

JAERI - M  
84-127

最小2乗法に基づくARモデル・フィッティング・アルゴリズムの評価

1984年7月

林 光二

JAERI-Mレポートは、日本原子力研究所が不定期に公刊している研究報告書です。  
入手の問合わせは、日本原子力研究所技術情報部情報資料課（〒319-11茨城県那珂郡東海村）あて、お申しこしください。なお、このほかに財団法人原子力弘済会資料センター（〒319-11茨城県那珂郡東海村日本原子力研究所内）で複写による実費頒布をおこなっております。

JAERI-M reports are issued irregularly.

Inquiries about availability of the reports should be addressed to Information Section, Division of Technical Information, Japan Atomic Energy Research Institute, Tokai-mura, Naka-gun, Ibaraki-ken 319-11, Japan.

©Japan Atomic Energy Research Institute, 1984

編集兼発行 日本原子力研究所  
印 刷 いばらき印刷(株)

最小2乗法に基づくARモデル・フィッティング・アルゴリズムの評価

日本原子力研究所東海研究所原子炉工学部

林 光二

(1984年6月25日 受理)

本報告は、最小2乗法に基づく自己回帰モデル・フィッティングのための2つのアルゴリズム (Marple および北川・赤池のアルゴリズム) の数値的評価に関するものである。

正弦波あるいは正弦波に白色雑音を付加した信号からサンプルされた3種類のデータを用いて、これらのアルゴリズムのシミュレーション・スタディを行った。この結果、両アルゴリズムは従来のもの (Yule-Walker, Burgのアルゴリズム) に対して以下の点で優れている事が明らかになった。

- (1) データ中に強い周期成分が含まれている場合でも安定な数値計算が行え、またその結果も妥当である。
- (2) 少数サンプル・データのフィッティングの際に起るARスペクトル・ピークの周波数バイアス誤差が非常に小さい。

Evaluation of Autoregressive Model Fitting Algorithms  
Based on Least Squares Method

Koji HAYASHI

Department of Reactor Engineering,  
Tokai Research Establishment, JAERI

( Received June 25, 1984 )

This report is concerned with the numerical evaluation of two algorithms, i.e. the Marple's and the Kitagawa-Akaike's for autoregressive model fitting based on least squares method.

Simulation study on these algorithms was performed using three types of test data which were sampled from sinusoidal signals with and without additive white noise. The results show that these algorithms are superior in the following points to the conventional algorithms, i.e. the Yule-Walker's and the Burg's.

- 1) Even in case where the sampled data contains the strong periodic components, its numerical results are stable and also physically consistent.
- 2) Frequency bias errors in AR spectral peaks which occur in case of fitting to small number of sampled data are very small.

Keywords: Autoregressive Model, Least Squares AR spectrum,  
Marple Algorithm, Kitagawa-Akaike Algorithm,  
Numerical Evaluation

## 目 次

1. はじめに	1
2. 理論およびアルゴリズム	3
2.1 ARモデルと観測値からのモデル同定	3
2.2 Yule-Walker 方程式	3
2.3 Burg のアルゴリズム	4
2.4 Marple のアルゴリズム	5
2.5 北川・赤池のアルゴリズム	10
3. シミュレーション・スタディ	15
3.1 試験データ	15
3.2 ケースⅡ（正弦波+白色雑音；その1）の解析	15
3.3 ケースⅢ（正弦波+白色雑音；その2）の解析	17
3.4 ケースⅠ（正弦波のみ）の解析	19
3.5 北川・赤池の場合のAIC	22
3.6 各方法のAICの考察	23
3.7 各方法の計算時間	25
4. 結 論	46
謝 辞	47
参考文献	47

## CONTENTS

1. Introduction .....	1
2. Theory and Algorithms .....	3
2.1 AR model and its identifications using the observations .....	3
2.2 The Yule-Walker equation .....	3
2.3 The Burg algorithm .....	4
2.4 The Marple algorithm .....	5
2.5 The Kitagawa-Akaike algorithm .....	10
3. Simulation Studies .....	15
3.1 Test data .....	15
3.2 Analysis of Case-II data ( sinusoid + white noise ; part 1 ) ...	15
3.3 Analysis of Case-III data ( sinusoid + white noise ; part 2 ) ..	17
3.4 Analysis of Case-I data ( sinusoid ) .....	19
3.5 AIC of the Kitagawa-Akaike's case .....	22
3.6 Consideration of AIC for Different Algorithms .....	23
3.7 Computation Time for Different Algorithms .....	25
4. Conclusion .....	46
Acknowledgement .....	47
References .....	47

## 1. はじめに

今日では、原子力の分野においても、原子炉雑音解析や原子炉システムの異常診断に自己回帰 (Autoregressive ; AR) モデルを用いた時系列解析手法が盛んに応用されるようになってきた。

このARモデルは数値モデルの1つであるが、炉雑音信号等の実測データからモデルが同定できれば、従来から行われている測定信号のスペクトル推定といった問題はもちろんの事、多変数ARモデルの場合にはそのモデルの構造を利用して、各信号間に存在する (フィードバック・ループも含めた) 信号伝達経路やシステムが内蔵する雑音源の特性解析や評価といった問題、さらには各信号の将来の変動予測といった問題に至る幅広い応用が可能である。さらに、スペクトル解析だけを見ても、FFT (高速フーリエ変換) 法やBlackman-Tukey法に比較して、分解能が良い、なめらかなスペクトルが得られる、少数サンプル・データの場合でも妥当な結果が得られる、等の特徴を持っている。

ただし、ARモデルを用いた解析にも幾つかの欠点のある事が知られている。適用する測定データがある程度の定常性を満足していなければならないという点は他の時系列解析手法にも言える事であるが、その他に、測定データはAR過程またはその近似で表現できる変動をしているという仮定を持つ事、モデルの構造上広い周波数範囲をカバーできない事などである。

ところで、原子炉システムなどで測定されるデータは必ずしもARモデルの適用に都合の良いデータばかりではないが、このようなデータこそシステムの動特性解析や異常診断にとって重要な情報源であるケースが多く見られる。例えば、出力上昇時の中性子束変動データの様に大幅なDC成分の変動とともに種々の雑音源からの成分が新たに加わってくる様な非定常な変動データの場合、測定データ中に振動成分等の線スペクトル状の強い成分が重畳している場合、あるいはシステム自体に非線型要素が含まれている場合、さらには測定データが極めて少ない場合、など枚挙にいと間がない。

このようなケースのデータを個々に解析する手法が無い訳ではないが、多変数フィードバック・システムを対象として信号伝達経路や雑音源の解析を行う場合にはどうしてもARモデルの適用が不可欠である。明らかに非定常であるとか非線形であるといった局端なケースは除くとしても、通常得られる測定データに対して (例えそれが近似的なものであったとしても) 適合するARモデルが同定できれば良いのであるが、現実にはモデル同定を行ってみるとその結果が妥当性を欠くとか、悪くすると計算処理が不安定となって結果すら得られないといった事態に出くわす例も珍しくない。

現在知られているARモデルの同定アルゴリズムの中では、相関関数の推定値を基にしてYule-Walker方程式を逐次解いていくYule-Walker法が最も一般的である。このアルゴリズムは取り扱う時系列データが完全に定常でしかも理想的な相関関数の推定値が得られる場合でも、正弦波の様な線スペクトル成分が強いデータに対してはモデルをうまく同定できないという欠点を持っている。その意味では、数あるアルゴリズムの中で最も制約条件の厳しい部類に属している。

これに対してBurgのアルゴリズムは、Yule-Walker法と同様にYule-Walker方程式を逐次

解いていくという方法であるが、方程式を解くのに必要な相関関数をも逐次計算が安定性を保つ様に推定していくという構造を持っている。このため、線スペクトル様の成分が重畳したデータに対しても妥当な結果が得られるという特徴がある。しかし、反射係数の計算に特別な漸化式を利用している関係上このアルゴリズムは多変数のモデル同定には使用できないという根本的な欠点を持っている。さらに、取り扱うデータのサンプル数が少ないと（Yule-Walker 法程ではないにしても）、スペクトル推定に応用した場合にピーク割れとかピーク周波数のずれといった問題を生ずる事も知られている。

ところで最近になって、Yule-Walker 法や Burg 法の欠点を改善した幾つかのアルゴリズムが提唱された。その一つは L. Marple によって 1980 年に発表されたもの<sup>[2]</sup>で、もう一つは北川・赤池らによって 1978 年に発表されたもの<sup>[4]</sup>である。両アルゴリズムとも最小 2 乗法を利用した方法であるが、Marple の場合内容的には Burg 法の発展形であり、正逆両方向のモデルの予測誤差エネルギーを最小化する事を前提として特別な方程式を構成する 1 変数 AR モデル用のアルゴリズムであるのに対して、北川・赤池らの場合はデータ個数個（正確に言えばデータの数からモデルの次数を引いた数分）の AR モデルの連立方程式に対してその残差共分散を最小化する方向で最小 2 乗法を適用するという方法である。Marple の方法はその構造から 1 変数 AR モデルにしか適用できないが、北川・赤池の方法は多変数の場合にも適用できるという特徴を持っている。

本報告は、以前に行った Yule-Walker 法と Burg 法の比較検討<sup>[1]</sup>の続編として、Burg 法の結果も交えながら Marple のアルゴリズムと北川・赤池らのアルゴリズムの特徴や性質を明らかにし、特に原子炉雑音解析や異常診断に不可欠な少数サンプル・データ解析への応用を検討するものである。



## 2. 理論およびアルゴリズム

### 2.1 ARモデルと観測値からのモデル同定

不規則な変動を示す平均値 0 の定常過程

$$\{x_k\} \dots\dots\dots (2.1)$$

が与えられたとき、時刻  $k$  の値  $x_k$  は下記の形式で表現できるとする。これを自己回帰 (AR) モデルと呼ぶ。

$$x_k = \sum_{m=1}^M a_{M,m} x_{k-m} + \epsilon_k \dots\dots\dots (2.2)$$

ここで、 $\{a_{M,m}; m=1, \dots, M\}$  は AR 係数

$M$  は AR モデルの次数

$\epsilon_k$  は白色雑音

AR モデルは (2.2) 式からわかる様に、時刻  $k$  以前の過去の値と線形関係にある部分と、過去の値とは全く相関を持たない白色雑音の部分から構成されている。

いま、実際に観測されたサンプル数  $N$  の平均値を除いた有限な時系列データ

$$\{x_k; k=1, \dots, N\} \dots\dots\dots (2.3)$$

から AR モデルを同定する問題を考える。(2.2) 式の右辺最終項  $\epsilon_k$  は、

$$\epsilon_k = x_k - \sum_{m=1}^M a_{M,m} x_{k-m} \dots\dots\dots (2.4)$$

ただし、 $k = M+1, \dots, N$

と書く事ができるが、これを残差項あるいは予測誤差と呼ぶ。(2.3) 式の時系列が定常な、かつ AR 過程であってサンプル数  $N$  が十分に大きい場合、残差系列  $\epsilon_k$  は

$$\begin{aligned} E\{\epsilon_k\} &= 0 \\ E\{\epsilon_k^2\} &= \sigma^2 \end{aligned} \dots\dots\dots (2.5)$$

を満たす白色雑音 (ガウス過程) と考えることができる。したがって、(2.3) 式のデータから AR モデルを同定する問題は、モデルの次数  $M$  と  $M$  個の AR 係数  $\{a_{M,m}\}$  および残差系列の共分散  $\sigma_M^2$  を決める問題に他ならない。

### 2.2 Yule-Walker 方程式

時系列  $\{x_k\}$  の相関関数  $r_l$  は下式に定義される。

$$r_l = \lim_{N \rightarrow \infty} \frac{1}{N-l} \sum_{k=1}^{N-l} x_k x_{k+l} \dots\dots\dots (2.6)$$

( $l = 0, \pm 1, \pm 2, \dots$ )

この相関関数  $r_l$  と AR 係数、残差系列の共分散  $\sigma^2$  の間には次の関係がある。これを Yule-

Walker方程式と呼ぶ。

$$R_M \cdot A_M = \Sigma_M \quad \dots\dots\dots (2.7)$$

ここで

$$R_M = \begin{bmatrix} r_0 & r_{-1} & \dots & r_{-M} \\ r_1 & r_0 & & \\ \vdots & & \ddots & \\ r_M & \dots & \dots & r_0 \end{bmatrix}, \quad A_M = \begin{bmatrix} 0 \\ a_{M,1} \\ \vdots \\ a_{M,M} \end{bmatrix}, \quad \Sigma_M = \begin{bmatrix} \sigma_M^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

ARモデル同定のアルゴリズムの中で古くから用いられてきたYule-Walker法やBurg法はこの方程式を解く形式のもので、既知数、未知数のとり方は多少異なるもののLevinsonの逐次計算構造を利用している点は全く同じである。

この方程式が成り立つ条件としては、前章でも述べた様に時系列  $\{x_k\}$  が定常なAR過程である事が必要であり、方程式を解くためには相関関数  $r_\ell$  ( $\ell=0, 1, \dots, M$ ) の精度良い推定値が必要である。しかし、一般にサンプル数  $N$  が有限な(2.3)式のようなデータの場合、 $N$  が小さければ小さい程その相関関数の推定値の偏りや分散は大きくなるため、Yule-Walker方程式を用いたARモデル同定は難しくなってくる。例えば、Yule-Walker法のように、推定された相関関数が(2.7)式の相関関数行列の正定性を満さないため方程式が解けなくなる場合や、Burg法の場合のようにあらかじめ他の条件から反射係数  $a_{M,M}$  を与えて安定な計算を行っても初期値として与えた  $r_0$  の推定値の精度が良くないために妥当な解が得られないという場合が起るのである。

**2.3 Burgのアルゴリズム**

後に述べるMarpleのアルゴリズムはBurg法の発展形である。その意味から、ここではBurg法を概観することにする。

Burgは線形予測フィルタとして係数  $\{a_{M,m}; m=1, \dots, M\}$  からなる下記のを考え、

$$x_k = - \sum_{m=1}^M a_{M,m} x_{k+m} \quad \dots\dots\dots (2.8)$$

実測データ  $\{x_k\}$  から係数を決定するに当り、このフィルタの正方向および逆方向の予測誤差系列  $\{f_k\}$ ,  $\{b_k\}$  の共分散の平均(または、エネルギー和)が最小となる様に決定する事を考えた。

$$\sigma_M^2 = \frac{1}{2(N-M)} \sum_{k=1}^{N-M} \{f_{M,k}^2 + b_{M,k}^2\} \quad \dots\dots\dots (2.9)$$

または

$$e_M = \sum_{k=1}^{N-M} \{f_{M,k}^2 + b_{M,k}^2\} \quad \dots\dots\dots (2.10)$$

このとき、

$$\frac{\partial \sigma_M^2}{\partial a_{M,m}} = 0 \quad \left( \frac{\partial e_M}{\partial a_{M,m}} = 0 \right)_{m=1, \dots, M} \quad \dots\dots\dots (2.11)$$

ただし、

$$f_{M,k} = x_{M+k} - \hat{x}_{M+k} = x_{M+k} + \sum_{m=1}^M a_{M,m} x_{M-m+k} \quad \dots\dots\dots (2.12)$$

$$b_{M,k} = x_k - \hat{x}_k = x_k + \sum_{m=1}^M a_{M,m} x_{k+m} \quad \dots\dots\dots (2.13)$$

$$(1 \leq k \leq N-M)$$

問題は (2.11) 式をいかに実現するかであった。ところで、Yule-Walker 方程式 (2.7) 式が成り立つと仮定すると Levinson の漸化式が成り立つから、AR 係数のうち最終項である反射係数  $a_{M,M}$  を決定すると他の係数は全て決まってしまうという関係がある。したがって (2.11) 式は、反射係数  $a_{M,M}$  のみに限定できる。

$$\frac{\partial \sigma_M^2}{\partial a_{M,M}} = 0 \quad \left( \frac{\partial e_M}{\partial a_{M,M}} = 0 \right) \quad \dots\dots\dots (2.14)$$

以上の考えから、(2.14) 式と (2.9) 式または (2.10) 式を利用して、実測データ  $\{x_k\}$  のみから反射係数  $a_{M,M}$  を計算する漸化式を導出し、この結果 Yule-Walker 法では実現できなかった、より最適な解を得る事に成功したのである。

要約すれば、Burg 法は Yule-Walker 方程式を用いる (すなわち、相関関数の推定値を要する) という制約条件下での狭義の最小 2 乗法を適用したアルゴリズムであるという事になる。

このため、データのサンプル数が少ない等の理由で相関関数 (初期値としてのデータの分散値) の精度良い推定値が得られなくなると、AR モデルの同定結果も妥当性に欠ける様になる。

## 2.4 Marple のアルゴリズム<sup>[2]</sup>

Marple は AR 同定アルゴリズムから相関関数の推定値を排除し (Yule-Walker 方程式の使用をやめ)、全ての AR 係数を最小 2 乗法を適用して決定する方法を考えた。

### (1) 解くべき方程式の導出

Burg と同様に正逆両方向の残差系列 (2.12), (2.13) 式を考え、残差系列のエネルギー和 (2.10) 式を最小にする様、各 AR 係数  $\{a_{M,m}; m=1, \dots, M\}$  を決定する問題を考える。まず、(2.11) 式に (2.12), (2.13) 式を代入する事により下式が得られる。

$$\sum_{j=0}^M a_{M,j} r_M(i,j) = 0 \quad \dots\dots\dots (2.15)$$

$$(i = 1, \dots, M)$$

ここで、

$$r_M(i,j) = \sum_{k=1}^{N-M} \{x_{k+M-j} x_{k+M-i} + x_{k+j} x_{k+i}\} \quad \dots\dots\dots (2.16)$$

$$a_{M,0} = 1$$

一方、(2.15) 式と (2.12) 式、(2.13) 式から次式が得られる。

$$\sum_{j=0}^M a_{M,j} r_M(0,j) = e_M \quad \dots\dots\dots (2.17)$$

以上を統合すると下記の新たな方程式が得られる。

$$R_M \cdot A_M = E_M \quad \dots\dots\dots (2.18)$$

ここで、

$$R_M = \begin{bmatrix} r_M(0,0) & \dots\dots\dots r_M(0,M) \\ \vdots & \vdots \\ r_M(M,0) & \dots\dots\dots r_M(M,M) \end{bmatrix}, \quad A_M = \begin{bmatrix} 1 \\ a_{M,1} \\ \vdots \\ a_{M,M} \end{bmatrix}, \quad E_M = \begin{bmatrix} e_M \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \dots\dots\dots (2.19)$$

この方程式こそが、AR係数の最小2乗解を得るための基となる式である。ここでは便宜的に Marple の方程式と呼ぶことにする。この方程式は形式的には Yule-Walker 方程式と同じ様にみえる。しかし、関数  $r_M(i, j)$  (これをその意味から便宜的にエネルギー相関関数と呼ぶ事にする) の行列  $R_M$  は Toeplitz 行列でないため、方程式の解法は全く異なったものとなっている。

(2) Marple の方程式の解法

(2.19) 式を解く最も単純な方法は、 $(M+1)$  元連立 1 次方程式であるから、これを強引に 1 つずつ解いて行く方法であろう。この方法の利点は原理が単純である事、欠点は次数  $M$  が大きくなると時間がかかる、精度が悪くなる等であろう。特にエネルギー相関行列  $R_M$  は通常の相関関数と異なり、次数  $M$  に依存する関数であるから、モデル次数を更新する毎に  $R_M$  の各要素を再計算する必要がある。

このため、方程式を効率良く解くには Levinson の逐次計算アルゴリズムにある様な  $(M)$  次の AR 係数と次段の  $(M+1)$  次の結果とを関係付ける漸化式、さらにエネルギー相関関数についても同様の漸化式が不可欠となる。

Levinson のアルゴリズムを念頭に入れて (2.19) 式のエネルギー相関行列の次数  $M+1$  の場合をみてみると、これは Toeplitz 行列の様な対称行列ではないけれども下式の様な関係を持っていることがわかる。

$$R_{M+1} = \begin{bmatrix} R'_M & \vdots & r_{M+1}(0, M+1) \\ \vdots & \ddots & \vdots \\ r_{M+1}(M+1, 0) & \dots\dots\dots & r_{M+1}(M+1, M+1) \end{bmatrix} = \begin{bmatrix} r_{M+1}(0, 0) & \dots\dots\dots & r_{M+1}(0, M+1) \\ \vdots & \ddots & \vdots \\ r_{M+1}(M+1, 0) & \dots\dots\dots & R''_M \end{bmatrix} \quad \dots\dots\dots (2.20)$$

ここで、

$$R'_M = \begin{bmatrix} r'_M(0,0) & \dots\dots\dots r'_M(0,M) \\ \vdots & \vdots \\ r'_M(M,0) & \dots\dots\dots r'_M(M,M) \end{bmatrix}, \quad R''_M = \begin{bmatrix} r''_M(0,0) & \dots\dots\dots r''_M(0,M) \\ \vdots & \vdots \\ r''_M(M,0) & \dots\dots\dots r''_M(M,M) \end{bmatrix} \quad \dots\dots\dots (2.21)$$

$$r'_M(i, j) = \sum_{k=1}^{N-M-1} (x_{k+M+1-j} x_{k+M+1-i} + x_{k+j} x_{k+i}) \quad \dots\dots\dots (2.22)$$

$$r''_M(i, j) = \sum_{k=1}^{N-M-1} (x_{k+M-j} x_{k+M-i} + x_{k+1+j} x_{k+1+i}) \quad \dots\dots\dots (2.23)$$

$$(0 \leq i, j \leq M)$$

上の  $r_M$  および  $r_M^*$  は (2.16) 式の時間の添字 (time-index) を片側だけ 1 点ずらした表現となっている。そこで, (2.10) 式の残差系列のエネルギー和  $e_M$  について同様の変形をほどこした場合を考える。

$$e_M' = \sum_{k=1}^{N-M-1} (f_{M, k+1}^2 + b_{M, k}^2) \quad \dots\dots\dots (2.24)$$

$$e_M^* = \sum_{k=1}^{N-M-1} (f_{M, k}^2 + b_{M, k+1}^2) \quad \dots\dots\dots (2.25)$$

これを最小化する問題を考えたとき, 下式が得られる。

$$R_M' A_M' = E_M' \quad \dots\dots\dots (2.26)$$

$$R_M^* A_M^* = E_M^* \quad \dots\dots\dots (2.27)$$

ここで,

$$A_M' = \begin{bmatrix} 1 \\ a_{M, 1}' \\ \vdots \\ a_{M, M}' \end{bmatrix}, \quad E_M' = \begin{bmatrix} e_M' \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad A_M^* = \begin{bmatrix} 1 \\ a_{M, 1}^* \\ \vdots \\ a_{M, M}^* \end{bmatrix}, \quad E_M^* = \begin{bmatrix} e_M^* \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \dots\dots\dots (2.28)$$

