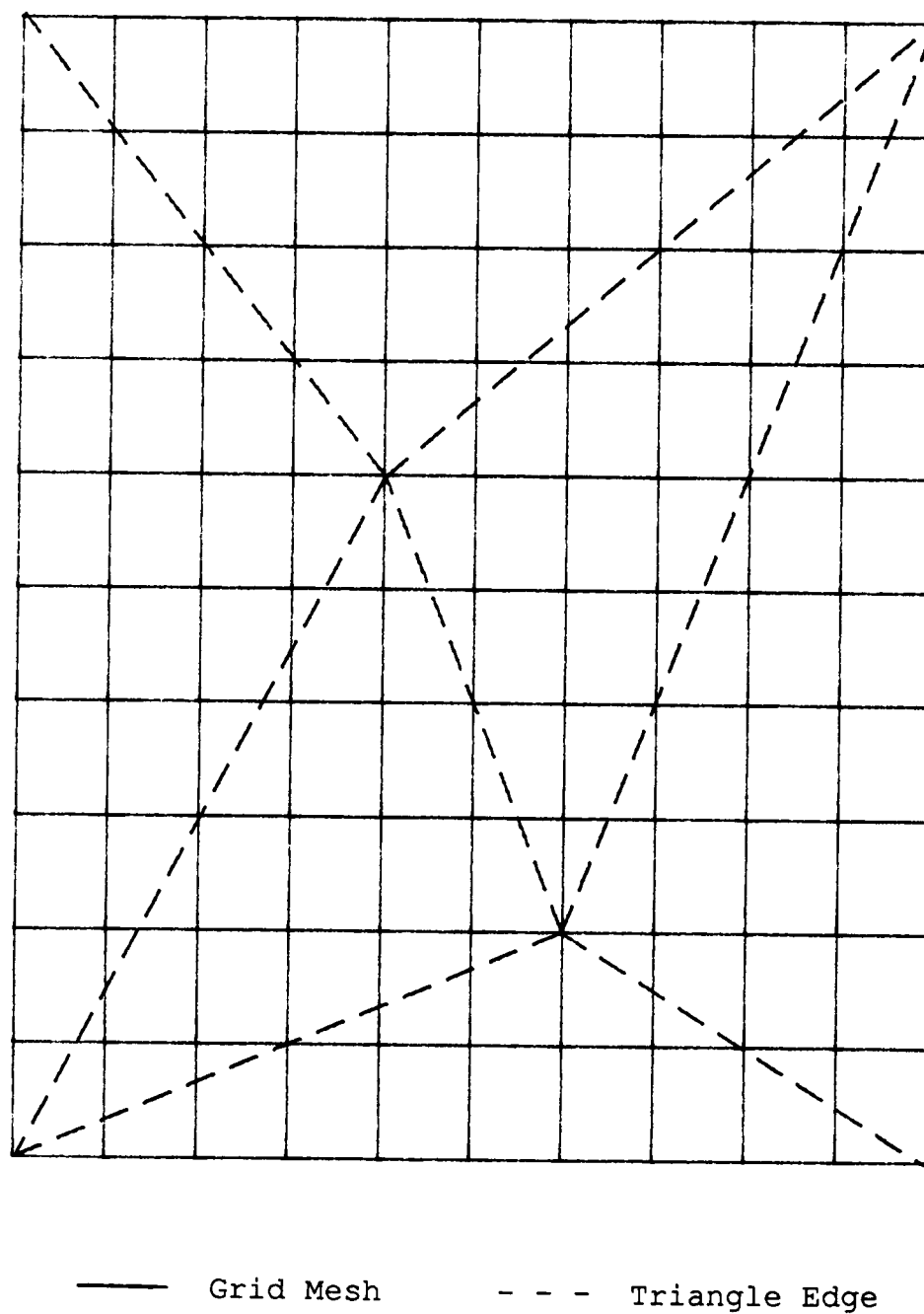


FIGURE 4.2

Grid Superimposed on a TIN Surface

integer format is acceptable and reduces storage requirements.

An Arc/Info procedure, LATTICEGRID, is used to perform the first part of this conversion. The resulting datafile, referred to as a GRID-SVF (grid single variable file), is downloaded to a personal computer. A program which the author created then converts the GRID-SVF file to one which contains only elevation values. Maintaining binary format, two bytes are used for each interpolated elevation. This simple structure enables analytical programs to efficiently access the data.

4.3 General Description of Evaluation Procedures

Methods were developed to characterize the original gridded datasets, the newly gridded datasets interpolated from the TINs, and the relationships between them. Some of these methods implement standard geomorphometric techniques that describe the terrain. Other methods, such as correlation analysis and ridge-form definition, were used to compare datasets to each other. This section provides a general description of the procedures which were used in the study.

Geomorphometry has been defined as the science which treats the geometry of the landscape.⁴ Two subdisciplines of this science are general and specific geomorphometry. General geomorphometry is the measurement and analysis of those characteristics of landforms which are applicable to any continuous rough surface. This can be distinguished from specific geomorphometry which measures and analyzes specific landforms such as cirques and drumlins.⁵ A strength of general geomorphometry is that it provides a basis for quantitative comparisons of qualitatively different landscapes. It has been found useful in providing global descriptions of the DEMs.

Most of the parameters in general geomorphometry are derived from altitude and are composed of descriptive statistics. The frequency distributions of these rather simple statistics successfully portray some of the complex characteristics inherent in terrain. These parameters attempt to characterize terrain in both horizontal and

⁴Mark, David M. (1975), "Geomorphometric Parameters: A Review and Evaluation", Geografiska Annaler, Vol. 57A, No. 3-4, pp. 165.

⁵Evans, Ian S. (1972), "General Geomorphometry, Derivations of Altitude, and Descriptive Statistics", in Spatial Analysis in Geomorphology. edited by R.J. Chorley, London : Methuen, pp. 18.

vertical dimensions. This is essential in describing topographic form.

Relief is one of the primary components used in characterizing terrain. Quite often relief has simply been defined as the range in elevation of a study site. The range is defined as the difference between minimum and maximum elevations contained in the dataset. Although this has merit, attention should focus on the fact that it relies on minimum and maximum values. These can be dramatically influenced by small shifts in study area size or location.

Evans⁶ suggested a more powerful measure of relief, that of the standard deviation of elevation. To describe the dispersion of a distribution, the standard deviation is a more effective and stable statistic than the range because it is based on all values, not only the two extremes. It is even more informative when used in conjunction with the mean and skewness of elevation. Still, range in elevation is more frequently used probably because it is easier to calculate. This study used the range as one means to describe the original datasets, but the standard deviation of elevation was employed for statistical comparisons.

⁶Ibid, pp. 31-33.

Studying the interrelationship between area and altitude is another method used to characterize terrain. Most measures of this interrelationship which describe aspects of the distribution of landmass with elevation are based on the hypsometric curve. One form of the curve is the absolute hypsometric curve. It is a plot of absolute or relative area above a certain elevation and is basically a cumulative frequency diagram of elevations.

The most widely used parameter based on the hypsometric curve is the hypsometric integral.⁷ Sometimes referred to as the coefficient of relative massiveness it attempts to depict the 'dissection' or 'aeration' of the landscape. It is calculated by subtracting the minimum elevation from the mean and dividing the result by the range. It's value will always fall between zero and one. Larger values are supposed to reflect youthful, unstable areas of landscape development.

A common technique used in raster based GISs has been applied which calculates the local difference between associated datasets. This involves literally subtracting corresponding pixels in one dataset from another and

⁷Strahler, A. N. (1952), "Hypsometry (area-altitude) Analysis of Erosional Topography", Geological Society of America Bulletin, Vol. 63, pp. 1121.

creating a new dataset of the differences. These differences show, on a pixel-by-pixel basis, the contrast or similarity between the two datasets. One can use descriptive statistics to convey information about a continuous surface of departure. It can be treated in the same manner as the original elevation data. Departures can also be mapped to help display locations of worst fit.

Correlation analysis was performed to display the association between datasets. Correlation is performed by matching pixels from an original gridded DEM to one interpolated from a TIN. One would expect high correlations but the insight comes in observing how correlation coefficients vary with TINs derived from different degrees of generalization.

The delineation of geographic features, such as ridges, is another method which was used to compare datasets. The previous measures of comparison rely on aggregate statistics. While these are informative, additional insight can be gained by examining the spatial distribution of features, such as ridges. Ridge patterns, as depicted in the original DEMs, can be compared to those contained in the TIN derived models. In this fashion, one can determine whether surface structure is acceptably preserved in the TIN models based on different levels of generalization.

In order to extract ridges from gridded models, a neighborhood function was used. This function depends on a combination of indices. One index is surface convexity, where the elevation of the center pixel in a 3x3 window is compared to the average elevation of its eight neighbors. If it is higher, the surface can be considered convex and associated with a ridge. One should be able to specify how much higher the center pixel need be before it is identified as representing a convex surface.

The other related index counts the number of neighboring pixels which are higher or lower than the center pixel. If a sufficient number of neighbors are lower (at least four), then the center pixel can be identified as belonging to a ridge. This number should be somewhat flexible and determined by the terrain being modeled.

The simultaneous use of these two indices creates an adaptable function which can define the ridge form of different types of terrain. For this study each pixel needs to satisfy both parameters before it is designated as part of a ridge. This was found to be effective on both rough and smooth surfaces, where ridges may or may not be sharply defined.

4.4 Specific Methodology

The above analyses were performed on the original DEMs and on the gridded matrices derived from TIN surfaces. All of the analytical procedures were achieved by implementing programs developed by the author for this research.⁸ TINs were made from one, three, five, seven, ten, fifteen, and twenty percent samples. Higher percentage samples would be unjustified, as a twenty percent TIN already occupies twice the computer storage space required by the original grid.

The original DEMs were obtained by extracting submatrices from three USGS gridded DEMs to yield four different study sites. Two 300x300 submatrices were taken from the 1:250,000 West-Knoxville DEM. One 300x300 submatrix was taken from the 1:24,000 Meadow, Tennessee DEM. One 300x300 submatrix was taken from the 1:24,000 Cades Cove, Tennessee/North Carolina DEM. The grid spacing used for the 1:250,000 scale models is three arc-seconds, being seventy five meters, east to west, and ninety two meters north to south. This approximates the actual ground distance between points at the middle latitude of the West-

⁸Programs were written in Turbo Pascal version 5.5 on a MS-DOS microcomputer with a 80286 CPU running at 12.5 mHz.

Knoxville DEM.⁹ The grid spacing used for the 1:24,000 scale models is a square thirty meters on a side.

All matrices were checked to see if they contained blunders.¹⁰ The only case where blunders were discovered was in the Cades Cove DEM. Because another DEM in an area of comparable relief and free from blunders could not be found, the author used the Cades Cove DEM. The areas containing the blunders in this DEM, in both the original and TIN derived models, were excluded from all calculations.

TINs were generated from these models at the varying resolutions and interpolated back to raster format. Afterwards, twenty five rows and columns were clipped from each edge of the original and interpolated datasets to yield 250x250 matrices. This was done to eliminate the inherent edge-effect in TIN models.

⁹The horizontal distance between points actually decreases, following the convergence of the meridians, towards the north. For the size of the sites used, the convergence is minimal and is therefore ignored. If larger areas were to be used, a spacing of three arc-seconds would be given. After points were selected from the grid file, they could then be re-projected to render the correct horizontal and vertical spacing needed to be used in a TIN.

¹⁰Carter, James R. (1989), "Relative Errors Identified in USGS Gridded DEMs", in Proceedings of the Ninth International Symposium on Computer Assisted Cartography. Falls Church, VA: American Congress on Surveying and Mapping, pp. 255-265.

To describe the original datasets the minimum, maximum, mean, range, hypsometric integral, and standard deviation of elevation were calculated. Hypsometric curves and contour plots were also plotted for each of the original models. These provide a general idea of the nature of terrain represented by the different models.

The mean and standard deviation of elevation were calculated and graphed for each TIN derived model. These values from the original models were placed on the charts as well for comparison. Correlations of each TIN derived model to its original were also performed and graphed.

Several graphs were prepared to evaluate the departure of each TIN to the original model. These were the maximum absolute departure, mean absolute departure, standard deviation of departure, and distribution of departure.

Maps of ridge form were created for all models. Both function indices were consistently used. The parameters which produced reasonable results for all models were that each pixel must be at least two meters higher than the average elevation of its eight neighbors, and at least higher in elevation than four of its neighbors.

All of these figures are presented and discussed in the following chapter. Used together, they provide valuable

information to evaluate the quality of TINs derived from gridded DEMs.

CHAPTER V

FINDINGS

The four DEMs used in this research were selected to provide contrasts in both scale and terrain characteristics. Two of the DEMs originate from 1:250,000 scale data and the two others from 1:24,000 scale data. At each scale, two distinctly different physiographic environments are examined. The relatively smooth, undulating terrain of the Ridge and Valley Physiographic Province in Tennessee is represented by one site while the other exhibits the rough, high relief characteristics of the Blue Ridge Mountains in North Carolina and Tennessee.

Models containing these different properties were used in order to evaluate the relative performance of the TINs. One assumption is that fewer points would be needed from gridded DEMs of low relief than of high relief. This is readily adapted in other sampling procedures, as the TIN is designed for this, but no current literature provides a quantified verification of this for sampling from gridded data. In addition, documentation should indicate whether TINs generated from larger scale data perform differently than TINs constructed from smaller scale DEMs.

5.1 Description of Original Datasets

Study sites 250LR and 250HR are derived from the 1:250,000 scale datasets with elevations sampled every three arc-seconds. Site 250LR is from the low relief ridge and valley while site 250HR is from the high relief Blue Ridge Mountains. Sites 24LR and 24HR are from the 1:24,000 scale datasets with elevations sampled at every thirty meters. Site 24LR exhibits the low relief of the ridge and valley while site 24HR is within the high relief Blue Ridge. Site 24HR is contained within the area encompassed by 250HR. All dataset matrices were 250x250 in size. The UTM coordinates of the southwest corner of these datasets are listed in table 5.1.

Table 5.1 UTM Coordinates of DEM Origins

Site	UTM Zone	Easting	Northing
250LR	17	230,745	3,962,238
250HR	17	237,948	3,922,416
24LR	16	750,769	3,948,985
24HR	17	241,334	3,935,116

A quantitative description of the datasets is the primary methodology employed. Table 5.2 lists the fundamental characteristics inherent in the different models.

Table 5.2 - Characteristics of Original Datasets

Site	Minimum	Maximum	Range	Mean	H.I.	S.D.
	m.	m.	m.	m.		m.
250LR	250	563	313	290	0.129	26.61
250HR	380	1676	1296	828	0.346	264.49
24LR	251	423	172	318	0.387	23.10
24HR	493	1450	957	876	0.401	259.44

H.I. - refers to the hypsometric integral

S.D. - refers to the standard deviation

m. - values are in meters of elevation

The values shown in the table display the contrast in landscapes between datasets. Note that 250LR and 24LR are similar in characteristic as are 250HR and 24HR particularly in terms of the mean, and standard deviation. The range and hypsometric integral are less valuable than the standard deviation in portraying relief for site 250LR. This is because only a few observations in the southeast corner of the dataset approach the maximum elevation. This exotic extreme gives an unrepresentative maximum, and thus an unrepresentative range and hypsometric integral. Hypsometric curves do a better job of portraying the distribution of landmass against elevation. These curves are provided in figures 5.1 through 5.4. Site 250LR, depicted in figure 5.1 shows a very small percentage of area above three hundred seventy five meters. This is why use of the hypsometric integral and range, used by themselves, are

