# Quantification of Terrain Variation in Mountainous Regions based upon Numerical Map Analysis by means of Electronic Computer (Ⅱ)

Masami SHIBA, Toshiaki YAMAMOTO and Isao SASAKI

# 電算機を用いた数値地形解析による山岳地域の地形変動の計量化について（Ⅱ）

芝　正己・山本　俊明・佐々木　功

## Résumé

Trend surface analysis is a procedure for separating the relatively large-scale systematic variations, or trend in mapped data from essentially non-systematic small-scale variations, or residuals due to local effects. This is accomplished by fitted a trend function to a set of data values. Trend function of contour-type map is currently based mainly on applications of the polynomial and double Fourier models. Both stem directly from the general linear model, but the structure of a single map observation is different in the two models, and the kinds of fitted surfaces obtained from a given set of data generally differ in the pattern of their contour lines.

To provide an effective comparison of the polynomial and Fourier models fitted to the complex configuration of the actual land surface, a number of statistical tests are discussed which should assist in determining the optimum amount of complexity to ascribe to trend.

It is proposed that the configuration of the Fourier model obtained by evaluating the double Fourier series to the original land surface reveals a remarkable similarity compared with that of the polynomial model. Conversely, the goodness of fit test of higher-order polynomial model does not always reflect the spatial correspondence between the polynomial surface and the actual surface.

Choice of the two types of surface-fitting model, polynomials or double Fourier series, depends partly on objectives, and no simple conclusion can be given as to which type of surface-fitting model to use.

The techniques of numerical taxonomy could be use to procedure an objective classification system for land surfaces employing Fourier or polynomial coefficients.

# 要　旨

傾向面分析は数値化された事象に対して，体系的・規則的変動に起因する空間領域と非体系的偶然的変動に起因する領域とを分離する手法であり，原理的には前者が分布数値の全域的成分を示す傾向面で，後者が局地的成分を示す残差である。

最小自乗法による傾向面への近似関数（傾向面モデル）として，多項式と二重フーリエ級数を取り上げ，数値地形データに対する両関数の適合度および付随した統計的処理に基づく理論性について検討した。

複雑な地表面に対する適合度は，関数の増次に伴い漸次向上し分散分析の結果も増次の有意性が認められる（次数間の不偏分散の有意性）。同係数量を持つ多項式モデルに比べて，フーリエモデルの適合度は飛躍的に増加し，数値地形データの様に複雑な分布に対してはフーリエモデルの方がはるかに適合しやすいことを示している。これはフーリエモデルが低次でも多項式モデルに比べて多くの係数を持っていることに基づいている。

これに対して，多項式モデルでは増次に伴う適合度の向上が認められるにもかかわらず，不偏分散の有意性が棄却されない場合がある。このことは，誤差尺度としての適合度検定が，分布数値と近似値の空間的分布の一致する程度を，必ずしも反映しないことを示している（関数による分布数値の空間構造の規定）。

数値地形データに対する近似関数として，いずれを導入するかは一義的には決定されないが，比較的単純な構造の地表面（少なくとも低次の傾向面で等分散性が棄却される）に対しては，多項式モデルでも充分に利用しうるし，複雑な地表面をより詳細に近似したい時はフーリエモデルがすこぶる有効であることがわかる。

この様にして得られた地形の数式モデルの有効な利用法として，地形形状の地域間の類似性や差異に関する計量的地域区分を，クラスター分析を併用して，モデル間の変数の係数群についての距離行列の形で，処理して分類することが可能であることがわかった。


## Introduction

Trend-surface analysis of geomorphologic structure can be used to delineate gross structural patterns and thus separate regional from local effect. Resolution of regional, or large-scale aspects, and local small-scale aspects usually has been restricted to graphic methods because numerical methods required extensive computational labor.

Now, however, the computer and suitable programs make it possible to analyze trends within sets of data quickly and easily. Information can be reduced to a series of easily recognizable patterns which, in turn, provide a means of isolating and studying specific characteristics within these data[1,2,3,4,5,6,7,8,9,10,11,12].

Trend-surface analysis is currently used in ore reserve calculations, geochemical prospecting, analysis of geologic structure maps, and analysis of lithofacies maps[13,14,15,16].

A growing number of investigators are applying polynomial approximation or "trend" surfaces to the description of geological and geophysical measurements. Among these are Miller[17], Krumbein[18], Mandelbaum[19], Whitten[20], Merriam and Harbaugh[21]. An excellent review of trend surface methods and the general problem of mapping geological is given by Miller and Kahn[22]. The predominant effort has been to use non-orthogonal polynomials. Up to now, however, little use has been made of Fourier series for data a representation in geomorphology and geology[23][24][25][26][27][28][29][30][31][32][33].

An exception is the work of Preston and Harbaugh[34] who applied double Fourier series to determine whether complex topography can be represented by interacting harmonic terms.

This report emphasizes two aspects of the subject. The first is a comparison of Fourier model with polynomial model on a goodness of fit basis that express the percentage reduction in total corrected sum of squares, accounted for by the fitted surface and second is a consideration that brings out some basic similarities in the Fourier and polynomial models.

## Statistical Measures

*Goodness of Fit*[35][36]

Computation of coefficients of a trend function is only part of trend analysis. In addition it is essential to compute measures that express the goodness of fit of the trend function to the data and then to determine whether the trend function components are statistically significant. One measure is the amount of a variable, Z, from its mean value, which is an index of the totla variation within the entire data set. This is calculated by summing the squares of deviations from the mean. The simplest expression is

$$SS_T = \sum (Z_{obs} - \bar{Z}_{obs})^2 \qquad (1)$$

where $SS_T$ is the total corrected sum of squares of deviations from the mean, $Z_{obs}$ is the observed value of variable Z at data points and $\bar{Z}_{obs}$ is the arithmetic mean of observed values of Z. An alternative, short-cut method is

$$SS_T = \sum Z^2_{obs} - \frac{(\sum Z_{obs})^2}{n} \qquad (2)$$

where $n$ is number of data points. The total sum of squared deviations from the mean, in turn, may be regarded as consisting of two sources of variation, namely, that part contributed by the trend function, $SS_R$ and that part due to deviations from the trend function, $SS_D$:

$$SS_T = SS_R + SS_D \qquad (3)$$

This assumes, of course, that the least-squares criterion has been satisfied and that the coefficients of the trend function are linear. The sum of squares contributed by the trend function, $SS_R$, represents the squared difference between the predicted (or trend) values of Z from the mean value of Z,

$$SS_R = \sum (Z_{trend} - \bar{Z}_{obs})^2 \qquad (4)$$

The sum of squares due to deviations is a reflection of the failure of the trend values to coincide with observed values:

$$SS_D = \sum (Z_{obs} - Z_{trend})^2 \qquad (5)$$

In assessing goodness of fit, *RSS*, it is convenient to express that part of the total sums of squares accounted for by the trend function as a percentage because this permits ready comparison of trend functions fitted to different set of data.

One widely used measure, the percentage reduction in total corrected sum of squares accounted for by the fitted surface or, simply, percent of total sum of squares, is given by the expression,

$$RSS = \left[1 - \frac{SS_D}{SS_T}\right] 100\% \qquad (6)$$

which may also be written

$$RSS = \left[1 - \frac{\sum (Z_{obs} - Z_{trend})^2}{\sum (Z_{obs} - \bar{Z}_{obs})^2}\right] 100\% \qquad (7)$$

A perfect fit of a trend function to the data points would yield a value of 100 percent. A fit of 100 percent is uncommon, there being no deviations of data points from the trend function with a fit of 100 percent. It should be pointed out that a perfect fit will be obtained if the number of terms in the trend function equals the number of data points. There is little reason to fit a trend function under such circumstances, a geomorphologically reasonable fit possibly being obtained when gridded data are used, but a very unrealistic evaluation is likely to occur with irregularly spaced data. When the goodness of fit (*RSS*) of a trend function is low, it is a signal that most of the variation present in the data is not represent by the trend function. That is not necessarily bad, but interpretation of results should be made with this fact clearly in mind. With a low percentage of total sum of squares, the deviations or residuals may be geomorphologically significant, however.

*Statistical Tests of Trends*

The goodness of fit (*RSS*) of a trend surface may be tested statistically, by comparing the variance due to regression or trend to the variance due to deviations from the trend. It will be recalled that tests of equality of variances involve the $F$ distribution and are valid only if the data satisfy certain conditions. If these assumptions can justifiably be made, it may regard the coefficients of the trend function found by least squares as estimates of the true population regression coefficients, and test hypotheses about their nature. Assumptions concerning the data are:

1. The observed values are not clustered into groups. They either occur on a regular grid across the area or they occur at random.
2. The observed values are statistically normally distributed.

Assumptions concerning the deviations are:

3. The deviations are statistically normally distributed about the trend surface.
4. The deviations are uncorrelated with each other, that is to say, they are trend free across the area, and therefore they do not display autocorrelation.

The significance of a trend or regression may be tested by performing an analysis of variance, which is the process of separating the total variation of a set of observations into components associated with defined sources of variation. This, of course, has been done by dividing the total variation of Z into two components, the trend (or regression) and the residuals (or deviations). The degrees of freedom associated with total variation in a trend analysis are ($n$-1), where $n$ is the number of observations. The degrees of freedom associated with regression are determined by the number of terms or coefficients in the function fit to the data. Degrees of freedom for deviation are the number of degrees of freedom associated with total variation minus those that are accounted for by regression.

