

ニューラルネットワークを用いた有義波高の予測について

太田 隆 夫*・木村 晃**

1. はじめに

従来の簡便な波浪予測法として、重回帰式や多変量自己回帰式を用いた統計モデルが挙げられる。これらの方法については、すでにその適用性が検討され、実用的な精度の予測結果を得られる場合があることも報告されている(たとえば小舟ら, 1990)。また、問題点として、高波浪時の立ち上り遅れが指摘されているが、その他に比較的多くのデータ(有義波高や気圧など)を必要とすることも問題の1つであるといえる。さらに、入力(説明変数)と出力(目的変数)との関係を線形関数で表しているという制約もある。波浪などの自然現象を対象とする場合、仮に回帰的なモデルでその変動が記述できるとしても、入出力間の関係は非線形関数で表現するほうがより一般的である。ただし、現象のメカニズムが明確でない場合は、ある特定の非線形関数を選択する根拠がないことになる。このような場合には、関数を陽に仮定せずに学習という操作で現象の特徴抽出を行っていくニューラルネットワークの適用が考えられる。本研究では、従来の統計モデルのような制約がなく、より少数のデータで作成できる予測モデルとして、ニューラルネットワークを用いた自己回帰的モデルに着目し、これによる有義波高の予測について検討する。

2. 予測モデルの作成

2.1 ネットワークモデル

本研究で用いる予測モデルは、基本的には次式で表されるものである。

$$y_s = f(x_n)$$

$$x_n = (x_1, x_2, \dots, x_n), \quad y_s = (y_1, y_2, \dots, y_s) \dots \dots (1)$$

この関数 f を N 個の入出力データ組 $(x_n^1, y_s^1), \dots, (x_n^N, y_s^N)$ が与えられたもとで、ニューラルネットワークにより同定することを考える。ただし、 $s \geq 2$ の場合、出力 y_s の各成分は $[-1, 1]$ に正規化されていることが前提となる。ニューラルネットワークは、人間の脳神経細胞を模擬した数理モデルであり、これにより任意の非線形関数を任

意の精度で近似可能であることが証明されている。最近では、入力層・中間層・出力層からなる階層型ネットワークと、誤差逆伝播法という学習法との組み合わせで用いられることが多い。ここで、学習とは与えられた学習データ(入力データとそれに対応する出力データの組)に合うように、ネットワークのパラメータを調整していく操作のことである。本研究では、ニューラルネットワークの一種であるラジアル基底関数ネットワーク(片山ら, 1992)を用いるが、その構造を図-1に示す。中間層の各ユニットには、入力層からのデータに対する出力を決める基底関数が組み込まれている。基底関数としては、図-2に示すようなガウス関数をベースにしたものを用いることが多く、その重ね合わせで任意関数を近似する。このネットワークモデルでは、出力値 $\bar{y}_j$ ($j=1, \dots, s$) は次のように与えられる。

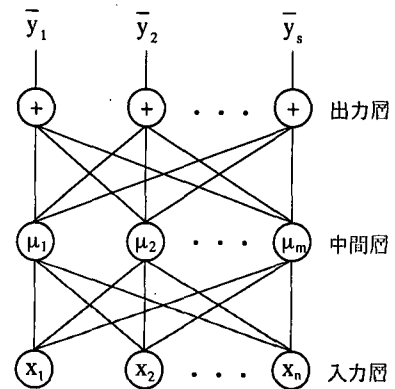


図-1 ネットワークの構造

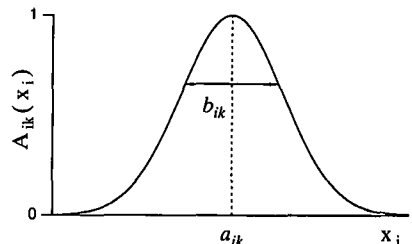


図-2 ガウス関数

* 正会員 工修 鳥取大学助手 工学部社会開発システム工学科

** 正会員 工博 鳥取大学教授 工学部社会開発システム工学科

$$\bar{y}_j = \sum_{k=1}^m w_{jk} \mu_k(x_n, a_k, b_k) \dots \dots \dots (2)$$

ここに、 $\mu_k(\)$: k 番目の基底関数、 m : 基底関数数、 w_{jk} : k 番目の基底関数からの出力値に対する重み係数、 x_n : 入力データのベクトル、 a_k : ガウス関数の中心値のベクトル、 b_k : ガウス関数の幅を表すパラメータのベクトル。

