

研究ノート

連続波長データへの因子分析法の適用の試み

醍醐元正・落合史生
(同志社大学経済学部教授)(帝塚山大学人文学部教授)
村松加奈子・藤原昇
(奈良女子大学理学部助教授)(奈良女子大学理学部教授)

はじめに

筆者等はこれまでパターン展開法(PDM)[藤原 96]と名付けた多波長衛星画像データを解析する手法を研究開発して来たが、この手法は広い範囲の波長について適用可能である事が判っている[醍醐 99, Muramatsu 00]

PDMは実体的な意味を与え得る座標軸を構成する為に斜交座標を利用している。同じように斜交座標を利用している多変量解析の手法に因子分析法がある。

この因子分析法を衛星画像データに適用して、PDMの基本スペクトルパターンに類似の因子を抽出する事が可能である事を前回に報告した[醍醐 01]。この事から意味ある座標軸を利用する為に斜交座標を利用するというPDMの手法が他の解析手法によっても支持されたとと言えるであろう。

PDMは350 nm–2500 nmの波長領域に対して10 nm分解能で適用可能である事が判っているが[醍醐 99]、前回の報告で因子分析に使用したデータはLANDSAT-5/TMの6バンドデータである。そこで今回は連続波長データに対して因子分析法の適用を試みたので、その結果について報告する。次節で使用したデータの概要、そして解析結果について述べる。なお解析手法は前稿と全く同様であり、また因子分析法アルゴリズムの詳細についても前稿と同様にふれないので、詳細については前稿[醍醐 02]やその参考文献[芝 79]等を参照されたい。

使用したデータ

今回使用したデータは、放射分光計により地上で測定された反射率データである。使用した放射分光計はField Spec FR (Analytical Spectral Devices Inc.)とMSR 7000 (Opto Research Corp.)である。どちらの放射分光計も測定波長域は350 nm から 2500 nm で、1 nm 毎のデータを出力するが、実際の分解能は3 nm から 10 nm である。

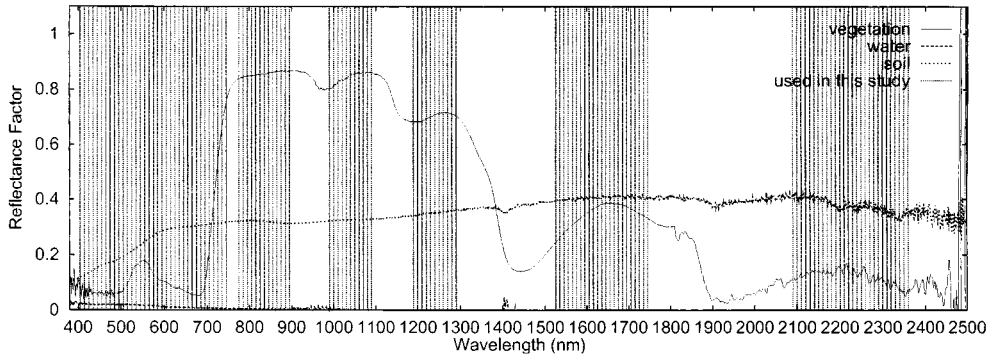


図1 この解析で使用したチャンネルを縦線で表示してある。三つの規格化されていない標準スペクトルパターンも同時に表示してある。

実際の衛星画像データでは波長分解能は 10 nm 程度である為、今回は放射分光計のデータを 10 データ点毎に平均して分解能 10 nm のデータとして使用した。またこうする事により、解析結果は個々の放射分光計の性能の差にそれほど依存しなくなると言う利点もある。

一つのサンプルからと標準白板 (Labsphere Inc. の Spectralon Reflectance Target) からのスペクトルを交互に測定し、サンプルのスペクトルと白板のスペクトルの比をもってそのサンプルの反射スペクトルとした。

サンプルは室外では太陽光で、室内ではハロゲンランプを使用して測定された。サンプルと受光器のあいだは約 50 cm の間隔をあけた。太陽光を使用している為、今回は解析するデータ領域を大気中の透過率が 80% 以上の波長領域に限っている。

最終的に図 1 に示す様な 10 nm 幅の 121 バンドのデータ領域を選択した。図には PDM の三つの標準スペクトルも描かれている。この解析に使用したデータのサンプル数は 1068 であった。

