

電磁気観測データのRobust stacking法について

東京大学地震研究所 上嶋 誠

On the robust stacking method of EM data
Makoto Uyeshima (Earthquake Research Institute)

1. はじめに

MT法やGDS法において、impedance tensorやtipper等、電磁場間の周波数伝達関数

$$Y(f) = H_1(f) \cdot X_1(f) + H_2(f) \cdot X_2(f) \quad (1)$$

$H_i(f)$; 周波数伝達関数

$X_i(f)$; 入力のFourierスペクトル (磁場の水平2成分)

$Y(f)$; 出力のFourierスペクトル (MT法では電場の1成分、GDS法では磁場の垂直成分)

を求める際、観測データをいくつかの幅で区切り、各々の区間でFFTを行い全体として平均的な伝達関数を求めるstacking法が一般的である。ところが、特に電場においてそうであるが、矩形状の異常変化やステップ (図3-4) が現れたり、昼間のノイズが大きかったりするために推定した伝達関数のエラーレベルが大きくなり、時には100%を越えることも少なくない。

これに対して、例えば、夜間値のみを使用したり、磁気嵐の時のデータのみを使用したり、データの出力を見てなるべくそういった異常変化の影響の少なそうなところを選択してstackingを行うというのが一般的であるが、そういったデータの選択を自動的に行う方法を考案したのでここに紹介する。

2. 周波数伝達関数の統計的性質

Bendat and Piersor (1971) によれば、もとの各々のデータのノイズがガウス分布に従うエルゴード的定常不規則過程であると仮定すると、そのパワーはカイ自乗分布に従い、その比である伝達関数は、F分布に従う。そしてその標準誤差 ε_H は、

$$\varepsilon_H = \frac{4}{n-4} (F_{n_1, n_2; \alpha}) \frac{[1 - \hat{\gamma}_{y \cdot x}^2(f)] \hat{G}_y(f)}{[1 - \hat{\gamma}_{i \cdot x}^2(f)] \hat{G}_i(f)} \quad (2)$$

で表される。ここでは、2入力1出力系 (式(1)) を考えており、

n : 各々のスペクトル密度の推定の自由度 (stacking数 * 2)

$F_{n_1, n_2; \alpha}$: 自由度 $n_1=4, n_2=n-4$ のF分布の100 α 百分比点 ($\alpha=0.05$ 、即ち、95%信用区間がよく用いられる)

$\hat{G}_i(f)$: 入力 $x_i(t)$ のパワースペクトルの推定値

$\hat{G}_y(f)$: 出力 $y(t)$ のパワースペクトルの推定値

$\hat{\gamma}_{y \cdot x}^2(f)$: 出力 $y(t)$ とすべての入力との間のmultiple coherencyの標本推定値

$\hat{\gamma}_{i \cdot x}^2(f)$: 入力 $x_i(t)$ と $x_i(t)$ 以外の全ての入力との間のmultiple coherencyの標本推定値

である。式(2)をみればわかるように、自由度が大きく (即ちstacking数が大きく、より

多くのデータから推定し)、異常変化が含まれていないほど(即ち、入出力間のcoherencyが高く、入力どうしのcoherencyが低いほど)、標準誤差は小さくなり、求められた伝達関数はよく決まる。

従って、全観測データをいくつかの区間に分け、stackingを行う際、上記の条件を満たす区間を可能な限り多く選び出せば良いことになる。問題は、いかにしてそういった区間を選び出すかである。

3. 自動的にstacking区間を選択する方法

まず全ての観測データをいくつかの区間に区切り、各々の区間でFFTを行い、全部の区間から平均的な伝達関数

$$\begin{pmatrix} H_1(f) \\ H_2(f) \end{pmatrix} = \begin{pmatrix} S_{11}(f) & S_{12}(f) \\ S_{21}(f) & S_{22}(f) \end{pmatrix}^{-1} \begin{pmatrix} S_{1y}(f) \\ S_{2y}(f) \end{pmatrix} \quad (3)$$

