

主成分解析による新たな地形変動分析法に関する理論的検討

北野利一*・中野 晋**・岡 彰紀***・間瀬 肇****

1. はじめに

漂砂による海底地形変化の現象を把握するために、深淺測量による地形の時系列データに主成分解析を適用した研究は、Aubrey(1979), Aubrey et al.(1980), 加藤・吉松(1984)等、これまで多数発表されている。また近年では、Liang・Seymour(1991), Bosma・Dalrymple(1996)および横木ら(1998)は、現地データに対して複素値による主成分解析を行い、海底地形変動の進行波成分の抽出を検討している。

基礎となる主成分解析は、複雑なデータの特徴を抽出するために工学のあらゆる分野で広く用いられてきている統計処理の1手法で、別名を経験的固有関数法として知られる。海底地形の変動分析のみならず、海岸工学においては、湾の形状や海底地形の影響を受けた現地波浪観測データからのエッジ波の抽出(例えば、加藤ら, 1980), スペクトルの発達特性の解析(Vincent・Resio, 1977)等の幅広い応用例がある。また、橋本ら(1995)は主成分解析の各固有関数の線形独立性に着目し、カルマンフィルターを用いて気圧情報から波浪統計予測を行う際に、主成分解析から得られる気圧データの固有関数を予測式の説明変数に選ぶことにより、線形回帰の多重共線性をうまく回避しており、異色な応用例であるといえる。

海底地形変動の現況把握に主成分解析を適用した従来の研究においては、各モードの固有関数が直交するという性質が強調され過ぎており、地形変化特性の把握に得られた固有関数が十分に活用されていない。極端に言えば、従来の研究では、各軸の固有関数の解析結果を例示しているだけであり、寄与率の大きい第1モードの固有関数を再合成し、元データとの再現性を確認するといった表面的な考察に終わっている。これは、解析により得られる各軸の固有関数は互いに掛け合わせて意味をもち、単独で利用するにはかなりの推測と注意が必要であることが原因である。特に、漂砂による砂移動特性の把

握という観点からは、単に各地点での堆積や侵食の判定をするだけでは不十分であり、その時間スケールや空間領域での伝播経路について明確にする必要がある。

本研究では、従来の主成分解析の理論に新たな数式表現を追加することにより、既往理論において陽的に表現されていない固有関数の特性を明らかにし、これまでの主成分解析を用いた研究で利用されていない固有関数の特性をより有効に利用する手法を提示する。

2. 主成分解析の理論的性質

2.1 主成分解析の理論骨子

従来の主成分解析の定式化においては、実際の数値データ処理を目的としているため、離散量のベクトルや行列が用いられる。つまり、解析対象となるデータは、離散化された時間や空間上の観測値で、行列で表現され、そのデータ行列の転置行列とデータ行列とを掛算して得られる相関行列の固有値・固有ベクトルを算出することにより数値的に主成分解析が行われる。本研究の理論考察で用いるデータは、時間的にも空間的にも有限な区間 $(x, t) \in [0, L] \times [0, T]$ で定義される連続関数とする。

いま、空間上での2地点 (x, y) の相関関数を

$$R(x, y) := \langle \eta(x, t), \eta(y, t) \rangle \dots\dots\dots (1)$$

と定義する。相関関数の固有値、固有ベクトルは以下の条件を満たすものである。

$$\left. \begin{aligned} \langle R(y, x), e_p(x) \rangle &= \lambda_p e_p(y), (1 \leq p \leq M) \\ \langle e_p(x), e_q(x) \rangle &= \delta_{pq} \end{aligned} \right\} \dots\dots\dots (2)$$