さらに、基底関数は式 (3)、(4) で与えられる。

$$\mu_k(x_n, a_k, b_k) = \prod_{i=1}^n A_{ik}(x_i, a_{ik}, b_{ik}) \dots \dots \dots (3)$$

$$A_{ik}(x_i, a_{ik}, b_{ik}) = \exp\left\{-\frac{(x_i - a_{ik})^2}{b_{ik}}\right\} \dots \dots \dots (4)$$

式 (2)~(4) で、学習によって決定されるパラメータは、ガウス関数の中心値 a_{ik} と分布幅のパラメータ b_{ik} 、および基底関数の出力値に対する重み係数 w_{jk} である。また、中間層ユニット数 (基底関数数) をいくつにするかということも問題となる。従来のニューラルネットワークでは、中間層ユニット数を変えて学習を行い、誤差が最小となるユニット数を採用するという試行錯誤的な方法が用いられてきた。本研究では、1 回の学習過程の中で自動的に基底関数数を増やす方法 (山ら, 1992) を用いる。この方法については次節に述べるが、これによる関数の同定は以下の 2 つの過程から構成される。

① 基底関数数が固定されたもとのパラメータの学習過程

② パラメータの学習による誤差の低減が小さくなった段階で、誤差の最大絶対値を与える入出力データを中心とする新たな基底関数を発生させる過程

2.2 関数の同定

基底関数数が m の場合、学習の対象となる決定パラメータベクトルを次のように定義する。

$$\left. \begin{aligned} a_k &= (a_{1k}, a_{2k}, \dots, a_{nk}) & (k=1, \dots, m) \\ a^m &= (a_1, a_2, \dots, a_m) \\ b_k &= (b_{1k}, b_{2k}, \dots, b_{nk}) & (k=1, \dots, m) \\ b^m &= (b_1, b_2, \dots, b_m) \\ w_k &= (w_{1k}, w_{2k}, \dots, w_{sk}) & (k=1, \dots, m) \\ w^m &= (w_1, w_2, \dots, w_m) \end{aligned} \right\} \dots \dots \dots (5)$$

このとき、 N 個の入出力データの組に対し、同定問題 $P(m)$ を次のように定式化できる。

$$\min_{(a^m, b^m, w^m)} E_m(a^m, b^m, w^m) = \frac{1}{2} \sum_{p=1}^N \sum_{j=1}^s (y_j^p - \bar{y}_j^p)^2 \dots \dots \dots (6)$$

ここに、 E_m : 評価関数、 y_j^p : p 組目の入力データ x_n^p に対する j 番目の出力データ、 $\bar{y}_j^p$: x_n^p を入力したときのネットワークによる j 番目の出力値。

$P(m)$ の最適解を (a^m, b^m, w^m) とすると、実現すべき誤差 $\varepsilon > 0$ が与えられたもとの全体の同定問題は、

$$E_m(a^m, b^m, w^m) < \varepsilon \dots \dots \dots (7)$$

を満足する最小の基底関数数 m と、そのときの (a^m, b^m, w^m) を求めることになる。このような問題は一般に勾配法によって解くことができる。評価関数 E_m の決定パラメータ a_{ik}, b_{ik}, w_{jk} に関する勾配は式 (8)~(10) で与えられる。

$$\frac{\partial E_m}{\partial a_{ik}} = \frac{-2}{b_{ik}} \sum_{p=1}^N \sum_{j=1}^s w_{jk} \mu_k(x_n^p, a_k, b_k) \cdot (y_j^p - \bar{y}_j^p)(x_i^p - a_{ik}) \dots \dots \dots (8)$$

$$\frac{\partial E_m}{\partial b_{ik}} = \frac{-1}{b_{ik}^2} \sum_{p=1}^N \sum_{j=1}^s w_{jk} \mu_k(x_n^p, a_k, b_k) \cdot (y_j^p - \bar{y}_j^p)(x_i^p - a_{ik})^2 \dots \dots \dots (9)$$

$$\frac{\partial E_m}{\partial w_{jk}} = - \sum_{p=1}^N \mu_k(x_n^p, a_k, b_k)(y_j^p - \bar{y}_j^p) \dots \dots \dots (10)$$