ところで, (2.22) 式および (2.23) 式の定義からわかる様に  $r_M$  と  $r_M^*$  との間には次の対応関係がある。

$$r_M(i, j) = r_M^*(M-i, M-j) \quad \dots\dots\dots (2.29)$$

以上を整理すると, Marple の方程式は Levinson の場合と同様な形式で下記のように表現できる。

$$\begin{bmatrix} r_{M+1}(0, 0) & \dots\dots\dots & r_{M+1}(0, M+1) \\ \vdots & & \vdots \\ r_{M+1}(M+1, 0) & \dots\dots\dots & r_{M+1}(M+1, M+1) \end{bmatrix} \begin{bmatrix} 1 \\ a_{M, 1}' \\ \vdots \\ a_{M, M}' \\ 0 \end{bmatrix} + \alpha'' \begin{bmatrix} 0 \\ a_{M, M}' \\ \vdots \\ a_{M, 1}' \\ 1 \end{bmatrix} = \begin{bmatrix} e_M' \\ 0 \\ \vdots \\ 0 \\ \Delta_{M+1} \end{bmatrix} + \alpha' \begin{bmatrix} \Delta_{M+1} \\ 0 \\ \vdots \\ 0 \\ e_M^* \end{bmatrix} \quad \dots\dots\dots (2.30)$$

この式と (2.18) 式から, 次の結果が得られる。

$$\Delta_{M+1} = \sum_{m=0}^M r_{M+1}(M+1, m) a_{M, m}' \quad \dots\dots\dots (2.31)$$

ただし,  $a_{M, 0}' = 1$

$$a_{M+1, M+1} = \alpha' = \frac{-\Delta_{M+1}}{e_M'} \quad \dots\dots\dots (2.32)$$

$$e_{M+1} = e_M' + \alpha' \Delta_{M+1} = e_M' (1 - a_{M+1, M+1}^2) \quad \dots\dots\dots (2.33)$$

$$a_{M+1, m} = a_{M, m}' + a_{M+1, M+1} \cdot a_{M, M+1-m}' \quad \dots\dots\dots (2.34)$$

(m = 1, \dots, M)

この様に, Levinson と同様の考え方で M 次の時間シフト AR モデル (ダッシュ付の AR モデル) から M+1 次の AR モデルを計算する漸化式が得られた。残す問題は M 次の AR モデルと M 次のダッシュ付 AR モデルとの関係式を導出する事である。

エネルギー相関関数  $R_M$  はその定義から容易に以下の関係式の成り立つ事がわかる。

$$R'_M = R_M - \begin{pmatrix} x_{M+1} \\ \vdots \\ x_1 \end{pmatrix} (x_{M+1}, \dots, x_1) - \begin{pmatrix} x_{N-M} \\ \vdots \\ x_N \end{pmatrix} (x_{N-M}, \dots, x_N) \quad \dots\dots (2.35)$$

$$R'_M = R_M - \begin{pmatrix} x_N \\ \vdots \\ x_{N-M} \end{pmatrix} (x_N, \dots, x_{N-M}) - \begin{pmatrix} x_1 \\ \vdots \\ x_{M+1} \end{pmatrix} (x_1, \dots, x_{M+1}) \quad \dots\dots (2.36)$$

また残差エネルギー和  $e_M$  は同次の AR モデルであるから、スカラー係数  $\alpha_M$  を介して次の関係にある。

$$E'_M = \alpha_M E_M \quad \dots\dots\dots (2.37)$$

すなわち

$$e'_M = \alpha_M \cdot e_M \quad \dots\dots\dots (2.38)$$

残る AR 係数は (2.26) 式に (2.37) 式を代入し、(2.35) 式や後に導入する関係式を用いて整理すると次の関係式で記述できる。

$$A'_M = \alpha_M (A_M + \beta_1 C_M + \gamma_1 D_M) \quad \dots\dots\dots (2.39)$$

ここで、 $C_M, D_M$  は以下に示す行ベクトルであり、(2.41) 式の関係を満たす様に定義される。

$$C_M = \begin{pmatrix} c_{M,0} \\ \vdots \\ c_{M,M} \end{pmatrix}, \quad D_M = \begin{pmatrix} d_{M,0} \\ \vdots \\ d_{M,M} \end{pmatrix} \quad \dots\dots\dots (2.40)$$

ただし

$$R_M C_M = \begin{pmatrix} x_{M+1} \\ \vdots \\ x_1 \end{pmatrix}, \quad R_M D_M = \begin{pmatrix} x_{N-M} \\ \vdots \\ x_N \end{pmatrix} \quad \dots\dots\dots (2.41)$$

また  $\beta_1, \gamma_1$  はスカラー係数である。(2.39) 式の 3 つの係数を求めるために、次のスカラー量を定義しておく。

$$q_M = (x_{M+1}, \dots, x_1) C_M = \sum_{m=0}^M x_{M+1-m} c_{M,m} \quad \dots\dots\dots (2.42)$$

$$w_M = (x_{N-M}, \dots, x_N) D_M = \sum_{m=0}^M x_{N-M+m} d_{M,m} \quad \dots\dots\dots (2.43)$$

$$h_M = (x_{N-M}, \dots, x_N) C_M = \sum_{m=0}^M x_{N-M+m} c_{M,m} \quad \dots\dots\dots (2.44)$$

$$s_M = (x_N, \dots, x_{N-M}) C_M = \sum_{m=0}^M x_{N-m} c_{M,m} \quad \dots\dots\dots (2.45)$$

$$u_M = (x_N, \dots, x_{N-M}) D_M = \sum_{m=0}^M x_{N-m} d_{M,m} \quad \dots\dots\dots (2.46)$$

$$v_M = (x_1, \dots, x_{M+1}) C_M = \sum_{m=0}^M x_{m+1} c_{M,m} \quad \dots\dots\dots (2.47)$$

このとき、式 (2.39) の両辺に  $R'_M$  をかけると、(2.37) 式との関係から、係数  $\alpha_M, \beta_1, \gamma_1$  が求まり、ダッシュ付 AR 係数は下式で得られる事がわかる。

$$a'_{M,m} = \alpha_M \left[ a_{M,m} + \frac{f_{M,1}(1-w_M) + b_{M,N-M} h_M}{DEN_M} c_{M,m} + \frac{b_{M,N-M}(1-g_M) + h_M f_{M,1}}{DEN_M} d_{M,m} \right] \quad \dots\dots\dots (2.48)$$

(m = 1, \dots, M)

ここで,

$$\alpha_M = \left[ 1 + \frac{f_{M,1}^2 (1-w_M) + b_{M,N-M}^2 (1-g_M) + h_M f_{M,1} b_{M,N-M}}{e_M \text{DEN}_M} \right]^{-1} \dots\dots\dots (2.49)$$

$$\text{DEN}_M = (1-g_M)(1-w_M) - h_M^2 \dots\dots\dots (2.50)$$

最後に行ベクトル  $C_M, D_M$  の計算が残った。この  $C_M, D_M$  も次数が1次下の結果から漸化式を用いて算出する事ができる。詳しい説明は省略するが、(2.41)式の  $R_M$  の代りに時間シフト・ARモデルの  $R'_M$  を用いて  $C'_M, D'_M$  を定義し、これを中介として次の漸化式が得られる。

$$\begin{aligned} c_{M,m} = & c_{M-1,m-1} + \frac{v_{M-1}(1-w_{M-1}) + h_{M-1}s_{M-1}}{\text{DEN}_{M-1}} c_{M-1,M-m-1} \\ & + \frac{s_{M-1}(1-g_{M-1}) + v_{M-1}h_{M-1}}{\text{DEN}_{M-1}} d_{M-1,M-m-1} \\ & + \frac{f_{M-1}a_{M,m}}{e_M} \quad (m=1, \dots, M) \dots\dots\dots (2.51) \end{aligned}$$

$$\begin{aligned} d_{M,m} = & d_{M-1,m-1} + \frac{s_{M-1}(1-w_{M-1}) + h_{M-1}u_{M-1}}{\text{DEN}_{M-1}} c_{M-1,M-m-1} \\ & + \frac{u_{M-1}(1-g_{M-1}) + s_{M-1}h_{M-1}}{\text{DEN}_{M-1}} d_{M-1,M-m-1} \\ & + \frac{b_{M,N-M}a_{M-m}}{e_M} \quad (m=1, \dots, M) \dots\dots\dots (2.52) \end{aligned}$$

ただし  $m=0$  については,

$$c_{M,0} = \frac{f_{M,1}}{e_M} \dots\dots\dots (2.53)$$

$$d_{M,0} = \frac{b_{M,N-M}}{e_M} \dots\dots\dots (2.54)$$

また、 $M=1$  に限り,

$$c_{1,1} = \frac{x_1}{e'_0} + \frac{f_{1,1}a_{1,1}}{e_0} \dots\dots\dots (2.55)$$

$$d_{1,1} = \frac{x_N}{e'_0} + \frac{b_{1,N-1}a_{1,1}}{e_0} \dots\dots\dots (2.56)$$

以上の漸化式を利用して、Marple の方程式を逐次計算で解いて行くことができる。なお、この他に初期値として、下式が与えられる。

$$\left. \begin{aligned} e_0 &= r_0(0,0) \\ \alpha_0 &= \text{DEN}_0 \\ g_0 &= v_0 = \frac{x_1^2}{e_0} \\ w_0 &= u_0 = \frac{x_N^2}{e_0} \end{aligned} \right\} \dots\dots\dots (2.57)$$

$$\left. \begin{aligned} h_0 = s_0 &= \frac{x_1 x_N}{e_0} \\ f_{0,1} &= x_1 \\ b_{0,N} &= x_N \end{aligned} \right\}$$

(3) さらに効率化のための漸化式

アルゴリズム中でエネルギー相関関数  $r_M(i, j)$  が使用されるのは (2.31) 式のみである。これを (2.16) の定義式どおりに計算を実行すると、モデル次数を1次上げる毎に  $M+1$  回の新たな計算が必要となる。ところで、エネルギー相関関数  $r_M(M, m)$  は1次下の結果と次の関係にある事が定義式からわかる。

$$r_M(M, m) = r_{M-1}(M-1, m-1) - x_{N-m+1} x_{N-M+1} - x_M x_m \quad \dots\dots (2.58)$$

$$(m = 1, \dots, M)$$

このため、次数を1次上げた場合、定義式 (2.16) から  $r_M(M, 0)$  を計算しさえすれば、残りは上の漸化式で求めることができる。

また、スカラー係数のうち、 $g_M$  と  $w_M$  については行ベクトル  $C_M, D_M$  の漸化関係から以下の漸化式が成り立つ。

$$g_M = g_{M-1} + \frac{f_{M,1}^2}{e_M} + \frac{v_{M-1}^2(1-w_{M-1}) + s_{M-1}^2(1-g_{M-1}) + v_{M-1} h_{M-1} s_{M-1}}{DEN_{M-1}} \quad \dots\dots (2.59)$$

$$w_M = w_{M-1} + \frac{b_{M,N-M}^2}{e_M} + \frac{u_{M-1}^2(1-g_{M-1}) + s_{M-1}^2(1-w_{M-1}) + s_{M-1} h_{M-1} u_{M-1}}{DEN_{M-1}} \quad \dots\dots (2.60)$$

(4) ARモデルの残差共分散とAIC

残差系列の共分散は、エネルギー和  $e_M$  を用いて次式で計算される。

$$\sigma_M^2 = \frac{1}{2(N-M)} e_M \quad \dots\dots (2.61)$$

また、モデル同定の評価関数であるAICは、

$$AIC = (N-M) \log \sigma_M^2 + 2(M+1) \quad \dots\dots (2.62)$$

と定義される。

### 2.5 北川・赤池のアルゴリズム

$N$ 個の観測データ  $\{x_k\}$  から次数  $M$  のARモデル (2.2式) を同定する問題を考える。ARモデルの残差系列  $\{\epsilon_k\}$  は (2.4) 式のように、観測値から作られる  $N-M$  元連立方程式で表現される。

このとき、残差系列  $\{\epsilon_k\}$  の2乗和 (エネルギー)

$$e_M = \sum_{k=M+1}^N \epsilon_k^2 \quad \dots\dots (2.63)$$

を最小とする様に各ARパラメータを決定する事を考える。Marpleの方法との考え方の相違点



は正逆両方向のモデルのエネルギー和を考えるか単一方向のエネルギー和を考えるかの違いだけである。

(1) 解くべき方程式

残差系列  $\{\epsilon_k\}$  の  $N-M$  元連立方程式 (2.4) の最小 2 乗法による解法は、下記の行列 (ベクトル) 表現

$$X = \begin{bmatrix} x_M & x_{M-1} & \cdots & x_1 \\ x_{M+1} & x_M & \cdots & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-1} & x_{N-2} & \cdots & x_{N-M} \end{bmatrix}, \quad Y = \begin{bmatrix} x_{M+1} \\ x_{M+2} \\ \vdots \\ x_N \end{bmatrix}, \quad A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} \quad \dots\dots\dots (2.64)$$

を用いると、ユークリッド・ノルムの 2 乗

$$\|Y - X \cdot A\|^2 \quad \dots\dots\dots (2.65)$$

を最小にする問題となり、係数  $\{a_m; m=1, \dots, M\}$  に関する偏微分を 0 と置いて得られる下記の正規方程式を解く問題に帰着する。

$$X^t X A = X^t Y \quad \dots\dots\dots (2.66)$$

(t ; 転置行列)

(2) 正規方程式の数値解法

式 (2.66) の解は、数式表現では逆行列を用いて、

$$A = (X^t X)^{-1} X^t Y \quad \dots\dots\dots (2.67)$$

と与えられるところであるが、計算機処理による数値計算では、変数を有限桁 (FORTRAN 単精度実数で 6 桁弱、倍精度実数で 16 桁強) で表現する都合上、加減乗除演算に伴う丸め誤差が発生し、特に (2.67) 式の様逆行列を含む演算を行う場合、行列の規模によっては解が得られないことも希ではない。

これを解決する方法に QR 分解による解法<sup>[5]</sup>がある。長方形行列  $X$  の QR 分解とは、

$$X = QR \quad \dots\dots\dots (2.68)$$

という分解で、 $Q$  は列ベクトルが正規直交系をなす  $(N-M) \times M$  長方形行列、 $R$  は  $M \times M$  の上三角行列である。ただし、

$$Q^t Q = I \quad \dots\dots\dots (2.69)$$

この様な分解ができれば、正規方程式 (2.66) は

$$R^t R A = R^t Q^t Y \quad \dots\dots\dots (2.70)$$

となり、 $X$  がフルランクなら  $R$  は正則で上式は

$$R A = Q^t Y \quad \dots\dots\dots (2.71)$$

と書くことができる。 $R$  は上三角行列であるから、上式は後退代入だけで解くことができる。

(3) Householder 法

QR 分解を実現する方法の 1 つに Householder 変換行列

$$P = I - 2 w w^t \quad \dots\dots\dots (2.72)$$

ただし、 $\|w\| = 1$

を用いた方法がある。行列  $X$  を列単位で処理し、 $m$  回の変換を行えば次式の様分解できる。

$$X = P^{(1)} P^{(2)} \dots\dots P^{(m)} \begin{bmatrix} R \\ - \\ 0 \end{bmatrix} = [Q : Q'] \begin{bmatrix} R \\ - \\ 0 \end{bmatrix} = QR \quad \dots\dots\dots (2.73)$$

ただし、上式では  $P^{(1)} \dots P^{(m)}$  で表わされる  $(N-M) \times (N-M)$  直交行列を  $(N-M) \times M$  行列  $Q$  と  $(N-M) \times (N-2M)$  行列  $Q'$  とに分解した形で表現した。

Householder 変換行列  $P^{(m)}$  は、具体的には以下の様に計算される。まず、 $m$  回目 (第  $m$  列) の変換を

$$X^{(m+1)} = P^{(m)} X^{(m)} \quad \dots \dots \dots (2.74)$$

(ただし  $X^{(1)} = X$ )

と記すこととし、行列  $X^{(m+1)}$  の  $(i, m)$  要素を  $x_{i,m}^{(m+1)}$  と表わすことにする。このとき  $P^{(m)}$  は

$$P^{(m)} = I - v_m v_m^t \frac{1}{h_m} \quad \dots \dots \dots (2.75)$$

ここで、

$$v_m^t = (0, \dots, 0, x_{m,m}^{(m)} \pm \tau, x_{m+1,m}^{(m)}, \dots, x_{N-K,m}^{(m)})^t \quad \dots \dots \dots (2.76)$$

$$\tau^2 = \sum_{i=m}^{N-K} \left\{ x_{i,m}^{(m)} \right\}^2 \quad \dots \dots \dots (2.77)$$

$$h_m = \tau^2 \mp \tau x_{m,m}^{(m)} \quad \dots \dots \dots (2.78)$$

式 (2.76) の  $\tau$  の複号は数値計算上の桁落ちを防ぐ意味から、 $x_{m,m}^{(m)}$  と同符号とする。

行列  $X$  に対して、第 1 列から第  $M$  列までの変換を行った結果として下式を得る。

$$X^{(M+1)} = Q^{-1} X = \begin{bmatrix} R \\ 0 \end{bmatrix} \quad \dots \dots \dots (2.79)$$

ただし、

$$Q^{-1} = P^{(M)} P^{(M-1)} \dots P^{(1)} \quad \dots \dots \dots (2.80)$$

ここで、(2.71) 式右辺を

$$Z = \begin{bmatrix} z_1 \\ \vdots \\ z_{N-M} \end{bmatrix} = Q^t Y \quad \dots \dots \dots (2.81)$$

と置き、上三角行列  $R$  の要素を

$$R = \begin{bmatrix} r_{11} & \dots & r_{1M} \\ & \ddots & \vdots \\ 0 & & r_{MM} \end{bmatrix}$$

と定義すると、係数  $\{a_{M,m}; m=1, \dots, M\}$  の最小 2 乗解は下記の線形方程式を解くことによって得られる。

$$\begin{bmatrix} r_{11} & \dots & r_{1M} \\ & \ddots & \vdots \\ 0 & & r_{MM} \end{bmatrix} \begin{bmatrix} a_{M,1} \\ \vdots \\ a_{M,M} \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_M \end{bmatrix} \quad \dots \dots \dots (2.82)$$

ただし、(2.81) 式の実際の計算は  $Q$  がユニタリ行列であるから  $Q^t = Q^{-1}$ 、したがって (2.80) 式を用いる。

このとき、残差系列  $\{\epsilon_k\}$  の分散は下式で与えられる。

$$\sigma_M^2 = \frac{1}{N-M} \sum_{i=M+1}^{N-M} z_i^2 \quad \dots \dots \dots (2.83)$$

(4) Householder法によるARモデル同定

一般に最小2乗法を適用する問題では、係数の決定が目的であるが、ARモデルの同定問題では2.1節で述べた様に残差系列の分散も求める必要がある。

このため、(2.81)式に定義したベクトルZは、(2.82)式の係数決定に用いられる要素 $(z_1, \dots, z_M)$ 以外の要素 $(z_{M+1}, \dots, z_{N-M})$ も(2.83)式の計算に必要となる。ここで、

$$Z^{(M+2)} = P^{(M+1)} Z = \begin{bmatrix} r_{1, M+1} \\ \vdots \\ r_{M+1, M+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_M \\ r_{M+1, M+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \dots\dots\dots (2.84)$$

という変換を行うと、(2.83)式は下式で表現できる。

$$\sigma_M^2 = \frac{1}{N-M} \cdot r_{M+1, M+1}^2 \quad \dots\dots\dots (2.85)$$

以上から、行列Xの変換(2.79)式とともにベクトルYの変換(2.81)、(2.84)式を同時に変換すれば計算の効率化がはかれる。いま、データ行列を

$$\left[ \begin{array}{cccc|c} X_M & X_{M-1} & \dots & X_1 & X_{M+1} \\ X_{M+1} & X_M & \dots & X_2 & X_{M+2} \\ \vdots & \vdots & & \vdots & \vdots \\ X_{N-1} & X_{N-2} & \dots & X_{N-M} & X_N \end{array} \right] \quad \dots\dots\dots (2.86)$$

と定義し、これにHouseholder変換の逐次適用によって以下の上側三角行列に変形したとする。

$$\left[ \begin{array}{cccc|c} r_{11} & \dots & r_{1M} & r_{1, M+1} \\ 0 & \dots & \vdots & \vdots \\ \vdots & & & \\ \hline 0 & & r_{MM} & r_{M+1, M+1} \\ \vdots & & & \vdots \\ 0 & & & 0 \end{array} \right] \quad \dots\dots\dots (2.87)$$

このとき、求めるARモデルは最大次数をMとしたときの任意の次数 $m (1 \leq m \leq M)$ のモデルについて、

$$\sigma_m^2 = \frac{1}{N-M} \sum_{i=m}^M r_{i+1, M+1}^2 \quad \dots\dots\dots (2.88)$$

$$a_{m, m} = r_{m, m}^{-1} \cdot r_{m, M+1} \quad \dots\dots\dots (2.89)$$

$$a_{m, i} = r_{i, i}^{-1} \left[ r_{i, M+1} - \sum_{j=i+1}^m a_{m, j} \cdot r_{i, j} \right] \quad \dots\dots\dots (2.90)$$

$$(i = m-1, \dots, 1)$$

という様に計算できる。

(5) 計算プログラム上での具体化<sup>[6]</sup>

式(2.86)のデータ行列は $(N-M) \times (M+1)$ 行列である。実際の計算処理ではNが大きいところの $(N-M) \times (M+1)$ 個の記憶容量の確保が意外と難しい。

そこで、データ行列を以下の実現可能な大きさに分解する。

$$\left[ \begin{array}{c|c} X & Y \end{array} \right] = \left[ \begin{array}{c} \overbrace{\begin{array}{c} X_0 \\ \hline X_1 \\ \hline X_2 \\ \hline \vdots \\ \hline X_{IB-1} \\ \hline X_{IB} \end{array}}^{M+1} \\ \hline L \\ \hline L \\ \hline \vdots \\ \hline L \\ \hline L' \end{array} \right] \left. \vphantom{\begin{array}{c} X \\ Y \end{array}} \right\} N-M \quad \dots\dots\dots (2.91)$$

ここで、 $X_0$  は  $(M+1) \times (M+1)$  行列  
 $X_1, \dots, X_{IB-1}$  は  $L \times (M+1)$  行列  
 $X_{IB}$  は  $L' \times (M+1)$  行列 とする。

$N$  が大きく、 $(N-M) \times (M+1)$  の記憶容量がとれない場合、 $(L+M+1) \times (M+1)$  が実現可能な大きさとなる様に  $L$  を選ぶ。次に (2.91) 式のようにデータ行列の上側  $(M+1) \times (M+1)$  行列  $X_0$  以外の残りの部分を  $L$  行ずつ区切り  $X_1, \dots$  をつくる。残りを  $L'$  行とおく。IB をブロック数とする。

以上の準備の後、 $X_0, X_1$  行列を三角行列に変形する。

$$\left[ \begin{array}{c} X_0 \\ \hline X_1 \end{array} \right] \rightarrow \left[ \begin{array}{c} R_1 \\ \hline 0 \end{array} \right] \quad \dots\dots\dots (2.92)$$

次に残る  $X_2, \dots, X_{IB}$  を逐次変形していく。

$$\left[ \begin{array}{c} R_{j-1} \\ \hline X_j \end{array} \right] \rightarrow \left[ \begin{array}{c} R_j \\ \hline 0 \end{array} \right] \quad (j=2, \dots, IB) \quad \dots\dots\dots (2.93)$$

