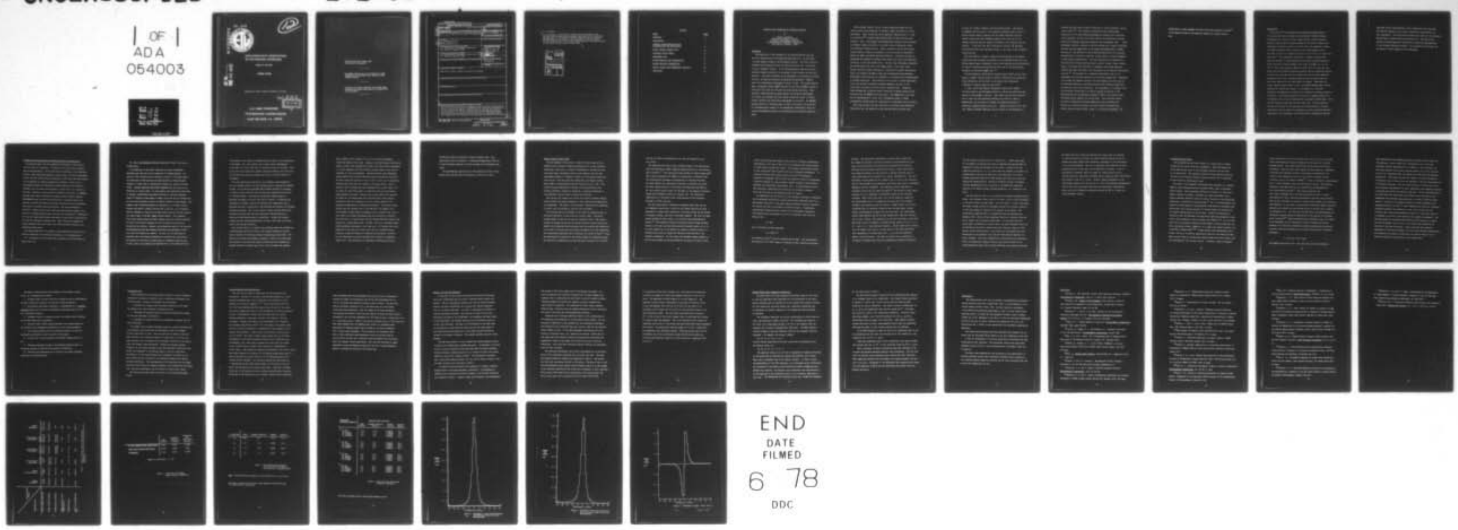
AD-A054 003 ARMY ENGINEER TOPOGRAPHIC LABS FORT BELVOIR VA F/G 8/2
ELEVATION DATA COMPACTION BY POLYNOMIAL MODELING.(U)
APR 78 J R JANCAITIS
UNCLASSIFIED ETL-0140 NL

1 OF 1
ADA
054003



END
DATE
FILMED
6 78
DDC

ETL - 0140

AD A054003

ELEVATION DATA COMPACTION BY POLYNOMIAL MODELING

James R. Jancaitis

APRIL 1978

AD No. ______
DDC FILE COPY

U.S. ARMY ENGINEER
TOPOGRAPHIC LABORATORIES
FORT BELVOIR, VA 22060

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (14) ETL - 0140	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9)
4. TITLE (and Subtitle) (6) ELEVATION DATA COMPACTION BY POLYNOMIAL MODELING.		5. TYPE OF REPORT & PERIOD COVERED Technical Report.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) James R. Jancaitis		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Army Engineer Topographic Laboratories Fort Belvoir, Virginia 22060		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (12) 47p.
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Engineer Topographic Laboratories Fort Belvoir, Virginia 22060		12. REPORT DATE (11) Apr[redacted]78
		13. NUMBER OF PAGES 45
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report details the status of ongoing research directed towards development of a near-term production implementation of digital data compression of terrain elevation information. The first section discusses the important data characteristics, the major applications, and the compression needs. The second section discusses the various published terrain representations, their capabilities and limitations. The third section presents an overview → next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

403 192 Gw

20. continued

of the Polynomial Terrain Model's characteristics and construction. The next section contained the development plan identified for production implementation of the polynomial modeling technique, and the remaining sections report on the status of various phases of this development. The results showed that the Polynomial Matrix method is the most promising of the various digital terrain formats (DFT).

CONTENTS

ELEVATION DATA COMPACTION BY POLYNOMIAL MODELING

by

James R. Jancaitis
Mapping Developments Division
Topographic Developments Laboratory
US Army Engineer Topographic Laboratories
Fort Belvoir, Virginia 22060

Background.

The feasibility of the automation of the stereocompilation process was first demonstrated by the Bausch and Lomb Optical Co. in 1950 under a US Army Engineer Research and Development Contract. The first practical equipment was produced soon after by Pickard and Burns, Inc., also under a Corps of Engineer contract. In the following years, Army and Air Force supported research resulted in a succession of more accurate and efficient equipment. This work culminated in the development of the production system entitled, The Universal Automatic Map Compilation Equipment,[1] UNAMACE, under a US Army Engineer development contract in 1964. The Defense Mapping Agency Topographic Center (DMATC) has had the work-horse UNAMACE systems in production since 1965 and currently has seven systems operating. These automatic stereocompilation systems produce 100 elevation measurements per second, and represent the most prolific production source of digital terrain elevation data from stereo photography in the world. The UNAMACE systems operate in a profiling mode in order to facilitate the production of orthophotographs, another of its considerable capabilities, which are required intermediate products in the defense map and database generation tasks.

These systems, however, are not without limitations based upon the current state-of-the-technology of automatic image correlation in stereo photography. These limitations can be generally characterized as "a lack of reliability,"[2,3] and, more specifically, result in the following three problems. First, the automatic stereocompilation process is more precisely an automatic image correlation of two small areas on photographs taken from different viewing stations. Camera orientation and non-zero ground slopes combine to distort the geometry in the two pictures considerably - making the image correlation much more difficult. Correction circuits or algorithms are employed to minimize these effects but they still definitely contribute to the total system errors. Second, these instruments frequently get unrecoverably lost due to poor image quality, cloud cover, water bodies, and other reasons resulting in a necessity for manual intervention. Third, the automatic equipment output has non-negligable high-frequency measurement errors that must be "smoothed" or filtered before orthophotograph or database generation. This problem is due not only to system measurement errors and the area correlation technique employed, but also to the correlation occurring on the tree tops or building roofs. Originally these problems were solved for contour production by manual extraction and smoothing from the UNAMACE drop line output. The UNAMACE drop line chart was an orthographically correct plot of greyshades keyed to change unambiguously through elevation transitions at a specified contour interval.

While some research has been directed at correcting these problems at compilation time, the majority of work has been directed toward their solution in a digital computer, post-processing fashion. This choice of emphasis was due in part to the downward spiraling costs of time on general purpose digital computers and the upward spiraling costs for special purpose electronic hardware systems, and in part due to the already recognized need for digital processing of the elevation data for contours. It was felt that data verification, editing, and smoothing could be most effectively performed during, or as a part of, the automated contouring task.

