INTERNATIONAL SOCIETY OF PHOTOGRAMMETRY

AND REMOTE SENSING

LEAST SQUARES POLYNOMIAL SURFACE FITTING

WITH OVERLAPPING DATA

W.P. SEGU
UNIVERSITY OF DAR ES SALAAM
TANZANIA

COMMISSION II

ABSTRACT

The author describes a method of least squares surface fitting, as a system for terrain elevation data processing, in which the elevation data are overlapped along common cell boundaries to avoid discontinuities of surfaces across the boundaries. The method has been tested using Chebyshev Polynomials and some test results are presented in this paper.

INTRODUCTION

Terrain elevation data can be presented in three main modes, namely discrete mode, graphical mode and functional mode. The functional mode of representing terrain elevation data gives rise to what can be described as Functional Digital Elevation Model (FDEM).

FDEM is a mathematical expression representing terrain elevation, Z, of a point as a function of its plan coordinates (x,y).

That is,

$$Z = f(x,y)$$

In expanded power series form, the expression can be written as:-

$$X = a_o + a_1 + a_2y + a_3x^2 + a_4xy + a_5y^2 + \ldots\ldots\ldots$$

A computer hardware and software are required for storing and processing data in FDEM.

Some examples of the most commonly used functional model forms of representing terrain elevation data as surfaces are: linear function (Allam, 1978), bilinear function (Leberl, 1973; de Manson d'Autume, 1979), hyperbolic paraboloid function (Grist, 1972), bicubic polynomial function (Schut, 1976), and double Fourier Series function(Maxwell and Turpin, 1968).

The main uses of FDEMs are found in computer-assisted engineering applications and in mapping. Doyle (1978) and Turner (1978) have highlighted some of these applications.

The usefulness of these applications notwithstanding, the mathematical formulation of the appropriate FDEMs and their solutions have some problems to overcome. The author has identified some of these problems at different operational levels as shown in Table 1.

Table 1: Problems Associated with FDEMs

LEVEL OF OPERATION	ASSOCIATED PROBLEMS
1. DATA CAPTURE	(a) Type of terrain (b) Accuracy of data (c) Density of data (d) Pattern of data
2. DATA PROCESSING	(a) Type of FDEM (b) Model constraints (c) Accuracy of output data (specifications) (d) Computer Algorithm
3. DATA OUTPUT	(a) Type of output data (b) Accuracy output data (c) Application of output data

Surface Fitting

In digital elevation model (DEM), functional mode has some particular advantages over discrete form or graphical form. In functional mode, a digital computer that processes the data also stores the elements of the surface either cellwise or globalwise. These elements can be of different types depending on the intended use of the terrain (Collins, 1975). In this work, the stored elements of a surface are the coefficients of the surface polynomial (Segu, 1985). Now during any subsequent use of the surface, such as in interpolation, surface data analysis can be done directly on that surface by use of these surface elements. This is of great advantage.

The elemens of a terrain surface stored in a computer are derived from field data by a process of surface fitting. Several techniques of surface fitting have been developed and tested both patchwise and globalwise. Jancaitis and Junkins (1973) describe some technique of surface fitting and its associated problems.

Least Squares

Least Squares technique is a very powerful tool for fitting experimental data to some mathematical model. Birge (1947)and Berztiss (1964) have covered this subject to great depths. The writer has adopted the principle of least squares in this work without prejudice.

Chevbyshev Polynomials

As pointed out earlier, there are many different techniques of carrying out surface fitting. The work carried out by the author has been based on Chebyshev polynomials (Segu, 1985).

Surface fitting by Chebyshev Polynomials requires that the data be in profile mode. The profiles should be at regular intervals but the elevation data along these profiles need not necessarily be at regular intervals.

PRINCIPLES OF OVERLAPPING

DEM data for terrain surface fitting are acquired invariably in discrete mode either photogrammetrically or by digitizing a contour map. The data can be in random, regular grid or profile form.

If the captured data are fitted with least squares surface polynomials cellwise, the surfaces so derived will not, in general, be continuous across cell boundaries unless certain precautions are taken (Jancaitis and Junkins, 1973). The technique of overlapping the data across fixed cell boundaries as discussed in this paper is an attempt to solve this problem.