こうして得られた三角行列  $R_{IB}$  は元の  $(N-M) \times (M+1)$  データ行列を一度に変換した場合と同じものが得られる。

(6) AIC

ARモデル同定の評価関数であるAICは、Householder法を用いた北川・赤池のアルゴリズムの場合、モデルの最大次数をMとして計算を行ったときm次のモデルのAICは、

$$AIC_m = (N-M) \log \sigma_m^2 + 2(m+1) \quad \dots\dots\dots (2.94)$$

と定義される。

### 3. シミュレーション・スタディ

この章では、前章に述べたMarple および北川・赤池の両アルゴリズムが実際のデータ処理で如何なる特徴や性質を持つかを明らかにする。具体的には、素姓の知っている典型的な試験データのモデル・フィッティングを通じて、Yule-Walker法やBurg法の結果も交えながら、データのサンプル数、モデルの次数等のパラメータと同定結果との関係性を評価し、特に原子炉異常診断への応用が考えられている少数サンプル・データのフィッティングやオン・ライン処理での問題点を検討する。

#### 3.1 試験データ

データとしてはTable 3.1, Fig. 3.1に示した3つのケースを用意した。これらのデータは以前Burgのアルゴリズムの評価を行った際に使用したのと同じである<sup>[1]</sup>。

ケースⅠは正弦波だけから作られている。スペクトル表現では $\delta$ 関数(線スペクトル)となり、ARモデル・フィッティングでは最も困難なデータのひとつである。周波数 $f_0$ は1.5 Hzに選び、そのサンプリング間隔 $\Delta t$ は0.061 secとした。振幅項 $A$ は約2.5とした。

ケースⅡはケースⅠの正弦波に白色雑音を重畳したものである。白色雑音の最大振幅は正弦波の振幅の10%前後となる様に設定されている。サンプリング間隔 $\Delta t$ も0.061 secと同じである。

ケースⅢはケースⅡと同じ正弦波+白色雑音の信号であるが、サンプリング間隔 $\Delta t$ をその $1/10$ (0.006 sec)に設定し、スペクトル上の $f_0$ が $\Delta t$ で決まる解析周波数の上限に対してケースⅠ、Ⅱよりも1ディケード低域側にずらしたものである。

各ケースの正弦波の1周期は時間にして約0.666 secであるが、サンプル数 $N$ ではケースⅠ、Ⅱが約12点、ケースⅢでは約111点で表現されることになる。

#### 3.2 ケースⅡ(正弦波+白色雑音; その1)の解析

##### 3.2.1 サンプル数 $N$ の効果

まず3つのデータの中で最もフィッティングが容易と思われるケースⅡのデータの結果からみていくことにする。フィッティングの結果を直観的に評価できる様に、得られたARモデルからパワー・スペクトルを推定した。ここでは、Burg, Marpleおよび北川・赤池の3つの方法による結果を示す。

Fig. 3.2はデータ・サンプル数 $N=4000$ (1.5 Hzの正弦波にして約366周期分)、モデルの次数 $M=50$ の場合のパワー・スペクトルである。各方法とも実に良く一致したスペクトルが得られている。

スペクトルの形状についてみると、正弦波の発振周波数 $f_0=1.5$  Hzを表わすピーク成分と、重畳された白色雑音による周波数領域全体にわたり平坦な成分から構成されている。ピーク周波

数は正しく 1.5Hz である。ピークの高さは正弦波の振幅が  $A=2.5$  であるから真値は  $P_0=3.125$  ( $\text{rms}^2/\text{Hz}$ ) であるが、これもほぼ正しい結果を得ている。なお、白色雑音については真値は不明である。データ・サンプリングの際にローパス・フィルター等を使用していないため、白色雑音成分については強力なエイリアシング (aliasing) が起っていると考えられ、後に示すケースⅢの結果に比べ白色雑音のレベルが約 10 倍高くなっている。デジタル化したデータの処理としては、この図の白色雑音のレベルは妥当な結果であると考えられる。

Fig. 3. 3 はデータ・サンプル数  $N=120$  (約 11 周期分)、モデルの次数  $M=16$  の結果である。各方法とも少しづつ差異がみられる。

ピーク周波数  $f_0$  は各方法とも一致し真値の 1.5Hz となっている。ところがピーク高は各方法とも異なる値をとり、しかもいずれの結果も妥当でない。真値からのずれは Marple、北川・赤池法が 2~3 倍で Burg 法では実に約 100 倍となっている。一方、白色雑音についてみると、前図 ( $N=4000$ ) の結果に比べてスペクトルの波打ちが大きくなり幾つかのピーク成分が現われている。これではもはや白色雑音とはいえない。モデルの次数を増してもスペクトルの波打ちが激しくなる (多数のピークが現われ、そのピーク高も大きくなる) だけで、平坦な表現は得られない。この理由は、サンプル数  $N$  が 120 と少ないため本来無相関である白色雑音の成分が (他の時刻のデータと) 相関を持ち始めるためであると考えられる。したがって、このスペクトル推定結果が全く間違っているという訳ではない。なお、この部分については Burg 法と Marple 法が割合に良く一致しており、北川・赤池法だけが少し異なる形状をしている。この原因は、アルゴリズムの正逆両方向の信号を取り扱うか、片側だけかという構造に由来するものと思われる。

Fig. 3. 4 はサンプル数を  $N=60$  (約 5.5 周期分) と前図の  $\frac{1}{2}$  にした場合の例で、モデル次数は  $M=5$  である。

ピーク周波数  $f_0$  は Marple、北川・赤池法ともに一致して 1.5Hz であるが、Burg 法では少し高目になっている。またピーク高はいずれの方法も真値の 10~30 倍程度と大きい。白色雑音の表現は前図よりも悪くなっており、モデル次数を前図と同様  $M=16$  とするとスペクトルの波打ちはより激しいものとなる。

### 3. 2. 2 サンプル数 $N$ とモデルの最適次数

サンプル数  $N=4000, 120, 60$  のケースⅡのデータに対するモデル・フィッティング時に得られた AIC (情報量規準) を Fig. 3. 5 に示す。

解析しようとする有限のデータ中に、デジタル化される前の信号が持っていた情報がどの程度反映されるかは、サンプル数  $N$  とサンプリング間隔  $\Delta t$  の大きさによって決まる。一方、AR モデルの場合、表現できる情報の量は AR 係数の数 (モデル次数) で決まる。したがって、ある物理現象を表現するための妥当なモデルを得るためには、その物理現象がうまく表現できる程度の大きさを持った  $N$  と  $\Delta t$  を選択する必要がある。逆に、 $N$  と  $\Delta t$  がすでに与えられている様なサンプル・データは、その物理現象の表現に制約が課せられる。したがって、このときのモデル次数の大きさも何らかの制約が存在する。モデルのフィッティングの良さを評価する AIC は、この意味で与えられたサンプル数  $N$  の基での最適次数を探す指針となる。

Fig. 3. 5 (下段) は  $N=4000$  (約 366 周期分) の結果である。サンプル数が十分大きいため、

20次前後から高次ではほとんど変化しない。MAICE (最小AIC評価)の次数は、Marple, Burgの各方法が $M=39$ , 北川・赤池法は $M=44$ であった。北川・赤池法は他に比べて少し高い次数となっているが、先にFig. 3. 2でみた様に $M=50$ でのパワー・スペクトルは各方法とも全く良く一致している。スペクトルの結果から、サンプル数が十分ある場合にはモデル次数が40前後でケースⅡのデータが十分表現できていることがわかる。

一方、Fig. 3. 5 (中段)は $N=120$  (約11周期分)の結果である。MAICEによると、Marple, 北川・赤池の各方法が $M=16$ であり、Burg法は $M=6$ である。Burg法の場合最適次数がかなり低くなっているのは、次数を上げて他の方法でフィッティング誤差( $\sigma_M^2$ )が小さくならないためである。この差異はアルゴリズムが最小2乗フィッティング (Marple, 北川・赤池法)か否 (Burg法)かという点から生じたものと考えられる。

最後に $N=60$  (約5.5周期分)の結果Fig. 3. 5 (上段)をみよう。MAICEによると、Burg, Marple, 北川・赤池法の順に $M=5, 5, 6$ となる。

以上から、モデル・フィッティングの目的をデータ中の正弦波成分の周波数 $f_0$ の推定だけに限定すれば、モデル次数は $M=5$ 前後あれば良く、正弦波成分のピーク高 $P_0$ や白色雑音のレベルの推定も目的にする場合は $M=40$ 前後が必要である。

### 3.3 ケースⅢ (正弦波+白色雑音; その2)の解析

今度は前節と同様の信号をサンプリング間隔 $\Delta t$ をその $1/10$ としてサンプリングしたデータについてみよう。一般にARモデルの場合、表現できる周波数範囲が狭い、ARスペクトルは他の手法の結果に比べて低域側の周波数でバイアスを持つなどといわれている。それらの性質を評価する意味からも、このケースⅢの結果は重要である。

#### 3.3.1 サンプル数 $N=4000$ のモデル

##### (1) パワー・スペクトル

Fig. 3. 6はサンプル数 $N=4000$  (約36.6周期分)、モデル次数 $M=100$ のパワー・スペクトルである。Marpleと北川・赤池の両方法は完全に一致している。Burgの場合はピークの先端部が若干異なるがその他の部分では良い一致をみる。これに対してYule-Walker (以下Y-W)法の場合は、高域では完全に一致しているものの、周波数が低くなると他の方法の結果との差異が大きくなり、低域側では完全なバイアス成分を持つことがわかる。なお、本報で用いたY-W法のフィッティング・プログラムは前報<sup>[1]</sup>のものに比較して一部倍精度化など数値計算の安定化をはかったものとなっている (したがって、その結果も前報と比べると良くなっている)。

まず、ピーク周波数 $f_0$ についてみると、各方法とも正しく1.5Hzとなっている。しかし、ピーク高はMarple, 北川・赤池, Burgとも真値の3倍程度大きくなっている。Y-W法はさらに大きい。ケースⅡの解析 (Fig. 3. 2)と比べると、サンプル数 $N$ は同じであるがデータ中に含まれる正弦波成分は約36.6周期という様に $1/10$ である。ピーク高を正しく推定するためには、さらに10倍以上のサンプル数が必要と考えられる。

一方、白色雑音については、Burg, Marple, 北川・赤池の各方法とも妥当な結果である様にみえる。

### (2) 反射係数

先のモデル・フィッティングの際に得られた反射係数（各モデル次数MにおけるAR係数の最終項） $a_M(M)$ をFig.3.7に示す。図から、Marple, Burgの方法は完全に一致していること、北川・赤池法もほとんどこれに近いこと、これに対しY-W法はバイアスを持つことがわかる。反射係数は次数の増加とともに非常に良く0に収束して行く様子が示されている。

### (3) AIC

AICをFig.3.8に示す。各方法とも全体の傾向としては良く似た変動を示している。しかし、詳細にみると方法によりそれぞれ異なった曲線を描いている。MAICEはY-W法でM=34, Burg法でM=51, Marple, 北川・赤池法がM=67という様に大きく異なる。ただし、Burg法とMarple, 北川・赤池法との差異はM=100のスペクトル（Fig.3.6）でみる限りほとんど無い。わずかにピーク成分の形状が異なるのみである。

## 3.3.2 サンプル数N=120のモデル

ここでは、サンプル数N=120（約1.1周期分）のデータについてモデル・フィッティングを行った結果を示す。

### (1) 反射係数

Fig.3.9に得られた反射係数 $a_M(M)$ を示す。前節でみたN=4000の場合（Fig.3.7）と比較してかなり大きく変動していることがわかる。この図ではY-W法を除いた3つの方法が示されている。MarpleとBurgの方法は20次付近まではほぼ一致した動きをしており、20次から高次にかけても多少のバイアス的な差はあるにしても同一の変動傾向を示すことがわかる。これに対して北川・赤池法は4次付近から他と異なる値をとり、高次になるとその変動幅も大きくなる傾向を持つ。

以上の様にサンプル数Nが小さいと反射係数の収束性は悪くなり、またその変動幅も大きくなる傾向をもつ。反射係数の変動幅が小さいという意味からは、正逆両方向の信号を取り扱うMarpleやBurgの方法が北川・赤池法より優れているといえる。

### (2) AIC

Fig.3.10に各方法のAICを示す。Burg, Marpleの方法が非常に明確な極値を持つ（ともにM=13）のに対し、北川・赤池法は減少傾向を持ち、そのMAICEの次数もM=49と非常に高く、計算範囲を拡げればさらに高くなる様子がうかがえる。

今回のケースⅢのデータでサンプル数N=120に含まれる正弦波成分は約1.1周期分しかない。このためモデル次数を上げると高々50次の範囲内では予測誤差（ $\sigma_M^2$ の値）は減少するばかりである。このため、ケースⅡに比べてどの方法のAICも次数に依存した増大傾向が弱まっている。

しかし、上記の理由だけでは北川・赤池法と他の方法との傾向の差は説明できない。これについては後に述べることとする。



### (3) モデル次数に依存するパワー・スペクトルの形状

$N=120$ のデータに対して行ったモデル・フィッティングで得られた各次数のモデルからパワー・スペクトルを推定したのが Fig. 3.11~12である。

Fig. 3.11はMarpleの方法による結果である。先に述べた様に、サンプル数 $N=120$ の場合データ中に含まれる主成分としての正弦波成分が約1.1周期分しかない。このため、モデル次数 $M$ の増加によるフィッティング誤差(残差共分散)の減少の割合はゆるやかである。

$M=5$ の場合、ケースⅡでは主成分である正弦波を表わすピークが明確に現われていたのに比べ、ケースⅢではピーク成分が全く現われず、その代りに低域側から高域側にかけて単調減少する形状となっている。

$M=10$ で $f_0=1.5\text{Hz}$ 付近に鈍いピーク成分が現われはじめ、次数を上げていくと次第に鋭さを増していく。それと同時に高域側でスペクトル成分の波打ちが始まり次第にその変動幅が大きくなっていく。

このフィッティングに用いたデータは先の反射係数やAICで用いたもの( $X_1, \dots, X_{120}$ )と異なり、 $X_{11}, \dots, X_{130}$ である。このためMAICE次数も異なり、 $M=12$ であった。

一方、Fig. 3.12は北川・赤池法の結果である。全体的な傾向としてはMarpleの場合と同じである。明確な差が現われるのは高い次数の場合である。例えば、 $M=43, 50$ をみるとわかる様に、 $f_0=1.5\text{Hz}$ のピーク高が低くなる点、高域でのスペクトルの波打ちの変動幅が大きい点である。

MAICE次数は $M=43$ であった。Marpleの場合と比較して、MAICE次数の異なる理由は後に述べることとする。

## 3.4 ケースⅠ(正弦波のみ)の解析

ここでは、ARモデル・フィッティングにとって最も困難なデータの一つである正弦波の解析についてみることにする。

### 3.4.1 サンプル数 $N=200$ のモデル

ケースⅠデータは正弦波だけをケースⅡと同じ条件でサンプリングしたデータであり、 $N=200$ は18.3周期分の時間長である。白色雑音成分が含まれてないから、サンプル・データの $S/N$ 比は非常に良く、極く短いデータでも精度の良いフィッティングが期待される。しかし、一方ではARモデルの予測誤差(残差系列)を仮定した構造からみると、フィッティングの困難が予想される。

ここでは、十分な大きさのサンプル・データ( $N=200$ )に対して、Yule-Walker, Burg, 北川・赤池, Marpleの各方法を適用した場合に、どの様な結果が得られるかをパワー・スペクトルの表現を通じてみてみよう。

#### (1) MAICE次数のパワー・スペクトル

Fig. 3.13に各方法の結果を示す。モデルの次数はAICが最小のものとなっている。ただし、北川・赤池法の場合は計算過程における最大次数 $M_{\max}$ を50としている。

どのスペクトルも正弦波の周波数  $f_0 = 1.5 \text{ Hz}$  は正しく推定している。ピーク高は Marple が非常に真値に近い値を示しているが、他の方法では全くでたらめな値となっていることがわかる。

図にはないが、北川・赤池法のスペクトルは Marple と同じモデル次数の場合は Marple の結果と良く似た形状を持つ。ただし、ピーク高はやや低目となる。

## (2) モデル次数とスペクトル・ピーク高の関係

Fig. 3.14 にモデル次数に依存した各方法のスペクトル・ピーク高の変化を示す。

Yule-Walker および Burg の方法では、モデル次数を高くしてもピーク高がほとんど一定であること、しかも真値 ( $P_0 = 3.15$ ) からかなりかけはなれた大きさであることがわかる。

Marple の場合、次数を高くして行くとフィッティングの予測誤差が非常に小さくなり、数値計算の精度が保てなくなる ( $M = 14 \sim 15$  以上)。この計算過程の途中で、MAICE 次数  $M = 8$  となる。このときのピーク高は、真値に近いものとなっている。

最後に北川・赤池法であるが、Marple と異なり数値計算上では安定している。しかし、得られた結果 (MAICE 次数  $M = 11$ ) はデータから期待される結果と大きく異なる。この場合、モデル次数を Marple と同じかそれ以下に選ぶ必要がある。これについても、後の AIC の項で述べることにする。

### 3.4.2 データの初期位相の効果

強い周期成分を含むデータの AR スペクトル推定では、サンプル数が少ないとデータ中の周期成分の初期位相の値によって推定されたスペクトル上のピーク周波数  $f_0$  が大きくずれることが知られている。

Burg や Marple の方法では、このピーク周波数のずれの問題はかなり明らかにされている<sup>[1][2]</sup>が、北川・赤池法については未知のままである。ここでは、表現の簡単な正弦波だけのサンプル・データ (ケース I) を用いて、データの初期位相の効果が AR 係数や予測誤差 (残差共分散) にどの様に反映するかを明らかにする。

Fig. 3.15 に Marple の結果、Fig. 3.16 に北川・赤池法の結果を示す。相互に比較しながら話を進めて行くことにする。

#### (1) データ

図の最上段に 19 点から成る正弦波のサンプル・データを示す。この正弦波は 1 周期が約 12 点で表わされている。したがって、サンプル・データの第 1 点目  $X_1$  から第 12 点目  $X_{12}$  までの間に位相は  $360^\circ$  変化する訳である。

そこでサンプル数  $N$  の任意の区間  $\{X_k, \dots, X_{k+N-1}\}$  についてフィッティングを行う場合、フィッティング用データの開始点  $k$  の選び方によりデータの初期位相が変化する。ここでは、Marple の場合  $N = 6$ 、北川・赤池法の場合  $N = 7$  とし、 $k = 1 \sim 12$  の 12 区間のフィッティングを行った。

#### (2) データの分散 $\sigma_x^2$ の推定値

各図の上から 2 段目は、各区間のデータから推定されたデータの分散値  $\sigma_x^2$  を示している。サンプル数が 6 と 7 では使用できるデータ区間が異なるため、推定値は多少量なる。全体の傾向

としては、データの初期位相が $0^\circ$ および $\pm 180^\circ$ 付近で最小、 $\pm 90^\circ$ 、 $\pm 270^\circ$ 付近で最大値をとる。真値 $\sigma_x^2 = 3.12$ に対する推定値のずれの程度は、 $N = 6$ の場合 $0^\circ$ および $\pm 180^\circ$ のとき最も大きく、 $\pm 90^\circ$ 、 $\pm 270^\circ$ ではほとんどずれはない。これに対して、 $N = 7$ の場合 $0^\circ$ および $\pm 180^\circ$ では同様に最も大きくずれが、 $\pm 90^\circ$ 、 $\pm 270^\circ$ のときも $N = 6$ のときよりも大きくずれが。

### (3) スペクトル・ピーク周波数 $f_0$ の推定

各図の上から3段目は、フィッティングしたモデルを用いてパワー・スペクトルを推定したときのピーク周波数 $f_0$ の位置を示している。各方法ともモデル次数は $M = 2$ と3の結果が示されている。

Marple、北川・赤池の両方法とも $M = 2$ の場合は1.5Hzから大きくずれることがあること、しかもこの変動はデータの分散値 $\sigma_x^2$ と一致した動きをすることがわかる。これに比較して $M = 3$ の場合はほとんど変化しないことがわかる。これは、Marple、北川・赤池法ともにBurg法の場合<sup>[1]</sup>に比べて初期位相の影響に対し強いという特徴を示している。

### (4) スペクトル・ピーク高の推定値と残差共分散

各図4段目は、スペクトル・ピークの高さの真値 $P_0 = 3.12$ からのずれを示している。また5段目はモデル・フィッティングの予測誤差（残差共分散 $\sigma_M^2$ ）を示している。

Marpleの場合、データの分散 $\sigma_x^2$ の推定値のずれが大きい（推定誤差が大きい）とモデルの予測誤差 $\sigma_M^2$ も大きくなること、さらに $\sigma_M^2$ が大きいとスペクトル・ピーク高は低くなり真値 $P_0$ に近づくことがわかる。

一方、北川・赤池の場合、データの分散 $\sigma_x^2$ の推定誤差が大きいとMarpleと同様にモデルの予測誤差 $\sigma_M^2$ も大きくなること、さらに $\sigma_M^2$ が大きくなるとMarpleの場合とは逆にスペクトル・ピーク高は高くなり真値 $P_0$ に近づくことがわかる。

### (5) AR係数

図の最下段にモデル次数 $M = 2$ および3のAR係数の変化を示す。Marpleの場合、 $M = 2$ ではデータの分散値 $\sigma_x^2$ の推定誤差の大きさに従って第1項 $a_2(1)$ が変化しているが、 $M = 3$ では全く安定していることがわかる。

一方、北川・赤池の場合、各次数ともかなり変化することがわかる。

以上の結果をまとめると、サンプル数 $N$ が小さいとデータの初期位相に依存してデータの分散 $\sigma_x^2$ に推定誤差が生じること、この大きさに依存してモデルの予測誤差 $\sigma_M^2$ が決まること、さらにスペクトル・ピーク $f_0$ のずれはデータの分散 $\sigma_x^2$ の推定精度で決まること、モデルの予測誤差 $\sigma_M^2$ の大きさは、データの分散 $\sigma_x^2$ の推定誤差の大きさに依存すること、そしてピーク高は予測誤差 $\sigma_M^2$ が大きいときに真値 $P_0$ に近づくことが示された。また、このときのAR係数は北川・赤池法よりもMarpleの方法の方が安定であることがわかった。

### 3.4.3 データ長と初期位相の効果との関係

ここでは、データの初期位相によって決定されるスペクトル・ピークの周波数やピーク高のずれが（サンプル数 $N$ とサンプリング間隔 $\Delta t$ によって決まる）データ長とどの様な関係にあるかを述べる。

(1) スペクトル・ピーク周波数  $f_0$  のずれ

$4t$  を一定としサンプル数を  $N=6$  (あるいは7) から増加していった場合のデータに対し、モデル・フィッティングを行った場合、スペクトルのピーク周波数がデータの初期位相に依存してどの様に変化するかを示したのが、Fig. 3.17~3.19である。Burg, Marple, 北川・赤池の3方法について評価を行った。モデル次数は  $M=2$  (図上段(a)),  $M=3$  (図下段(b)) である。参考のため一部ケースⅡの白色雑音を付加したデータについても図示してある。

まず Burg の場合 (Fig. 3.17) についてみてみよう。各次数の結果とも  $N=12$  (約1周期) および  $N=17$  (約1.5周期) ではピーク周波数の推定値のずれが極めて少ないことがわかる。さらにサンプル数  $N$  を増加していくと、ずれの大きさは0に収束していくことがわかる。ただし、Burgの場合モデルの次数が  $M=2$  と3とではほとんど差異がない。ケースⅡ (白色雑音を含む) のデータについても試みたが、Burg法では正弦波だけの場合とほとんど差異はみられない。

Marpleの場合 (Fig. 3.18),  $M=2$  では約1周期以上,  $M=3$  では1周期なくてもほとんどピーク周波数のずれはない。ただし、白色雑音が含まれる場合は  $S/N$  比悪化のために比較的大きいずれを生ずることがわかる。

北川・赤池の場合 (Fig. 3.19), その傾向は Marple とほぼ同じであるが、若干ずれの大きいことが示されている。

## (2) ピーク高のずれ

Fig. 3.20~3.22に Burg, Marple および北川・赤池の各方法によるスペクトル推定値のピーク高の変化を示す。

ピーク高の変化はピーク周波数  $f_0$  の結果とは異なり、真値  $P_0$  への収束性はみられず、各次数  $M=2, 3$  ともフィッティングの方法によらず増大傾向にあることがわかる。このことは、一見奇異にみえるが、ピーク周波数がモデルの次数にほぼ不変な性質をもっているのに対してピーク高は Fig. 3.14 でみた様に強くモデル次数に依存していることを考えれば説明のつくことである。

ピーク高の精度の良い推定を行うためには、ケースⅡやⅢでの結果からもわかる様に十分なサンプル数  $N$  と適切なモデル次数  $M$  を選択する必要がある。

## 3.5 北川・赤池の場合の AIC

ケースⅠ~Ⅲのフィッティングで明らかな様に、データ・サンプル数  $N$  が小さいと北川・赤池法の場合 Marple に比べ MAICE によるモデル次数が高く選ばれることが多い。しかも、結果をみると妥当性に欠く場合も少なくない。例えば、ケースⅢのデータ  $N=120$  のフィッティング (Fig. 3.10) の場合がそうであった。この様な現象は何故起きたのであろうか?