The first efforts at digital computer post-processing of the digital terrain data were performed by personnel at the Army Map Service (now the Defense Mapping Agency Topographic Center) which demonstrated the feasibility of digital contouring and contained the first use of the least-squares technique for smoothing UNAMACE data.[4,5]

The Electromagnetic Compatibility Analysis Center (ECAC) was the first agency to augment terrain elevation profile's definition with stream and ridge information, and the first to use digital data compression techniques to compact elevation data grids.[6]

In 1969, the US Army Engineer Topographic Laboratories (USAETL) initiated research and development efforts to improve the efficiency and quality of the automatic contours.[7] This effort was the first reported work in which the area smoothing technique was utilized to smooth UNAMACE data, and the first to consider the digital mosaicking of separately compiled elevation data sets. During the period of 1971 thru 1974, DMATC and USAETL supported contract efforts at the University of

Virginia directed toward automatic production of near-cartographic quality contour lines.[9,12] This research contained the first simultaneous application of the least-squares approximation technique with functional terrain modeling, as well as the first efforts in elevation data editing based upon separately digitized planimetric and hydrographic data. These development contracts resulted in terrain modeling and automated contouring software entitled, Contouring via the Surface Averaging Concept, CONSAC. The basis of this software's terrain modeling algorithm is the sequential determination of a matrix of numerous, locally valid low order polynomials by least-squares approximation. Low order polynomials are preferred so that solution of terrain characteristics, such as, contour lines, terrain slope, or new grids from the model is optimized; and because empirical evidence has shown that higher order polynomials do not improve interpolation accuracy.[13] The necessity for a sequential algorithm is due to the requirement for efficient modeling over arbitrarily large areas. Reasonable core storage and computer processor time requirements can only be met if the modeling technique is sequential and has local definition based upon processing of only local information. The requirement for a globally valid and consistant terrain model is achieved through utilization of the Weighting Function Interpolation Technique[12] (WIT) on the locally valid polynomials which comprise the terrain model. This technique assures reasonable interpolation between neighboring local polynomials and assures the continuity necessary for smooth terrain representations, either contours, profile lines, slopes, or reflectance greyshades. The significance of these arguments was also given great emphasis by Leberl[13] in his empirical study of interpolation schemes for digital terrain data.

Introduction.

In the early 70's the automated data gathering systems employed in related areas such as Landsat and projections of digital database requirements for archival storage, weapon systems and environmental studies at DMA focussed increased emphasis on the need for digital data compression techniques. Motivated by this trend, the compaction inherent the CONSAC polynomial terrain model was first investigated in 1973 and found to result in a compaction ratio of 30 to 1 over a 75 square mile test area.[14] This technique's compression of digital elevation data grids was compared to linear prediction and fourier transform methods in seven, thirty-five square kilometer test areas by Crombie, et. al. in 1974.[15] This study achieved compaction ratios of 27 to 1 over the test areas using the polynomial modeling technique; and for comparable modeling accuracy this ratio was over four times better than the best compaction achieved by the fourier transform method and was over seven times the best ratio achieved by the linear prediction method. These were the latest research results when, in 1975, the Defense Mapping Agency shifted emphasis and funding from research to development of a near-term production implementation of digital data compression of terrain elevation information. This report details the status of an ongoing research and development effort at the US Army Engineer Topographic Laboratories,. which has been directed toward this end since that time. The first section includes a discussion of the important data characteristics, its major applications, and the needs for compression, as well as those areas where compression is not warranted. The second section discusses the various published terrain representations, their capabilities and limitations with specific emphasis in the areas of application compatibility and data compression. The third section presents an overview of the Polynomial Terrain Model's characteristics and construction. The next section contains the development plan identified for production implementation of the polynomial modeling technique. The remaining sections report on the status of various phases of this development.

Automatic Stereocompilation Data Characteristics and Applications.

As mentioned above, the most significant limitation of this type of data is a lack of reliability. If taken at face value, this statement can be very misleading. Of the three factors mentioned in the background section, namely photograph geometry, system losses, and system measurement errors, only the latter affects the majority of the data, and these effects can be minimized by rigorous mathematical treatment. Errors due to photographic geometry differences and system losses are more difficult problems, but occur much less frequently, usually only 3 to 5 percent of the time. Currently, these problems are handled in two ways; by UNAMACE hardware and by post processing on a general purpose digital computer. The UNAMACE allows for both the specification of adverse areas by the operator (in which the machine just holds elevation), and for intervention and retrace of operator specified profiles. The post-processing is currently a simple surface slope filtering algorithm which requires user input of a maximum allowable slope over the area considered.[12] Combination of these two capabilities reduces the number of gross errors significantly but does not completely eliminate them. Small interactive systems are being considered which will permit cost effective manual editing of the remaining problem areas.

The characteristics of the automatic stereocompilation elevation data, which are important to the modeling techniques, can be summarized as:

(1) is obtained on a very dense grid, typically over 5,000 points per square inch, and

(2) has a non-negligible randomly distributed "noise" level due to system errors.

It is important to note that, currently, the major requirement dependent upon this data is the production of orthophotographs. The density of the elevation data is directly a result of this application; the production of high accuracy and high resolution orthophotographs depends on a dense grid of continuously varying (e.g. smooth) elevation values. Another important application dependent upon this data is the production of contour lines. At first glance this application would seem to require a considerably less dense grid than is currently employed. Leberl's[13] empirical study seems to support this view since he showed that for most cases elevation accuracy is linearly related to grid spacing. For example, most 1:50,000 map sheets are contoured at a ten meter contour interval and Leberl's results indicate that a grid spacing of 50 meters would be adequate to insure the one-half contour interval accuracy required for class A maps in the most rugged terrain studied. Also, numerous other study results quoted by Leberl indicate that the currently employed 12.5 meter and 25 meter spacings are far too dense for the accuracy required in the contours. However, as was mentioned earlier, the automatic stereocompilation data has randomly distributed non-negligable "noise" even after the blunders are removed. In order for the least-squares algorithm used to accurately determine the elevations for contouring in the presence of this noise, as many points as reasonably possible must be used. This is true because the expected error at any given grid point is decreased as the number of neighboring points used in its determination is increased. For these reasons, the optimal contour determination occurs when all of the dense grid of elevation values are utilized. This would seem to run contrary to Leberl's empirical evidence only until this and the fact that his study assumes zero error in the grid values are considered.

The other applications impacted by the automatic stereocompilation data are standard formats for data exchange, weapon's systems and archival storage. The currently established DMA standard format[16] for exchange of digital terrain elevation data includes the uncompacted grid written in a profile format. This mode was determined optimum because it minimizes the impact of partial data loss in transit, it minimizes the complexity of software required by different user communities, and it maximizes the data's compatibility with existing user hardware and software application systems. Most current weapon's support systems depend upon elevation data grids and the costs associated with modification of existing fielded systems would be prohibitive. Formats for archival storage are still being formulated and, in fact, this study is directed toward their optimal definition.

From initial analyses it appears that existing production systems for orhtophotographs and contour lines, data exchange formats, and current weapons systems are operating optimally using elevation data grids. The hardware and software developments necessary, as well as the additional time required to pack and then unpack the data, more than outweigh any possible benefits of having fewer bits to carry through these systems.

