

地球磁場観測ネットワークデータの解析手法について（その1）

藤井郁子（東京大学地震研究所）、A.シュルツ（ケンブリッジ大学理論地球物理研究所）

On Data Processing Methods for the Geomagnetic Observatory Network: Part I

Ikuko Fujii(ERI, Univ. of Tokyo) and Adam Schultz (ITG, Univ. of Cambridge)

1、はじめに

MT 法や GDS 法のように周波数領域で電磁場の応答関数を求めて地下の電気伝導度構造を推定する場合、問題を難しくしているのが応答関数の地下構造に対する感度の悪さである。例えば、地下で電気伝導度が 10 倍変化していても応答関数が 10 倍変化することはない。逆に言えば、応答関数の微少な変化を精度よく計算しなければ、電気伝導度の大きな変化でさえ捕らえられないことになる。ところが、データ処理手法の違いから、求められた応答関数に大きな差が現れることが知られている。

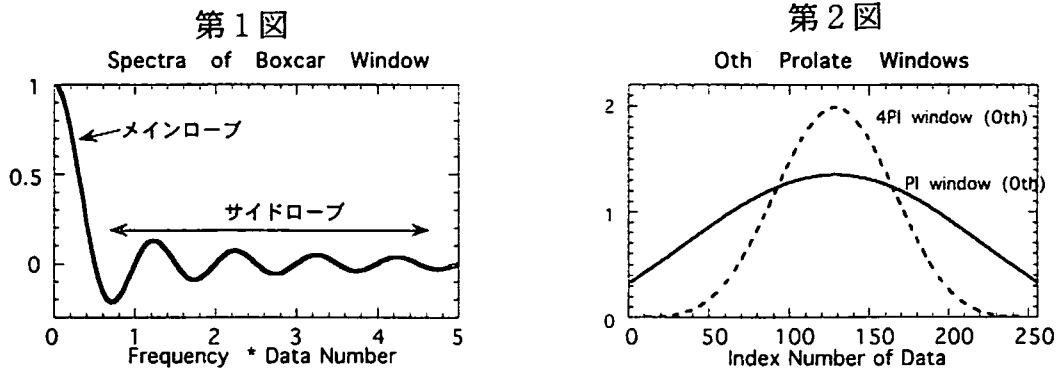
我々は全球での地球磁場観測ネットワークのデータを使って地磁気応答関数を計算する研究を行ってきたが、その過程でわかった様々な処理法による結果の違いについて報告する。ここで取り上げるのは定常時系列解析であり、バックグラウンドスペクトルを求めることを目的とし、最小自乗法による応答関数の計算にロバスト推定を取り入れる。本論文では紙面の都合上、一観測点の時系列から最も正確なスペクトルを求めるためのデータ処理手法についてのみ述べる。観測点が多数ある時にデータの空間構造を知るための手法については続編に譲る。本論文の構成は、第 2～6 章で各手法の背景説明等をし、第 7 章で応答関数にあらわれる処理による違いを概観する。我々の結果の要約は、（1）目的とするタイプのシグナルを決めたらそれにあつた解析法を選び、それ以外のタイプのシグナル・ノイズは取り除いたほうがよい、（2）不要なものには様々なタイプがあるのでそれぞれにあつた方法で取り除く必要がある、（3）不要なものの中では振幅の大きいものほど影響が大きい、である。

2、ウィンドーの選択

どのスペクトル解析の教科書にも書いてあるが、 $|t| \leq N/2$ で定義される時系列データ $x(t)$ をフーリエ変換すると、それは幅 N の箱形ウィンドーを作用させた時系列のフーリエ変換を行うことに等しい。箱形ウィンドーの周波数特性を第 1 図に示す。箱形ウィンドー付きのスペクトル $x(\omega)$ は、主に周波数 ω （第 1 図では原点）を中心としたある範囲（メインローブ）のスペクトルを重ね合わせたものを反映しているが、サイドローブを通じて離れた周波数のスペクトルの影響も無視できないほど取り込んでいることがわかる。