を求めるところから始める。ここで、

$S_{iy}(f)$: 入出力間の相互スペクトル密度関数の推定値

$S_{ij}(f)$: 入力間の相互スペクトル密度関数の推定値

$S_{ii}(f)$: 入力のパワースペクトル密度関数の推定値

である。こうして、各々のstacking区間について

$$\hat{X}_y(f) = H_1(f) * X_1(f) + H_2(f) * X_2(f)$$

を求め、 $\hat{X}_y(f)$ は、各stacking区間の生の入力スペクトル $X_i(f)$ (FFTの結果)と平均的な伝達関数 $H_i(f)$ より計算される生の出力スペクトルの推定値である。あるstacking区間の時系列上に矩形形状やステップなどの白色的でないノイズがのっているとき、あるいはあるstacking区間における応答の仕方が異常であるとき、そのstacking区間について、この推定値 $\hat{X}_y(f)$ と実際にFFTによって計算された生の出力スペクトル $X_y(f)$ とは大きく食い違ってくるはずである。そこで、この予測誤差 $\varepsilon = |\hat{X}_y(f) - X_y(f)|$ が最も大きなstacking区間を除去し、残ったものからまた平均的な伝達関数 $H_i(f)$ を求めると、最初の推定に比べてcoherencyが向上し、求められた伝達関数は最初のものよりよく決まっているはずである。こうして、次には残ったstacking区間の中から最も誤差 ε が大きなものを除去し、また、その残ったものから $H_i(f)$ を求める。そして、このような、最も ε が大きいstacking区間を除去してはその残りから $H_i(f)$ を求めるという作業を次々に行っていく。なお、実際の計算においては、上記の誤差 ε を入力のスベクトル強度でnormalizeしたものを周波数にわたって平均化した

$$\sum_{k=1}^N |X_y(f_k) - \hat{X}_y(f_k)| / \sqrt{|X_1(f)|^2 + |X_2(f)|^2} / 2 / N \quad (4)$$

を用いて上記の除去stacking区間の選択を行っている。ここで入力のスベクトル強度でnormalizeを行うのは、誤差 $|X_y(f_k) - \hat{X}_y(f_k)|$ が入力のスベクトル強度に比例して大きくなることが予想されるからである。

こうした除去を無限におこなえば、常に最も誤差 ε が大きいstacking区間は存在するのでしまいにはデータがなくなってしまうが、どこで除去を中止するかを判断するのに2. でのべた標準誤差(式(2))を用いる。先に述べたように、一般的に異常変化を含んだ

stacking区間を除去していくと、入出力間のcoherencyが急激に増大することにより、標準誤差が減少し、周波数伝達関数 $H_i(f)$ の推定値はより正しくなっていく。しかし、異常変化を含んだstacking区間の除去が完了すると、それ以降いくら誤差 ε の大きなデータを取り除いていっても統計的性質はあまり変わらずcoherencyの向上はあまりみられなくなる。この場合には、stackingの数が減少していくことによる伝達関数推定の不定性の方が大きくなる。従って、stackingの数を上記のcriterionのもとに減少させていくと同時に2. で述べた標準誤差のチェックを行うと、最初のうちは標準誤差が減少して行くが、異常変化が完全に取り除かれた時点から標準誤差はかえって増大していく。この、標準誤差が最小になったときの伝達関数をもって、最も良い推定値であるとするのである。この方法は、全体のデータセットの中からある統計的性質に従う部分を自動的に選択し取り入れるという点でrobust推定法の一つであるといえる。もちろん、全体にわたってノイズレベルが非常に低い場合は、一番始めに得られた伝達関数が最も標準誤差が小さく最も良い推定値となる。

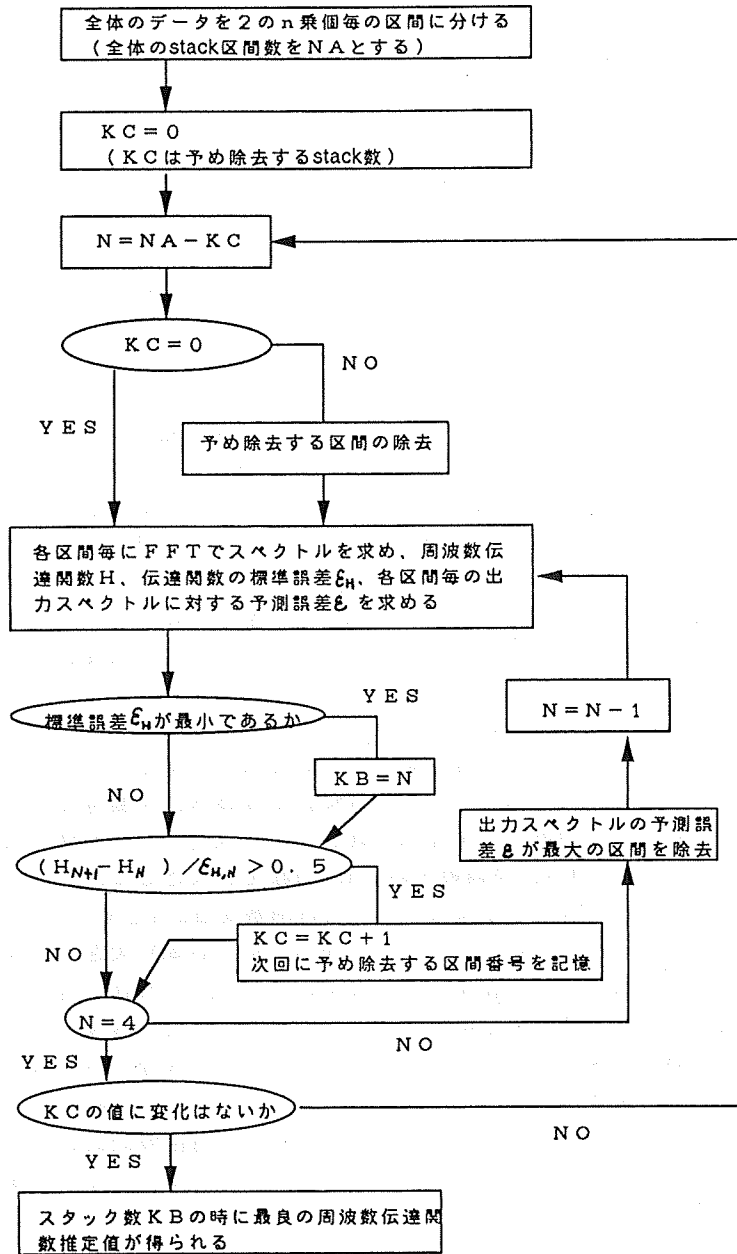
ところが、この方法を実際のデータにこのままの形で適用すると問題が生じる。それは、上で述べたように入力のスペクトル強度でノーマライズするところに起因している。すなわち、入力が大きいく所所に異常な変化がのっているときにその異常の発見が遅れるわけである（これは、磁気嵐などシグナルの振幅が大きいく所に異常があってもそれを発見しにくいということで感覚的には正当なことのように思われる）。このため、そういった異常変化がある場合、かなりstacking数が減少した時点でそういった異常が検知され、それを除いたところで標準誤差がステップ的に小さくなるといった事態が生じることになる。時には、その時点で標準誤差が最小になることもある。そういった異常はstacking除去の最初の段階に取り除いておけばより良い伝達関数の推定値が得られるはずである。このため、stacking除去の過程で次のようなチェックをも行うことにした。すなわち、 i 個目のstacking除去における伝達関数の推定値および標準誤差をそれぞれ H_N 、 $\varepsilon_{H,N}$ （ここで N は i 個目のstacking除去段階におけるstack数、即ち、最初の全stack区間数を N_A として、 $N = N_A - i$ ）、その前の $i - 1$ 個目のstacking除去の結果得られた伝達関数を H_{N+1} として、

