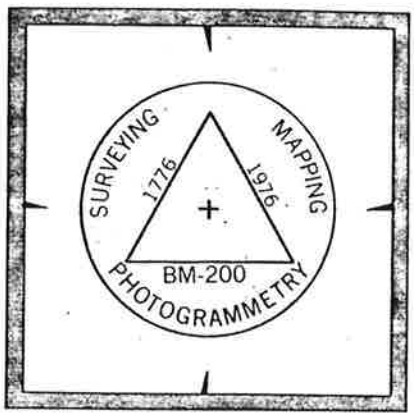

TWO CENTURIES OF SERVICE

PROCEEDINGS OF THE AMERICAN SOCIETY OF PHOTOGRAMMETRY

42ND ANNUAL MEETING
FEBRUARY 22-28, 1976
WASHINGTON, D.C.

(76-269)

INTERPOLATION OF A SURFACE FROM SETS OF DISCRETE HEIGHT DATA OF DIFFERENT STATISTICAL CHARACTERISTICS

Dr. Franz Leberl
NRC Resident Research Associate
Space Sciences Division
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91103

BIOGRAPHICAL SKETCH

Franz Leberl graduated in 1967 as a geodetic "diploma engineer" from the Technical University of Vienna, Austria. After military service in Austria in 1968, he joined the scientific staff of the International Institute for Aerial Survey and Earth Sciences (ITC), Netherlands, in 1969. In 1972, he obtained a doctorate from the Technical University of Vienna, Austria. His research concerned geometric aspects of non-conventional imaging sensors, interpolation techniques, digital terrain models, and reseau photography.

ABSTRACT

This paper presents and analyzes a method for the interpolation of a unique surface from two sets of independent digital height data of differing statistical characteristics. This method is based on linear prediction and thus relies on the concepts of auto- and cross-covariance functions.

The linear prediction algorithm for two sets of digital height measurements is first derived and then evaluated using the method of moving averages and bilinear interpolation for comparison. It is found that the overall root mean square interpolation errors of linear prediction are similar to those from moving averages and bilinear interpolation. This accuracy performance, together with the well known potential for controlled filtering of measuring errors and good-behavior in areas of poor control, makes linear prediction a versatile and general method for interpolating a unique surface from two sets of digital height data, with applications in photogrammetric mapping, remote sensing, and other fields.

1. INTRODUCTION

It is not uncommon in topographic mapping for surface heights to be sampled in two or more sets of measurements of differing origin and accuracy. In photogrammetry, such a case can occur in the overlapping area of two stereomodels, where each stereomodel provides a different set of measurements for the relief of a mapping surface. Another example is photogrammetric acquisition of digital terrain data by profiling along meanders: profiles scanned in opposite directions provide different sets of measurements. Two samples of surface heights are also available when tacheometric and photogrammetric data or data from different sensors (imaging radar and altimeters) have to be combined.

Kraus (1973) was the first to address the problem of merging two sets of photogrammetric measurements using linear prediction and filtering. The formulation chosen by Kraus required a restrictive a priori assumption that makes the prediction algorithm applicable to only two sets of measurements of identical statistical properties. The current paper will extend the linear prediction algorithm to applications to two sets of measurements of differing statistical behavior.

In the terminology of statistics, the different sets of measurements are taken from a random function. In the general case, the random function could be n-dimensional, and defined on an m-dimensional space. In the case of topographic mapping, this random function (the terrain surface) is one-dimensional, and defined on a two-dimensional reference space (Leberl, 1975). The present study will limit itself to this case of topographic mapping.

A linear prediction algorithm will first be derived for the case of more than one set of measurements and then evaluated by comparing it to other interpolation methods using simulated surfaces and measurements. It will be shown that root mean square interpolation errors from linear prediction are about equal to or only slightly smaller than those from a moving average algorithm or bilinear interpolation.

2. STATISTICAL MODEL OF A SURFACE AND OF ITS MEASUREMENTS

A surface $z = z(x,y)$ can be represented as the sum of a trend $t(x,y)$, a signal $s(x,y)$, and almost uncorrelated component $r(x,y)$:

$$z(x,y) = t(x,y) + s(x,y) + r(x,y) \qquad (1)$$

Figure 1 illustrates the concepts of trend, signal, and component r in one rather than two dimensions. Trend t is not considered to be a random function but a deterministic entity. Signal s is a random function, whose statistical properties are described by a covariance function $C_s(d)$, with d being a distance. The covariance of r is denoted by $C_r(d)$. The following relations hold:

$$\lim_{d\to\infty} C_s(d) = 0$$

$$C_r(d) = 0 \text{ for } d \neq 0$$

$$C_r(d) \geq 0 \text{ for } d = 0$$

For the benefit of a simplified notation, the argument of the covariance function will not be explicitly indicated (thus $C_s = C_s(d)$). The concepts of trend, signal, noise, and covariance function have been explained in detail in the photogrammetric literature by Kraus and Mikhail (1972). This explanation is, therefore, not repeated here.