本来は第 1 図が $\delta(\omega)$ になるのが望ましく、少しでも $\delta(\omega)$ に近付けるため時系列にウィン

ドローを作用させてフーリエ変換を行う。ウィンドーの設計にはエネルギー保存の要請から、(1)メインローブの幅を広くするかわりにサイドローブの振幅をできるだけ小さくする（分解能は失うが外のスペクトルの影響はほとんど受けない）、(2)メインローブの幅をできるだけ狭くしサイドローブの振幅はそこそこにする（分解能を確保した上で外のスペクトルの影響を抑える）、の2つの考え方がある。



最も強力なウィンドーは Thomson(1982)が実用化した prolate ウィンドーシリーズと考えられている。箱形ウィンドーの周波数領域での関数形から(1)の考え方に沿うように決められたのが 4π prolate ウィンドーで、(2)の場合が π prolate ウィンドーである。それぞれのウィンドーは固有ベクトル空間を成すが、最もバイアスカットに優れている0次の 4π ウィンドーと π ウィンドーを第2図に示す ($N=256$ の場合)。これらのウィンドーの特徴を、歴史的に有名なハニング、ハミング各ウィンドーを例にして比較すると、分解能の良さでは π ウィンドー \approx ハニング/ハミングウィンドー $>$ 4π ウィンドー、サイドローブの影響カットの良さでは 4π ウィンドー $>$ π ウィンドー $>$ ハニング/ハミングウィンドー、となる。その他のウィンドーとの比較は Percival&Walden(1993)に詳しい。Thomson はさらに、数次のウィンドーをそれぞれ時系列に作用させて重ね合わせる multiple windowing によって、ウィンドーの効果がより優れたものになり、かつメインローブの幅（ガード周波数）を解析者が設計できることを示している。

本研究では、時系列を多数の小区間に分けてそれぞれの区間のスペクトルを平均操作する場合には0次 π ウィンドーを用い、時系列全体を1区間として扱う場合には多重 4π ウィンドー（8次まで）を用いることにする。

3、異常値の除去

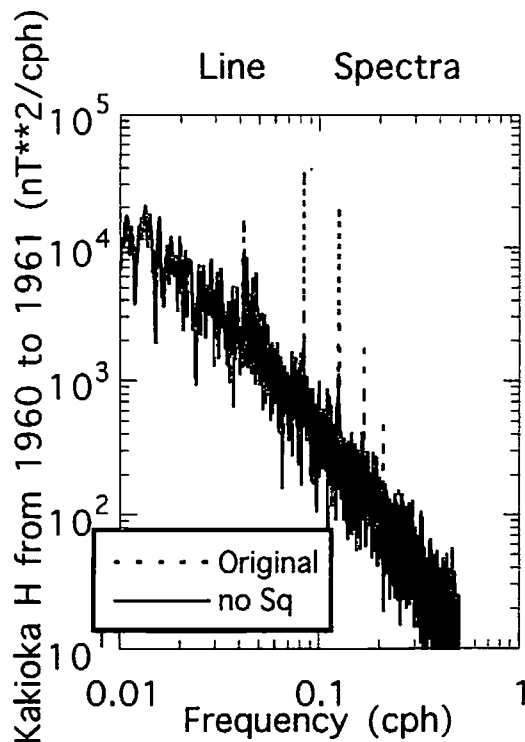
ここでいう異常値は、スパイクや階段状ギャップなど時間領域で検出・訂正がしやすいもののことである。あとでロバスト演算をする場合でも、明らかな異常値は最初に除いておいたほうがよい。というのは、振幅が著しく大きい異常値が一つでもあると、ロバスト判定の基になる最初の応答関数の計算が大きく影響を受けてしまい、その後の異常値の判