$(H_{N+1} - H_N) / \varepsilon_{H,N}$ があるレベル以上（我々は0.5にしている）大きいとき、 i 個目に取り除いたstacking区間には異常が存在するとみなすのである。これは、異常stacking区間を除去したときの周波数伝達関数の推定値の変化がその段階の誤差レベル $\varepsilon_{H,N}$ に比べてある程度大きい場合に、その除去されたstacking区間に特に寄与の大きなノイズが含まれているという判断を下すことを意味する。そうしたチェックを1回目のstacking除去の段階で行い、次に、チェックされたstacking期間を予め除去しておいて2回目のstacking除去を行っていく。こうして、上記のチェックにひっかかったstacking区間がなくなるまで除去の過程を繰り返す。こうしたことを繰り返すと、また、仕舞にはデータがなくなってしまうのではないかという危懼を抱かれるかも知れないが、実際にこの方法でいくつかの実際のデータを解析してみた結果、ほとんどのデータで1回のチェックを行う（すなわちstacking除去の過程を2回繰り返す）だけでことが足りるようであった。

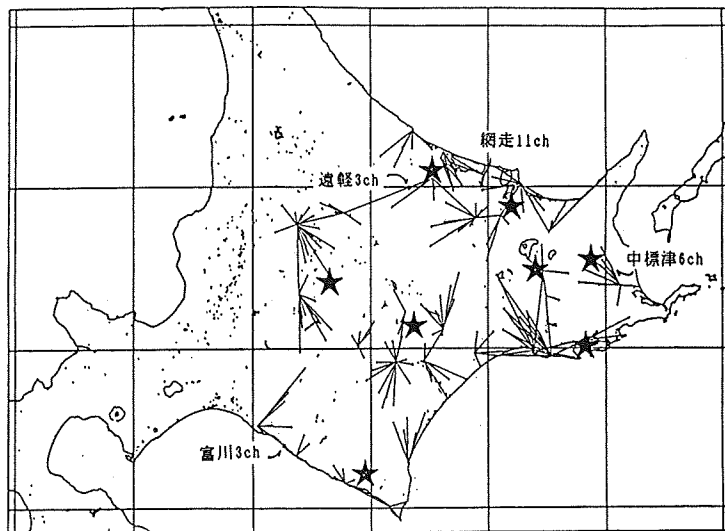
最後に全体のアルゴリズムを第1図に示しておく。

4. 実際のデータへの適用例

上記の方法を、1989年7月から11月にかけて行われた、北海道中東部におけるNTT通信



第1図、全体のアルゴリズム図 (stack数4個になるまで除去を繰り返すことにしている)



第2図 北海道中東部におけるNETWORK-MT法の観測網。★は磁力計の設置点。

用アースと回線網を用いた地電位差観測 (Network-MT、第2図) データの解析に用いた。いくつかの測線では図3-4で示したような異常変化がみられたり、観測途中で回線に支障をきたして微妙に振幅がおかしくなったりした。