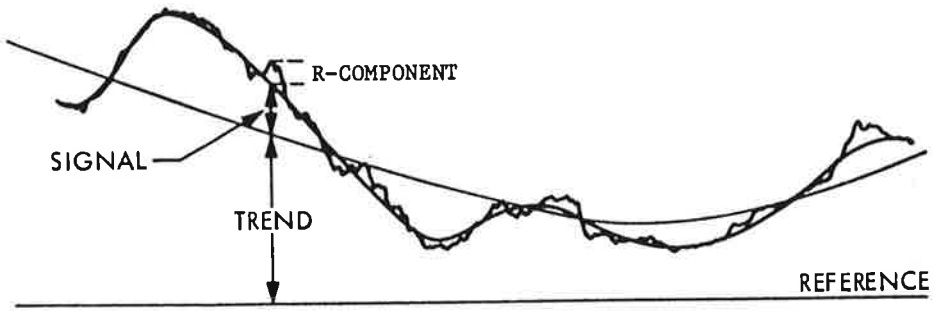

Figure 1. Illustration of the decomposition of the terrain surface into a trend (long-term variation), signal (random function with extended auto-covariance function), and a random function with auto-covariance function approaching zero for distances $d > 0$.

A discrete measurement $z_i = z'(x_i, y_i)$ of the surface $z(x,y)$ is by necessity erroneous. If the measurement is photogrammetric, then it is indirect; namely, through a stereoscopic model. In this case of indirect measurement, measuring errors can in the general case consist of a trend, signal, and noise:

$$z'_i = z_i + t'_i + s'_i + r'_i \qquad i = 1,2, \ldots n \tag{2a}$$

$$= t_i + s_i + r_i + t'_i + s'_i + r'_i$$

A second set of independent measurements z''_j consists of

$$z''_j = z_j + t''_j + s''_j + r''_j, \qquad j = 1,2, \ldots m \tag{2b}$$

The statistical properties of s'_i, r'_i, s''_j, r''_j are again described by covariance functions C'_s, C''_s, C'_r, C''_r.

Linear prediction will here be applied to random functions, but not to the deterministic components t, t', t". The components e', e" of the measurements after trend elimination are composed of

$$e'_i = s_i + s'_i + r_i + r'_i \tag{3a}$$

$$e''_j = s_j + s''_j + r_j + r''_j \tag{3b}$$

and the covariances of e', e" are C'_e, C''_e. Linear prediction should now provide estimates of s on the basis of e', e". For this purpose, the relationships among covariance functions have to be established. It is assumed that s is not cross-correlated with s' and s", and that r, r', r" is completely uncorrelated.

One thus obtains

$$C'_e = C_s + C'_s + C'_r + C_r \tag{4a}$$

$$C''_e = C_s + C''_s + C''_r + C_r \tag{4b}$$

$$C_{ee'} = C_s + C_{ss'} + C_r \tag{4c}$$

$C_{ee'}(d)$ is the cross-covariance among the random components of the measurements, and $C_{ss'}(d)$ is the cross-covariance of the signal of the measurement error. C'_e, C''_e and $C_{ee'}$ can be computed from the measurements; C_s, C'_s, C''_s, $C_{ss'}$ are unknown.

The above represents the case of two sets of measurements. If a third set is available, then one must add to Eqs. (4):

$$C'''_e = C_s + C'''_s + C'''_r + C_r$$

$$C_{ee''} = C_s + C_{ss''} + C_r$$

$$C_{e'e''} = C_s + C_{s's''} + C_r$$

Generalization to more than three data sets is straightforward. One thus obtains a set of constraints among those covariance functions which enter into a linear prediction algorithm.