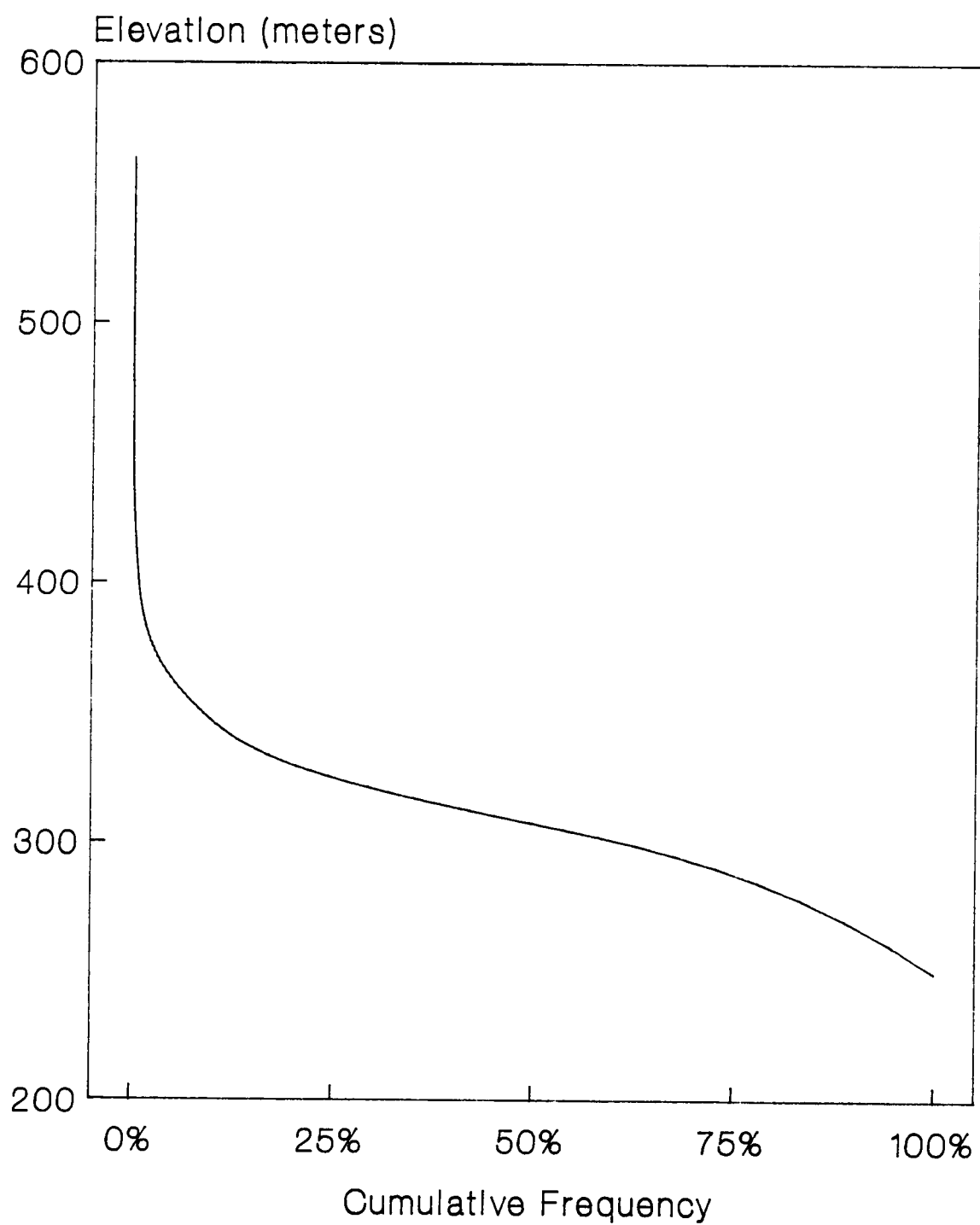


FIGURE 5.1

Hypsometric Curve - Site 250LR

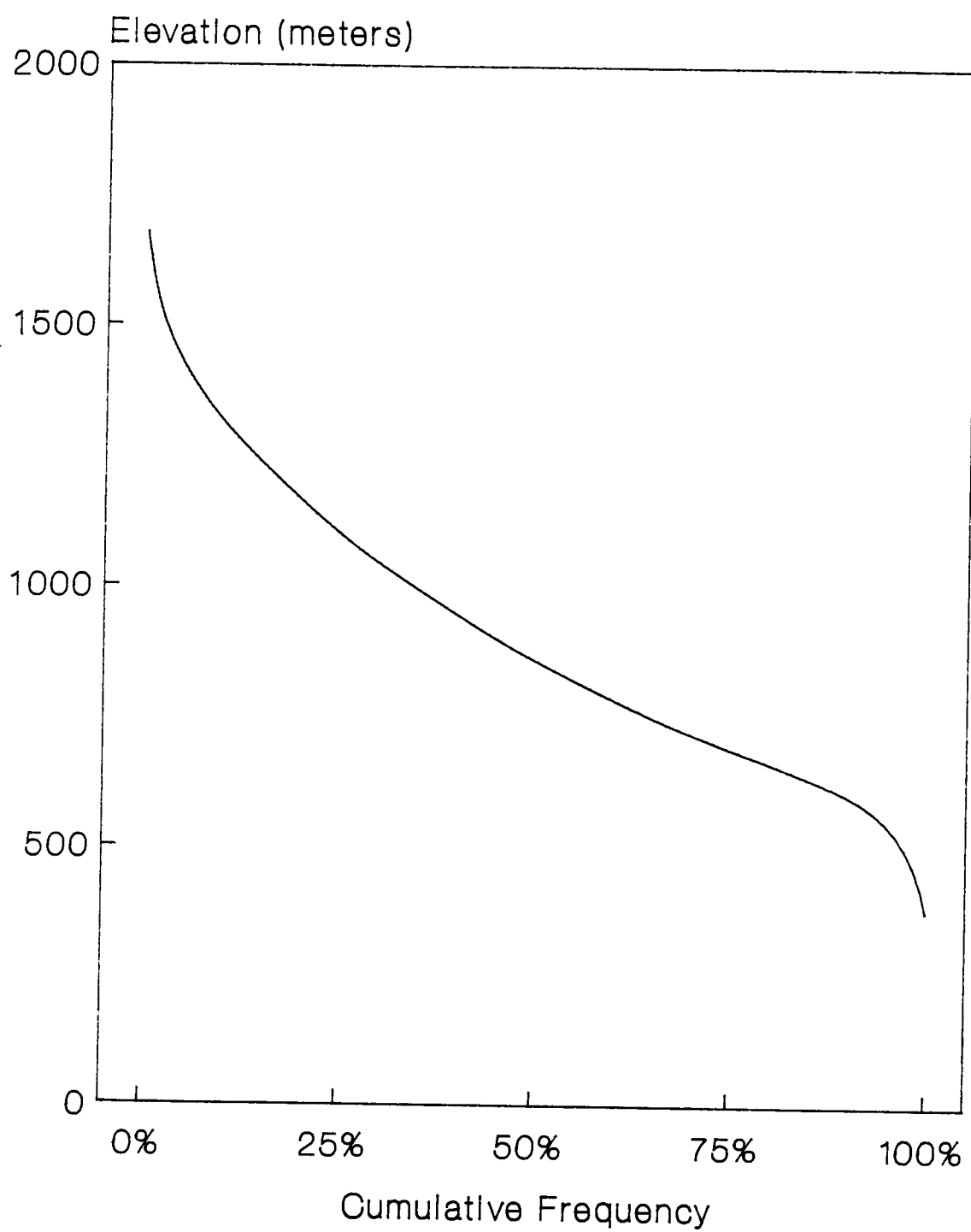


FIGURE 5.2

Hypsometric Curve - Site 250HR

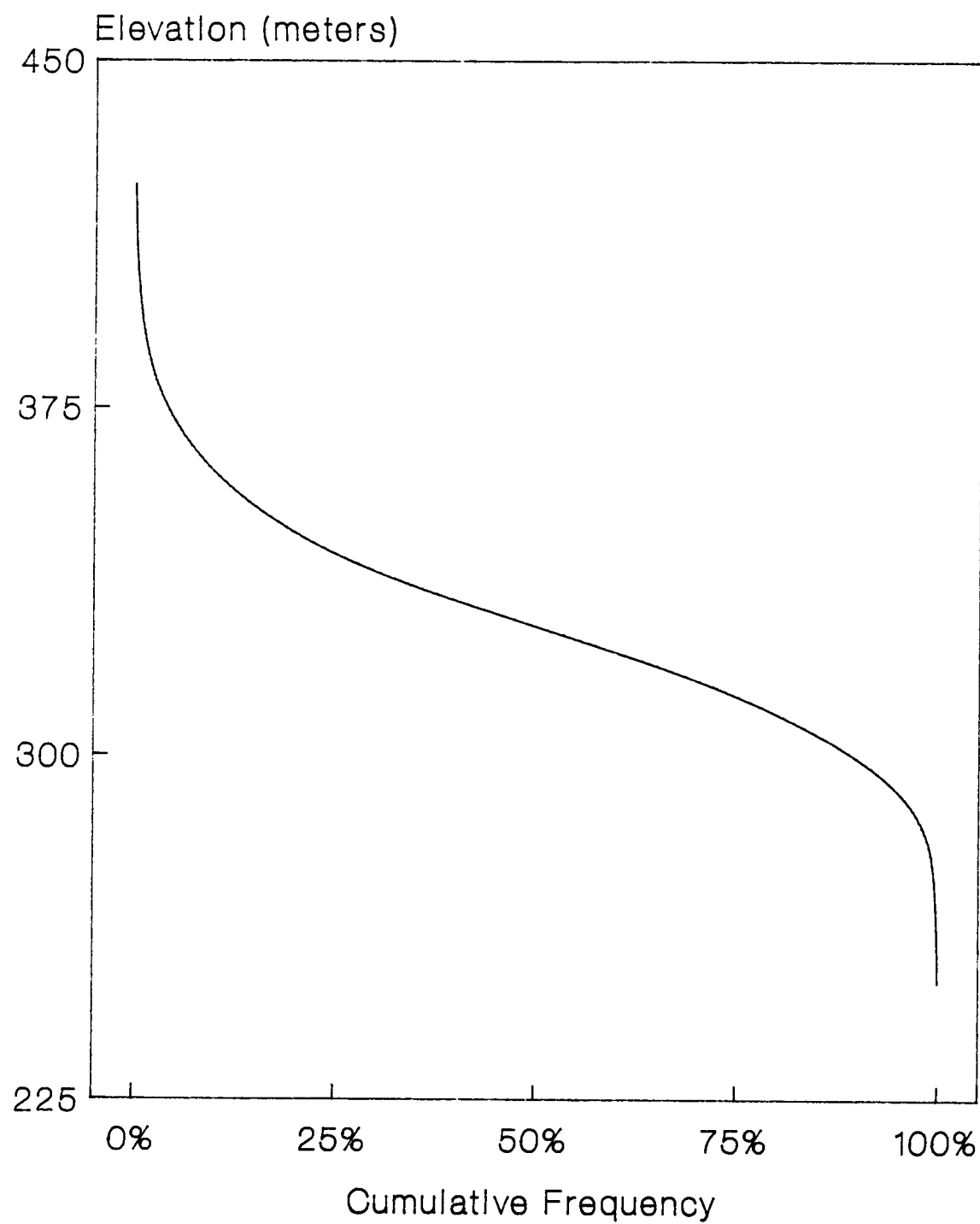


FIGURE 5.3

Hypsometric Curve - Site 24LR

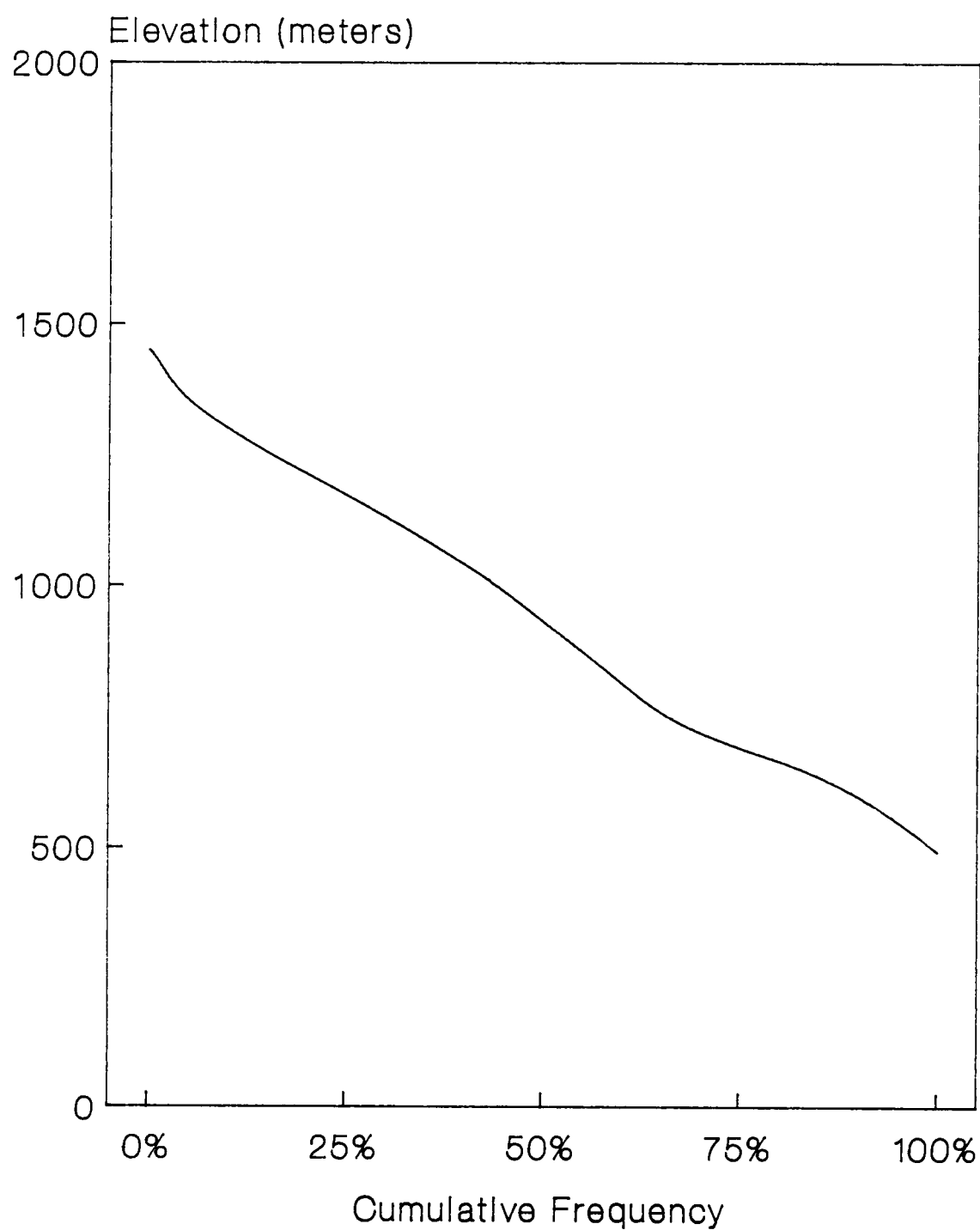


FIGURE 5.4

Hypsometric Curve - Site 24HR

inappropriate to effectively describe the relief of 250LR. Sites 250HR and 24LR portray what are considered classic s-shaped curves, as a few elevations are slightly above and below the rest of the distribution. The curve for site 24HR does not display a classic shape because of the presence of a fairly large, flat cove. The cove contains the lowest elevations and prevents the hypsometric curve from forming the s-shape tail at the seventy five to one hundred percent cumulative frequency.

Contour plots are an additional means used to characterize the landscapes of the four study sites (figures 5.5 through 5.8). A twenty five meter contour interval was used with the low relief sites, and a one hundred meter contour interval was used with the high relief sites.

Figure 5.5 from site 250LR shows a fairly level terrain with several ridges oriented northeast to southwest. A small area of higher relief can be seen in the southeast corner. Figure 5.6 from site 250HR displays the high relief characteristics of mountain topography. A large cove is found in the northwest separated from a lake to the southeast by a major ridge. Figure 5.7 from site 24LR presents low relief characteristics in the northwest and southeast. A complex ridge network traverses from northeast to southwest. Figure 5.8 from site 24HR displays

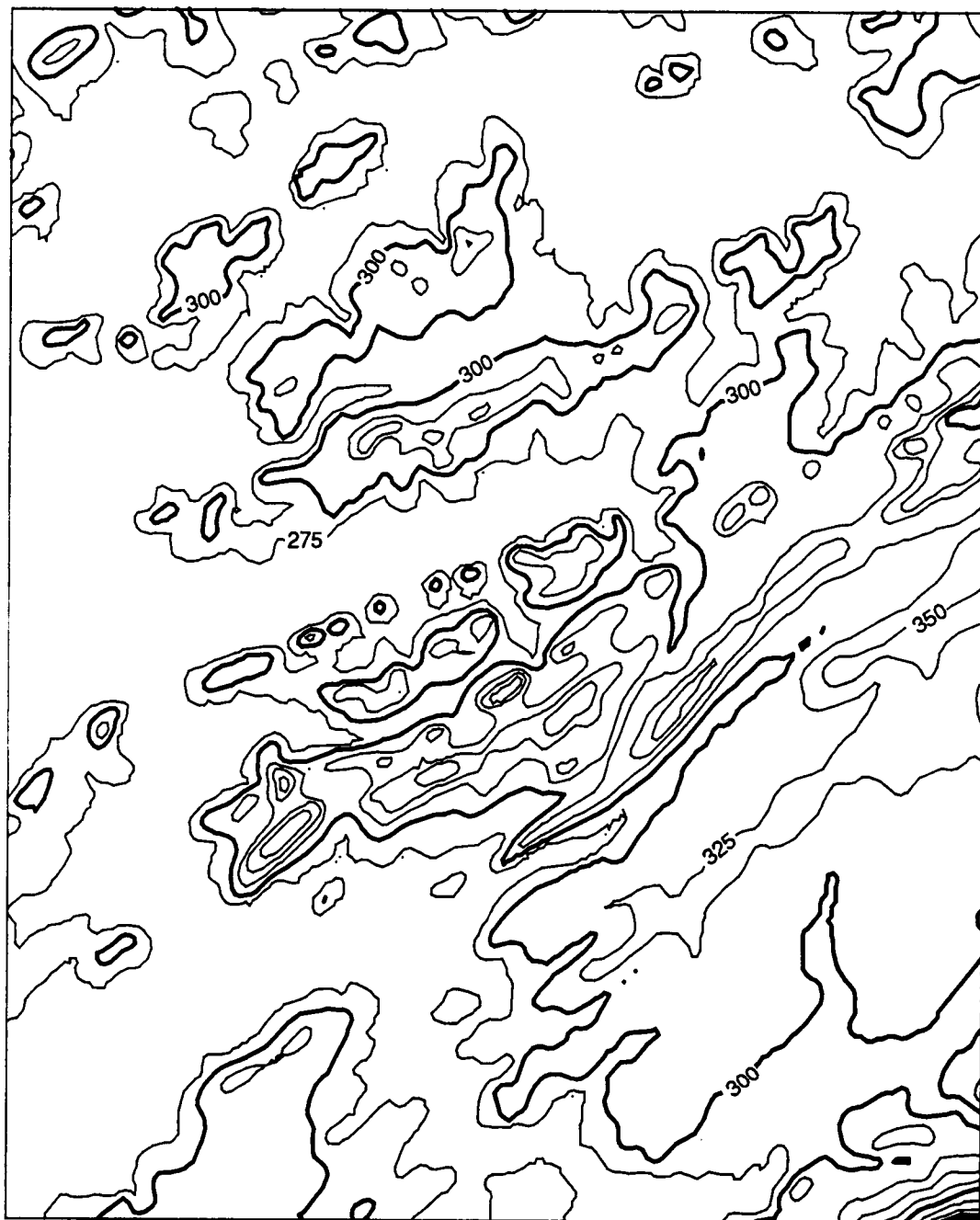


FIGURE 5.5

Contour Plot - Site 250LR (25 meter interval)