A formal analysis of variance (ANOVA) table is shown in Table 1. The mean squares are found by dividing the variances sum of squares by the appropriate degree of freedom. By reducing sums of squares to mean squares, they have been converted to estimates of variance and may be compared using a $F$ probability distribution. The MSd is the variance about the regression line; MSr is the variance of the regression line about the mean. If the regression is significant, the deviation about the regression will be small compared to the variance of the regression itself. In a general test of a trend-surface equation, the ratio of interest is that between variance due to regression and variance due to deviation. The $F$ test gives a probabilitistic answer to be the equation of whether the variances being examined have been obtained by random sampling from the same population. Or, is the regression effect not significantly different from the random effect? An affirmative answer may be interpreted as meaning that (a) the distribution of Z is random and independent of values of $X_1,\ldots,X_m$, or (b) the distribution of Z may be in part a function of $X_1,\ldots,X_m$, but the wrong functional model has been fit to the data.

In more formal terms, the $F$ test for significance of fit is a test of the hypothesis and alternative

$$H_0 : \beta_1=\beta_2=\cdots\cdots=\beta_m=0$$
$$H_1 : \beta_1, \beta_2, \cdots\cdots, \beta_m \neq 0$$

The hypothesis to be tested is that the partial regression coefficients are equal to zero, or in other words, there is no regression. If the computed value of $F$ exceeds the

Table 1. General ANOVA for Significance of Regression of $K$th-Degree Polynomial Trend Surface. Number ofC oefficients in Trend-Surface Equation, not Counting the $bo$ Coefficient, is $m$; Number of Data Points is $n$

| Source of variation | Sums of squares | Degrees of freedom | Mean square | F-test |
|---|---|---|---|---|
| Polynomial regression | SSr | m | MSr | |
| Deviation from polynomial | SSd | n-m-1 | MSd | MSr/MSd |
| Total variation | SSt | n-1 | | |

T $SSr=\sum(Ztrend-\bar{Z}obs.)^2$  MSr=SSr/m
TT $SSd=\sum(Zobs.-Ztrend)^2$  MSd=SSd/n-m-1
TTT $SSt=\sum(Zobs.-\bar{Z}obs.)^2$

table value of $F$, this hypothesis is rejected and the alternative, $H_1$ is accepted.

In trend-surface analysis, it is customary for some investigators to fit a series of equations of successively higher degrees to the data. In such an analysis, a number of regression sums of squares will be produced, each larger than the preceding sum. The analysis of variance table may be expanded to analyze the contribution of the additional partial regression coefficients and give a measure of the appropriateness of increasing the order of the regression. The test is developed by finding the difference in sums of squares due to regression of the higher polynomial or Fourier equation minus the regression sums of squares due to fitting the lower-order equation. This difference is divided by the difference in regression degrees of freedom, giving the mean square of regression due to increasing the degree of the polynomial. This mean square is then divided by the mean square due to deviation from the higher polynomial. If the resulting $F$ value is significant, the deleted order was contributing to the regression and should be retained. If the value is not significant, nothing has been gained by fitting the higher-degree polynomial. An ANOVA table for testing the significance of a higher-degree trend function is Table 2.

The $F$ test for significance of added terms is a test of the hypothesis and alternative

$$H_0 : \beta_{k+1}=\beta_{k+2}=\cdots\cdots=\beta_m=0$$
$$H_1 : \beta_{k+1}, \beta_{k+2}, \cdots\cdots, \beta_m \neq 0$$

The null hypothesis states that partial regression coefficients after the $k$th term are all equal to zero, or, they do not contribute to the regression caused by the 1 through $k$th term. Remember that the polynomial trend surface of degree $p$ contains $k$ coefficients, whereas the polynomial equation of the $(p+1)$ trend contains $m$ coefficients.

Table 2. General ANOVA for the Significance of Increasing the Degree of a Polynomial Trend from $p$- to $(p+1)$-Degree; Polynomial Equation of Degree $p$ has $k$ Coefficients, not Counting the $b_0$ Term; Equation of Degree $(p+1)$ has $m$ Coefficients, not Counting the $b_0$ Term; Number of Observations is $n$

| Source of variation | Sum of squares | Degrees of freedom | Mean squares | F-test |
|---|---|---|---|---|
| Regression of degree $(p+1)$ | SSrp+1 | m | MSrp+1 | MSrp+1/MSdp+1 [T] |
| Deviation from degree $(p+1)$ | SSdp+1 | n-m-1 | MSdp+1 | |
| Regression of degree $p$ | SSrp | k | MSrp | MSrp/MSdp [TT] |
| Deviation from degree $p$ | SSdp | n-k-1 | MSdp | |
| Regression due to increase from $p$ to $(p+1)$-degree | SSri= SSrp+1 - SSrp | m-k | MSri | MSri/MSdp+1 [TTT] |
| Total variation | SSt | n-1 | | |

T Test of significance of the $(p+1)$-degree trend surface.
TT Test of significance of the $p$-degree trend surface.
TTT Test of significance of increase in fit of the $(p+1)$-degree over $p$-degree.

*Interpretation of Confidence Intervals on Trend Surface*

In addition to determining the statistical significance levels of regression components in trend analysis, it is also possible to compute an envelope that represents a confidence level of some specified significance. Confidence lines can be fitted on each side of a trend line; for example, confidence surfaces can be fitted to trend surfaces, and confidence hypersurfaces to trend hypersurfaces. Krumbein[37] has outlined the mathematical details of geologic application of confidence surfaces to trend surfaces.

The distance from the trend surface to either of the two confidence surfaces that envelope it depends, in part, on the significance level that has been specified. If the significance level is low, the two confidence surfaces will be relatively close to the trend surface. Conversely, if the stated significance level is high, the confidence surfaces will be farther from trend surface. Contour values on confidence surfaces may be in the same units as the trend surface itself. Confidence surface may be defined either for all points on a trend surface, considered simultaneously for a given specified significance level, or instead they may pertain to points on the trend surface considered at only one geographic location at a time. For simultaneous consideration of all points, the confidence surfaces are farther removed from trend surface for a given significance level.

It slould be pointed out that in the calculation of significance levels and confidence surfaces pertaining to trend surfaces some of the basic assumptions that apply to these probability measures remain unsatisfied. One assumption is that repeated measurements at the same point will yield a frequency distribution of values of the dependent variable, or, in other word, values of Z according to notation. The variance of this distribution is termed that the "error variance", and is assumed to be the same at all points on the surface, or, in other words, at all values of the independent variable. Finally, the deviations from trend surfaces are assumed to be mutually uncorrelated. The other assumption, that the deviations be mutually uncorrelated, may or may not be satisfied. Generally, however, the deviation from trend surface will be correlated with each other, perhaps strongly so. A general effect is to overstate the significance level.

## Comparison of the Structure of High-Order Polynomial Functions and Double Fourier Series as the Quantitative Model

*The General Linear Model*[38]

The general linear model in its conventional form can be stated as follows:

$$Z=\beta_0+\sum_{i=1}^{k}\beta_i X_i+\varepsilon \qquad (8)$$

Where Z is an observable random variable ; $X_1$, $X_2$, ...,$X_k$ represent observable independent variables measured without error; the $\beta$'s are unknown parameters; and $\varepsilon$ is an unobservable random variable with mean zero and variance $\sigma^2$. For map analysis the general linear model is expressed in its two-dimensional form, as follows:

$$Z=\beta_{00}+\sum_{i=1}^{m}\sum_{j=1}^{n}\beta_{ij}X_i Y_j+\varepsilon \qquad (9)$$

Here $Z$ is an observable random variable (the mapped variable in this content), $X_1$, $X_2$, $\ldots, X_m$ and $Y_1$, $Y_2, \ldots, Y_n$ represent observable independent variables measured without error (these become functions of geographic coordinates in the present context), the $\beta$'s are unknown parameters (the coefficients of the fitted surfaces), and $\varepsilon$ is an unobservable random variable with mean zero and variance $\sigma^2$, representing the residuals on the fitted trend surface. The polynomial version of equation (9) involves simply the change of the subscripts of $X$ and $Y$ to superscripts:

$$Z=\beta_{00}+\sum_{i=1}^{m}\sum_{j=1}^{n}\beta_{ij}X^iY^j+\varepsilon \tag{10}$$

Here $X^i$ and $Y^j$ represent successive powers of the map coordinates $X$ and $Y$.

The Fourier model can be expressed in the general form of equation (9) as followed[40)41)42)43)]:

$$Z=\beta_{00}+\sum_{i=1}^{M}\sum_{j=1}^{N}F(\beta_{ij}P_iQ_j)+\varepsilon \tag{11}$$

where $P_i=2\pi_iX/M$ and $Q_j=2\pi_jY/N$. Here $M=m+1$, and $N=n+1$. The $\beta_{ij}$ of the general linear model of equation (9) now become a series of sine and cosine terms, yielding generally four Fourier coefficients for each $i$, $j$:

$$F(\beta_{ij}P_iQ_j)=CC_{ij}COSP_iCOSQ_j+CS_{ij}COSP_iSINQ_j+SC_{ij}SINP_iCOSQ_j +SS_{ij}SINP_iSINQ_j \tag{12}$$

Thus, instead of having simply $\beta_{23}$, for example, the corresponding Fourier coefficients are $CC_{23}$, $CS_{23}$, $SC_{23}$, and $SS_{23}$.