解析として、女満別地磁気観測所の磁場水平2成分データ $H(f)$ 、 $D(f)$ に対する各測線の電位差 $V(f)$ のtransfer function $T_{UH}(f)$ 、 $T_{UD}(f)$;

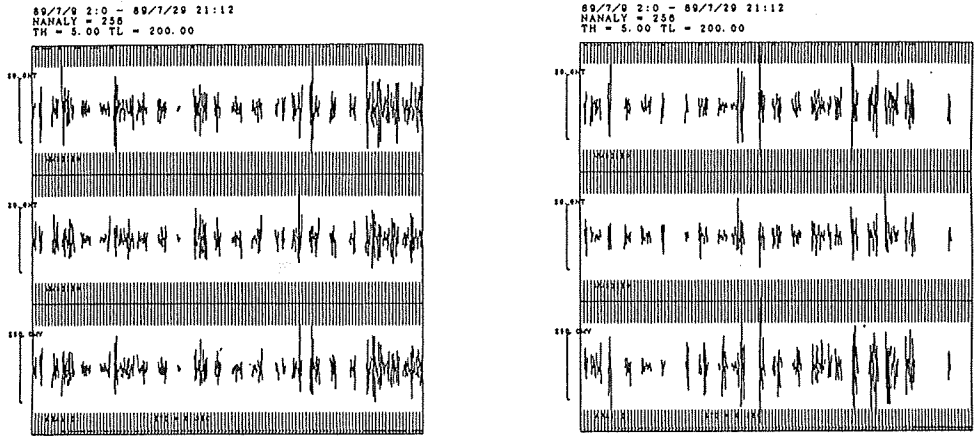
$$V(f) = T_{UH}(f) \cdot H(f) + T_{UD}(f) \cdot D(f)$$

の算出を行った。

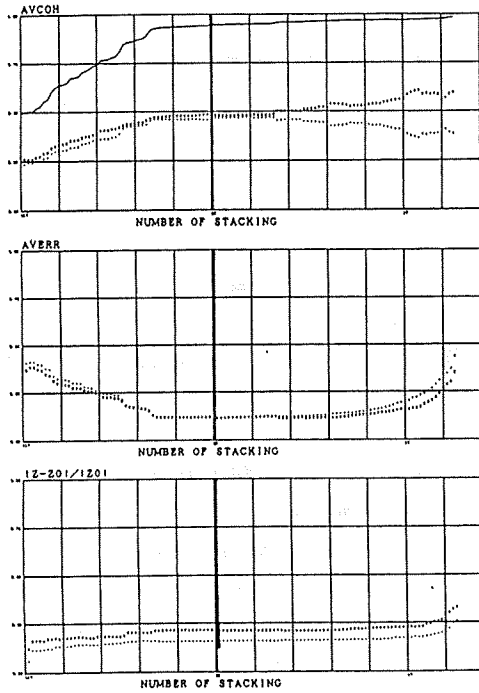
具体的な異常区間除去の様子を示すため、最後の段階 (標準誤差が最小) でtransfer function推定に使われたstacking区間、除去され使われなかったstacking区間の時系列データを女満別の磁場水平2成分と共に示し、stacking数が減少していくときのcoherency、標準誤差、transfer function初期推定値 Z_0 からのずれの様子、パワースペクトル、transfer functionの変化の様子を図で示すことにする。

第3図は網走の11チャンネルのものである。第3-1図では余りに時間を縮めてみて本当に異常変化が取り除かれているのかわからないため時間軸を延ばして描いたものが第3-4図であり、目でみて判断のつきにくい磁気嵐の時の異常変化も取り除かれている様子がわかる (太線内が取り除かれたstacking区間である)。また、第3-2図ではこれまで述べてきたことを視覚的に確かめることが出来る。すなわちstacking除去が進むにつれ、はじめのうちはcoherencyが急激に増大し標準誤差の改善がみられるが、あるところからcoherencyの改善がみられなくなりそこから標準誤差はかえって増大していくことが確かめられる。なお、太い点線は T_{UH} に関する情報、細い点線は T_{UD} に関する情報であり、縦の太線は、最も標準誤差を小さくするところである。第3-3図では、絶対値についてはlogスケールで描いているためにあまり変化が明瞭ではないが、誤差が減っていく様子、位相が次第にまとまっ