定の手続きが正しい方向に行かなくなる可能性があるからである。

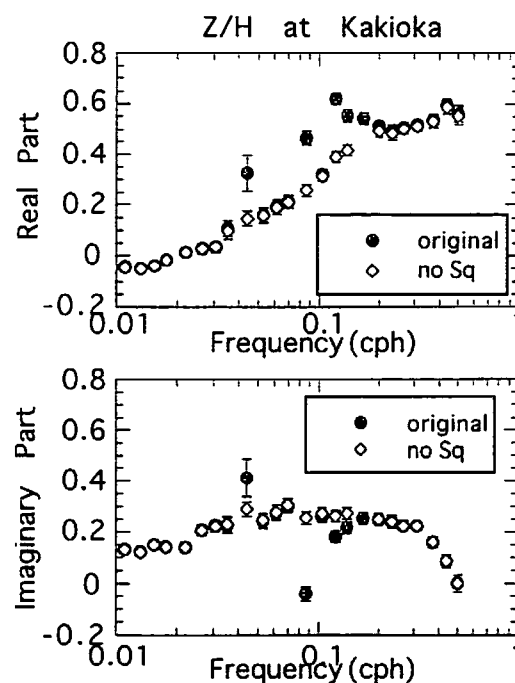
異常値の検出法には、(1) 近くの観測点との比較、(2) その観測点の過去の値との比較、などが考えられる。地球磁場観測ネットワークのデータ (のべ 241 観測点での 1901-1995 年の磁場 3 成分の一時間値) を (1) の方法で検証したところ、一般に中・低緯度では 100nT 以上の異常値は検出できることがわかった。内訳は、スパイク、階段状ギャップに加えて、人為的異常値 (ミスタイプ、符合まちがい、フォーマット変換ミス等)、原因不明のものも多かった。

4、ラインスペクトルの除去

第 1 図からわかるとおり、離散フーリエ変換は、特定の周波数 ω_0 に極在するシグナルを求めるのに相応しい方法ではない。特に極在するシグナルがバックグラウンドに対して著しく大きな振幅を持っていた場合 (いわゆるラインスペクトル)、離散フーリエ変換ではメインローブ内に ω_0 を含む周波数域で本来の値から大きくずれたスペクトルが得られる。また、サイドローブを通じて離れた周波数のスペクトルの計算にも影響を与えるが、その振幅が大きいだけに被害も甚大である。



第 3 図



柿岡の地球磁場の 1 時間値を使って、Sq・潮汐を除いた場合とそのままの場合の違い

をパワースペクトル（Hのみ）と応答関数 Z/Hについて第3図に示す。パワースペクトルから 24h、12h、8h、6h、4.8h にラインスペクトルがあることがわかるが、1年分のデータを一区間で使っているので分解能が高く周辺に対する影響は少ないように見える。しかし、同じデータを使って一区間の長さを 454~14 時間に変えながら Z/H を求めると、ラインスペクトルを事前に除くかどうかで広い範囲に渡って顕著な差が出る。ラインスペクトルを除いたところで、その周波数ではもはや応答関数の計算を行うことはできないが、周囲の計算値が影響を受けるのを防げることがわかる。

ラインスペクトルの存在を検出するため、Thomson は F テストを提案している。これは F 検定により、ある周波数帯の平均レベルからスペクトルが大きすぎてどうかを判定する。周波数 ω_0 がわかれば時間領域でデータに対する $A\sin(\omega_0 t + \theta)$ の最小自乗解を計算して、それをデータから引くことでラインスペクトルを取り除ける。第3図で使ったのは、このようにして補正したデータである。バンドパスフィルターを適用することもできるが、残った時系列に与える影響が小さいのは最小自乗法だろう。また、 ω_0 周辺の周波数帯を使わないことでラインスペクトルの影響を除くこともできるが、第3図からわかるようになりに広い範囲でスペクトルを失うことになり、最善とは言えないと思われる。