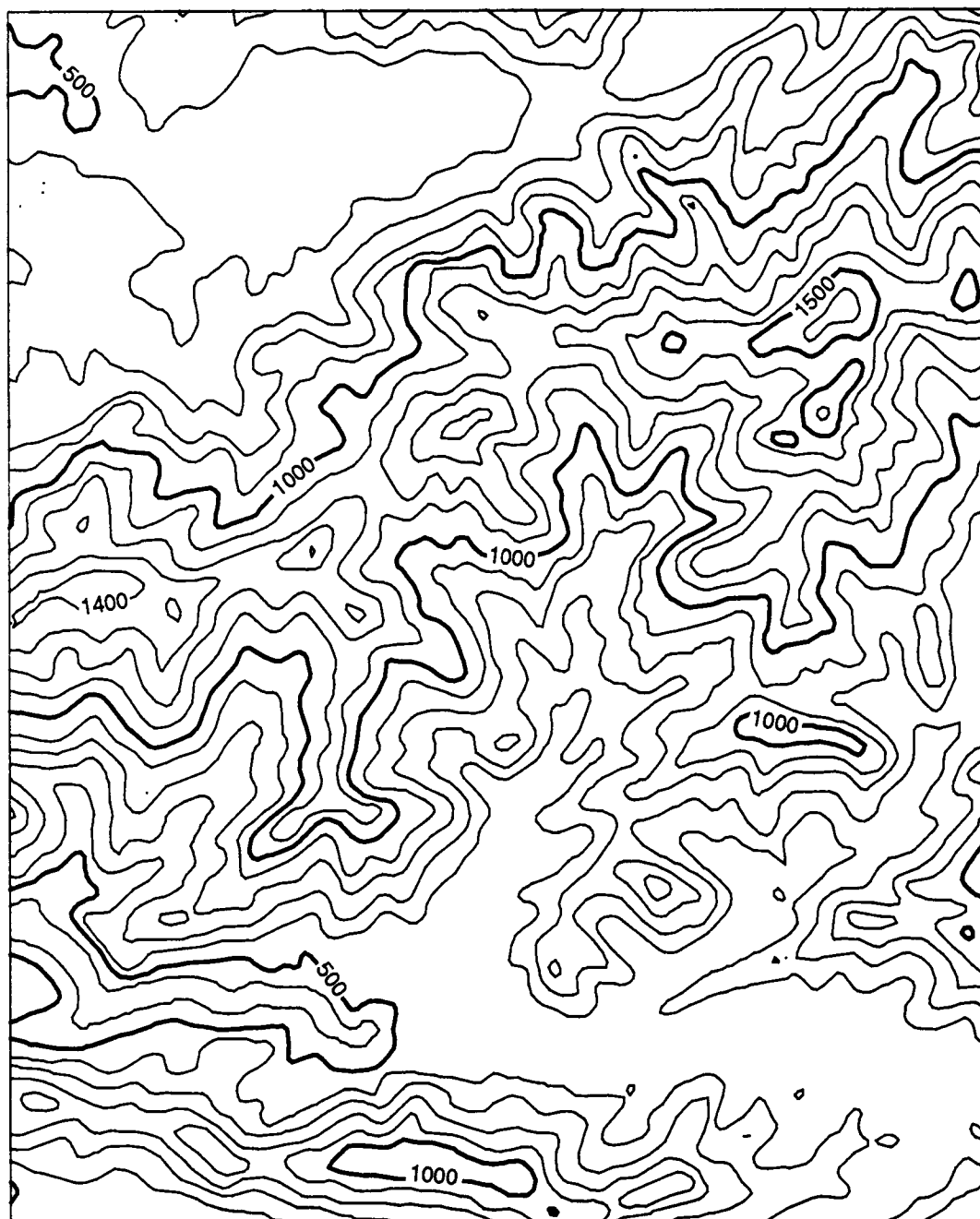


FIGURE 5.6

Contour Plot - Site 250HR (100 meter interval)

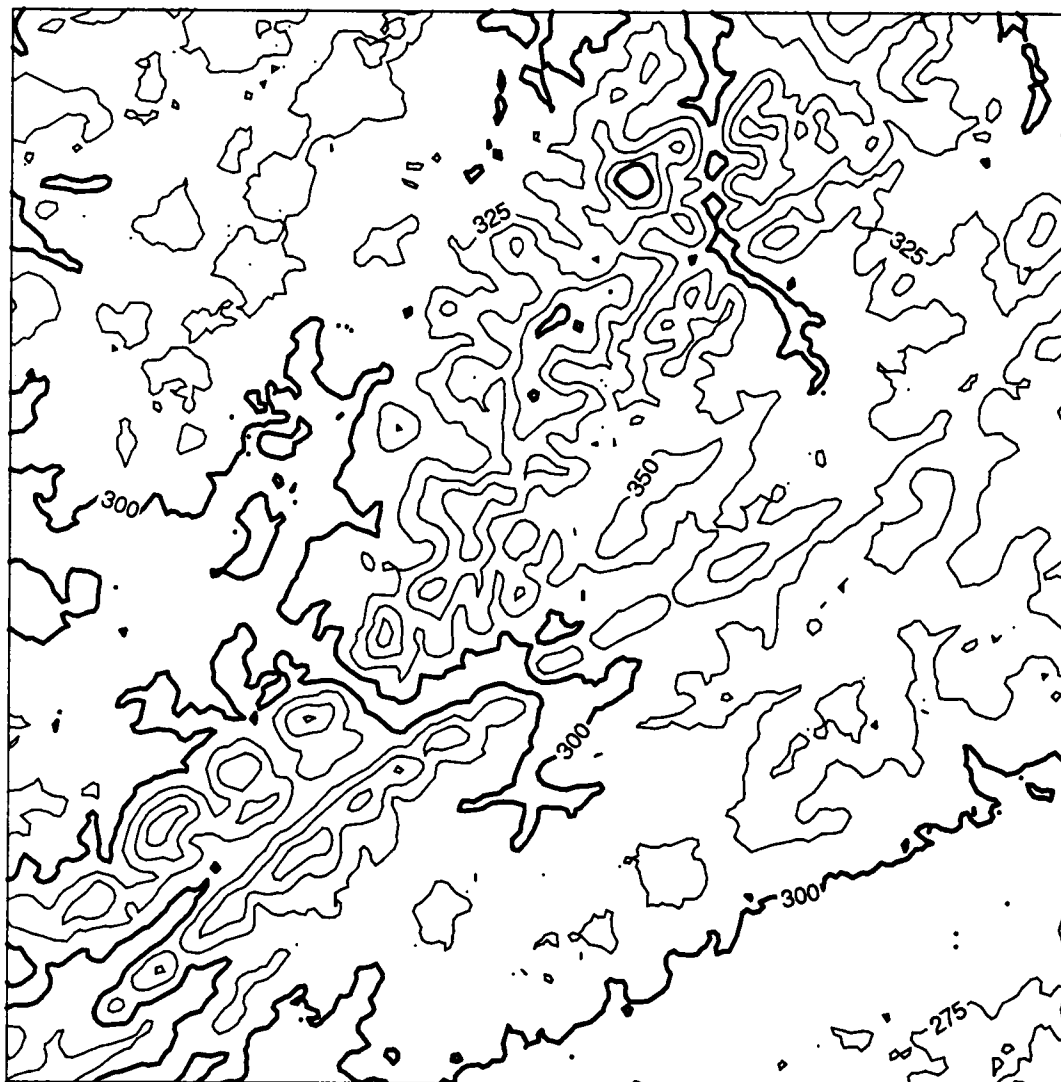


FIGURE 5.7

Contour Plot - Site 24LR (25 meter interval)

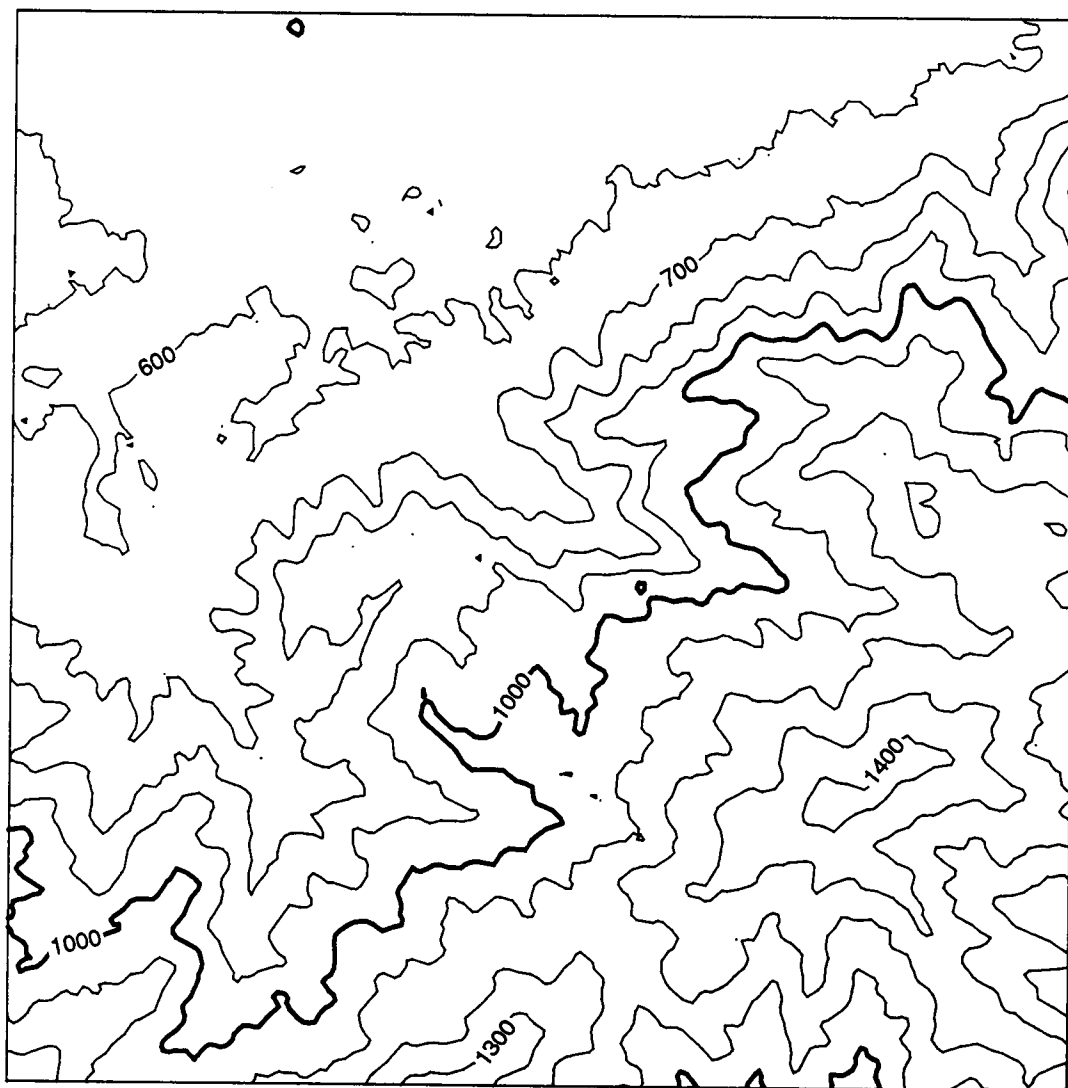


FIGURE 5.8

Contour Plot - Site 24HR (100 meter interval)

mountainous terrain to the south and southeast with the same cove as in site 250HR to the north occupying roughly a third of the site.

5.2 Similarity Between Original and Interpolated DEMs

To gain an understanding of the overall relationship of the TIN derived models to the originals, one can view the graphs of the mean and standard deviation of elevation plotted for TIN samplings of one, three, five, seven, ten, fifteen, and twenty percent. These are provided in figures 5.9 through 5.12. Mean and standard deviations were calculated for each of these TIN samplings. The percentage sampling is given along the x-axis with the mean or standard deviation provided along the y-axis. Dual y-axes were used so that the two sites from each scale could be compared on the same graph. The dual axes use the same vertical increment; they differ only in the position of the sea-level base.

For both the mean and standard deviation, the TIN models from the sites of lower relief came closer to the original values for most of the sampling percentages than did the sites of higher relief. It is interesting to note how the means of the 1:250,000 scale TIN datasets were