3. LINEAR PREDICTION WITH MORE THAN ONE SET OF MEASUREMENTS

3.1 Prediction algorithm

Linear prediction is based on those components e_i', e_i'', . . ., of the original measurements which are left after trend elimination (see Eqs. 3). The prediction algorithms provide an estimate $\bar{s}_i = \bar{s}(x_i, y_i)$ of the surface signal s, using the following expression in matrix notation (Yaglom, 1962; Kraus and Mikhail, 1972):

$$\bar{s}_i = \underline{q} \cdot \underline{Q}^{-1} \cdot \underline{e}^T \tag{5}$$

where

$$\underline{e} = (\underline{e}', \underline{e}'', \ldots) = (e_1', e_2', \ldots e_n'; e_1'', e_2'', \ldots e_m''; \ldots)$$

Vector $\underline{e}$ contains the different sets of measurements reduced for trend. Matrix $\underline{Q}$ contains the covariances of vector $\underline{e}$ and thus specific values of the functions C_e', C_e'', C_{ee}'''. Vector $\underline{q}$ consists of specific values of the cross-covariance functions C_{se}', C_{se}'', ... , between signal s_i and vectors $\underline{e}'$, $\underline{e}''$, respectively.

3.2 Covariance functions

Linear prediction requires knowledge of covariance functions C_e', C_e'', ..., C_{ee}''', ... ; C_{se}', C_{se}'', ... The first two of these three groups of functions can be computed in a straightforward manner, employing the standard algorithm for covariance computation as described, for example, by Kraus and Mikhail (1972). Inputs to this computation are the sets of e_i', e_j'', ..., values obtained after trend elimination. The empirical covariance values obtained for a number of distance classes $d_i - \Delta d \leq d < d_i + \Delta d$ are used to fit a positive definite function with two variables C_0 and C_1. An example of such a function is:

$$C(d) = \frac{C(0)}{1 + d^2/C_1^2} \tag{6}$$

The discontinuity of C(d) for d=0 represents the sums $(C_{r'} + C_r)$, $(C_{r''} + C_r)$ (see Kraus and Mikhail, 1972).

Functions $C_{se'}, C_{se''}, \ldots$ cannot be computed directly. However, since it can be assumed that s is not cross-correlated with s', s" ...; r', r", ...; r, one obtains, from Eq. (3),

$$C_{se'} = C_{se''} = \ldots = C_s \tag{7}$$

A pair of equations (4a,b) allows the formulation of an explicit expression for C_s:

$$C_s = \frac{(C_{e'} - C_{r'} - C_r) \cdot C_{s''} - (C_{e''} - C_{r''} - C_r) \cdot C_{s'}}{C_{s''} - C_{s'}} \tag{8}$$

If more than two sets of measurements are available, then Eq. (8) can be formulated for any pair of these. Covariance functions $C_{s'}$, $C_{s''}$, ..., are required for the determination of C_s but are not accessible to direct computations. They describe the statistical properties of the signals s', s", ..., of the errors of measurement. If these functions cannot be

determined in a calibration procedure of the measuring process, then there is only the alternative of an a priori assumption. Postponing the discussion of the specific statistical meaning of such an assumption, it appears convenient to postulate proportionality of the covariance functions:

$$C_{s'} = a \cdot C_{s''} \tag{9}$$

Equation (8) can be evaluated for any assumed proportionality factor $a \neq 1$. For $a = 1$, however, C_s cannot be solved from Eq. (8). In that particular case, C_s can be found from Eq. (4c) if an assumption can be made concerning C_r and $C_{s's''}$, e.g.:

$$C_{s's''} = b \cdot C_{s'} = c \cdot C_{s''} \tag{10}$$

It should be stressed, that Eqs. (9) and (10) are at this time arbitrary choices from an infinite number of possible relationships. One could at this point expand to the case of more than two sets of measurements. The discussion will, however, be limited from here on to the particular case of only two sets of measurements.

3.3 Interpretation of covariance relations (9) and (10)

In order to attach some statistical relevance to parameters a, b, and c in Eqs. (9) and (10), it is helpful to examine the case of two sets of measurements, of which one is more accurate than the other. The measurements should thus be weighted appropriately:

$$C_{s'} = p^2 \cdot C_{s''} \tag{11}$$

If one would interpolate two surfaces, each from one of the two independent measurement groups, and then take a weighted mean, one would imply:

$$s' = -p \cdot s'' \tag{12}$$

Obviously Eq. (12) would in turn imply the validity of Eq. (11) and:

$$C_{s's''} = -p \cdot C_{s''} = -\frac{1}{p} \cdot C_{s'} \tag{13}$$

Thus

$$a = p^2; \qquad b = -p; \qquad c = \frac{1}{b}$$

However, one is not free to assume a value p intuitively if use is to be made of the computed covariance functions $C_{e'}$, $C_{e''}$, $C_{e'e''}$. The value must be chosen in such a way that Eqs. (4) are simultaneously fulfilled by Eqs. (11) and (13). This leads to

$$p = \frac{C_{e'} - C_{r'} - C_r - C_{e'e''}}{C_{e''} - C_{r''} - C_r - C_{e'e''}} \tag{14}$$

