

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

橋本典明*・永井紀彦**・清水勝義***
菅原一晃***・久高将信****・田中 聡****

1. はじめに

港湾工事や荷役の安全確保、最適な施工管理あるいは航行船舶の安全確保等のためには的確な波浪予測が不可欠である。波浪予測手法としては、波浪推算法を用いる方法と統計モデルを用いる方法がある。この内、波浪推算法を用いる方法では、広域の気象情報が必要で、計算に際しては比較的大規模なシステムが必要であって多くの費用や専門知識を要することが一般である。一方、統計モデルを用いる方法は、専門知識を必要とせず、比較的容易に波浪予測を実施できる利点がある。

既往の統計的波浪予測手法としては、重回帰式(須田・湯沢, 1983; 小舟ら, 1987)や多変量自己回帰式(小舟ら, 1990)等の回帰式を用いる方法、分割表を用いる判別型モデル(小舟ら, 1990; 駒口ら, 1994)など種々の方法が提案されている。しかしながら、既往の回帰式を用いる方法では予測波高が観測波高より遅れて変動する重大な欠点を有し、また、判別型モデルでは予測波高を具体的な数値として推定できない等の難点がある。

本研究は、既往の統計的波浪予測手法の問題点を種々検討し、実用的な統計的波浪予測手法として、新たに主成分分析とカルマンフィルタを組み合わせたモデルを作成し、本方法の妥当性、適用性を検討したものである。

2. 主成分分析とカルマンフィルタを用いた波浪予測

本手法は、図-1に示すフローで波浪を予測する。すなわち、最初にカルマンフィルタによるトレンドモデルによって、気圧に含まれる長周期成分と短周期成分をリアルタイムで分離する。さらに、これらの成分を予測対象期間以前のデータで主成分分析して得られたそれぞれの固有ベクトル上に展開し、別のデータに変換する。この時に固有ベクトルにかかる重み(係数)を説明変数として、カルマンフィルタによる時変係数型波浪予測式を用いて波浪を予測する。

気圧を長周期成分と短周期成分に分離したのは、数日先を対象とする短期波浪予測式が、気圧の長周期成分の影響を強く受け、バイアスが生じるのを軽減するためである。気圧情報を主成分分析した固有ベクトル上に展開したのは、格子上の気圧情報を直交する他の情報に変換することによって、気圧値相互間の多重共線性(線形独立性が低い)の問題を軽減するためである。また、時変係数型波浪予測式としたのは、波浪場の季節変動等のゆるやかな変化を予測誤差から検知し、この影響をモデルの中に組み込むためである。

(1) カルマンフィルタの概要(北川, 1993)

y_n を l 次元の観測ベクトルの時系列とし、この時系列を次のような状態空間モデルで表現する。

$$x_n = F_n x_{n-1} + G_n v_n \quad (\text{システム方程式}) \quad \cdots (1)$$

$$y_n = H_n x_n + w_n \quad (\text{観測方程式}) \quad \cdots (2)$$

ここに、 x_n は状態変数で、直接には観測できない k 次元のベクトル、 v_n はシステムノイズで、平均ベクトル 0 、分散共分散行列 Q_n に従う正規白色雑音、 w_n は観測ノイズで平均ベクトル 0 、分散共分散行列 R_n に従う正規白色雑音、 F_n, G_n, H_n はそれぞれ $k \times k, k \times m, l \times k$ の行列である。

カルマンフィルタは y_n の観測値に基づいて状態 x_n の推定を行う逐次計算アルゴリズムで、以下に示す1期先予測とフィルタを交互に繰り返すことにより、状態 x_n が逐次求められる。