In Chebyshev polynomials, the data to be fitted are in profile form. The method of surface fitting in a cell is based on fitting observed profile elevation data, Z, along a fixed axis, say y, to a degree k in y line by line and to a degree l in x. By overlapping data from adjacent cells at varying percentages of overlap, the resulting accuracy of fit were compared.

The basic principle guiding mathematical modelling with overlapping data is that a cell surface function is allowed to be influenced by data outside the original cell boundary. The effect of this is that such a surface will acquire some characteristics of the adjacent cell. Assuming cells of equal sizes and similar terrain types, the maximum percentage overlap has been assumed intuitively to be 50%. Over 50% overlap, the surface so defined can be assumed to have degenerated into a completely different surface function from the original.

The results presented in this paper have been based on calculations of percentage overlap from 0% to 30% in steps of 5%.

EXPERIMENTAL DATA

The results of this work are based on an experiment carried out on a test site measuring 2000m by 8000m. The original field data were captured photogrammetrically as (x, y, z) at regular intervals of 50m in x and y. The test area was later subdivided into cells of sizes 1000m by 10000m. There were thus 2 x 8 = 16 cells in the whole test area. And a typical cell had 21 x 21 = 441 data points.

Each cell's data were fitted with a least squares Chebyshev polynomial surface to a degree of fit that gave a Root Mean square Error (RMSE) of fit closest to and less than 5m. The data were then overlapped progressively across common cell boundaries from 0% to 30% in steps of 5%. The low accuracy of fit specified for this project, as also evidenced from the presented results, is a result of two main factors:

(1) low density of data and

(2) rough terrain.

To test the accuracy of fit at different percentage overlaps, two approaches were adopted:

(1) For each cell, nine fixed "standard" points were defined as shown in Fig.1. After each surface fitting routine, residuals were computed at the standard points of each cell and the RMSE calculated at the corresponding percentage overlap. Table 2 is a summary of these results.