解析結果について

今回の解析では地上での測定データを使用するのであるから Rayleigh scattering 分を差し引く必要はない。また抽出した因子を比較する対象は絶対値和で規格化された PDM の標準スペクトルパターンである (図 2)。そこで今回はデータを絶対値和で規格化してそれに因子分析の手法を適用する事にした。

U_{ij} を計算してその固有値、寄与率と累積寄与率を計算すると表 1、又グラフ化すると図 3 の様になった。但し 10 番目以降の固有値はその寄与率があまりに小さいので表にも図にも表していない。前稿と同様に今回も因子数は 3 とし以降の解析を行った。

解析結果の因子パターンを図 4 に示す。図には $\alpha = 0.0, 0.5, 0.7, 0.8, 0.9, 1.0$ の場合を示した。これらの図を見ると α が大きくなるとある程度各因子の独立性が出て来るが、 $\alpha = 1.0$ に

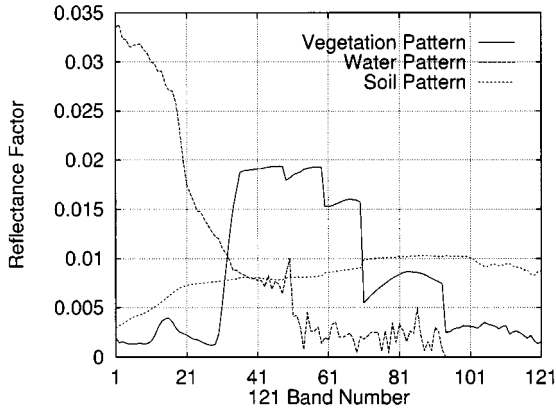


図2 絶対値和で規格化した PDM の標準スペクトルパターン

表1 固有値，寄与率と累積寄与率

	固有値	寄与率	累積固有値
1	0.011134357	0.93670	0.9367
2	0.000393036	0.03306	0.9698
3	0.000222909	0.01875	0.9885
4	3.46 E-05	0.00291	0.9914
5	2.15 E-05	0.00181	0.9932
6	1.22 E-05	0.00103	0.9943
7	1.09 E-05	0.00091	0.9952
8	5.36 E-06	0.00045	0.9956
9	5.08 E-06	0.00043	0.9961
10	4.26 E-06	0.00036	0.9964