本研究では最も基本的な最急降下法を用いるが、その場合の学習則は以下のようになる。

$$a_{ik}(h+1) = a_{ik}(h) - \alpha_h \frac{\partial E_m}{\partial a_{ik}} \dots \dots \dots (11)$$

$$b_{ik}(h+1) = b_{ik}(h) - \alpha_h \frac{\partial E_m}{\partial b_{ik}} \dots \dots \dots (12)$$

$$w_{jk}(h+1) = w_{jk}(h) - \alpha_h \frac{\partial E_m}{\partial w_{jk}} \dots \dots \dots (13)$$

ここに、 h : 学習の繰り返し回数、 α_h : 繰り返し回数 h の時の学習係数。

つぎに、基底関数数 m を固定して学習を進めても評価関数 E_m の減少効果がほとんど見込めなくなった段階で、中間層ユニットを 1 つ増やしてそれに対応する新規の基底関数を発生させる。 N 個の学習データのうち、ネットワーク出力値との誤差の絶対値が最大となるデータに着目し、この学習データの入力値を中心とする基底関数を発生させる方法をとる。この方法に基づいた全体の同定問題に対する手順を以下に示す。

- ① N 個の学習データに対して、最急降下法により $(a^m(h), b^m(h), w^m(h))$ から $(a^m(h+1), b^m(h+1), w^m(h+1))$ を計算する。ここで、 h は中間層ユニット数が m のときの学習回数を表す。
- ② 学習回数が $(h+1)$ のときの評価関数値が ε より小さければ、そのときの m が式 (7) を満たす最小の基底関数数であり、 $(a^m(h+1), b^m(h+1), w^m(h+1))$ を (a^m, b^m, w^m) として採用し終了する。
- ③ 評価関数の減少率が許容値 ε_1 より小さければ④へ進み、 ε_1 以上であれば①へ戻る。
- ④ 入出力データ (x_n^p, y_s^p) ($p=1, \dots, N$) のうち、すでにガウス関数の中心値として選択されたもの以外で、絶対値誤差が最大となる入出力データを (x_n^p, y_s^p) とするとき、 $(m+1)$ 個目の基底関数を以下のように発生させる。

$$\begin{aligned} a_{i,m+1} &= x_i^? & (i=1, \dots, n) \\ b_{i,m+1} &= b_0 & (i=1, \dots, n) \\ w_{j,m+1} &= y_j^? - \bar{y}^? & (j=1, \dots, s) \dots\dots\dots (14) \end{aligned}$$

ここで b_0 は、ガウス関数のパラメータの初期値として与えられている定数である。($m+1$) 個目の基底関数を上記のように設定し、①へ戻る。新規の基底関数以外の基底関数に含まれるパラメータには、直前の学習過程で得られている値を使用する。

3. 学習データの作成

本研究で用いるデータは、運輸省港湾局による全国港湾海洋観測資料のうち、鳥取港における1984年1月から1989年10月までの2時間毎の有義波高データである。この期間のデータ取得率は99.6%である。また、欠測の影響を小さくするために、連続8回以下の欠測に対しては線形補間を行い、それ以上の欠測は観測波高をゼロとした。2.に述べたように、ネットワークモデルの作成のためには、まず学習データが必要となる。本研究では、自己回帰的な予測モデルを考えているので、入力データ、出力データともに有義波高となる。一組の学習データは、ある時刻の有義波高データを中心に、それを含めた過去の n 個を入力データ、ある時刻以降の s 個を出力データとして構成される。また、2.1に述べたように、出力データは $[-1, 1]$ であることが前提条件となっている。そのため、有義波高の常用対数をとって学習データを作成する。有義波高が10 mを越えることは稀であり(鳥取港では1979年～1996年の間に記録されていない)、10 cmを下回ることも少ないことが、常用対数を採用した理由である。上述のデータから、欠測が少なく、10 cm未満の観測値がない期間を選択して以下の学習データを作成した。

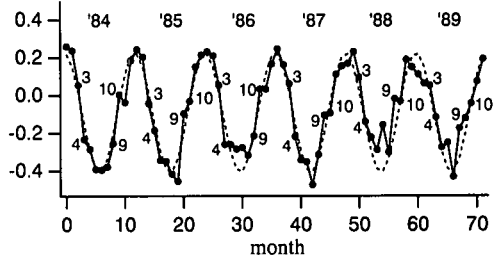
- ① 1987年12月と88年1月、88年3月と4月、6月と7月、9月と10月の各2ヵ月分のデータから500組の学習データ
- また、比較のために次の2通りの学習データも作成した。
- ② 有義波高データから周期変動成分を除いたものを用いた学習データ
- ③ 1987年1月と88年1月、87年4月と88年4月、87年7月と88年7月、87年10月と88年10月の有義波高データから各500組の学習データ
- ②については、次節に述べる。