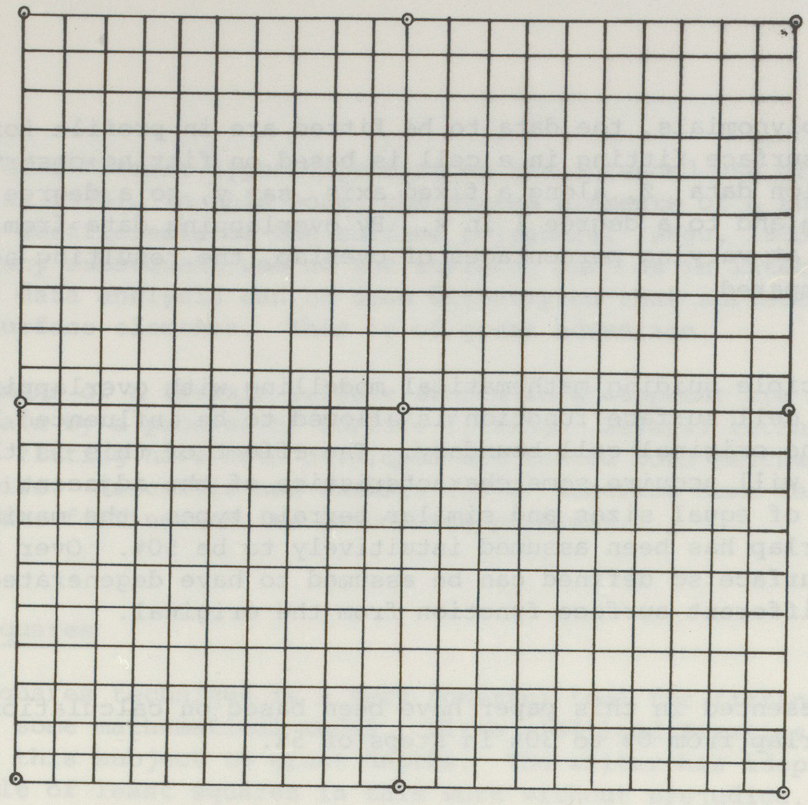

FIG. 1. TYPICAL CELL WITH 9 "STANDARD" CHECK POINTS

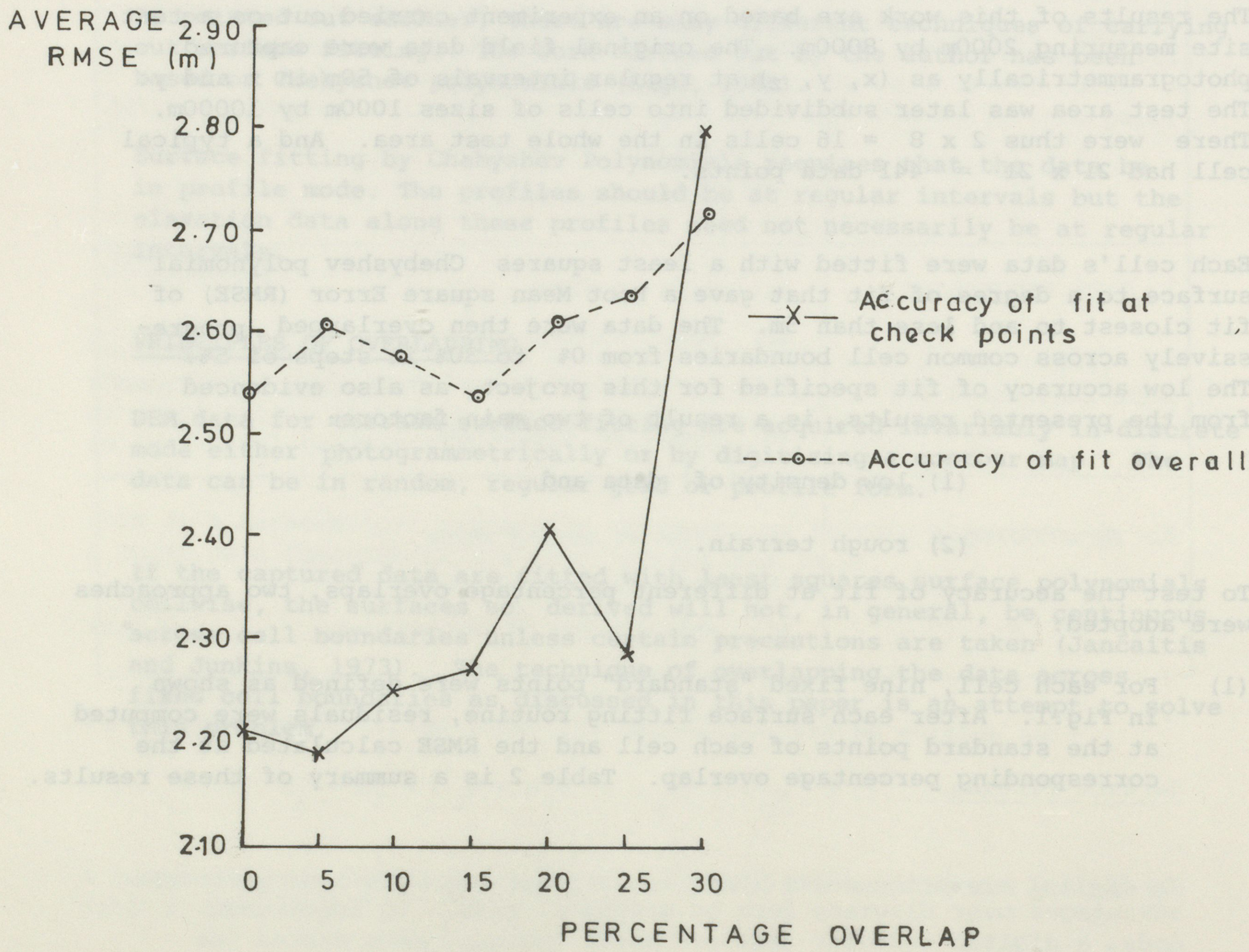

FIG. 2. COMBINED RESULTS

TABLE 2: SUMMARY OF RMSE AT CHECK POINTS

CELL NO	RMSE (M) % OVERLAP						
	0%	5%	10%	15%	20%	25%	30%
A1	1.90	2.32	1.80	1.64	1.83	2.45	3.02
A2	4.04	3.75	1.63	1.71	1.80	1.87	1.93
B1	0.96	1.63	1.69	1.63	2.08	2.33	2.71
B2	3.73	0.21	0.60	1.73	2.08	2.21	3.00
C1	3.89	2.82	2.23	2.06	2.05	2.14	2.20
C2	2.82	2.33	2.06	1.93	0.86	1.10	3.01
D1	3.13	2.98	2.85	3.23	3.71	2.41	2.41
D2	3.35	3.22	3.19	2.29	2.75	2.15	2.56
E1	0.76	0.85	0.83	0.63	1.28	-	-
E2	1.27	1.20	-	-	-	-	-
F1	1.29	1.93	2.36	2.83	2.43	1.49	1.49
F2	2.35	0.91	0.64	0.87	3.01	1.05	-
G1	1.67	2.04	2.61	2.67	2.77	2.82	3.43
G2	1.83	2.08	2.58	3.24	3.43	3.70	3.88
H1	0.12	1.86	3.13	2.65	3.71	4.03	3.97
H2	-	4.95	4.50	4.94	-	-	-

TABLE 3: SUMMARY OF RMSE OVERALL

CELL NO	0%	5%	10%	15%	20%	25%	30%
A1	2.34	2.69	1.85	2.11	2.36	2.57	2.77
A2	2.73	2.87	2.27	2.28	2.28	2.27	2.27
B1	2.87	2.00	2.18	2.42	2.70	2.30	2.45
B2	1.85	2.28	1.89	2.18	2.41	2.60	2.78
C1	2.39	2.49	2.56	2.60	2.62	2.65	2.69
C2	2.28	2.43	2.61	2.83	2.45	2.52	2.59
D1	2.55	2.70	3.00	2.52	2.89	2.27	2.57
D2	1.88	2.15	2.56	1.65	1.78	2.78	2.77
E1	2.80	2.75	2.93	2.97	2.96	-	-