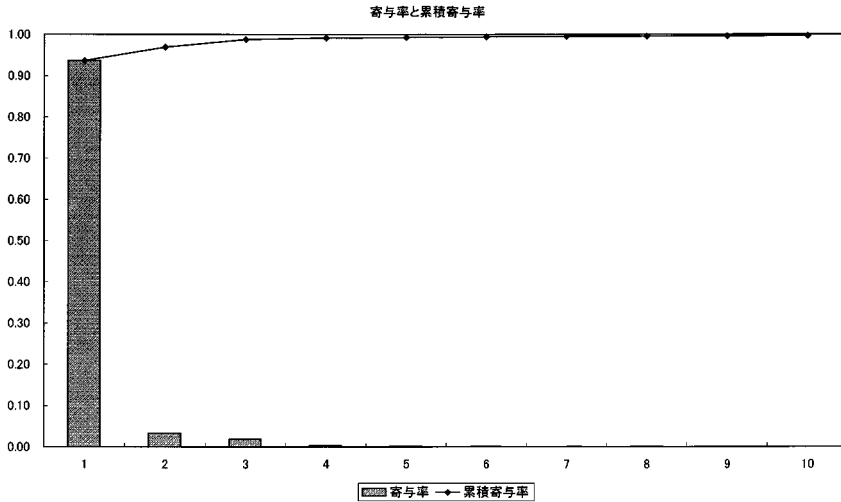
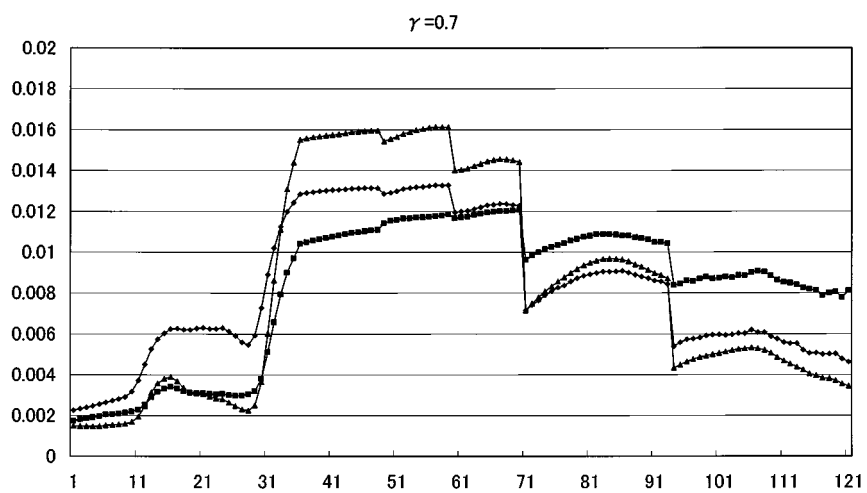
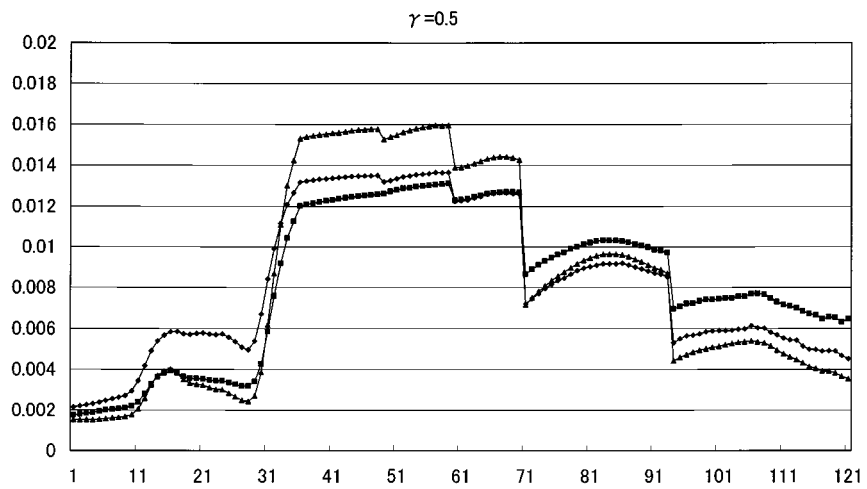
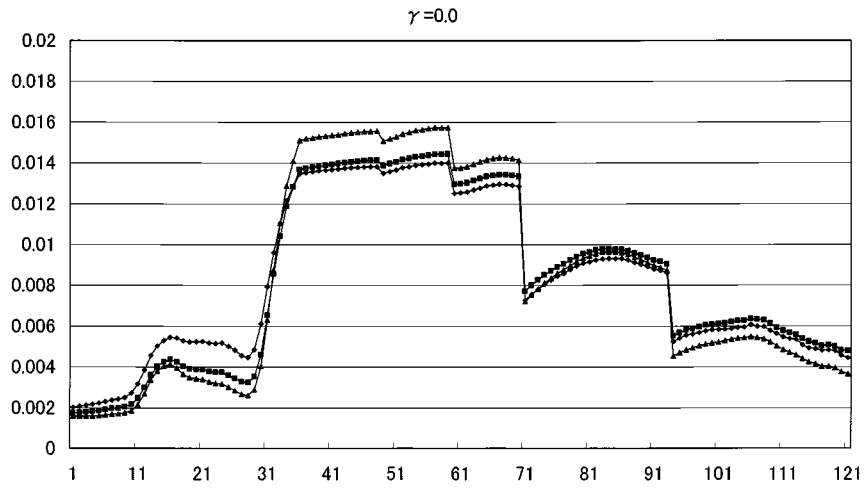


図3 寄与率と累積寄与率

なってもまだ不十分である様な印象を受ける。

LANDSAT データでは が 0.7 程度以上になるとあまり因子は変動しなくなるが、121 チャンネルのデータでは が 0.7 から 1.0 になるに従って因子も変化している。また PDM の水の標準パターンは短波長域になるにつれて反射率が増加して行くが、今回のデータでは 21 チャンネルあたりから短波長域にかけて反射率が減少して行く。また水域・土壌域両因子とも植生域の因子に引きずられて 31 チャンネルあたりで大きく値を変化させているのである。これから見る限り水域・土壌域の因子は植生域の因子に引きずられており、その独立性が前稿の場合より小さくなっているという印象を受ける。