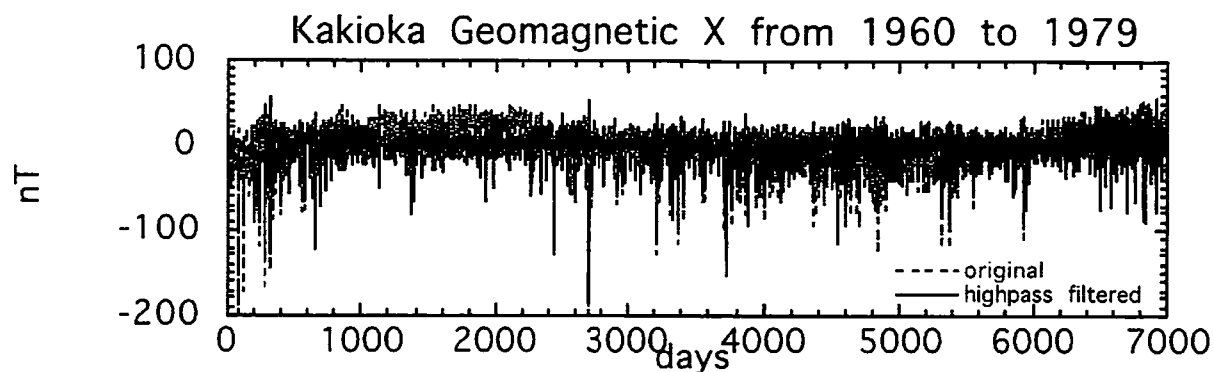
地球電磁場の周期帯 1 時間から 2 年に含まれるラインスペクトルには、Sq とその高調波（24h、12h、8h、6h、4.8h、4h、3.4h、3h）、潮汐、年周、半年周がある。27 日とその高調波はラインスペクトルとはいえないようである。こうしてみると、ラインスペクトルの原因となる物理現象はバックグラウンドとは違っており、ソースの波長も違っていることがわかる。異質のものを混ぜないという意味でも、ラインスペクトルの除去は重要だろう。

Sq については、振幅が季節変化等することを考慮し周波数 ω_0 に加え $\omega_0 \pm 2\pi/(1\text{year}/n)$ も除くとよいようである。n は、24 時間のスペクトルの場合だと 1 から 7 くらいまで同定できる。また、Sq は太陽活動の静穏日の現象なので、磁気嵐を除いて計算したほうが結果がよかった。我々はロバスト最小自乗を適用することで、事実上の磁気嵐はずしを行った。

5、ローパス/ハイパスフィルター

低周波のシグナルだけに興味がある場合、時系列にある種のローパスフィルターをかけてリサンプリングすると便利である。我々の興味は周期 2 日以上に応答関数なのでカットオフ周期を 2 日にして 1 時間値を 1 日値にリサンプリングした（第4図の点線）。World Data Center 等で公開されている日平均を使わなかったのは、（1）24 個の 1 時間値を平均して作った日平均では Sq が正確に 24 時間周期でないことの影響が残る、（2）単純平均より周波数領域でのふるまいを設計できるフィルターのほうがフィルタリングの影響をコントロールしやすい、という理由からである。また、最終的に 1 日値にするにもかかわらず Sq や潮汐を先に取り除いたのは、振幅の大きいラインスペクトルがカットオフ周波数近辺にあるとフィルタリングの際にエイリアジングを起こす可能性が高いからである。

デジタルフィルターは、巡回形 (recursive filter) と非巡回形 (nonrecursive filter) に分けることができる。よく使われる斎藤フィルター (1978) は巡回形で、ある時刻 t のデータ $x(t)$ をフィルタリングした値 $y(t)$ を計算するのに $x(t-\tau)$ だけでなく $y(t-\tau)$ を用い、フィルターの長さを短くしている。そのかわり位相が回るのを逆向きにもう一度フィルターをかけることによって補正する。ただ、時系列の両端はうまく補正できないから、適当な長さのデータを捨てねばならない。非巡回形フィルターでは $x(t-\tau)$ だけを使うので位相の回転は起こらないが、巡回形よりもフィルターが長くなる。



第4図

我々ここでは非巡回形フィルターの一種(Hamming, 1977)を使った。巡回形では両端の使えない部分の長さを試行錯誤して決めなければならず、大量のデータがあっいちいち確認するのが面倒な時は非巡回形のほうが無難と思われる。非巡回形フィルターが長いことで失うデータの量と、巡回形フィルターが位相を回転することで失うデータの量は、経験的には大差ない。