(1期先予測)

$$x_{n|n-1} = F_n x_{n-1|n-1} \quad \cdots (3)$$

$$V_{n|n-1} = F_n V_{n-1|n-1} F_n^t + G_n Q_n G_n^t \quad \cdots (4)$$

(フィルタ)

$$K_n = V_{n|n-1} H_n^t (H_n V_{n|n-1} H_n^t + R_n)^{-1} \quad \cdots (5)$$

$$x_{n|n} = x_{n|n-1} + K_n (y_n - H_n x_{n|n-1}) \quad \cdots (6)$$

$$V_{n|n} = (I - K_n H_n) V_{n|n-1} \quad \cdots (7)$$

なお以下では、時系列 y_n の次元が $l=1$ であることから、 $R_n = \sigma_w^2$ で一定とし、観測ノイズとシステムノイズを新たに $\bar{R}=1, \bar{Q}_n = \sigma_v^2 / \sigma_w^2 I$ として、トレードオフパラメータ σ_v^2 / σ_w^2 のみを制御パラメータとして解析を行った。

(2) カルマンフィルタを用いたトレンド成分の分離

いま観測値 y_n はトレンド成分 t_n に様々な変動成分 w_n

* 正会員 工博 運輸省港湾技術研究所海洋エネルギー利用研究室長
** 正会員 運輸省港湾技術研究所海象調査研究室長
*** 正会員 運輸省港湾技術研究所水工部主任研究官
**** 正会員 (株)エコー 第一技術部 波浪プロジェクト室

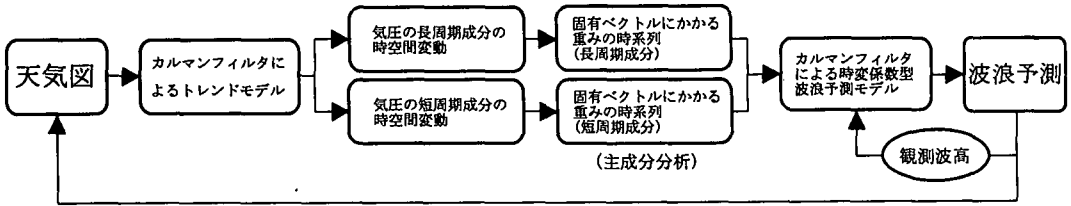


図-1 主成分分析とカルマンフィルタを組み合わせた波浪予測フロー

が加わったものと仮定すれば,

$$y_n = t_n + w_n \quad (\text{観測方程式}) \quad \dots\dots\dots (8)$$

と表される。ここに、 w_n は平均 0、分散 σ^2 の正規白色雑音である。トレンド成分 t_n は時間差分オペレータを Δ^k とすれば次式で表される。

$$\Delta^k t_n = v_n \quad (\text{システム方程式}) \quad \dots\dots\dots (9)$$

ここに、 v_n は平均 0、分散 σ^2 の正規白色雑音である。

時間差分オペレータ Δ^k は 1 次および 2 次の場合を検討し、ここでは 2 次の差分オペレータを用いた。

(3) 主成分分析を用いた気圧の時空間データの変換

気圧の時空間分布を平均値と標準偏差で正規化し、これを $P_{z,t} = P(x, y, t)$ で表す。また、 $P_{z,t}$ がある固有ベクトル $e_{n,z}$ 上で表されるとすると $P_{z,t}$ は次式で表される。

$$P_{z,t} = \sum c_{n,t} e_{n,z} \quad \dots\dots\dots (10)$$

ここに、

$$\sum e_{n,z} e_{m,z} = \delta_{n,m} \quad \dots\dots\dots (11)$$

である。気圧の時空間分布の共分散分散行列を A とすると、その i, j 成分は次式で表される。

$$a_{i,j} = \frac{1}{n_z n_t} \sum_{t=1}^{n_t} P_{i,t} P_{j,t} \quad \dots\dots\dots (12)$$