These arguments will continue to be true for the high throughput production system of the future. However, the future weapons and tactical support systems could substancially benefit from digital data compression if it were compatible with the archival storage formats of the day and "designed in" from the beginning. Possible benefits from compression include smaller unit size and increased ground coverage and accuracies. One often quoted and incorrect "fact of life" is that use of compression technique must slow down the response of any given system. This is generally true in systems that have negligable data access times or do not allow for some "look ahead" or prediction with pre-access and/or preprocessing of the data. However, data access times can be improved by compaction for applications with relatively long access times by allowing for the use of smaller but faster storage systems if the compaction regeneration is sufficiently efficient. This point is made clearer when one considers that on the CDC-6400, for example, over 150 different 16-term bicubic polynomials can be accessed from a set of over 6,000 core resident polynomials, and each evaluated at different and arbitrary X, Y values during only the average disk head movement time required of a single disk access. Further, (as will be discussed later) 6,000 polynomials represents a very large area. For systems that allow prediction of the next set of data need, the use of advance access techniques along with parallel or pipeline microprocessors provides for the additional time and processing power necessary to regenerate compact data. The utilization of compression techniques would most beneficially impact the archival storage of digital data. This application is not as sensitive to compression/regeneration time and is quite favorably impacted by reduced storage size requirements and costs.

The environmental applications of this data have not been covered because they are much less well defined at this point in time.

Digital Terrain Formats (DTF).

For the purposes of this study, a Digital Terrain Format will be defined as any consistant digital representation of terrain elevation values over a given area. A digital representation includes not only the actual binary information stored in the computer, but also the algorithms necessary to reference, unpack, and continuously define the data for use. Note that using this definition, the digital terrain model (DTM) or equally-spaced elevation data grid is not a DTF. This digital storage strategy has an implied access algorithm but only becomes a DTF when a method for consistant, continuous definition is also specified, such as linear or spline interpolation.

This section will present and discuss the currently proposed digital terrain formats with special emphasis on their application compatibility and compression. The digital terrain format's characteristics summarized in table 1 will be utilized to motivate this discussion.

The elevation matrix is included in the table for comparison and, as discussed previously, is the currently utilized format. The binary storage format in this case is just the elevation values. The storage overhead is minimal because the horizontal location of each elevation is implied by its position within the grid; only the position and orientation of one point in the grid and the spacing are needed to use the information. The interpolation algorithm used is arbitrary, as proven by Leberl,[13] since the interpolation accuracies of the various schemes are essentially equal. The application compatibility has been discussed previously for this DTF, and the low relative implementation cost and zero compaction ratio are obvious.

The Coded Elevation Matrix uses a Huffman coding of the Differential Pulse Code Modulation (DPCM) technique applied directly to the elevation matrix profiles.[17] This approach has been applied to digital image data[17] and is being investigated at USAETL for application to elevation data. This investigation represents an application of existing well-known errorless digital data compression techniques to elevation data grids. The other characteristics of this format are essentially the same as for the elevation matrix except for the eight to ten compaction achieved (based upon preliminary results of reference 18). The application compatibility is only given a good rating because of the unpacking required to utilize the data.

The next three formats considered are commonly based upon the use of irregularly spaced data, a quite old and logical approach for data gathering in a manual mode, especially if the data extrema are the points of greatest interest, and will be discussed together. The first of these, termed irregularly spaced data-A, is currently being investigated by Peucker.[19] This approach involves the greatest binary storage overhead because it not only requires three values per elevation point (x, y, z), but also requires up to six pointers or links per elevation point in a secondary data structure needed for efficient access and use of the elevation information. In this format the overhead storage required to use the data exceeds the storage needed for the data by a factor of two.

Little work has been done todate on the effect of different interpolation algorithms for this type of data, but it is suspected that results similar to those found for grids will result. The irregular spaced data-B refers to work by Whitten[31] which involves the use of orthogonal polynomials. The irregularly spaced data-C format refers to Hardy's[20] multiquadric equations of topography, which defines the elevation surface by determining a function composed of a sum of quadric surface equations (the polynomial equations for paraboloids, hyperboloids, etc). In these formats (B and C) only functional coefficients are stored and the overhead is minimal since the structure of the modeling function is implied.

Of significance are the facts that the orthogonal polynomial coefficients can be determined analytically (closed form solution equations) while the multiquadric equations of topography requires numerical solution (inversion of a matrix the size of the number of elevation points). Considerable computational savings could be realized by the multiquadric approach if the elevations were constrained to lie on an orthonormal grid, then the bilinear form

$$H = XAY^T$$

and its efficient solution algorithm

$$A = X^{-1}H(Y^{-1})^T$$

as presented by Schut[21] could be extended and utilized. The interpolation algorithms for the B and C formats are defined by their respective functional approach. The application compatibility of these three formats were all judged poor because of the poor efficiency of grid generation, poor area to area continuity, and the difficulty encountered in producing this data from the current elevation data sources. Hardy does address the continuity problem, but his solution of redoing adjoining areas as a new larger model is not feasible for a global representation. The poor efficiency of grid generation for formats B and C is due to the evaluation time for the extremely long polynomials that result when areas of reasonable (mapsheet) size are considered. The compaction ratios achievable by irregularly spaced data are based upon Mark's[22] and Peucker's works, which represent the only comparison of irregularly spaced points and grids available. Mark's work falls short of an actual empirical comparison, and instead makes use of the assumption that grid errors are linearly related to grid spacing to get his results. Further, Mark's critera for comparison were not elevation values but were the mean slope, terrain relief, and hypsometric integral over the test areas. The relationship between these measures and the one most desired here, elevation accuracy, is not clear, however, they represent the only data available at this time. Mark found ratios of irregular to gridded storage space of 47.2 to 1, 2.3 to 1, and 1.4 to 1 using Peucker's approach. Peucker notes quite correctly that the highest ratio (due to terrain releif) is the most unrealistic since measurement of extreme value differences is not an appropriate measure of grid precision. The other two ratios can be translated to irregular to gridded points ratios by removing the overhead of Peucker's and this results in ratios of 14 to 1 and 8.4 to 1. These ratios must not be weighted too heavily since they do represent projections based on comparative critera not suitable for this study. Peucker's one test comparison resulted in a 1.3 to 1 ratio when the overhead is included. The implementation costs are all high due to the complexity of the basic database generation required by these formats, and their poor source and application compatibilities. Peucker's technique received the highest relative implementation cost rating due to the additional complexity required in the binary storage overhead definition, e.g. finding all the links.

The last entry in table 1 is the previously discussed Polynomial Matrix format which defines a sparce grid of low order, locally valid polynomials as the basic digital storage format. As in all matrix or grid approaches, the binary storage overhead is negligable because the horizontal position of each data element is implied by its position in the matrix. The interpolation algorithm used is the Weighting Function Interpolation Technique (WIT), which utilizes weighting functions to produce a globally valid, smooth, and continuous terrain model. At any given point the global model is defined by a locally valid, low order polynomial which can be efficiently solved for a smooth data grid, contours, slope, or other desired terrain characteristic. The accuracy of this approach has been demonstrated on an extensive test area (over 600 square kilometers) in Cache, Oklahoma. The source compatibility is good because the polynomial matrix is sequentially computed directly from profile subsets of the stereocompilation output grid using the efficient least-squares algorithm.