興味のある周波数帯の外側に振幅の大きいシグナルがある場合、その影響をおさえる処理をする必要がある。これは prewhitening と呼ばれる。平均値やトレンドを取り除くのもバンドパスフィルターをかけるのも prewhitening の一種である。我々は地球磁場の周期帯 2 日から 2 年に興味があるが、時系列データのプロットを見ると 2 年より長周期の変化も存在し特に 11 年周期は大きな振幅を持っていることがわかる。ここでは、カットオフ周期 128 日のハイパスフィルターをかけた (第4図の実線)。

他の手段としては、 4π ウィンドーを使うことでガード周波数以外の影響をカットする、AR フィルターを使う、などがある。これらが採用されなかった理由は、地球磁場にはソースのメカニズムが違う異質なスペクトルが混在することが多いのである程度の周波数分解能は確保したほうがよいと判断したこと、実際の時系列が完全には周期的でないために AR フィルターによって位相が少しずれることがあるとわかったこと、の2点である。

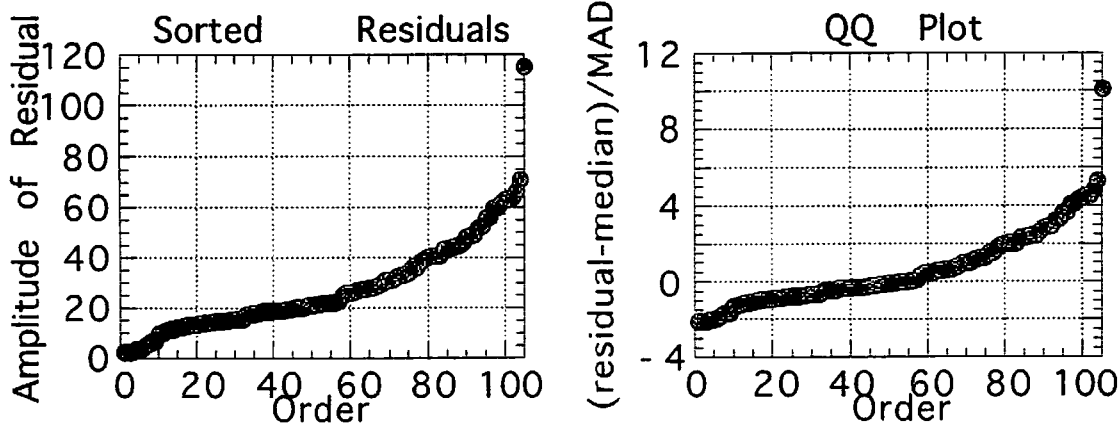
6、ロバスト推定

異常値の影響をうけないように値を推定する方法をロバストという。例えば、平均値は

ロバストな推定値ではないが、中央値はロバストな推定値である。単純な最小自乗は平均のようなものであるから、そのようにして求められた応答関数は異常値の影響を受けて真の値から大きくずれている可能性がある。

ロバスト推定にはいろいろな流儀があるが、ここでは我々が使った方法(cf. Chave et al,1987)について述べる。その他の方法の詳細は Huber(1981)などを参照されたい。ハイパスフィルター処理された磁場 2 成分から計算した周波数 ω の N 個のサンプル、余緯度成分(X_1, X_2, \dots, X_N)と鉛直成分上向き成分(Z_1, Z_2, \dots, Z_N)を使って、応答関数 Schumucker $C (=0.5R \tan \theta B_r / B_0, R$:地球半径, θ :観測点の緯度)を求める場合を想定する。

第5図



まず問題になるのは異常値の検出法である。観測値に含まれるノイズが白色雑音ならば、観測値と予測値の残差($\epsilon_1, \epsilon_2, \dots, \epsilon_N$)は正規分布に従い、残差の振幅を小さい順に並べると直線上にのることがわかっている(第5図 a)。よって、直線から大きくはずれた残差を検出することにする。判定基準を定量化するため小さい順に並んだ残差の振幅から中央値 ϵ_{med} を引いてばらつき σ_{robust} で規格化し、一定の値を越えたものを異常とみなす(第5図 b)。この時、平均値と標準偏差ではなく、ロバストな推定値 ϵ_{med} と σ_{robust} を使う。しきい値の大きさは第5図 b だと最大でも6だろう。経験的には大きさ3~12を採用している。