抽出された因子は因子間の独立性が小さく、PDM の基本スペクトルパターンと大きく異なる、特に水と土壌のパターンが異なると見えるのであるが、この問題の原因としては手法の性質とデータの性質と言う二つの理由が考えられる。手法の性質として考えられるのは、オブリ



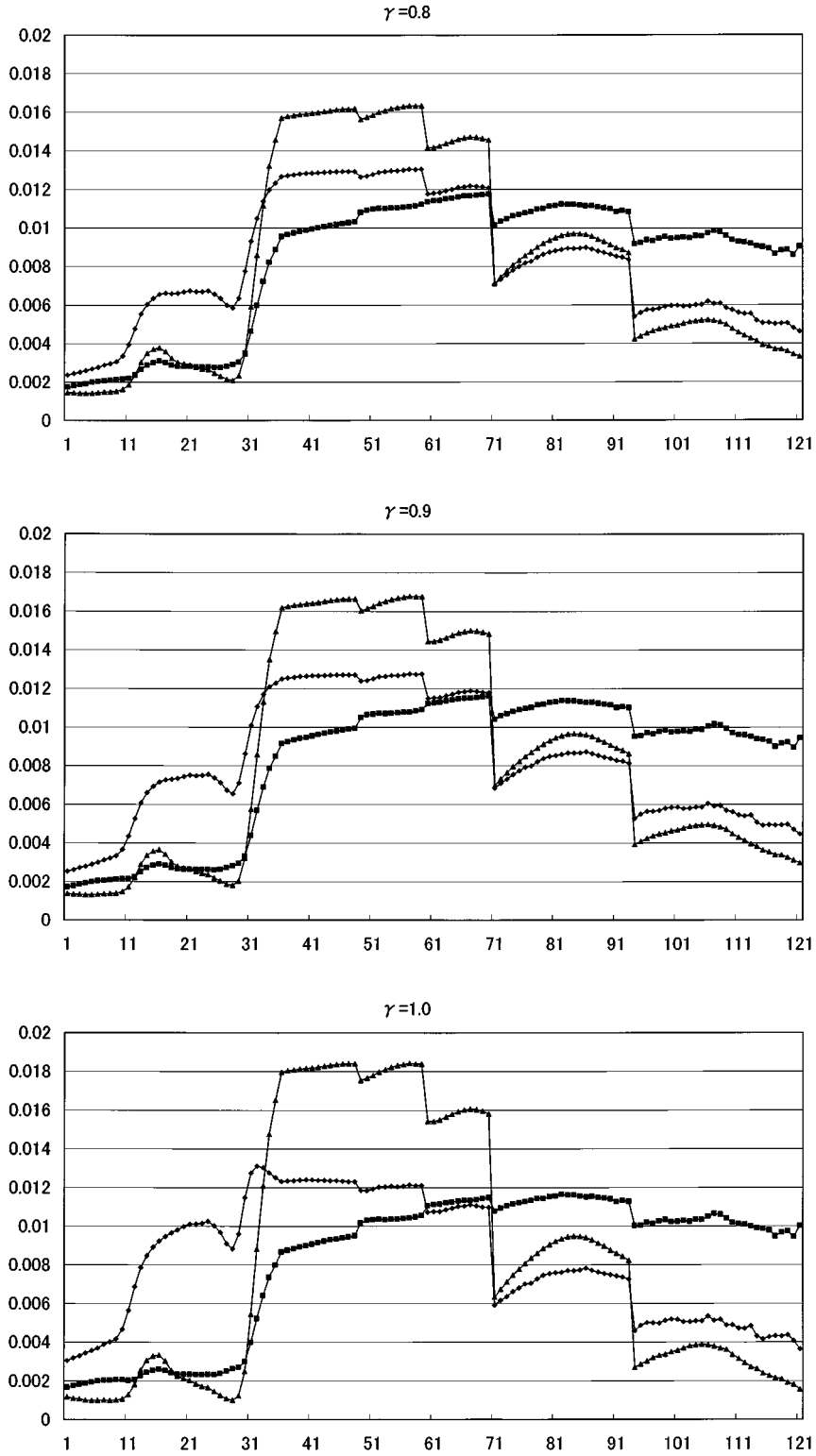


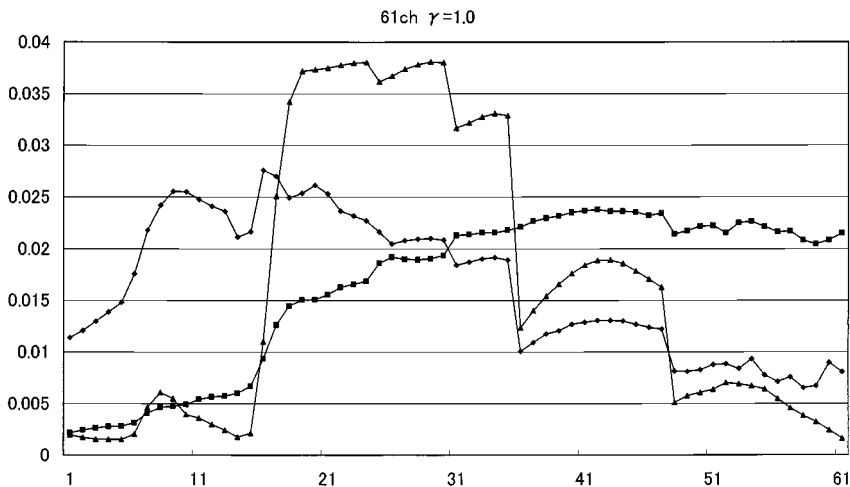
図4 121チャンネルデータを因子分析した結果。 γ は0.0, 0.5, 0.7, 0.8, 0.9, 1.0。

ミン法は扱う自由度が大きくなると因子間の独立性が失われていくのではないかという疑いである。データの性質というのは、我々は主に植生域の研究に重点をおいているので、今回のデータには植物の葉からの反射スペクトルが最も多く含まれているという問題である。特に水域のデータは実験室での測定は不可能であり、室外での測定も大変困難で、水域として区分されるデータのサンプル数は15しかない。

因子間の独立性が小さく見える原因を探る為に、先ずデータのチャンネル数を減らして行ってその影響を見る事にした。使用するチャンネルを、1チャンネルおき、2チャンネルおき、4チャンネルおき...