このときのフィッティングではモデルの最大次数  $M_{\max}$  を50として計算したものであった。この  $M_{\max}$  を1~50まで逐次増加しながら、それぞれの場合の MAICE 次数を調査したのが Fig. 3.23 である。これによると、最大次数  $M_{\max}$  を幾つに設定するかによってモデルの最適次数が全く変わってしまうことがわかる。

この原因は北川・赤池のアルゴリズムにあるのである。北川・赤池法ではモデルの最大次数  $M_{\max}$  を与えフィッティングを行えば、 $M_{\max}$  より低い次数のモデルの情報も上三角行列の中に含まれている。しかし、このアルゴリズムでは、解くべき方程式の数はサンプル数  $N$  と最大次数  $M_{\max}$  の差  $N - M_{\max}$  で決まるから、 $M_{\max}$  より下位の次数  $m$  でも方程式の数は同じである。逆にいうと、 $m$  次のフィッティングをする場合、 $M_{\max} - m$  個のデータ  $\{X_1, \dots, X_{M_{\max} - m}\}$  は使用していないことになる。すなわち、 $M_{\max}$  より下位のフィッティングの結果はサンプル数  $N - (M_{\max} - m)$  のデータについて行ったものということになる。

Fig. 3.24に Fig. 3.10に示した  $M_{\max} = 50$  の AIC と、フィッティングに使用するサンプル数が常に  $N = 120$  となる様に逐次  $M_{\max}$  を変えていった場合の AIC を示す。この結果、北川・赤池法の場合ケースⅢの  $N = 120$  のデータについてフィッティングをすると、AIC は明確な極値 ( $M = 14$ ) を持ち、モデル次数の増加とともに上昇傾向になることがわかる。この MAICE 次数  $M = 14$  は Burg, Marple の結果ともほぼ一致する (Fig. 3.10) ことがわかる。

以上から、北川・赤池法ではサンプル数  $N$  が小さいデータの解析を行う場合、最大次数  $M_{\max}$  を設定してフィッティングした結果からそれより下位の次数のモデルを使用する場合には上述の注意が必要であることがわかる。

### 3.6 各方法の AIC の考察

#### (1) 各方法の AIC のより正確な定義式

ケースⅠ～Ⅲのフィッティングを通じて各方法の AIC の計算値をみてきたが、前節の北川・赤池法の問題点を別にしても Burg, Marple の両方法の MAICE 次数は低すぎる傾向にある。というのは、両方法の残差共分散  $\sigma_M^2$  の値は一般に北川・赤池よりも小さく推定されるのが普通だからである。何故  $\sigma_M^2$  が北川・赤池法より小さいのに MAICE 次数は北川・赤池法と同程度、あるいはそれより低い結果しか得られないのであろうか。この原因は、実は AIC の計算式 (2.62) にあるのである。

AIC は以下の様に定義される。

$$AIC_M = -2 \ell(\hat{\theta}) + 2p \quad \dots\dots\dots (3.1)$$

ここで、 $\ell(\hat{\theta})$  はパラメータのベクトル  $\theta$  の対数尤度であり、 $\hat{\theta}$  のとき最大となる。また  $p$  は  $\theta$  の次元 (自由度) である。AR モデルの場合、尤度関数  $L(\theta)$  は、

$$L(\theta) = \prod_{k=1}^N (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{\varepsilon_k^2}{2\sigma^2}\right) \quad \dots\dots\dots (3.2)$$

ただし、 $\sigma^2$  は白色雑音の分散  
 $\varepsilon_k$  は残差系列 ( $k = 1, \dots, N$ )

と近似される。これより対数尤度  $\ell(\theta)$  は、

$$\ell(\theta) = -\frac{N}{2} (1 + \log 2\pi) - \frac{N}{2} \log \sigma_M^2 \quad \dots\dots\dots (3.3)$$

ただし、 $\sigma_M^2$  は残差共分散

となる。第 1 項目は次数  $M$  に関しては定数であるから省略して、Yule-Walker 法の場合は

AICとして下式を得た。

$$AIC_M = N \log \sigma_M^2 + 2(M+1) \quad \dots\dots\dots (3.4)$$

Burg, Marpleおよび北川・赤池法の場合、データ長は次数Mまたは最大次数 $M_{max}$ に依存するため(2.62)または(2.94)式となった。しかし、(3.3)式の第1項もサンプル数Nが小さいときは無視できなくなる。第1項も含めた場合、Burg, MarpleのAIC(2.62)式は下式となる。

$$AIC_M = (N-M) \{ \log \sigma_M^2 + 1 + \log 2\pi \} + 2(M+1) \quad \dots\dots\dots (3.5)$$

(M=0, 1, ...)

また、北川・赤池法の場合(2.94)式は下式となる。

$$AIC_M = (N-M) \{ \log \sigma_m^2 + 1 + \log 2\pi \} + 2(m+1) \quad \dots\dots\dots (3.6)$$

(m=0, 1, ..., M)

北川・赤池法の場合(N-M)はMが最大次数に設定されるため定数となる様にみえるが、前節で述べた様に同一のデータ長NについてMAICEを適用するときは最大次数Mを逐次変えて計算する必要があることから、結局は定数としてはあつかえないことになる。

以上をまとめると、Yule-Walker法以外は(3.3)式の第1項は無視できないことがわかる。

(2) 数値計算上の問題

AIC計算におけるもう一つの問題は数値計算上の精度の問題である。データ $\{x_1, \dots, x_N\}$ についてフィッティングした場合と、同一のデータを適当な倍率 $\alpha$ で増幅したときの $\{\alpha x_1, \dots, \alpha x_N\}$ についての場合を考える。両データのARモデルのAICは

$$AIC_M = N \log \sigma_M^2 + 2(M+1) \quad \dots\dots\dots (3.7)$$

および

$$\begin{aligned} AIC'_M &= N \log \sigma_M'^2 + 2(M+1) \\ &= N \log (\alpha \cdot \sigma_M)^2 + 2(M+1) \\ &= N \log \sigma_M^2 + 2(M+1) + N \log \alpha^2 \\ &= AIC_M + N \log \alpha^2 \quad \dots\dots\dots (3.8) \end{aligned}$$

となる。理論的には、両者の違いは単にバイアス成分だけであり、MAICEの適用には何らさしつかえはない。しかし、数値計算上ではNが大きい場合で係数 $\alpha$ が比較的大きい(または小さい)場合、または $AIC_M$ の次数による差異がほとんどない場合、計算精度が失われてしまう危険がある。このバイアス成分の大きさに無関係に常に正しくMAICEを適用するためには、フィッティングされたモデルの残差共分散 $\sigma_M^2$ の推定値の有効桁がかなり大きくなければならぬが、現実的には計算処理を倍精度化(FORTRAN言語での)してもほとんど改善されない。

一番効果的な精度改善法は、サンプル・データに対して適当な係数 $\alpha$ を選択して、(3.8)式の第1項 $N \log \sigma_M'^2$ と第2項 $2(M+1)$ の絶対値を近づける方法である。具体的には、データをデータの持つ標準偏差 $\sigma_x$ で規準化してやると、 $N=200, M=0\sim 100$ で残差共分散 $\sigma_M^2$ が0.1~0.01付近であったとすると第1項は-460~-920付近となり、第2項は2~200付近となる。絶対値の最大桁数がだいたい一致するから数値計算上は十分な精度が期待できる。

(3) 各方法の再計算

ケースI~IIIのフィッティングでは数値計算上の問題は無視できるが、(1)の尤度計算のいわ

ゆる定数項は無視できないため、AICの再計算を試みた。

この結果、ケースⅡのFig. 3.5では、 $N=60$ の場合Burg法は変更はなかったが、Marple法は $M=13$ のとき最小となった。 $N=120$ のときは、Burgは $M=16$ 、Marpleは $M=17$ となった。 $N=4000$ の場合、Burgは $M=44$ 、Marpleは $M=91$ となった。

また、ケースⅢ  $N=4000$ の場合 (Fig. 3.8), Marple法は変更はなかったが、Burg法は $M=67$ となった。さらに、 $N=120$  (Fig. 3.10)の場合、Burg法は変更はなかったが、Marple法は $M=21$ となった。

ケースⅠのFig. 3.14では各方法とも変更はなかった。

北川・赤池法の最大次数 $M_{\max}$ を逐次変えて計算したAIC (Fig. 3.24)では、AICの傾斜は変わったがMAICE次数に変更はなかった。

以上をまとめると、定数項として無視できない項を含めたAICによる次数決定では、Burg法、Marple法ともにその残差共分散 $\sigma_M^2$ が他の方法に比べて小さく推定される様な場合でも矛盾の無いMAICE次数が得られることがわかった。

### 3.7 各方法の計算時間

最後に各方法のフィッティングに要する時間の測定結果を示す。使用した計算機はSystems社製のミニコンSEL-3200  $77/62$ であり、使用言語はFORTRAN 77+である。各モデル次数に対応したARパラメータのディスク装置への転送時間が各方法とも一率に含まれている。計算処理の絶対時間は計算機により異なるが、各方法の相対的な時間はどの計算機でも再現性があると思われる。

Fig. 3.25はサンプル数 $N$ に対する処理時間の測定値である。(a)に北川・赤池法を、(b)にMarple法の結果を示す。北川・赤池法の場合、各次数においてサンプル数 $N$ に比例した時間を費すことがわかる。一方、Marpleの場合ほぼ $(N)^{\frac{1}{2}}$ に比例しかなり有利であることがわかる。

Fig. 3.26はモデル次数 $M$ に対する処理時間の測定値である。(a)北川・赤池法、(b) Marple法、(c) Yule-Walker法、(d) Burg法である。北川・赤池法はモデル次数の2乗 $M^2$ に比例すること、他の方法は $M$ に比例することが示されている。特にMarpleの場合、Yule-Walkerと並んで一番早く、Burgの場合はほぼ2倍の処理時間を要する。

以上から、Marpleの方法は処理時間の点で有利なこと、北川・赤池法はサンプル数 $N$ やモデル次数 $M$ の大きいときはかなりの時間を要することがわかった。

Table 3.1 Test data for simulation study

	Signal	Sampling interval $\Delta t$	Number of data samples in a period ( 0.666 sec )
Case-I	$A \sin(2\pi f_0 t + \phi_0)$ ; $A=2.5, f_0=1.5$ Hz	0.061 sec	12
Case-II	$A \sin(2\pi f_0 t + \phi_0) + w(t)$ ; $A=2.5, f_0=1.5$ Hz	0.061 sec	12
Case-III	$A \sin(2\pi f_0 t + \phi_0) + w(t)$ ; $A=2.5, f_0=1.5$ Hz	0.006 sec	111

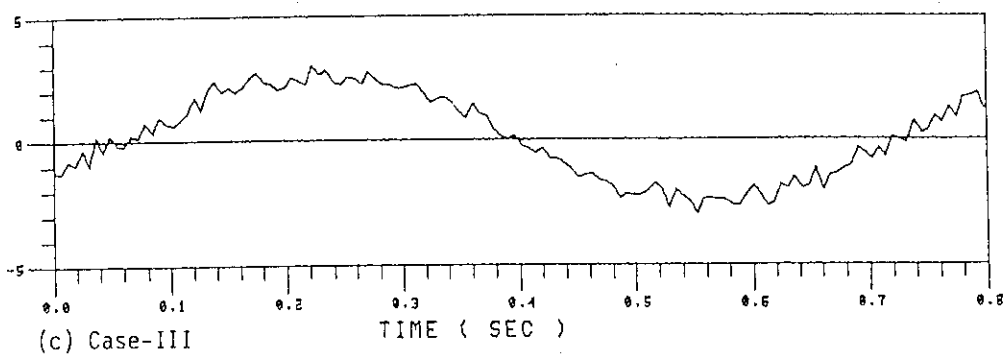
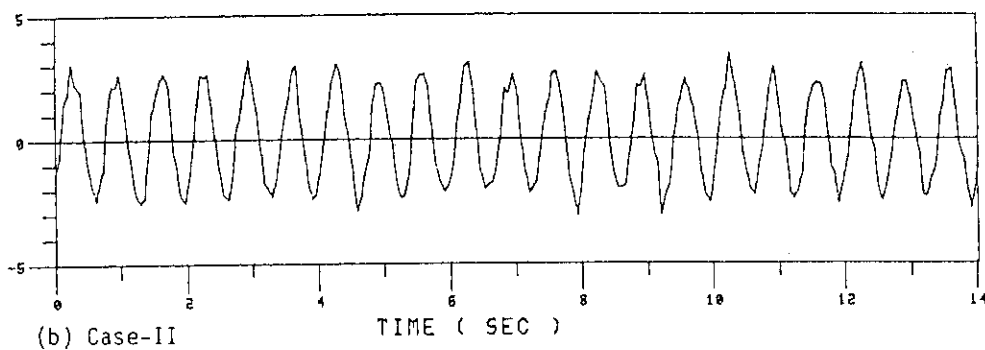
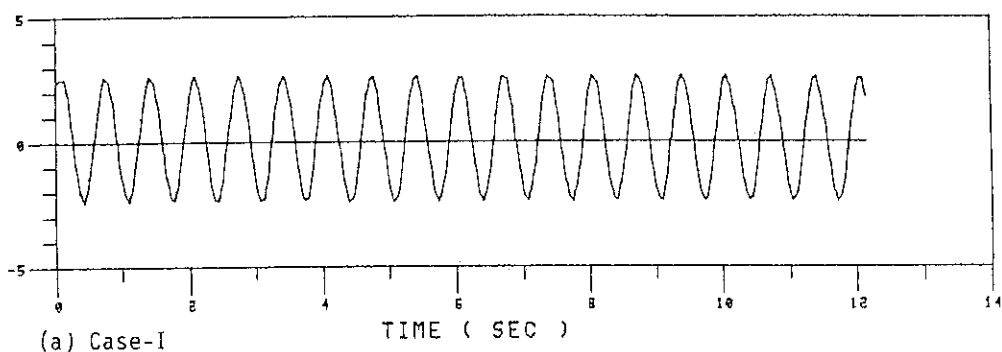


Fig. 3.1 Examples of the test data



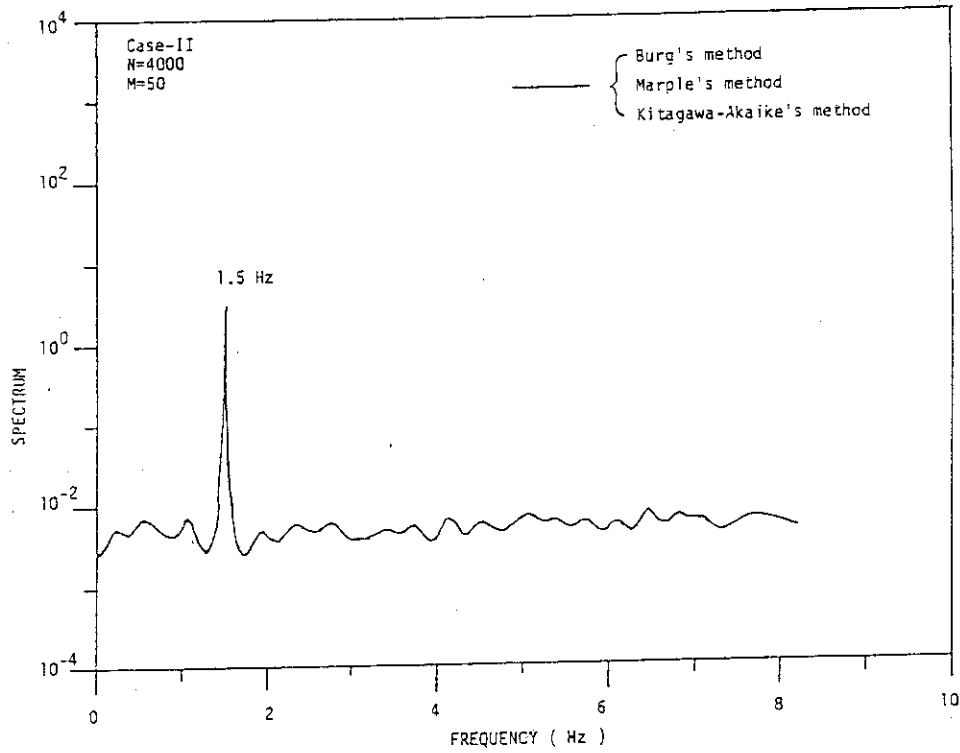


Fig. 3.2 Power spectra estimated by three different methods with model order M=50 using a 4000-points data set taken from the Case-II data

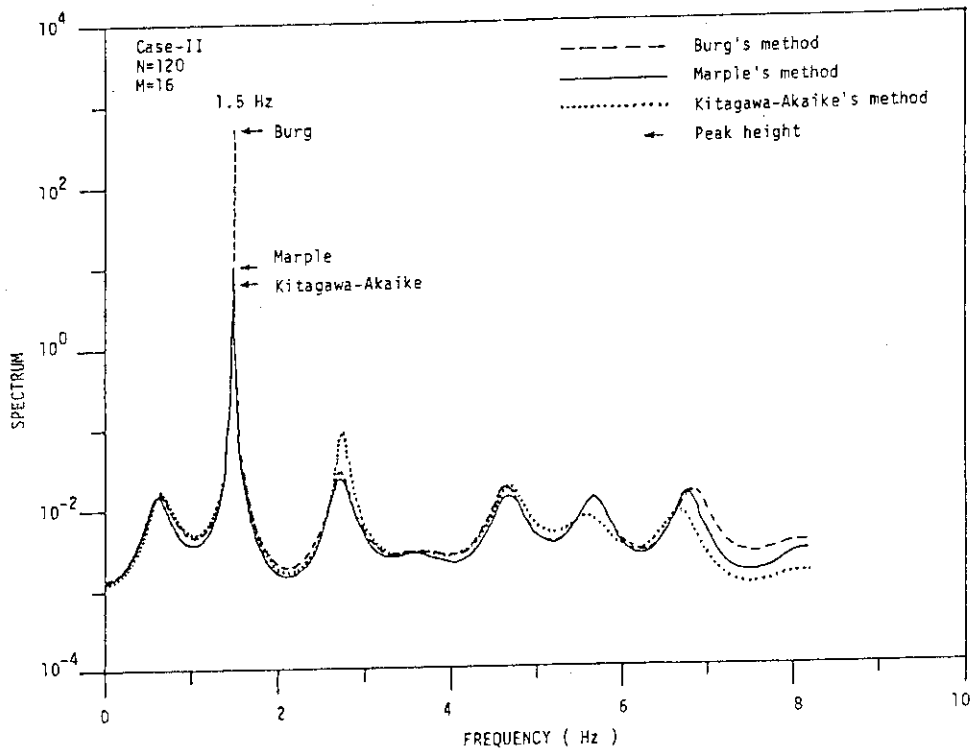


Fig. 3.3 Power spectra estimated by three different methods with model order M=16 using a 120-points data set taken from the Case-II data

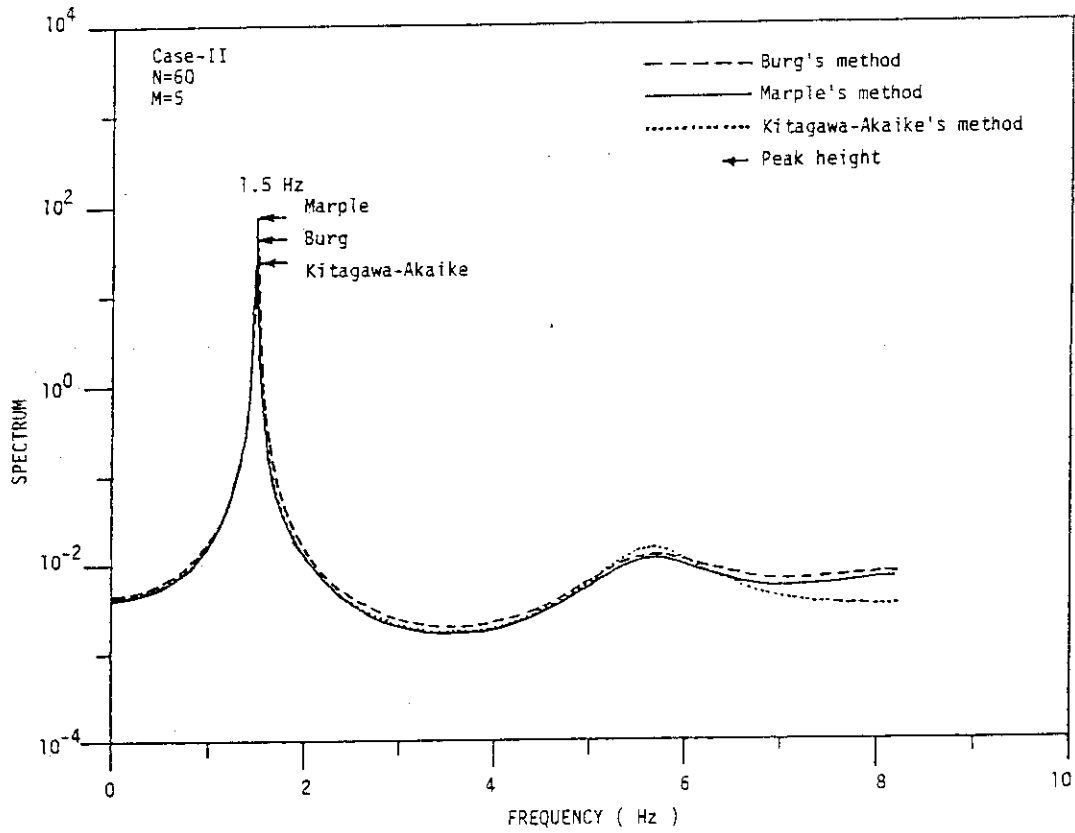


Fig. 3.4 Power spectra estimated by three different methods with model order  $M=5$  using a 60-points data set taken from the Case-II data