空間軸上での固有関数が決定できれば、その直交性を利用することにより、時間軸上での固有関数は、

$$c_p(t) := \langle e_p(x), \eta(x, t) \rangle \dots\dots\dots (3)$$

により決定される。以上により、関数 $\eta(x, t)$ は次式のような変数分離された M 個の固有関数を重ね合わせた表現を得る。

$$\eta(x, t) = \sum_{1 \leq p \leq M} e_p(x) c_p(t) \dots\dots\dots (4)$$

式(1)で定義される空間2地点での相関関数 $R(x, y)$ を基にした関数 $\eta(x, t)$ の分解は、2時刻 (t, s) での相関関数

$$R(t, s) := \langle \eta(x, t), \eta(x, s) \rangle \dots\dots\dots (5)$$

についての固有値問題を解くことによって、同様の結果

* 正会員 工修 徳島大学助手 建設工学科
** 正会員 工博 徳島大学助教授 建設工学科
*** (株) 西池組
**** 正会員 工博 京都大学助教授 防災研究所

を得ることができる。また、式 (1), (2), (3) および (5) で用いられている内積は、次のように定義される。

$$\left. \begin{aligned} \langle f(t), g(t) \rangle &:= T^{-1} \int_0^T \bar{f}(t)g(t)dt \\ \langle f(x), g(x) \rangle &:= L^{-1} \int_0^L \bar{f}(x)g(x)dx \end{aligned} \right\} \dots\dots\dots (6)$$

ここで、関数が複素値をとる場合、 \bar{f} は f の共役複素数である。式 (6) で表される内積は、つまるところ、時間軸および空間軸での平均操作である。このことから、主成分解析とは、時間平均による各地点の相関行列の固有値問題 (あるいは、空間平均による各時刻の相関行列の固有値問題) である。そこでは、単一軸の相関関数の考察に帰着させるために、2変数のいずれか一方の軸で“平均”して消去するという操作を必要とする。その結果、解析の対象となる関数を変数分離の形式で表現されるということが主成分解析の最大の特徴である。

ところで、主成分解析を成立させるためには、最低2元以上の変数をもつデータを対象としなければならない。なぜなら、上述の主成分解析の理論骨子において、例えば時間変数を固定すると、

$$\left. \begin{aligned} \eta(x) &= \sum_{1 \leq n \leq k} c_n \langle e_n(x), \eta(x) \rangle \\ \langle e_n(x), e_m(x) \rangle &= \delta_{nm} \end{aligned} \right\} \dots\dots\dots (7)$$

となり、相関関数による固有値計算は不要となる。式(6)で定義された内積空間の基底関数は三角関数であるので、式(7)は Fourier 級数展開の定義に縮退する。従って、1変数での解析では、固有関数はデータの特性に基づいたものでなく、内積の定義のみに依存したものとなる。以上の考察から、2変数での解析においても内積の性質を優先させて、解析対象の変動データをあらかじめ Fourier 級数展開しておけば、特定の条件下では、相関関数および固有関数を陽的に表現することが可能になり、理論展開に適すると考えられる。

2.2 定在波と進行波の固有関数

本節では、最も基本的な波である定在波と進行波について、実数値で扱った場合と複素値で扱った場合の主成分解析の計算例を示し、その違いについて考察する。

実数関数の定在波と進行波は以下のように表される。

$$\eta_s^s(x, t) = \cos(kx) \cos(\omega t) \dots\dots\dots (8)$$

$$\eta_f^s(x, t) = \cos(\omega t - kx) \dots\dots\dots (9)$$

各々の波の時間平均による空間2地点の相関関数は、

$$R_s^s(x, y) = \frac{1}{2} \cos(kx) \cos(ky) \dots\dots\dots (10)$$

$$R_f^s(x, y) = \frac{1}{2} \{ \cos(kx) \cos(ky) + \sin(kx) \sin(ky) \} \dots\dots\dots (11)$$

となるので、直ちに定在波の固有関数は以下のように得られる。

$$\{ \lambda, e_s^s(x), c_s^s(t) \} = \left\{ \frac{1}{4}, \sqrt{2} \cos(kx), \frac{1}{\sqrt{2}} \cos(\omega t) \right\} \dots\dots\dots (12)$$

