

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 (Ⅲ)

—AIC 最小化推定法 (MAICE) による
地形計量モデル化への可能性—

芝 正己・古谷士郎・山本俊明・佐々木 功

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

—Accessibility to the Modelling of Terrain Configuration
employing the Minimum AIC Estimation (MAICE)
procedure—

Masami SHIBA, Shiroo FURUTANI, Toshiaki YAMAMOTO and Isao SASAKI

要 旨

計量モデルが離散的な有限個のデータから構造と誤差とを分離する過程を通して構成されると考えるならば (統計的誤差処理), これには [1] データ特性に関するアプリオリな知識に基づいてモデルが的確に構成されること, [2] 想定されるモデルの中から適切なモデルを比較・評価するための客観的基準が提示されること, [3] 得られたモデルによって何らかの予知・予測が可能になること,

が不可欠な条件となる。これらのいずれかの条件が不備であれば, そのモデルの有効性は損われる。

著者等は, 数値地形データに対して多項式あるいは2重フーリエ級数をモデルとして採用し [1], 統計的仮説検定によって多項式の次数およびフーリエ級数の有限項の打切り項数を決定し [2], これらのモデルを通して, クラスター分析による小地形単位の地形形状の計量的区分法, 残差分布と相対開析量・卓越要素分布の対応関係に基づく開析量を指標とした地形特性の把握法, 林道開設における地形と土工量の問題等 [3], を論じてきた。

しかしながら, 前報で指摘した様に [2] の条件については適合度判定上の問題として課題を残しつつ, これにかわる誤差尺度の導入あるいは開発の必要性を説いてきた。

何らの統計数値表も主観的判断さえも必要としない AIC 最小化推定法 (Minimum AIC Estimation MAICE) は, この問題に対する極めて実用的で有効な接近法である。このことは, 数値実験および芦生の数値地形データ解析の結果から明らかになった。

KEYWORDS

尤度比検定 (Significant Test of Likelihood) ; 適合度検定 (Significant Test of Goodness of Fit) ; 最尤推定法 (Maximum Likelihood Estimation) ; マルコフ過程 (Markov Process) ; K-L情報量 (Kullback-Leibler Information Quantity) ; 漸近的正規性 (Asymptotic Normality) ; 赤池の情報量基準 (Akaike's Information Criterion AIC) ; AIC 最小化推定法 (Minimum AIC Estimation MAICE)

I 序 論

著者等は¹⁻²⁾, これまで地形変動の計量化の問題を「定常確率過程と見做した離散的な標高配列を多項式あるいはフーリエ級数³⁻⁴⁾で連続関数近似し, 点情報から必要とされる地形の線情報・面情報のスケールをこれらのモデルを通して推定すること」として議論してきた。

実際, 複雑な地表形態を厳密に数式化して記述・再現することが不可能であるにしても, 有限個の離散化された数値地形データ (地形図より抽出された等高線は連続量として定義された地形情報であるが, 計量化においてはこれを離散的な点情報の集合体として digital 化して処理しなければならない) から一般的利用に適する様な単純な数学的モデルとしての性格を備えた地形の現実的モデルを作ることは考える⁵⁻⁷⁾。

これは広義には, 有限個のデータから構造と誤差 (残差) を分離するいわばデータ解析における統計的誤差処理の過程として考えることができる。このモデル構成においては少なくとも次の条件が満たされる必要がある⁸⁾。

〔1〕: データの特性に関するアプリオリな知識に基づいてモデルが的確に構成されること

〔2〕: 想定されるモデル群の中から適切なモデルを選択する客観的評価基準が示されること

このモデルを通して地形的特徴による空間の分類を行なう場合

① モデルの波長や振幅などの周期性・非周期性に着目してスペクトル, 自己相関関数によって分析する調和解析 (Harmonic Analysis)^{1-2), 9-11)}

② 変動を傾向面 (変動面にできる限り近似した数学的曲面つまりは一般の分布傾向を表現するモデル) と残差に分離して, 全体的傾向や局所的分布を解析する傾向面分析 (Trend Surface Analysis)^{1-2), 10-13)}

の手法が考えられる。

これらの応用として「傾向面と分離された残差分布と, 起伏・傾斜・谷密度の地形要素を複合させた相対開析量・卓越要素分布との対応関係から, 開析量を指標として地形特性を検討する問題」¹²⁾, 「小地形単位の地形形状の類似性や差異に関する計量的区分を, クラスタ分析を併用してモデル間の係数についての距離行列として分類する手法」¹²⁾, 「傾向面からの残差分散を変動面の粗さ (Roughness) を与えるパラメータと見做して, 林道開設の作業難易の程度を残差分散と切取土工量との相互関係から比較する手法」¹³⁾, また本報では割愛したが「林道網の幾何特性を地形との相互相関関数によって定量的に評価する手法」¹⁰⁾ 等を検討してきた。

平野・横田^{9), 14)} 等が指摘する様に, 地形計量モデル構成の意義および有効性は「明瞭な地形学的問題意義 (目的意識) に裏付けられること」・「モデルによって何らかの意味での予知・予測が可能になること」にあると思われるが, その意味においてまず地形変動をパラメータとする事象を統一的に説明するためのシステムとしての普遍性が前提となり, そこから得られた解析結果

の妥当性・合理性の判断は、当然適用分野の専門的知識に委ねられるべきものであると考えられる¹⁵⁻¹⁸⁾。

前報で著者等^{20, 10)}は、多項式およびフーリエモデルを導入する場合の適合度判定が空間分布の構造的な一致度を必ずしも反映しないこと、これを評価するための何らかの誤差尺度の導入あるいは開発の必要性を強調した。実はこの問題は、モデル選択上の従来からの統計的仮説検定の問題¹⁹⁾（統計的推論を尤度だけに依存して有意性検定をモデル選択基準の拠所とする立場）と同義的なものと解釈できる。

赤池は²⁰⁻²⁴⁾尤度 (Likelihood) という概念を情報量の視点から見なおすことによって、AIC (赤池の情報量基準 Akaike's Information Criterion) と呼ばれる統一的モデル評価基準を提案しその有効性は広く理解されている²⁵⁻²⁷⁾。

本報では数値実験によりこの方法・基本的考え方を紹介し、2重フーリエ級数の有限項の部分和の打ち切り項数決定の問題に AIC を導入して計量化への接近性を探る。

II 統計的検定論^{19), 22), 28), 30)}と情報量基準 AIC²¹⁾⁻²⁶⁾

モデルの近似を高めるために高次項を stepwise して加える場合、あるいは多変量解析における変数増減・変数取捨選択により統計的に最適なモデルを構築する場合の様に、与えられたデータに対して自由パラメータ数の異なる複数のモデル群が想定される時、その中からデータを最も良く説明するモデルを選択することが重要となる³¹⁻³³⁾。

一般にデータ数に較べて用意される自由度 (自由パラメータ) が大きいと、モデルは誤差に過剰に反応し真の構造ではなく測定誤差を忠実に再現するだけになってしまい、誤差と構造とを分離して信頼性のある情報を引き出すことが不可能になってしまう。逆にパラメータが少なすぎると、モデルが対象の構造を十分に再現できなくなり得られる情報に偏りが生じるという困難がある。その中間に最適なモデルが存在するはずで、それを選び出す客観的基準が必要となる。

最小二乗法の立場からみると、残差二乗平均が小さくなる程モデルはデータの微細な構造を反映し見かけのあてはまりは良くなる。極端な場合、多項式モデルにおける Lagrange 補間式の様にデータ数と同数のパラメータ導入すればすべてのデータを完全に合わせることすらできる³⁴⁾。

従来、モデルをあてはめる場合に回帰分析や因子分析に見られる様に、その自由度決定には統計的検定法 (尤度比検定) が用いられてきたが、検定の有意水準の設定を主観的判断に委ねており複雑なモデル群が想定された場合には、妥当な有意水準を定めることが極めて困難であるという実用上の欠点を持っていた³⁵⁾。

これに対して、赤池は²⁰⁻²⁴⁾ エントロピー最大化原理 (Entropy Maximization Principle)²³⁾、³⁵⁾ という情報量概念を統計的モデル決定の領域に導入することによって、モデルの自由度とパラメータの推定を同時に行ない多くの構造の異なる統計モデルがあっても、一定の適用範囲内で最適なモデルを自動的に選択させることを可能にした³⁶⁾。モデル選択基準として提案されたこの AIC を用いると検定のための何らかの統計的数値表も主観的議論も必要としないために大量データを対象とした統計解析の自動処理化への実用性は極めて高い³⁷⁻⁴⁸⁾。

II-1 エントロピーと AIC^{23), 35)-36), 41), 45)-46)}

データを説明する統計モデルは一般に密度関数 (モデルが尤度で表現される) の族によって表現される。すなわちデータ $x_1, x_2, \dots, x_N$ は独立で同一の密度関数 $g(x; \theta)$ に従う確率変数 X_i

($i=1, 2, \dots, N$) の実現値と見做される。最尤推定法ではデータ $\{x_i\}$ に基づいて $\prod_{i=1}^N g(x_i; \theta)$

を最大にするようにパラメータ θ の値を定める。ところが現実には x_i を支配する確率が未知であり、最尤推定量は任意に選んだ関数族 $g(x; \theta)$ で定まる確率分布の中で、ある尺度で測って最も近いものを選び出す方法であると考えられる。

真の分布が密度関数 $f(x)$ を持つものとし、 $g(x; \theta)$ と $f(x)$ の近さを **Kullback-Leibler** の情報量^{34), 49)} で定義すると、エントロピーはその負値

$$B(f(\cdot); g(\cdot; \theta)) = - \int \log \left(\frac{f(x)}{g(x; \theta)} \right) f(x) dx \quad (1)$$

で与えられる。これは想定したモデル $g(x; \theta)$ から $f(x)$ が得られる確率の逆数の対数ともいうべきものであり、 $B(f(\cdot); g(\cdot; \theta))$ は $f(x) \equiv g(x; \theta)$ の時は 0 となりそれ以外では必ず正となる。

従って、このエントロピーという量はモデル $g(x; \theta)$ と真の構造 $f(x)$ との近似の度合を表わす量と考えることができるので、「エントロピー最大化」とは $-B(f(\cdot); g(\cdot; \theta))$ を最小にするモデル $g(\cdot; \theta)$ を求めることに帰着する。

赤池は^{20-23), 35)} 平均的な意味で、このエントロピーを最大にするモデルとして k 次元の最尤推定値、 $\hat{\theta}_r = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r; r=1, 2, \dots, k)$ なるパラメータを持つ確率分布 $g(\cdot; \hat{\theta}_r)$ の負エントロピーの期待値 $E[-B(f(\cdot); g(\cdot; \theta))]$ の推定値を、最小にするモデルの採用法 **MAICE** (**AIC** 最小化推定法 **Minimum AIC Estimation**) を提案している。この値が **AIC** で、赤池によれば Kullback-Leibler 情報量の 2 倍が定数項を別にして

$$2B(f(\cdot); g(\cdot; \theta)) \sim \text{AIC} = (-2) \log (\text{最大尤度 Maximum Likelihood}) \\ + 2(\text{自由パラメータ数 Number of Free Parameters}) \quad (2)$$