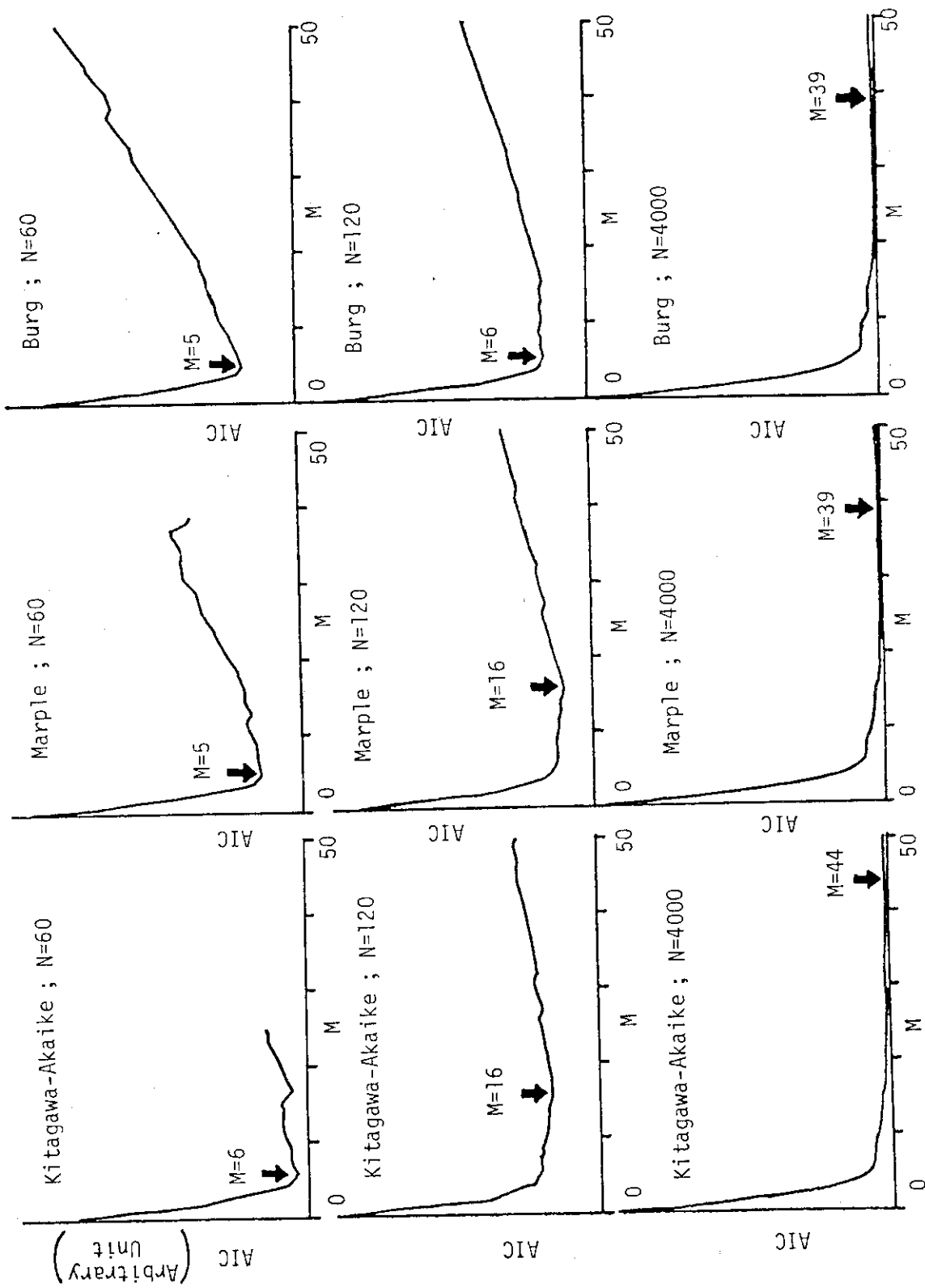


Fig.3.5 AIC functions for the AR models fitted by three different methods for the 60, 120 and 4000-points data sets taken from the Case-II data

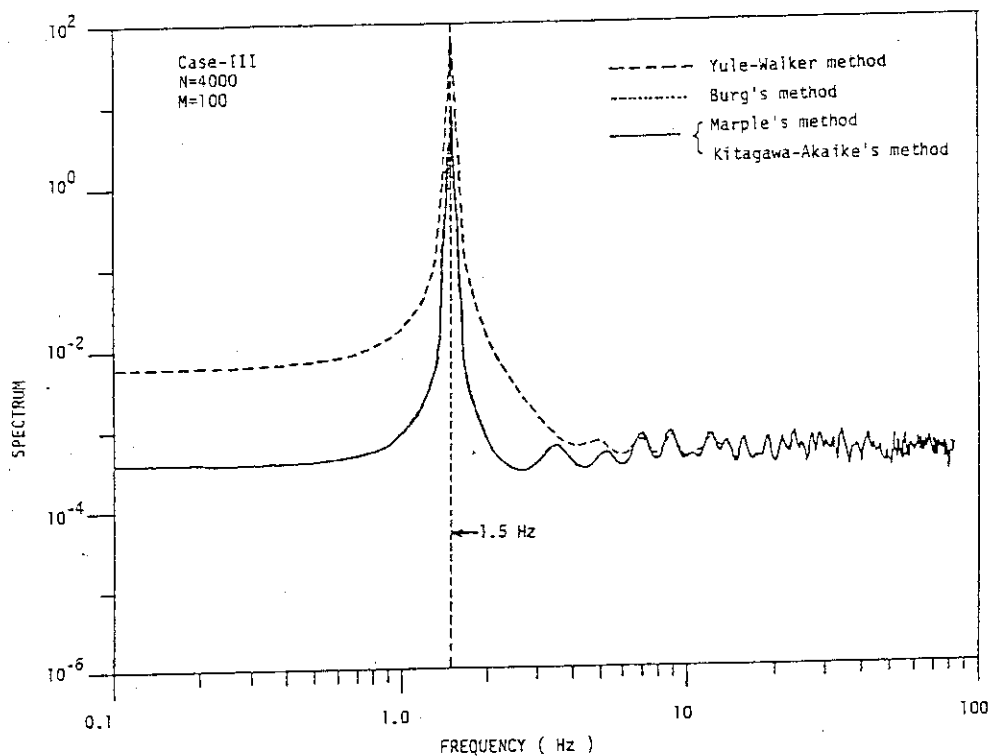


Fig. 3.6 Power spectra estimated by four different methods with model order  $M=100$  using a 4000-points data set taken from the Case-III data

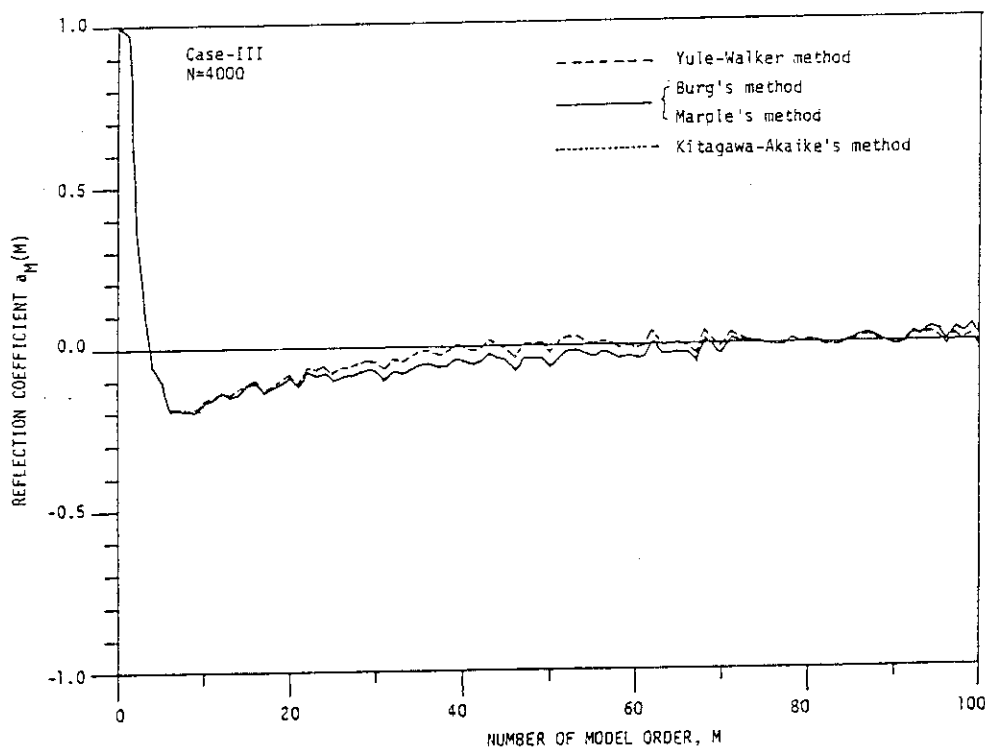


Fig. 3.7 Reflection coefficient sequences of the AR models fitted by four different methods using a 4000-points data set taken from the Case-III data

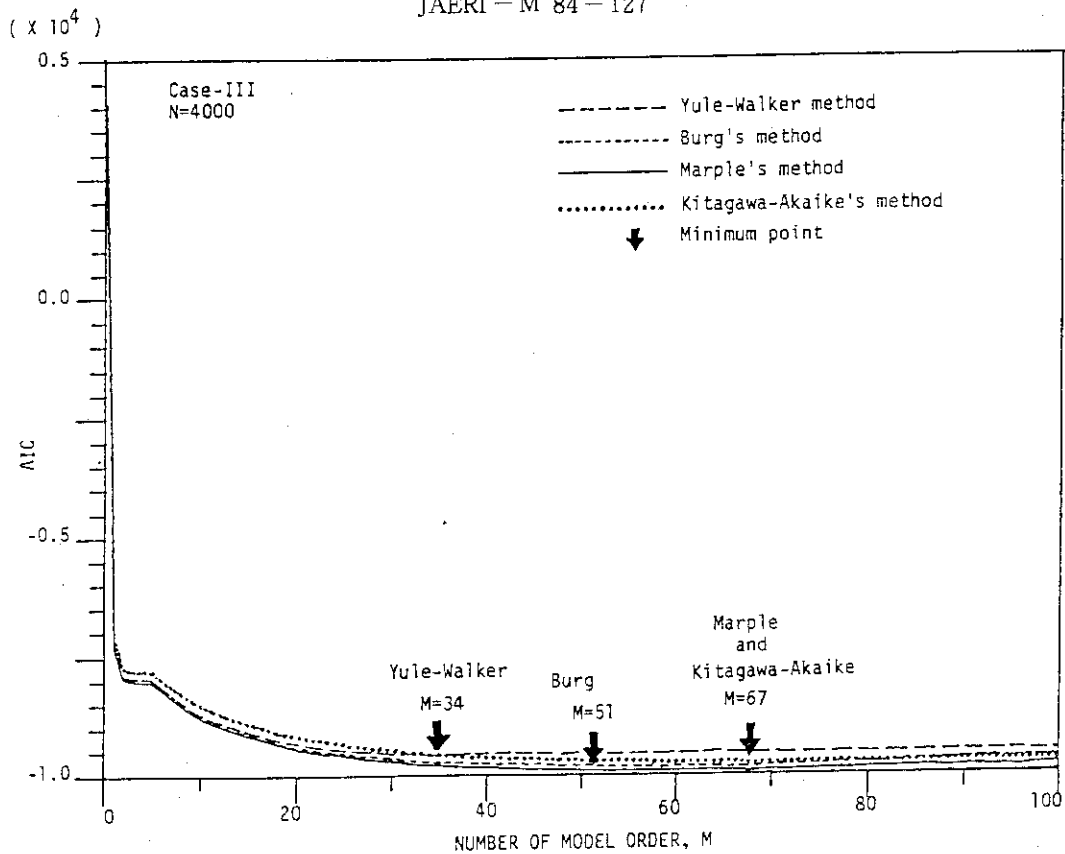


Fig. 3.8 AIC functions of the AR models fitted by four different methods for a 4000-points data set taken from the Case-III data

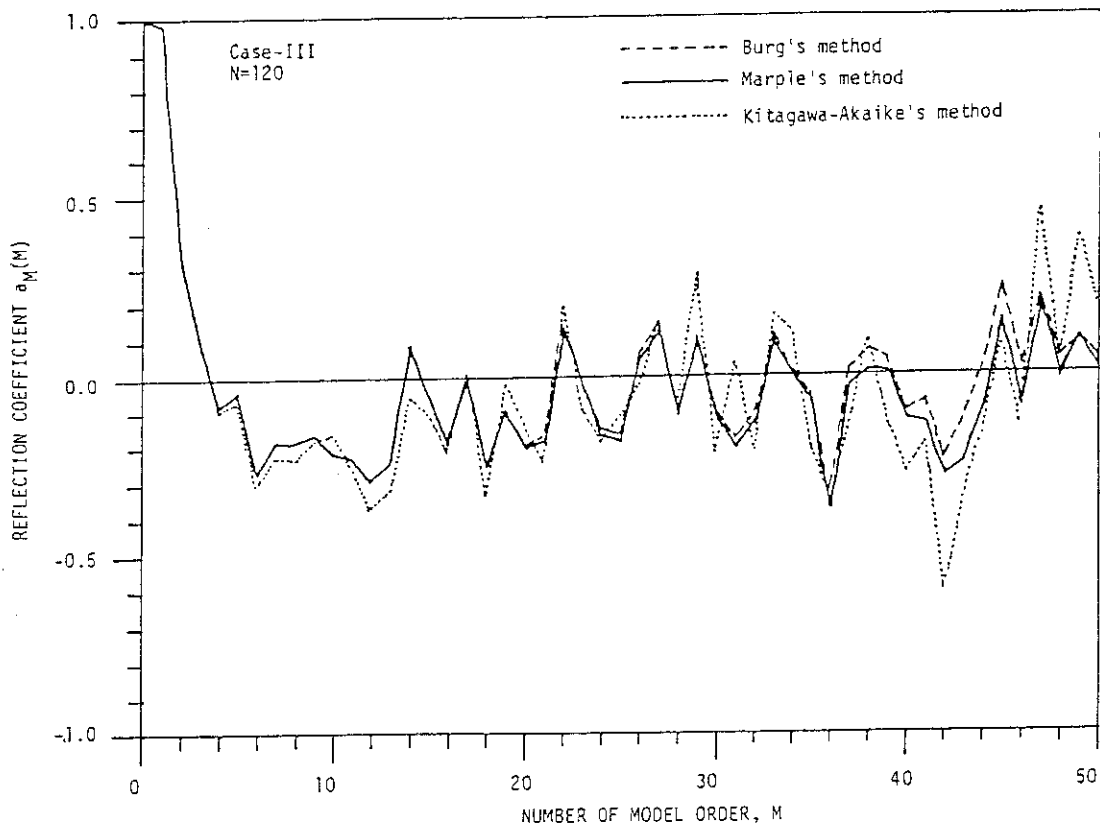


Fig. 3.9 Reflection coefficient sequences of the AR models fitted by three different methods for a 120-points data set taken from the Case-III data

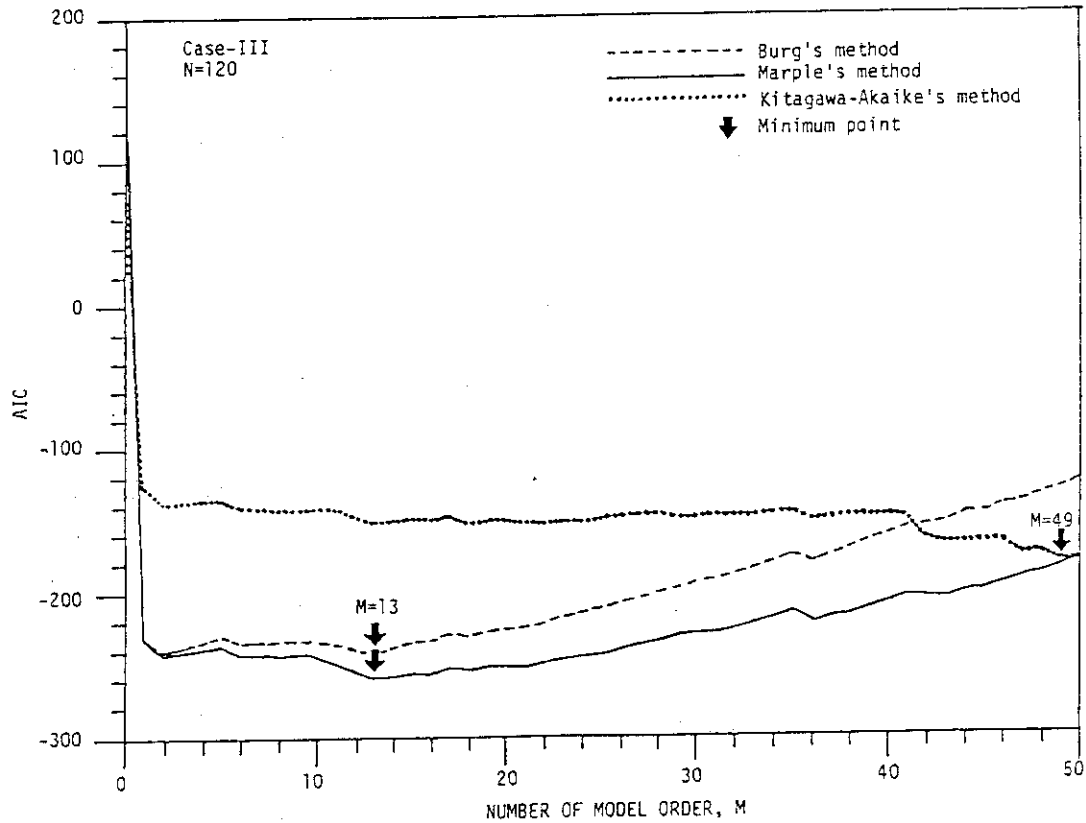


Fig. 3.10 AIC functions of the AR models fitted by three different methods for a 120-points data set taken from the Case-III data

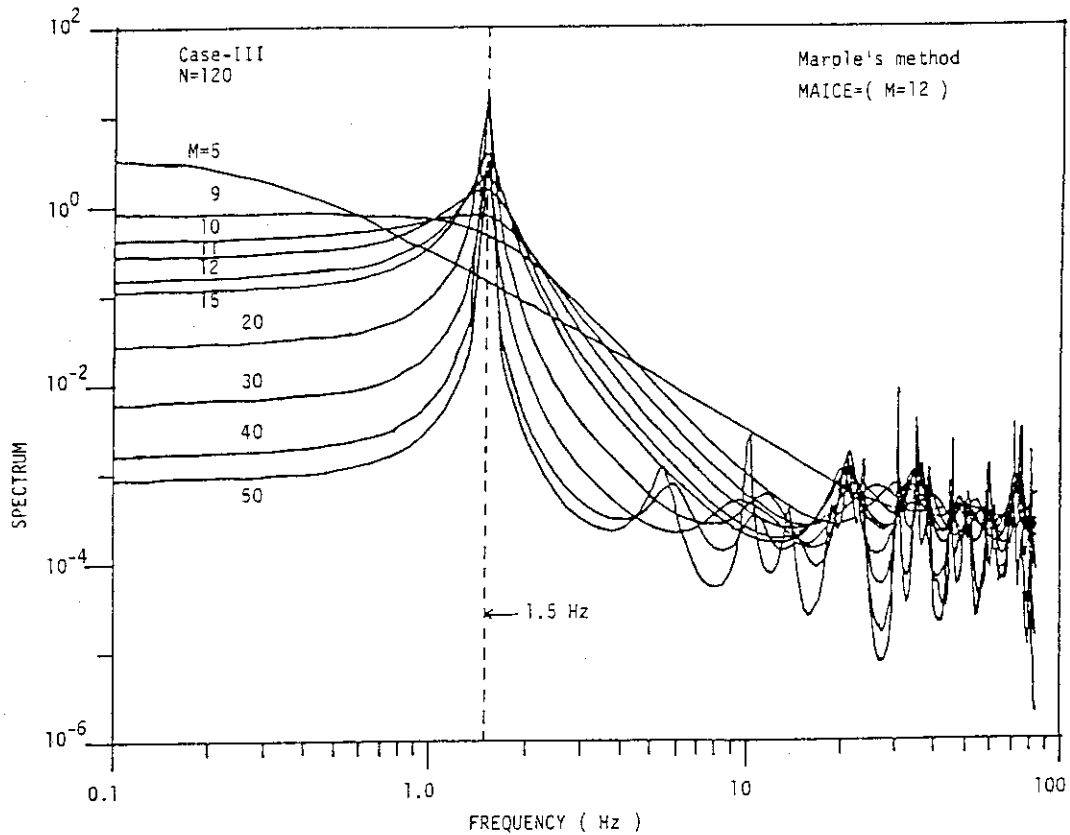


Fig. 3.11 Comparison of the power spectra estimated by the Marple's method with various model orders using a 120-points data set taken from the Case-III data

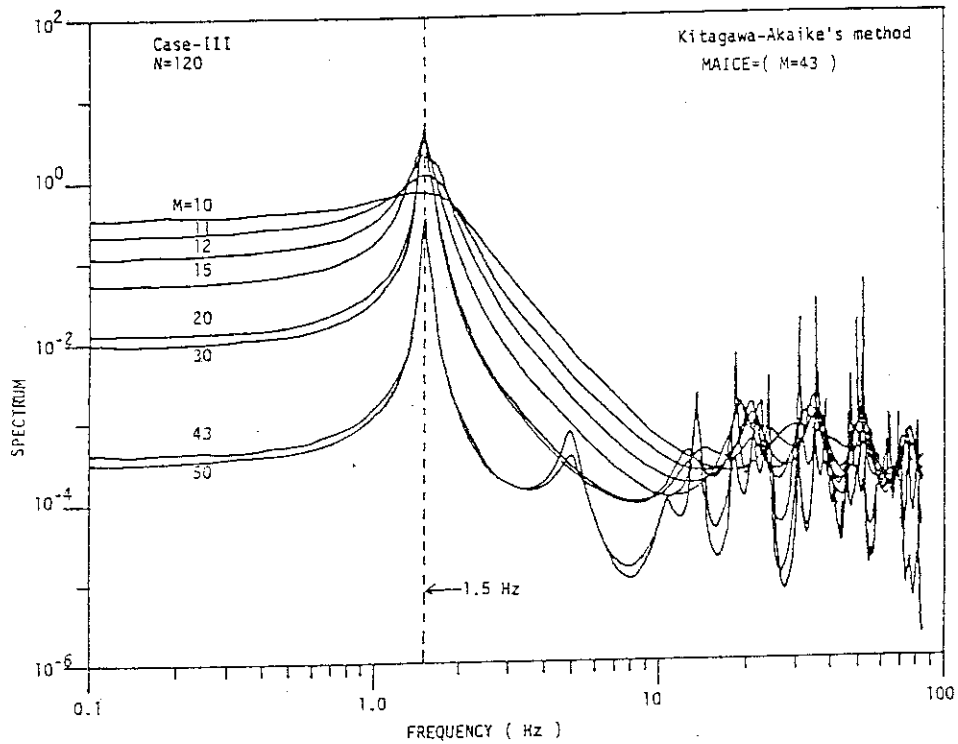


Fig. 3.12 Comparison of the power spectra estimated by the Kitagawa-Akaike's method with various model orders using a 120-points data set taken from the Case-III data