また、進行波の固有関数は、

$$\begin{aligned} \{ \lambda, e_f^s(x), c_f^s(t) \} &= \left\{ \frac{1}{4}, \sqrt{2} \cos(kx), \frac{1}{\sqrt{2}} \cos(\omega t) \right\}, \\ &\left\{ \frac{1}{4}, \sqrt{2} \sin(kx), \frac{1}{\sqrt{2}} \sin(\omega t) \right\} \dots\dots\dots (13) \end{aligned}$$

となり、実数値表現での進行波は2つの固有関数に分離されて、1つの固有関数で表現できないことがわかる。

一方、複素関数の進行波と定在波

$$\eta_f^c(x, t) = \exp i(\omega t - kx) \dots\dots\dots (14)$$

$$\eta_s^c(x, t) = \exp i(\omega t - kx) + \exp i(\omega t + kx) \dots\dots (15)$$

の空間2地点の相関関数は、それぞれ、

$$R_f^c(x, y) = \exp ik(x - y) \dots\dots\dots (16)$$

$$R_s^c(x, y) = \{ \exp(iky) + \exp(-iky) \} \times \{ \exp(ikx) + \exp(-ikx) \} \dots\dots (17)$$

となる。従って、直ちに各々の固有関数は、

$$\{ \lambda, e_f^c(x), c_f^c(t) \} = \{ 1, \exp(-ikx), \exp(i\omega t) \} \dots\dots\dots (18)$$

$$\begin{aligned} \{ \lambda, e_s^c(x), c_s^c(t) \} &= \left\{ 2, \frac{1}{\sqrt{2}} \{ \exp(ikx) + \exp(-ikx) \}, \sqrt{2} \exp(i\omega t) \right\} \\ &\dots\dots\dots (19) \end{aligned}$$

と得られる。以上から、複素値での解析では、定常波も進行波も1つの固有関数で表現できるのに対し、実数値解析では、固有関数が定常波しか表現できないため、進行波は2つの定常波の和として表現されることがわかる。このことは、複素値と実数値で基底関数が異なり、べき乗数に純虚数をとる exp 関数では $\omega t \pm kx$ といった位相の移動を表現できる変数に書き換えることができるのに対し、sin および cos 関数では、1組の空間関数と時間関数の積では表現できないことに起因する。

2.3 Hilbert 変換による複素化

複素表現であれば定在波も進行波も1つの固有関数として得られるため、変動データの主成分解析を行う際は、データを複素化した方が有効である。本節では、実数値変動データの複素化手法について検討する。

Hilbert 変換は、実数軸上で定義された関数を複素平面上に解析接続するための虚数関数を与えるものであり、以下のように定義される。

$$H[f](z) = p.v. \int_{-\infty}^{\infty} \frac{f(y)dy}{\pi(z-y)} \dots\dots\dots (20)$$

ただし、p.v.は積分の主値を表す。三角関数に対しては単

純に、位相を $\pi/2$ だけ進める変換である。対象とする変動関数は時間と空間の2変数の関数であり、いずれの変数に対して Hilbert 変換を適用すればよいかについて検討する必要がある。以下では、実際に両変数に Hilbert 変換を適用して考察する。実数値変動関数を、以下のように空間と時間の2重 Fourier 級数に展開する。

$$\eta_0^-(x, t) = \sum_{m,n} \{ a_{mn}^{cc} \cos mkx \cos n\omega t + a_{mn}^{cs} \sin mkx \cos n\omega t + a_{mn}^{sc} \cos mkx \sin n\omega t + a_{mn}^{ss} \sin mkx \sin n\omega t \} \dots\dots\dots (21)$$

時間および空間軸に対する Hilbert 変換を H^t および H^x と表し、変動関数 $\eta_0^-(x, t)$ を複素化すると次のようになる。

$$\eta_0^+ + iH^t[\eta_0^-] = \sum_{m,n} \{ c_{mn}^{(+)} \exp i(n\omega t + mkx) + c_{mn}^{(-)} \exp i(n\omega t - mkx) \} \dots\dots\dots (22)$$