*Computation of Equation Coefficients*

The method of estimating the coefficient of trend functions fitted by least squares is basically the same for any variant of the general linear model regardless of the degree and number of variables involved. To utilize the general linear model, the experimenter selects a set of $X$'s, one from each of the $X_i$ either at random or by predetermined design. The first set of $X$'s will be denoted by $X_{11}$, $X_{12}, \ldots, X_{1k}$; then a $Z$ value is drawn at random and is denoted by $Z_1$. When these observed values are substituted into equation (8), we get

$$Z_1=\beta_0+\sum_{i=1}^{k}\beta_iX_{1i}+\varepsilon_1 \tag{13}$$

Then the experimenter selects another set of $X$'s one from each of the $X_i$, either at random or by predetermined design. This is repeated until $n$ sets are observed. The entire set of equations is written as

$$\begin{aligned} Z_1&=\beta_0+\beta_1X_{11}+\beta_2X_{12}+\cdots\cdots+\beta_kX_{1k}+\varepsilon_1\\ Z_2&=\beta_0+\beta_1X_{21}+\beta_2X_{22}+\cdots\cdots+\beta_kX_{2k}+\varepsilon_2\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ Z_n&=\beta_0+\beta_1X_{n1}+\beta_2X_{n2}+\cdots\cdots+\beta_kX_{nk}+\varepsilon_n \end{aligned} \tag{14}$$

or, more compactly, as

$$Z_j = \beta_0 + \sum_{i=1}^{k} \beta_i X_{ji} + \varepsilon_j \qquad j = 1, 2, \cdots\cdots, n \tag{15}$$

The least-squares estimates of the $\beta_i$ are obtained by minimizing the sum of squares of errors $\sum_{j=1}^{n} \varepsilon_j^2$. We obtain

$$L = \sum_{j=1}^{n} \varepsilon_j^2 = \sum_{j=1}^{n} (Z_j - \beta_0 - \sum_{i=1}^{k} \beta_i X_{ji})^2 \tag{16}$$

The values of the $\beta_i$ that produce the minimum sum of squares of errors are obtained by setting equal to zero the derivatives of $L$ with respect to each $\beta_i$, which are

$$\begin{aligned}
\frac{\partial L}{\partial \beta_0} &= -2\sum_{j=1}^{n} (Z_j - \hat{\beta}_0 - \sum_{i=1}^{k} \hat{\beta}_i X_{ji}) = 0 \\
\frac{\partial L}{\partial \beta_1} &= -2\sum_{j=1}^{n} (Z_j - \hat{\beta}_0 - \sum_{i=1}^{k} \hat{\beta}_i X_{ji}) X_{j1} = 0 \\
\frac{\partial L}{\partial \beta_2} &= -2\sum_{j=1}^{n} (Z_j - \hat{\beta}_0 - \sum_{i=1}^{k} \hat{\beta}_i X_{ji}) X_{j2} = 0 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\frac{\partial L}{\partial \beta_k} &= -2\sum_{j=1}^{n} (Z_j - \hat{\beta}_0 - \sum_{i=1}^{k} \hat{\beta}_i X_{ji}) X_{jk} = 0
\end{aligned} \tag{17}$$

If these equations are each divided by 2 and the term involving the Z's is transferred to the right-hand side of each equation, the normal equations are obtained:

$$\begin{aligned}
& n\hat{\beta}_0 + \hat{\beta}_1 \sum_{j=1}^{n} X_{j1} + \hat{\beta}_2 \sum_{j=1}^{n} X_{j2} + \cdots\cdots + \hat{\beta}_k \sum_{j=1}^{n} X_{jk} = \sum_{j=1}^{n} Z_j \\
& \hat{\beta}_0 \sum_{j=1}^{n} X_{j1} + \hat{\beta}_1 \sum_{j=1}^{n} X_{j1}^2 + \hat{\beta}_2 \sum_{j=1}^{n} X_{j1} X_{j2} + \cdots\cdots + \hat{\beta}_k \sum_{j=1}^{n} X_{j1} X_{jk} = \sum_{j=1}^{n} X_{j1} Z_j \\
& \hat{\beta}_0 \sum_{j=1}^{n} X_{j2} + \hat{\beta}_1 \sum_{j=1}^{n} X_{j2} X_{j1} + \hat{\beta}_2 \sum_{j=1}^{n} X_{j2}^2 + \cdots\cdots + \hat{\beta}_k \sum_{j=1}^{n} X_{j2} X_{jk} = \sum_{j=1}^{n} X_{j2} Z_j \\
& \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
& \hat{\beta}_0 \sum_{j=1}^{n} X_{jk} + \hat{\beta}_1 \sum_{j=1}^{n} X_{jk} X_{j1} + \hat{\beta}_2 \sum_{j=1}^{n} X_{jk} X_{j2} + \cdots\cdots + \hat{\beta}_k \sum_{j=1}^{n} X_{jk}^2 = \sum_{j=1}^{n} X_{jk} Z_j
\end{aligned} \tag{18}$$

These normal equations are system of $k+1$ equations in $k+1$ unknown (the unknowns being $\hat{\beta}_0$, $\hat{\beta}_1, \ldots, \hat{\beta}_k$). The solutions of these equations for the $\hat{\beta}_i$ are also maximum-likelihood estimators of the unknown parameters $\beta_i$ when the error $\varepsilon_j$ satisfy condition B; the solutions are best linear unbiased estimators when the error $\varepsilon_j$ satisfy condition A. The normal equations can be written in matrix form as follows:

$$\begin{pmatrix}
n & \sum_j X_{j1} & \sum_j X_{j2} & \cdots\cdots \sum_j X_{jk} \\
\sum_j X_{j1} & \sum_j X_{j1}^2 & \sum_j X_{j1} X_{j2} & \cdots\cdots \sum_s X_{j1} X_{jk} \\
\cdots & \cdots & \cdots & \cdots \\
\sum_j X_{jk} & \sum_j X_{jk} X_{j1} & \sum_j X_{jk} X_{j2} & \cdots\cdots \sum X_{jk}^2
\end{pmatrix} \times \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \cdots \\ \hat{\beta}_k \end{pmatrix} = \begin{pmatrix} \sum_j Z_j \\ \sum_j X_{j1} Z_j \\ \cdots\cdots \\ \sum_j X_{jk} Z_j \end{pmatrix} \tag{19}$$

In this equation the $\hat{\beta}_i$ -vector multiplied by the $X$ matrix is equal to the vector containing $Z$. This may be done by multiplying the $Z$ vector with the inverse of the $X$ matrix.

According the all the theory in the general linear model, it is necessary to expand the preceding situation into a polynomial or Fourier model in two variables. These models are special cases of the general linear model. We shall call these variables U and V, describe a cubic polynomial equation (20) for the general expression.

$$Y=a+bU+cV+dU^2+eUV+fV^2+gU^3+hU^2V+jUV^2+kV^3 \tag{20}$$

The principal item that deserves description is the form of the [U, V] matrix and its column vector [Y], and the coefficient vector is to be determined by obtaining the inverse of the [U, V] matrix, $[U,V]^{-1}$ matrix, and multiplying it by the [Y] column vector. Table 3 shows the complete $10\times10$ matrix and its column vector is based on

Table 3. Cubic and Column Vector for Orthogonal Polynomial Analysis of Map Data The Complete $10\times10$ Matrix and Its Column Vector is based on the General Expression for a Cubic Polynomial as follows;

$$Y=a+bU+cV+dU^2+eUV+fV^2+gU^3+hU^2V+jUV^2+kV^3$$

$$\begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ j \\ k \end{pmatrix} = \begin{pmatrix} \sum Y \\ \sum UY \\ \sum VY \\ \sum U^2Y \\ \sum UVY \\ \sum V^2Y \\ \sum U^3Y \\ \sum U^2VY \\ \sum UV^2Y \\ \sum V^3Y \end{pmatrix} \times \begin{pmatrix} N & \sum U & \sum V^{T} & \sum U^2 & \sum UV & \sum V^{2\,TT} & \sum U^3 & \sum U^2V & \sum UV^2 & \sum V^3 \\ \sum U & \sum U^2 & \sum UV & \sum U^3 & \sum U^2V & \sum UV^2 & \sum U^4 & \sum U^3V & \sum U^2V^2 & \sum UV^3 \\ \sum V & \sum UV & \sum V^2 & \sum U^2V & \sum UV^2 & \sum V^3 & \sum U^3V & \sum U^2V^2 & \sum UV^3 & \sum V^4 \\ \sum U^2 & \sum U^3 & \sum U^2V & \sum U^4 & \sum U^3V & \sum U^2V^2 & \sum U^5 & \sum U^4V & \sum U^3V^2 & \sum U^2V^3 \\ \sum UV & \sum U^2V & \sum UV^2 & \sum U^3V & \sum U^2V^2 & \sum UV^3 & \sum U^4V & \sum U^3V^2 & \sum U^2V^3 & \sum UV^4 \\ \sum V^2 & \sum UV^2 & \sum V^3 & \sum U^2V^2 & \sum UV^3 & \sum V^4 & \sum U^3V^2 & \sum U^2V^3 & \sum UV^4 & \sum V^5 \\ \sum U^3 & \sum U^4 & \sum U^3V & \sum U^5 & \sum U^4V & \sum U^3V^2 & \sum U^6 & \sum U^5V & \sum U^4V^2 & \sum U^3V^3 \\ \sum U^2V & \sum U^3V & \sum U^2V^2 & \sum U^4V & \sum U^3V^2 & \sum U^2V^3 & \sum U^5V & \sum U^4V^2 & \sum U^3V^3 & \sum U^2V^4 \\ \sum U^2V^2 & \sum U^2V^2 & \sum UV^3 & \sum U^3V^2 & \sum U^2V^3 & \sum UV^4 & \sum U^4V^2 & \sum U^3V^3 & \sum U^2V^4 & \sum UV^5 \\ \sum V^3 & \sum UV^3 & \sum V^4 & \sum U^2V^3 & \sum UV^4 & \sum V^5 & \sum U^3V^3 & \sum U^2V^4 & \sum UV^5 & \sum V^6 \end{pmatrix}^{-1}$$

T The Linear Portion of the Cubic Matrix
TT The Corresponding Quadratic Matrix

the general expression for a cubic polynomial equation (20). The linear portion of the cubic matrix is indicated by the L-shaped line in Table 3 that blocks out a $3\times3$ matrix. The corresponding quadratic matrix is indicated by the L-shaped line that blocks out a $6\times6$ matrix.