σ_{robust} には Median Absolute Deviation (MAD) と Inter Quantile の2種類があるが、MADのほうが、扁平な分布の場合にも安定して計算でき、よりロバストだと思われる。MADは残差の振幅を ϵ_{med} からの距離に変換して0を除き小さい順に並べたときの中央値、Inter Quantileは第5図 a での $3N/4$ 番目と $N/4$ 番目の差で定義される。

異常値が検出されたとして、それに対する対処法は大きく分けて2つある。異常値を以降の計算から取り除く hard rejection と、重みをかけて影響を軽くして使う soft rejection である。どちらが良いか一概には言えないが、我々は、球関数展開のように観測点を失うと展開次数に響く場合には soft rejection、応答関数の計算のように同じものを表現するデータの組が多数ある場合には hard rejection というように使い分けた。

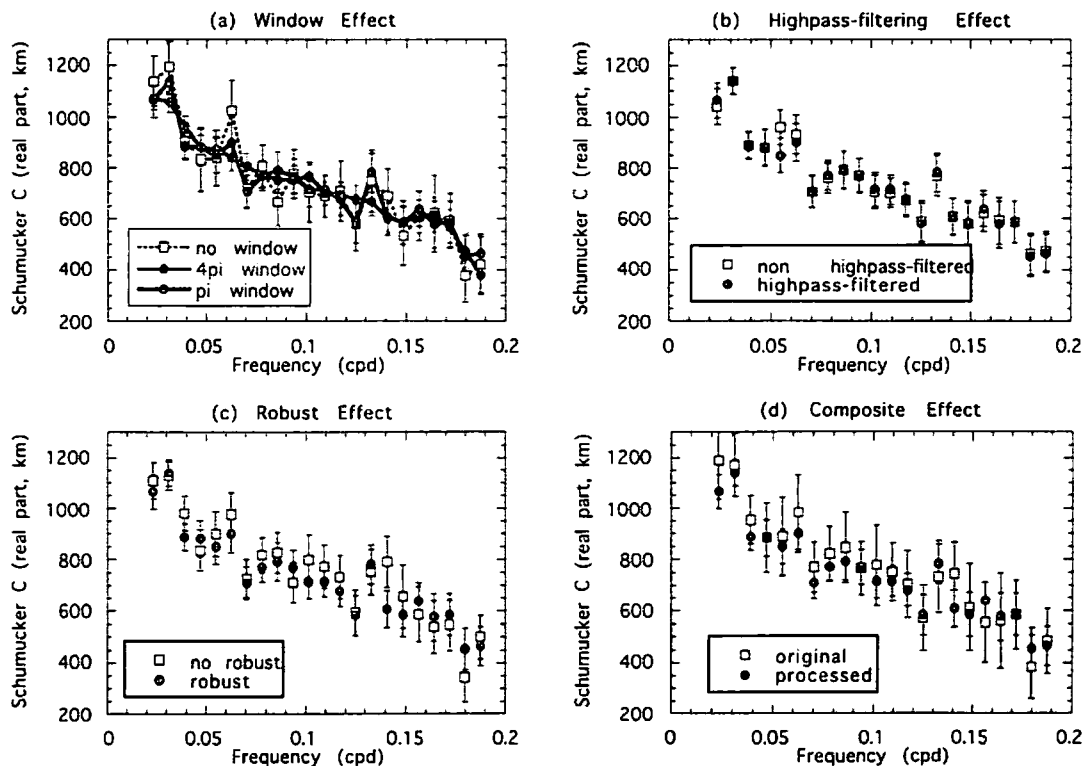
柿岡のケースでは、一度目の判定(第5図)で異常とみなされたのは激しい磁気嵐(振

幅 100~200nT) を含む区間であった。10 nT の矩形状の異常を人為的に2箇所挿入してみたが一度目では全く見えず、小さな異常を検出するためにはロバスト推定の繰り返しが必要なことがわかった。