と間引いて61, 31, 16, 8チャンネルのサンプルデータを作り、それを因子分析法にかけてその結果を見た。図5に結果を示すが、これから見る限り8チャンネルまでバンド数を減らしてもやはり因子間の独立性は小さく、水域・土壌域の因子は植生域の因子の影響を受けている様に見える。

そこで、次に水域の15個のデータをそれぞれ9回ずつ使用して、少ない水域のデータに重みをつけて解析を行った。結果は図6に示す様に因子間の独立性は回復しており、特に31チャンネル辺りで植生域因子に引きずられて大きく値を変化させていた水域・土壌域の因子は、なだらかな変化に改善されている。またこの解析では水域のデータだけに重みをつけたのであるが、土壌域の因子の形もそれにつられてかなりの改善をみている。この事から因子間の独立性の問題はデータサンプルの偏りに問題があるのであり、特に水域のデータ量を増やす必要があると言う事が判る。

しかし水域因子が短波長域で小さな値を取るという問題は依然として存在する。そこでこのデータサンプルに含まれている水域データを実際に眺めてみる事にした。15の水域スペクトルデータを図示したのが図7である。これを見ると実際のデータもやはり短波長域で小さな反射率を持っており、因子はそれをよく反映しているという事が判る。即ち手法には問題はなく、データに問題があると言う事である。