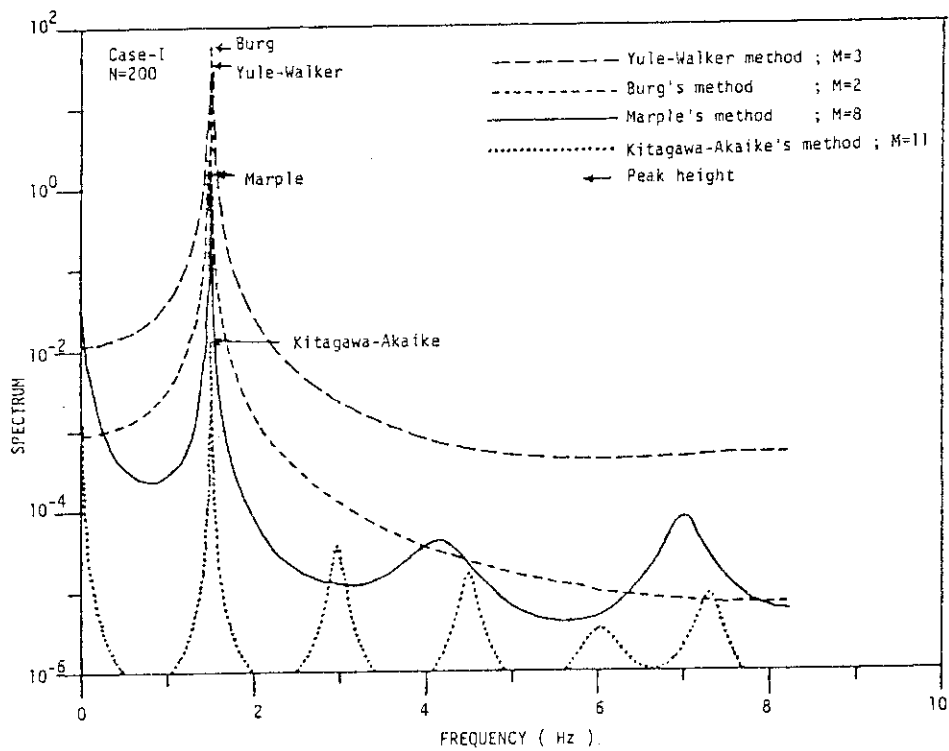


Fig. 3.13 Power spectra estimated by four different methods with the MAICE model orders for a 200-points data set taken from the Case-I data

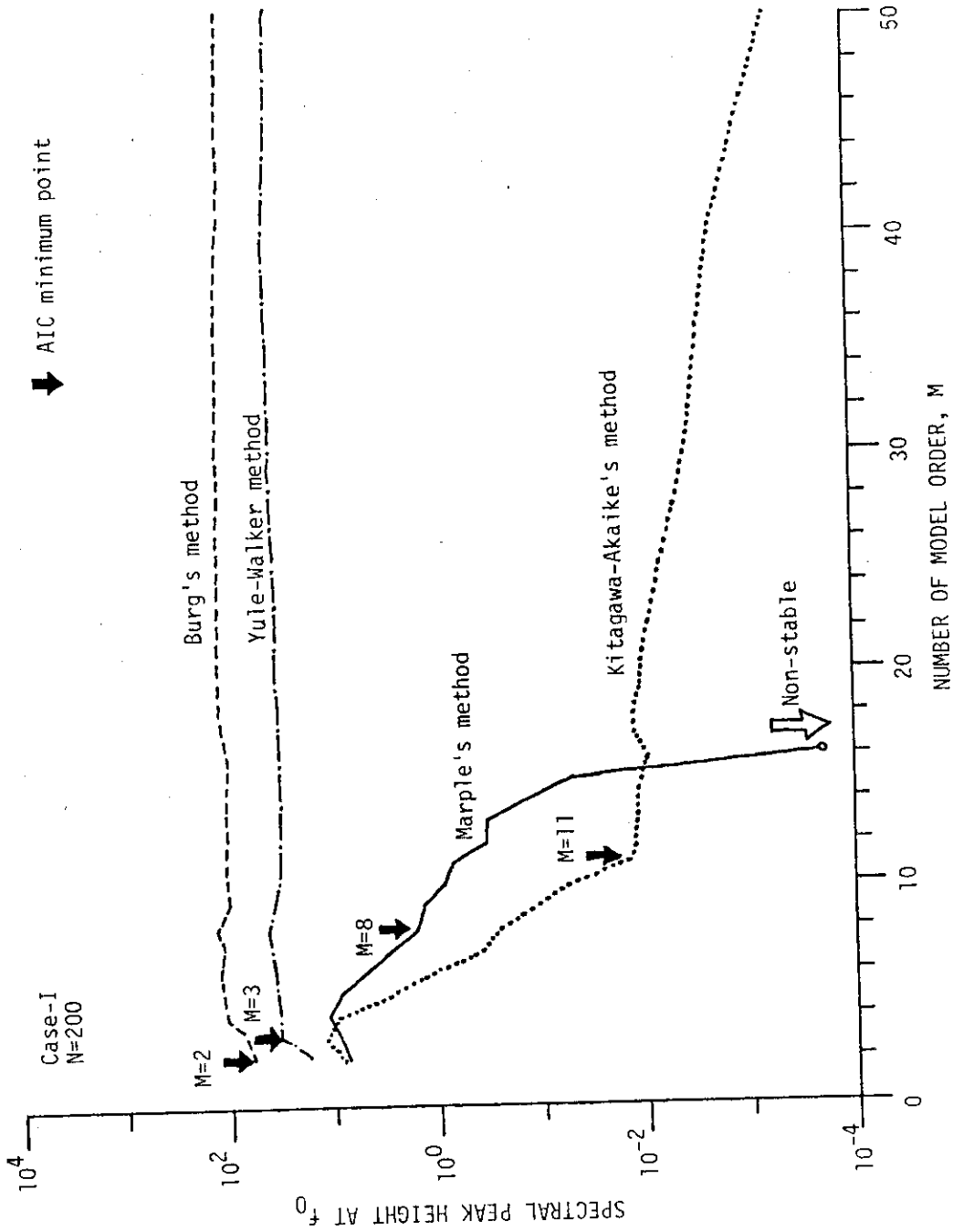


Fig. 3.14 Comparison of the spectral peak heights estimated by four different methods using a 200-points data set taken from the Case-I data



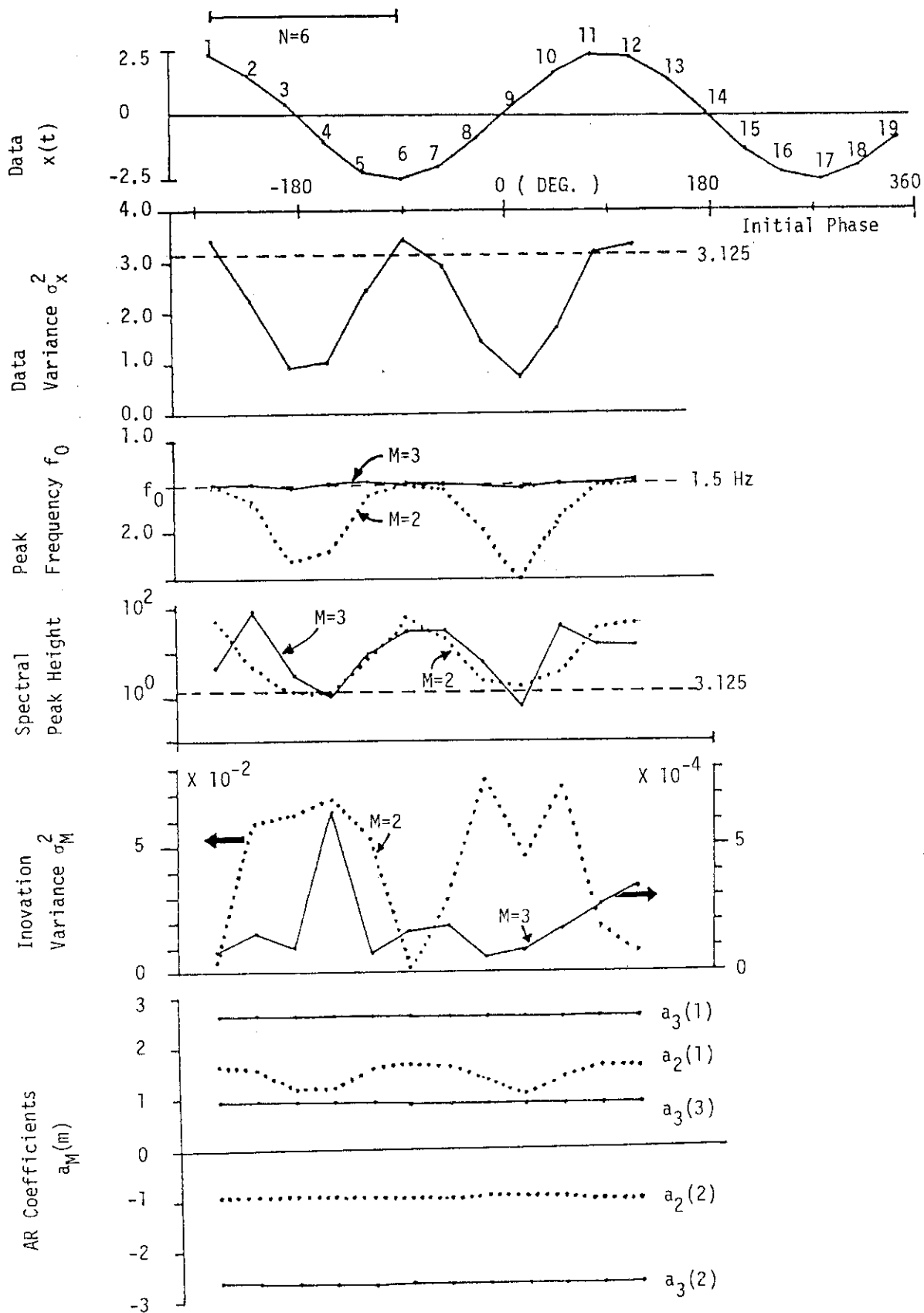


Fig. 3.15 The changes in the spectral peaks and the AR parameters depending on the initial phases of the sampled data in case of the fitting by the Marple's method

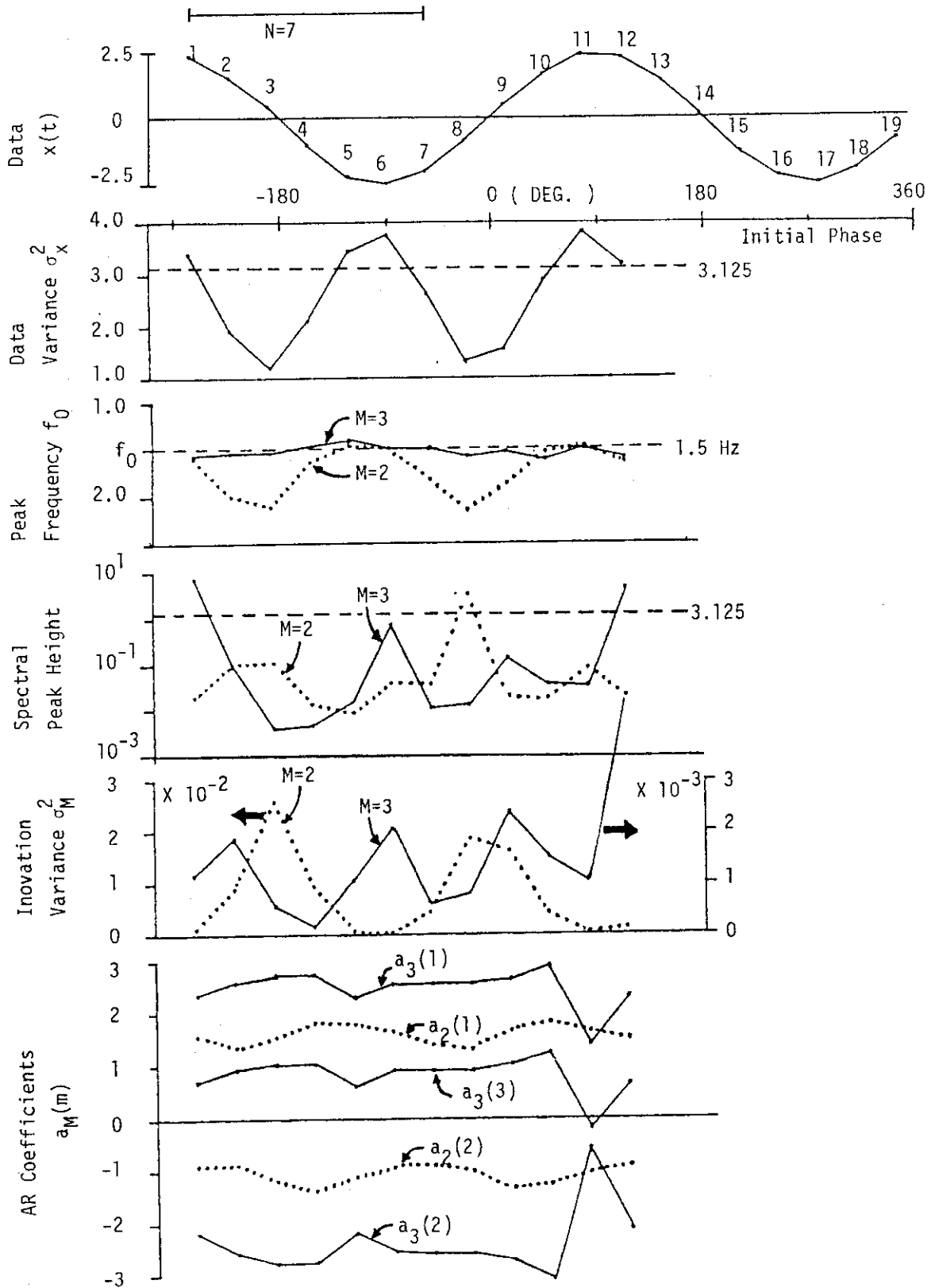


Fig. 3.16 The changes in the spectral peaks and the AR parameters depending on the initial phases of the sampled data in case of the fitting by the Kitagawa-Akaike's method

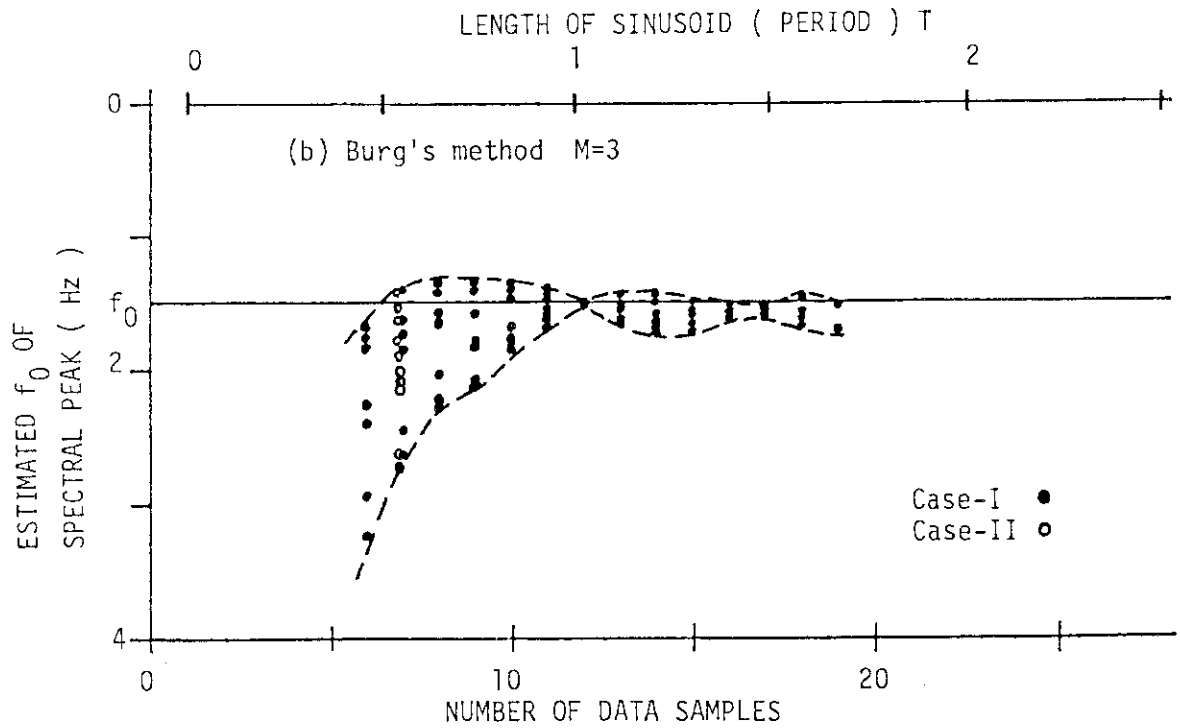
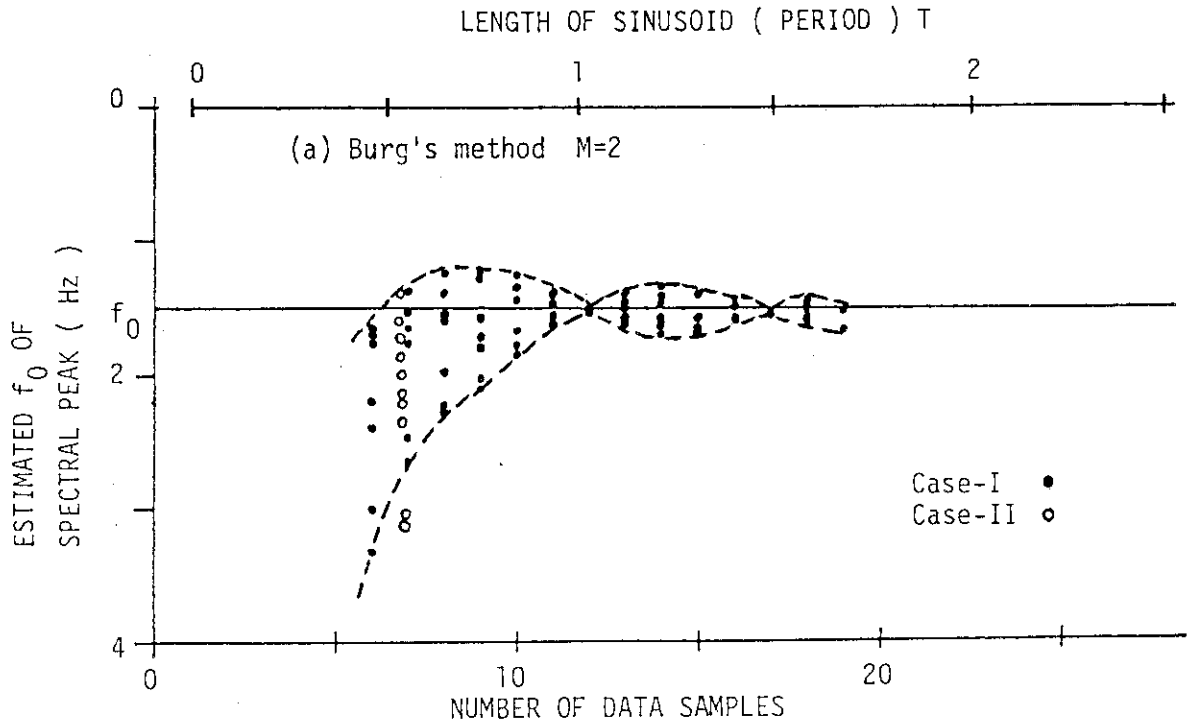


Fig. 3.17 The bias errors in the peak frequencies of the power spectra estimated by the Burg's method  
(a) Model order M=2 (b) M=3

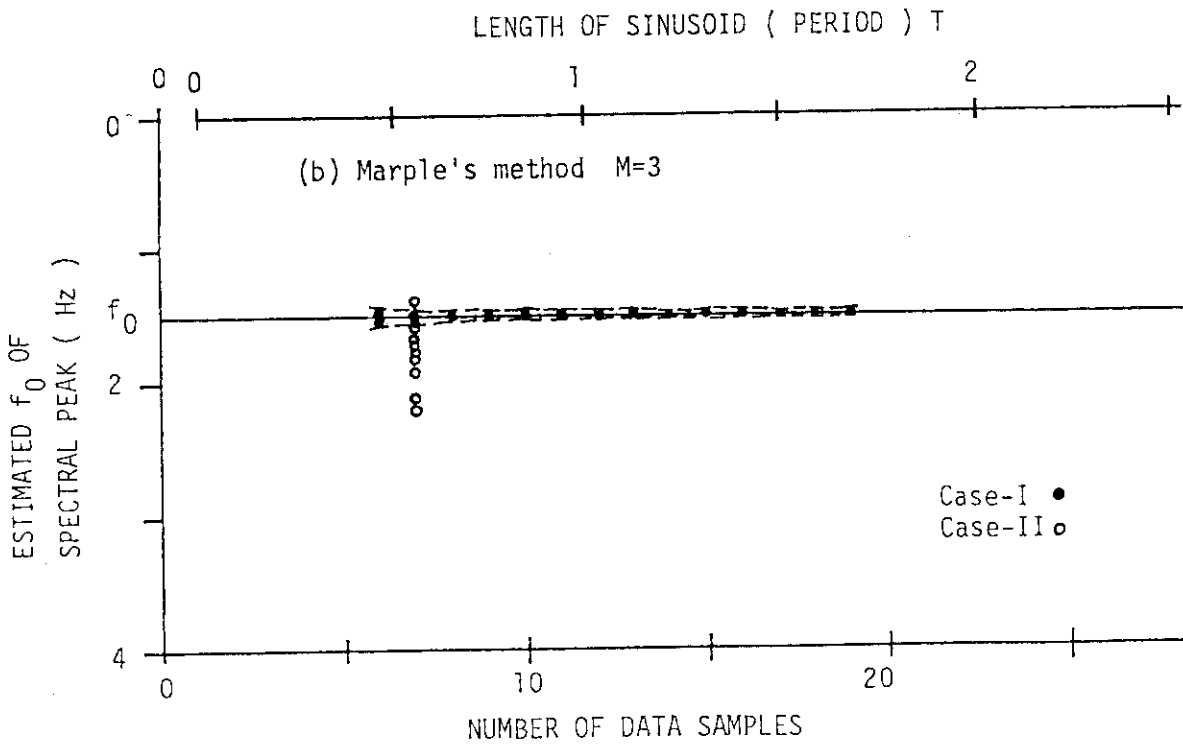
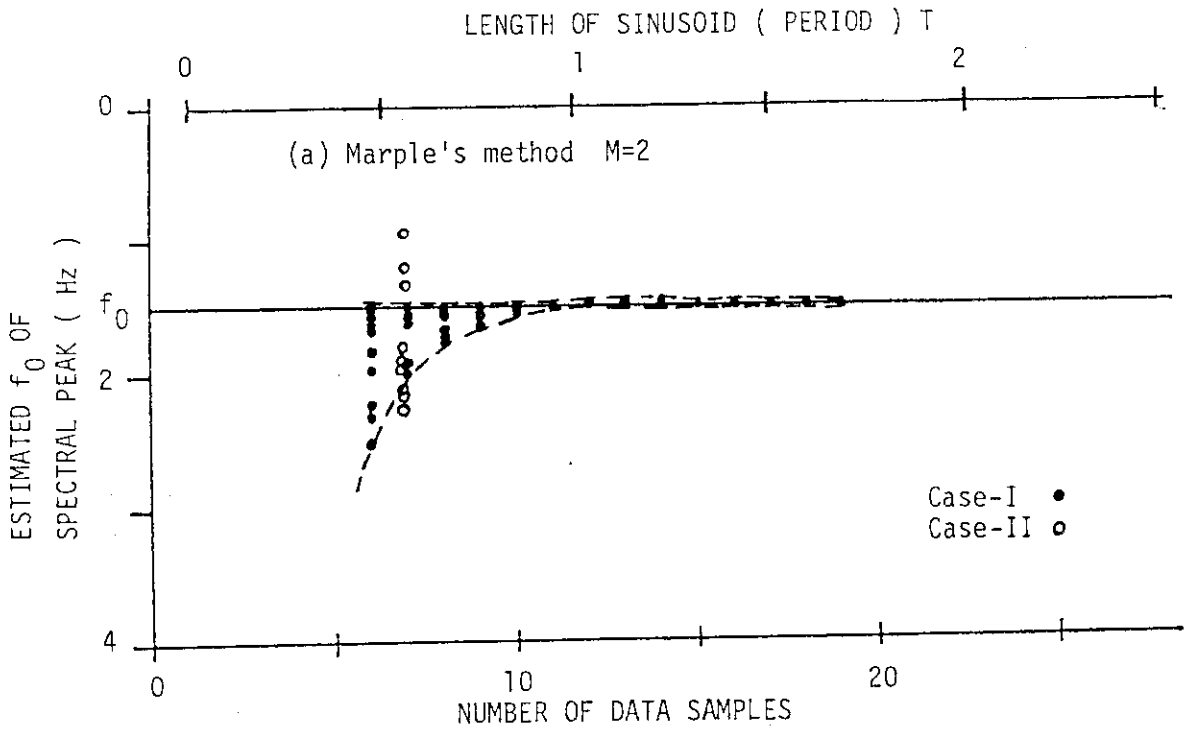