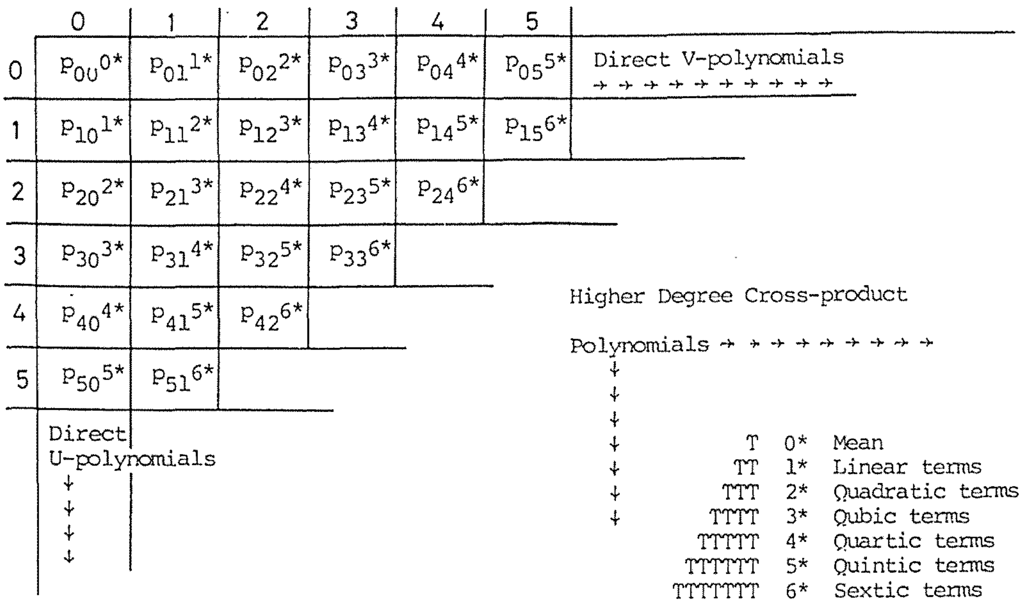
A finite definition of the double Fourier series for surface to gridded data is given in equation (11). The series is linear with respect to its coefficients and thus the least-squares method may be used to calculate the coefficient. Inasmuch as, this matrix may written as shown in Table 4.

Table 4. Matrix Equation for Least-Squares Determination of Fourier Series Coefficients

$$
\begin{pmatrix} CC_{00} \\ CC_{10} \\ \vdots \\ CS_{31} \\ \vdots \\ SS_{mn} \end{pmatrix}
=
\begin{pmatrix} \sum YA_0C_0 \\ \sum YA_1C_0 \\ \vdots \\ \sum YA_3D_1 \\ \vdots \\ \sum YB_mD_n \end{pmatrix}
\times
\begin{pmatrix}
\sum (A_0C_0)^2 & \sum A_1C_0A_0C_0 & \cdots & \sum B_mD_nA_0C_0 \\
\sum A_0C_0A_1C_0 & \sum (A_1C_0)^2 & \cdots & \sum B_mD_nA_1C_0 \\
\vdots & \vdots & \vdots & \vdots \\
\sum A_0C_0A_3D_1 & \sum A_1C_0A_3D_1 & \cdots & \sum B_mD_nA_3D_1 \\
\vdots & \vdots & \vdots & \vdots \\
\sum A_0C_0B_mD_n & \sum A_1C_0B_mD_n & \cdots & \sum (B_mD_n)^2
\end{pmatrix}^{-1}
$$

TT Notation: $A_i = COS(2\pi iU/M)$ $B_i = SIN(2\pi iU/M)$ Y = Mapped variable
$C_i = COS(2\pi iV/N)$ $D_i = SIN(2\pi iV/N)$
M = Maximum U value plus one (Fundamental Wavelength in U direction)
N = Maximum V value plus one (Fundamental Wavelength in V direction)
m = Maximum sine harmonic in U direction
n = Maximum sine harmonic in V direction

The coefficients associated with the polynomial model are commonly shown in diagonal arrangement, in which each succeeding polynomial surface (linear, quadratic, cubic, and higher-ordered surfaces) occupies a diagonal in the matrix of coefficients. James[30] developed a block arrangement for Fourier coefficients, in which successive blocks (Fourier surfaces) contain wavelengths of diminshing magnitude. Figure 1 (upper) shows the diagonal arrangement conventionally used for the polynomial model, in which the diagonal contours represent the usual sequence of linear, quadratic, and higher-ordered surfaces. The lower diagram, representing the block arrangement used for Fourier coefficients, has reverse L-shaped contours with values that define the successive blocks.



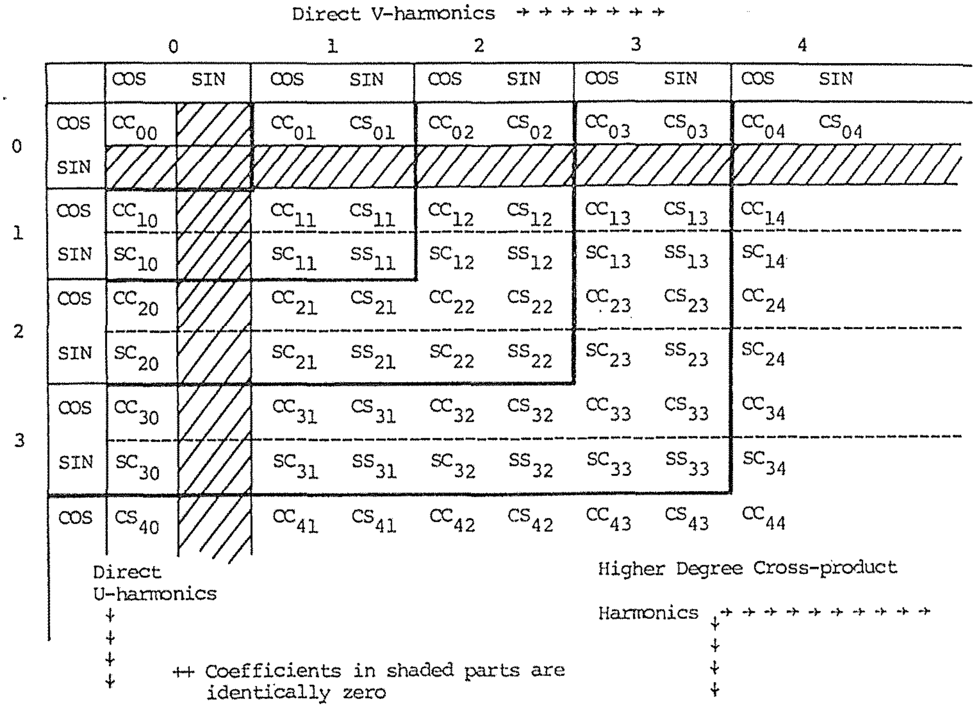



Figure 1. Diagrams Showing Arrangement of Polynomial Coefficient Terms (Upper) and Grouping of Double Fourier Series Coefficients according to Wavelength (Lower)

*Comparison of Geometric Properties of Fourier and Polynomial Models*[40)41)42)]

Because analysis employing double Fourier series models is a form of trend surface analysis, just as analysis with polynomials, it is important to compare differences as

Table 5. Comparison of Geometric Properties of Fourier with Polynomial Model

| **Polynomial Model** | | | | **Fourier Model** | | |
|---|---|---|---|---|---|---|
| Order | | Number of independent variables | Maximum number of extrema | Number of harmonics | Number of independent variables | Maximum number of extrema |
| Number $k$th | Name | $k(k+3)/2$ | $(k-1)^2$ | $k$th | $(2k+1)^2-1$ | $(2k)^2$ |
| 1 | Linear | 2 | 0 | | | |
| 2 | Quadratic | 5 | 1 | | | |
| 3 | Cubic | 9 | 4 | 1 | 8 | 4 |
| 4 | Quartic | 14 | 9 | | | |
| 5 | Quintic | 20 | 16 | | | |
| 6 | Sextic | 27 | 25 | 2 | 24 | 16 |
| 7 | Septic | 35 | 36 | | | |
| 8 | Octic | 44 | 49 | 3 | 48 | 36 |
| 10 | — | 65 | 81 | 4 | 80 | 64 |
| 12 | — | 90 | 121 | 5 | 120 | 100 |
| 14 | — | 119 | 169 | 10 | 440 | 400 |
| 16 | — | 152 | 225 | 20 | 1680 | 1600 |
| 17 | — | 170 | 256 | 30 | 3720 | 3600 |
| 19 | — | 210 | 324 | 40 | 6560 | 6400 |

similarities between the two forms of series when employed in surface fitting. Table 5 provides a comparison of two important geometric properties of the two types of surfaces, namely the maximum number of extrema (maxima and minima) on each surface, and the maximum number of inflection points on any profile through a surface. Direct comparison of double Fourier series surfaces with polynomial surfaces on a degree-by-degree basis is not feasible, however, because Fourier surface of a given degree contain more terms and are geometrically more complicated than polynomial surface of comparable degree.