および、

$$\eta_0^- + iH^x[\eta_0^+] = \sum_{m,n} \{ c_{mn}^{(+)} \exp i(mkx + n\omega t) + c_{mn}^{(-)} \exp i(mkx - n\omega t) \} \dots\dots\dots (23)$$

ここで、複素係数およびは次のようになる。

$$\left. \begin{aligned} c_{mn}^{(+)} &= \frac{1}{2} \{ (a_{mn}^{cc} - a_{mn}^{ss}) - i(a_{mn}^{cs} + a_{mn}^{sc}) \} \\ c_{mn}^{(-)} &= \frac{1}{2} \{ (a_{mn}^{cc} + a_{mn}^{ss}) - i(a_{mn}^{cs} - a_{mn}^{sc}) \} \end{aligned} \right\} \dots\dots\dots (24)$$

得られた2通りの結果を比較すると、時間軸に関して Hilbert 変換して得られた式 (22) の複素値関数は、各成分波の周波数は $+n\omega t$ のみであるのに対し、空間軸に関して Hilbert 変換した式 (23) の複素値関数では $\pm n\omega t$ あり、絶対値が等しく符号が異なる2種類の周波数が現れている。このことは、後者は未来から過去へも時間が流れることを許すことになり、物理的におかしくなる。従って、実数値関数を複素化する際には、時間軸に関して Hilbert 変換することが妥当である。

2.4 一般的な固有関数

時間軸に沿った変動に対して Hilbert 変換を適用することにより、空間と時間の複素表現による2重 Fourier 級数展開した変動関数を主成分解析により固有分解した結果は、形式的に以下のように表現できる。

$$\left. \begin{aligned} \eta(x, t) &= \sum_{1 \leq n \leq M} \lambda_n C_n(t) e_n(x) \dots\dots\dots (25) \\ c_n(t) &= \sum_p a_{n,p} \exp(i\omega_n p t) \\ e_n(x) &= \sum_q \{ \beta_{n,q} \exp(ik_{n,q} x) + \chi_{n,q} \exp(-ik_{n,q} x) \} \end{aligned} \right\} \dots\dots\dots (26)$$

ただし、式 (26) は、実数値固有関数をそのまま主成分解析を行って得られる固有関数を複素化したものとは異なる。また、主成分解析により分解される固有関数の個数 M は、多くても波データを Fourier 級数展開した時の周波数・波数 (正負を別と数える) の個数の大きい方の個数 M_0 であり、さらに、周波数・波数の組み合わせやその振幅比の組み合わせが特殊な場合、 M_0 より減少することが

知られている (北野・間瀬; 1999)。

式 (25) のように表される主成分解析による時間と空間の変数分離関数は、微分方程式の変数分離法とは異なり、時間の固有関数 $c_n(t)$ も単一周期の関数ではなく、一般に固有な (複雑な) 変動関数が得られる。従って、主成分解析は、変動のエネルギーが非定常に変化する波データを対象とするのに向いている。しかし、変動の周期が複雑で変動エネルギーが定常でないため、得られる固有関数をそのまま考察することは非常に困難である。この難点を避ける工夫については、3.での具体的なデータ解析例を通して説明する。また、空間の固有関数は、符合の異なる波数を含む変動関数となる。そこで、正・負の両方向に進行する波を単一方向への進行波

$$e_n^+(x) = \sum_q \beta_{n,q} \exp(ik_{n,q} x) \dots\dots\dots (27)$$

および

$$e_n^-(x) = \sum_q \chi_{n,q} \exp(-ik_{n,q} x) \dots\dots\dots (28)$$

に分離することにより、変動が激しい領域に挟まれて変動がほとんど無いような地点でも、変動の伝播経路の通過点として変動が現れるはずである。つまり、実数値の主成分解析では定在波が各成分関数となるのに対し、複素値による主成分解析では部分重複波の固有関数が得られるので、各成分関数として $e_n^+(x)$ および $e_n^-(x)$ で表される進行波に細分しようとするものである。