このとき、式 (10) 中の固有ベクトル $e_{n,z}$ は、

$$A = \lambda_n e_n \quad \dots\dots\dots (13)$$

から求められる。ここに、 λ_n は固有値である。係数 $c_{n,t}$ は固有ベクトルの直交性を利用して次式で求められる。

$$c_{n,t} = \sum P_{z,t} e_{n,z} \quad \dots\dots\dots (14)$$

以下に示すカルマンフィルタを用いた時変係数型波浪予測式では式 (14) の $c_{n,t}$ が説明変数として用いられる。ただしここでは、指定した範囲内の格子上の気圧値 3~4 日分をまとめて行列 $P_{z,t}$ として主成分分析する。すなわち、指定範囲内の気圧の空間変動の 3~4 日分の変動を説明する固有ベクトル $e_{n,z}$ と係数 $c_{n,t}$ を求め、この係数 $c_{n,t}$ が説明変数として用いられる。

(4) カルマンフィルタを用いた時変係数型波浪予測式

有義波高 $H_{1/3}$ を予測するための説明変数を z_i として、波高予測式を次式で与える。

$$\log_{10} H_{1/3} = a_0 + \sum_{i=1}^N a_i z_i + \varepsilon \quad \dots\dots\dots (15)$$

ここに、 a_i は未知係数、 ε は予測誤差で平均 0、分散 σ^2 の正規白色雑音であると仮定する。なお、有義波高 $H_{1/3}$ を対数変換して予測式を作成したのは、式 (15) の右辺は説明変数の線形結合で表現されており、中心極限定理により正規分布に近い分布を示すと考えられ、また、有義波高 $H_{1/3}$ は一般に対数正規分布あるいはワイブル分布に近い分布であることを考慮したものである。ここで、 $y_n = \log_{10} H_{1/3}$ 、 $x_n = (a_0, a_1, \dots, a_N)^t$ 、 $H_n = (1, z_1, z_2, \dots, z_N)$ 、 $w_n = \varepsilon$ と置くと、式 (15) は次式で表される。

$$y_n = H_n x_n + w_n \quad (\text{観測方程式}) \quad \dots\dots\dots (16)$$

また、波高予測式 (15) 中の係数 a_n が時間的に緩やかに変化するとすれば、

$$\Delta^k x_n = v_n \quad (\text{システム方程式}) \quad \dots\dots\dots (17)$$

で表される。ここに、 v_n は平均 0、分散共分散行列 Q に従う正規白色雑音である。

時間差分オペレータ Δ^k は 1 次および 2 次の場合を検討し、ここでは 1 次の差分オペレータを用いた。

3. 気象・海象データ解析

統計的波浪予測手法の適用性を検討するため、1980 年~1984 年の地上天気図から読み取られた気圧値(図-2 に示す 5° 間隔の格子上の気圧値、9 時および 21 時)と日本沿岸の 6 地点(むつ小川原、弾崎、鳥取、波浮、油津、中城)の波浪観測データを用いた。解析に際しては、1980 年~1981 の 2 年分のデータを事前処理(学習)に用

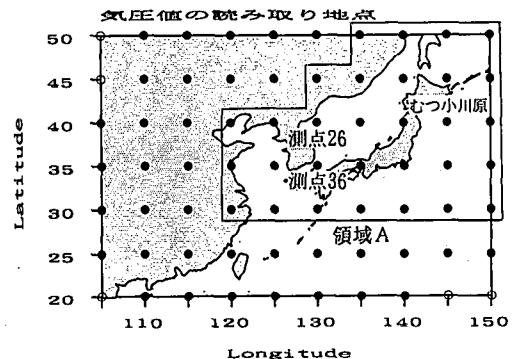


図-2 気圧の読み取り地点

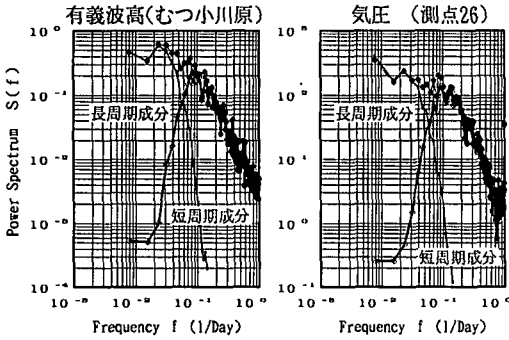


図-3 有義波高の対数変換値と気圧値のスペクトル

い、1982 年～1984 年の 3 年間を対象として予測精度を検討した。

図-3 は有義波高の対数変換値と気圧値をカルマンフィルタのトレンドモデル ($\sigma^2/\sigma_w^2=0.001$) により短周期成分と長周期成分とに分離し、スペクトル解析した結果である。分離結果は σ^2/σ_w^2 の大小によって変化し、小さな σ^2/σ_w^2 を指定すれば、長周期成分のスペクトルはより低周波数側へ移動する。図-3 の例では十日周期程度の短周期成分とそれよりも長い長周期成分とに分離されていることが分かる。