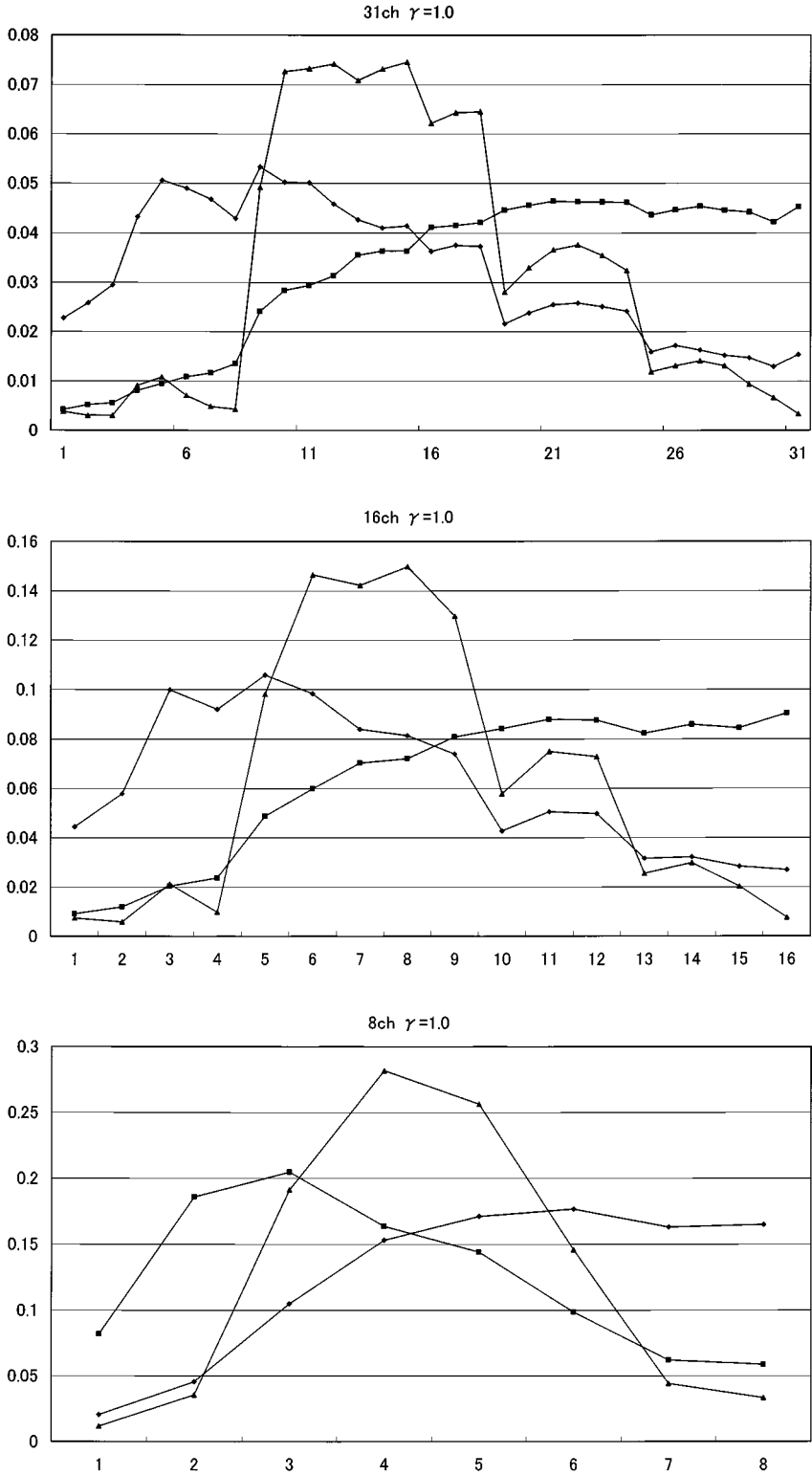


図5 チャンネル数を間引いて61 ch, 31 ch, 16 ch, 8 ch と減らしていった結果の因子。 $\gamma = 1.0$ 。

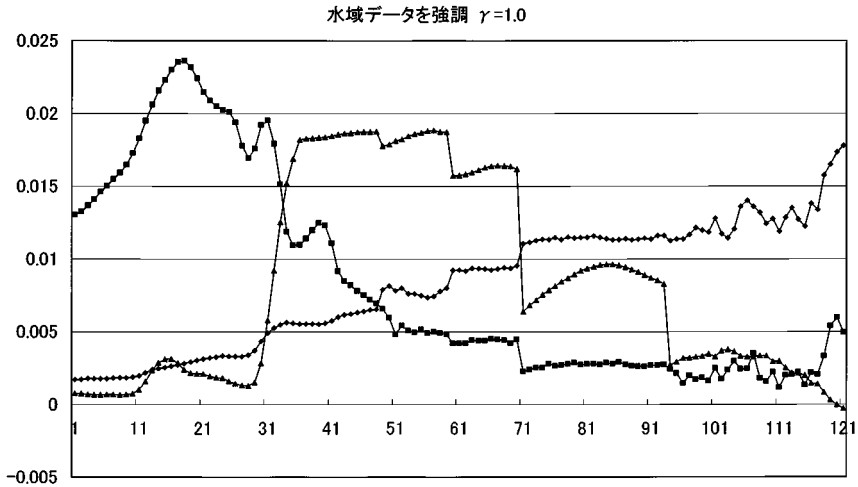


図6 水域データに重みをつけて解析した結果の因子。 $\gamma=1.0$ 。

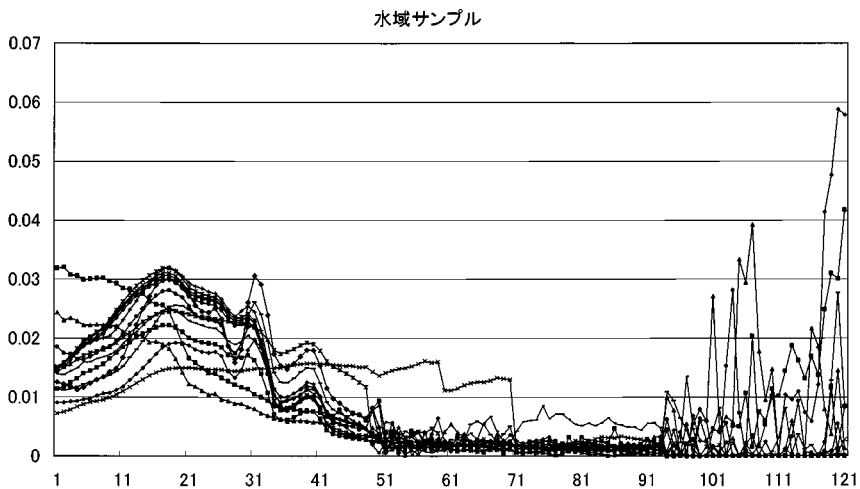


図7 今回使用した水域データの反射スペクトル。絶対値和で規格化してある。

水域のデータになぜこのような問題点があるかという点、それは水域データの測定が大変困難だからである。衛星高度からの水域データと言うと、それは普通、海等の深度の大きい所からの反射スペクトルである。そうでなければ水の底からの反射光が混じって純粋な水からの反射スペクトルとは言えなくなってしまふ。しかし地上でその様なデータを取るのは大変困難な事は容易に想像がつく。実験室では先ず不可能であり、屋外でも普通の海岸や湖岸ではやはり水深は十分ではない。港の栈橋やボート上から放射分光計の受光部を水面上に突き出す様にして測定する等の工夫が必要となるのである。実際、図7には二つだけ短波長域において反射率が大きいデータがあるが、それは和歌山県加太港の栈橋においてその様にして測定されたデータである。また、PDMの水域標準スペクトルパターンとして採用しているデータもこの加太港での測定データである。

ま と め