Instead of the two types of surface may most effectively be compared according to approximate number of terms. Thus a third-degree polynomial surface that contains 10 terms is most nearly equivalent in number of terms and geometric properties to a first-degree double Fourier series surface containing a total of 9 terms (one zeroth and eight first-degree). Likewise, a sixth-degree polynomial surface (28 terms) is most nearly equivalent to a second-degree Fourier surface (25 terms) and an eight-degree polynomial surface (45 terms) to a third-degree Fourier surface (49 terms). A surface of specified degree in these example is defined as containing all possible terms pertaining to that degree, and terms of all lower degree, including the zeroth degree.

## Example of Numerical Terrain Data from Two Kinds of Topographic Map of Different Scale, 1/25000 and 1/50000

In recent studies for assessing the intricate landform variation and for recognizing the spatial structure of a dissected relief, these areas were analyzed by the authors with double Fourier series to determined whether complex configuration of the actual surfaec can be represented by interacting harmonic terms.

Therefore, a suitable data set and computer programs were already available a project which evaluated same sampling patterns. A rectangle with topographic map coordinates of X (east-west direction) and Y (north-south direction) and in meters measured from the borders was digitized every 5 mm on both maps of different scale which yielded an altitude matrix of 60 rows and 60 columns. Figure 2 is computer-drawn terrain block diagram and contour map produced by an automatic contouring program that uses X-Y plotter, for digital terrain data of Kanmuri District (1/50000 scaled map, area shown is 15×15 Km in extent) respectively.

A double Fourier series containing 0 to 30 terms in vertical and horizontal dimension of map was fitted to the data. The zero-zero term (arithmetic mean) is of no significance because the surface was leveled before the Fourier coefficints and power-spectrum square-root value were computed. Thus scanning of the power-spectrum square-root value provides a ready method of ascertaining the contribution of term of specified degree.

Figure 3 is manually contoured diagram of power-spectrum square-root values obtained according to fundamental waveform containing for four harmonics fitted to original data.

Comparison of the Fourier surface obtained by evaluating the double Fourier series with the original surface reveals a remarkable similarity for higher-degree terms. Despite the extremely complex configuration of the actual surface, the Fourier surface provides an approximation of the actual surface that represents more than 99 percent of the

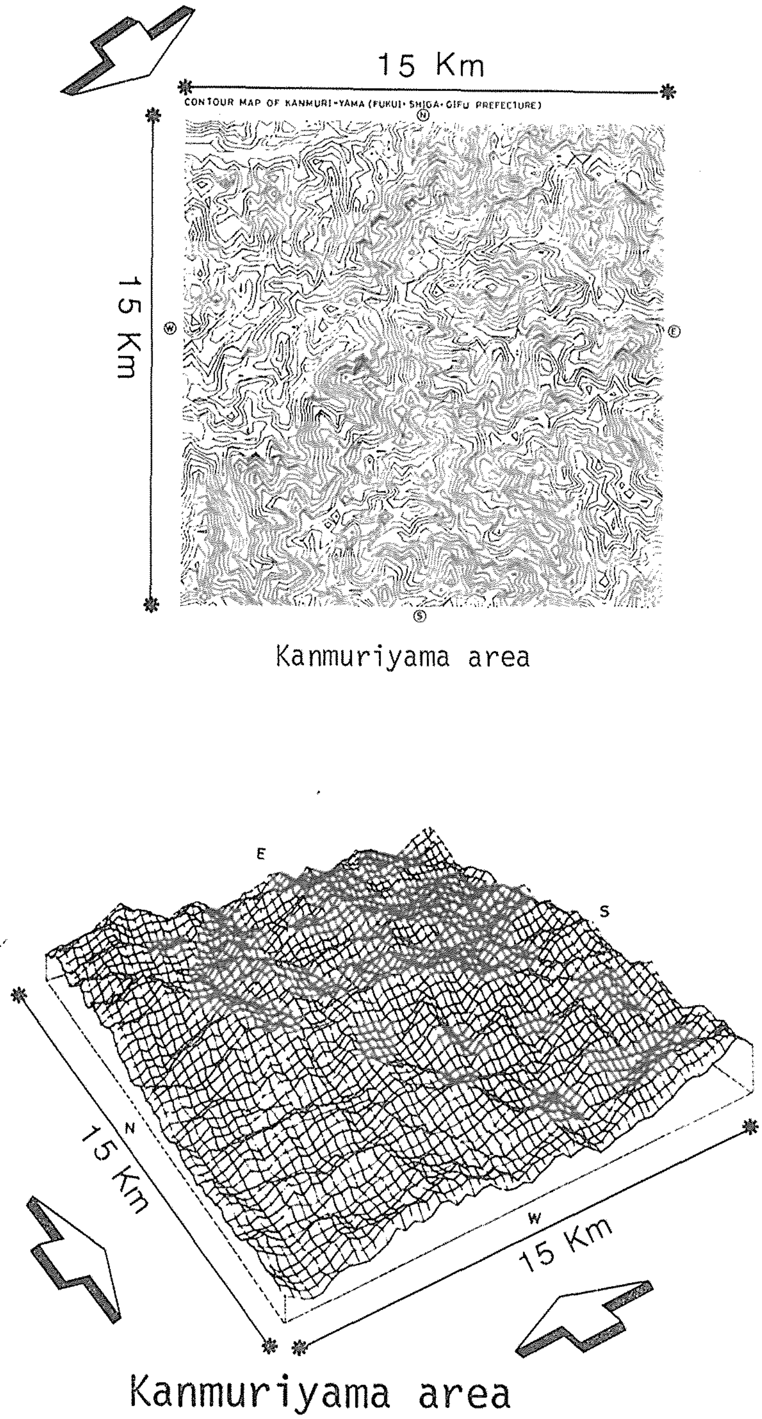



Figure 2. Computer-Drawn Contour Map (Upper) and Terrain Block Diagram (Lower) produced by an Automatic Contouring Program that used X-Y Plotter for Digital Terrain Data of Kanmuri District (1/50000 Scaled Topographic Map). Area shown is 15×15 Km in Extent. Contour Interval is 50m, Profiles drawn North-South and East-West.

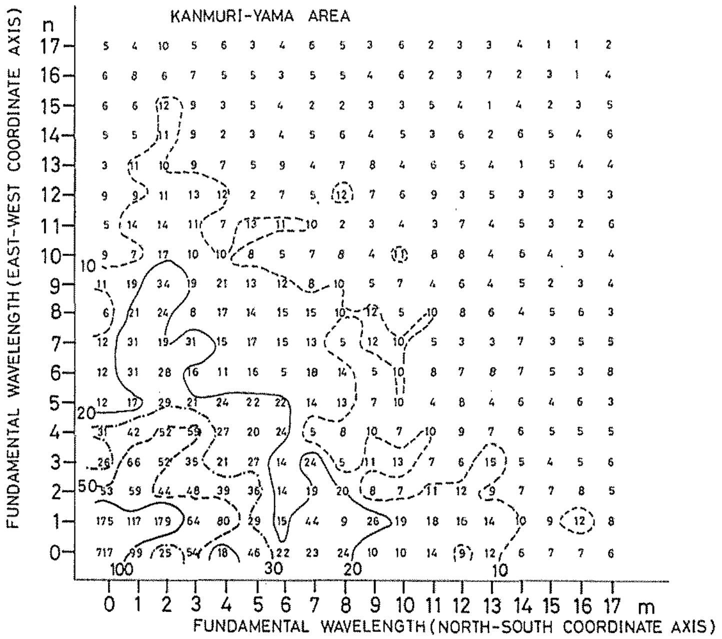



Figure 3. Manually Contoured Diagram of Power-Spectrum Square-Root Values obtainded according to Fundamental Waveform containing for Four Harmonics fitted to Original Data. Contour Interval are 100m, 50m, 30m, 20m, and 5m. The Dominant Wavelength or Periodic Component which is orientated along the Coordinates characterizes the Altitude Variation in Area.

total sum of squares of the original data.

The Fourier surface, however fails to accord with the actual surface at places along the edge of the map[43)44)]. This is because the map-edge values of the Fourier surface, prior to retilting, are the same at opposite points on any two edges. Leveling of the original surface, fitting of the Fourier surface, and subsequent retilting of the Fourier surface reduces, but does not eliminate, the edge effect. In such circumstances, there are almost no constraints on the form of the trend surface near the edge of the map. If high-order trends being fitted to data, extrapolated values near the edge of map may reach astronomical proportions. Minor edge effects will exist even if the entire map area is uniformly covered with control points up to the boundary. Therefore, the deleterious effect of edge distortions have evoked the authors to apply a "buffer region" which is an area in excess of the size of the area to be mapped.