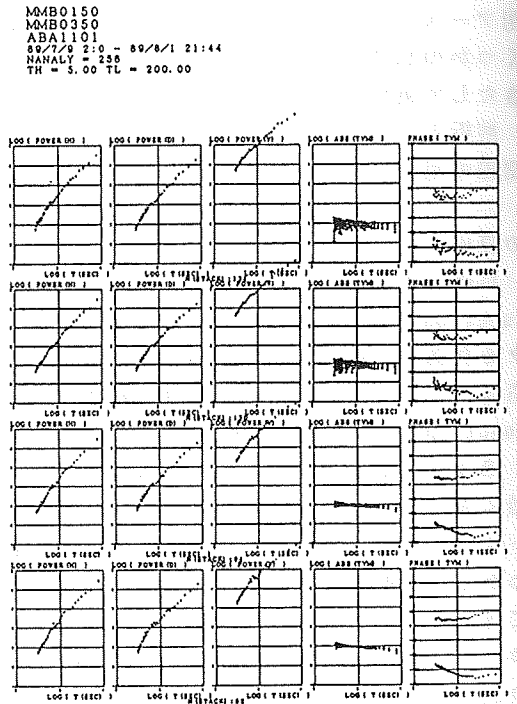
第3図 網走の11チャンネルについての除去stacking区間選択過程



第3-1図

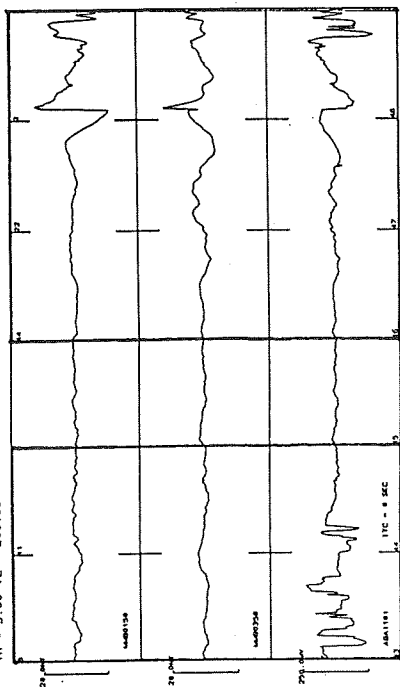


第3-2図

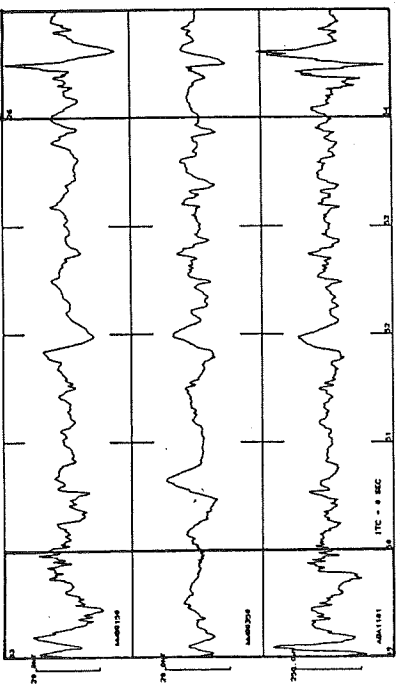


第3-3図

09/16 13:12 - 09/17 14:40
 MANLY - 256
 TH - 5.00 TL - 200.00



09/17 14:40 - 09/18 16:24
 MANLY - 256
 TH - 5.00 TL - 200.00



第 3 - 4 図

第 3 - 1 図 最良 (transfer function の標準誤差 ϵ_H が最小) の推定を行うのに使われた stack 区間 (左) と異常変化があるとして使われなかった区間 (右)。縦線の間隔は 256 データ (256 分)、上から女満別の磁場 NS 成分、EW 成分、網走 11ch 地電位差の各データ。

第 3 - 2 図 異常 stack 区間除去過程における種々の統計量の変化。左端が除去 0 であり、右端のデータは stack 数 4 に対応する。上から順に、partial coherency、transfer function の標準誤差、stack 数 0 の時に得られた transfer function からのずれを表している。太い点線は T_{UH} 、細い点線は T_{UD} に関する情報であり、coherency の図で実線は multiple coherency を示す。太い縦線は、 ϵ_H を最小にする stack 数の所にひかれてある。

第 3 - 3 図 各 stack 区間除去段階における、(左から) H のパワー、D のパワー、V のパワー、transfer function (太い点線は T_{UH} 、細い点線は T_{UD}) の絶対値 (エラーバーつき)、位相。横軸は周波数。上から順に除去 stack 数 0、前もって異常 stack 区間を除去した段階 (第 4 - 3 図では描かれていない)、中間的段階、 ϵ_H 最小の最良の段階。

第 3 - 4 図 第 3 - 1 図の一部を時間軸を引き延ばして描いた図。太線で囲まれた区間が最終的に除去された stack 区間。

ていく様子などが読み取れる。図は上から、除去stack数0、 $(H_{N+1} - H_N) / \varepsilon_{H,N}$ チェックで引かかったものを除去した段階、stack除去の中間段階、最良の推定値を与えた段階を順に示している。第4図は、中標津の6チャンネルのもので昼間のノイズが著しい。第4-1図をみれば夜間と磁場変動の大きいときのみを使って最良の予測値が得られたことがわかる。第5図は、途中で回線の支障を来した遠軽の3チャンネルの例で、最後の方のデータがことごとく使用されていない。最後に良いデータの例として富川の3チャンネルの例を第6図に示す。除去stack数0での標準誤差が最も小さくなっていた。