3.1 有義波高の周期的変動

表一1に1984年から88年までの1月・12月、3月・4月、6月・7月、9月・10月の有義波高月平均値を示す。このうち、特に3月から4月と、年によっては9月と10月にかけて、月平均値が大きく変化している。これが予測モデルに影響を及ぼすかどうかを検討するために、こ

表一1 有義波高の月平均値 (単位: m)

	'84	'85	'86	'87	'88
1月	(2.02)	1.93	1.90	2.01	1.74
12月	1.79	1.85	1.70	1.67	(1.62)
3月	1.30	1.01	1.36	1.37	1.39
4月	0.80	0.81	0.66	0.77	0.87
6月	0.44	0.55	0.62	0.55	0.57
7月	0.45	0.44	0.61	0.40	0.75
9月	0.63	0.91	0.70	1.03	1.14
10月	1.21	1.12	1.27	1.11	1.19



図一3 月平均値の変動

の変動を取り除くことを考える。図一3の実線は有義波高の常用対数の月平均値を示しており、その変動が周期的であることがわかる。ここでは、Soares・Ferreira (1995)の方法にもとづいて、周期変動成分を取り除く。図一3に示した月平均値のうち、1984年1月から87年12月までをフーリエ変換した結果、12ヵ月周期の成分が卓越したことから、その係数値を用いて次式のようにモデル化した。

$$\begin{aligned} m(t) = & -0.0920 + 0.311 \cos\left(\frac{2\pi}{12}t\right) \\ & - 0.0547 \sin\left(\frac{2\pi}{12}t\right) \dots\dots\dots (15) \end{aligned}$$

図一3の破線は式(15)の値を示しているが、88年1月以降も含めて、観測値との対応はよい。特に3月と4月は、どの年もよく一致している。ただし、式(15)は月平均値を与えるものであるから、これをもとに各有義波高データに対する周期変動成分を考える。例えば、1年(365日)分のデータ個数 n_d は4380であり、これで1周期(2π) 分になる。式(15)の位相で $t = \tau n_d$ とくと、 $\tau = 12/4380$ となる。式(15)は上述の48ヵ月の月平均値データから得たもので、この期間の有義波高データ数は17532である。したがって、 $\tau = 0.002738$ となり、各有義波高データに対する周期変動成分は次式で与えられる。

$$\begin{aligned} m(n_i) = & -0.0920 + 0.311 \cos\left\{\frac{2\pi}{12} \times 0.002738(n_i - n_c)\right\} \\ & - 0.0547 \sin\left\{\frac{2\pi}{12} \times 0.002738(n_i - n_c)\right\} \dots\dots (16) \end{aligned}$$

ここで、 n_c は 1 月 1 日 0 時のデータをゼロとした番号、 n_c は式 (15) の月平均値がその月の中心になるようにする数で、平均的な半月分のデータ数 (182 または 183) である。有義波高の常用対数から式 (16) の値を差引くことにより、周期変動成分を除くことができる。この方法で得られるデータのうち、88 年 3 月と 4 月のデータから学習データを作成した。

4. 学習および予測の結果

3. で述べた①の各 2 ヶ月間の有義波高データから、入力データ数 n を 3~6、出力データ数 s を 3 とした 4 通りの学習データを作成した。また、87 年 12 月・88 年 1 月のデータからは、 n と s をともに 4~6 とした 3 通りの学習データも予備的に作成した。学習過程においては、中間層ユニット数 (基底関数数) m を $[1, 2n]$ の範囲で増加させ、学習終了後の評価関数値が最小となる m を選択する。各学習データに対して 1 つ選択されるネットワークモデルにより、6 時間後および 12 時間後の有義波高予測を行う。予測対象期間は、87 年 12 月・88 年 1 月のデータに対して 89 年 1 月、88 年 3 月・4 月に対して 89 年 4 月、88 年 6 月・7 月に対して 89 年 7 月、88 年 9 月・10 月に対して 89 年 10 月である。この予測モデルでは、2 時間毎の有義波高を入出力データとして学習を行っているので、ネットワークの出力値 $\bar{y}_3$ は最新の入力データから 6 時間後の値となる。たとえば $s=3$ の場合、1 回目のネットワーク出力を次の入力 (の一部) とすると、その出力値 $\bar{y}_3$ が 12 時間後の値となる。図-4 に予測値の計算手順 (模式図)