The compaction ratio listed is based upon new results that are reported in later sections of this paper, but which basically involve the use of standard errorless digital data compression techniques on the coefficients of the local polynomials. The 30 to 1 compaction ratio reported for this technique earlier in this paper can be explained as follows; a grid of polynomials is computed, which is roughly 90 times as sparce as the original data grid. Since each polynomial has an average of three coefficients, this results in the 30 to 1 ratio. Further, digital compression of the coefficients results in at least an 80 to 1 ratio. The moderate implementation cost is due to the relatively high source and application compatibility. Grids can be very economically generated because the DTF is a simple low order polynomial locally.

Polynomial Matrix Format.

The polynomial matrix format consists of a sparce grid of locally determined and locally valid node polynomials. These polynomials are called node because there is one polynomial centered over each node in the sparce grid. The processes involved in their generation and use (as well as this format's important characteristics) will be discussed in this section in chronological order.

The output of the automatic stereocompilation equipment is a computer magnetic tape containing numerous contiguous DTM's, each in sequential profile format. Batch processing software is utilized on a large scale digital computer to mosaick, transform, and regrid these DTM's into a single consistent DTM which covers the area of interest in the desired map projection or cartesian co-ordinate system.[23] For a 1:50,000 mapsheet this results in a DTM composed of approximately 1,000 profiles containing 2,200 points each, with a ground spacing of 25 meters between the profiles and 12.5 meters between points in the profiles. Further processing may be utilized to edit the DTM based upon separately digitized planimetric detail, such as lake boundaries, rivers, drainage, and ridgelines.[12] It is at this point that the polynomial modeling process begins, currently based upon the batch processing software, CONSAC II, on a large scale digital computer, e.g. a CDC-6400 or UNIVAC-1108.[24,25] Profile subsets with approximately fifty percent overlap of the DTM are sequentially read into core memory and then square ground areas are sequentially defined along each profile subset, also with approximately fifty percent overlap. (Currently, these overlapping square ground areas contain rectangular data sets of 17 by 33 of the DTM points representing 0.16 square kilometers on the ground for 1:50,000 mapsheet areas). Each square ground area then receives the following processing. (Note that since each subarea receives identical, independent processing this algorithm contains a high degree of parallelism that could be taken advantage of on one of the parallel processor digital computers, such as the Goodyear STARAN).[26] First, the subarea is approximated with a "node" polynomial using the least-squares critera. The weighting employed in this approximation is as specified by the Maximum-Error Local Approximation Theorem.[27] This theorem specifies the most liberal (maximum) error distribution allowed in local or "node" approximation so that a specified approximation tolerance can be achieved when the WIT algorithm is utilized to form the final, smooth, and continuous approximation. At this point, surface slope filtering and simple statistical measures are utilized to locate extreme errors in the data and to measure the goodness of fit. The effects of the bad data on the approximation are efficiently removed through the application of the inverse of the Kalman Filter algorithm.[28] Program modifications are underway so that the length of the polynomial can be automatically varied to achieve the desired goodness of fit and error distribution based on the statistical analyses. Currently, the local polynomials are the most efficient four coefficient polynomial,

$$Z = C_0 + C_1x + C_2y + C_3xy$$

the simplest non-trivial case. Note that this order node polynomial, when combined with the weighting functions specified by WIT, allows for the modeling of terrain forms with a wavelength as low as the node polynomial spacing--approximately 0.2 kilometers. This resolution has been found adequate in all tests conducted thusfar, but in any case this will not be a limitation after the completion of the previously mentioned modification. When these processes are complete, the node polynomial's coefficients can be written out as the database, or used for grid generation or automated contouring. The first step necessary in further use of the node polynomial approximations is their combination with weighting functions as specified by the WIT interpolation algorithm. The WIT algorithm defines the final model over each square area defined by the centroids of four neighboring node polynomial approximations as a single, different simple function. Currently, the simplest and most efficient polynomial ratio weighting functions are used.[29] These weighting functions have a quadratic polynomial numerator and denominator, which, when combined with the currently used linear node polynomials, results in a final model that is a polynomial ratio with a cubic numerator and quadratic denominator. The globally smooth and continuous model defined by WIT results in a simple low order polynomial ratio in each of the square areas defined by the centroids of the node polynomials. These areas have been optimally determined as 0.05 square kilometers for the current software configuration. Grid generation and analytic solution for the contour lines are very efficient because of the simplicity of the global model.

The general characteristics and strengths of this digital terrain format can be summarized as follows:

1. automatic model accuracy definition through the use of approximation theorems, statistical analyses, and variable length polynomials.

2. the modeling algorithm is ameanable to implementation on a parallel processor because of the identical independent processing steps for the local polynomial subareas.

3. automatic removal of large data errors via surface slope filtering and inverse Kalman Filter algorithms.

4. efficient model format requiring minimal core storage because of the sequential and independent determination of the local polynomials.

5. low storage overhead since the horizontal position of the local polynomials are implied by their location in the matrix.

6. a smooth and continuous globally valid model through the use of WIT.

7. efficient geographic access of the database because access to a horizontal position is given by the structure of the matrix.

8. efficient grid generation due to the local low order polynomial definition of the global model.

Development Plan.

Having identified the Polynomial Matrix Format as the most promising compression technique for archival use in a production environment, the following steps or stages of development were identified.

1. Determine the largest areas accurately modelable by the lowest and highest order node polynomials considered for use.

2. Determine the minimum precision of digital representation needed for the node polynomial coefficients.

3. Determine the optimal digital data compression technique for the node coefficients.

The largest area accurately modelable under all terrain conditions with the polynomial range chosen must be determined so that the automatic accuracy algorithm will work. The area size must be chosen so that the most efficient four coefficient model is utilized most of the time, and small enough so that the highest order polynomial available can accurately model the roughest areas encountered. Knowing the minimum precision necessary for the digital storage of the node polynomial coefficients will permit the assignment of the minimum number of bits per coefficient. Determination of the optimal digital data compression technique for the node coefficients will provide for the reasonable minimum total digital storage requirements. The reasonableness of the approaches will be judged based upon such factors as compaction achieved, data regeneration processing time and power requirements, and the effects of partial data losses.

The following sections report the status of our ongoing development effort.

Optimum Modeling Area Determination.