によって定義される量となっている。

II-2 多項式およびフーリエモデルにおける AIC

A: 多項式モデル (1 次元)^{20), 33)-34), 50)}

$$y_i = a_0 + a_1 x_i + \dots + a_m x_i^m + \varepsilon_i \quad (3)$$

ε_i は正規分布 $N(0, \sigma^2)$ に従う独立な確率変数で、このモデルは確率変数 y_i の変動を確定変数 x_i の多項式と偶然誤差の和として(3)式のように表現する。 m 次の多項式モデルは x_i が与えられた時、 y の確率分布 f が平均 $a_0 + a_1 x_i + \dots + a_m x_i^m$ 、分散 σ^2 の正規分布で与えられると仮定する条件分布モデル(4)である。

$$f(y_i | a_0, \dots, a_m, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m)^2 \right\} \quad (4)$$

n 組のデータ $\{(x_i, y_i); i=1, \dots, n\}$ が与えられた時の尤度は(5)式となる。

$$\begin{aligned} L(y_1, \dots, y_n | a_0, \dots, a_m, \sigma^2) \\ &= \prod_{i=1}^n f(y_i | a_0, \dots, a_m, \sigma^2) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m)^2 \right\} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m)^2 \right\} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \varepsilon_i^2 \right\} \end{aligned} \quad (5)$$

与えられた n 組のデータに対して、この尤度を最大にする $a_0, a_1, \dots, a_m$ および σ^2 が多項式モデル Model(m) の係数および残差分散の最尤推定量となる。この尤度関数 L の自然対数を取り対数尤度 l を(5)式より求めると

$$\begin{aligned} l(y|a_0, \dots, a_m, \sigma^2) &= \log L(y_1, \dots, y_n|a_0, \dots, a_m, \sigma^2) \\ &= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - a_0 - \dots - a_m x_i^m)^2 \end{aligned} \quad (6)$$

が得られる。最尤推定量を求めるためには、対数尤度 l を最大とする $a_0, a_1, \dots, a_m$ および σ^2 を求めればよい。この対数尤度を最大とする係数を求めるためには(7)式を最小にする $a_0, a_1, \dots, a_m$ を算定すればよいことがわかる。すなわち多項式モデルの最尤法は最小二乗法に帰着する。

$$S = \sum_{i=1}^n (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m)^2 \quad (7)$$

$a_0, a_1, \dots, a_m$ が(7)式の S を最小とするためには

$$\begin{aligned} \frac{\partial S}{\partial a_0} &= -2 \sum_{i=1}^n (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m) = 0 \\ \frac{\partial S}{\partial a_1} &= -2 \sum_{i=1}^n x_i (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m) = 0 \\ &\vdots \\ \frac{\partial S}{\partial a_m} &= -2 \sum_{i=1}^n x_i^m (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m) = 0 \end{aligned} \quad (8)$$

となり、最尤推定値 $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_m$ は(9)式の正規方程式を解いて得られる。

$$\begin{bmatrix} n & \sum x_i & \dots & \sum x_i^m \\ \sum x_i & \sum x_i^2 & \dots & \sum x_i^{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_i^m & \sum x_i^{m+1} & \dots & \sum x_i^{2m} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^m y_i \end{bmatrix} \quad (9)$$

$\hat{\sigma}^2$ が対数尤度(6)を最大にするためには

$$\frac{\partial l}{\partial \sigma^2} \Big|_{\sigma^2 = \hat{\sigma}^2} = -\frac{1}{2\hat{\sigma}^2} + \frac{1}{2(\hat{\sigma}^2)^2} \sum_{i=1}^n (y_i - \hat{a}_0 - \dots - \hat{a}_m x_i^m)^2 = 0 \quad (10)$$

であり、これより残差分散の最尤推定量は

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (y_i - \hat{a}_0 - \hat{a}_1 x_i - \dots - \hat{a}_m x_i^m)^2 \\ &= \frac{1}{n} \left\{ \sum_{i=1}^n y_i^2 - 2 \sum_{i=0}^m \hat{a}_i \sum_{j=1}^n x_j^i y_j + \sum_{i=0}^m \hat{a}_i \sum_{j=0}^m \hat{a}_j \sum_{k=1}^n x_k^{i+j} \right\} \\ &= \frac{1}{n} \left\{ \sum_{i=1}^n y_i^2 - \sum_{i=0}^m \hat{a}_i \sum_{j=1}^n x_j^i y_j \right\} \end{aligned} \quad (11)$$

となる。ここで Model(m) の残差分散 $\hat{\sigma}^2$ を $D(m)$ とすると(6), (11)式より最大対数尤度は

$$l(y|\hat{a}_0, \dots, \hat{a}_m, \hat{\sigma}^2) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log D(m) - \frac{n}{2} \quad (12)$$

Model(m) は $a_0, a_1, \dots, a_m$ および分散 σ^2 の $m+2$ 個のパラメータを持ち、(2)の AIC の定義式に(12)を代入すると

$$\begin{aligned} \text{AIC}(m) &= -2 \left\{ -\frac{n}{2} \log 2\pi - \frac{n}{2} \log D(m) - \frac{n}{2} \right\} + 2(m+2) \\ &= n \log 2\pi + n \log D(m) + n + 2(m+2) \end{aligned} \quad (13)$$

となるが, $n \log 2\pi$ および n はデータ数に関する定数であるので簡略化して

$$\text{AIC}(m) = n \log D(m) + 2(m+2) \quad (14)$$

従って AIC 計算の問題はモデルの最小二乗推定量を求めさえすればよい。

B: フーリエモデル (2次元)^{1)-3), 10)}

x, y を 2 説明変数, z を目的変数とする 2 次元のフーリエモデルを $Z = f(x, y) + \varepsilon$ と定義する。 ε は残差で平均 0, 分散 σ^2 の正規分布に従うと仮定し $f(x, y)$ を 2 重フーリエ級数として (15) 式で表わす。

$$f(x, y; \alpha, \beta, \gamma, \delta) = \sum_{m=0}^M \sum_{n=0}^N \lambda_{mn} \left(\alpha_{mn} \cos \frac{m\pi x}{L} \cos \frac{n\pi y}{H} + \beta_{mn} \cos \frac{m\pi x}{L} \sin \frac{n\pi y}{H} \right. \\ \left. + \gamma_{mn} \sin \frac{m\pi x}{L} \cos \frac{n\pi y}{H} + \delta_{mn} \sin \frac{m\pi x}{L} \sin \frac{n\pi y}{H} \right) \quad (15)$$

$\alpha, \beta, \gamma, \delta$ はモデル $f(x, y)$ の自由パラメータで $f(x, y)$ を構成するパラメトリゼーションとする。

$$\begin{aligned} \alpha &= (\alpha_{00}, \alpha_{01}, \dots, \alpha_{MN}) \\ \beta &= (\beta_{00}, \beta_{01}, \dots, \beta_{MN}) \\ \gamma &= (\gamma_{00}, \gamma_{01}, \dots, \gamma_{MN}) \\ \delta &= (\delta_{00}, \delta_{01}, \dots, \delta_{MN}) \end{aligned} \quad (16)$$

このモデルに従う $f(x_i, y_j)$ の密度関数 $F(f(x_i, y_j); \alpha, \beta, \gamma, \delta, \sigma^2)$ は

$$\begin{aligned} F(f(x_i, y_j); \alpha, \beta, \gamma, \delta, \sigma^2) \\ = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[f(x_i, y_j) - \sum_{m=0}^M \sum_{n=0}^N \lambda_{mn} \left(\alpha_{mn} \cos \frac{m\pi x_i}{L} \cos \frac{n\pi y_j}{H} \right. \right. \right. \\ \left. \left. + \beta_{mn} \cos \frac{m\pi x_i}{L} \sin \frac{n\pi y_j}{H} + \gamma_{mn} \sin \frac{m\pi x_i}{L} \cos \frac{n\pi y_j}{H} \right. \right. \\ \left. \left. + \delta_{mn} \sin \frac{m\pi x_i}{L} \sin \frac{n\pi y_j}{H} \right) \right]^2 \right\} \end{aligned} \quad (17)$$

で与えられる。従って $n_1 \times n_2$ 組のデータ $[x_i, y_j, f(x_i, y_j); i=1, \dots, n_1, j=1, \dots, n_2]$ が与えられた時, その尤度は

$$\begin{aligned} L(f(x_1, y_1), \dots, f(x_{n_1}, y_{n_2}) | \alpha, \beta, \gamma, \delta, \sigma^2) \\ = \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} F(f(x_i, y_j) | \alpha, \beta, \gamma, \delta, \sigma^2) \\ = \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[f(x_i, y_j) - \sum \sum \lambda_{mn} (\dots) \right]^2 \right\} \\ = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n_1 n_2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left[f(x_i, y_j) - \sum \sum \lambda_{mn} (\dots) \right]^2 \right\} \\ = \left(\frac{1}{2\pi\sigma^2} \right)^{n_1 n_2 / 2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \varepsilon_{ij}^2 \right\} \end{aligned} \quad (18)$$

となる。データ $[x_1, y_1, f(x_1, y_1)], \dots, [x_{n_1}, y_{n_2}, f(x_{n_1}, y_{n_2})]$ に対してこの尤度を最大にする $\alpha, \beta, \gamma, \delta$ および σ^2 がモデルの係数, 残差分散の最尤推定量を与えるので (18) 式の対数尤度は,

$$\begin{aligned} l(f(x, y) | \alpha, \beta, \gamma, \delta, \sigma^2) &= \log L(f(x_1, y_1), \dots, f(x_{n_1}, y_{n_2}) | \alpha, \beta, \gamma, \delta, \sigma^2) \\ &= -\frac{n_1 n_2}{2} \log 2\pi - \frac{n_1 n_2}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left[f(x_i, y_j) - \sum \sum \lambda_{mn} (\dots) \right]^2 \end{aligned} \quad (19)$$

この対数尤度を最大にする $\alpha, \beta, \gamma, \delta$ を推定するためには、(20)式の Q を最小にすればいいから

$$Q = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left[f(x_i, y_j) - \sum \sum \lambda_{mn}(\dots) \right]^2 \quad (20)$$

$$\frac{\partial Q}{\partial \alpha} = \frac{\partial Q}{\partial \beta} = \frac{\partial Q}{\partial \gamma} = \frac{\partial Q}{\partial \delta} = 0 \quad (21)$$

(21)式を満たすことになる。又同様に $\hat{\sigma}^2$ が(19)式を最大とするためには、

$$\frac{\partial l}{\partial \sigma^2} = \left|_{\sigma^2} \right. \hat{\sigma}^2 = -\frac{n_1 n_2}{2\sigma^2} + \frac{1}{2(\hat{\sigma}^2)^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left[f(x_i, y_j) - \sum \sum \lambda_{mn}(\dots) \right]^2 = 0 \quad (22)$$

従って残差分散の最尤度推定量は(23)式で求められる。

$$\hat{\sigma}^2 = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left[f(x_i, y_j) - \sum \sum \lambda_{mn}(\dots) \right]^2 \quad (23)$$

打ち切り項数 m, n のフーリエモデル Model(m, n) の残差分散 $\hat{\sigma}^2$ を $D(m, n)$ と表わすことにする。この時 $\sigma^2 = \hat{\sigma}^2$ として(19)式に代入すると、最大対数尤度は

$$l(f(x, y) | \hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}, \hat{\sigma}^2) = -\frac{n_1 n_2}{2} \log 2\pi - \frac{n_1 n_2}{2} \log D(m, n) - \frac{n_1 n_2}{2} \quad (24)$$

このフーリエモデル Model(m, n) に含まれる自由パラメータ数は $(2m-1) \times (2n-1) + 1$ であるから、 $k = n_1 \times n_2$ として AIC の定義式(2)に代入すると

$$\text{AIC}(m, n) = k \log 2\pi + k \log D(m, n) + k + 2[(2m-1)(2n-1) + 1] \quad (25)$$