It should be noted here, however, that Eq. (14) can produce different values of p for different values of the independent variable d (distance) entering the covariance functions. Only if functions $C_{e'}$, $C_{e''}$ and $C_{e'e''}$ are strictly proportional can a single value p = constant be obtained. Figure 2 is an example of actually computed functions $C_{e'}$, $C_{e''}$, $C_{e'e''}$. These functions pertain to numerical experiments described in Section 5. The p-value resulting from the functions of Fig. 2 is 2.7. In Fig. 2 $C_{e'}$, $C_{e''}$ and $C_{e'e''}$ have their maximum at d=0. This is not always the case for $C_{e'e''}$. This in itself already would create a dependency of p on d.

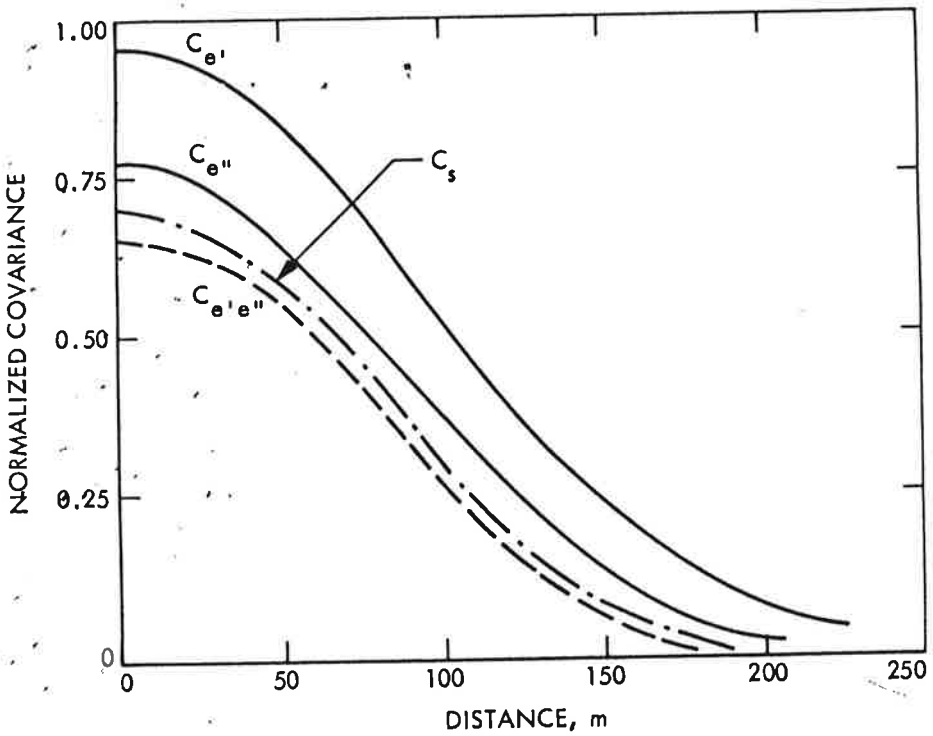

Figure 2. Covariance functions C_e', C_e'', C_{ee}''', and C_s as obtained for two simulated sets of measurements.

If indeed the covariance functions $C_{e'}$, $C_{e''}$, $C_{e'e''}$ are not proportional, then one must either choose an average value of $\bar{p}$ taken over a range of definition of the covariance functions, or C_s is evaluated separately for each distance d from Eq. (8).

It might be of interest to point out here that the case treated by Kraus (1973) is a specialization of the discussion presented in this paper. Postulating interpolation of a surface that represents the arithmetic mean of the two measurements, Kraus' formulation implies $p = 1$, and thus also $a = 1$, $b = c = -1$. Equation (8) can not be evaluated in this case. Equations (4), however, show that then

$$C_s = C_{e'e''} - C_{e'} = C_{e'e''} - C_{e''} \tag{15}$$

3.4 The concept of trend and trend elimination

Linear prediction is theoretically applicable to data which fulfill the criteria of statistical stationarity (mean is constant in different parts of a random function) and homogeneity (covariance function is constant). Actual data, however, only very rarely can be called stationary or homogeneous. Here the concept of trend is helpful. By breaking up an originally large data set into many subsets, computing a trend, and eliminating it from each data subset, one can hope to achieve stationarity and homogeneity of the residuals. These can then be further processed by linear prediction.