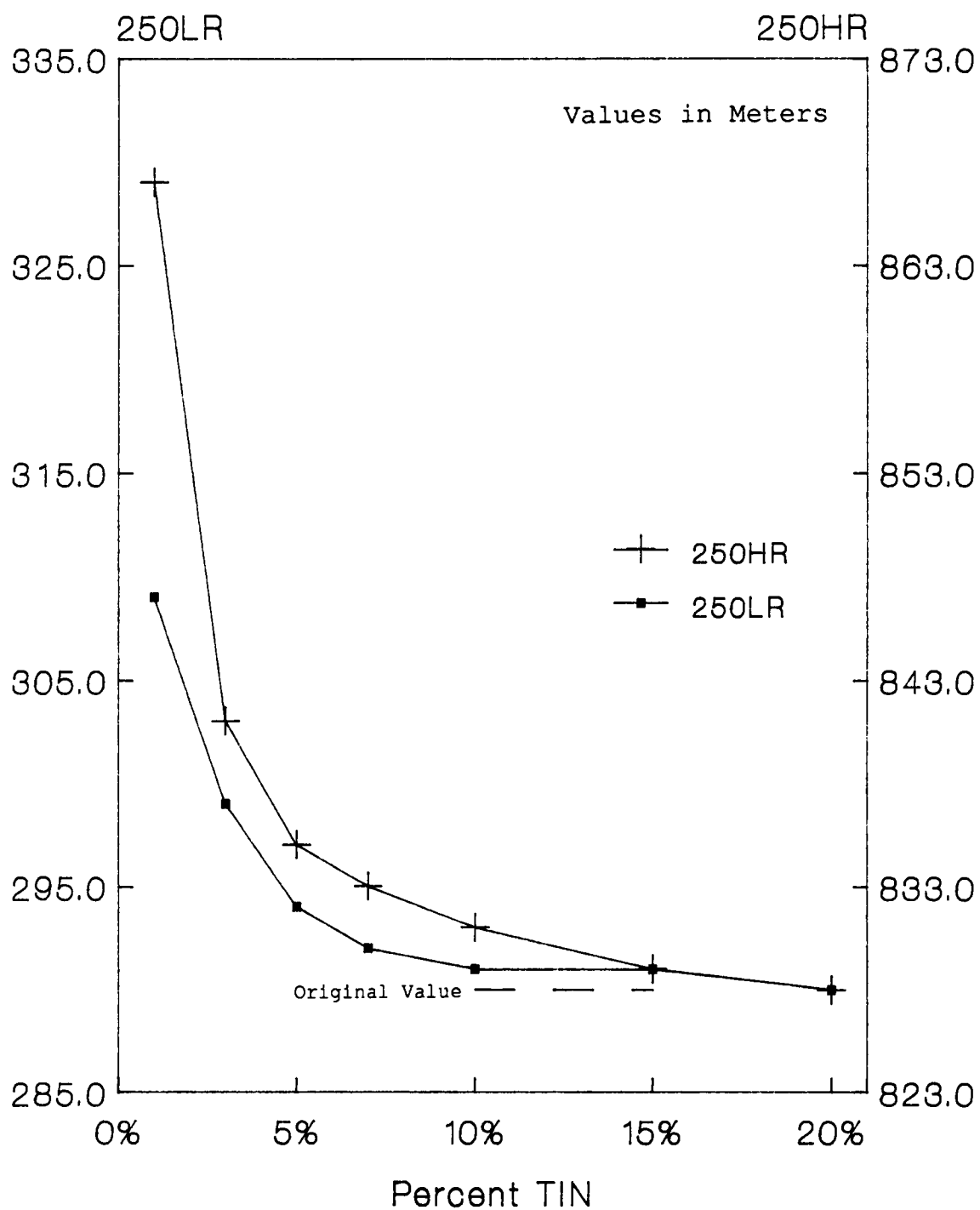


FIGURE 5.9

Mean Elevation - 1:250,000 Scale Study Sites

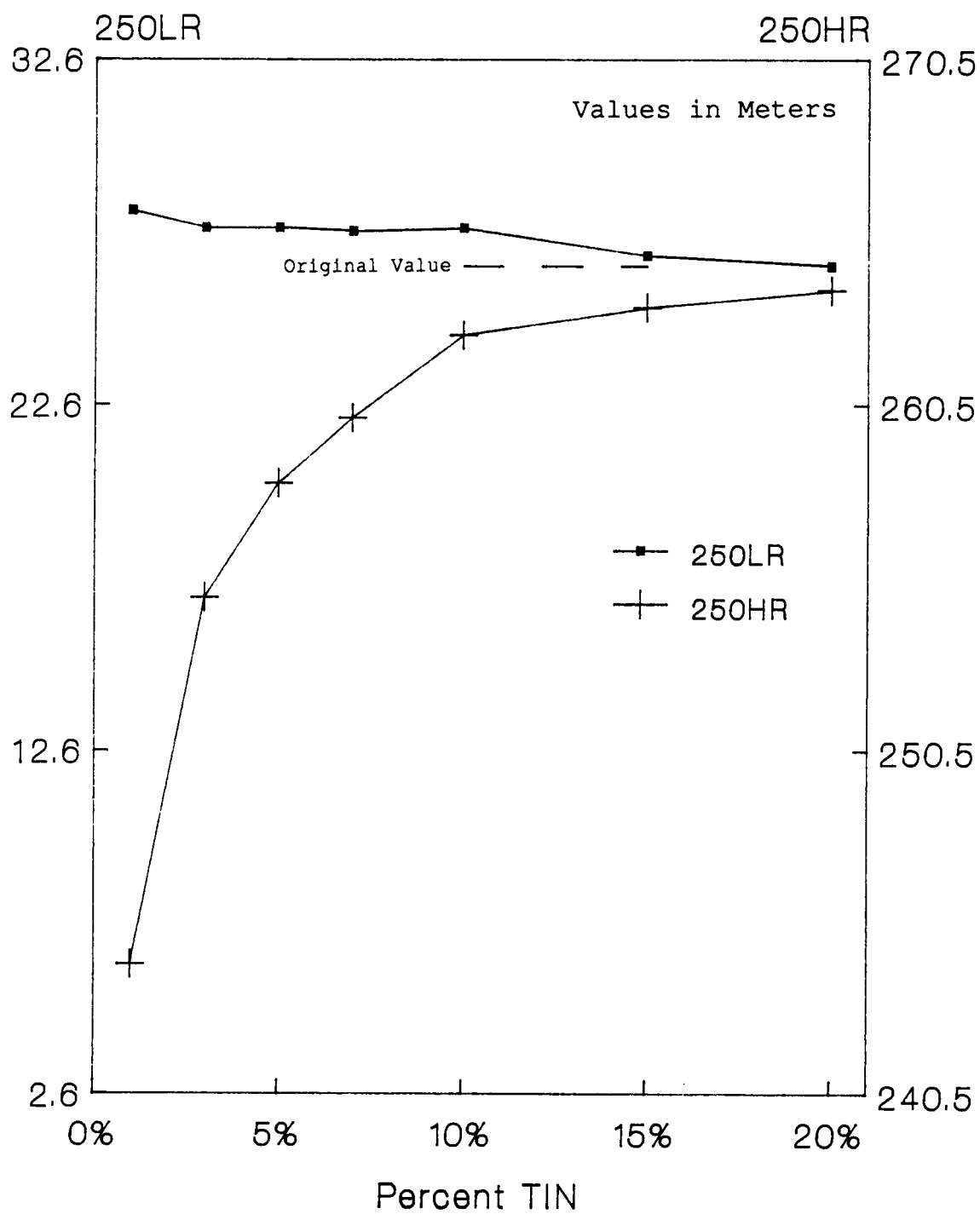


FIGURE 5.10

Standard Deviation of Elevation - 1:250,000 Scale Study Sites

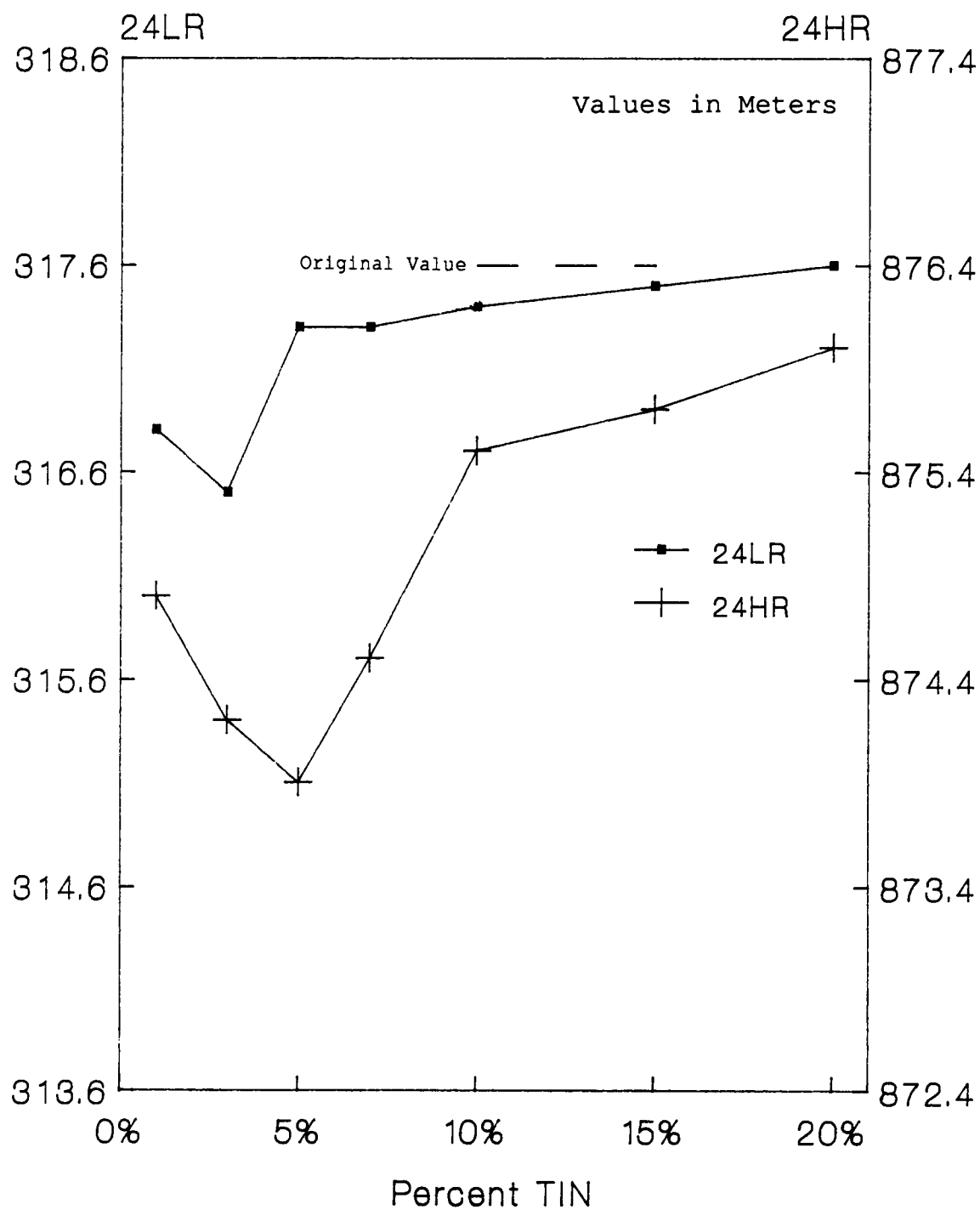


FIGURE 5.11

Mean Elevation - 1:24,000 Scale Study Sites

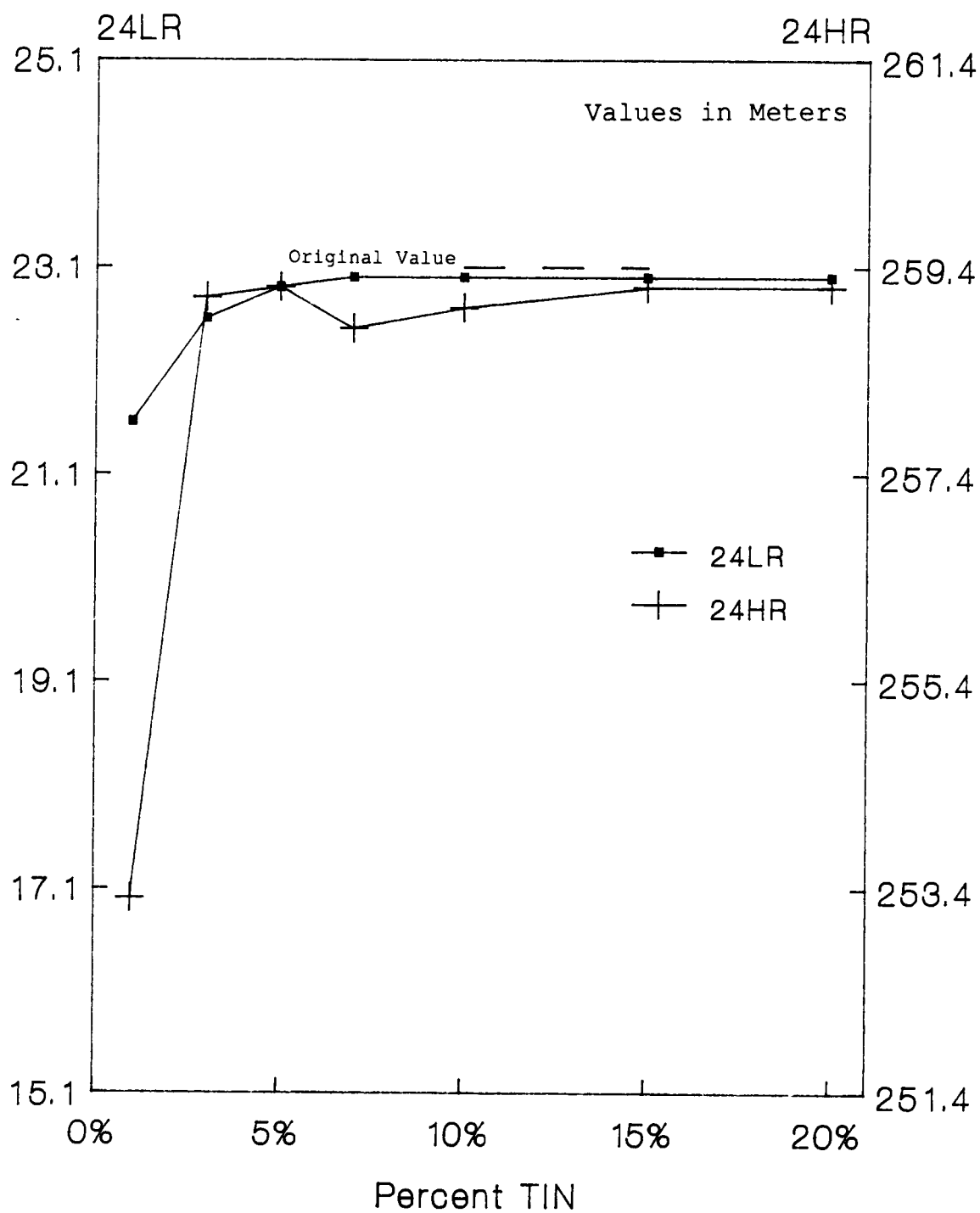


FIGURE 5.12

Standard Deviation of Elevation - 1:24,000 Scale Study Sites

initially overestimated whereas the means of the 1:24,000 scale models were initially underestimated.

Based on the means and standard deviations, the larger scale 1:24,000 TIN derived datasets displayed a closer correspondence to the original DEMs than did the smaller scale. Values ranges on the y-axis illustrates a striking difference. The range of values needed to portray the data on the y-axis is much larger for the 1:250,000 scale data. This holds true for both the mean and standard deviation. This stands in contrast to the contour maps where the larger interval was used for the high-relief areas, irrespective of the scale of the original data.

Correlation is another method to determine the relationship between the models. Correlations were calculated between gridded elevations taken from each TIN model and the original gridded elevations from which the TIN surface derived. Figures 5.13 and 5.14 show plots of the correlation coefficients plotted against the level of TIN sampling.

Although one would expect high correlations it was surprising to see such significant r-squared values. At both scales, the models quickly approached perfect correlations. The low relief sites, at both scales, started with lower correlations than did the high relief sites. It

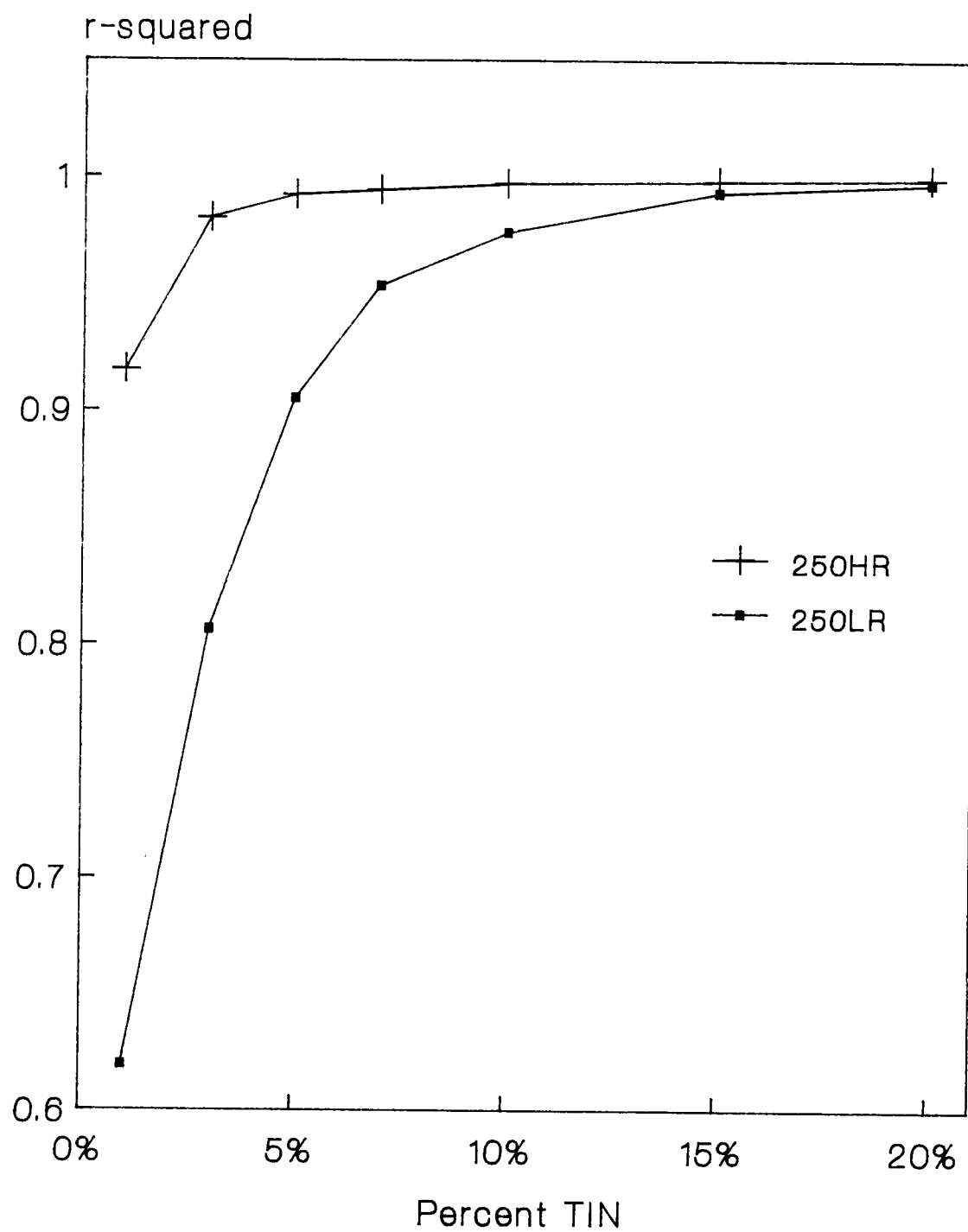


FIGURE 5.13

Correlation Between Original and Interpolated DEMs
1:250,000 Scale Study Sites

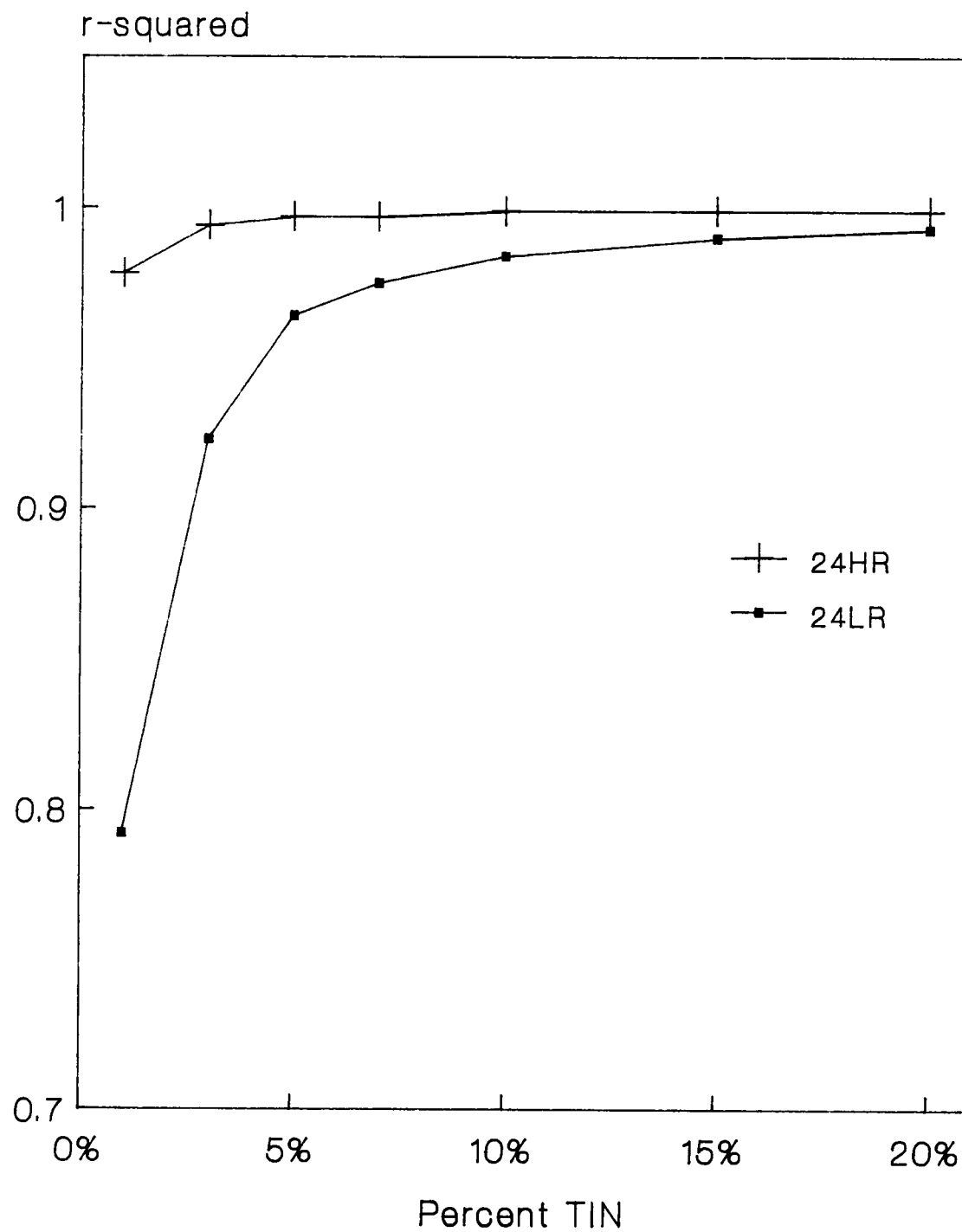


FIGURE 5.14

Correlation Between Original and Interpolated DEMs
1:24,000 Scale Study Sites

appears that these differing correlations are depicting some aspect of surface variation. Where high variations of relief in the original models exist, a matched variation exists in the interpolated models. Although the association is inexact it is calculated as a closer relationship than with the low relief models, where small departures between the interpolated and original models appear more significant.

5.3 Departures Between Original and Interpolated DEMs

Surfaces of departure were created by subtracting the original models from the TIN derived models. These new datasets were examined with similar statistics as those used previously. The absolute maximum, mean, and standard deviation were calculated from each departure surface and graphed. Figures 5.15 through 5.20 show these distributions.