The need for this study is based upon the following empirical observation. The size of the square area optimally modeled by a final, weighted polynomial when a four coefficient node polynomial is used depends on the terrain roughness.[11,15] This result, coupled with the fact that this modeling algorithm requires that the size of the square area remain constant, dictates that node polynomials of varying complexity be used. Unfortunately, research has also shown that the overall modeling efficiency is inversely related to the length of the node polynomial. However, this penalty can be minimized if Dr. Rauhala's Array Algegra algorithm is used.[30] For this reason, the approach taken here is to limit our analyses to consideration of the biquadric as the largest node polynomial. Other factors contributing to this decision are as follows. The current optimum square ground area for the final polynomials is 0.05 square kilometers (based on four coefficient node polynomials), over a 664 square kilometer test area (the area covered by the 1:50,000 Cache, OK sheet) which contains all representative terrain types. The optimum was empirically determined as the largest square areas which resulted in 90 percent of the modeled points being within ± 5 meters (one-half the contour interval) of the stereocompilation data (extreme errors excluded). The contours produced from this model are too generalized in the rough terrain areas, fit the noise in the flat areas, and are good in the rolling terrain areas. Note that, although 90 percent of the elevation values are within one-half contour interval, 90 percent of the contours are not! Further, three and four coefficient node polynomials have been investigated and there was good correspondence between the number of coefficients used in the node polynomials and the size of the square area modeled to within a given accuracy. The best compaction previously reported, 30 to 1, was based upon use of the three coefficient node polynomials. When the length of the node polynomials was increased from three to four the area covered to within the same accuracy increased so that the node polynomials are 160 times as sparce resulting in a 40 to 1 compression ration, (12500-4 coefficient polynomials per 2×10^6 elevations). When these arguments are all combined they result in a reasonably good expectation of modeling square areas as large as .14 square kilometers with a single final polynomial. This is based on the assumption that four or less coefficient node polynomials could be accurately used in the flat areas and that nine coefficient node polynomials would accurately model the roughest terrain encountered. Work has been initiated on implementation of these strategies and results of the use of these hypotheses should be available by the end of this fiscal year.

Minimum Precision Determination

At the start of this study the precision with which the floating point node coefficients must be stored to maintain model accuracy was unknown. The only information available was that the terrain modeling software had been successfully run on large scale digital computers with 36 and 60 bit word sizes. In order to investigate the effects of varying coefficient precision, four measures of model accuracy were identified; the mean of the differences, the normalized root mean square of the differences, the standard deviation of the differences and the contour line locations. The mean of the differences is the mean value of the model minus grid elevation values and the root mean square of the differences is the square root of the summed squared differences divided by the number of samples. The standard deviation has its normal meaning.

The test area chosen is a well surveyed 644+ square kilometer region covered by the 1:50000 Cache, Oklahoma map sheet. As mentioned in the last section, previous empirical testing had shown that the current modeling software achieved a modeling accuracy of $\pm$ 5 meters for 90% of the stereo compilation data (over 2 million points). The histogram of residuals, model values minus measured values, for this case is shown in Figure 1 and the model statistics are contained in Table 2.

At present the only precision test completed is a simple truncation integerization of the node polynomial coefficients. The histogram of residuals which resulted are plotted in Figure 2 and the model statistics are contained in Table 2. Figure 3 shows the histogram of the differences.

The results of this first simple test is suprisingly quite good. As would be expected from truncation integerization the mean changed from essential zero to essentially one half meter, this shift should be almost completely removed if rounding the numbers to produce integers were used instead of truncation. The normalized root mean square residuals differ by only .078 meters and the standard devaitions by only .038 meters. The contour line plots were indistinguishably different.

These are very critically important results in the consideration of the feasibility and compaction of the polynomial matrix method. All comparisons of compaction to date assumed that the polynomial coefficients would require no more precision than that currently used for the elevation values, namely 16 bits. If the node polynomial coefficients had to be stored in a floating point, e.g., mantissa and exponent, format then not only might this have resulted in greatly intreased data compression and regeneration times but also might have required more than 16 bits per coefficient. This could have seriously adversely effected the compression ratio for this technique.

The precision requirements were further investigated by a statistical study of the coefficients generated over the Cache test area. The means and standard deviations obtained are summarized in Table 3. The terrain over this test area varied from 320 to 740 meters. The mean of the Co term can be interpreted as the average terrain elevation since it is the average of the elevation predicted by the 12,500 node polynomials at their centroids. The C_1 and C_2 term's mean can be interpreted as the average slopes, and the C_3 term's mean can be interpreted as the average cross-slopes, or interaction of the X and Y slopes, e.g., this term can be understood as either the change in X slope with movement along the y axis or vice versa. The importance of these numbers is in their magnitude. The magnitude of the C_0 term is determined by the range of elevation values in the area modeled, and so the use of 16 bits for this term is as reasonable as the current use of 16 bits for the elevation values in the stereocompilation grids. The other three coefficients are bounded by $\pm 2^7 = \pm 128$ so that if only their integer value is used eight bits storage for these is more than adequate. This is an important result because our compaction ratio of 40 to 1 was based upon each coefficient being represented with 16 bits. By lowering the required length of 3 of the 4 coefficients the compaction ratio achievable becomes: 12500 polynomials with 1-16 bit coefficient and 3-8 bit coefficients equals 500,000 bits divided into 2 million grid elevation values at 16 bits each gives a compaction ratio of 64 to 1.

Optimal Digital Data Compression Techniques

As stated before the data compression philosophy chosen in this study is that all significant data smoothing will be accomplished in the mathematical modeling step via application of the rigorous least squares criteria. For this reason only errorless digital data compression algorithms will be considered for further compaction of the integerized mode polynomial coefficients.

The techniques identified for initial investigation is the differential pulse code modulation, (DPCM), technique. This technique was chosen as the first to be investigated because it is one of the simplist both conceptually and in terms of implementation difficulty. The DPCM technique involves storage of only the difference from the last parameter value encountered. For example, the string of numbers:

10, 12, 13, 15, 14, 12, 11, 9

could be stored using DPCM as the first value plus the differences from the previous value, e.g.

10, +2, +1, +2, -1, -2, -1, -2

The important thing to note is that the magnitude of numbers encountered in the second case is significantly reduced and therefore would require fewer bits to represent them in the digital computer. In order to study the applicability of the DPCM technique to the polynomial matrix coefficients the statistics of the absolute value differences between neighboring polynomials were computed. The absolute value differences were used because it is the magnitude of the differences which is of the greatest importance in this case. The differences were computerd along rows, columns and diagonals and the results are in table 4.

The much smaller range of values for the C_0 differences when compared to its original range is very significant. The largest number encountered is within $2^7 = 128$ so that 8 bits could be saved by using the DPCM technique for the C_0 coefficient. The range of values of differences for C_1, C_2 and C_3 are essentially equal and all exceed the $2^7 = 128$ range that bound the values of the coefficients themselves. Therefore, based upon these numbers, use of the DPCM technique for the C_1, C_2 and C_3 coefficients does not appear to be beneficial. These numbers, however, include some areas of dense, large magnitude errors which are negatively affecting the range, so that further analyses will be performed.

If the DPCM technique is used for only the C_0 coefficient then it can be represented with only eight bits instead of 16, this results in a new compaction ratio of 80 to 1, since:

12500 node polynomials with 4-8 bit coefficients each equals 400,000 bits divided into the 2×10^6 - 16 bit elevations gives the 80 to 1 ratio.

Other techniques are being investigated which take advantage of the low mean and standard deviation absolute value differences. The fact that these are absolute value difference means coupled with their relatively small magnitude indicates that the numbers are tightly grouped at the low end of the scale. Histograms of the differences were prepared which showed that the magnitude of 90% of all the differences were within twice the standard deviations.

Conclusions.

The characteristics and uses of automatic stereocompilation equipment's terrain elevation data have a significant effect on the suitability of the various digital terrafn formats, DTF. The most important requirements of the DTF and its generating algorithms are that it a) must remove large magnitude errors, b) smooth the non-negligable frequency noise by making optimum use of the redundancy in the data, c) allow for efficient geographical access of the data, d) permit efficient grid generation for use in the applications and c) result in good compression with reasonable regeneration penalties.