This procedure is quite straightforward and rather unambiguous for measurements with trend-free errors. For a single set of measurements, trend computation can even be carried out simultaneously with prediction in a process of least squares collocation (Moritz,1973). It turns out that results from linear prediction are not actually overly sensitive to the theoretical requirement of stationarity and homogeneity, and that no special care needs to be taken in the trend computation, which is generally just the definition of a largely overdetermined polynomial of low order (1st, 2nd, or 3rd order).

But when trend elimination concerns trend in the measuring errors, then its purpose is entirely different. If a significant trend is present in the measuring errors, then this means a low absolute accuracy of the measurements. Therefore, trend elimination represents a way of improving the absolute accuracy of the measurements prior to the actual interpolation. In the case of a single set of measurements, there is absolutely no means of defining this trend and eliminating its effect. If two independent sets of measurements are available, then some modest means is given to eliminate a trend from measuring errors and to improve the absolute accuracy of the measurements.

According to the model of topographic height measurements presented in Eq. (2), measurements z'_i, z''_j can in general contain three different trends: t, t', t". Trend elimination could start by computing a polynomial pol(x,y) using z' and z" simultaneously, and then defining two other polynomials pol'(x,y), pol"(x,y) using z' and z" separately.

If then, for example, measurements z' have trend-free errors, one can expect that (see Fig. 3)

$$t\,(x,y) = pol'(x,y)$$

$$t'(x,y) = 0$$

$$t''(x,y) = pol''(x,y) - pol'(x,y)$$

It should be obvious that elimination of a trend in the measuring errors will in most cases require intuitive judgment, particularly if data points are not uniformly distributed and their density varies from one data set to another.

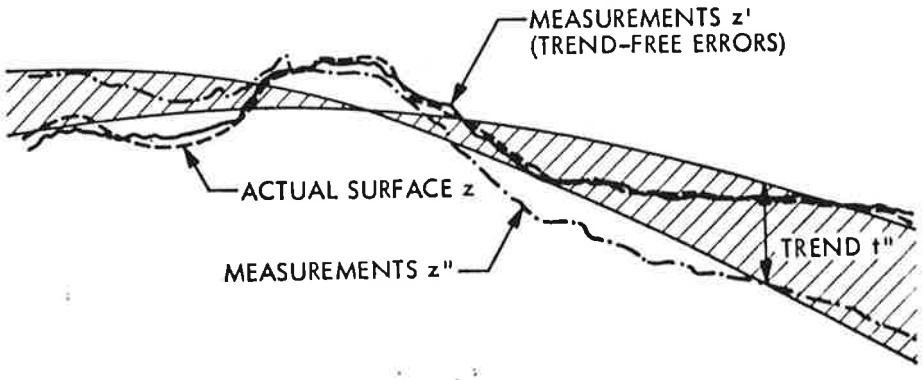

Figure 3. Illustration of trend and its elimination from the measurements. If one set of measurements has trend-free errors (is of high absolute accuracy), then a trend in the errors of the other measurements can be computed and eliminated (shaded area represents trend t" of measuring error in z").

4. OTHER METHODS OF MERGING TWO SETS OF DIGITAL HEIGHT DATA

Linear prediction is but one of many methods of merging two sets of measurements for interpolation of a unique surface. The methods of moving averages and bilinear interpolation are alternatives used in the present context to evaluate the overall performance of linear prediction.

A moving average algorithm selects the n measurements closest to the x_i, y_i location at which z is to be estimated (see Fig. 4). A polynomial of order m (usually m = 1, 2, or 3) is fit through the n data points, giving each of them a weight according to the distance from the estimation point. A moving average can be applied to two sets of measurements by appropriately weighting each of the sets. The weights w(d) attributed to the measurements can be of the type

$$w(d) = \frac{c_0}{1 + d^k} \tag{16}$$

where parameters c_0 and k must be chosen subjectively.

Bilinear interpolation is used in an algorithm selecting the four closest data points for each new computation. A bilinear polynomial is used to compute the new height independently in each set of measurements, thus resulting in two independent estimates $\bar{z}'$, $\bar{z}''$, from which a weighted mean $\bar{z}$ is computed.