図-4 はむつ小川原の有義波高の対数変換値と図-2 の測点 36 の気圧値との相互相関係数を示したものである。(a) は長周期成分を除去しない場合の相互相関係数、(b) は長周期成分を除去した場合の相互相関係数である。(a) では相互相関係数はタイムラグの広い範囲で負の相

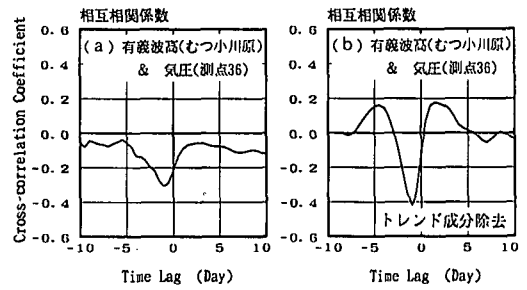


図-4 有義波高の対数変換値と気圧値との相互相関係数

関を示し、極値はあまり明確ではない。(b) では相互相関係数の極値が明確に現れ、タイムラグの増加とともに相互相関係数の絶対値が小さくなっていることが分かる。

3～5 日先の短期波浪予測を実施する場合、より精度の良い予測値を得るためには、対象とする時間スケールの現象を抽出して予測式を作成する必要がある。しかしながら、冬季の日本海側の高波浪に見られるように、季節的な長周期成分も無視できず、実際には両者を同時に考慮した予測が必要である。これがトレンドモデルで短周期成分と長周期成分を分離した理由であり、それぞれの成分が独立の要因として予測式が作成される。

図-5 は主成分分析して得られた気圧の時空間変動の第 1～4 主成分の固有ベクトルを示したものである。実際の計算では、図-2 の 1/4～1/2 程度の狭い範囲の気圧の時空間変動データを用い、短周期成分と長周期成分のそ