In terms of the maximum absolute departure, sites 250LR and 24LR, the low relief areas, showed lesser departures. The ten percent samples taken from the 1:250,000 scale datasets had equal departures because of an uncharacteristically high maximum absolute departure value for site 250LR. The trend of lesser departures with the low relief areas is not surprising as there is less room for

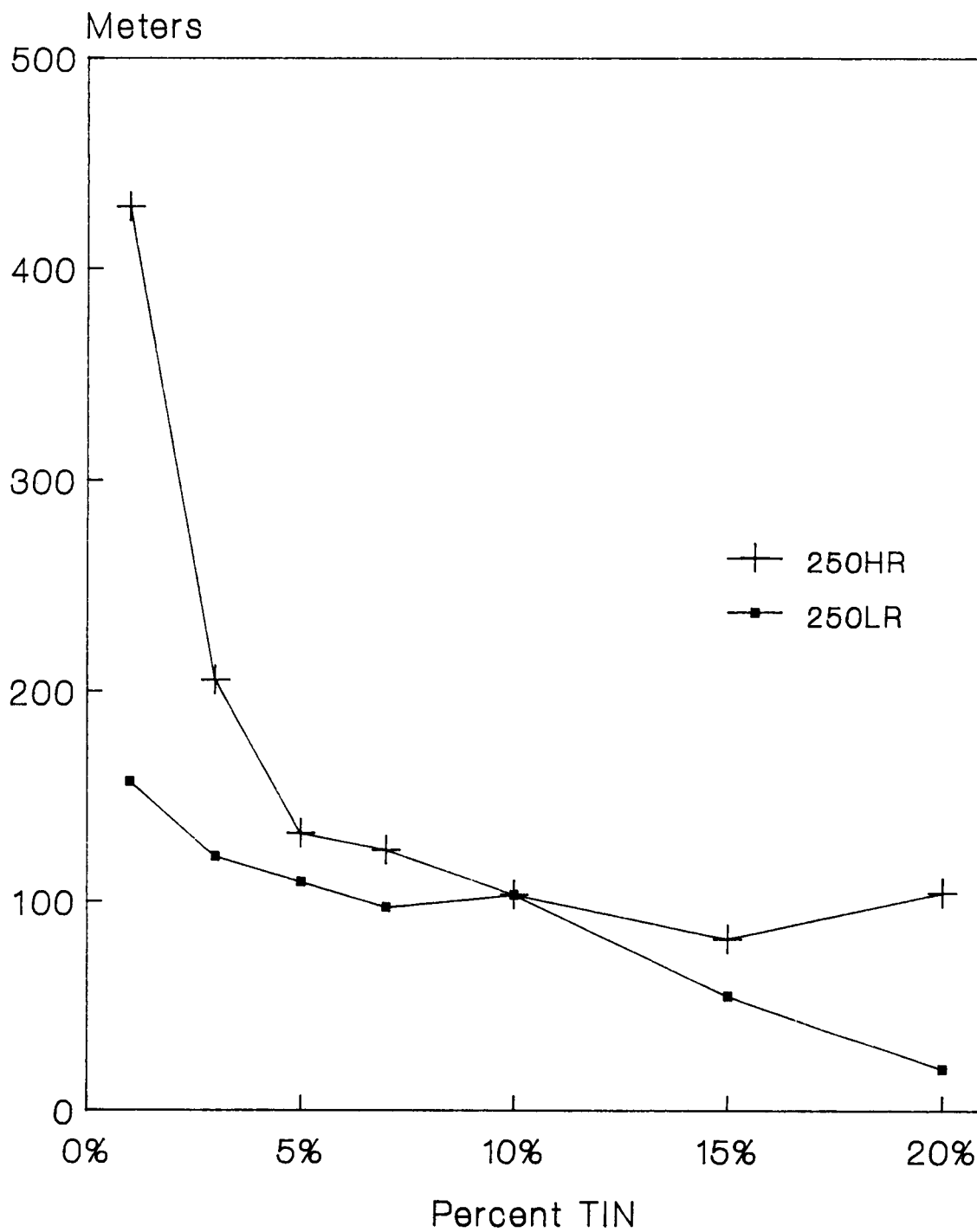


FIGURE 5.15

Maximum Absolute Departure - 1:250,000 Scale Study Sites

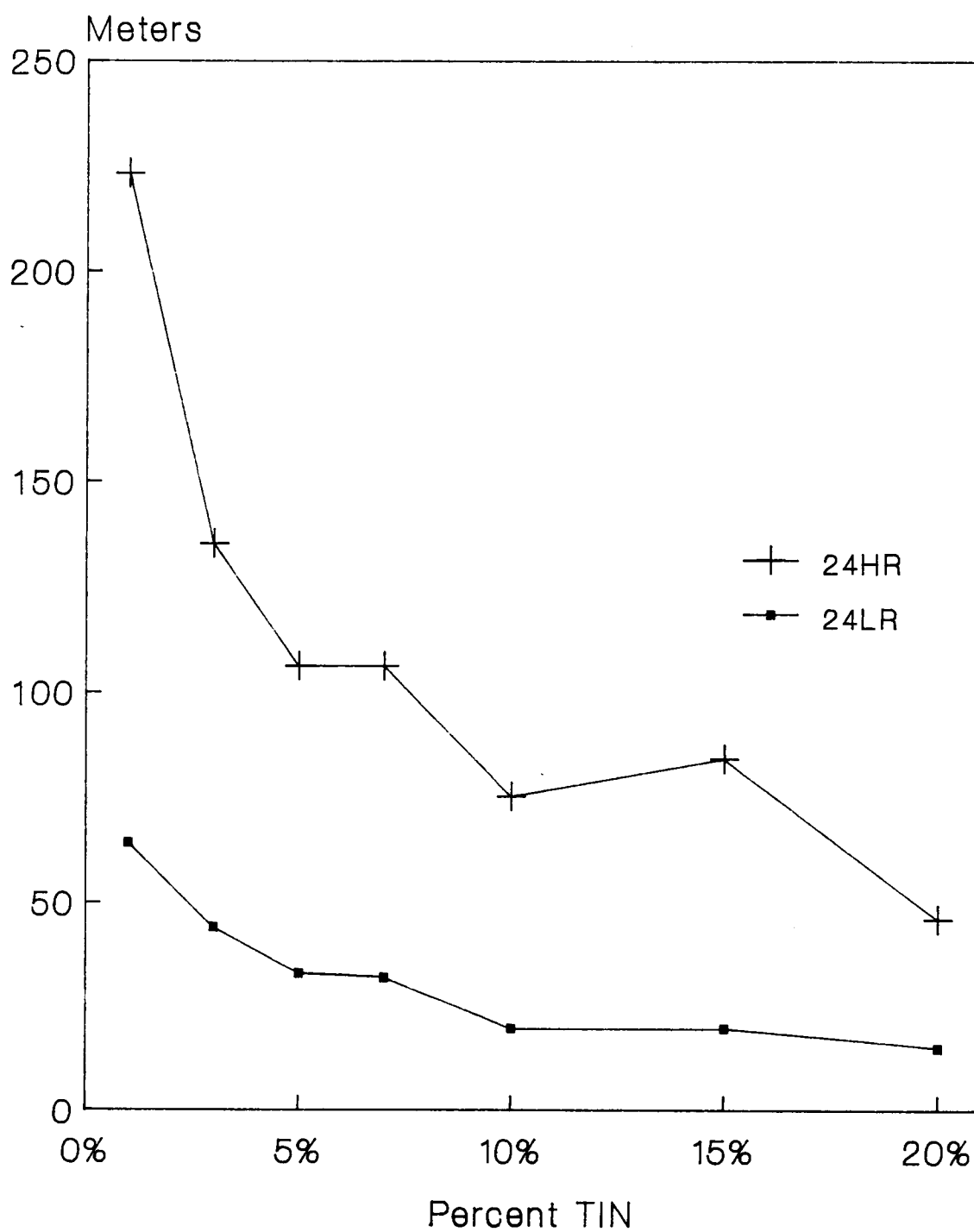


FIGURE 5.16

Maximum Absolute Departure - 1:24,000 Scale Study Sites

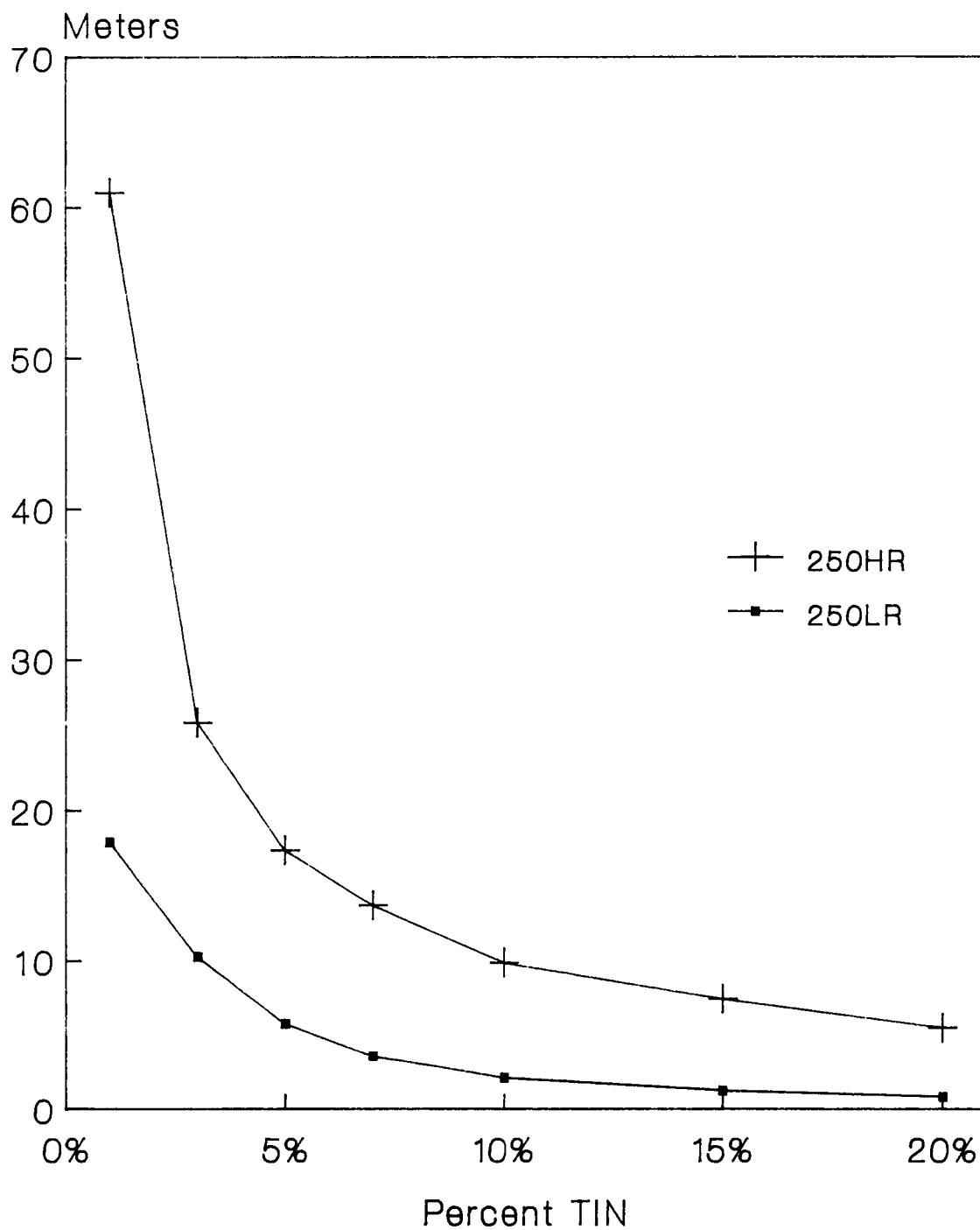


FIGURE 5.17

Mean Absolute Departure - 1:250,000 Scale Study Sites

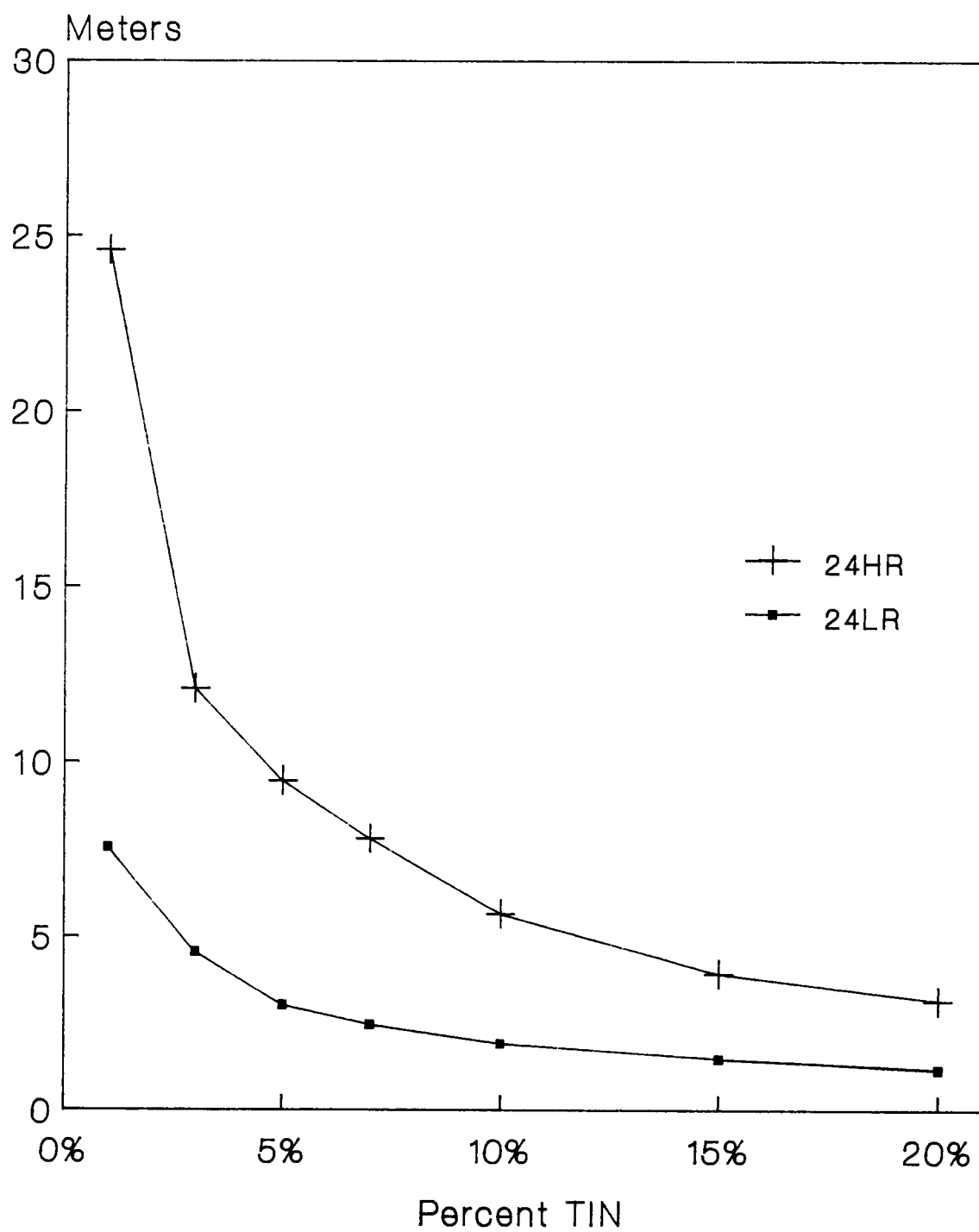


FIGURE 5.18

Mean Absolute Departure - 1:24,000 Scale Study Sites

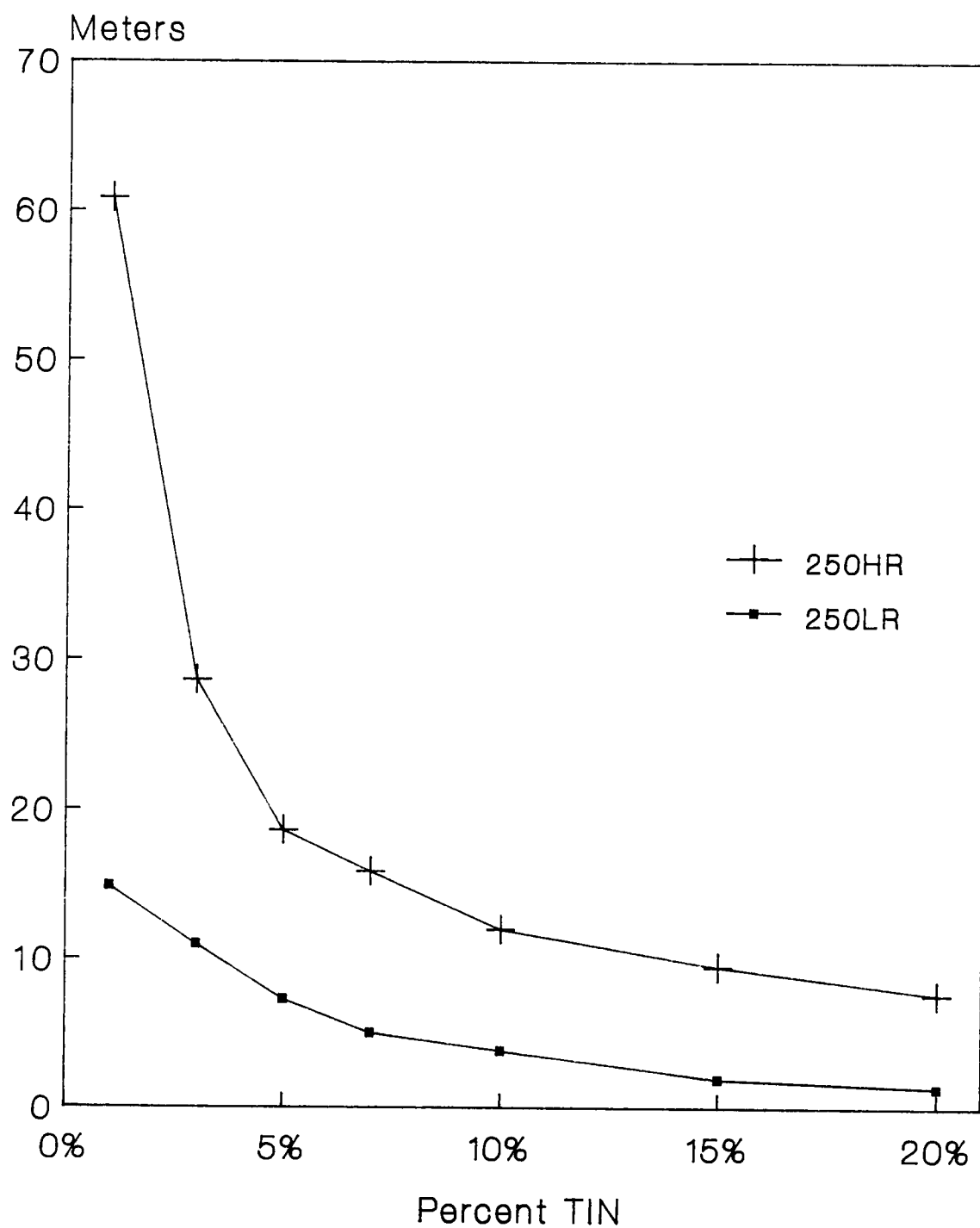


FIGURE 5.19

Standard Deviation of Departure - 1:250,000 Scale Study Sites

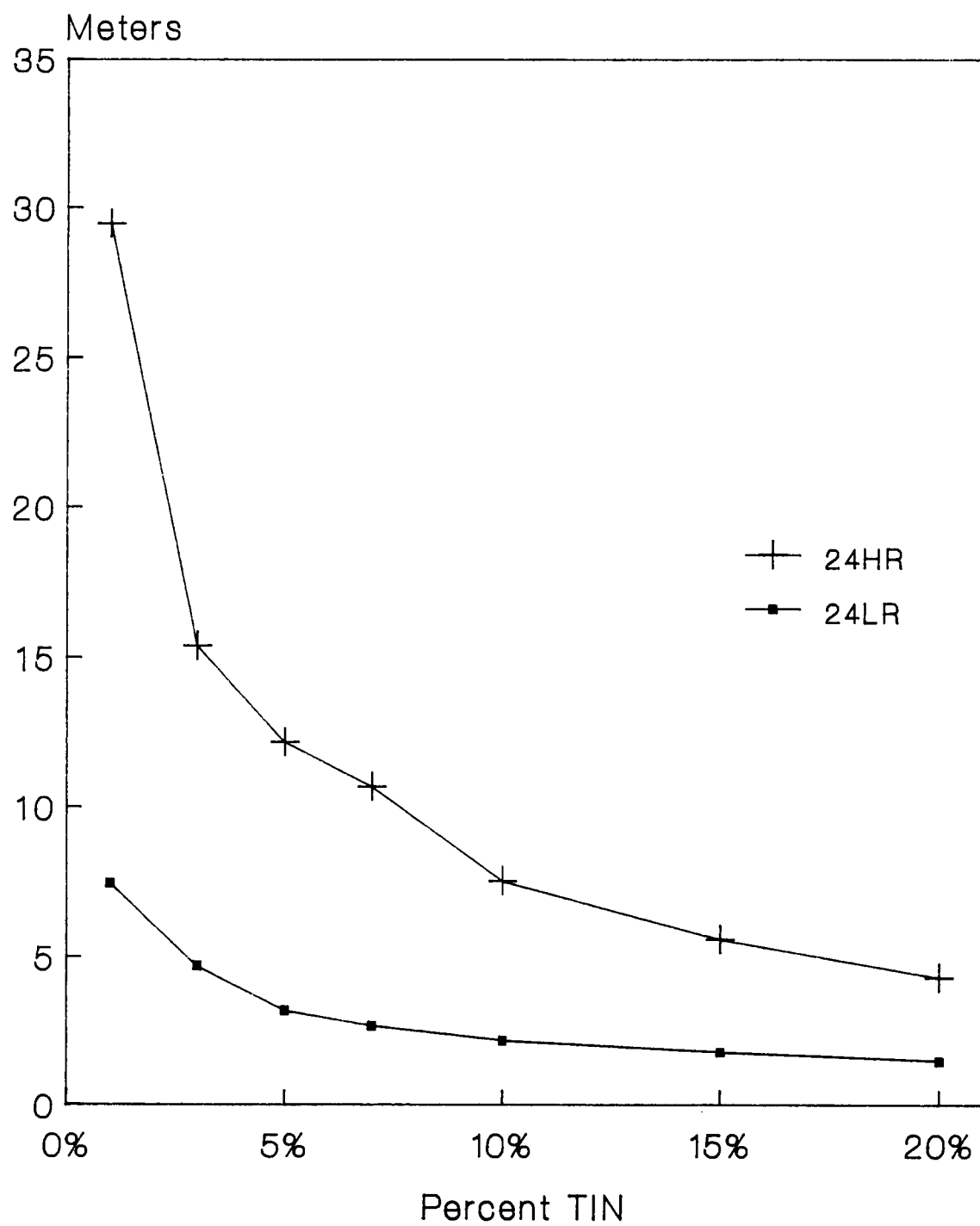


FIGURE 5.20

Standard Deviation of Departure - 1:24,000 Scale Study Sites

error in these datasets when compared to the two datasets from sites of high relief. The range in elevation is not as great.

The 1:24,000 scale models had lower departures than those from the 1:250,000 scale data for both smooth and rough terrain. This should be expected because the distance between sample points from the larger scale data is less than the smaller scale data. The variation in elevation between sample points with the 1:24,000 scale data would therefore also be decreased providing less margin for error.

Although the charts of absolute departure show that generally less departure exists with a larger percentage of points employed to define the TIN, the trend is somewhat erratic. This is because the measures of absolute departure are dependant on the extreme values on the surface defined by the gridded elevations. The charts of the absolute mean departure, found in figures 5.17 and 5.18, and absolute standard deviation of departure, found in figures 5.19 and 5.20, show a more consistent trend. The graphed curves of the low relief sites are flatter than the areas of higher relief.

The previous graphs of departure were based upon absolute values and provided insight to the overall degree of departure. An additional set of charts was generated to

display the distribution of departure as related to over or under estimation of original values. Only the one, five, ten, and twenty percent sampling statistics are shown in figures 5.21 through 5.24 to keep from having too many lines.

The 1:250,000 scale study site distributions of 250LR and 250HR are markedly different in some aspects and similar in others. The low relief site has a much tighter distribution pattern, especially when approaching higher sampling percentages. Over fifty thousand observations are found to place within plus or minus a one meter departure with the twenty percent sample. In contrast, site 250HR has only about half that at the same sample percentage. Both models display a large degree of positive departure with one and five percent samples. This trend continues to a lesser degree with higher percentage samples.

The 1:24,000 scale study sites exhibit a more normal distribution pattern than the 1:250,000 scale sites. Study areas 24LR and 24HR are not as skewed as their 1:250,000 scale counterparts. Evidence of skewness may still exist but it appears to have only a slight tendency of negative departure. The low relief site 24LR did not fare as well with ten and twenty percent samples as did site 250LR, but was better at the lower sampling percentages. The high