今回の解析により、121チャンネルという自由度の大きいデータに対してもオブリミン法は有効であろうと言う事が予想出来る。ただ今回はデータに問題があって、きれいな因子を得る事は出来なかった。しかし上にも述べた様に放射分光計で多数の水域データを測定する事は困難であり、今回の解析を更に進めるには航空機搭載の多波長放射分光計からのデータを使用する等、何らかの手段を考える必要があると思われる。

参考文献

- [Muramatsu 00] Muramatsu, K., Furumi, S., Fujiwara, N., Hayashi, A., Daigo M., and Ochiai, F., (2000) " Pattern decomposition method in the albedo space for Landsat TM and MSS data analysis ", INT. J. Remote Sensing, Vol. 21, No. 1, pp. 99-119.
- [芝 79] 芝祐順, (1979) 『因子分析法 第2版』東京大学出版会.
- [醍醐 99] 醍醐元正 (1999) 「連続波長データに対するパターン展開法の適用可能性」『同志社大学経済学論叢』第50巻第4号, pp. 138-149.
- [醍醐 02] 醍醐元正 (2002) 「因子分析法とパターン展開法」『同志社大学ワールドワイドビジネスレビュー』第4巻第1号, pp. 1-12.
- [藤原 96] 藤原昇, 他 (1996) 「衛星データ解析のためのパターン展開法の開発」『日本リモートセンシング学会誌』第16巻第3号, pp. 17-34.