ここで $k \log 2\pi$ および k は項数に関係しない定数なので、 $\text{AIC}(m, n)$ は(20)式の様になる。

$$\text{AIC}(m, n) = k \log D(m, n) + 2[(2m-1)(2n-1) + 1] \quad (26)$$

III 数 値 実 験

前章で定義した(14), (20)式の AIC を導入して、1次元の多項式モデルの次数およびフーリエモデルの打ち切り項数の決定に関する簡単な数値実験を示す。

A：多項式モデルの次数決定

真のモデルを3次式 $F(x) = 4.0x^3 - 2.0x^2 - 0.3x + 3.0$ とし、 x が -1.0 から 1.0 までの閉区間にあるものとし $x = -1.0, \dots, 1.0$ (20点) に対し真の値に $\sigma = 1.0$ の正規乱数を加えてデータを作成した。このデータに対して単精度で1次式から順次あてはめた。最小二乗計算および乱数は、SSL II のサブルーチンプログラム LESQ 1, RANN 2 を使用している。図-1 中の●印はデータを、実線が真の $F(x)$ 、破線が各次数のモデルを表わしている。図中のパラメータは、

$$\begin{array}{ll} \text{DEGREE: 回帰式の次数 } q & \text{AIC: 次数 } q \text{ の AIC} \\ \text{GOF: 適合度 (\%)} & \text{SUM: 残差二乗和} \end{array}$$

を示している。

結果は、図中に示した様に AIC が最小となる (MAICE モデル) のは3次のモデルで真のモデルとした $F(x)$ を的確に選択している。すなわち AIC は構造と誤差とを明確に分離していることがわかる。これに対して、高次で推定されたモデルは適合度、残差二乗和の漸近的向上は認められるにもかかわらず真の構造を充分に表現しきっておらず、むしろ誤差を含むデータの動きに引かれすぎてモデルとして明らかに不適であり「みかけのあてはまりを良くしている」にすぎないことが理解される。

正しい次数 $q=3$ に対する結果にはこれらの傾向は見られず、的確な推定値を得るために次数の決定がいかに重要であることを示している。

EXAMPLE 1

†† The design procedure and test run for the function check of the program for the one-dimensional *polynomial trend*

Conditions

Polynomial function; $F(x)$

$$F(x) = 4.0x^3 - 2.0x^2 - 0.3x + 3.0$$

Data; D_i

$$D_i = F(x_i) + \varepsilon_i \quad i = 1, 2, \dots, 21$$

$[x_i: -1, 1]$

ε_i ; normal random number (0,1.0)

Outputs

DEGREE : q

AIC : AIC ($[D_i] : q$)

GOF : Goodness of fit (%)

SUM : Sum of square due to deviation

Subroutines (SSL II)

CALL RANN 2, CALL LESQ 1

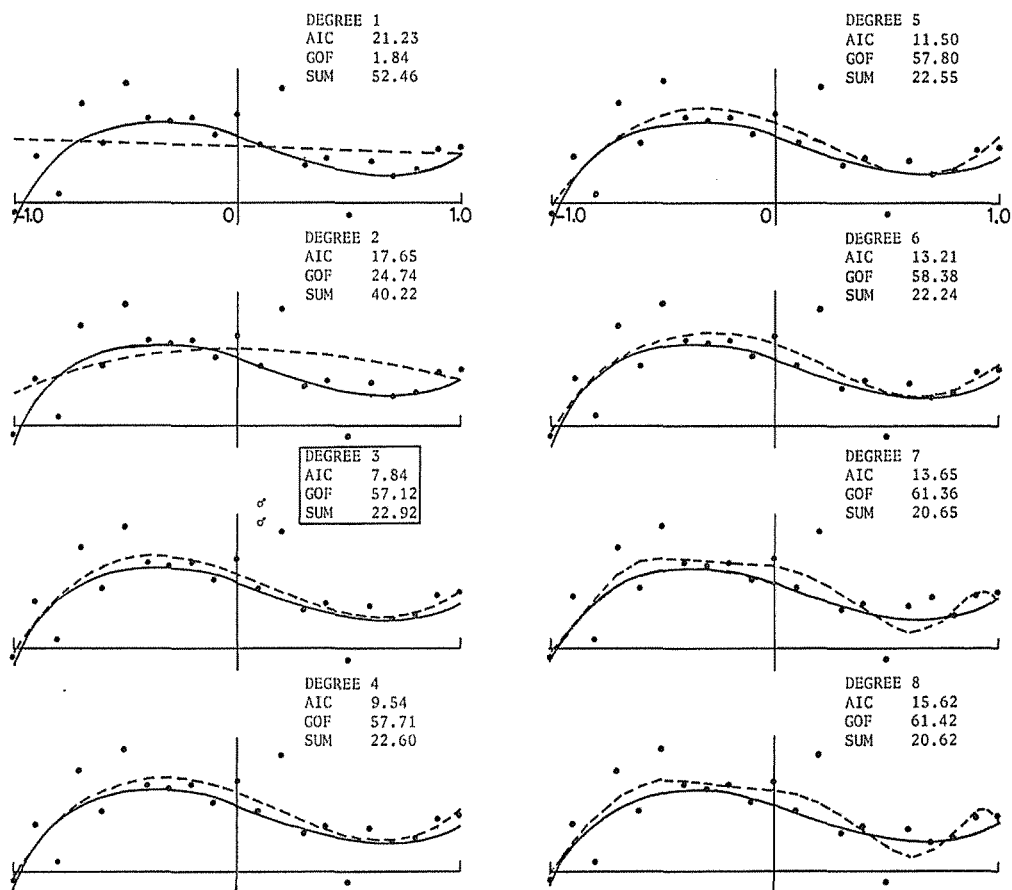


Fig. 1. The design procedure and test run for the function check of the program for the one-dimensional polynomial model: theoretical model (solid line), approximate model (broken line), and data (dots).

EXAMPLE 2

†† The design procedure and test run for the function check of the program for the one-dimensional *Fourier trend*

Conditions

.....

Fourier series; $F(t)$

$$F(t) = 0.5 + 0.4 \cos \pi t + 0.4 \sin \pi t - 0.3 \cos 2\pi t - 0.3 \sin 2\pi t + 0.5 \cos 3\pi t + 0.5 \sin 3\pi t - 0.1 \cos 4\pi t - 0.1 \sin 4\pi t - 0.2 \cos 5\pi t - 0.2 \sin 5\pi t$$

Data; D_t

$$D_t = F(t_i) + \varepsilon_i \quad i = 1, 2, \dots, 32$$

[t_i : -1, 0.9375]

ε_i ; normal random number (0, 1.0)

Outputs

.....

DEGREE : q

AIC : AIC ($[D_t] : q$)

GOF : Goodness of fit (%)

SUM : Sum of square due to deviation

Subroutines (SSL II)

.....

CALL RANN 2, CALL RFT

EXAMPLE 2-1

FITTED MODEL "FOURIER MODEL"

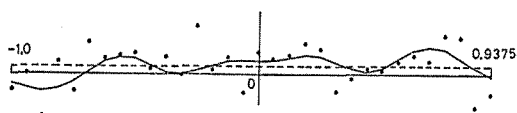
Conditions

.....

Same the EXAMPLE 2

ε_i ; normal random number (0, 0.1)

Kth-order	Sum of square	Goodness of fit	AIC
1	17.79	0.00	-14.78
2	13.06	26.61	-20.69
3	10.00	43.80	-25.23
4	1.87	89.48	-74.84
5	1.46	91.78	-78.72
6	0.32	98.23	-123.82 ⁺⁺
7	0.28	98.42	-123.45
8	0.26	98.52	-121.56
9	0.22	98.75	-122.91
10	0.22	98.76	-119.32
11	0.18	98.98	-121.54
12	0.16	99.11	-121.81
13	0.16	99.11	-117.90
14	0.16	99.12	-114.21
15	0.11	99.38	-121.53
16	0.00	100.00	-339.53

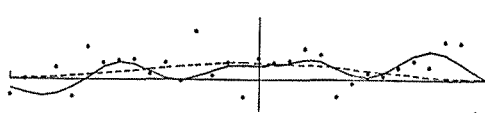


DEGREE 1 (ARITHMETIC MEAN)

AIC 22.39

GOF 0.00

SUM 56.85



DEGREE 2

AIC 23.83

GOF 7.70

SUM 52.48

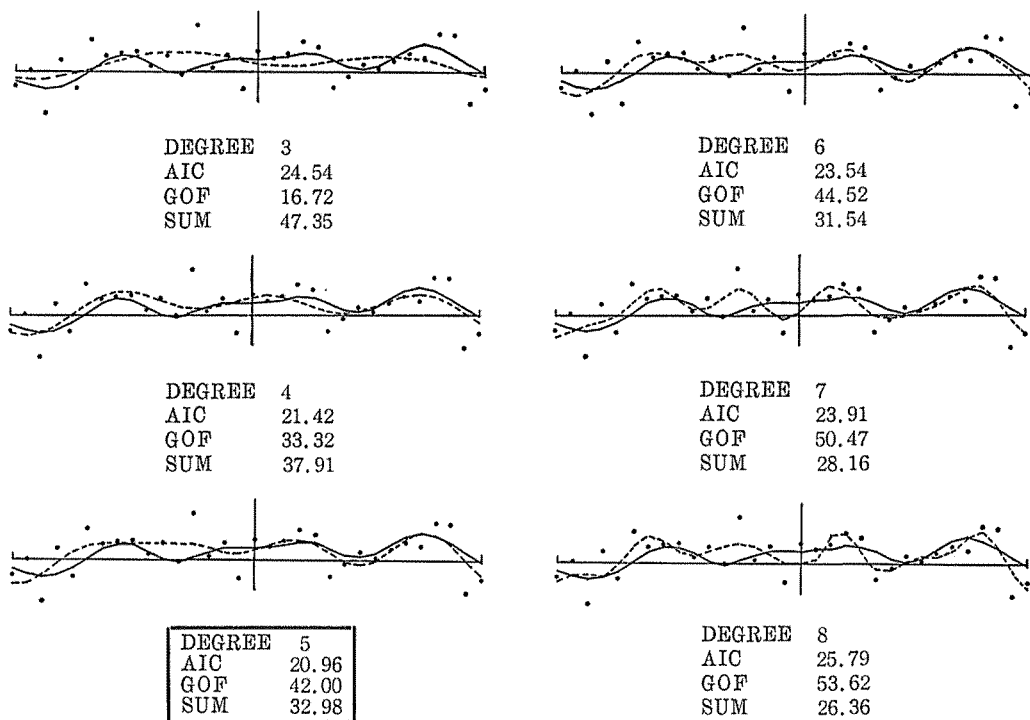


Fig. 2. The design procedure and test run for the function check of the program for the one-dimensional Fourier model: theoretical model (solid line), approximate model (broken line), and data (dots).

B: フーリエモデルの打ち切り項数の決定⁵¹⁾

Aと同様に真のモデル $F(t)$ を1個の平均成分と5個の周波数成分から合成されているものとして (B-1) 式で定義する

$$F(t) = 0.5 + 0.4 \cos \pi t + 0.4 \sin \pi t - 0.3 \cos 2\pi t - 0.3 \sin 2\pi t + 0.5 \cos 3\pi t + 0.5 \sin 3\pi t - 0.1 \cos 4\pi t - 0.1 \sin 4\pi t - 0.2 \cos 5\pi t - 0.2 \sin 5\pi t \quad (\text{B-1})$$

$t = -1.0, \dots, 0.9375$ (32点) に対し, 真の値 $F(t)$ に $\sigma = 1.0$ の正規乱数を加えてデータとした (SSL II: RANN 2)。フーリエ変換・逆変換は FFT (SSL II: RFT)⁵²⁾⁻⁵³⁾ を使用し, 単精度で1項から全項までの AIC を算定した。図-2 はその結果を示している。

MAICE モデルは5項までのフーリエ係数を合成したモデルとなっており, 真のモデルとした6項の $F(t)$ を選択していない。そこで実線で示した真のモデルに対する **MAICE** モデルと6次のモデルを比較すると, 6次のモデルは明らかに $F(t)$ を表現しきっておらず誤差の影響を過度に受けている。**MAICE** モデルあるいは4次のモデルは, 打ち切り項数が低く押えられているにもかかわらず $F(t)$ に十分に近似している。

このデータは $F(t)$ に対して $\sigma = 1.0$ の乱数を加えたが, 実は真のモデルが誤差に隠れてしまう様に意図的にフーリエ係数を小さくしておいた結果であり, 誤差の多い精度の低いデータに対しては AIC の意味での最適モデルと真のモデルとが必ずしも一致しないことを示している。しかしながら前述した様に, モデル選択基準としての AIC は単に統計学的に最適なモデルを選択するのではなく, その存在を想定している真の関係を最も適切に表現するモデルを見出すことにあり, その意味において, **MAICE** モデルとして5次のモデルが選択された結果は何らの理論

Table 1. Results of run test for the function check of the program applied to two models polynomial and Fourier fitted to theoretical function.