5. NUMERICAL EVALUATION OF THE METHODS USING SIMULATED DATA

5.1 Surface measurements

A rather simple method for simulating trend, signal, and noise was devised for the study. This method is based on the fact that a differentiation of trend, signal, and noise is basically only a matter of

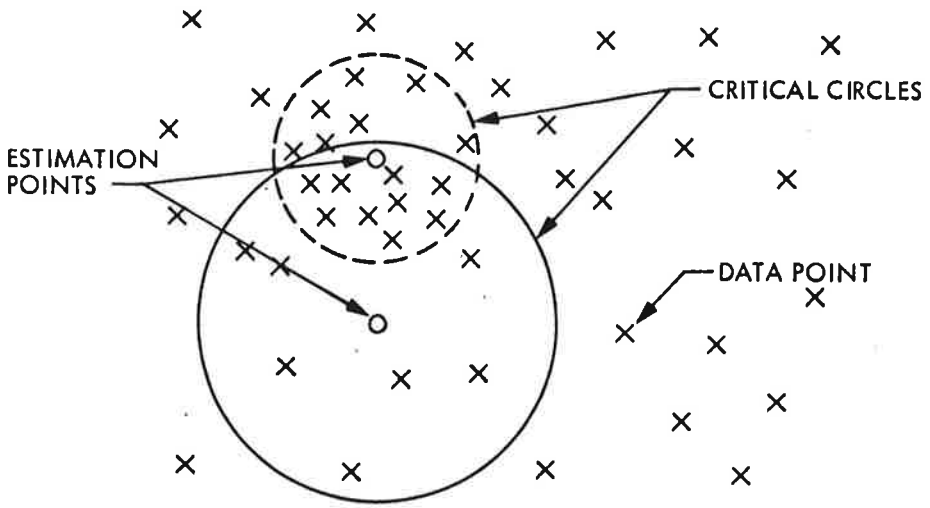

Figure 4. "Moving average" interpolation selects the n (n = 16 here) closest data points per estimation and fits a polynomial through the data using appropriate weights.

scale: what seems to be a trend at a certain scale might be a signal at a smaller scale, and be interpreted as noise if scale is further reduced. Recognizing this, one method can produce all components of a surface. Normally distributed values are generated in the intersection points of a square grid (see Fig. 5). Next, the slope of the simulated surface is computed in the directions of the gridlines, using adjacent height values according to Fig. 5. For each grid mesh, four height and eight slope values define a polynomial with 12 coefficients, using an approach devised by Jancaitis and Junkins (1973).

Input to the evaluation is then a set of simulated measurements of a surface, given on a grid of 10 x 10 points to be used for linear prediction. The center of each grid mesh serves as checkpoint. Figure 6 presents a set of profiles to illustrate an example of a pair of measurements of a surface.

5.2 Results

Tables 1 and 2 present the root mean square (rms) differences between the interpolated and the given height at 81 checkpoints. The computations used the simulated measurements illustrated in Fig. 6. The rms interpolation errors are shown as fractions of the rms errors which result from bilinear interpolation.

Interpolation with the moving average method employed a weight w(d) for each data point according to Eq. (16), using k = 4. Linear prediction requires four covariance functions, namely $C_{e'}$, $C_{e''}$, $C_{e'e''}$, obtainable directly from the measurements e', e", and C_g, which must be computed from Eqs. (14), (11), (8), or (15). Figure 2 illustrates the computed and smoothed covariance functions pertaining to the data shown in Fig. 6.

An obvious conclusion may be drawn from Tables 1 and 2: rms interpolation errors are similar for the methods of linear prediction, moving averages, and bilinear interpolation. Differences amount to only about 15%. This overall result confirms experimental data obtained in other studies on interpolation (Leberl, 1975). The main loss of information in a digital description of a surface occurs in the process of sampling

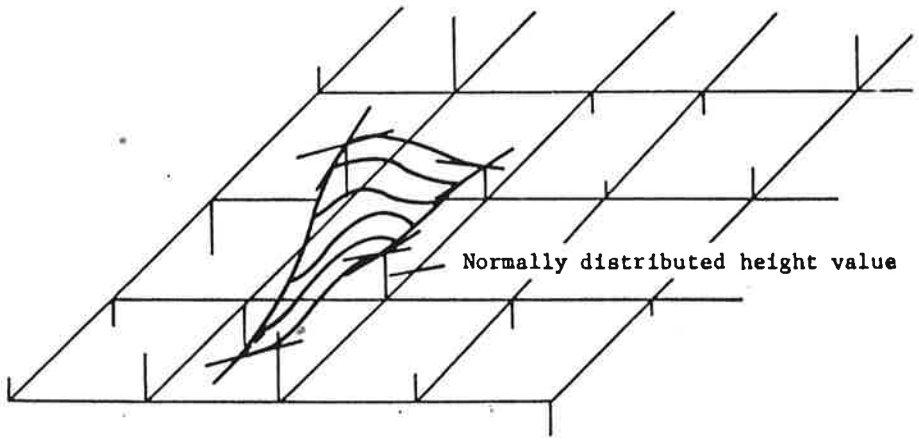

Figure 5. Generation of a random function using discrete random variables x generated on a regular grid.