Results of fitting polynomial and Fourier surfaces are shown in Figure 4 and 5. Here, the plots show goodness of fit expressed as the percent of total sum of squares of equivalent successive polynomial (upper) and Fourier (lower) surface fitted to original data. Goodness of fit of each sampling area becomes successively greater with increasing the number of terms in lower-order models, however, there being only a slight improvement in its value from higher-order models. First, second, and third-degree Fourier models are compared with polynomial models that most nearly accord in number of terms (third-, sixth-, and eighth-degree, respectively). Direct comparison of Fourier model with polynomial model on comparable degree-by-degree basis is that Fourier models provide better

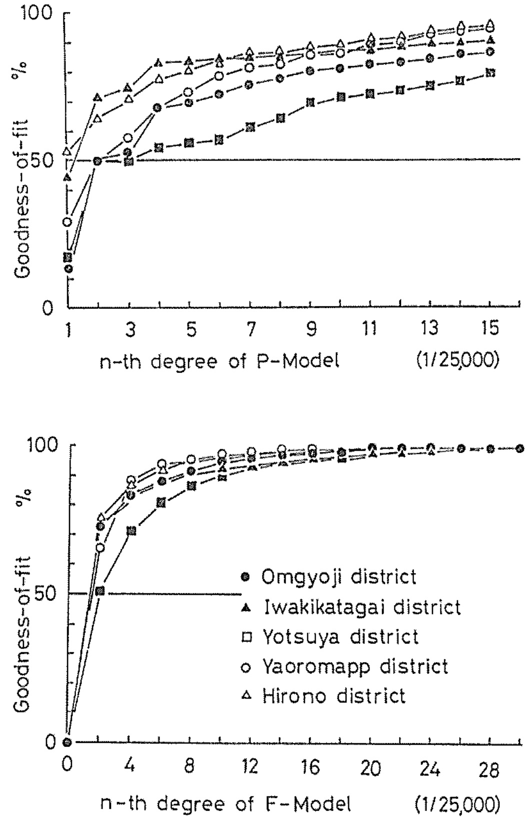



Figure 4. Plot of Goodness of Fit expressed as the Percent of Total Sum of Squares versus Successive Polynomial (Upper) and Fourier (Lower) Models fitted to Five Areas of 1/25000 Scaled Topographic Maps.

percentage fit to data than the corresponding polynomial model, and the Fourier models appear more nearly to represent the complexities present in the actual landform. A similar comparison my be made between the higher-order models.

Although by including only a few terms with high harmonics, the Fourier model yields a fitted surface with many undulations and for data that are nearly periodic, should provide a fit with fewer coefficients than that are needed for polynomial model. An important disadvantage from the convergence of functional approximation point of view, however, is that the linear, quadratic and cubic trend components dominates the polynomial model. Thus, predictions based only on these components of a trend that have higher terms may have limited the spatial structure because the predictions do not take account of higher systematic effects.

It also suggests that complex land surface can be quantitatively and objectively compared with each other by transforming data representing the surfaces to double Fourier series coefficients, and, in turn, using the Fourier coefficients as descriptors of the surfaces with the techniques of numerical taxonomy. It should be possible to produce an objective classification system for land surfaces employing the Fourier coefficients.

The authors have already tried to use coefficients of fifth-degree polynomial trend surfaces fitted to 39 DTMs which yielded an altitude matrix of 20 rows and 20 columns,

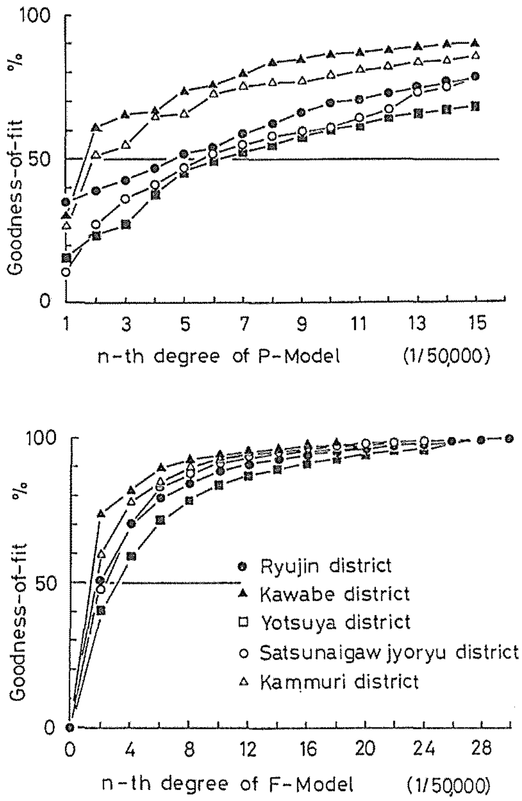



Figure 5. Plot of Goodness of Fit expressed as the Percent of Total Sum of Squares versus Successive Polynomial (Upper) and Fourier (Lower) Models fitted to Five Areas of 1/50000 Scaled Topographic Maps.

as numerical descriptors of land surfaces which permit calculation of similarity coefficients between different pairs of surfaces[45][46][47][48][49][50].

Fourier models may become confused, however, if there are linear trends in the data, because the sine and cosine functions have no linear components, and the fit is therefore distorted. For extrapolation Fourier models are usually worse than polynomial models, because they must repeat the pattern established over the region of data. In short, Fourier models are most useful for numerical terrain data that are periodic or oscillatory but have no linear trend.

The statistical significance of a trend function may be tested by separating the source of variation into components. The objective is to determine if components of a trend function are statistically significant, or whether they probably reflect chance alone. Thus, the authors indicated two analyses of variance to compare the statistical validity of both types of trend functions to the same data set. Table 6 and 7 show the ANOVA obtained as a result of trend surface fitting to the Hirono District and the Kanmuri District respectively.

In the analysis of variance format employed, the degree of trend surface; the coefficient of determination (the percentage reduction in total corrected sum of squares accounted for by the fitted surface); the confidence interval corresponding to a sig-

Table 6. Comparison of the Error Measures of Polynomial (Upper) and Fourier (Lower) Trend Surfaces fitting for the Hirono District (Fukui, Shiga and Gifu Prefectures) came from the Standard Topographical Map 1/25000. DTM Data by 3600 Grid-Cells digitized every 5mm on the Map (125m on the Ground) which yield a Matrix 60 Rows and 60 Columns.

POLYNOMIAL MODEL

| Degree of trend surface | Coefficient of determination (%) | Confidence interval at the 99% level $_nZ_{ij}\pm$ (m) | Two succeeding degree of trend surface $n$-order | $(n+1)$-order | Variance ratio between two succeeding degree of trend surface | Degree of freedom |
|---|---|---|---|---|---|---|
| 1 | 53.33 | 7.189 | mean | 1 | 1978.4521** | ( 2/3597) |
| 2 | 63.92 | 6.323 | 1 | 2 | 351.7264** | ( 3/3594) |
| 3 | 70.64 | 5.708 | 2 | 3 | 205.2330** | ( 4/3590) |
| 4 | 77.50 | 4.995 | 3 | 4 | 220.4166** | ( 5/3584) |
| 5 | 80.44 | 4.665 | 4 | 5 | 88.3404** | ( 6/3579) |
| 6 | 83.37 | 4.306 | 5 | 6 | 89.8628** | ( 7/3572) |
| 7 | 85.86 | 3.975 | 6 | 7 | 78.4300** | ( 8/3564) |
| 8 | 86.55 | 3.881 | 7 | 8 | 20.5180** | ( 9/3555) |
| 9 | 88.00 | 3.671 | 8 | 9 | 42.8028** | (10/3545) |
| 10 | 89.15 | 3.497 | 9 | 10 | 33.8327** | (11/3534) |
| 11 | 90.29 | 3.313 | 10 | 11 | 34.6959** | (12/3522) |
| 12 | 91.50 | 3.106 | 11 | 12 | 38.1971** | (13/3509) |
| 13 | 93.17 | 2.791 | 12 | 13 | 60.8463** | (14/3495) |
| 14 | 94.22 | 2.571 | 13 | 14 | 42.5473** | (15/3480) |
| 15 | 94.77 | 2.453 | 14 | 15 | 22.5804** | (16/3464) |

FOURIER MODEL

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 | 75.46 | 5.229 | mean | 2 | 148.9166** | ( 24/3574) |
| 4 | 87.20 | 3.806 | 2 | 4 | 57.6366** | ( 56/3518) |
| 6 | 92.61 | 2.929 | 4 | 6 | 28.5252** | ( 88/3430) |
| 8 | 94.69 | 2.527 | 6 | 8 | 10.8113** | (120/3310) |
| 10 | 95.98 | 2.252 | 8 | 10 | 6.6545** | (152/3158) |
| 12 | 96.90 | 2.037 | 10 | 12 | 4.8119** | (184/2974) |
| 14 | 97.63 | 1.851 | 12 | 14 | 3.8974** | (216/2758) |
| 16 | 98.09 | 1.737 | 14 | 16 | 2.5124** | (248/2510) |
| 18 | 98.45 | 1.667 | 16 | 18 | 1.7744** | (280/2230) |
| 20 | 98.84 | 1.553 | 18 | 20 | 2.0834** | (312/1918) |
| 22 | 99.12 | 1.495 | 20 | 22 | 1.4453** | (344/1574) |
| 24 | 99.35 | 1.467 | 22 | 24 | 1.1558** | (376/1198) |
| 26 | 99.60 | 1.417 | 24 | 26 | 1.2124** | (408/790 ) |

** Test of significance of increase in fit of the $m$-degree over $k$-degree trend surface at the 99% level of significance
Polynomial Model : $k=n$, $m=n+1$ Fourier Model : $k=2n$, $m=2(n+1)$

nificance level of 99 percent for all points on a trend surface; two succeeding degrees of trend surfaces from $n$- to $(n+1)$-order; the variance ratio between two succeeding degrees of trend surface, in which double star is significant at 99 percent level, single star, at 95 percent level; and the number of degrees of freedom.