Fig. 3.18 The bias errors in the peak frequencies of the power spectra estimated by the Marple's method  
 (a) Model order  $M=2$  (b)  $M=3$

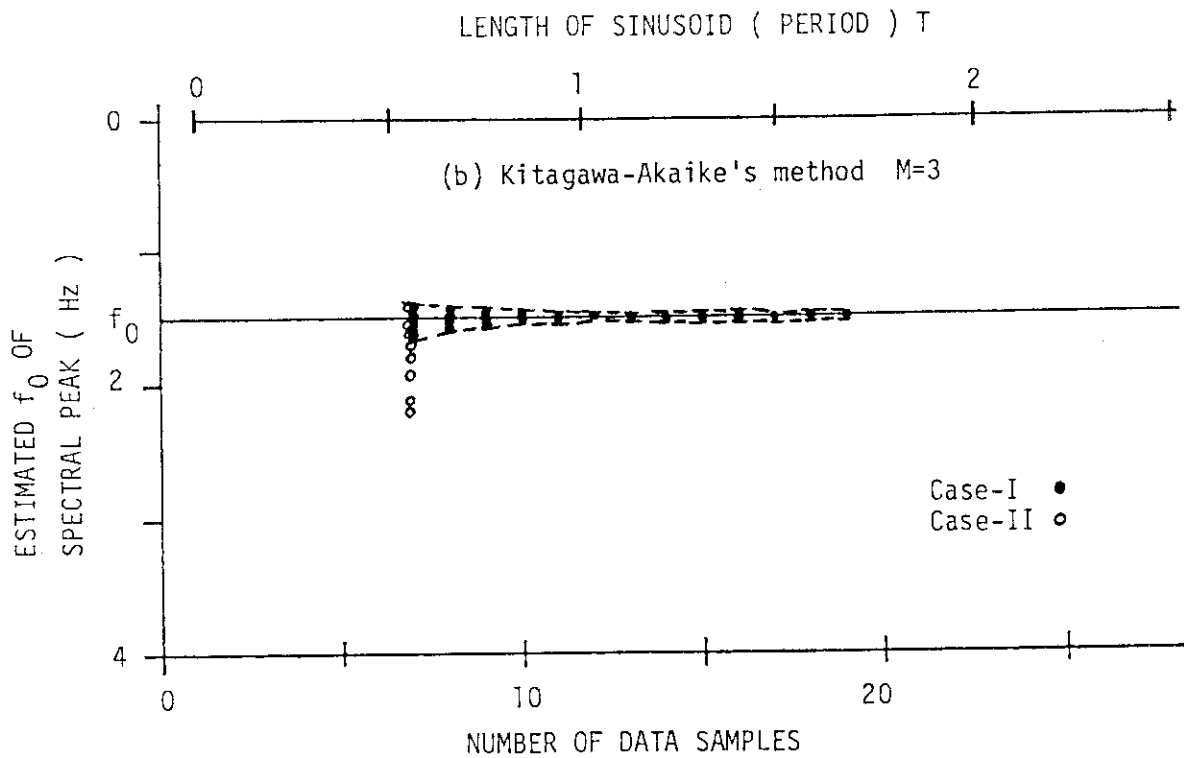
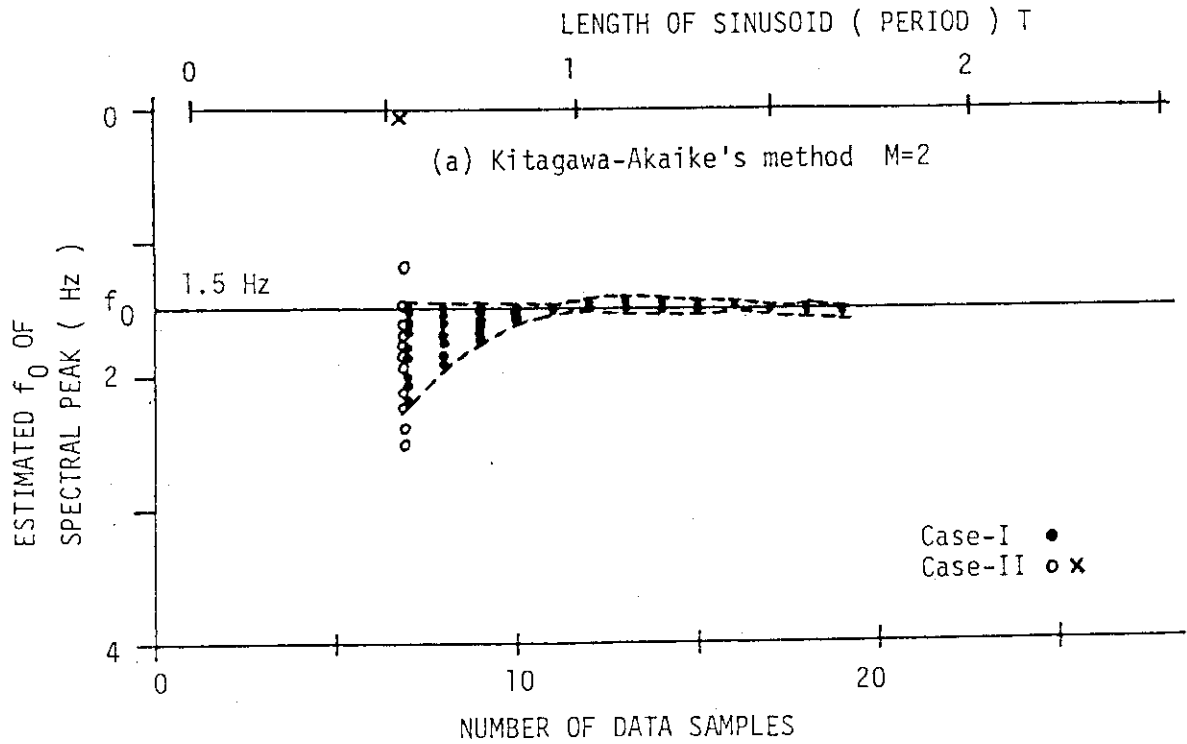


Fig. 3.19 The bias errors in the peak frequencies of the power spectra estimated by the Kitagawa-Akaike's method  
(a) Model order M=2 (b) M=3

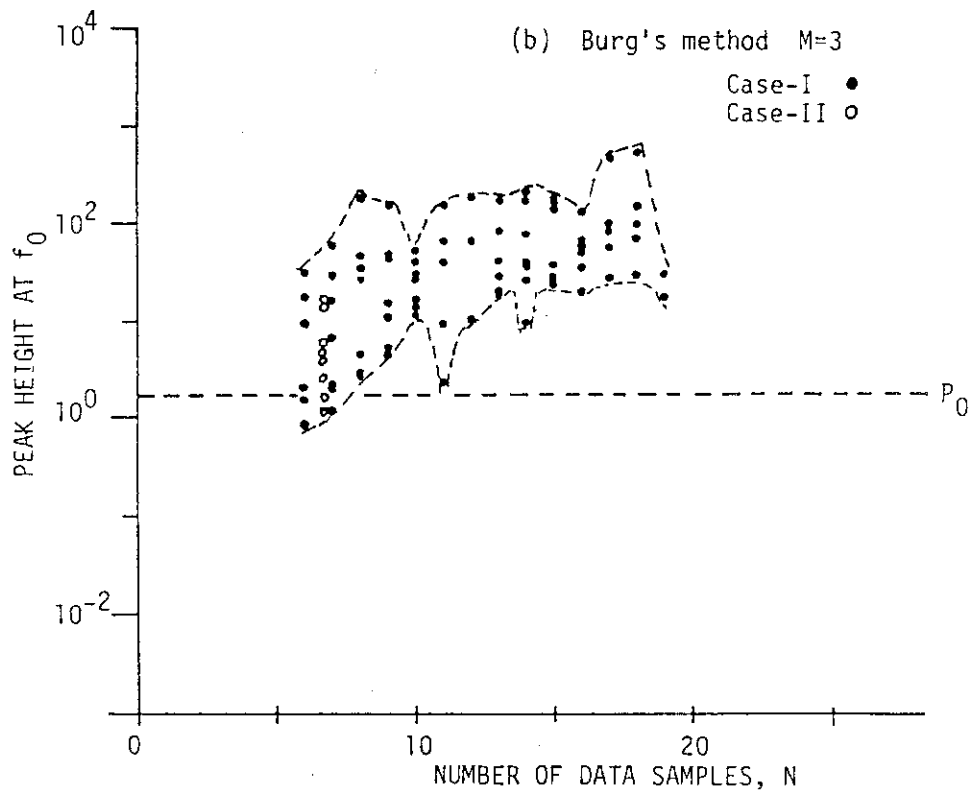
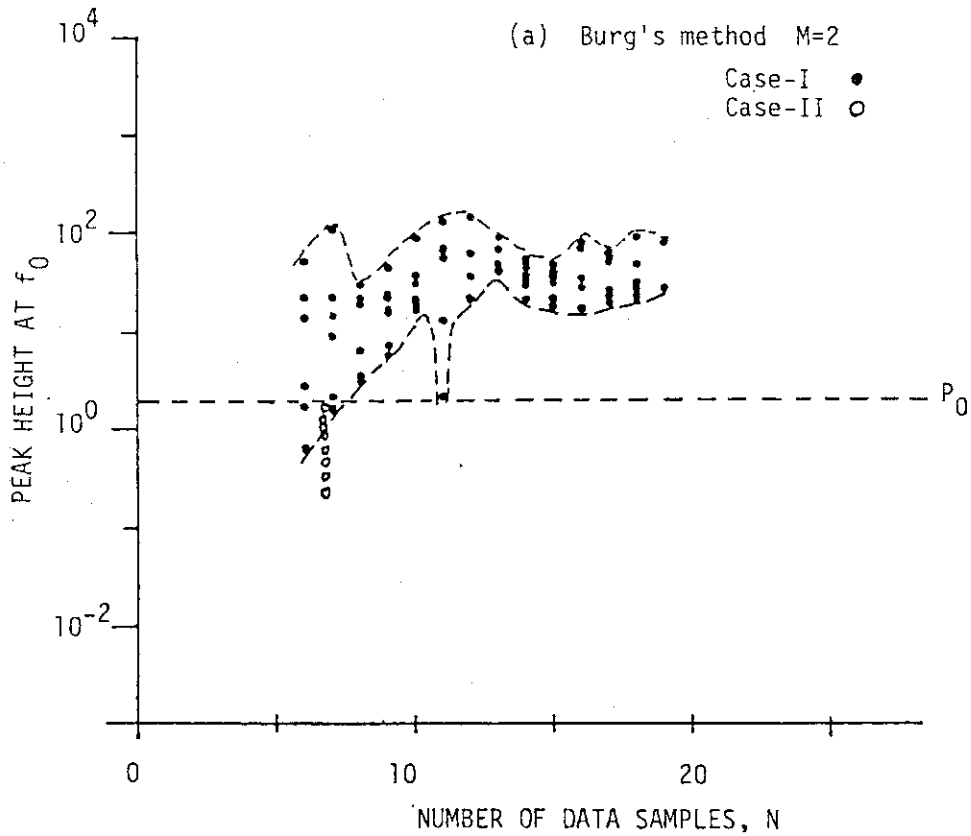


Fig. 3.20 The bias errors in the peak heights of the power spectra estimated by the Burg's method  
(a) Model order M=2 (b) M=3

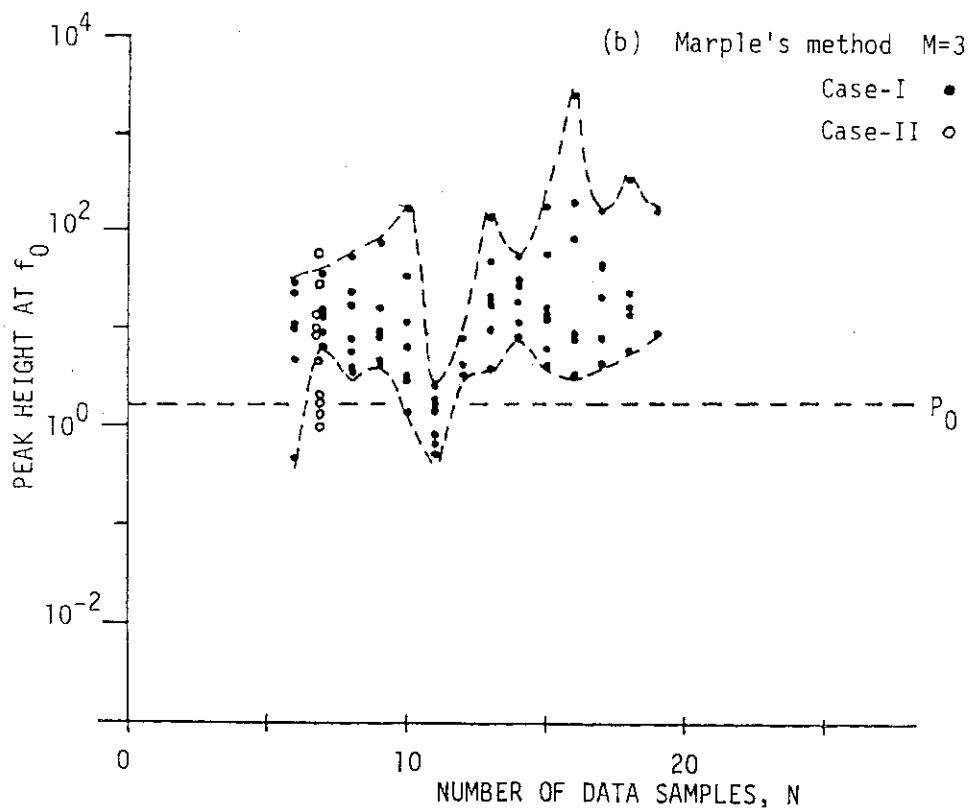
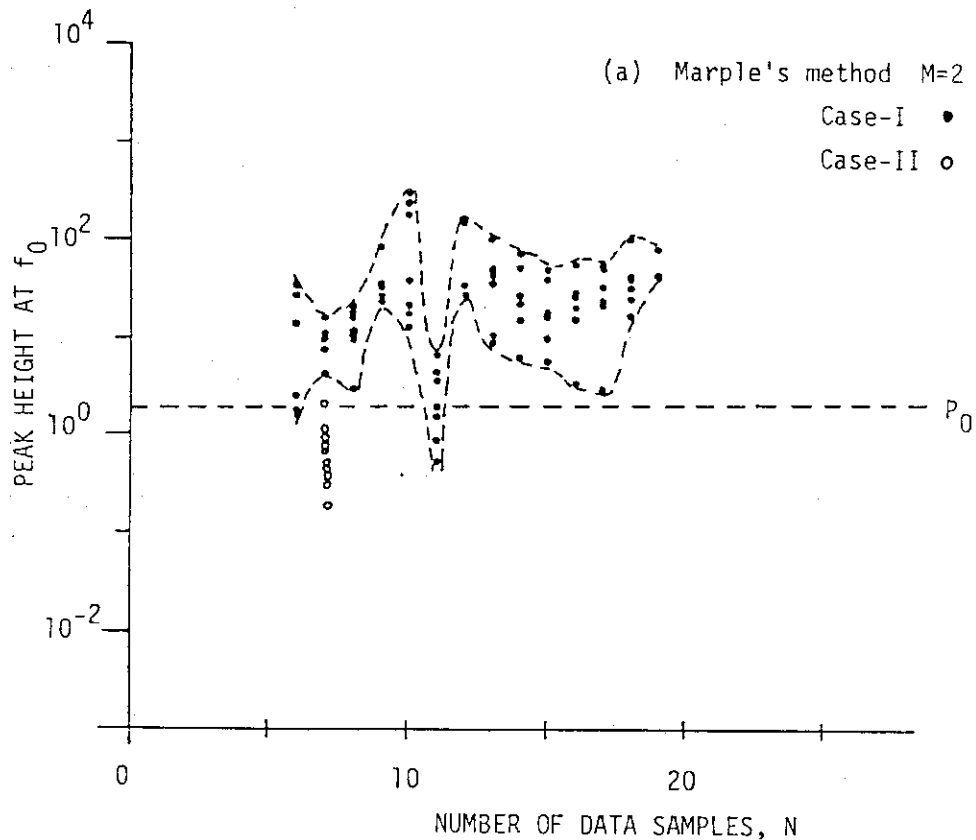


Fig. 3.21 The bias errors in the peak heights of the power spectra estimated by the Marple's method  
(a) Model order M=2 (b) M=3

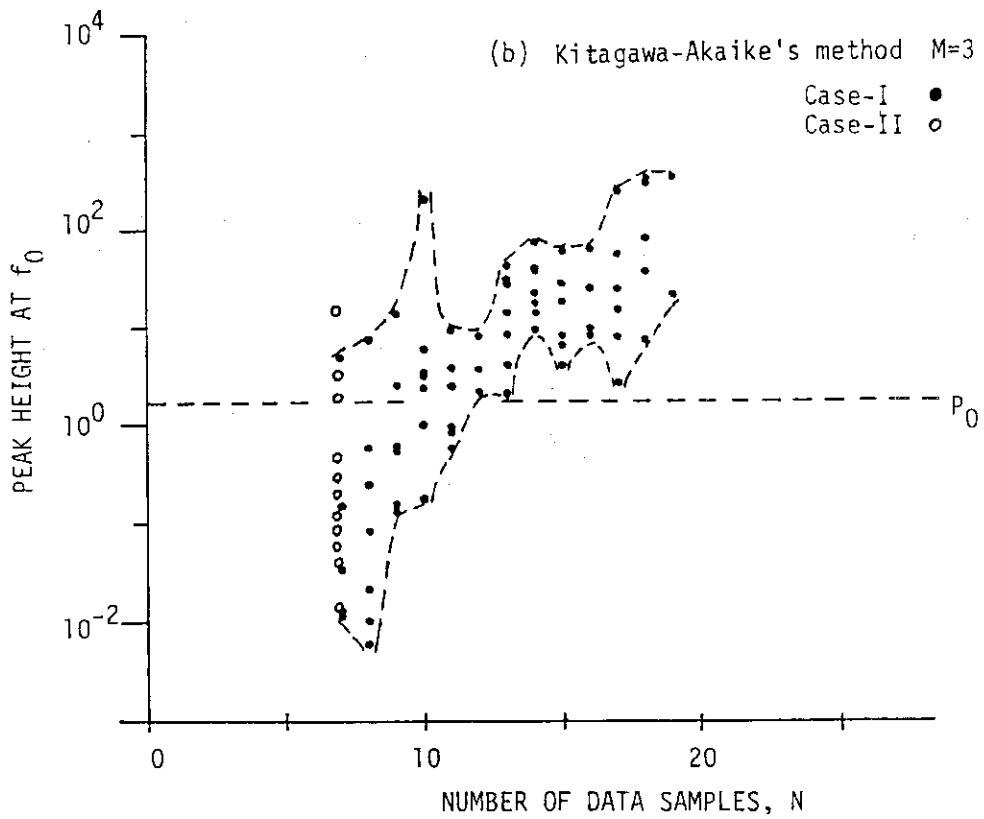
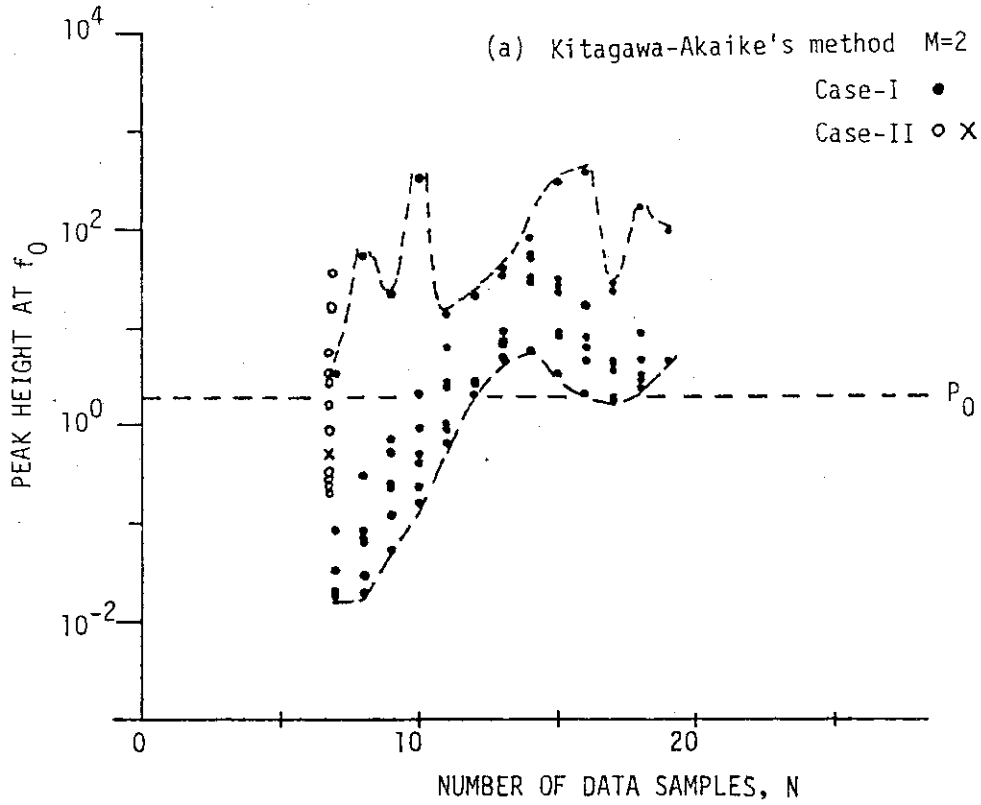