FITTED MODEL "POLYNOMIAL MODEL"

Kth-order	Sum of square ¹	Goodness of fit ¹¹	AIC
1	52.46	1.84	21.23
2	40.22	24.74	17.65
3	22.92	57.12	7.84 ⁺⁺
4	22.60	57.71	9.54
5	22.55	57.80	11.50
6	22.24	58.38	13.21
7	20.65	61.36	13.65
8	20.62	61.42	15.62
9	17.43	67.39	14.08
10	16.92	68.33	15.47
11	16.91	68.35	17.46
12	13.95	73.91	15.40
13	13.55	74.65	16.80
14	13.23	75.25	18.30
15	—	—	—
16	—	—	—
17	—	—	—
18	—	—	—
19	—	—	—
20	—	—	—

FITTED MODEL "FOURIER MODEL"

Kth-order	Sum of square	Goodness of fit	AIC
1	56.85	0.00	22.39
2	52.48	7.70	23.83
3	47.35	16.72	24.54
4	37.91	33.32	21.42
5	32.98	42.00	20.96 ⁺⁺
6	31.54	44.52	23.54
7	28.16	50.47	23.91
8	26.36	53.64	25.79
9	22.30	60.78	24.44
10	22.02	61.27	28.04
11	18.12	68.12	25.81
12	15.86	72.11	25.54
13	15.81	72.19	29.44
14	15.66	72.45	33.14
15	10.99	80.67	25.81

! : Sums of squares due to deviation
 !! : Percentage (%)
 ++ : Minimum AIC model (theoretic model)

AIC (Akaike's Information Criterion) and FPE (Final Prediction Error)

AIC is defined as follows:

$$\text{AIC} = (-2) \log (\text{maximum likelihood}) + 2(\text{number of free parameter}) \quad (1)$$

AIC measures both the fit of a model and the unreliability of a model; the former is measured by the first term of the definition, which decreases as the number of parameters within the model is increased, and the latter by the second term, which increases with the number of parameter. Akaike (1973a) introduced the **MAICE** (minimum AIC estimation) procedure which selects the model whose structure with its associated parameter value gives the minimum of AIC. Before the introduction of AIC, for the fitting of autoregressive models, Akaike (1970a) introduced a statistic called **Final Prediction Error (FPE)** in the general form of equation (2) as follows, which is an estimate of the one step ahead prediction error variance of an autoregressive model of order p , where $\hat{\sigma}^2(p)$ is the estimated residual variance of the model and N is the number of observations. From the equation (2), noted that (3).

$$\text{FPE}(p) = \hat{\sigma}^2(p) \left[1 + \frac{p+1}{N} \right] \left[1 - \frac{p+1}{N} \right]^{-1} \quad (2)$$

$$\log \text{FPE}(p) = \log \hat{\sigma}^2(p) + [2(p+1)/N] + \sigma(N^{-2}) \quad (3)$$

Since (-2) times the log-likelihood of a Gaussian autoregressive model of order p is approximately by $N \log \hat{\sigma}^2(p) + \text{constant}$ and the parameter involved are the p autoregressive coefficients, the residual variance and the mean of the series, we obtain the equation (4).

$$N \log \text{FPE}(p) = \text{AIC} + \text{constant} + \sigma(N^{-1}) \quad (4)$$

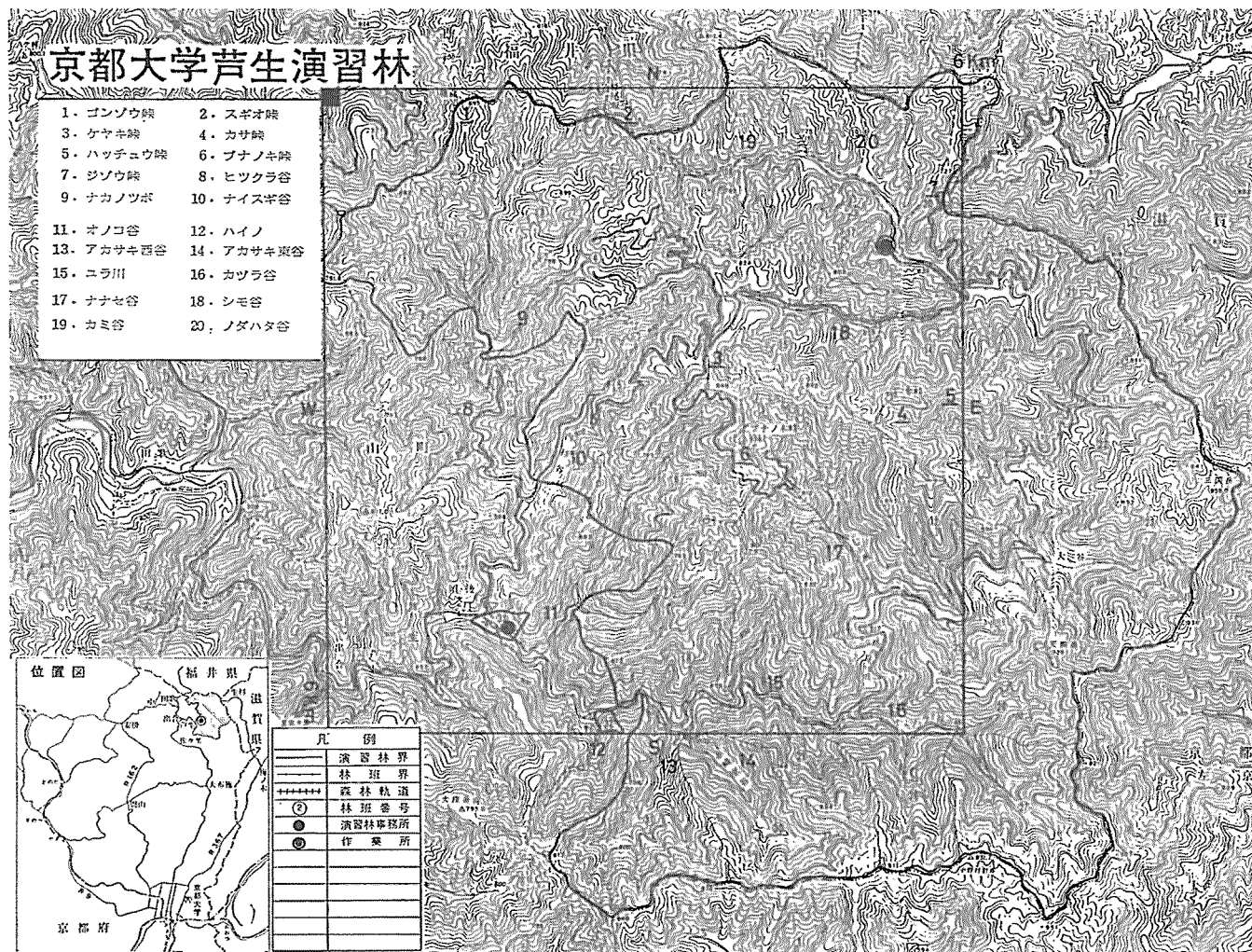
Thus the **MAICE** procedure for the autoregressive model fitting is asymptotically equivalent to the minimum **FPE** procedure.

的矛盾も与えない。

図-2の右上の表の **EXAMPLE 2-1** は $F(t)$ に $\sigma=0.1$ の正規乱数を加えてデータとし **AIC** を算定した結果である。**MAICE** モデルは6次のモデルとなり、真のモデル $F(t)$ の項数とまったく一致する。すなわち測定誤差が小さくなるとよりの確に選択される。

表-1に数値実験A, Bにおける残差二乗和(SUM) および寄与率(GOF) と **AIC** の関係を示す。

A, Bの数値実験より、モデルの次数が増加すると真の分布からの偏差は小さくなるものの推定値のばらつきは大きくなり、従って「偏差を小さくすること」と「ばらつきを小さくすること」という2つの背反的要求を適切に妥協させることが、良好な推定値を得るための point になることが理解される。**AIC** はこの意味において、何らの主観的判断も必要としない^{20)-22), 24)}。



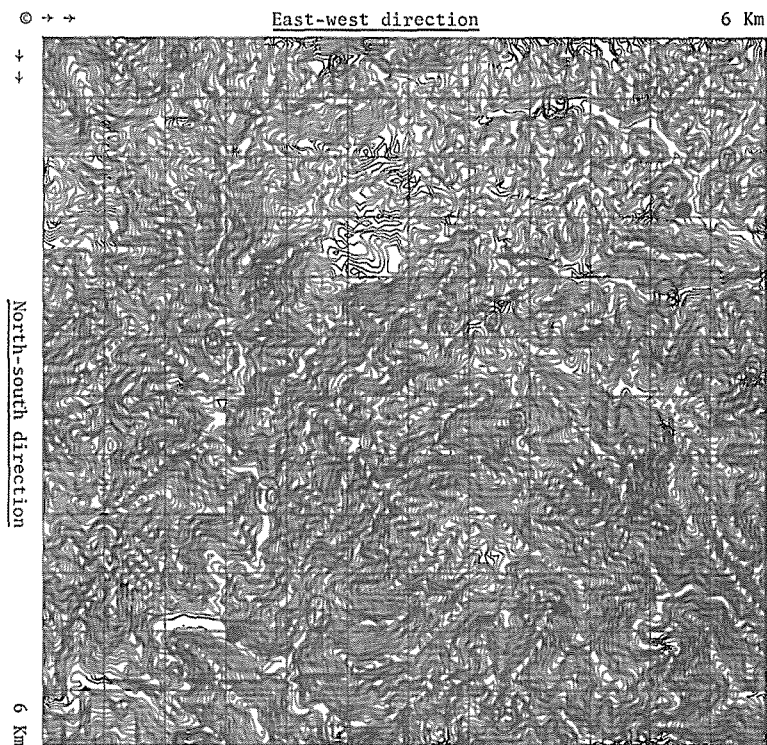


Fig. 3. An effective comparison of the base map (1/40,000 topographical map with 20 meters contour interval) and subsequent contour map produced by an automatic contouring program that uses X-Y plotter (contour interval is 10 meters). This area locates north-western part of University Forest in Ashiu and is 6×6 kilometers extent. Data consist of topographic elevations observed at points on a grid containing 120 rows and 120 columns (120×120 altitude matrix based on a square grid interval of 50 meters) and a total of 14,400 data points are used. It shows how little information has been lost by taking approximate altitude at discrete points.