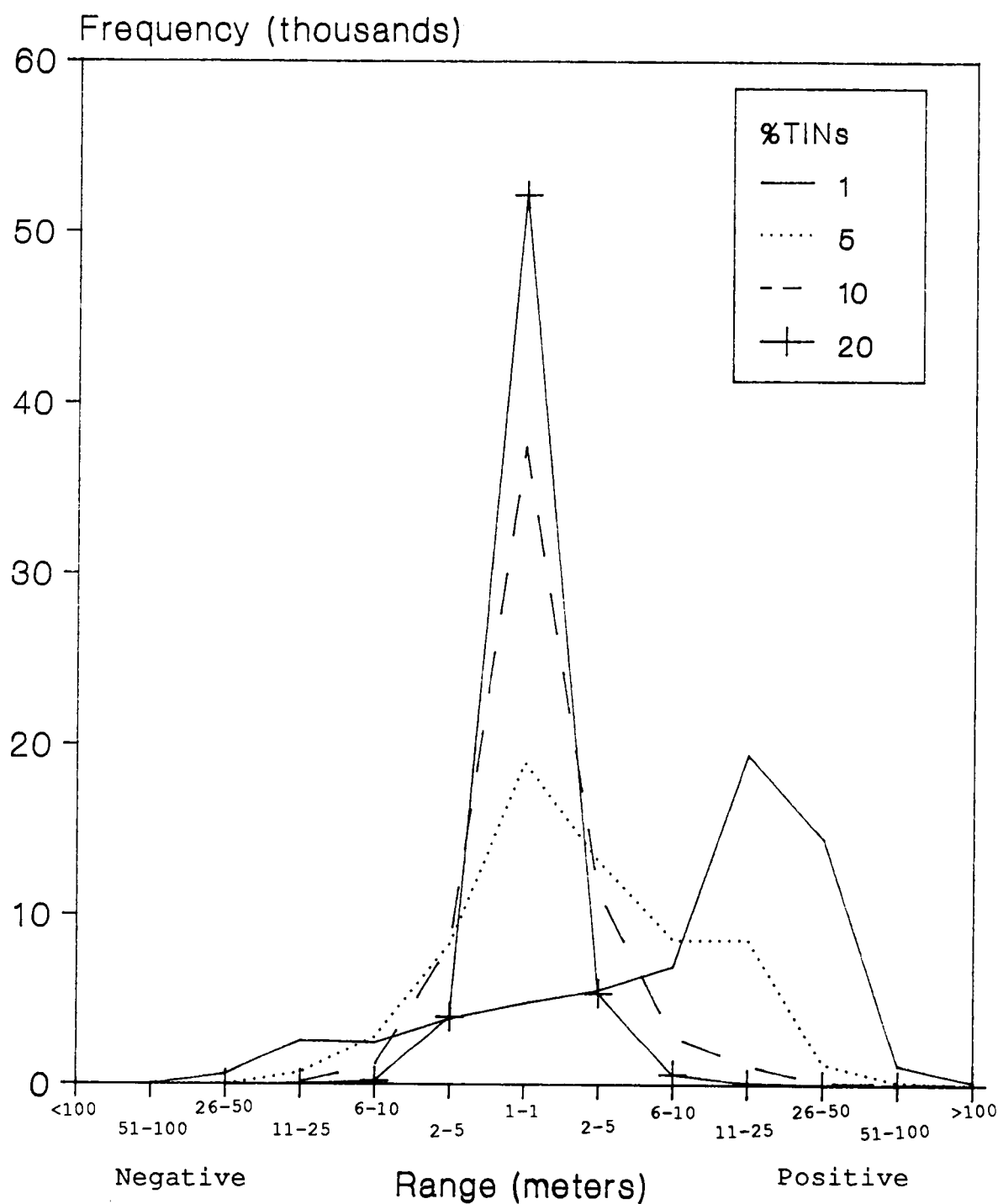


FIGURE 5.21

Distribution of Departure Between Original
and Interpolated DEMs - Site 250LR

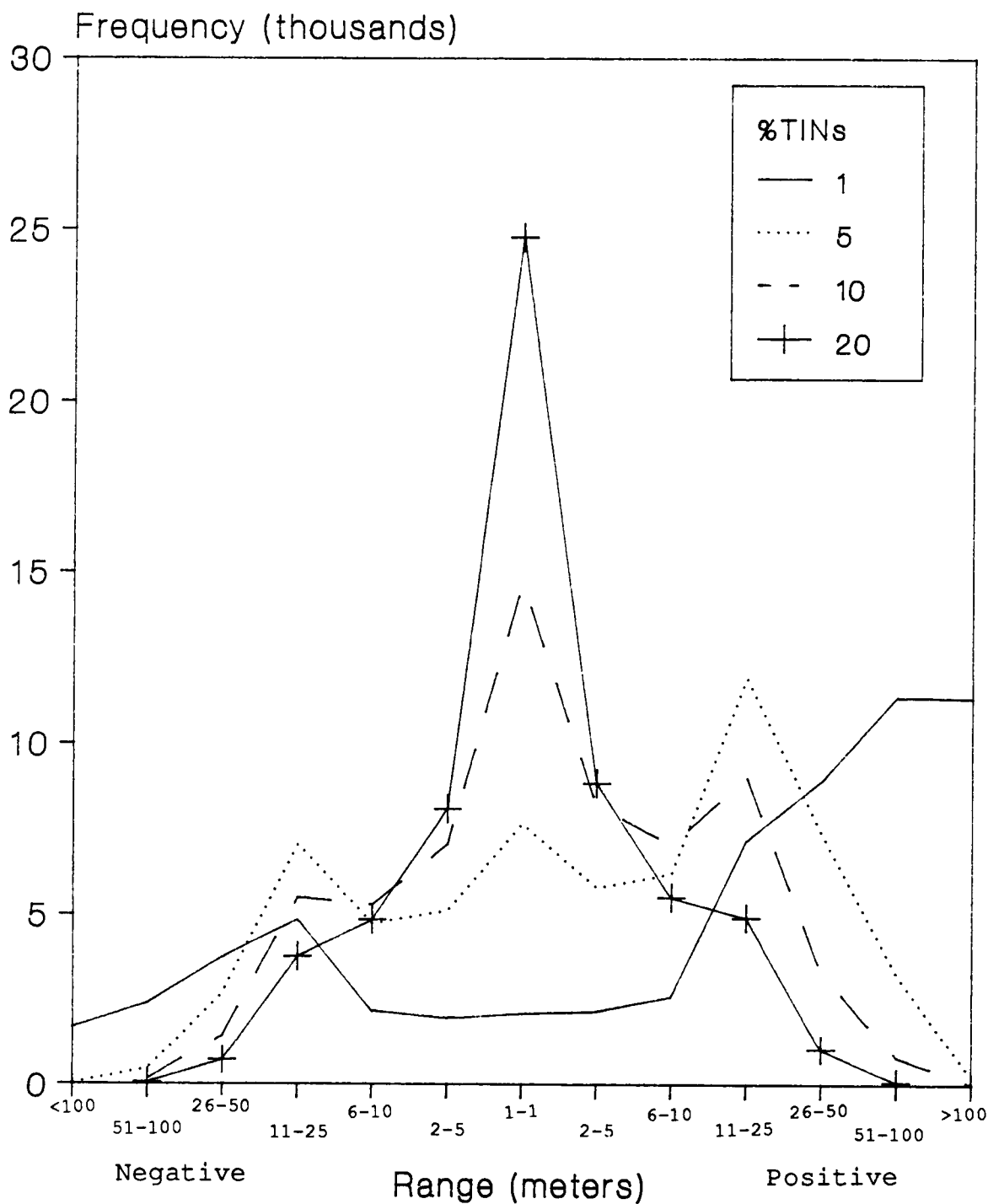


FIGURE 5.22

Distribution of Departure Between Original
and Interpolated DEMs - Site 250HR

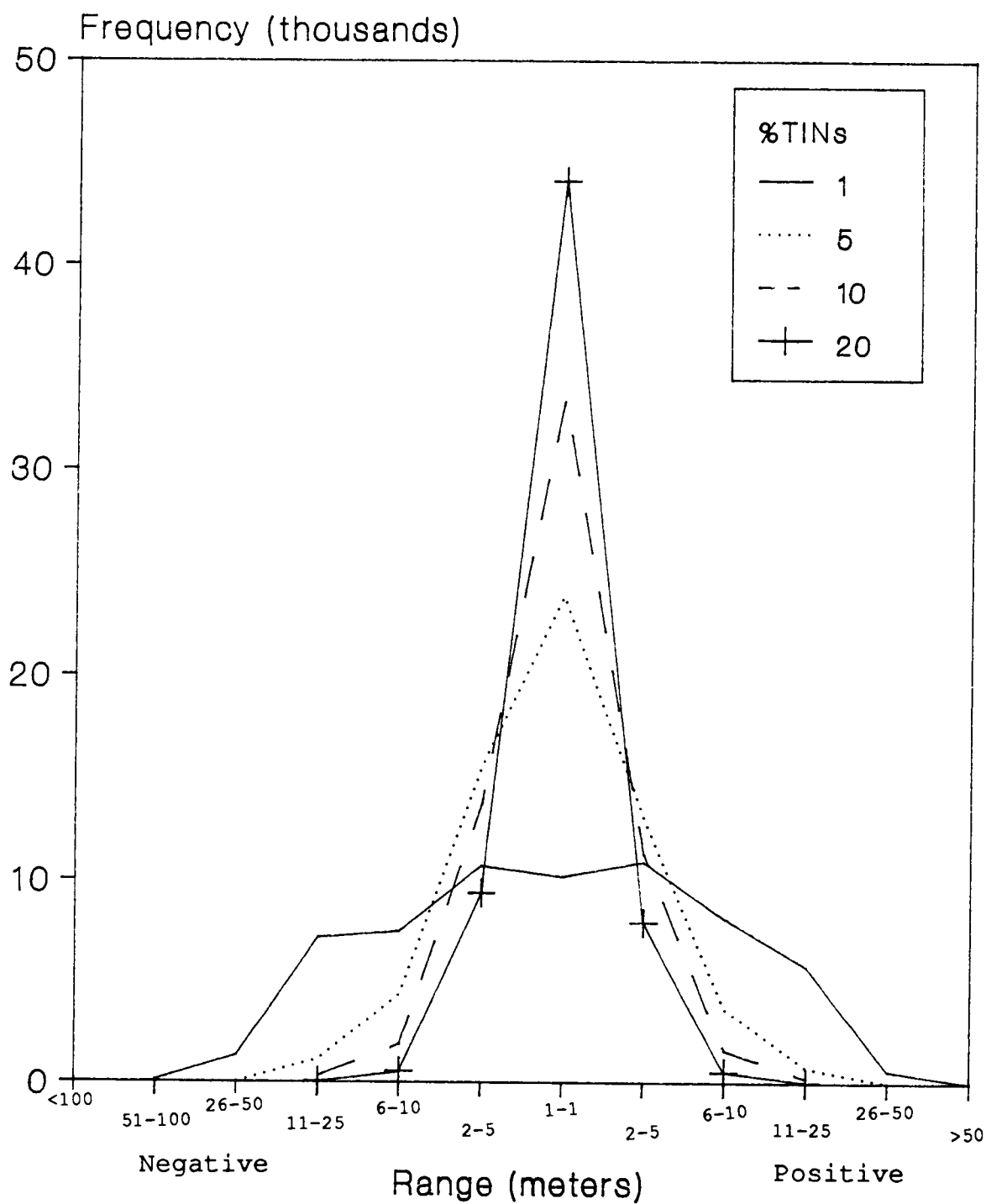


FIGURE 5.23

Distribution of Departure Between Original
And Interpolated DEMs - Site 24LR

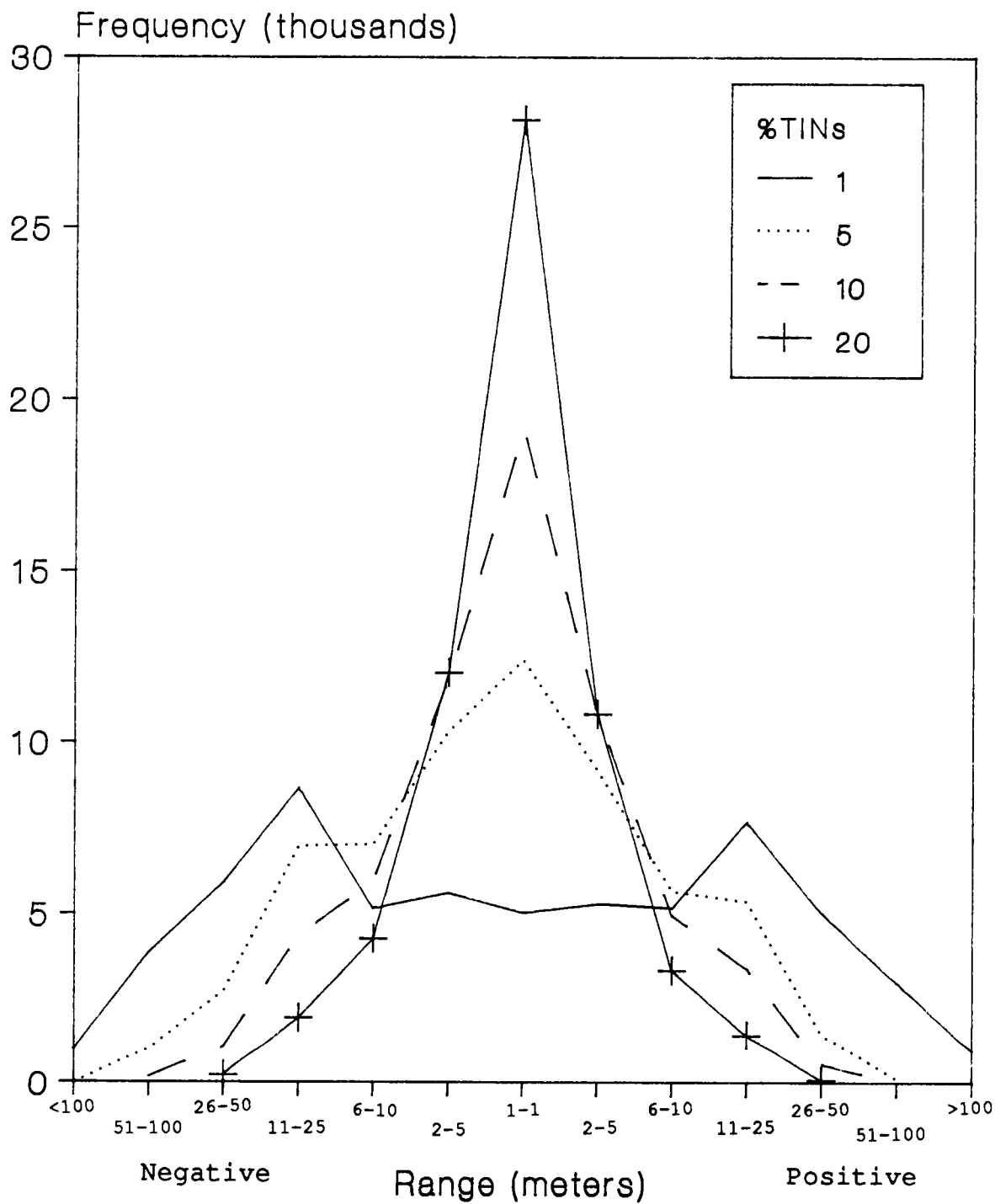


FIGURE 5.24

Distribution of Departure Between Original
and Interpolated DEMs - Site 24HR

relief site 24HR had a superior distribution compared to its counterpart, site 250HR, in all sampling percentages.

5.4 Comparison of Surface Structure

In order to evaluate how well the TIN derived models maintained the surface structure as defined by the original models, maps of ridge form were created. While the previous two sections dealt with aggregate values, this methodology permits one to measure how well the recreated grids depict these geographical features. When comparing the ridge form of the TIN derived grids to their sources, one can make a qualitative judgement as to which sampling percentages are acceptable and which are not.

Maps of ridge form were created for each of the original DEMs and their TIN derived counterparts. After reviewing the results of all TIN derived models, the author decided to include only three, five, seven, and ten percent samples. This was partly because all one percent samples showed unacceptable results. Ridges were either poorly delineated or were often oriented in different directions than those defined by the originals. In addition, ridge form maps made from fifteen and twenty percent samples showed little improvement over the ten percent samples. The

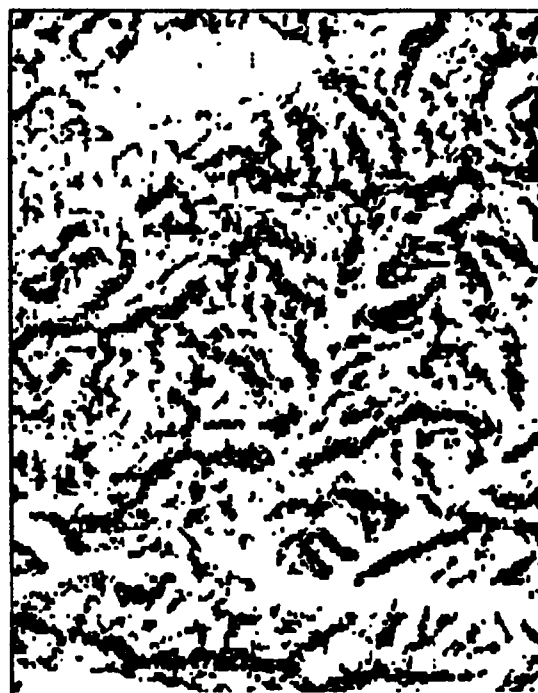
originals are located in figure 5.25 with the TIN derived models following in figures 5.26 through 5.29.

Site 250LR contains a series of ridges traversing from the northeast to the southwest. The larger ridges present in the center of site 250LR are apparent in all TIN derived models (figure 5.26). They are less defined in the one and five percent samples. These two models also have ridges in inappropriate locations and directions and are therefore unacceptable. The ten percent sample is very similar to the original leaving no question that it adequately maintains the ridge structure of the original. The seven percent model is similar and has a closer relationship to the ten percent model than the five percent model. Seven percent is therefore considered satisfactory.

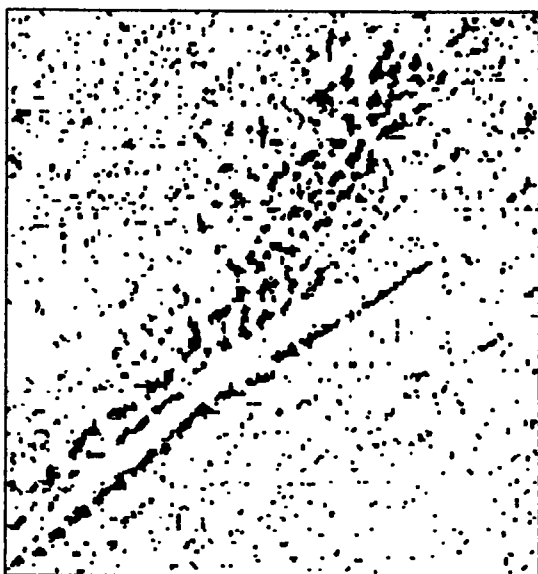