E2	2.40	3.00	-	-	-	-	-
F1	2.54	2.71	2.88	2.91	2.93	2.97	2.88
F2	2.61	2.67	2.70	2.66	2.71	2.89	-
G1	2.70	2.79	2.86	2.67	2.83	2.72	2.91
G2	2.75	2.92	2.98	2.86	2.84	2.98	2.98
H1	2.88	2.52	2.76	2.48	2.58	2.77	2.93
H2	2.99	2.81	2.73	2.96	2.87	2.60	2.72

TABLE 4: SUMMARY OF DEGREE OF FIT

CELL	DIGREE OF FIT AT DIFFERENT % OVERLAP						
NO.	0%	5%	10%	15%	20%	25%	30%
A1	4	4	5	5	5	5	5
A2	2	2	3	3	3	3	3
B1	1	1	2	2	2	2	2
B2	1	1	2	2	2	2	2
C1	2	2	2	2	2	2	2
C2	1	1	1	1	2	2	2
D1	1	1	1	2	2	3	3
D2	1	1	1	2	2	4	5
E1	14	16	17	18	19	-	-
E2	19	20	-	-	-	-	-
F1	7	8	9	10	13	17	19
F2	5	11	14	16	18	20	-
G1	8	8	9	10	10	11	11
G2	9	9	9	10	10	9	9
H1	8	9	9	10	10	10	10
H2	3	4	5	5	6	7	7

(2) Each new cell, defined by the appropriate percentage overlap, is tested for accuracy of fit by computing the residuals at all points and the corresponding RMSE calculated. Table 3 is a summary of these results.

The corresponding degrees of fit at different percentage overlaps is summarised in Table 4. In some cells, the terrain is so rugged that the available data were not sufficient to define a polynomial surface to the required specifications of accuracy of fit. These cells are shown with a dash (-) in Table 3. at the corresponding percentage overlap.