From these analyses of variance (Table 6 and 7), it may be concluded that the con-

Table 7. Comparison of the Error Measures of Polynomial (Upper) and Fourier (Lower) Trend Surfaces fitting for the Kanmuri District (Fukui, Shiga and Gifu Prefectures) came from the Standard Topographical Map 1/50000. DTM Data by 3600 Grid-Cells digitized every 5mm on the Map (250m on the Ground) which yield a Matrix 60 Rows and 60 Columns.

POLYNOMIAL MODEL

| Degree of trend surface | Coefficient of determination (%) | Confidence interval at the 99% level $_nZ_{ij}\pm$ (m) | Two succeeding degree of trend surface $n$-order | $(n+1)$-order | Variance ratio between two succeeding degree of trend surface | Degree of freedom |
|---|---|---|---|---|---|---|
| 1 | 26.49 | 8.857 | mean | 1 | 2381.8100** | ( 2/3597) |
| 2 | 51.66 | 7.185 | 1 | 2 | 1214.6223** | ( 3/3594) |
| 3 | 54.70 | 6.959 | 2 | 3 | 60.3143** | ( 4/3590) |
| 4 | 65.03 | 6.119 | 3 | 4 | 211.6355** | ( 5/3584) |
| 5 | 66.31 | 6.011 | 4 | 5 | 22.7822** | ( 6/3579) |
| 6 | 73.07 | 5.379 | 5 | 6 | 128.0092** | ( 7/3572) |
| 7 | 75.40 | 5.147 | 6 | 7 | 42.1704** | ( 8/3564) |
| 8 | 76.80 | 5.005 | 7 | 8 | 23.9514** | ( 9/3555) |
| 9 | 77.17 | 4.972 | 8 | 9 | 5.5912** | (10/3545) |
| 10 | 79.19 | 4.754 | 9 | 10 | 31.3596** | (11/3534) |
| 11 | 81.04 | 4.546 | 10 | 11 | 28.4580** | (12/3522) |
| 12 | 82.01 | 4.436 | 11 | 12 | 14.6188** | (13/3509) |
| 13 | 83.95 | 4.199 | 12 | 13 | 30.1630** | (14/3495) |
| 14 | 84.61 | 4.120 | 13 | 14 | 9.9100** | (15/3480) |
| 15 | 85.52 | 4.005 | 14 | 15 | 13.6357** | (16/3464) |

FOURIER MODEL

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 | 59.97 | 6.556 | mean | 2 | 148.9171** | ( 24/3574) |
| 4 | 77.92 | 4.908 | 2 | 4 | 51.0631** | ( 56/3518) |
| 6 | 84.68 | 4.141 | 4 | 6 | 17.1896** | ( 88/3430) |
| 8 | 89.93 | 3.418 | 6 | 8 | 14.3552** | (120/3310) |
| 10 | 92.84 | 2.951 | 8 | 10 | 8.4470** | (152/3158) |
| 12 | 94.68 | 2.619 | 10 | 12 | 5.5870** | (184/2974) |
| 14 | 95.85 | 2.403 | 12 | 14 | 3.5735** | (216/2758) |
| 16 | 96.66 | 2.260 | 14 | 16 | 2.4614** | (248/2510) |
| 18 | 97.48 | 2.083 | 16 | 18 | 2.5811** | (280/2230) |
| 20 | 98.02 | 1.992 | 18 | 20 | 1.6723** | (312/1918) |
| 22 | 98.52 | 1.903 | 20 | 22 | 1.5289** | (344/1574) |
| 24 | 98.89 | 1,879 | 22 | 24 | 1.1106** | (376/1198) |
| 26 | 99.31 | 1.835 | 24 | 26 | 1.1426** | (408/790 ) |

** Test of significance of increase in fit of the $m$-degree over $k$-degree trend surface at the 99% level of significance
Polynomial Model : $k=n$, $m=n+1$ Fourier Model : $k=2n$, $m=2(n+1)$

fidence interval corresponding to a significance level of 99 percent and the variance ratio between two succeeding degree of double Fourier trend surface, for each area, diminish rapidly as increasing number of terms, reflecting increased structural complexities of land surface. However, the confidence interval on the polynomial trend surface is rather wide as compared with that on Fourier trend surface, and the spatial distribution of residuals is far from randomness. In addition, the higher-order polynomial trend is not always accompanied by a decrease in the variance ratio between two succeeding degrees.

These results suggest that the goodness of fit of the polynomial trend surface by coefficient of determination does not always reflect the spatial correspondence between the observed values and the computed values of trend surface, and that even the higher order trend surface analysis by polynomials is not appropriate for extracting the spatial structure of the extremely complex configuration of the actual surface. In fitting trend surfaces, spatial patterns of control points, for example, spacing or distribution of data points and maximum number of extrema, are critical. It is desirable that the data points be more or less evenly distributed within the mapped area. They should not be clustered in some places and spread far apart elsewhere, because clustered data points give undue influence to the areas containing them relative to areas in which the points are far apart.

Therefore, the advanced numerical experiment for evaluating the reliability and pertinence of trend surface analysis of digital terrain data have to be performed on statistical models.

## Residual from Fitted Trend Surfaces[51]

When a polynomial or double Fourier equation is fitted to the numerical terrain data, the resulting surface seldom if ever corresponds exactly to the actual observations at the sample points, because the fit is not perfect. Rather, between the surface and the observations there is a residual variation, measured by the vertical distance at each point between the elevation of the point and the elevation of the fitted surface, both with reference to the data plane. This residual variation, that is the residual is conventionally designated positive if the observation is above the fitted surface and negative if the observation is below. Residual may represent random noise or they may contain geomorphological significant information.

They represent random noise if the chosen mathematical model truly represents the observation. If the residuals contain geomorphologically significant information, as is more common, the basic underlying behavior of the dependent variable may be easier to recognize once a generalized trend is removed.

Whitten[8] has already discussed the relation between the degree of equation fitted and the pattern of residuals. Also, the authors recently studied deviation maps of the regional disparity as the spatial extent of complex configuration of the land surface in terms of a dissection model with the compound mesh map.

It was the author's impression that positive residuals generally corresponded with structural "high" or anticlines and negative residuals coincided with structural "low" or syncline. On the other hand, if the variance changes radically from place to place, in practice it was often deleterious to separate the regional and local effects with lower-order polynomial trends. In this situation, the higher-order polynomial or double Fourier trend might be applied mathematically with the advantage.

## Concluding Remarks

Mathematically computed trend surface, plane or the extremely complex surfaces that have regular trends and fitted by the least-squares criterion, may represent large-scale or "regional" structural features on the land form, whereas the residuals, the remainder found by subtracting the actual from the computed value, may represent small-scale or "local" structure. Trend-surface analysis does not reveal features that cannot be perceived in the original data with close scrutiny; however, trend-surface analysis does strongly accentuate structural features that are of less than regional magnitude, and in this way bring out details that may have gone unnoticed previously.

When choosing a trend, the question to consider is: what geomorphological model is appropriate, and how is it related to the real world as well as to a statistical model that can be implemented at the present state of statistical knowledge and of the computer technique?

The choice of a geomorphological model should be based on theoritical knowledge, on experience with similar situations, and on interval evidence in the data if they have already been collected. A simple model, such as a linear or quadratic polynomial, is to be preferred if the geomorphological model is poorly known or unknown as well as its form is linear or quadratic. A more complicated model, such as higher-order polynomial or Fourier series, may be useful for a well-defined geomorphological model that corresponds to that mathematical form, or if local variability is low and well controlled.

Here, to compare the goodness of fit of the trend function to the numerical terrain data, the two types of surface-fitting models, polynomials and double Fourier models are derived from the general linear model. Configuration of the Fourier model obtained by evaluating the double Fourier series to the original land surface reveals a remarkable similarity compared with that of the polynomial model. In spite of the extremely complex configuration of the actual surface, the Fourier surface provides such a good approximation of the actual surface, that is represents more than 99 percent of the total sum of squares of the original data. The Fourier surface, however, fails to accord with the actual surface at places along the edge of map. This is due to the fact that the map-edge values of the Fourier surface must be the same at opposite points on any two edges. The edge effect is minimized by a "buffer region".

On the contrary, the goodness of fit test of higher-order polynomial model does not always reflect the spatial correspondence between the polynomial surface and the actual surface. Choice of the two types of surface-fitting models depends partly on objective. If the objective is to isolate an extremely simple trend, the low-order polynomial are superior because they yield simple surfaces than the simplest, or first-degree Fourier surface. On the other hand, many natural features, including these land forms, are more realistically represented by Fourier model, reflecting periodicities inherent in the actual features.

In their capability of representing complexities, however, the higher-order polynomials have certain advantage over Fourier series, that is the higher degree polynomials are more capable of representing complicated surface, on a per-term basis than are double Fourier surfaces.

Finally, Fourier models are probably better when the objective is to isolate periodicities in digital terrain data, whereas polynomial models are more effective when either simple or complicated representation is desired and the isolation of underlying periodicities is not an objective.

## Acknowledgments

The authors wish to express our indebtedness to Professor K. Kanzaki of Tottori University for long-continued encouragement, ranging from discussions in the early stages of this work. We are also much indebted to Dr. Y. Fujii, Mr. T. Sakai, and Mr. S. Furutani of Kyoto University and Mr. K. Saito of Japan International Cooperation Agency for helping formulate many of the ideas that are present here, and for helping as our best stimulators over the years.