The original ridge form map of 250HR (figure 5.25b) is significantly different from all of its derivatives (figure 5.27). The ridge areas are broader and less linear than the ones depicted in all of the four TIN derived models. This is a function of the triangle edges in the TIN models defining the ridge lines and the nature of the algorithm used to define ridge-form. Because of this, and the fact that the pattern is so complex, a comparison between them is difficult to make. As with 250LR, all of the major ridges are represented in each TIN model. They are better



a. 250LR



b. 250HR



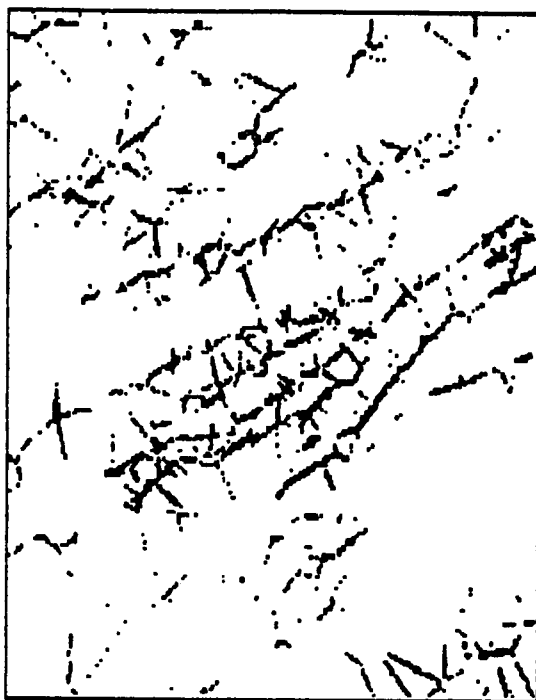
c. 24LR



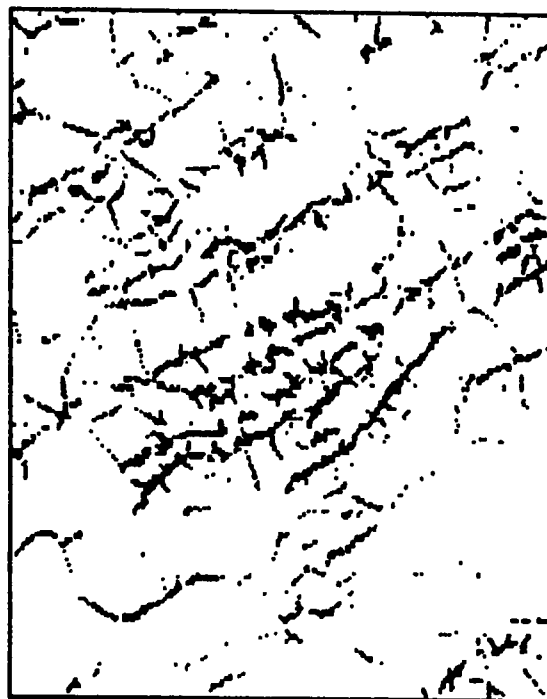
d. 24HR

FIGURE 5.25

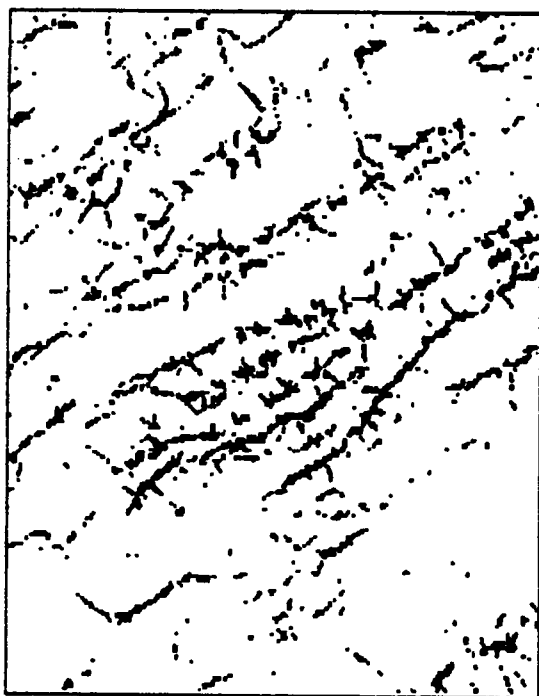
Ridge-Form of Original DEMs



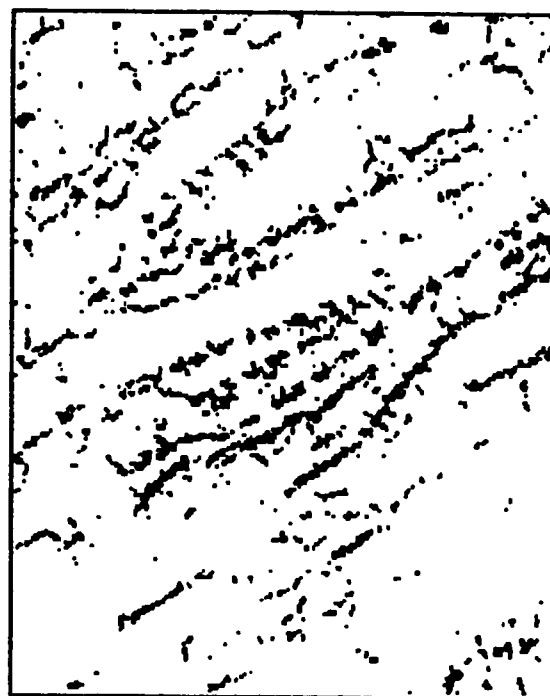
a. 3%



b. 5%



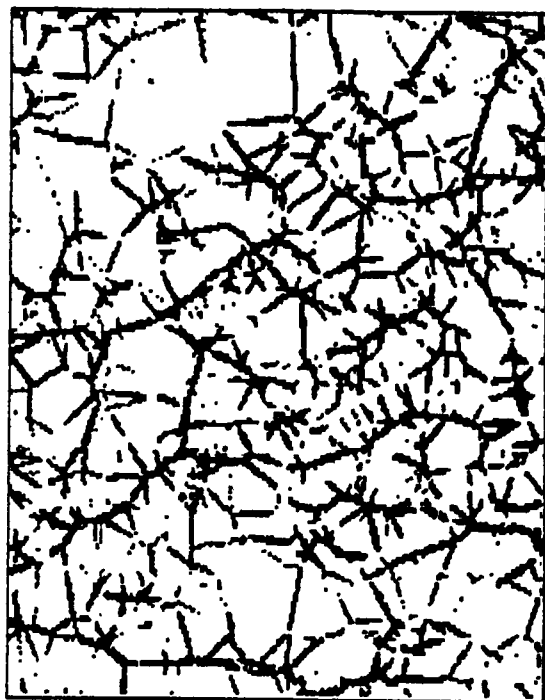
c. 7%



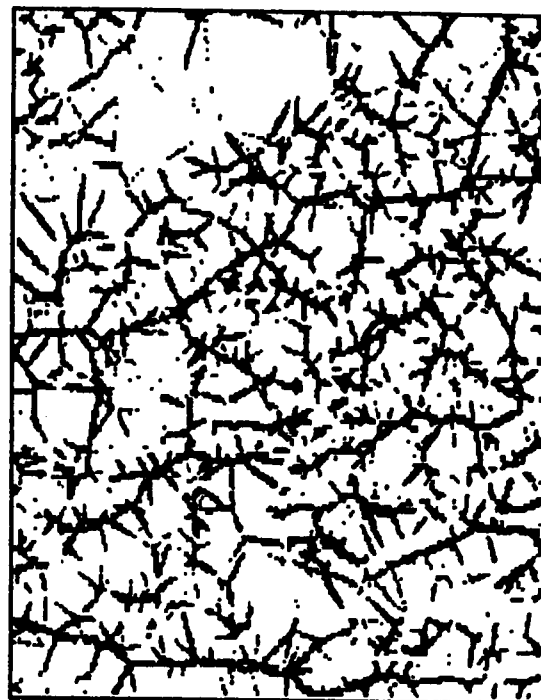
d. 10%

FIGURE 5.26

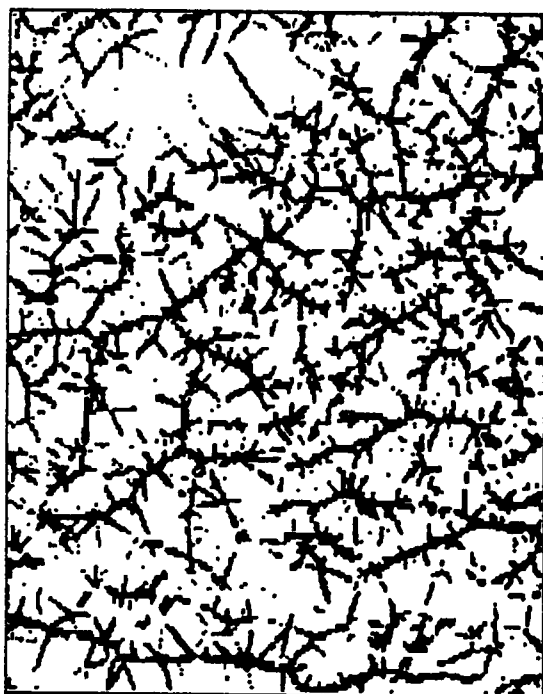
Tin Derived Ridge-Form of Site 250LR



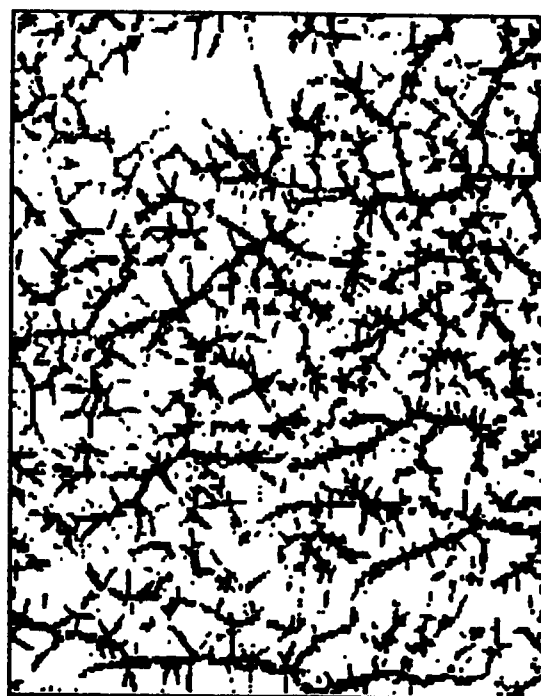
a. 3%



b. 5%



c. 7%



d. 10%

FIGURE 5.27

TIN Derived Ridge-Form of Site 250HR



a. 3%



b. 5%



c. 7%



d. 10%

FIGURE 5.28

TIN Derived Ridge-Form of Site 24LR



a. 3%



b. 5%



c. 7%



d. 10%

FIGURE 5.29

TIN Derived Ridge-Form of Site 24HR

delineated here than in the low-relief model. The ridges are more pronounced in this area of high relief. The most important differences between these models is level of detail. It decreases as sampling percentage is reduced. The seven or ten percent samples would be acceptable based on detail of ridge form with ten percent containing the most detail. Anything below seven percent would be missing at least some of the second order ridges.¹

For the most part, site 24LR exhibits a poorly defined ridge form pattern. Only one sharp, well defined ridge, can be found oriented from the northeast to southwest. The other ridges, slightly north of this and oriented in basically the same direction, are not well connected but are highly associated. They belong to a broad raised area of land. The remainder of the study area contains scattered, isolated cells, flagged as ridges to the northwest and southeast. All TIN derived models (figure 5.28) show the two major ridge areas. The sharper ridge to the south is acceptably represented by all four models but the intricate pattern to the north of this is not well articulated. As the ten percent sample provides the best representation of a derived model available here, a comparison to it can be

¹The major ridges in a study site are considered first order. Ridges extending from those are second order.

made. The seven percent sample is very similar and therefore acceptable. The five percent model is borderline and the three percent sample is too generalized.