3. 主成分解析の理論的性質

3.1 解析例 1

図-1(a)の濃淡図で示す変動データ $\eta^A(x, t)$ は、複素主成分解析により、固有関数 $e^A(x)$ および $c^A(t)$ に分離できる。ここで得られた固有関数は、全変動エネルギーに対する比率が99.3%であり、複素値を動径表示して、図中各軸に平行に表示している。

まず、時間の固有関数 $c_n(t)$ に着目する。比較のため、図-1(b)に実数主成分解析により得られた時間の固有関数 (エネルギー比率95.4%)を示す。本例のように、漂砂による地形変動の時間周期は一般に長く、観測期間に収まらないことが多い。このような場合、実数値の固有関数では変動性は不明であり、複素値を動径表示したものではこまかな変動のため卓越周期が不明瞭となる。また、各軸の固有関数は両者を掛け合わせてはじめて意味を持つものであり、各軸の固有関数にそれぞれ大きさ1の複素数定数 $e^{i\gamma}$ (γ は定数) を乗除することができるので、固有関数の位相は取り扱いに注意を必要とする (固有関数の任意性の問題、北野・間瀬 (1999) を参照)。しかし、 $e^{i\gamma}$ を乗除しても包絡線関数は一致するので、各軸の固有関数を単独で考察する場合には、包絡線関数を利用する。図-1(c)に、時間の固有関数の包絡線関数およびその実・虚関数を示す。包絡線を導入することにより、

変動関数が包絡線関数と接する時間間隔から卓越周期が明確に得られる。この考え方は、北野ら (1997, 1998) で提案されている個別波分割法としてのオービタル定義と同じ意味になっている。なお、各時刻での変動エネルギーも変化するため、平均エネルギーで一定振幅の卓越周期で振動する三角関数 $\bar{c}_n(t)$ を図-1(d) に示す (図中 (c), (d) において、長・短の破線は複素変動の実・虚値を、実線は変動の包絡線を表す)。

空間の固有関数を複素値 FFT 解析により、進行方向に分離した。その結果を図-2(a, b) に示す。主成分解析から直接得られる空間の固有関数 $e_n(x)$ からは、地点 A, B での変動エネルギーが局在することしかわからなかったのに比べ、 $e_n^+(x)$ と $e_n^-(x)$ に分離することにより、 $e_n^-(x)$ による進行波は地点 C で伝播速度の異なる境界があることがわかる。また、各地点での進行方向別のエネルギー分担率を図-2(c) のように図示することにより変

動の空間特性がさらに明確になる。地点 C を境界とし地点 B を含む領域では両進行方向でのエネルギー分布が対称となることが読み取れ、定在波の特性をもつものに対し、地点 A を含む領域では、 e^+ 方向に進行する波は局在性が強く、 e^- 方向に進行する波との波長が異なる。ここでやっている進行方向の分離法は、空間 2 地点での時系列の入反射分離法の考え方を拡張したものになっている。

3.2 解析例 2

図-3(a) の濃淡図で示すようなソリトン包絡線をもつ変動データ $\eta^B(x, t)$ を検討する。主成分により得られる固有関数を図-3(b) に示す。この例では、全変動エネルギーが 42.3, 29.7, 16.6% となる複数の成分に分散している。そこで、本例では、各地点・各時刻の局所的な変動エネルギー $|\eta^B + iH^*[\eta^B]|$ の主成分解析を行った。その結果は、図-3(c) に示すとおり、過半数の変動エネルギーが 1 つの固有関数に集中したものとなっている。変動データ $\eta^B(x, t)$ は非常に単純明快なものであるため、図-3 で表される結果の方が図-2 より適切に変動の特徴を抽出しているといえる。この理由は、式 (6) の内積は全時間および全空間での平均として定義されているため、 $\eta^B(x, t)$ のような局在したエネルギーそのものが伝播するような変動関数に対しては、各地点で異なる時刻にエネルギーが局在するにもかかわらず、各軸上での平均的な変動を捕らえようとするに無理が生じることである。つまり、本例のような変動データは、