Fig. 3.22 The bias errors in the peak heights of the power spectra estimated by the Kitagawa-Akaike's method  
(a) Model order  $M=2$  (b)  $M=3$



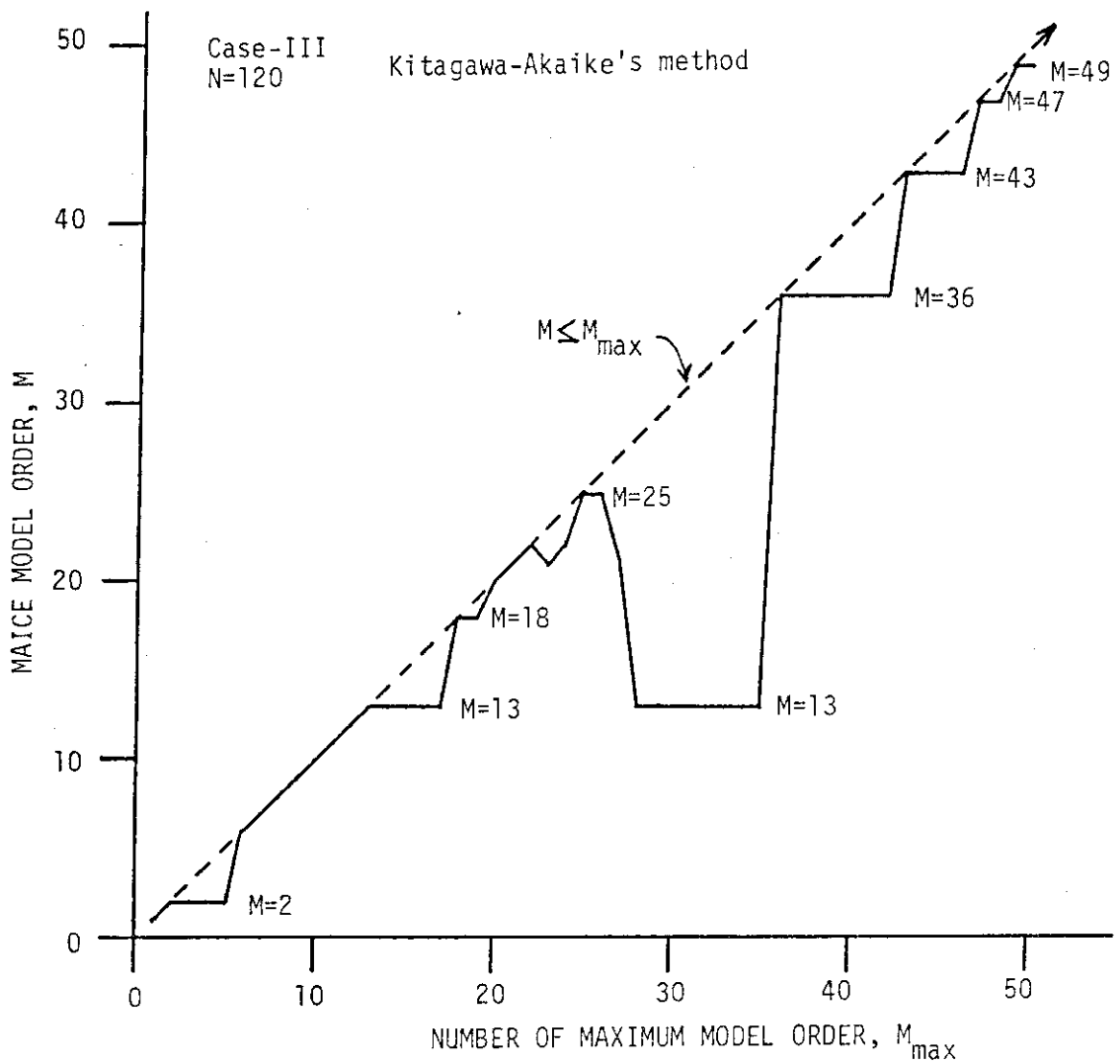


Fig. 3.23 The relationship between the MAICE model orders and the maximum model orders of the AR models fitted by the Kitagawa-Akaike's method using a 120-points data set taken from the Case-III data

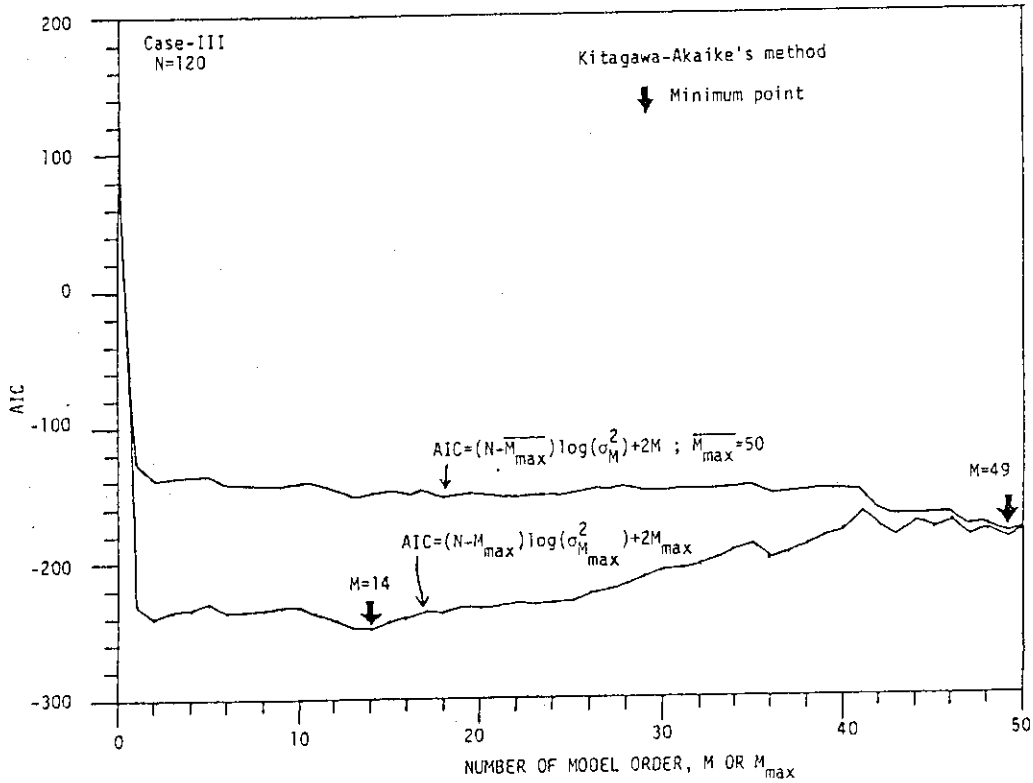


Fig. 3.24 Comparison of the AIC values for the AR models fitted by the Kitagawa-Akaike's method using a 120-points data set taken from the Case-III data : the cases with fixed and variable maximum model orders

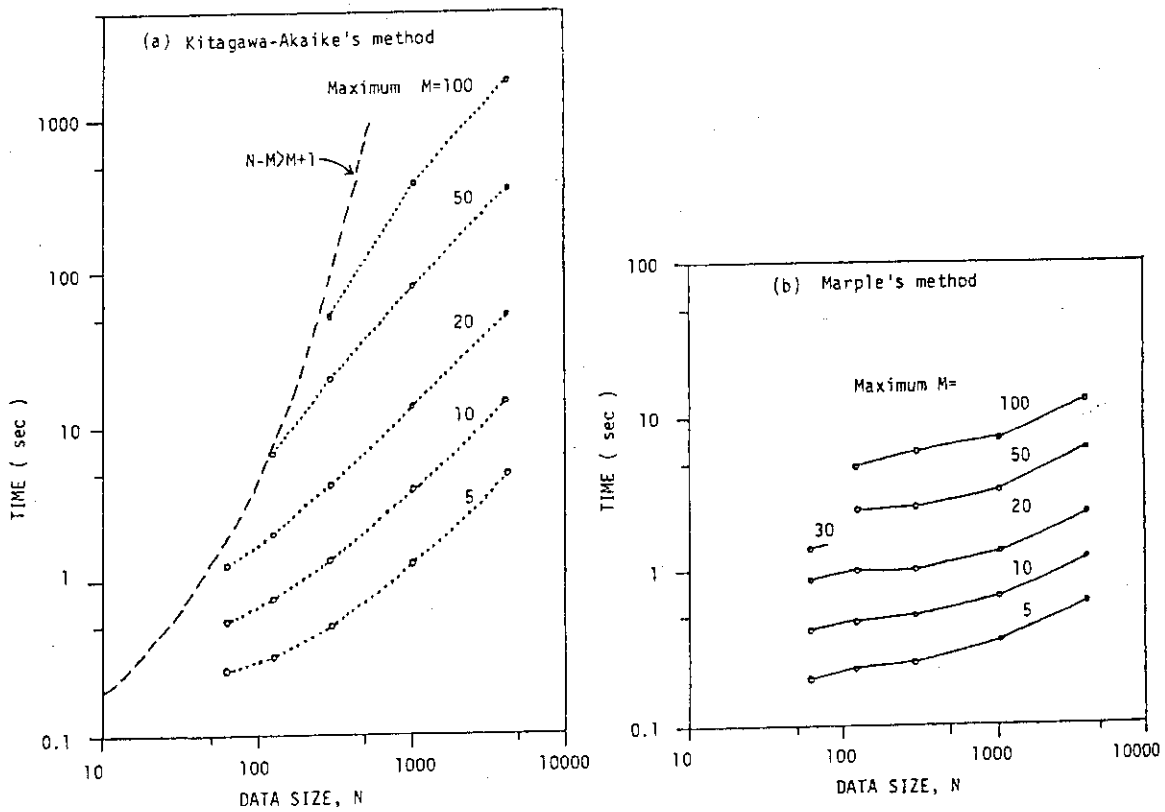


Fig. 3.25 Computation time vs. data size N of the AR model fittings

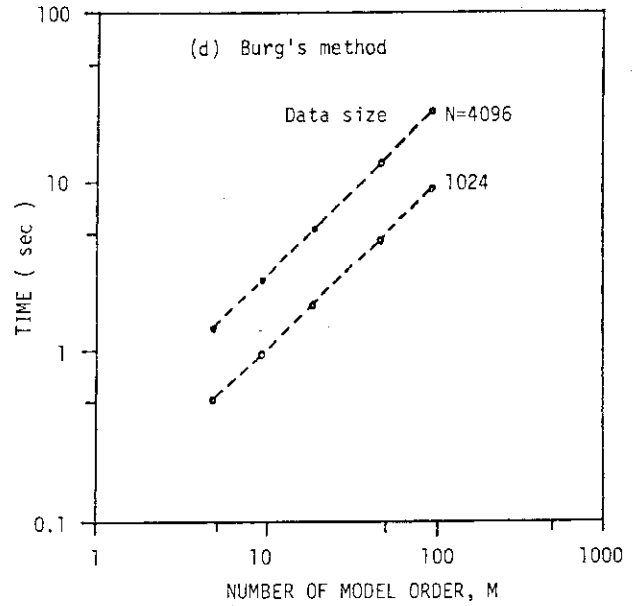
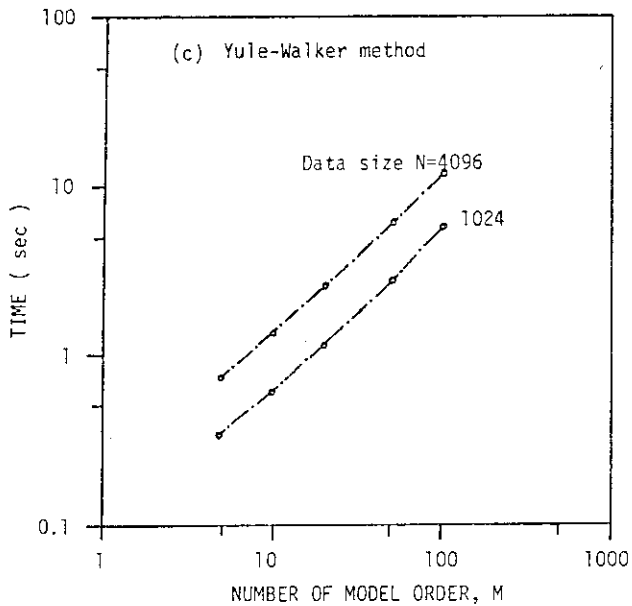
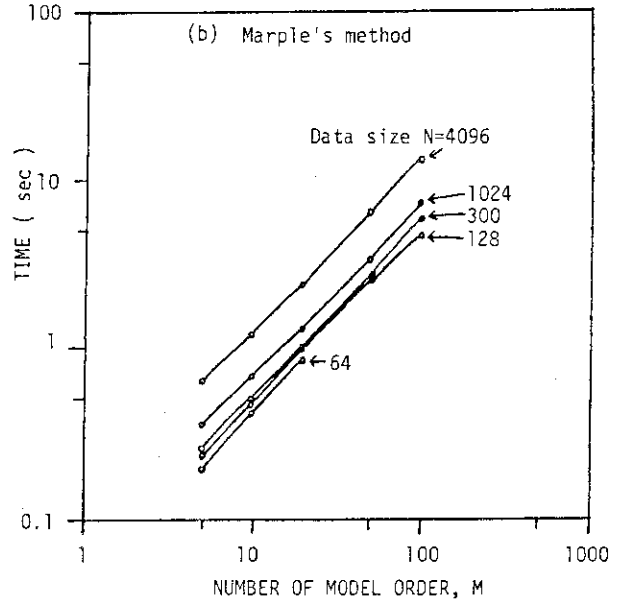
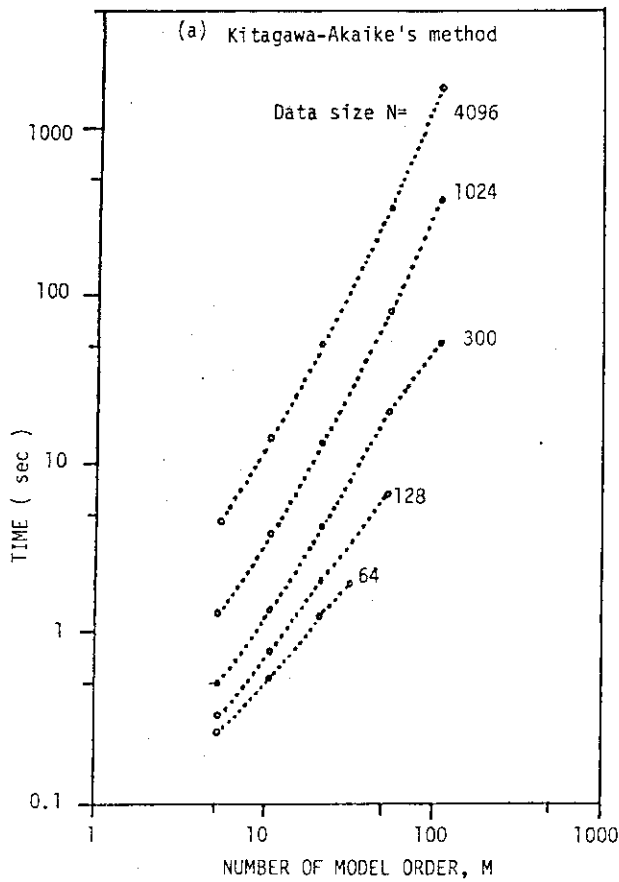


Fig. 3.26 Computation time vs. model order M of the AR model fittings  
 (a) the Kitagawa-Akaike's method (b) the Marple's method  
 (c) the Yule-Walker method (d) the Burg's method

## 4. 結 論

Marple および北川・赤池がそれぞれ独立に提唱した最小2乗法に基づくAR（自己回帰）モデル・フィッティングのための2つの異なるアルゴリズムの実験的評価を目的として、シミュレーション・スタディを行った。

この結果両方法とも従来の方法（Yule-Walker 法やBurg 法）に比較して以下の点で優れている事が確認された。

- (1) データ中に強い周期成分が含まれている場合でも安定な数値計算が行え、その結果も妥当である。
  - (2) 短いデータ（サンプル数の少ないデータ）からフィッティングを行いスペクトルを推定したときに生じる、スペクトル・ピークの周波数バイアスが非常に小さい。
- また、両方法は相互の比較を通じて以下の特徴や性質をもつことが明らかになった。
- (1) 十分大きいサンプル・データでは両方法とも得られた結果はほぼ一致する。
  - (2) サンプル数の小さいデータのフィッティングで問題となるスペクトル・ピークの周波数バイアス誤差は北川・赤池の方法よりもMarpleの方法の方がさらに小さい。
  - (3) 同じくサンプル数が小さい場合に問題となるスペクトル・ピークの高さの推定誤差は従来の方法と同様に両方法ともかなり大きい。
  - (4) 北川・赤池の方法はそのアルゴリズムの構造から、最大次数 $M_{\max}$ を設定してサンプル数 $N$ データについてフィッティングを行うと、同時に算出される $M_{\max}$ より低い次数 $m$ の結果はサンプル数 $N-(M_{\max}-m)$ のデータから計算したものが得られる。このため、AIC等によるモデルの最適次数の決定には注意を要する。また、データの有効利用の見地からは $M_{\max}$ より低次の結果は使用せず、 $M_{\max}$ を逐次変えて再計算を行う方が良い。
  - (5) 計算時間に関しては、Marpleの方法はYule-Walker法と同程度、Burg法よりも2倍近く速く処理できることがわかった。これに対して、北川・赤池の方法は、 $N \times M_{\max}^2$ に比例するため他に比較してかなり遅いことがわかった。このため、アルゴリズム上はARパラメータ等の新しいデータによる更新などが簡単に実行できる構造を持ちながらも、原子炉の異常診断等の高速性を要求されるオンライン処理への応用にはかなりの困難が予想される。

この他、モデルの次数決定に関して各フィッティング方法の最小AIC評価（MAICE）について検討した。この結果、

- (1) Burg および Marple の両方法のAICは、通常定数として無視される項が次数 $M$ に関する変数となる（3.5式）ため、これを含めた形で計算する必要がある。
- (2) 北川・赤池の方法の場合も異なる最大次数 $M_{\max}$ のAICを比較する場合は、この定数項が無視できない。
- (3) MAICEによって次数決定を行う場合、データの持つ標準偏差の大きさによっては計算精度上の問題が生ずる。そこで、数値計算上の精度を確保するためにデータの標準偏差を適当な係数で規準化する必要がある。

以上を総合すると、Marpleおよび北川・赤池の方法は、計算時間は別として、従来の方法より優れたフィッティング方法であると結論される。

## 謝 辞

本研究の遂行に当り有益な助言を与えて下さった原子炉制御研究室の篠原慶邦室長、島崎潤也氏に、そして報告書の作成に当り種々の検討をして下さった尾熊律雄氏に謝意を表します。

## 参考文献

- (1) 林 光二, 尾熊律雄; Burgのアルゴリズムに基づくARモデルの同定とその原子炉異常診断への応用, JAERI-M 82-009 (1982)
- (2) L. Marple; A New Autoregressive Spectrum Analysis Algorithm, IEEE Vol. ASSP-28 p. 441 (1980)
- (3) 添田 喬, 中溝高好, 大松 繁; 信号処理の基礎と応用, 日新出版 (1979)
- (4) G. Kitagawa and H. Akaike; A Procedure for the Modeling of Non-Stationary Time Series, Ann. Inst. Statist. Math., Vol. 30, Part B, p. 351 (1978)
- (5) 小柳義夫; 最小二乗法の新しいアルゴリズム, 情報処理, Vol. 23, No. 2, p. 99 (1982)
- (6) H. Akaike, et al.; TIMSAC-78: Computer Science Monographs, The Institute of Statistical Mathematics (1979)

以上を総合すると、Marpleおよび北川・赤池の方法は、計算時間は別として、従来の方法より優れたフィッティング方法であると結論される。

#### 謝 辞

本研究の遂行に当り有益な助言を与えて下さった原子炉制御研究室の篠原慶邦室長、島崎潤也氏に、そして報告書の作成に当り種々の検討をして下さった尾熊律雄氏に謝意を表します。

#### 参考文献

- (1) 林 光二, 尾熊律雄; Burgのアルゴリズムに基づくARモデルの同定とその原子炉異常診断への応用, JAERI-M 82-009 (1982)
- (2) L. Marple; A New Autoregressive Spectrum Analysis Algorithm, IEEE Vol. ASSP-28 p. 441 (1980)
- (3) 添田 喬, 中溝高好, 大松 繁; 信号処理の基礎と応用, 日新出版 (1979)
- (4) G. Kitagawa and H. Akaike; A Procedure for the Modeling of Non-Stationary Time Series, Ann. Inst. Statist. Math., Vol. 30, Part B, p. 351 (1978)
- (5) 小柳義夫; 最小二乗法の新しいアルゴリズム, 情報処理, Vol. 23, No 2, p. 99 (1982)
- (6) H. Akaike, et al.; TIMSAC-78: Computer Science Monographs, The Institute of Statistical Mathematics (1979)

以上を総合すると、Marpleおよび北川・赤池の方法は、計算時間は別として、従来の方法より優れたフィッティング方法であると結論される。

## 謝 辞

本研究の遂行に当り有益な助言を与えて下さった原子炉制御研究室の篠原慶邦室長、島崎潤也氏に、そして報告書の作成に当り種々の検討をして下さった尾熊律雄氏に謝意を表します。

## 参考文献

- (1) 林 光二, 尾熊律雄; Burgのアルゴリズムに基づくARモデルの同定とその原子炉異常診断への応用, JAERI-M 82-009 (1982)
- (2) L. Marple; A New Autoregressive Spectrum Analysis Algorithm, IEEE Vol. ASSP-28 p. 441 (1980)
- (3) 添田 喬, 中溝高好, 大松 繁; 信号処理の基礎と応用, 日新出版 (1979)
- (4) G. Kitagawa and H. Akaike; A Procedure for the Modeling of Non-Stationary Time Series, Ann. Inst. Statist. Math., Vol. 30, Part B, p. 351 (1978)
- (5) 小柳義夫; 最小二乗法の新しいアルゴリズム, 情報処理, Vol. 23, No. 2, p. 99 (1982)
- (6) H. Akaike, et al.; TIMSAC-78: Computer Science Monographs, The Institute of Statistical Mathematics (1979)