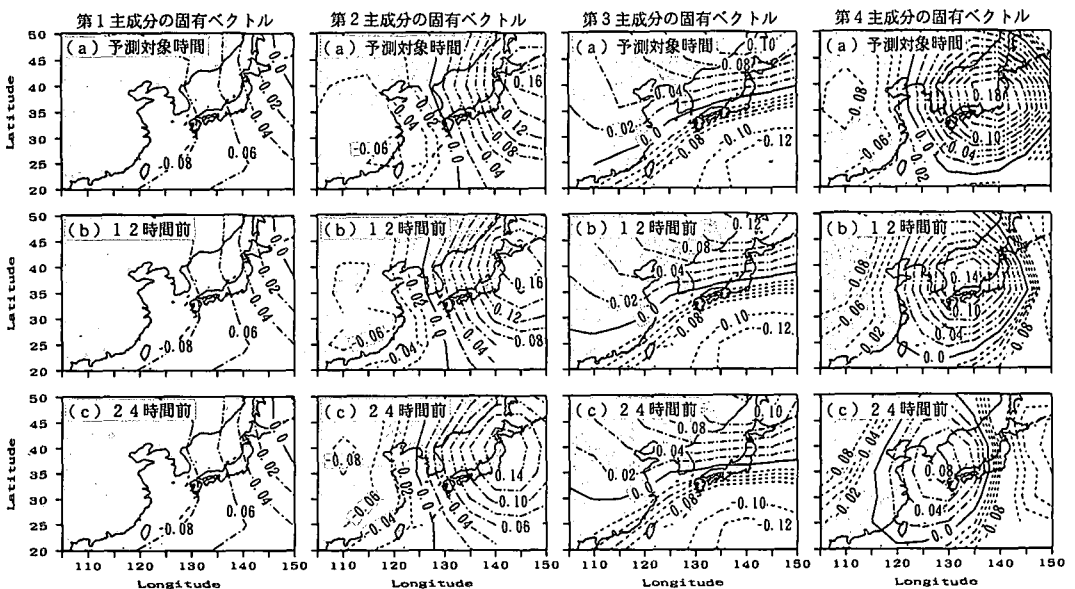


図-5 気圧の時空間変動の固有ベクトルの一例

れぞれについて主成分分析した結果を用いて波浪予測するが、ここでは便宜上、図-2の全領域のデータを用い、トレンド成分を分離せずに主成分分析した結果を示した。第1主成分は気圧場が時間的に変動しない平均的パターン、第2主成分は気圧場が発達しながら西から東へ移動するパターン、第3主成分は南から北へ移動するパターン、第4主成分は発達しながら南西から北東へ移動するパターンを表している。このように気圧の時空間変動は、図-5に示すような幾つかの独立した気象パターンの重ね合わせとして表現され、後で示す時変係数型波浪予測式では、図-5に例示した固有ベクトルにかかる重み(係数 $c_{n,i}$)が説明変数として用いられている。なお、長周期成分の変動はわずか3~4個の主成分で全変動の95%以上が説明でき、短周期成分も元の情報の半数程度の成分で同等の変動が説明でき、元の情報よりも説明変数の数を減らすことが出来る。また、結果として予測精度は向上する。

4. 波浪予測結果の検討

図-2で各港毎に適当な領域を定め、この領域内の気圧値を用いてトレンド分離および主成分分析を行い、時変係数型波浪予測式を用いて予測した。図-2中の領域Aは、むつ小川原を対象とした領域で、30の気圧測点が含まれている。なお、波浪予測に際しては、波浪推算と同様に、波浪観測値は現時刻まで、天気図は予測対象時刻まで与えられると仮定して予測した。また、主成分分析に際しては、指定領域内の3日分の気圧変動の95%を説明する主成分を求め、この重みを波浪予測に用いた。

図-6は σ_v^2/σ_w^2 を変えた場合の式(17)の状態 x_n の経時変化の一例である。 σ_v^2/σ_w^2 に大きな値を設定すれば x_n は激しく変動するが、 σ_v^2/σ_w^2 に小さな値を設定すれば、学習初期の激しい変動後、 x_n は平均値回りに緩やかに変動する。そこでここでは、 σ_v^2/σ_w^2 を変えて予測し、 σ_v^2/σ_w^2 の相違による予測値の相違を検討した。なお、予測精度の検討方法には種々の方法があるが、ここでは単に予測波高と有義波高の相関係数および予測波高と有義波高の誤差