The various DTF approaches have been identified and compared with the result that the Polynomial Matrix method is the most promising at this time.

A plan for development of production algorithm's and software has been formulated and work initiated. The preliminary results have proven that compaction ratios of 80 to 1 can be guaranteed and much higher ratios can be expected.

The major tasks remaining are the evaluation of the applicability of standard errorless digital data compression techniques and an analysis of the tradeoffs between compaction achieved and the costs associated with compaction and regeneration for use.

References:

[1]Bertram, S., "The Universal Automatic Map Compilation Equipment (UNAMACE)," Photogrammetric Engineering, XXXI, No. 2, March 1965, page 244.

[2]Thompson, M.M., Manual of Photogrammetry, Third Edition, Volume II, 1966 Chapter XV, Automation of Stereocompilation, George Banta Printing Co., Menasha, Wisconsin, pages 764 and 800.

[3]Blacuhut, T. J. and M. C. van Wijk, "Results of the International Orthophoto Experiment 1972-76," Photogrammetric Engineering and Remote Sensing, Vol. XLII, No. 12, December 1976, pp. 1483-1498.

[4] Light, D. L., "Ranger Mapping by Analytics," Photogrammetric Engineering, September 1966, pages 792-800.

[5]Vitelleo, D., Biggin, M. J. and Middleton, G., "Automatic Contouring at the Army Map Service," Photogrammetric Engineering, October 1968.

[6]Gahr, J. A. and J. Iseli, "Use of Topographic Data in Communications-Electronics," IIT Research Institute, Anapolis, Md., November 1965.

[7]Fleshel, B., Stewart, A. J., and M. Gilman, "CONPLOT I, A Contour Generating Program," ETL-CR-70-2, USAETL Final Contract Report, May 1970, 170 pages.

[8]Noble, B., Applied Linear Algebra, Prentice-Hall Inc., Englewood Cliffs, N.J., page 142.

[9]Junkins, J. and J. R. Jancaitis, "Mathematical Terrain Analysis," presented to the 1971 ASP-ACSM Annual Meeting, Washington, DC.

[10]Jancaitis, J. R. and J. Junkins, "Modeling Irregular Surfaces" Photogrammetric Engineering, 1973, pp 413-420.

[11]Jancaitis, J. R. and J. Junkins, "Mathematical Techniques for Automated Cartography," USAETL Contract Report ETL-CR-73-4, February 1973, 108 pages.

[12]Jancaitis, J. R., "Modeling and Contouring Irregular Surfaces Subject to Constraints," USAETL Contract Report ETL-CR-74-19, January 1975, 171 pages.

[13]Leberl, F., "Interpolation in Square Grid DTM," The ITC Journal, 1973 #5, pp 756-807.

[14]Jancaitis, J. R. and J. Junkins, "Weighting Function Techniques for Storage and Analysis of Mass Remote Sensing Data," Conference on Machine Processing of Remotely Sensed Data, Laboratory for Applications of Remote Sensing, West Lafayeete, Indiana, Oct 1973, paper published in proceedings.

[15]Crombie, J., May, T., and E. Thorp, "Compaction of UNAMACE Profile Data, "ETL Technical Report, July 1974, 44 pages.

[16]"DMA Standard for Digital Terrain Elevation Data File," (no stated author), DMA working paper, 26 November 1974, 8 pages.

[17]Crombie, M., "Errorfree Compression of Digital Imagery," USAETL Technical Report, ETL-0079, November 1976, 26 pages.

[18]Crombie, M., personal communication on unfinished investigation of compression of UNAMACE elevation data using digital compression techniques, January 1977.

[19]Peucker, T. K., et.al, "Digital Representation of Three-Dimensional Surfaces By Triangulated Irregular Networks (TIN)," ONR Technical Report No. 10 on Contract N00014-75-C-0886, 1976, 63 pages.

[20]Hardy, R. L., "Analytical Topographic Surfaces by Spatial Intersection," <u>Photogrammetric Engineering</u>, Vol. 38, No. 5, 1972.

[21]Schut, G. H., "Review of Interpolation Methods for Digital Terrain Models, "Commission III invited paper, XIIIth Congress of the International Society for Photogrammetry, Helsinki, 1976.

[22]Mark, D.M., "Computer Analysis of Topography: A Comparison of Terrain Storage Methods, Geografiska Annaler, 57A(1975) 3-4, pp 179-188.

[23]Jancaitis, J. R., "Basic Digital Terrain Elevation Database Software, DTEDS,"USAETL Statement of Work on Contract DAAG53-76-C-0149, February 1976, 82 pages.

[24]Jancaitis, J. R., "User's Manual for CONSAC II; Software for CONtouring Via the Surface Averaging Concept," prepared for Defense Mapping Agency Topographic Center under Contract DMA-800-74-C-0038, March 1975, 56 pages.

[25]Jancaitis, J. R., "Programmer's Reference Manual for CONSAC II, Software for CONtouring via the Surface Averaging Concept," prepared for the Defense Mapping Agency Topographic Center under Contract DMA-800-74-C-0038, March 1975, 56 pages.

[26]Batcher, K. E., "STARAN Parallel Processor System Software, 1974 National Computer Conference, AFIPS Conference Proceeding, Vol. 43, pp 17-22.

[27]Jancaitis, J., "The Weighting Function Interpolation Technique: Further Results," presented to the American Geophysical Union, 1974 Fall Annual Meeting, San Francisco, California, Dec 12-17.

[28]Gura, I. A., "An Algebraic Approach to Optimal State Estimation," Hughes Aircraft Company, Space Systems Division, SSD 70072R, March 1967, 93 pages.

[29]Jancaitis, J. R., "Extending Weighting Functions for Interpolation and Approximation," presented to the 51st Annual Meeting, Virginia Academy of Sciences, Williamsburg, Virginia, May 1973.

[30]Jancaitis, J. R. and R. L. Magee, "Investigation of the Application of 'Array Algebra' to Terrain Modeling," presented to the 1977 ASP-ACSM Joint Annual Spring Convention Washington, DC, March 1977.

[31]Whitten, E.H.T., "Orthogonal Polynomial Trend Surfaces for Irregularly Spaced Data," Mathematical Geology, Vol. 2, No. 2, 1970, pp. 141-152.