以上のように、各地のデータで今回の方法の有用性が実証されたわけである。

5. おわりに

誤差レベルの評価をするのに、周波数にわたっての平均値を使ったために、数の多い短周期の誤差を小さくするように除去stack区間が選択されるようであった。これに対しては、電場のノイズにはステップ状、パルス状など、短周期により影響を及ぼすものが多いことから妥当であろうと判断したのであるが、もっと厳密に推定を行うには周波数を限って、極端にはDFTの手法に今回の方法を適用すれば良いと思われる。また、SOMPI法においてその波素を選択する際にも今回の考え方が適用でき、より良い周波数変換関数推定値が得られるであろう。いずれの方法に適用するにせよ、客観的に異常変化の除去を行えるところにここで述べてきた方法の特長がある。

また、従来のrobust stacking法は(例えば中川・小柳(1982)、Chave et. al. (1987))、予測誤差 ε の大きさに応じて適当な重みをかけ(これは予測誤差 ε の分布としてGauss的ではなくもっと裾の拡がりの大きな分布を仮定することに対応する)、適切な重み関数をみいだすために ε を最小にすべくiterationを繰り返すというものであった。これに対して、ここで述べられている方法は、誤差がGauss的であるという仮定は捨てず周波数伝達関数がF分布に従うと仮定し、 ε_H を用いてその条件に適合するデータを全体のなかから選択するという点で従来の方法とは異なる。

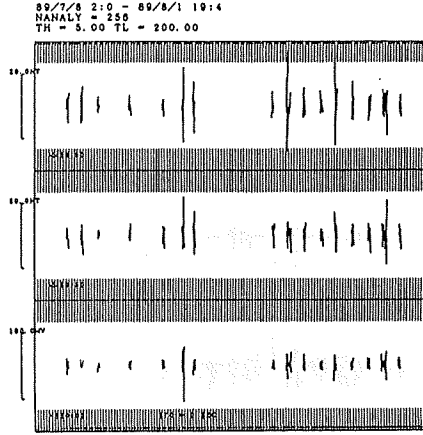
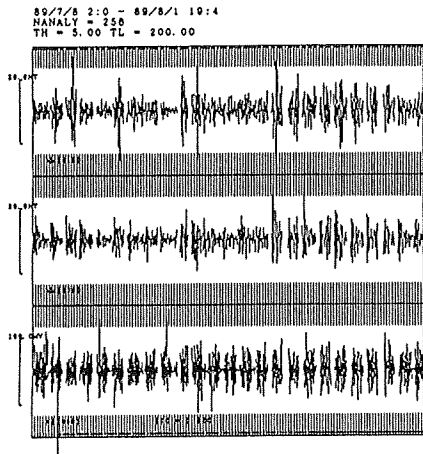
6. 謝辞

観測データの取得には各地のNTTの方々、北海道大学理学部の西田泰典助教授、大島弘光助手、清水久芳氏、東海大学海洋学部の上田誠也教授、木下正高助手、東京大学地震研究所の歌田久司助手、川瀬隆治氏に多大なるご援助を得ました。この中の西田泰典助教授、歌田久司助手には有益なご助言を頂きました。また、磁場のデータを快く送付頂きました気象庁地磁気観測所の方々にも厚く謝意を表します。

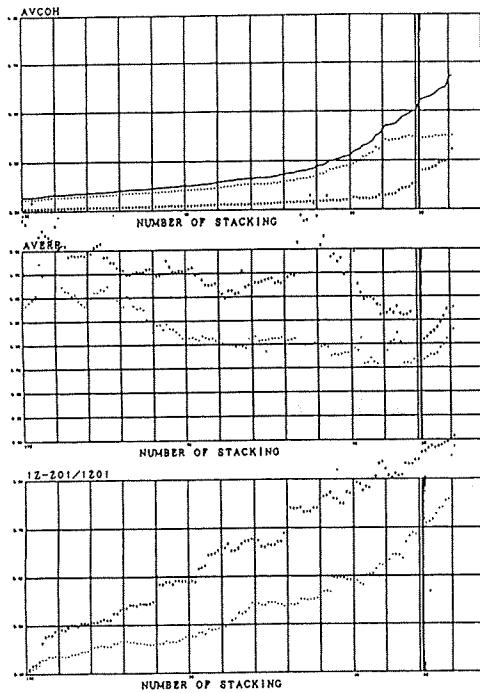
7. References

- Bendat, J. S. and Piersol A. G. (1971), Random Data; Analysis and Measurement Procedures, (日本語訳) ランダムデータの統計的処理, 培風館, 401ページ
中川 徹、小柳義夫 (1982), 最小二乗法による実験データ解析—プログラムSALS—, 東京大学出版会 (UP応用数学選書 7), 206ページ
Chave, A. D., Thomson, D. J. and Ander, M. E. (1987), On the Robust Estimation of Power Spectra, Coherences, and Transfer Functions, J. Geophys. Res., 92, 633-648