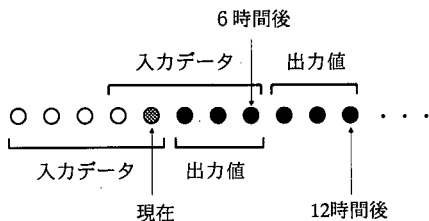


図-4 予測値の計算手順 (模式図)

順を模式的に示す (図は $n=5, s=3$ の場合)。また表-2 に、学習誤差として学習終了後の評価関数値を、予測誤差として予測値と観測値との 2 乗誤差の平均値を示す。学習誤差の欄の () 内は中間層ユニット数であり、予測誤差の欄の上段は 6 時間後予測、下段は 12 時間後予測に対する値である。この結果より、同一期間における学習および予測の結果に大きな差がないことがわかる。図-5, 6 には学習結果と予測結果の波形を示した。図-5(a) は、87 年 12 月・88 年 1 月のデータによる学習の結果であり、 $n=5, s=3, m=10$ の場合である。図中の実線は出力データの y_3 、破線はネットワークの出力値 $\bar{y}_3$ を示す。図-5(b), (c) はこのネットワークモデルによる 89 年 1 月の予測結果で、(b) が 6 時間後予測、(c) が 12 時間後予測である。図-6(a) は 88 年 3 月・4 月のデータによる学習の結果で、 $n=4, s=3, m=8$ の場合である。図-6(b), (c) は 89 年 4 月の予測結果を示している。図-5, 6 の (b), (c) で、実線は観測値、破線は予測値を表している。図-5(a), 図-6(a) ともに、ネットワークの出力値は、学習データの出力データとよく一致している。図-5, 6 の (b), (c) に示す予測結果では、従来の統計モデルと同様に高波浪の立ち上り遅れが見られる。この原因として、予測モデルにおいて最新の入力データの影響が大きく、これに引きずられる形で予測値が遅れることが考えられる。また、表-3 には、3. に述べた②, ③の学習データに対する結果を示した。周期変動を除いた 88 年 3 月・4 月のデータに関して、学習誤差が表-2 の結果 ($n=4$ の場合) とほぼ同じであることから、図-3 に示したように平均値が大きく変化するデータでも予測モデルの作成に影響しないといえる。ただし、予測誤差の点では周期変動を除かない方がよい結果を与えている。予測対象の月と同じ月の過去 2 年のデータを用いた場合については、学習誤差は 1 月を除いて表-2 の値よりも大きくなっている。予測誤差は、7 月と 10 月の場合に表-2 よりも小さくなっているが、全体的に有意な差があるとはいえない。

表-2 学習および予測の誤差 (その 1)

	'87. 12, '88. 1		'88. 3, 4		'88. 6, 7		'88. 9, 10		'87. 12, '88. 1		
n	学習誤差	予測誤差	学習誤差	予測誤差	学習誤差	予測誤差	学習誤差	予測誤差	n, s	学習誤差	予測誤差
3	11.0 ($m=6$)	0.264 0.566	8.50 ($m=6$)	0.129 0.250	9.51 ($m=6$)	0.0282 0.0392	5.86 ($m=5$)	0.320 0.553	3	11.0 ($m=6$)	0.264 0.566
4	10.7 ($m=8$)	0.255 0.556	7.95 ($m=8$)	0.123 0.248	9.12 ($m=7$)	0.0286 0.0442	5.85 ($m=8$)	0.320 0.572	4	17.8 ($m=8$)	0.267 0.525
5	10.3 ($m=10$)	0.253 0.546	8.00 ($m=10$)	0.126 0.253	9.17 ($m=10$)	0.0334 0.0487	6.07 ($m=8$)	0.360 0.603	5	25.0 ($m=10$)	0.273 0.551
6	9.96 ($m=11$)	0.265 0.546	7.93 ($m=12$)	0.132 0.263	9.24 ($m=10$)	0.0396 0.0469	5.88 ($m=12$)	0.369 0.604	6	35.4 ($m=12$)	0.297 0.576