図-3は当該地域を含む1/40,000地形図と、解析に用いたベースマップ(1/10,000地形図)から作成された数値地形データをX-Yプロッタにより自動図化して再現したコンターマップを比較したものである。

C: フーリエモデルの構成とAIC算定のアルゴリズム^{11), 53)}

フーリエモデル構成・AIC算定の general flowchart (図-4) に示される様に、まず x , y , z のデータが読み込まれ原データに関する図化処理が実行され、次に多次元離散型複素フーリエ変換 (SSL II: CFT) によりフーリエ係数、パワースペクトラムが算定される。打ち切り項 (m, n) に対応するフーリエ係数を用いてフーリエ逆換を施し推定値を算定し、 z とこの推定値より統計量が印刷され引き続き推定値に対する図化処理が行われる。次に②式で定義した AIC を算定しこのモデルの解析を終り、打ち切り項のすべての組み合わせについて順次このループを繰り返し実行を終了する。

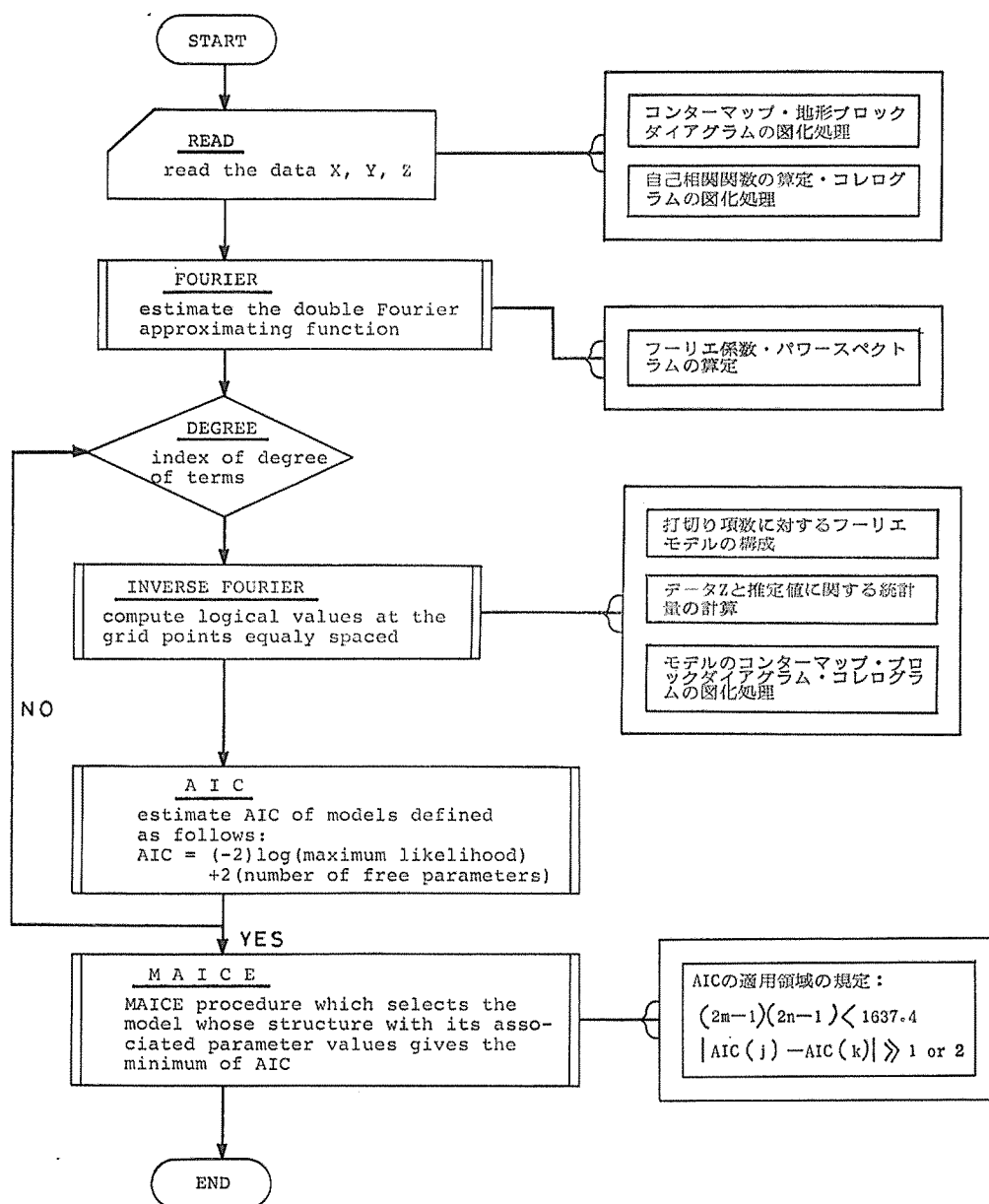


Fig. 4. A general flowchart for double Fourier series evaluation and the statistical model identification based on the minimum AIC procedure which selects a Fourier model whose AIC is the minimum. Fourier models various orders are fitted to one set of data and the corresponding values of $AIC(m,n)$ are computed, where $AIC(m,n)$ is the value of AIC for the model with m order to x -direction and n order to y -direction.

D: MAICE モデルの決定と適合性の検証

MAICE モデル決定において採用される $AIC(m, n)$ の適用範囲は、田辺³³⁾・坂元⁵⁴⁾等の指摘にある次の2制約条件を満足することとした。

[1]: MAICE は最尤推定量の漸近正規性理論 (*Asymptotic Normality Principle*)²⁸⁾に基づいており、モデルの自由パラメータ k はデータ数 N に対して、 $\frac{k+1}{N} < \frac{1}{10}$ とする。

残差分散を含むフーリエモデルの打ち切り項 (m, n) に対する自由パラメータ数は $(2m-1) \times (2n-1) + 1$ で与えられ、データ数 $N = k \times l$ に対してこの不等式は

$$\frac{(2m-1)(2n-1)}{16384} < \frac{1}{10} \quad (29-1)$$

となり、この不等式を満たす自然数 (m, n) の中で AIC を最小にするものが MAICE モデルとなる。

[2]: $AIC(m, n)$ により推定できるのはモデルの期待平均対数尤度であって *Kullback-Leibler* 情報量ではない。すなわち AIC とエントロピーの関係から

$$|AIC(m, n) - AIC(m1, n1)| \geq 1 \text{ or } 2 \quad (29-2)$$

ならば有意と考える。

この2条件を満足する Model (m, n) の統計量および $AIC(m, n)$ の値をまとめたのが表-2で、この結果より東西方向に29項、南北方向に14項を採用した Model $(29, 14)$ の AIC が最小で、同時に最も高い寄与率を示し全変動の95%がこのモデルによって説明されている。自由パラメータの最も多い Model $(16, 26)$ のあてはまりは2番目に良好であるが残差分散はむしろ増加している。Model $(45, 9)$ および Model $(60, 7)$ はほぼ同量の自由パラメータを含んでいるにもかかわらず、モデルの安定性に欠けることが理解される。

Table 2. Results of the MAICE procedure applied to $m \times n$ Fourier models fitted to numerical terrain data. This results shows that the model $(29, 14)$ will be adopted if we apply the MAICE procedure to the set of possible Fourier models.

Model	F.P.	T.S.S.	S.S.R.	S.S.D.	G.O.F.	AIC
(16, 26)	1581	0.93485 E09	0.88372 E09	0.51126 E08	94.53	293975
(29, 14)	1539	0.93485 E09	0.88797 E09	0.46875 E08	94.99	292469 (MAICE)
(45, 9)	1513	0.93485 E09	0.86057 E09	0.74273 E08	92.06	299958
(60, 7)	1547	0.93485 E09	0.82630 E09	0.10854 E09	88.39	306242

F.P.: Free Parameter T.S.S.: Total Sums of Squares

S.S.R.: Sums of Squares due to Regression

S.S.D.: Sums of Squares due to Deviation

G.O.F.: Goodness of Fit (%)

Set of possible Fourier models by asymptotic normality principle:

$$(2m-1) \times (2n-1) < 1637.4$$

図-5は(27)~(29)式に基づいて計算されたパワースペクトラムのスケールルート S_{mn} (単位はメートルで各波長に対する振幅を表わす) と Model (m, n) の打ち切り項数の関係を示している。

図に示す様に、高次項の値は無視できる程小さくモデルの情報としてほとんど寄与していない。比較的少数の低次の打ち切り項数によって安定したモデルを決定するいわばケチの原理 (*Prn-*

Direction containing terms whose degree is indexed with n and points indexed with j, l being maximum index of points, there being l+1 grid lines in Y-direction

being maximum index of points, there being k+1 grid lines in X-direction

MAICE model

	1	.	.	.	5	.	.	.	.	10	.	.	.	.	15	.	20					.	.	.	.	25	.	.	.	.	30		45		60	...		
1+	559	57	27	33	41	34	30	26	22	18	16	15	11	9	4	3	0	2	4	6	6	8	8	8	8	7	7	6	4	3	-----	3	-----	3	-----	6	-----	
•	71	9	14	6	12	6	12	2	2	4	4	2	3	2	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0	-----	0	-----	0	-----	0	-----	
•	42	11	9	7	4	3	2	3	2	3	2	0	1	1	1	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	30	6	0	5	1	7	1	2	2	1	1	2	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
5+	27	5	4	2	5	3	4	3	2	3	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	30	4	3	4	2	3	2	5	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	24	2	5	1	3	2	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	24	3	3	1	2	1	3	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	21	1	1	2	2	3	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
10+	17	2	2	1	1	2	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	17	2	1	1	1	2	1	0	0	2	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----
•	12	2	1	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----
•	10	0	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----
•	6	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
15+	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
20+	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
25+	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
30+	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-----	0	-----	0	-----	0	-----	
•																																						
•																																						

Fig. 5. Matrix of discrete power-spectrum square-root values (*harmonic vector magnitude*) obtained according to fundamental waveform containing for four harmonics fitted to original numerical terrain data, where the corresponding harmonic vector magnitude value $S(m,n)$ is the value for the model with m order to x -direction and n order to y -direction. Numbers printed on the grid are counted fractions of 0.5 and over as a unit and cutted away the rest. Array could be extended to include orders of successively higher degree, but these results omitted at some point.

Original terrain block diagram

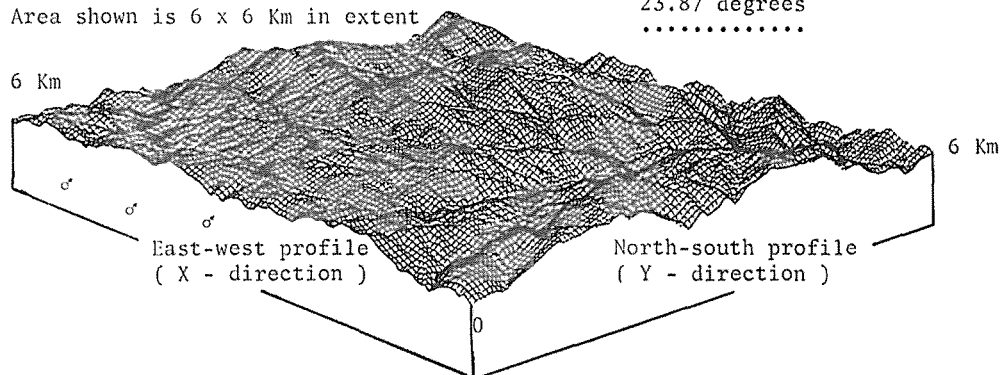
Area shown is 6 x 6 Km in extent

Angle of depression

23.87 degrees

.....