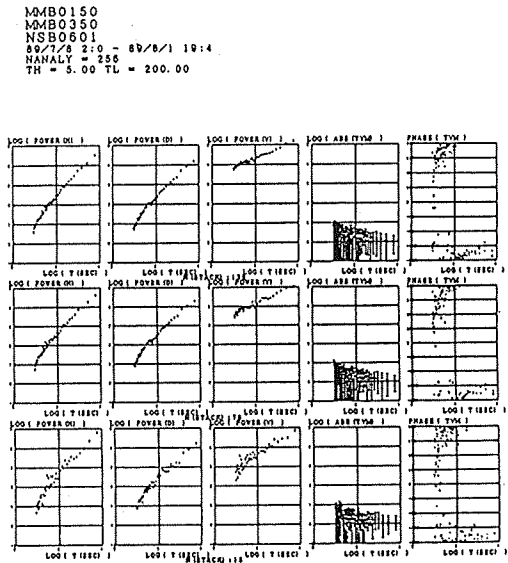
第4図 中標津の6チャンネルについての除去stacking区間選択過程
 (昼間のノイズが著しい例)
 (-1、-2、-3図で表示している内容は第3図と同じ)



第4-1図

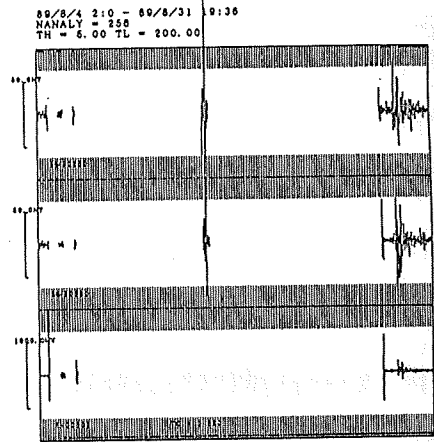
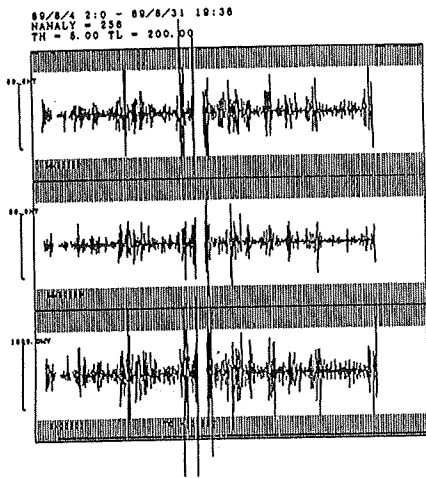


第4-2図

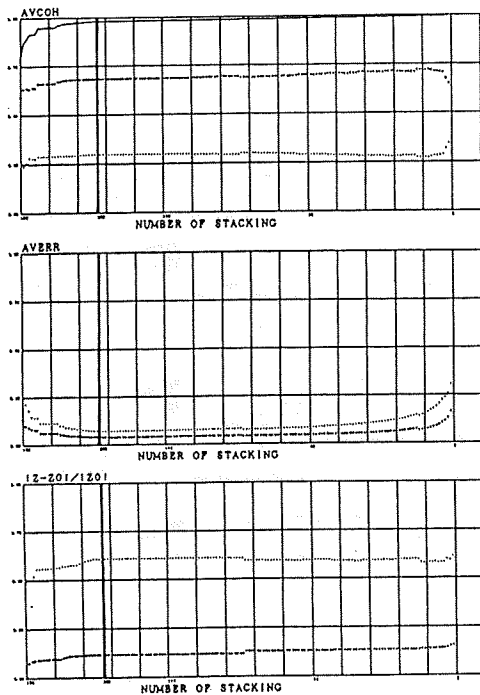


第4-3図

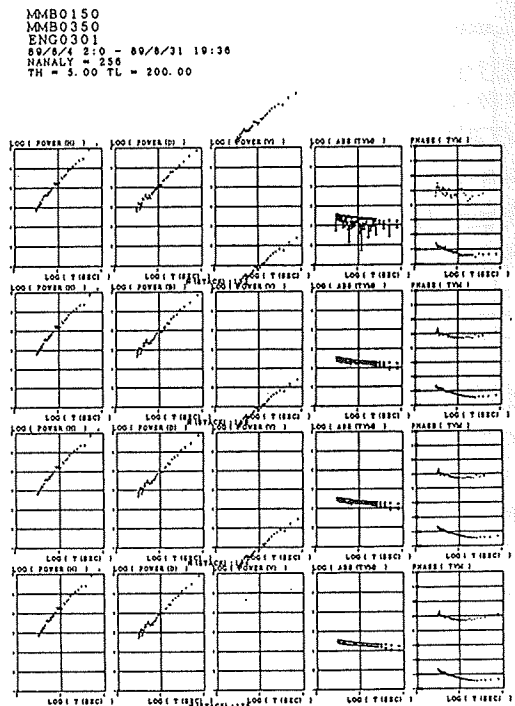
第5図 遠軽の3チャンネルについての除去stacking区間選択過程
 (観測途中で回線に異常をきたした例)
 (-1、-2、-3図で表示している内容は第3図と同じ)