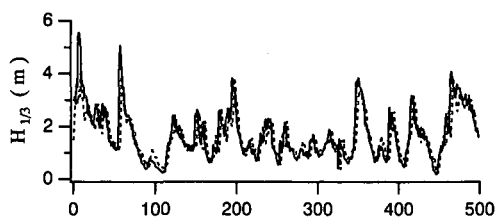


図-5(a) 学習結果 (1987. 12, 1988. 1)

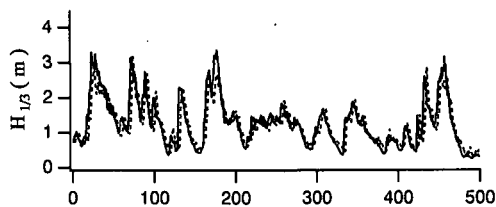


図-6(a) 学習結果 (1988. 3, 4)

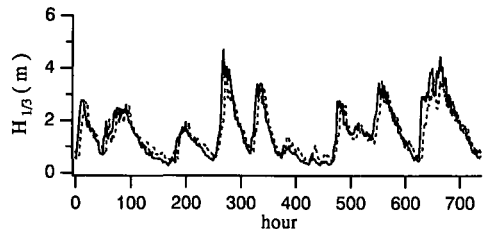


図-5(b) 6時間後予測 (1989.1)

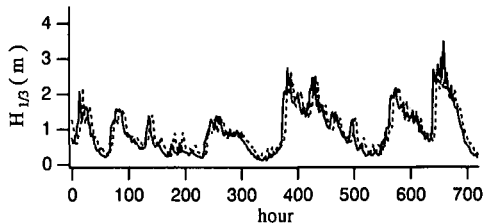


図-6(b) 6時間後予測 (1989. 4)

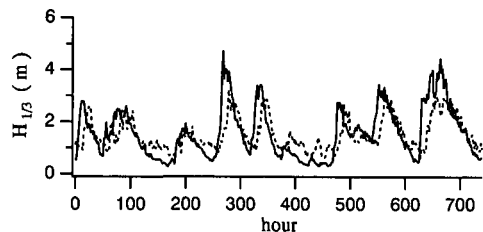


図-5(c) 12時間後予測 (1989. 1)

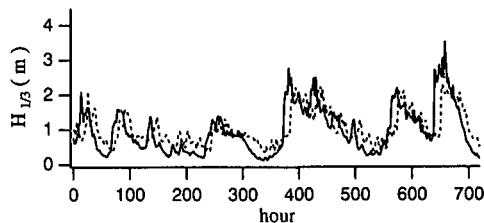


図-6(c) 12時間後予測 (1989. 4)

表-3 学習および予測の誤差 (その2)

'88. 3, 4 (n=4, s=3)		'87. 1, '88. 1 (n=5, s=3)		'87. 4, '88. 4 (n=4, s=3)		'87. 7, '88. 7 (n=5, s=3)		'88. 10, '88. 10 (n=3, s=3)	
学習誤差	予測誤差	学習誤差	予測誤差	学習誤差	予測誤差	学習誤差	予測誤差	学習誤差	予測誤差
8.12	0.182	7.94	0.247	10.7	0.143	14.6	0.0207	11.3	0.277
(m=8)	0.326	(m=10)	0.572	(m=8)	0.281	(m=6)	0.0284	(m=5)	0.509

5. おわりに

本研究では、ニューラルネットワークを用いて有義波高データに対する自己回帰的モデルを作成し、これによる有義波高の予測について検討した。その結果は、従来の統計モデルと同様の問題を有するものであり、自己回帰的モデルの限界を示しているともいえる。しかし、極めて少ないデータで作成したモデルであることを考慮すれば、ネットワークの予測モデルとしての可能性を示していると考えられる。予測性能の向上のためには、他の因子を取り込んだ多変量モデルとすることが挙げられ、これについては今後検討していきたい。

参 考 文 献

片山 立・梶谷雄治・鍛田海平・西田行輝 (1992): 自己増殖型ラジアル基底関数による非線形ダイナミカルシステムの同定と予測, 計測自動制御学会システム3部会合同シンポジウム講演論文集, pp. 187-194.

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, pp. 1-42.

Guedes Soares, C., A. M. Ferreira (1995): Analysis of Seasonality in Non-Stationary Time Series of Significant Wave Height, Computational Stochastic Mechanics, ed. P. D. Spanos, Balkema, pp. 559-568.