$$\eta^B(x, t) = H(x-ct) \exp i(kx - \omega t) \dots\dots\dots (29)$$

で表現される (H は実数値関数) ようなものであり、このような変動関数は、単純な変数分離関数で表現することはできないことが原因である。

以上にとりあげた 2 つの解析例以外にも、多様な主成分解析例が考えられる。例えば、変動関数が複数の成分に分解され、変動エネルギーの等しい固有関数が得られる場合、各固有関数が 1 通りに決定できず、異なる包絡線関数をもつ固有関数の無数の組み合わせに分解される

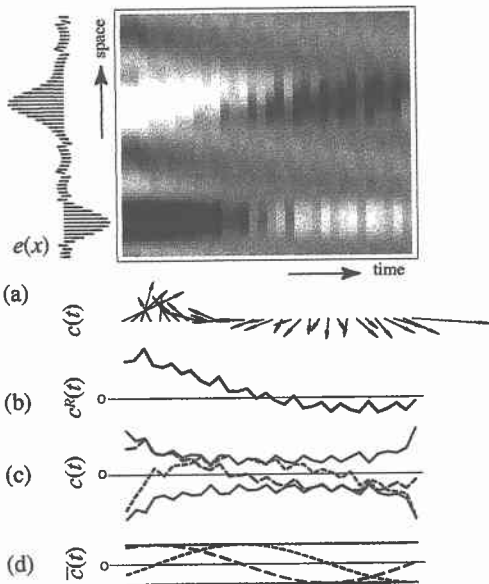


図-1 変動データ $\eta^A(x, t)$ およびその固有関数

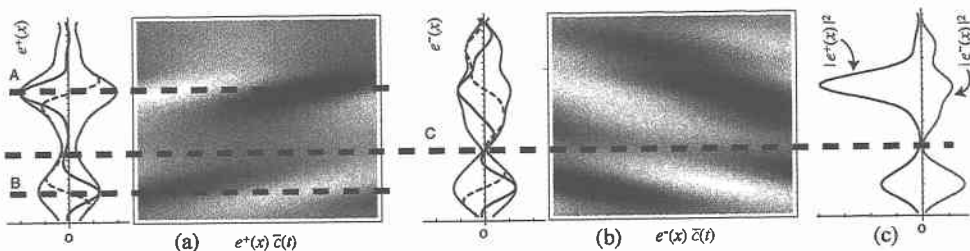


図-2 空間の固有関数 $e(x)$ の進行方向により分離: $e^+(x)$ および $e^-(x)$
(図中 (a), (b) において、長・短の破線は複素変動の実・虚値を、実線は変動の包絡線を表す)

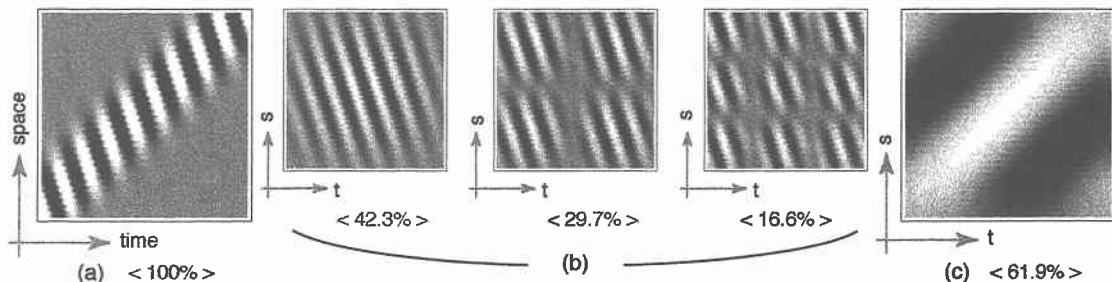


図-3 変動データ $\eta^p(x, t)$ およびその主成分解析における工夫
($\langle \rangle$ 内の数字は全エネルギーに対する比率を表す)