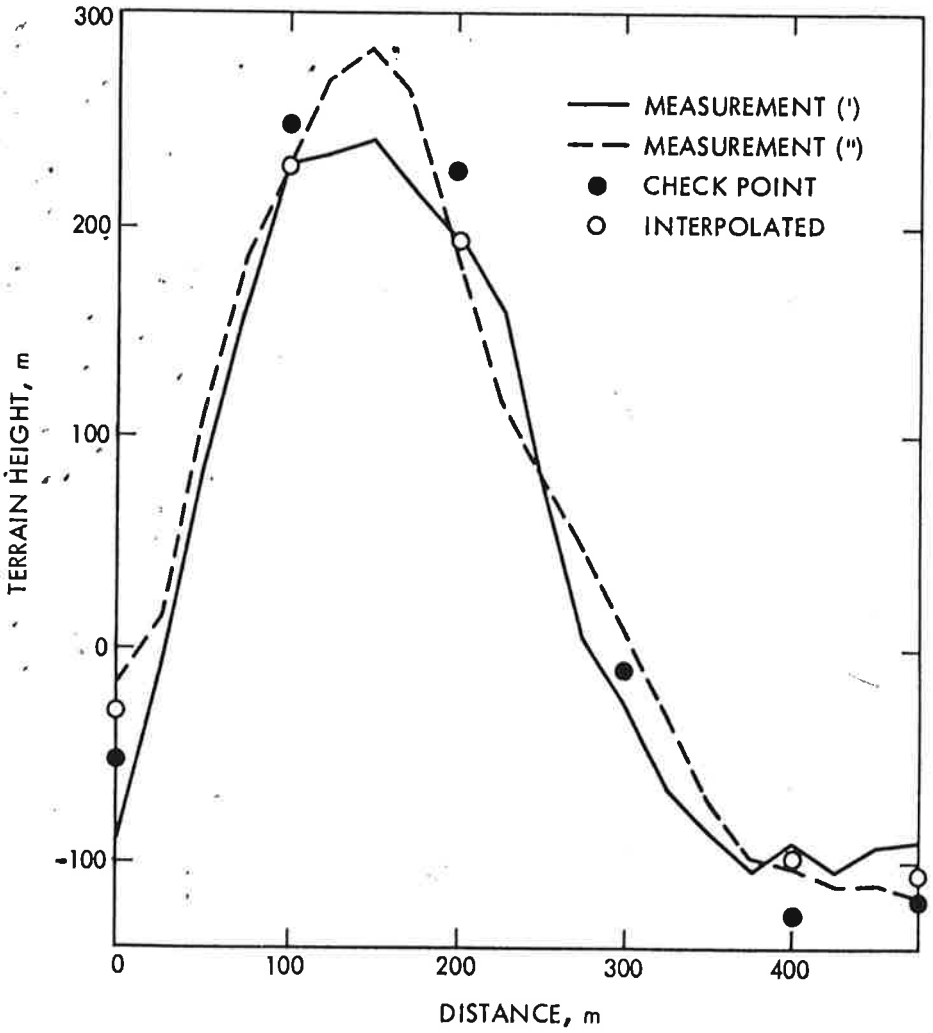

Figure 6. Profiles through two sets of simulated surfaces.

the surface by discrete measurements. It is a well established fact that the process of interpolation cannot compensate for this loss of information.

A number of different covariance functions of the type shown in Eq. (6) were used to study the sensitivity of linear prediction to the choice of covariance functions. The conclusion drawn from Table 2 is that linear prediction is not very sensitive to the choice of covariances. Similarly, the moving average method is not very sensitive to the choice of weights.

In any application of linear prediction, care must be taken in the conditioning of matrix $\underline{Q}$ in Eq. (5). This matrix approaches singularity if a large number of data points are used. In order to avoid singularity problems, the dimension of $\underline{Q}$ should not exceed a certain limit. Not only the dimension of $\underline{Q}$ but also the choice of covariance functions can contribute to ill-conditioning of $\underline{Q}$, as demonstrated by row 5 of Table 2.

Table 1. Results of moving average interpolation of 81 checkpoints from two sets of measurements. Values are rms discrepancies between interpolated and given heights in checkpoints, divided by the rms discrepancy obtained from linear interpolation.