A further analysis of the results is shown by Fig. 2. In this figure, the RMSEs in all cells at a corresponding percentage overlap are added together and the average value computed both for the check points and for the overall. Since the RMSEs apply to the same corresponding cells, these averages at different percentage overlaps do reflect the differences in accuracy of fit in the test area at the different overlaps.

Judging from the accuracy of fit of a fixed cell, the optimum percentage overlap for this test site at the given specification (i.e. accuracy of fit) is 15%. Overall, the optimum percentage overlap is 25%. These figures, however, will have to be taken with caution since the graphs in Fig.2 do not show definite minima.

DISCUSSION AND RECOMMENDATIONS

The method of least squares polynomial surface fitting with overlapping data has been developed with an assumption that one is working with a fixed grid (i.e. cell) size pattern. Therefore an overlap will have the meaning of an expansion of the cell size by an appropriate percentage in cell size. The data within a new cell area will be used to define the corresponding mathematical surface.

Differences in RMSE of fit across cell boundaries is an indication of a discontinuity of surfaces across those cell boundaries. These differences are not zero as shown in the tables above; but rather they are kept to a minimum within working accuracy. Where the terrain is broken and the observed data are not of high enough density, this method fails-as shown in Table 4.

The main factors influencing the use of this method for surface fitting are:

(1) type of terrain: different cells of different terrain types give different percentage overlap for the same accuracy of fit. This aspect requires further investigation on the type of terrain and the corresponding percent of overlap.

............/6

(2) density of data: some cells could not be fitted to the required accuracy of fit for lack of sufficient data. This implies a need for differential density for different types of terrain for a required accuracy of fit and fixed cell size or a variable cell size for a fixed density of data.

(3) Percentage overlap in relation to terrain type and density of data. This will have to be investigated further in order to establish their relationships.

CONCLUSION

The method of least squares polynomial surface fitting with overlapping data has been shown to be a working alternative in functional digital elevation modelling. However the actual amount of data overlap in percent is a function of many factors the main ones being terrain type, data density and accuracy of fit. The author is still working on this to establish these parameters and their relationships.

REFERENCES

Allam, M.M., 1978. DTM application in topographic mapping. Photogrammetric Engineering and Remote Sensing, Vol. 44(12): 1513 - 1520.

Berztiss, A.T., 1964. Least squares fitting of polynomials to irregularly spaced data. SIAM Review, vol. 6 (3): 203 - 227.

Birge, R.T., 1947. Least - squares' fitting of data by means of polynomials. Reviews of Modern Physics, vol. 19 (4): 298 - 347.

Collins, S.H., 1975 Terrain parameters directly from a digital terrain model. The Canadian Surveyor, vol. 29(5): 507 518.

Doyle, F.J., 1978
Digital terrain models: An overview. Photogrammetric Engineering and Remote Sensing, vol. 44 (12): 1481 - 1485

Grist, M.W., 1972. Digital ground models: An account of recent research. Photogrammetric Record, vol. 7 (40): 424 - 441.

........../7

Jancaitis, J.R. and Junkins, J.L., 1973

Modelling irregular surfaces. Photogrammetric Engineering, vol. 39 (4): 413 - 420.

Leberl, F., 1973. Interpolation in square grid DTM. I.T.C. Journal, 1973-5: 756 - 807.

De Masson d'Autume, G., 1979. Surface modelling by means of an elastic grid. Photogrammetria, vol. 35 (2): 65 - 74.

Maxwell, D.A. and Turpin, R.D., 1968

Numeric Group Image (NGI) Systems Design. Texas Transportation Institute, Texas A & M University, Research Report No.120-1, 30 pages.

Schut, G.H., 1976. Review of interpolation methods for digital terrain models. XIIIth Congress of ISP, Commission III, Helsinki, invited paper, 24 pages.

Segu, W.P., 1985. Terrain approximation by fixed grid polynomial. Photogrammetric Record, vol. 11(65): 581 - 591.

Turner, A.K., 1978. A decade of experience in computer aided selection. Proceedings of ASP Symposium on digital terrain models (DTM), St. Louis: 318-340.