7、応答関数に現れるデータ処理の違い

データ処理の各過程における選択の違いがどのように応答関数 C に現れるか計算する。ラインスペクトルの影響は第3図で見たので、ここでは(a)ウィンドーによる違い、(b)ハイパスフィルターの有無の違い、(c)ロバストの適用の有無の違い、(d)複合要因、について標準計算値 (●) との比較を示す (第6図)。標準計算値は、ラインスペクトルを除いてハイパスフィルターをかけた1日値 (第4図の実線) を128日の小区間に分け、それぞれに0次 π ウィンドーを作用させてFFTをし、hard rejection ロバスト推定を複数回適用して最小自乗的に求めたものである。比較する計算値は、(a)~(c)では一つだけ標準と違う処理をしたもの、(d)では1時間値に手を加えずにローパスフィルターをかけ1日値にしたものからウィンドーもロバスト推定も使わずに求めたもの、である。エラーはいずれもジャックナイフ法で計算した。

第6図



第6図(a)では、ウィンドーを使わなかった場合 (□)、0次 4π ウィンドー (薄い黒丸)

を使った場合と標準の計算値を比較した。ウィンドーを使わないと周波数領域で計算値がばたつくこと、図ではわかりにくいエラーバーの長さも有意に大きく、各区間の計算値もばたつくことがわかる。薄い黒丸と●の比較からは、分解能の違いが周波数領域での滑らかさの違いとなって現れているのがよくわかる。ここには示さなかったが、ハニングウィンドーを使うと、ウィンドーなしの場合より滑らかでエラーバーの小さい計算値が得られる。標準と比べると、エラーバーの大きさはほとんど同じで、応答関数の値は長周期で大きめ短周期で小さめとなりサイドロープの影響が推察される。

第6図(b)からは、ハイパスフィルターの影響は長周期に限られること、ハイパスフィルターを使うと長周期でもエラーバーを小さくできること、が見て取れる。これは、例えば11年周期のどの部分を含むかによって区間のトレンドが変わり、長周期のスペクトルの計算に影響がでると説明できる。

柿岡のデータが世界有数の質の高さをもつことを考えると第6図(c)は興味深い結果となった。ロバストで除かれたのが磁気嵐の区間であること、27日とその高調波に影響が出ていないことも考慮すると、磁気嵐の扱いを慎重に考える必要を示唆していると思われる。

第6図(d)では、本研究で適用された処理はエラーバーを有意に小さくする働きがあること、データ処理の違いだけで時にはエラーバーの範囲外にでるような大きな差が作り出せることが示されている。感度の悪い応答関数の場合、エラーバーを小さくすることは最適構造のモデルに非常に大きな制約を与えることになり、十分検討に値すると考える。

第6図のいずれの場合でも、27日とその高調波の周期ではバックグラウンドスペクトルの計算ができていないことが見てとれる。4 π ウィンドーの場合では平滑化されて振幅が減らされているが、これはバックグラウンドスペクトルが求められたのとは違う。27日問題は多数の観測点を使った解析で取り扱うことを考えており、続編で論じる予定である。

謝辞：この論文は著者の一人（藤井）が A.D.Chave, D.J.Thomson 両博士から学んだことが基本となっている。この場をかりて厚く御礼申し上げます。

参考文献：

Thomson, D.J., Spectrum Estimation and Harmonic Analysis, Proc. IEEE, 70, 1055-1096, 1982

Percival, D.B. and A.T. Walden, Spectral Analysis for Physical Applications, Cambridge University Press, Cambridge, 1993

斎藤正徳、漸化式デジタル・フィルターの自動設計、物理探鉱、31、240-263、1978

Hamming, R.W., Digital Filters, Prentice-Hall, Englewood Cliffs, 1977

Chave, A.D., D.J. Thomson, and M.E. Ander, On the Robust Estimation of Power

Spectra, Coherences, and Transfer Functions, J. Geophys. Res., 92, 633-648, 1987

Huber, P.J., Robust Statistics, John Wiley, New York, 1981