Moving Average Method						
Weights	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
	$\frac{1}{1+d^4}$	$\frac{1}{1+d^4}$	$\frac{0.8}{1+d^4}$	$\frac{1}{1+d^4}$	$\frac{1}{1+d^4}$	$\frac{0.8}{1+d^4}$
Results	0.90		0.88		0.92	

Table 2. Results of linear prediction of 81 checkpoints from two sets of measurements. Values are rms discrepancies between interpolated and given heights checkpoints, divided by the rms discrepancy obtained from bilinear interpolation. Covariance functions are in the form shown in eq. (6). All C1 - values are assumed to be 100 m.

Linear Prediction Method					
Row	$C_{e'}-C_{r'}-C_r$	$C_{e''}-C_{r''}-C_r$	$C_{e'e''}$	C_s	Results
1	0.950	0.800	0.700	0.750	0.85
2	0.950	0.900	0.800	0.850	0.86
3	0.800	0.950	0.700	0.750	0.84
4	0.900	0.900	0.700	0.800	0.92
5	0.900	0.900	0.900	0.900	1.24*

*Covariance matrix (eq. 5) is near-singular.

The fact that the two sets of measurements are of different accuracy is reflected in the non-identical covariances $C_{e'}$, $C_{e''}$. The more accurate measurements have a smaller covariance of measuring errors. It should be stressed that, therefore, the overall covariances are also smaller than the ones of the less accurate measurements. Figure 2 thus shows that the measurements denoted by (") are more accurate. As shown in rows 3 and 4 of Table 2, interpolation errors in the case of the particular simulated data are not distinctly different, even if the covariances are not introduced properly into the linear prediction.

6. CONCLUSIONS

The method of linear least squares interpolation and filtering (linear prediction) is expanded in the present paper to apply to problems of interpolating a unique surface using two sets of discrete height data. Kraus (1973) was the first to address the photogrammetric problem of merging two sets of height measurements; however, he did so only for the special case of identical statistical characteristics of the two sets. The present paper derives an interpolation algorithm valid for

measurements of different properties. The statistical model includes correlated and uncorrelated random components in both sets of measurements and the surface itself. A method is presented for computing the auto- and cross-covariances required as input for linear prediction.

Evaluation of the method of linear prediction with the help of simulated height measurements led to the conclusion that the accuracy performance is about the same as with other interpolation methods. It was furthermore confirmed that in this expansion of linear prediction, too, the particular choice of covariances for input to the interpolation is not very critical. Even though the method of linear prediction is not very sensitive to the covariances used, an interpolation using this method can only be successful if the dimensions of the covariance matrix among data points are smaller than about 50 x 50 to avoid near-singularity of that matrix.

In conclusion, linear prediction in the extension presented is a valuable, flexible method of interpolation with an overall accuracy performance similar to other algorithms. But the added advantages typical for linear prediction, namely, well controllable filtering, statistical significance of the parameters of the method, and good behavior in areas of low data point density make the method of linear prediction a tool well suited to most interpolation problems.

ACKNOWLEDGEMENT

I am grateful to Prof. R.Rapp and Dr.R. Rommel of Ohio State University for a very critical review of the manuscript. Unfortunately time did not allow to do complete justice to their thoughtful comments, so that this conference preprint only contains the effect of part of their observations.

REFERENCES

Jancaitis, J., and J. Junkins (1973), "Modeling Irregular Surfaces," Photogrammetric Engineering, Vol. XXXIX, No. 4.

Kraus, K., and E.M. Mikhail (1972), "Linear Least Squares Interpolation," Photogrammetric Engineering, Vol. XXXVIII, No. 10.

Kraus, K. (1973), "Prädiktion und Filterung mit zwei verschiedenen Stützpunkt-Gruppen," Zeitschrift für Vermessungswesen, Vol. 98, No. 4, pp. 146-153, Stuttgart, West Germany.

Leberl, F. (1973), "Interpolation in Square Grid DTM," ITC-Journal, 1973-5, pp. 756-807, Enschede, The Netherlands.

Leberl, F. (1975), "Photogrammetric Interpolation," Photogrammetric Engineering and Remote Sensing, Vol. XLI, No. 5.

Moritz, H. (1973), "Least Squares Collocation," Publications of the German Geodetic Commission, Series A, No. 75, Munich, W. Germany.

Yaglom, A. (1962), An Introduction to the Theory of Stationary Random Functions, Prentice Hall, New Jersey.