We hasten to add, however, that any errors and shortcoming of this paper are our sole responsibility and not that of our friends. We also would like to thank Dr. K. Yamamoto of Department of Geology and Mineralogy, Kyoto University, for his many important suggestions.

The work was supported by FACOM M-200 computer and the 925/1036 Calcomp plotter at Data Processing Center Kyoto University.

## References

1) Harbaugh, J. W. : Mathematical simulation of marine sedimentation with IBM 7090/7094 computers. Kansas Geological Survey Computer Contr. **1**, 1-51 (1966)

2) O'leary, M., Lippert, R. H., and Spitz, O. T. : FORTRAN IV and map program for computation and plotting of trend surfaces for degree 1 through 6. Kansas Geological Survey Computer Contr. **3**, 1-47 (1966)

3) Sampson, R. J. and Davis, J. C. : Three-dimensional response surface program in FORTRAN II for the IBM 1620 computer, Kansas Geological Survey Computer Contr. **10**, 1-19 (1967)

4) Fox, W. T. : FORTRAN IV program for vector trend analyses of directional data. Kansas Geological Survey Computer Contr. **11**, 1-35 (1967)

5) Koch, G. S. : A simulation of ghost stratigraphy. Kansas Geological Survey Computer Contr. **12**, 22-25 (1967)

6) Hobson, R. D. : FORTRAN IV programs to determine surface roughness in topography for the CDC 3400 computer. Kansas Geological Survey Computer Contr. **14**, 1-29 (1967)

7) Cole, A. J., Jordan, C., and Merriam, D. F. : FORTRAN II program for progressive linear fit of surfaces on a quadratic base using IBM 1620 computer. Kansas Geological Survey Computer Contr. **15**, 1-54 (1967)

8) Whitten, E. H. T. : FORTRAN IV CDC 6400 computer program to analyze subsurface fold geometry. Kansas Geological Survey Computer Contr. **25**, 1-45 (1968)

9) Krumbein, W. C. : FORTRAN IV computer program for simulation of transgression and regression with continuous-time Markov models. Kansas Geological survey Computer Contr. **26**, 1-37 (1968)

10) Miesch, A. T. and Connor. J. J. : Stepwise regression and nonpolynomial models in trend analysis. Kansas Geological Survey Computer Contr. **27**, 1-39 (1968)

11) Merriam, D. F. : Simposium on computer application in petroleum exploration. Kansas Geological Survey Computer Contr. 40, 1-40 (1970)
12) Sakamaki, K., Yamamoto, K., and Nishiwaki, N.: Data Analysis in Geology. Journal of the Japanese Association of Petroleum Technologists Vol. 41 No. 4, 10-18 (1976)
13) Robinson, J. E., Charlesworth, A. K., and Ellis, M. J. : Structural analysis using spatial filtering in interior plans of south-central Alberta. Bull. American Assoc. Petroleum Geologists Vol. 35, 2341-2367 (1969)
14) Hewlett, R. F. : Polynomial surface fitting using sample data from an underground copper deposit. BuMines Rept. of Inv. 6522, 1-22 (1964)
15) Link, R. F., Koch, G. S., and Gldfelter, G. W. : Computer methods of fitting surfaces to assay and other data by regression analysis. BuMines Rept. of Inv. 6508, 1-69 (1964)
16) Link, R. F., Yabe, N. N., and Koch, G. S. : A computer method of fitting surfaces to assay and other data in three dimensions by quadratic regression analysis. BuMines Rept. of Inv. 6876, 1-42 (1966)
17) Miller, R. L. : Trend surface, their application to analysis and description of environments of sedimentation. Jour. Geology Vol. 64, 425-446 (1956)
18) Krumbein, W. C. : Trend surface analysis of contour-type maps with irregular control-point spacing. Jour. Geophysical Research Vol. 64, 823-834 (1954)
19) Mandelbaum, H. : Statistical and geological implications of trend mapping with non-orthogonal polynomials. Jour. Geophysical Research Vol. 68, 505-519 (1963)
20) Whitten, E. H. T. : Sampling and trend surface analysis of granites, a reply. Geol. Soc. America Bull. Vol. 73, 415-418 (1962)
21) Merriam, D. F. and Harbaugh, J. W. : Trend-surface analysis of regional and residual components of geologic structure in Kansas. Kansas Geol. Survey Spec. Dist. Pub. No.11, 1-27 (1964)
22) Miller, R. L. and Kahn, J. S. : Statistical analysis in the geological sciences. John Wiley and Sons Inc., 390-439 (1962)
23) Preston, F. W. : Two dimensional power spectra for classification of land forms. Kansas Geological Survey Computer Contr. 8, 64-68 (1967)
24) Hirano, M. and Yokota, S. : Numerical analysis in geomorphology by means of electronic computer with reference to examples from southwest Japan. Geographical Review of Japan 49-7, 440-454 (1976)
25) Yamamoto, K. : Computer Application in Geology. Master's Thesis, Department of Geology and Mineralogy, Kyoto University (1972)
26) ―――― : The trend surface analysis by computer and its application. Jour. Geol. Vol. 79, 349-362 (1973)
27) Lee, W. H. K. and Kaula, W. H. : A spherical harmonic analysis of the earth's topography. Jour. of Geophysical Research Vol. 72 No. 2 (1967)
28) Bhattacharyya, B. K. : Two-dimensional harmonic analysis as a toll for magnetic interpolation. Geophysics 30 (5), 849-859 (1965)
29) Yamamoto. T., Saito, K., Shiba, M., and Sasaki, I. : Quantification of terrain variation in mountainus regions, based upon numerical map analysis by means of electronic computer I. Bulletin of the Kyoto University Forests 53, 184-202 (1981)
30) Shiba, M., Saito, K., and Sasaki, I. : Application of spectral conception to the terrain analysis (VIII). The principles and availability of the quantification of geomorphological variation by applying two-dimensional harmonic analysis. Trans. 92rd Meet. Jap. For. Soc., 475-480 (1981)
31) ―――――― : Application of spectral conception to the terrain analysis (IX). The quantitative significance and applicability of standard deviation of residuals as the landform variation. Trans. 92rd Meet. Jap. For. Soc., 481-486 (1981)
32) Rayner, J. N. : Correlation between surfaces by spectral methods. Kansas Geological Survey Computer Contr. 12, 31-37 (1967)
33) Esler, J. E. E. and Preston, F. W. : FORTRAN IV program for the GE 625 to compute the power spectrum of geological surfaces. Kansas Geological Survey Computer Contr. 16, 1-22 (1967)
34) Presston, F. W. and Harbaugh, J. W. : BALGOL programs and geologic application for single and double Fourier series using IBM 7090/7094 computers. Kansas Geol. Survey Spec. Dist. Pub. No. 24, 1-68 (1965)
35) Koch, G. S. and Link, R. F. : Statistical analysis of geological data. John Wiley and Sons Inc.,

2-77 (1971)
36) Krumbein, W. C. and GrayBill, F. A. : An introduction to statistical models in geology. McGaw-Hill Inc., 319-354 (1965)
37) Krumbein, W. C. : Confidence intervals on low-order polynomial trend surfaces. Jour. Geophysical Research Vol. **68**, 5869-5878 (1963)
38) ―――――― : Classification of map surfaces based on the structure of polynomial and Fourier coefficient matrices. Kansas Geological Survey Computer Contr. **7**, 12-18 (1967)
39) James, W. R. : FORTRAN IV program using double Fourier series for surface fitting of irregular spaced data. Kansas Geological Survey Computer Contr. **5**, 1-19 (1966)
40) Davis, J. C. : Statistics and data analysis in geology. John Wiley and Sons Inc., 299-411 (1973)
41) Chorley, R. J. : Spatial analysis in geomorphology. Methuen and Co. Ltd., 221-302 (1972)
42) Harbaugh, J. W. : Computer application in stratigraphic analysis. John Wiley and Sons Inc., 61-155 (1968)
43) Brigham, E. O. : The Fast Fourier Transform. Prentice-Hall Inc., 2-77 (1971)
44) Papouls, A. : The Fourier integral and its applications. McGraw-Hill Inc., (1962)
45) Shiba, M. and Sasaki, I. : Application of spectral conception to the terrain analysis (XI). Q-mode cluster analysis applied to polynomial trend-surface coefficients for numerical terrain data. Rep. 93rd Meet. Jap. For. Soc. (1982)
46) Miyake, I. and Yamamoto, K. : SPSS statistical package II. Toyokeizaishinpo Inc., 218-254 (1977)
47) Parks, J. M. : FORTRAN IV programs for Q-mode cluster analysis on distance function with printed dendrogram. Kansas Geologcal Survey Computer Contr. **46**, 1-31 (1970)
48) Wishart, D. : FORTRAN II programs for 8 methods of cluster analysis. Kansas Geological Survey Computer Contr. **38**, 1-111 (1968)
49) Bonham-Carter, G. F. : FORTRAN IV program for Q-mode cluster analysis of nonquantitative data using IBM 7090/7094 computers. Kansas Geological Survey Computer Contr. **17**, 1-27 (1967)
50) Fujitsu Co. ltd. : FACOM FORTRAN SSL II manual. 202-278 (1978)
51) Shiba, M. and Sasaki, I. : Application of spectral conception to the terrain analysis (X). Comparison of Fourier model with polynomial model on a goodness of fit basis to complicated land surfaces. Rep. 93rd Meet. Jap. For. Soc. (1982)