CHARACTERISTICS \ DIGITAL TERRAIN FORMAT	ELEVATION MATRIX	CODED ELEVATION MATRIX	IRREGULARLY SPACED DATA-A (Peucker)	IRREGULARLY SPACED DATA-B (Whitten)	IRREGULARLY SPACED DATA-C (Hardy)	POLYNOMIAL MATRIX
BINARY STORAGE ALGORITHM	Grid Values	Huffman of DPCM	Irregular x,y,z	Coeff's of Ortho Poly	Coeff's of Multiquad	Coeff's of Local Poly
BINARY STORAGE OVERHEAD	Negligable	Negligable	High-need Directory	Negligable	Negligable	Negligable
INTERPOLATION ALGORITHM	Arbitrary	Arbitrary	Arbitrary	Orthogonal Polynomials	Multiquad Equations	WIT
SOURCE AND APPLICATION COMPATIBILITY	Excellent	Good	Poor	Poor	Poor	Good
COMPACTION RATIO	0	>8 to 10	1.4 to 2.3*	8 to 14*	8 to 14*	>80
RELATIVE IMPLEMENTATION COST	Lowest	Moderate	Highest	High	High	Moderate

Table 1. Digital Terrain Format's Characteristics (*based upon projections from insufficient data, see text)

	Mean (meters)	Standard Deviation (meters)	Normalized Root Mean Square (meters)
60 BIT FLOATING POINT COEFFICIENTS	0.043	3.809	3.8087
TRUNCIATED INTEGER COEFFICIENTS	-0.558	3.8467	3.887
DIFFERENCES	0.601	0.0377	0.0783

Number of elevations = 2×10^6

Table 2. Cache Test Area Model Characteristics Comparison

COEFFICIENT	MEAN	STANDARD DEVIATION	MINIMUM	MAXIMUM*
C_0	408.6	58.7	320.6	721.9
C_1	6.1	9.2	0.0008	122.2
C_2	5.5	8.8	0.0002	105.9
C_3	5.1	7.4	0.0010	96.9

Table 3. Node Coefficient Statistics for Cache Test Area (based on Model Equation = $C_0+C_1X+C_2Y+C_3XY$)

NOTE: Statistics shown are absolute value statistics for C_1, C_2, and C_3

*This data contained areas of dense large magnitude errors which were not removed prior to processing.

COEFFICIENT DIFFERENCE DIRECTION	ABSOLUTE VALUE STATISTICS			
	MEAN (meters)	STANDARD DEVIATION (meters)	MINIMUM (meters)	MAXIMUM* (meters)
C_0				
(a) Row	5.3	8.0	0.000736	122.3
(b) Column	4.9	7.9	0.000005	99.5
(c) Diagonal	6.8	10.1	0.001285	117.5
(d) Averages	5.7	8.7	0.001169	113.1
C_1				
(a) Row	8.5	12.4	0.000241	164.1
(b) Column	8.0	13.4	0.000283	192.6
(c) Diagonal	8.5	12.2	0.000271	146.4
(d) Average	8.3	12.7	0.000265	167.7
C_2				
(a) Row	5.9	11.5	0.000065	156.4
(b) Column	6.4	10.9	0.000487	124.9
(c) Diagonal	6.6	11.0	0.000239	118.6
(d) Average	6.3	11.1	0.000264	133.3
C_3				
(a) Row	7.2	11.1	0.000200	163.6
(b) Column	7.2	10.7	0.000262	151.2
(c) Diagonal	7.2	10.4	0.000276	128.6
(d) Average	7.2	10.7	0.000246	147.8

Table 4. Cache Test Data Coefficient Difference Statistics

*this data contained areas of dense large magnitude errors.

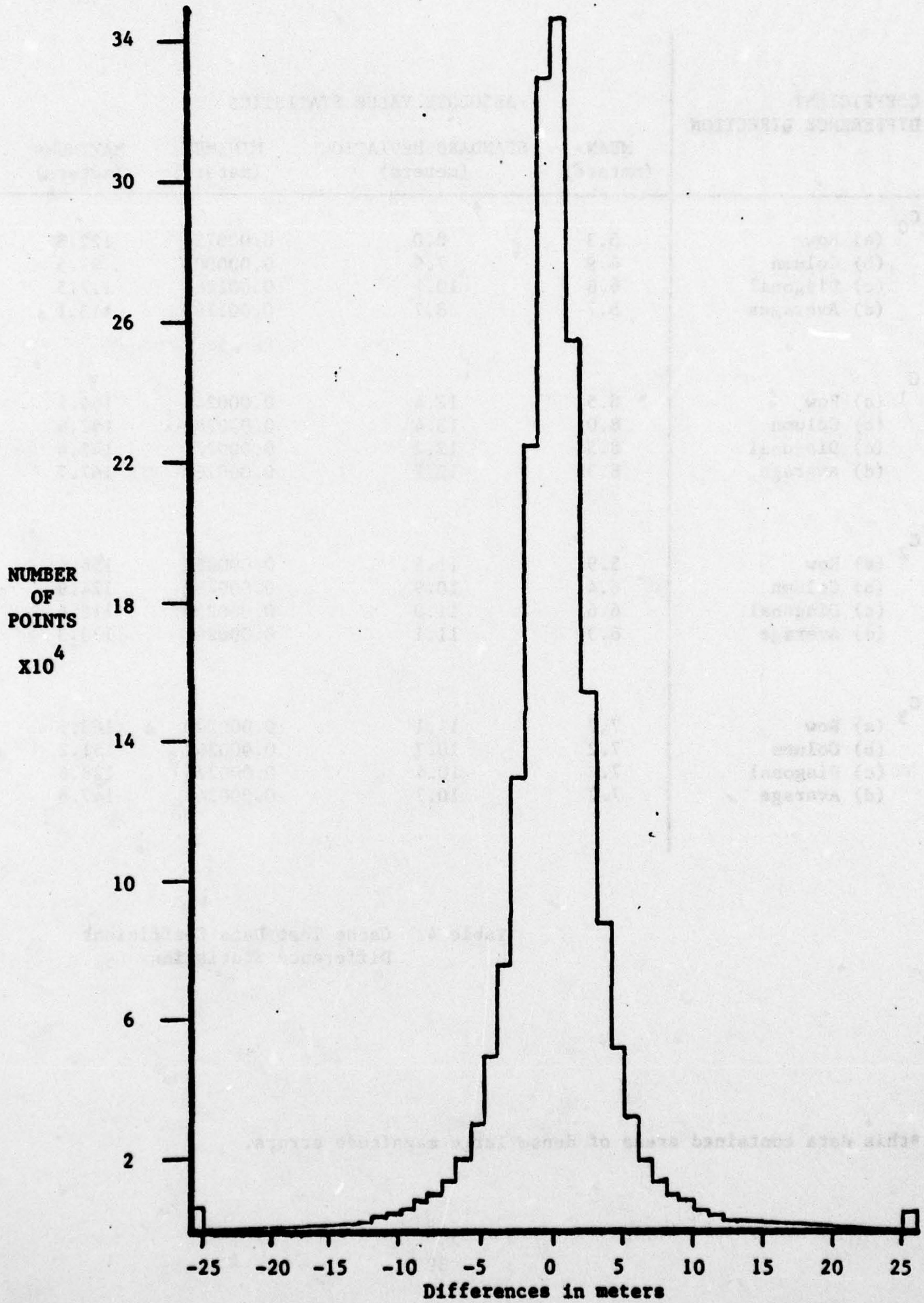


Figure 1. Histogram of Cache Test Differences with Floating Point Coefficient Representation