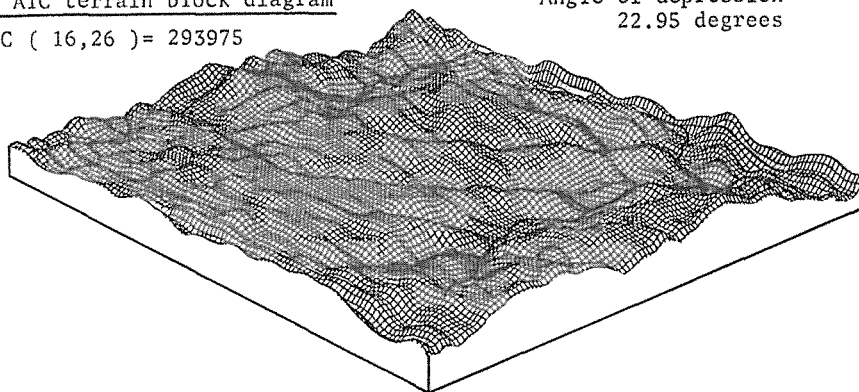
6 Km

16 AIC terrain block diagram

AIC (16,26) = 293975

Angle of depression

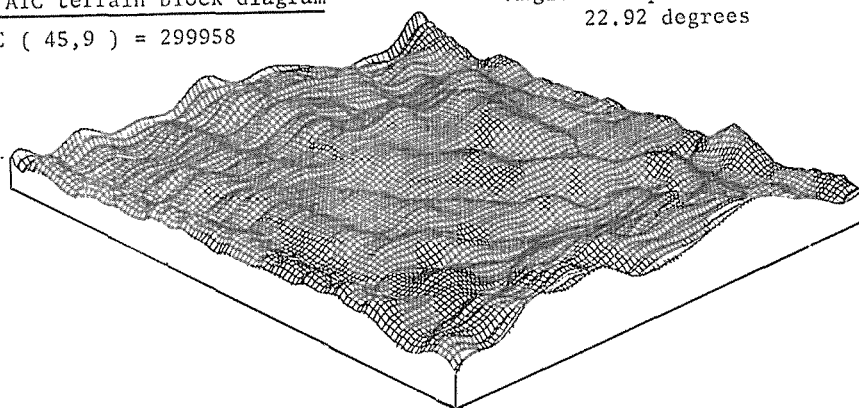
22.95 degrees

45 AIC terrain block diagram

AIC (45,9) = 299958

Angle of depression

22.92 degrees



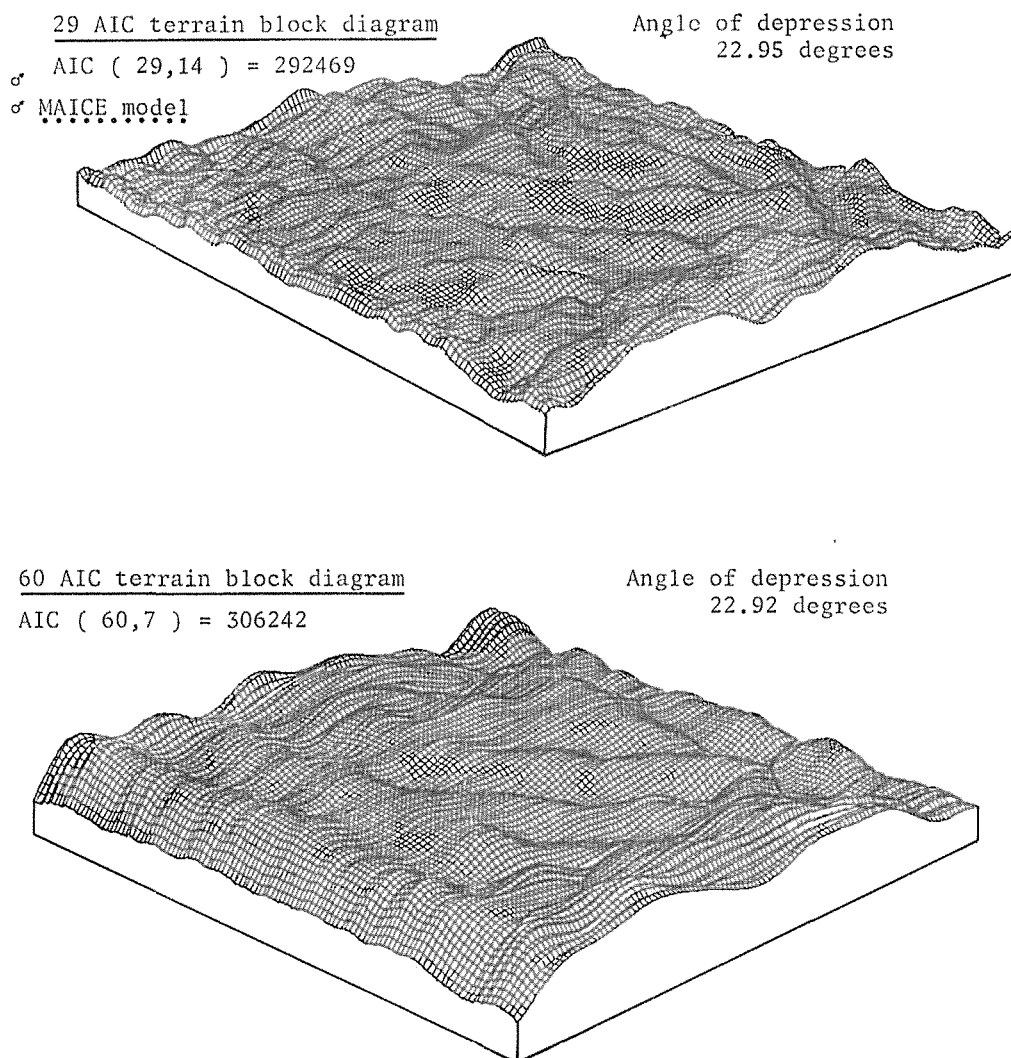
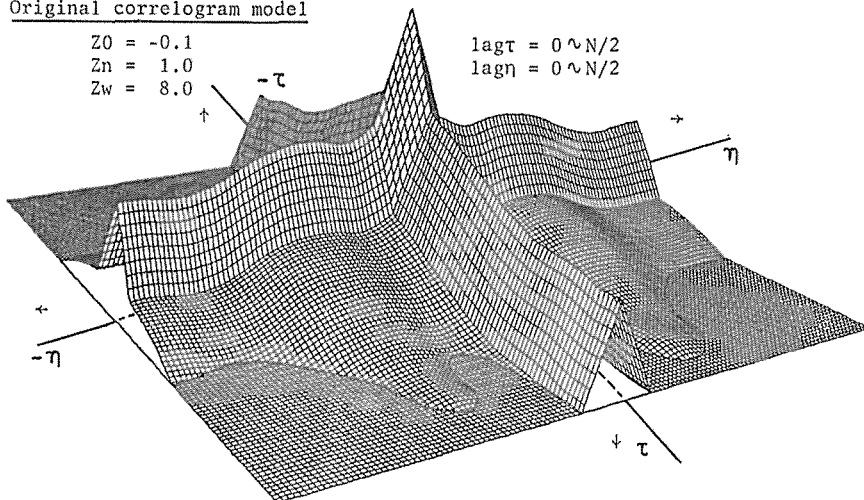


Fig. 6. Comparison of smoothed terrain block diagrams by Fourier models and original terrain block diagram(*upper*), view from the north-west. Profiles drawn north-south and east-west in parallel perspective. The **MAICE** procedure adopts the **model(29, 14)** as the best among the set of Fourier model(m,n). It looks as though both the **model(29,14)** and the **model(16,26)** would clearly pass the test. There is, of course, some loss in detail in models compared with the actual surface but is no great difference among them in its configuration(*typical ridge and valley complex*). However, the **model(45,9)** and the **model(60,7)** has a remarkable periodicity toward the north-south direction.

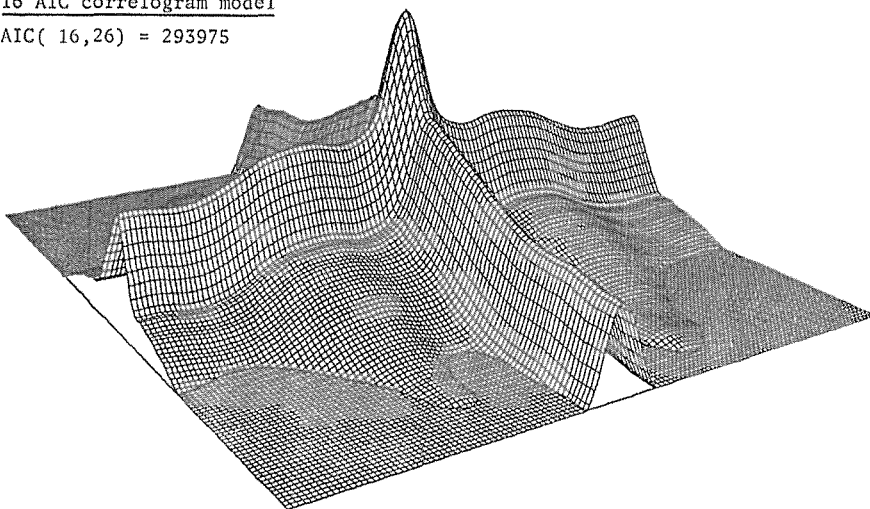
Original correlogram model

$Z_0 = -0.1$
 $Z_n = 1.0$
 $Z_w = 8.0$

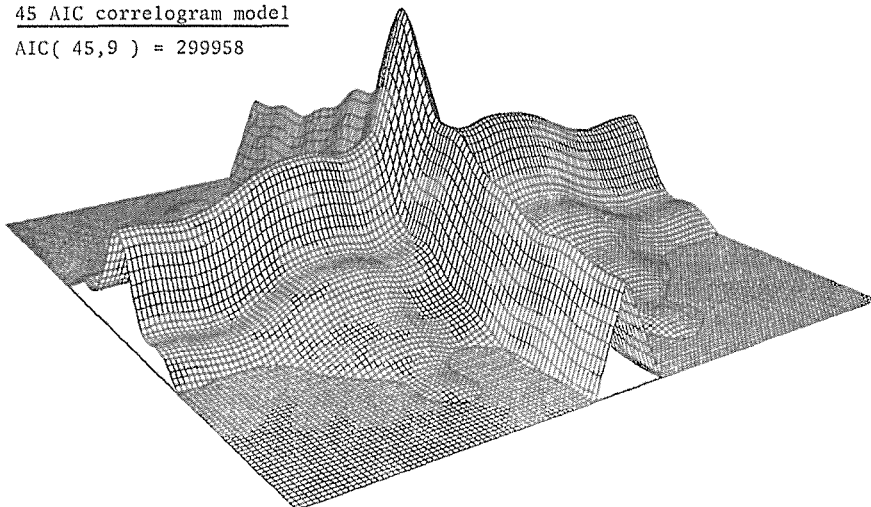
$\text{lag}\tau = 0 \sim N/2$
 $\text{lag}\eta = 0 \sim N/2$