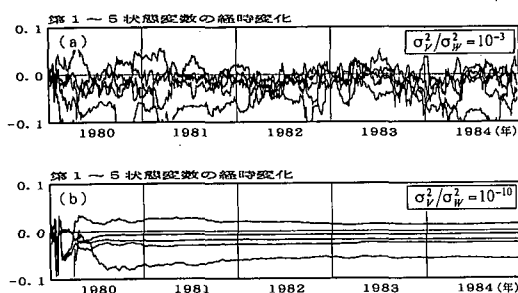


図-6 状態変数 x_n の経時変化の一例

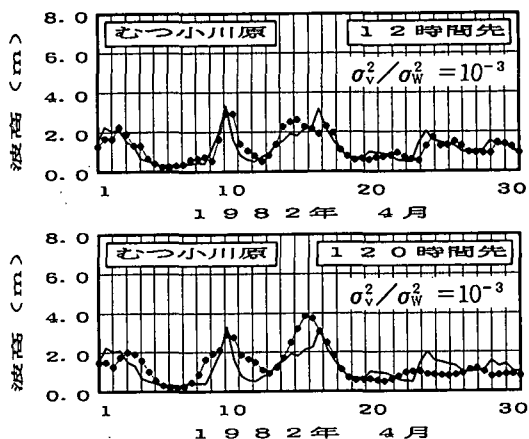


図-7 $\sigma_v^2/\sigma_w^2=10^{-3}$ の場合の波浪予測結果の一例

の標準偏差(予測誤差)を用いて検討した。

図-7は $\sigma_v^2/\sigma_w^2=10^{-3}$ の場合の12時間後と120時間後の予測波高と観測波高の時系列図を示したもので、図中の●が予測波高である。図-7に見られるように、12時間後の予測波高は観測波高に遅れて変動する傾向がある。また、120時間後の予測波高は実際の変動傾向をある程度捉えているが、精度はあまり良くない。

図-8は $\sigma_v^2/\sigma_w^2=10^{-10}$ とした場合の12時間後と120時間後の予測波高と観測波高の時系列図である。 σ_v^2/σ_w^2 をこのように小さく設定すると、 x_n の変動は緩やかで、120時間の範囲では x_n はほとんど変化しない。このため、結果として気圧情報のみで波高が予測されることになり、図-8に見られるように12時間後と120時間後の予測値はほぼ同じ値を示し、精度が良い。なお、12時間後の予測値に、予測波高が遅れて変動しているように見える箇所があるが、この変動は既に120時間前にも予測されており、この意味では予測が遅れている訳ではない。

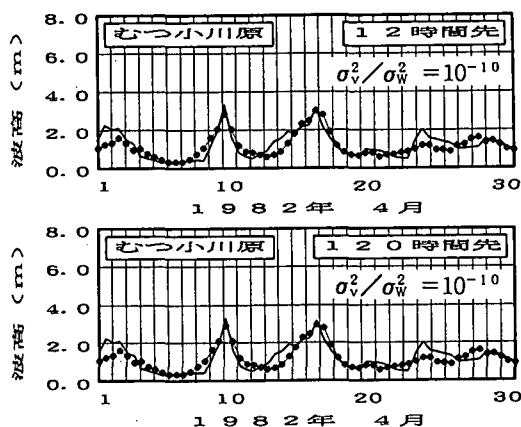


図-8 $\sigma_v^2/\sigma_w^2=10^{-10}$ の場合の波浪予測結果の一例

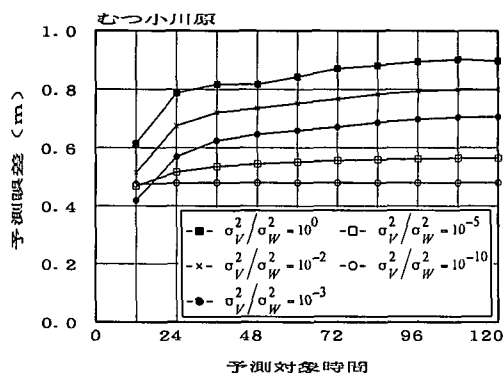
図-9 σ_v^2/σ_w^2 の相違による予測精度の相違

図-9 は、むつ小川原を対象として σ_v^2/σ_w^2 の値を種々変えて波浪予測し、予測対象時間毎に予測誤差を示したものである。図-9 に見られるように、 σ_v^2/σ_w^2 の値によって予測精度は異なり、12 時間先の予測誤差は $\sigma_v^2/\sigma_w^2=10^{-3}$ の場合に最小となるが、数日先の予測では $\sigma_v^2/\sigma_w^2=10^{-5}$ や 10^{-10} の場合の方が精度が高くなる。これは $\sigma_v^2/\sigma_w^2=10^{-3}$ では状態 x_n が観測波高にオーバーフィッティング (過学習) する傾向があるためである。すなわち、数日先の波浪予測を実施するためには、目先 (12 時間先) の予測誤差を最小化することは、必ずしも好ましくなく、事例のどの点にモデルをうまく適合させるべきかというような事前の検討が重要であることを示唆している。