Site 24HR, like 250HR, portrays the complex ridge pattern associated with mountainous terrain. An interesting difference between site 250HR and 24HR is how the ridges are portrayed. Site 24HR's ridges tend to be more distinct. This may be a function of the larger scale model's ability to provide a more refined definition of the ridges. As with site 250HR, it is difficult to make comparisons between the original 24HR site (figure 5.25d) and its TIN derivatives (figure 5.29). None of the recreated grids exhibit the detail present in the original model. They only approximate it by capturing the major features, with even some of them being weeded out at lower sampling percentages. The models below seven percent would have to be deemed unacceptable in this light. Seven percent is borderline with noticeably more detail still present in the ten percent sample.

5.5 Summary of Findings

Arc/Info TINs derived from USGS gridded DEMs exhibit varying degrees of quality. This quality is dependant primarily on sampling percentage. The ability to recreate a topographic surface diminishes as the number of sample

elevations used to define the surface decreases. Terrain morphology and data source scale influence quality as well, but to a lesser degree. Smoother surfaces and grid points sampled at closer intervals yield better results. All of this is consistent with sampling theories.^{2,3}

A TIN sampled from a relatively smooth gridded surface will contain fewer deviations from the original model than one using the same percent sample from a rough surface. This is because less variation is found in the smoother surface. This should be expected, as TINs are designed to use fewer points in defining a surface where fewer significant breaks in slope exist.

Data gathered at larger scales, with points sampled at closer intervals, produce TIN models with less deviation. This is because elevations between neighboring points in large scale models are more similar. There is a higher degree of spatial autocorrelation between points which are closer together. In the examples used in this research, the

²Ayeni, O. O. (1982), "Optimum Sampling for Digital Terrain Models: A Trend Towards Automation", Photogrammetric Engineering and Remote Sensing, Vol. 48, No. 11, pp. 1687-1694.

³MacEachren, Alan M. and Davidson, John V. (1987), "Sampling and Isometric Mapping of Continuous Geographic Surfaces", The American Cartographer, Vol. 14, No. 4, pp. 299-320.

1:24,000 scale data had points sampled at every thirty meters. The 1:250,000 scale data had points sampled at approximately every seventy five meters east to west and ninety two meters north to south. Thus, triangles formed by points an equal number of observations apart cover different areas when comparing one scale to the other. Triangles from the larger scale data will cover smaller areas with lower variations in elevation thus providing smaller departures from the original surface.

At both scales and degrees of relief in the research, TIN quality as measured by surface variation, departure, and structure improved dramatically with an increase in sampling percentage. One percent samples displayed poor results with improvements apparent in each step up in sampling percentage. Improvement tended to diminish in strength, especially in excess of ten percent samplings. The twenty percent sample showed significantly similar surface variations, low levels of departure, and similar definitions of ridge form with all models. This can be viewed as the point of diminishing returns with the TIN models faithfully representing the original surfaces with little generalization. Samples above this would show little improvement. When some generalization is desired, lower sampling percentages should be used. The fact that a twenty

percent sample occupies twice the computer storage space as an original model also warrants use of a smaller sample percentage.

The question arises as to what level of generalization is acceptable. How low a percentage can be used before the TIN no longer satisfactorily represents the original surface? To answer this, an examination of a combination of previous results is necessary. No one evaluative technique can answer the question. The most revealing and meaningful measures of TIN quality appear in the graphed distributions of both signed and absolute departures and the maps of ridge form.

The author suggests that both a normal, non-skewed, distribution of departure and preservation of ridge form would be necessary for a model to be acceptable. Departures should not be biased, positively or negatively, and the recreated grids should be able to recover significant aspects of the original landscape. The following presents the minimum percentages required for each site based on this criteria. Site 250LR would require at least a seven percent sample, site 250HR a ten percent sample, site 24LR a five percent sample, and site 24HR no lower than a seven percent sample. It should be noted that the standard used here to define a model's level of acceptance is arbitrary and

subjective. Criteria would best be established within a context relative to a particular need.

TIN models derived from USGS gridded DEMs sampled at low percentages exhibited high degrees of quality. The quality varied between different scales and degrees of surface variation. The author feels that sampling percentages from gridded DEMs can be kept to a minimum while maintaining much of the integrity of the original models. Lower sampling percentages are needed from smooth surfaces than rough surfaces. If study areas comprise both types of terrain, a percentage should be picked which will satisfy the minimum accuracy requirements of the rougher surface. The actual percentages used would be best decided upon with knowledge of specific study criteria.

CHAPTER VI

CONCLUSION

It is important to evaluate the effects on the quality of data in all conversion steps. The majority of elevation models in existence are in gridded format but the TIN model is rapidly gaining acceptance. The conversion of one model to the other is sometimes desirable and appropriate. While efforts have been made to identify errors contained within elevation models, little attention has focussed on determining the effects of converting from one model to another.

TIN models are arguably superior to gridded models in terms of computational efficiency and topographic representation. Growing support of the TIN in automated systems is being observed. Many systems provide gridded DEM to TIN conversion routines. While a TIN cannot possibly define a surface any better than the gridded DEM from which it was derived, it approximates the source and will do so with varying degrees of success. This success is closely associated with the number of critical points taken from the gridded models used to construct the TINs. To justify the use of a minimal number of critical points in creating a

TIN, some measure of the resulting quality of TINs is essential.

This research has presented evaluations made on the quality of Arc/Info TINs derived from USGS gridded DEMs. For this study, four gridded matrices of elevation were used: two from the 1:250,000 scale DEMs in the three arc-second format and two from the 1:24,000 scale DEMs at a thirty meter grid spacing. TIN models at varying levels of generalization were created from these matrices using Arc/Info routines. These models were then converted back to gridded matrices of elevation using a simple linear interpolation. The TIN derived gridded surfaces were then compared, pixel-by-pixel, to the original gridded DEMs to determine data losses in the conversion process. Because a linear interpolation procedure was used, the potential introduction of error in this step was considered to be minimal. Variance statistics were then employed to provide an overall assessment of the quality of each interpolated matrix derived from a TIN at each level of generalization. Further, maps of ridge-form were created to evaluate how well each TIN derived model captured significant features of the landscape in comparison to the original gridded DEM. The determination of whether a given model was acceptable for use was based on these comparisons.

Results from the four study sites indicate that samples of critical points from the gridded DEMs of between five and fifteen percent produced acceptable TINs depending on the complexity of the land surface. The small number of points needed proves the redundancy inherent in the gridded models. The sites containing relatively smooth surfaces required fewer points to define adequately the topography in the TIN format than did the sites containing rougher surfaces.

While the methodology performed in this research produced meaningful results, areas of further study are still warranted. Much of the evaluation depended on statistics which provided a single descriptive number for the entire surface of any one particular dataset. For example, the standard deviation of elevation was used as one measure of comparison. This provided a comparison of total surface variation between models, but did not address variations in local effects. No spatial assessment of quality could be presented with such global measures. To assess spatial variation in data quality the maps of ridge-form were generated. It is felt the assessment of TIN quality over space should be an area of more research.

The assessment of TIN quality in this study was subjective, with no criteria being established to enable the user to accept or reject any particular model. Acceptable

levels of TIN quality will vary between different types of applications. If a user were able to specify the minimal landform features to be retained in the terrain model, acceptance/rejection criteria could be made. It is possible that the measures of TIN quality, as presented in this work, may be insufficient to identify all quality concerns for some applications. On the other hand, a creative combination of these measures may indeed be effective, sufficient to give a user a basis to accept or reject a particular level of reduction in a TIN model.

This research has presented a set of tools which is useful in comparing elevation models. It was the intent of the author to provide a number of generic routines that could be incorporated into an integrated system, or used independently. Previous work performed in the field of general geomorphometry provided the framework for the development of the procedures. The traditional descriptions of terrain models were extended to provide insight on the differences between terrain models.

The successful integration of diverse data sources in geographic information systems requires that data documentation and integrity be maintained. Studies on the effects of analytic operations and conversion processes on spatial data still are not well developed. It is hoped this

research provides additional knowledge and insight into this area.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Akima, Hiroshi (1978), "A Method of Bivariate Interpolation and Smooth Surface Fitting for Irregularly Distributed Data Points, ACM Transactions on Mathematical Software, V. 4, No. 2, pp. 148-159.
- Ayeni, O. O. (1982), "Optimum Sampling for Digital Terrain Models: A Trend Towards Automation", Photogrammetric Engineering and Remote Sensing, Vol. 48, No. 11, pp. 1687-1694.
- Allder, W. R., Caruso, V. M., Pearsall, R. A., and Troup, M. I. (1983), "An Overview of Digital Elevation Model Production at the United States Geological Survey", in Proceedings of the Fifth International Symposium on Computer Assisted Cartography, Falls Church, VA: American Congress on Surveying and Mapping, pp. 23-32.
- Bentley, Jon L. and Shamos, Michael I. (1976), "Divide-and-Conquer in Multidimensional Space", in Proceedings, Eighth Annual ACM Symposium on Theory of Computing. New York: Association for Computing Machinery, pp. 220-230.
- Boehm, B. W. (1967), "Tabular Representation of Multivariate Functions With Application to Topographic Modeling", Association for Computing Machinery, 22nd. National Conference, pp. 403-415.
- Brassel, Kurt E. and Reif, Douglas (1979), "A Procedure to Generate Thiessen Polygons", Geographical Analysis, Vol. 11, No. 3, pp. 289-303.
- Carter, James R. (1987), "Evaluation of the 1:250,000 Digital Elevation Model for Use in a GIS for the Great Smoky Mountains National Park", in Proceedings of the International Geographic Information Systems Conference, Vol. 2. Crystal City, VA: NASA, pp. 43-51.
- _____. (1988), "Digital Representations of Topographic Surfaces", Photogrammetric Engineering and Remote Sensing, Vol. 54, No. 11, pp. 1577-1580.

- _____. (1989), "Relative Errors Identified in USGS Gridded DEMs", in Proceedings of the Ninth International Symposium on Computer Assisted Cartography. Falls Church, VA: American Congress on Surveying and Mapping, pp. 255-265.
- _____. (1989), "Indices of the Planar Nature of Gridded DEMs", unpublished manuscript.
- Chen, Zi-Tan and Guevara, Armando (1987), "Systematic Selection of Very Important Points (VIP) from Digital Terrain Models for Constructing Triangular Irregular Networks", in Proceedings of the Eighth International Symposium on Computer Assisted Cartography. Falls Church, VA: American Congress on Surveying and Mapping, pp. 50-56.
- Chen, Zi-Tan (1988), "Break Lines on Terrain Surface", in Proceedings, GIS/LIS '88. Falls Church, VA: American Society for Photogrammetry and Remote Sensing, pp. 781-790.
- Delaunay, B. (1934), "Sur la Sphere Vide", Bulletin of the Academy of Sciences of the USSR, Classe Mat. Nat., pp. 793-800.
- DeLotto, Joseph (1989), "Elevation Matrices as a Data Source for Automated Terrain Classification", Manuscript to be Published in Proceedings of the Fourteenth International Cartographic Association Conference. Budapest, Hungary: International Cartographic Association.
- DeLotto, Joseph and Weibel, Robert (1988), "Automated Terrain Classification for GIS Modelling", in Proceedings, GIS/LIS '88. Falls Church, VA: American Society for Photogrammetry and Remote Sensing, pp. 618-627.
- Elfic, M. H. (1979), "Contouring by Use of a Triangular Mesh", Cartographic Journal, Vol. 16, No. 1, pp. 24-28.
- Environmental Systems Research Institute (1988), TIN Training Course Class Materials. Redlands, CA: ESRI.

- Evans, Ian S. (1972), "General Geomorphometry, Derivations of Altitude, and Descriptive Statistics", in Spatial Analysis in Geomorphology. edited by R.J. Chorley, London : Methuen, pp. 17-90.
- _____. (1980), "An Integrated System of Terrain Analysis and Slope Mapping", Zeitschrift fur Geomorphologie, N.F., Supp. Bd. 36, pp. 274-295.
- Franklin, Steven E. (1986), "An Approach to Processing Dense Grid Digital Elevation Models for Use in LANDSAT Terrain Classification", in 1986 ASPRS-ACSM Fall Convention Technical Papers. Falls Church, VA: ACSM, pp. 214-225.
- G. & C. Merriam Company (1973), Webster's New Collegiate Dictionary, Springfield, MA : G. & C. Meriam Company.
- Hammond, E. H. (1964), "Analysis of Properties in Land Form Geography: An Application to Broad-Scale Land Form Mapping", Annals of the Association of American Geographers, Vol. 54, pp. 11-19.
- Houston, J. M. (1973), Terrain Evaluation, London : Longman Group Limited.
- MacEachren, Alan M. and Davidson, John V. (1987), "Sampling and Isometric Mapping of Continuous Geographic Surfaces", The American Cartographer, Vol. 14, No. 4, pp. 299-320.
- Mark, David M. (1975), "Geomorphometric Parameters: A Review and Evaluation", Geografiska Annaler, Vol. 57A, No. 3-4, pp. 165-178.
- _____. (1975), "Computer Analysis of Topography: A Comparison of Terrain Storage Methods", Geografiska Annaler, Vol. 57A, No. 3-4, pp. 179-188.
- _____. (1983), "Automated Detection of Drainage Networks From Digital Elevation Models", in Proceedings of the Sixth International Symposium on Computer Assisted Cartography, Vol. 2. Canada: Steering Committee, Sixth International Symposium on Automated Cartography, pp. 288-298.

McCullagh, Michael J. and Ross, Charles G. (1980), "Delaunay Triangulation of a Random Data Set for Isarithmic Mapping", Cartographic Journal, Vol. 17, No. 2, pp. 93-99.

_____. (1979), "Triangular Systems in Surface Representations", in Proceedings of the Fourth International Symposium on Computer Assisted Cartography, Vol. 1. Falls Church, VA: American Congress on Surveying and Mapping, pp. 146-153.

McEwen, Robert B. and Jacknow, Harry (1979), "The USGS Digital Cartographic Database", in Proceedings of the Fourth International Symposium on Computer Assisted Cartography, Vol. 1. Falls Church: American Congress on Surveying and Mapping, pp. 225-235.

Miller, C. L. and Laflamme, R. A. (1958), "The Digital Terrain Model: Theory and Application", Photogrammetric Engineering, Vol. 25, pp. 433-442.

Moellering, H. and Tobler, W. (1972), "Geographical Variances", Geographical Analysis, Vol. 4, No. 1, pp. 34-50.

Nordbeck, Stig (1967), Computer Cartography Point in Polygon Programs. Lund, Sweden: C.W.K. Gleerup Publishers.

Pike, Richard J. (1987), "Toward Geometric Signatures for Geographic Information Systems", in Proceedings of the International Geographic Information System Symposium, Vol. 3. Crystal City, VA: NASA, pp. 15-26.

Peucker, Thomas K. (1977), "Data Structures for Digital Terrain Models: Discussion and Comparison", in First International Advanced Study Symposium on Topological Data Structures for Geographic Information Systems, Vol. 5. Cambridge, MA: Harvard University.

_____. (1979), "Digital Terrain Models: An Overview", in Proceedings of the Fourth International Symposium on Computer Assisted Cartography, Vol. 1. Falls Church, VA: American Congress on Surveying and Mapping, pp. 97-107.

- Peucker, Thomas K. and Chrisman, Nicholas (1975), "Cartographic Data Structures", The American Cartographer, Vol. 2, No. 1, pp. 55-69.
- Peucker, Thomas K. and Fowler, Robert J. and Little, James J. and Mark, David M. and (1978), "The Triangulated Irregular Network", in Proceedings of the Digital Terrain Models (DTM) Symposium. Falls Church, VA: American Society of Photogrammetry, pp. 516-540.
- Rhynsburger, D. (1973), "Analytic Delineation of Theissen Polygons", Geographical Analysis, Vol. 5, pp. 133-144.
- Shamos, Michael I. (1975), "Geometric Complexity", in Proceedings, Seventh Annual ACM Symposium on Theory of Computing. New York: Association for Computing Machinery, pp. 224-233.
- Shamos, Michael I. and Bentley, J. L. (1978), "Optimal Algorithms for Structuring Geographic Data", in Harvard Papers on Geographical Information Systems, Vol. 6.
- Strahler, A. N. (1952), "Hypsometry (area-altitude) Analysis of Erosional Topography", Geological Society of America Bulletin, Vol. 63, pp. 1117-1142.
- Taylor, Peter J. (1977), Quantitative Methods in Geography: An Introduction to Spatial Analysis. Prospect Heights, Ill.: Waveland Press.
- Thiessen, A. H. and Alter, J. C. (1911), "Precipitation Averages for Large Areas", Monthly Weather Review, Vol. 39, pp. 1082-1084.
- Tobler, Waldo R. (1967), "Geographical Filters and Their Inverses", Geographical Analysis, Vol. 1, pp. 234-253.
- Tobler, Waldo R. (1979), "Lattice Tuning", Geographical Analysis, Vol. 11, No. 1, pp. 36-44.
- USGS (1987), Digital Elevation Models: Data Users Guide. Reston, VA: Department of the Interior.
- Yoeli, Pinhas (1977), "Computer Executed Interpolation of Contours into Arrays of Randomly Distributed Height-Points", Cartographic Journal, Vol. 14, No. 2, pp. 103-108.

Zobrist, Albert L. (1979), "Data Structures and Algorithms for Raster Data Processing", in Proceedings of the Fourth International Symposium on Computer Assisted Cartography, Vol. 1. Falls Church, VA: American Congress on Surveying and Mapping, pp. 127-137.

VITA

Clayton A. Crawford was born in New York City on November 16, 1962. After attending elementary school in Englewood, New Jersey and high school in Bronxville, New York he attended Dean Junior College in Franklin, Massachusetts. He later transferred to The State University of New York at Geneseo where he graduated with a Bachelor of Arts (Hons.) Degree in Geography with a minor in Environmental Studies. In September of 1987, Clayton accepted a teaching assistantship with the Department of Geography at the University of Tennessee in Knoxville. He received his Master of Science Degree in Geography in May 1990.