16 AIC correlogram model

$\text{AIC}(16, 26) = 293975$

45 AIC correlogram model

$\text{AIC}(45, 9) = 299958$

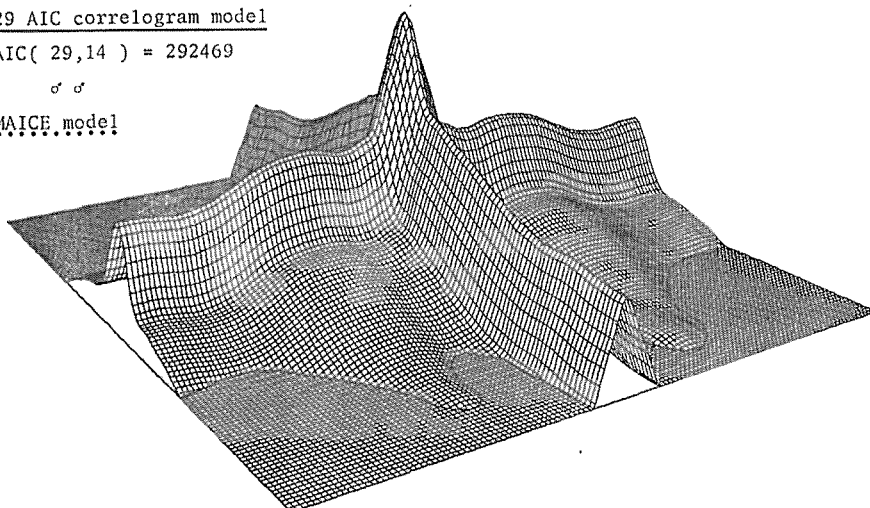


29 AIC correlogram model

AIC(29,14) = 292469

σ σ

MAICE.model



60 AIC correlogram model

AIC(60,7) = 306242

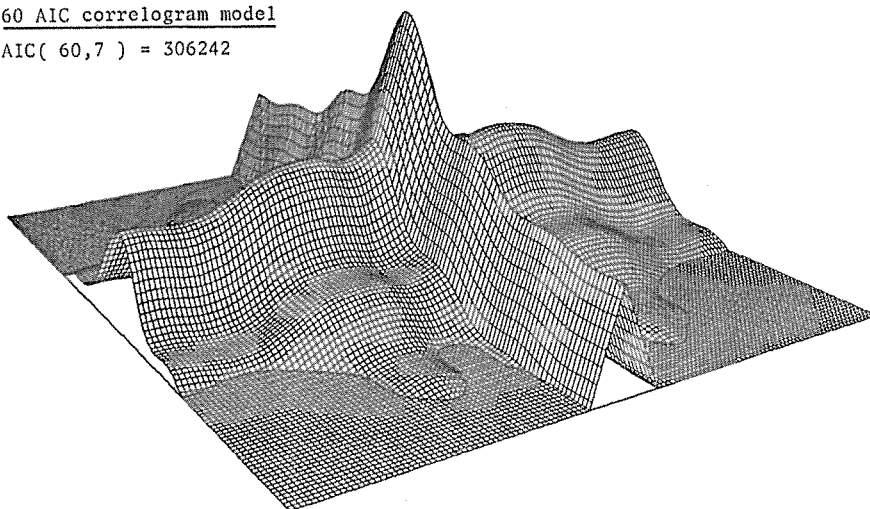


Fig. 7. Comparison of the behaviour of autocorrelation applied to the set of possible Fourier models and the original surface. The rather slow damping of the correlogram at some lag indicates that original sequence is fairly periodic with superimposed linear trend in y dimension of map and a typical pattern of non-stationary series in x dimension, where the observations increase steadily in value along the sequence. The correlogram shows steadily decreasing correlation. Estimated autocorrelation of each Fourier model, the behaviour of which provides us an important clue in the identification stage of the MAICE procedure. By the MAICE procedure, the model(29,14) is adopted as the best if we consider the set of possible Fourier models.

ciple of Parsimony)^{20)~22)} の妥当性がうかがわれる。東西方向の高次項を含む Model(45, 9), Model(60, 7) に比較して MAICE モデルとして選択された Model(29, 14) あるいは Model(16, 26) の方が良好なモデルであるという AIC の評価を裏付けている。

図-6 はモデルの推定値を用いて、地形ブロックダイアグラムによって原データと各モデルとを 3 次元的に比較したものである。

Model(45, 9), Model(60, 7) は東西方向の高次項の採用にもかかわらず、南北方向の採用項数が少ないために極だった周期性のみが強調されて原データの地形構造の再現性は充分であるとはいえない。これに対して、Model(29, 14), Model(16, 24) ではモデルの曲面が原面に対して十分に近似されており地形変動を極めて正確に捉えていることがわかる。

ちなみにブナノキ峠、傘峠、ケヤキ峠の高峰、中山盆地、七瀬、小野子、内杉、櫃倉、由良川の水系等の基本地形が良く再現されている。

この結果をさらに明確に定量化して評価するために、2 次元の自己相関関数を導入して原データとモデルの地形変動特性を自己相関関数を正規化したコレグラムで比較した^{1), 10), 52)}。

原データおよび各モデルに対する 2 次元の自己相関関数および係数の数値計算において、 $x-y$ 平面上の標高データの不規則変動を $f(x, y)$ とする時、 τ, η 隔たった 2 つの変動の積の平均値として自己相関関数 $R(\tau, \eta)$ を(30)式で定義する。

$$R(\tau, \eta) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} f(x_i, y_j) f(x_i + \tau, y_j + \eta) \quad (30)$$

τ, η はそれぞれ x, y 方向の distance lag で、 $f(x, y)$ の標本点間隔と等しい $\Delta\tau, \Delta\eta$ おきの点で与えられ座標系の標本数 N_1, N_2 に対して lag を $\frac{1}{2}$ まで採用する。

$$\begin{aligned} \tau &= 0, \Delta\tau, \dots, (N_1/2-1)\Delta\tau \\ \eta &= 0, \Delta\eta, \dots, (N_2/2-1)\Delta\eta \end{aligned} \quad (30-1)$$

$\tau = \eta = 0$ の時、(30)式の自己相関関数 $R(0, 0)$ は二乗平均 (平均パワー) を示し(30)式を(31)式で正規化すると自己相関関数 $\phi(\tau, \eta)$ が求められる。

$$R(0, 0) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} f^2(x_i, y_j) \quad (31)$$

$$\phi(\tau, \eta) = \frac{R(\tau, \eta)}{R(0, 0)} = \frac{\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} f(x_i, y_j) f(x_i + \tau, y_j + \eta)}{\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} f^2(x_i, y_j)} \quad (32)$$

(32)式に基づいて算定した $\phi(\tau, \eta)$ のコレグラムが図-7 で、図中の座標系の東西方向の lag が τ 、南北方向が η に対応している。原データのコレグラムの左上の各々のパラメータは、

Z_0 : コレグラムの基底面の関数値で、自己相関係数 $\phi(\tau, \eta) = -0.1$

Z_n : ピークの関数値 $\tau = \eta = 0$ の自己相関係数 $\phi(0, 0) = 1.0$

$Z_w: \overline{Z_n - Z_0}$ のスケールで、図形座標で 8cm に表現してある。

原データのコレログラムは、座標軸に対して異方的な地形変動が存在していることを示している。南北方向あるいは北西-南東方向を軸としてコレログラムは周期性が現われているが、これは対象地域における中山盆地-由良川水系、櫃倉-内杉-小野子谷水系の関係として理解できる。AIC モデルにおける南北方向 (y 軸) の打ち切り項数が Model(16, 26) をのぞいていずとも東西方向 (x 軸) に較べて少なく ($n=7\sim 14$) 採用されているのは、コレログラムに見られる周期的な地形の起伏特性を反映した結果であると考えられる。

これに対して、東西方向の変動はランダム性が強くコレログラムは単純マルコフ型を示している。AIC モデルにおいて東西方向に採用項数が多くなっているのはこの原因によるものである。またこの解析地域に存在している線形トレンドを最小二乗法により抽出した結果、

$$z = 657.03 + 1.49x - 1.28y$$

として求められ、地域全体が北東—南西方向にわずかに傾斜しておりこれも大きな地形特性となっていることが理解される。

以上の結果から判断しても、AIC モデルのコレログラム(AIC 決定による採用項数)は原データのコレログラム(全項を採用項とする)と大差が認められず、MAICE モデルとして選択された Model(29, 14) のコレログラムは原データのそれに最も良く近づけられたものであることが確認される。

V 結 語

多項式モデルの次数の決定、フーリエモデルの有限項の打ち切り項数の決定に MAICE が役立つことを数値実験で示し、数値地形データ解析の結果は MAICE モデルが「点情報から必要とされる地形の線情報・面情報のスケールをモデルを通して推定する」という意味で、計量モデルとして充分に有効であることを示唆したものであるといえる。これらの結果を要約すると次のようになる。

① 離散データの連続関数近似あるいは多重回帰分析・因子分析によるデータ構造決定における主要な問題は、データ特性を表現する統計モデルを分析目的に応じていかに的確に構成するか、また想定されるモデルの良さをいかなる基準に従って客観的に評価・比較するかである。

従来の統計的仮説検定(尤度比検定)法は、検定の有意水準の設定を主観的判断に委ねており複雑なモデル群が想定された場合には妥当な有意水準を定めることが困難であるという実用上の重大な欠点を持っている。

赤池はエントロピー最大化という情報量の概念を統計的モデル決定の領域に導入することによって、モデルの決定とパラメータの推定を同時に行ない異なるモデルも尤度に基づく客観的規準によって比較可能であることを証明している。この規準が AIC で、モデルを統一的に評価・比較する過程で何らの統計数値表も仮説検定上の主観的判断も必要としない。

② 数値実験から明らかな様に、必要以上のモデルの次数増加は仮定した真の分布からの偏差を小さくしているにもかかわらず、推定値のばらつきを大きくし誤差を含むデータの動きに引かれすぎてみかけのあてはまりの良さを示しているにすぎない。従って、的確な推定値を得るためには、偏差を小さくすることとばらつきを小さくするという2つの背反的な要求を満たす必要があり、MAICE はその判定の指針を与える。

③ フーリエ級数を導入してモデルを構成する実用上の問題点はフーリエ級数が各点収束ではあるが一様収束ではないことであり、このために級数を有限項で打切ると両端あるいは不連続点の付近に Gibbs 現象⁵²⁾を生ずることである。これを減少させるための手法として Féjer 和、Lancyes の σ -factor による平滑化が適用されているが充分ではない⁵²⁾。実際の計算においてはフーリエ級数を有限項で打ち切りざるを得ないのであるから安定したモデルをあてはめるためには、フーリエ級数の最初の部分(打ち切りフーリエ級数)と誤差項における一様収束しない部分を分離する誤差限界の判定基準が必要である。

この問題は広義には、統計的誤差処理の過程と見做すことができるので、この判定基準として

AICを導入することが同様に可能である。

④ フーリエ級数の有限個の部分和について AIC を算定し、原データに対するフーリエモデルの適合性を比較した結果、MAICE モデルとして採用された Model (29, 14) が適用地域の地形変動を最も的確に再現していることが、推定値から算定した自己相関関数および地形ブロックダイアグラムから認められた。

以上、本論の大部分が情報量基準 AIC に関する記述によって占められた。それは、AIC の応用やそれに言及した論文が内外を問わず膨大なものになりその有効性が実際の応用において卓抜な成果を上げているにもかかわらず、実験・観測データの解析や統計処理がかなりの部分で共通の基盤となって我々の研究領域でほとんどその実例応用例を目にすることがなく、統計学に精通しない著者等自身が AIC に関する論文を取り上げたゼミの中で、また実際のいくつかのデータ解析を通して、あらためてその有効性を身にしみて感じてきたからである。

又再三取り上げてきた多項式モデルあるいはフーリエモデルによる地形データの構造決定における適合度検定の問題^{2), 10)} (データに対するモデルの適合度は両者の残差二乗平均あるいは決定係数で判定されるが、いくつかの適用事例で認められた様にそれらの適合度判定は両者の空間分布間の適合度を必ずしも反映しない。従って両者の空間的分布の近似度を評価する誤差尺度を導入あるいは開発する必要がある)は、赤池²⁰⁾⁻²⁴⁾・坂元⁵⁴⁾等が指摘してきた「従来の統計的仮説検定の有効性の問題」と同様に解釈して議論できるものと考えたからである。