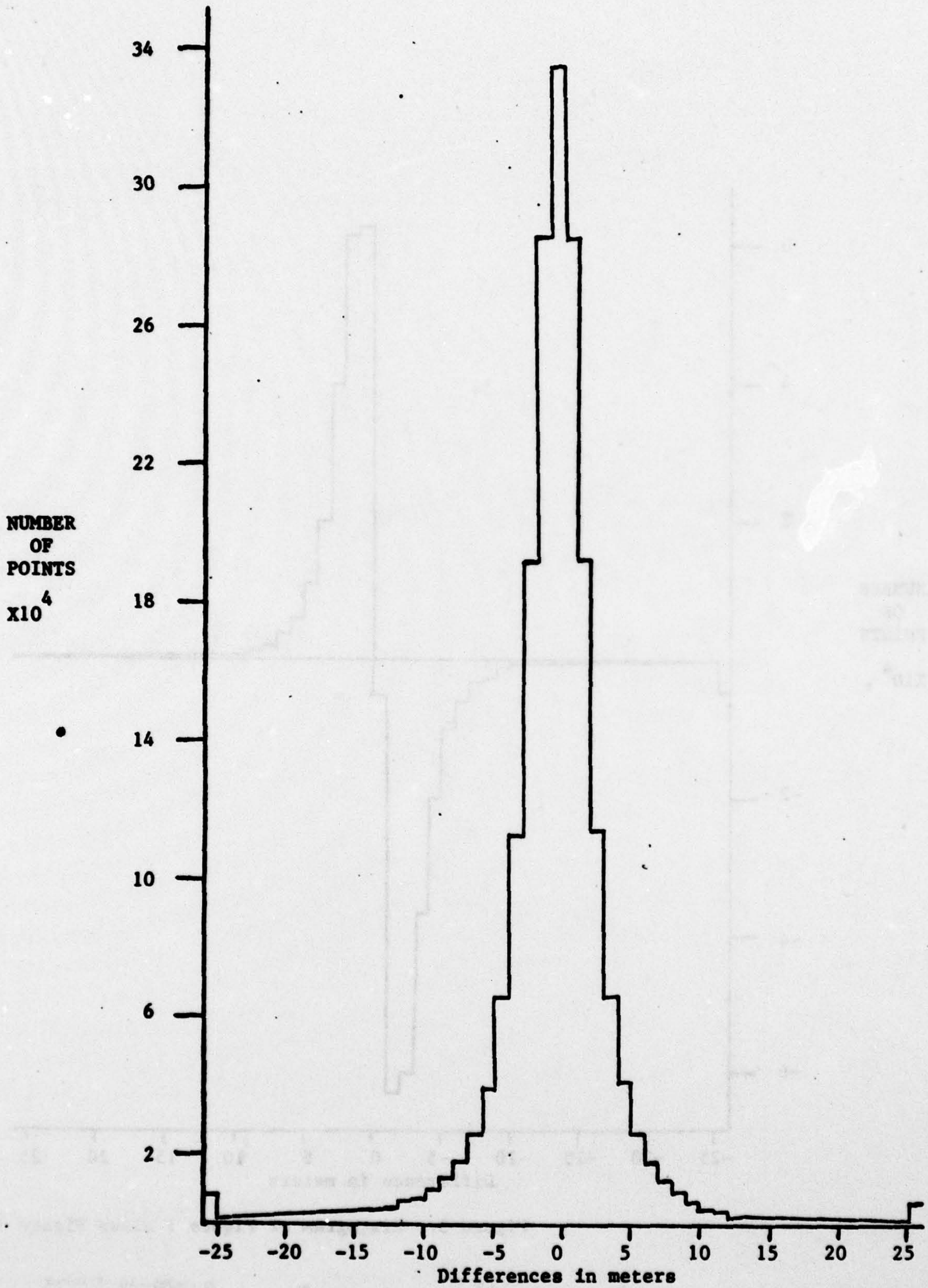


Figure 2. Histogram of Cache Test Differences with Truncated Integer Coefficient Representation

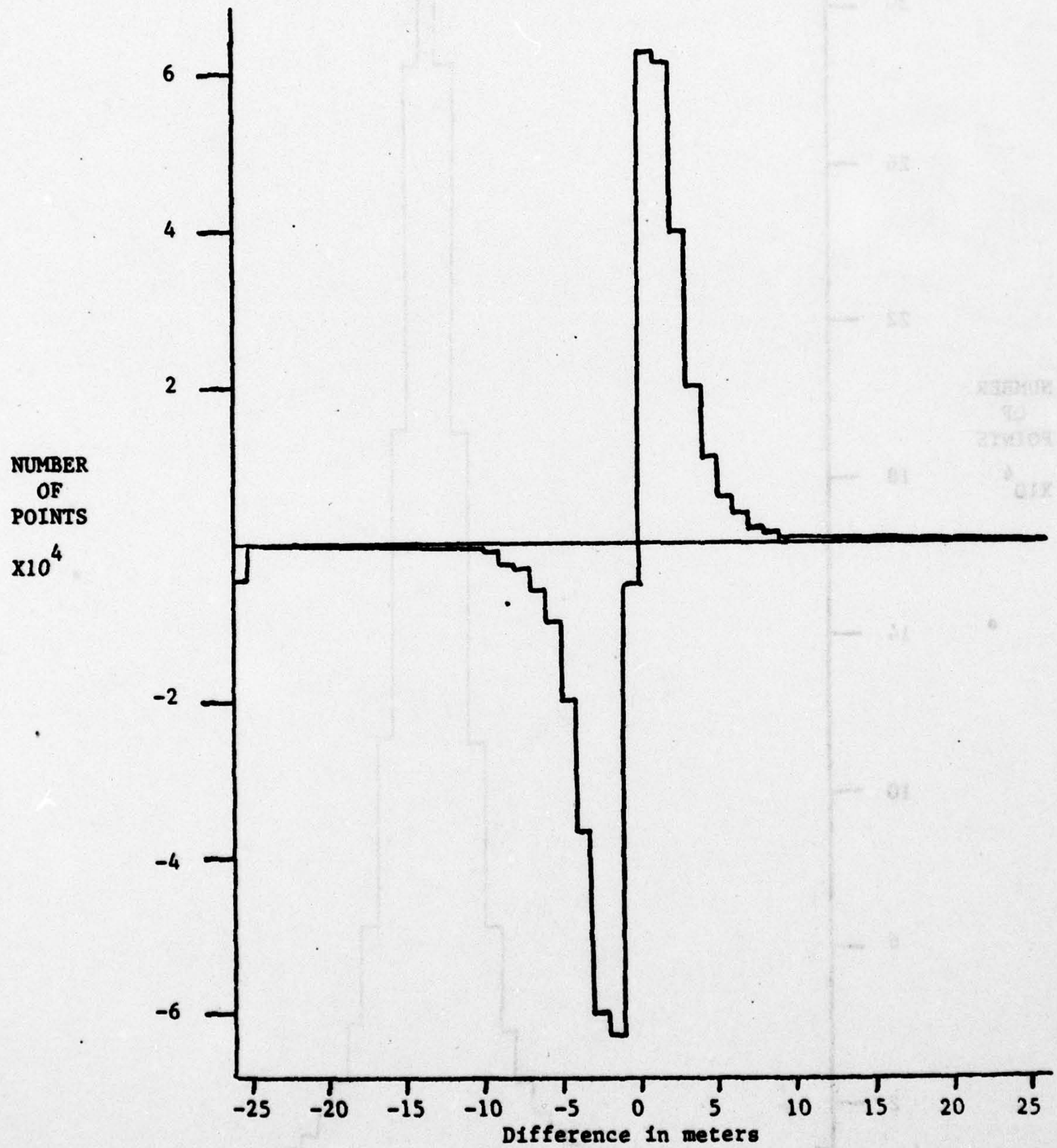


Figure 3. Histogram of Figure 1 minus Figure 2