(固有関数の不定性の問題；北野・間瀬 (1999) 参照)。

また、一意に分解されても観測データそのものの誤差、および解析が有限範囲であることに起因するリンク効果等の問題 (北野・間瀬, 1999) を考慮に入れると上述の不定性の危険性を犯す可能性がある。つまり、変動エネルギーの比率が同程度となる複数の固有関数の重ね合わせで表現される場合は、全変動エネルギーの大部分が1つの固有関数に集中するように、本節で用いたような何らかの工夫をすべきであると考えられる。この考え方は、変数分離法による微分方程式の解法での適切な座標変換を行うという発想に類似したものである。

4. 結 語

海底地形変動の現況把握に用いられる主成分解析の理論背景を従来とは異なる視点で整理し、得られる固有関数の特性について検討した。さらに、主成分解析の第一の特徴が変数分離形であることに注目し、1) 不規則波浪の統計処理に用いられる波別解析法や 2) 不規則波の反射率測定に用いられる入反射分離法を導入することにより、漂砂による海浜地形の変化を把握する上で重要となる変動の時間的スケールやその伝播経路等の特性について抽出することが可能であることを示した。

最後に、理論構築に自由度の含まれる3変数の変動データの解析法、他の内積定義の採用、固有値分解における条件の付加等、主成分解析にはまだまだ発展の余地が残されているが、それらは今後の検討としたい。

謝辞：主成分解析の今後の改良点に関して、運輸省港湾技術研究所の橋本典明博士との有意義な議論は、本研究をまとめるにあたって非常に参考になり、ここに感謝致します。また、本研究の一部は文部省科学研究費基盤

研究(C)(2) (No. 10650509) および前田記念工学振興財団によったことを付記し、謝意を表します。

参 考 文 献

- 加藤一正・灘岡和夫・田中則男 (1980): Empirical eigenfunction によるエッジ波の解析, 第27回海岸工学講演会論文集, pp. 178-182.
- 加藤一正・吉松 晃 (1984): 三次元の経験的固有関数による深淺図解析法, 港湾技術研究所報告, 第23巻, 第2号, pp. 24-47.
- 北野利一・間瀬 肇 (1999): 複素主成分解析の変動データへの応用, 京都大学防災研究所年報, (印刷中).
- 橋本典明・永井紀彦・清水勝義・菅原一晃・久高将信・田中 聡 (1995): 主成分分析とカルマンフィルタを用いた統計的波浪予測手法の適用性について, 海岸工学論文集, 第42巻, pp. 336-340.
- 横木裕宗・三村信男・佐藤圭輔 (1998): 主成分解析を用いた大洗港周辺の海浜地形解析, 海岸工学論文集, 第45巻, pp. 571-575.
- Aubrey, D. G. (1979): Seasonal patterns of onshore/offshore sediment movement, J.G.R., Vol. 84, No. C10, pp. 6347-6354.
- Aubrey, D. G., D. L. Inman and C. D. Winant (1980): The statistical prediction of beach changes in Southern California, J.G.R., Vol. 85, No. C6, pp. 6264-3276.
- Bosma, K. F. and R. A. Darlymple (1996): Beach profile analysis around Indian River Inlet, Delaware, USA, Proc. 25th Coastal Eng. Conf., ASCE, pp. 401-405.
- Liang, G. and R. J. Seymour (1991): Complex principal component analysis of wave-like sand motions, Proc. Coastal Sediments '91, ASCE, pp. 2175-2186.
- Liang, G., T. E. White and R. J. Seymour (1992): Complex principal component analysis of seasonal variation in near-shore bathymetry, Proc. 23rd Coastal Eng. Conf., ASCE, pp. 2242-2250.
- Vincent, C. L. and D. T. Resio (1977): An eigenfunction parameterization of a time sequence of wave spectra, Coastal Eng., Vol. 1, pp. 185-205.