フーリエモデルの最大尤度を誘導するための数式や記述法の多くは赤池^{20)-21), 35)}・坂元⁵⁴⁾・田辺³³⁾・中川³⁴⁾等の論文から引用させていただいたものであるが、浅薄な理解や曲解による部分があるとすれば、いうまでもなくそれは著者等の統計に対する浅学さと日頃の勉強不足によるものであり、率直な御批判を賜われれば幸いである。

なお AIC は、ソフトウェアパッケージあるいはサブプログラムライブラリーの統計量として TMSAC⁵⁵⁾、SALS^{34), 56)}、SSL II、CAPDAP⁵⁷⁾⁻⁶¹⁾等にすでに導入されている。

本研究遂行にあたり深い御理解と有益な御指摘を頂いた鳥取大学神崎康一教授、ゼミの討論の中で多くの示唆を受けた藤井禧雄助教授ならびに林業工学研究室の諸先生方に深謝したい。東南アジア研究センターの内田晴夫、大型計算機センターの高井孝之の各氏には参考文献の紹介・収集について御協力を得た、併せて謝意を表したい。又、川那辺三郎教授をはじめとする当演習林の教職員の方々にも多くの励ましを頂戴した、心から御礼申し上げたい。

なお本研究の一部は昭和57年度文部省科学研究費補助金(奨励研究A)によったこと、計算処理は京大大型計算機センター FACOM M-200 を用いたことを付記する。

引用文献

- 1) 山本俊明・斉藤克郎・芝正己・佐々木功：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、京大演報，53，184-201，(1981)
- 2) Shiba, M., Yanamoto, T. and Sasaki, I.: Quantification of Terrain Variation in Mountainous Regions based upon Numerical Map Analysis by means of Electronic Computer (II), Bulletin of the Kyoto University Forest, 54, 140-163, (1982)
- 3) Preston, 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)
- 4) Frederiksen, P.: Terrain analysis and accuracy prediction by means of the Fourier transformation.

- Photogrammetria, **36**, 145-157, (1981)
- 5) 平野昌繁: 平衡形の理論, 地理評, **45**, 703-715, (1972)
 - 6) Craig, R. G.: A computer program for the simulation of landform erosion, Computer and Geosciences, **6**, 111-142, (1980)
 - 7) Freibergen, W. and Grenander, U.: Surface patterns in theoretical geography, Computer and Geosciences, **3**, 547-578, (1977)
 - 8) 近藤次郎: 数学モデル, 丸善, 2~134, (1976)
 - 9) 平野昌繁・横田修一郎: 西南日本に例をとった電子計算機による地形数値解析, 地理学評論, **49-7**, 440-454, (1976)
 - 10) 芝正己・佐々木功: 地形解析に対するスペクトル概念の応用 (XII) Model 選択規準としての適合度検定と AIC, 日林講, 1430, 275, (1983)
 - 11) 山本嘉一郎: 電算機による Trend Surface Analysis とその使用例, 地質学雑誌, **79**, 5, 349-362, (1973)
 - 12) 芝正己・佐々木功: 地形解析に対するスペクトル概念の応用 (X)~(XI), 日林講, 1213, 1214, 158, (1982)
 - 13) 奥野隆史: 計量地理学の基礎, 大明堂, 21-351, (1977)
 - 14) 平野昌繁・横田修一郎: 電子計算機を利用した地形数値解析の実際と問題点, 地理評, **48**, 648-650, (1975)
 - 15) 中野尊正: 地域の計量と評価, 鹿島出版会, 4-478, (1970)
 - 16) 石水照雄・奥野隆史: 計量地理学, 共立出版, 2-143, (1975)
 - 17) 奥野隆史・西岡久雄: 地域の統計的分析, 大明堂, 67-272, (1976)
 - 18) 坂牧和博・山本嘉一郎・西脇二一: 地質学におけるデータ解析, 石油技術協会誌, **41**, 4, (1976)
 - 19) Frieden, B. R.: Probability, Statistical Optics, and Data Testing, Springer-Verlag, 70-339, (1983)
 - 20) Akaike, H.: A new look at the statistical model identification, IEEE Trans. Autom. Contr., **AC-19**, 716-723, (1974)
 - 21) 赤池弘次: 情報量規準 AIC とは何か, 数理科学, **153**, 5-11, (1976)
 - 22) —: 統計的検定の新しい考え方, 数理科学, **198**, 51-57, (1979)
 - 23) —: エントロピーとモデルの尤度, 日本物理学会誌, **35**, 7, 608-614, (1980)
 - 24) —: モデルによってデータを測る, 数理科学, **213**, 7-10, (1981)
 - 25) 石黒真木夫: 情報量規準: モデルの良さをはかるためのものさし, BASIC 数学, **11**, 5, 40-45, (1978)
 - 26) 柴田里程: 統計的推測方式とモデル選択, 数理科学, **213**, 45-50, (1981)
 - 27) 竹内啓: 情報量統計量の分布とモデルの適切さの規準, 数理科学, **153**, 12-18, (1976)
 - 28) Shibata, R.: Asymptotically efficient selection of the order of the model for estimating parameters of a linear process, Ann. Inst. Statist. Math., **8**, 147-167, (1980)
 - 29) 赤池弘次: 統計的推論のパラダイムの変遷について, 統計数理研究所彙報, **27**, 5-12, (1980)
 - 30) 竹内啓・大橋靖雄: 統計的推測 2 標本問題, 入間「現代の数学」[11], 日本評論社, 4-164, (1981)
 - 31) Akaike, H.: Fitting autoregressive models for Prediction, Ann. Inst. Statist. Math., **21**, 243-247, (1969)
 - 32) —: Statistical prediction identification, Ann. Inst. Statist. Math., **22**, 203-217, (1970)
 - 33) 田辺國士: 統計的データの誤差処理, bit 臨時増版, 1383-1395, (1975)
 - 34) 中川徹・小柳義夫: 最小二乗法による実験データ解析プログラム SALS, 東大出版会, 11-196, (1982)
 - 35) Akaike, H.: On entropy maximization principle, In Applications of Statistics (P. R. Krishnaiah, ed.), Amsterdam: North-Holland, 27-41, (1977)
 - 36) Shimizu, R.: Entropy maximization principle and selection of the order of an autoregressive Gaussian process, Ann. Inst. Statist. Math., **30A**, 263-270, (1978)
 - 37) 赤池弘次・中川東一郎: ダイナミックスシステムの統計的解析と制御, サイエンスライブラリ情報電算機 9, サイエンス社, (1972)
 - 38) 北川源四郎・赤池弘次・大津皓平: 局所定常自己回帰モデルによる適応制御系の実現, 統計数理研究所彙報, **27**, 1, 95-105, (1980)
 - 39) Akaike, H.: Autoregressive model fitting for control, Ann. Inst. Statist. Math., **23**, 163-180, (1971)
 - 40) —: A Bayesian extension of the minimum AIC procedure of autoregressive model fitting, Biometrika, **66**, 237-242, (1979)
 - 41) Ogata, Y.: Maximum likelihood estimators of incorrect Markov models for time series and derivation of AIC, Journal of Applied Probability, **17**, No.1, 59-72, (1980)
 - 42) Tanabe, K.: Fitting Regression Curves and Surfaces by Akaike's Information Criterion, Inst.

- Statist. Math. Research Memorandum, No. 62, (1974)
- 43) — : Statistical regularization regularization of a noisy ill-conditioned system of linear equations by Akaike's information criterion, *Computation and Analysis*, **6**, No.4, 2-25, (1975)
 - 44) Sugiura, N.: Further analysis of the data by Akaike's information criterion and the finite corrects, *Commun. Statist.* **A7**(1), 13-26, (1978)
 - 45) Akaike, H.: A new look at the Bayes procedure, *Biometrika*, **65**, 53-59, (1978)
 - 46) — : Likelihood and Bayes procedure, *Trab. Estadist.*, **31**, 143-166, (1980)
 - 47) Ozaki, T.: Marked point processes and non-linear systems modelling of daily rainfall and riverflow time series in stochastic hydrology, Technical Report No.148, 1-16, (1981)
 - 48) Tong, H.: Determination of the order of a Markov chain by Akaike's information criterion, *J. Appl. Prob.*, **12**, 488-497, (1975)
 - 49) 有本卓：確率・情報・エントロピー，森北出版，28-217，(1980)
 - 50) 田辺國士：不適切問題への統計的アプローチ，数理科学，**153**，60-64，(1976)
 - 51) 秦野和郎・秦野常世・二宮市三：複合多項式による関数近似，情報処理学会論文誌，**23**，6，617-624，(1982)
 - 52) 日野幹雄：スペクトル解析，朝倉書店，10-286，(1977)
 - 53) 宮川洋・今井秀樹：高速フーリエ変換，科学技術出版，1-218，(1979)
 - 54) 坂元慶行・石黒真木夫・北川源四郎：情報統計学，情報統計学講座A・5・4，共立出版，1-154，(1983)
 - 55) Akaike, H., Arahata, E. and Tada, F.: TIMSAC-78, Computer Science Monograph, **11**, (1979)
 - 56) 中川徹・小柳義夫：最小二乗法標準プログラムSALS(第2版)利用の手引き，東大大型計算機センター，(1979)
 - 57) Sakamoto, Y. and Akaike, H.: Analysis of cross-classified data by AIC, *Ann. Inst. Statist. Math.*, **30**, B, 185-197, (1978)
 - 58) Sakamoto, Y.: A model for the optimal pooling of categories of the predictor in a contingency table, *Inst. Statist. Math. Research Memorandum*, No. 119, (1977)
 - 59) 坂元慶行：カテゴリカルデータにおける変数選択 — プログラム CAPDAP を中心に，統計数理研究所彙報，**28**，1，135-155，(1981)
 - 60) Katsura, K. and Sakamoto, Y.: CAPDAP, Computer Science Monograph, **14**, (1980)
 - 61) — : CAPDAP A categorical data analysis program package, Computer Science Monograph, **14**, (1980)

Résumé

Fourier series provide convenient means for interpolation, and limited extrapolation of numerical data those are oscillatory. A possible general use of double Fourier series would produce numerical indexes that are useful in classification. For example, an objective numerical system for classifying the configuration of landforms could be developed by means of double Fourier series analysis.

However, additional consideration needs to be done which includes (1) assessment of the desirability of leveling of the data prior to Fourier analysis, (2) establishing the guides for grid point spacing and for the number of columns and rows in the grid, and (3) arriving at suitable guides for determining the "cut-off" point with respect to the numbers of terms (index of harmonic degree of terms) which are used in the Fourier series.

The Fourier model fails to accord with the actual landform at places along the edge of the map. This is because the map-edge values of the Fourier model, prior to retilling, are the same at opposite points on any two edges.

Computation of coefficients of a model is a part of Fourier analysis. In addition it is essential to compute measures that express the goodness of fit of the model to the data and then to determine whether the model are statistically significant or not (statistical

hypothesis testing procedure as a method of statistical model construction or identification).

In a practical situation, it was suggested that there was some gap between the theoretical results and the practical procedures of adequate model constructing through the analysis of the digital terrain data.

From the information criterion's point of view, Akaike (1973a) introduced the **MAICE** (**Minimum AIC Estimation**) procedure which selects the model whose structure with its associated parameter values gives the minimum of **AIC**. Akaike's Information Criterion (**AIC**) is defined as follows:

$$\text{AIC} = (-2) \log (\text{maximum likelihood}) + 2(\text{number of free parameter}).$$

AIC examines both the fit of a model and the unreliability of a mode; the former is measured by the first term of the definition, which decreases as the number of parameters within the model is increased, and the latter by the second term, which increases with the number of parameters.

AIC, which is an estimate of the Kullback-Leibler information quantity, provides a powerful and almost automatic procedure for the identification of models free from the ambiguities inherent in the application of conventional hypothesis testing procedure.