図-10 はむつ小川原を対象として、種々のモデルで波浪予測した結果 (120 時間先) と観測波高との散布図を示したものである。(a), (b) は気圧値のトレンド分離および主成分分析をせずに、気圧値をそのまま説明変数として予測した例で、(a) は $\sigma_v^2/\sigma_w^2=10^{-3}$ 、(b) は $\sigma_v^2/\sigma_w^2=10^{-10}$ とした場合の結果である。また、(c), (d) は気圧値のトレンド分離および主成分分析を導入した例で、それぞれ $\sigma_v^2/\sigma_w^2=10^{-3}$ 、 $\sigma_v^2/\sigma_w^2=10^{-10}$ の場合の結果である。(d) に見られるように、本論文で示した方法で適当なパラメータ σ_v^2/σ_w^2 を選択することにより、実用上許容できる誤差範囲内で数日先の波高を予測し得ることが分かる。

この他、弾崎、鳥取、波浮、油津、中城を対象として同様な検討をしたが、いずれの地点でも予測精度は良好で、特に弾崎、鳥取、油津では 120 時間先の予測値と観測波高の年間の相関係数が 85 % 以上を示す例も見られた。

5. おわりに

主成分分析とカルマンフィルタを組み合わせた波浪予測手法は、3~5 日先の短期波浪予測を実施する場合、

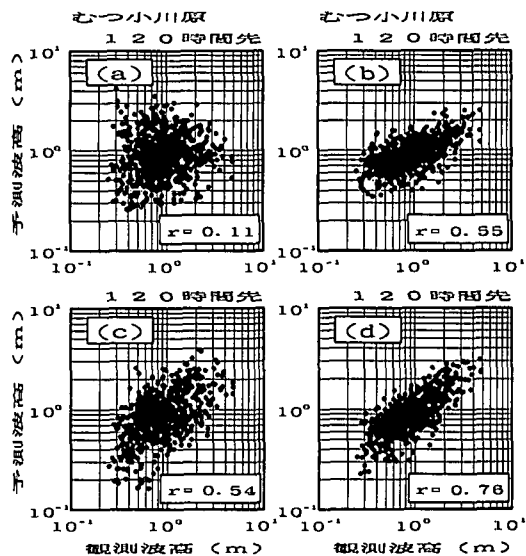


図-10 波浪予測法の相違による予測精度の相違

同じ情報を説明変数とする既往の統計的波浪予測手法よりも予測精度が格段に向上する。本手法で予測される波高は、カルマンフィルタのトレードオフパラメータの大きさにより異なった特性を示し、妥当なパラメータの大きさを設定すれば、一期先の予測誤差は最小ではないものの、予測誤差の予測ステップ数への依存度が小さくなり、数ステップ先 (3 日~5 日先) の予測値も実用上許容できる誤差範囲内で有義波高を予測可能である。

なお、今回の検討は 5 時間隔の粗い格子間隔で読み取られた気圧値を用いて検討しており、予測精度にも限界がある。今後、波浪推算と同程度の細かい格子間隔で読み取られた気圧値を用いて更に検討する予定である。

参考文献

- 須田 照・湯沢 昭 (1983): 波浪予測に基づく外海シーバースの待ち行列に関する基礎的研究, 土木学会論文報告集, 第 339 号, pp. 177-185.
- 小舟浩治・橋本典明・亀山 豊・久高将信 (1987): 重回帰式を用いた波浪予測手法の適用について, 第 34 回海岸工学講演会論文集, pp. 167-171.
- 小舟浩治・橋本典明・亀山 豊 (1990): 統計モデルを用いた波浪予測手法の適用性に関する検討, 港湾技研資料, No. 673, 42 p.
- 駒口友章・佐藤勝弘・後藤智明・青野利夫 (1994): カテゴリカルモデルを利用した波浪予測の適用性, 海岸工学論文集, 第 41 巻, pp. 211-215.
- 北川源四郎 (1993): 時系列解析プログラミング, 岩波書店, 390 p.