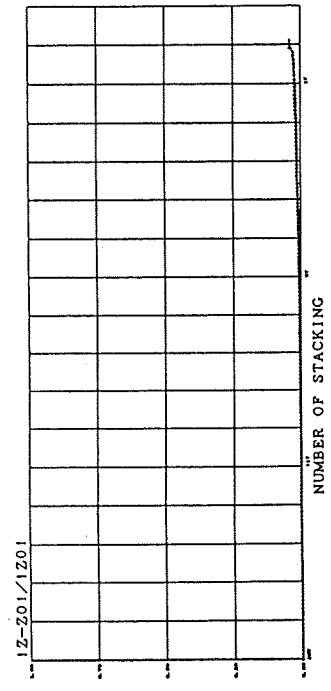
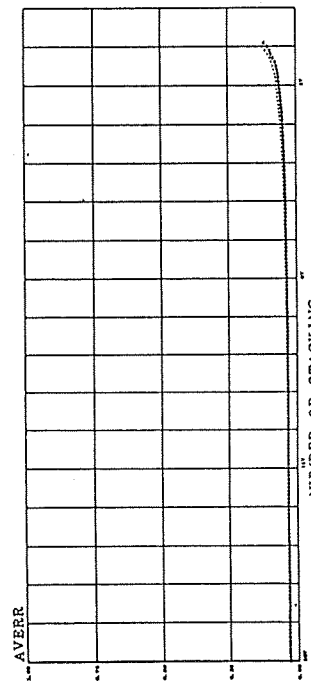
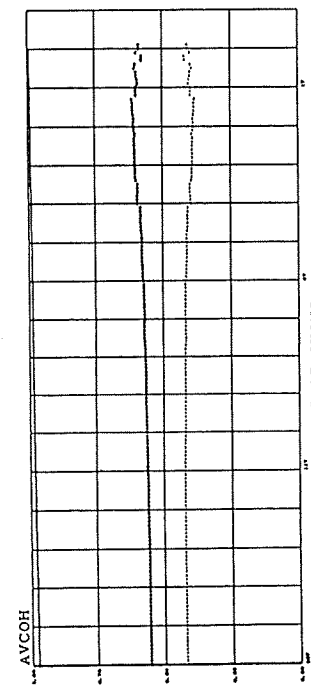
第5-1図



第5-2図

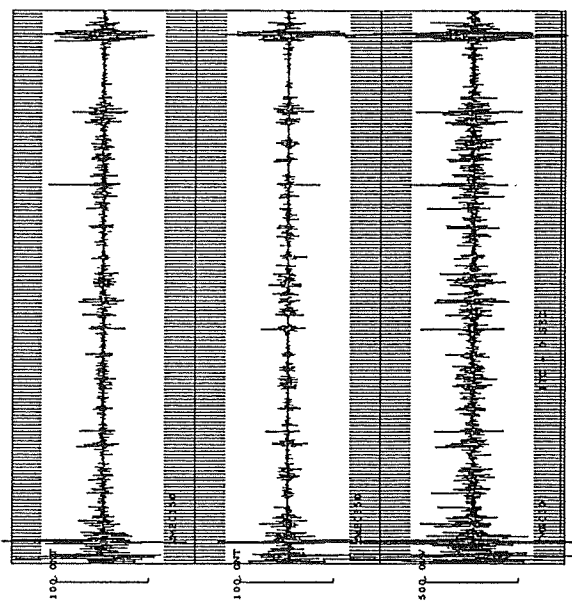
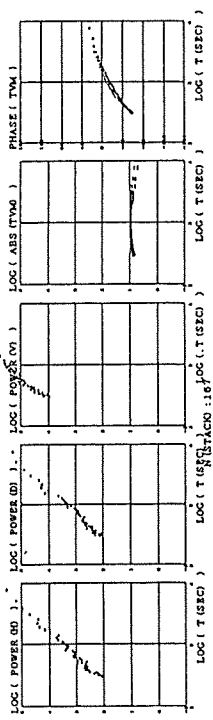


第5-3図



MMB0150
MMB0350
TMK0301
89/10/20 17:5 - 89/11/19 9:37
NANALY = 256
TH = 5.00 TL = 200.00

第6-3 図



第6-1 図

第6-2 図

第6 図 富川の3チャネル (除去stack数0の例、-1、-3 図は1つか描かれていない。)