

多変量解析法による水質データの解析と赤潮予測について

平 山 勇*・村 上 和 男**

1. ま え が き

海域の環境問題を論じる際に、非常に数多くの水質指標があり、水質調査においても数多くの水質項目についての測定が実施されている。これは、水域の環境という問題が、人間の視覚的なものから、物理的なもの、化学的なもの、更には生物学的なものもあり、その内容も複雑多岐にわたっているためと考えられる。

海域の水質汚濁の典型的な現象として赤潮というものがある。赤潮とは、ある海域でのプランクトンの異常発生による海の色の変化現象を表すもので、時として大きな漁業被害をもたらすことがある。赤潮の発生は、大雨のち高温で日照が続く海面が穏やかなとき、または異質の水塊が混合するような場合に多く起こると言われている。しかし、このような条件が満たされても必ずしも赤潮が発生するとは限らず、まだ不明確な点も多い。

赤潮の発生の予察に関しては、近年かなりの研究例¹⁾がある。しかし、赤潮の発生に関連する因子が多いうえに、プランクトンの生態も複雑なことから、まだ十分に解明されたものとは言えない。また、我国の沿岸域においても数多くの水質調査が実施されているが、これらの調査は月に1回、あるいは週に1回程度の頻度での調査が多く、突発的な現象の赤潮予測にはあまり適さないものと思われる。

本研究は、赤潮予測を目的として実施された水質の時系列データを用いて、統計手法および多変量解析手法による水質データの解析を行うとともに、赤潮の予測手法についての検討を行ったものである。

2. 自動観測機器による水質の連続観測

突発的な赤潮を予測するためには、赤潮に関連のある水質項目の時系列データが必要である。運輸省第三港湾建設局では、このような水質データを取得するために、水質の自動観測装置の開発調査²⁾を実施した。この調査において、次に示す12の項目についての自動観測を西宮沖

の防波堤の脇において実施した。その調査地点の位置を図-1に示す。測定項目は、水温、塩分、pH、溶存酸素、濁度、アンモニア態窒素、硝酸態窒素、化学的酸素要求量、リン酸態リン、総リン、ケイ酸、クロロフィル *a* である。なお、表-1にこれらの水質指標の記号を示し、以下、表記の長いものに対してはこれを用いるものとする。ここに示した水質項目は、赤潮に関連が深く、かつ自動計測が可能なものが選定されている。調査は、昭和58～60年度にかけて実施されているが、ここでは、主として60年度のデータについての解析を実施した。

図-2に示すのは60年度に実施された chl-*a* の変化曲線である。測定は、表層と4m層とで1時間毎に交互に実施されている。観測期間は7月10日～8月17日までの40日間である。図中の矢印は、調査時に観察者によって赤潮が発生していると認められた期間を示している。他の項目に対しても chl-*a* と同様の測定であるが、水温、塩分、DO に関しては、両層で1時間毎の測定を行っている。

なお、水質自動観測装置による調査結果の詳細については参考文献(2)を参照されたい。

3. 多変量解析法による水質データの解析

多変量解析とは、互いに相関のある多変量(多種類の特性値)のデータのもつ特徴を要約し、かつ所与の目的に応じて総合するための手法である³⁾。水質データには

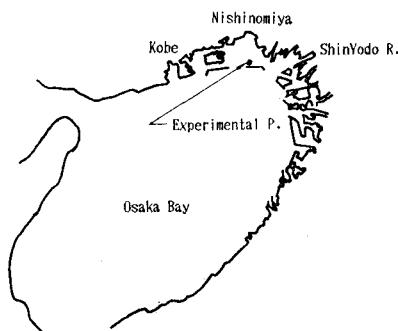


図-1 調査地点の位置

* (前)運輸省第三港湾建設局 神戸機械整備事務所
** 正会員 工修 運輸省港湾技術研究所 海洋水理部

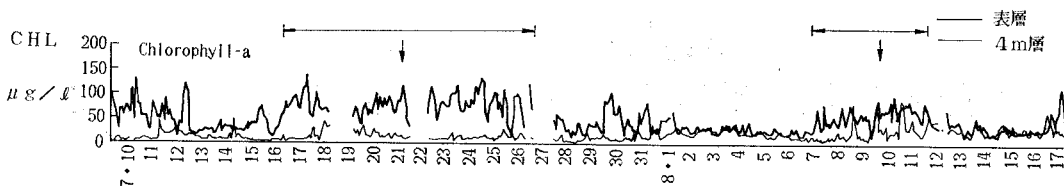


図-2 chl-aの測定結果(太線: 表層の値, 細線: 4m層の値)

数多くの項目がある。これらの項目の中には、水の濁りの指標、有機物濃度の指標、あるいは自然環境に類する指標などがある。また、互いに正の相関が強いもの、逆に負の相関が強いもの、あるいは互い全然相関のない項目もある。このような数多くの水質項目から、水質汚濁というある漠然とした概念を取り出すのに多変量解析法は適した方法であると考えられる。

(1) 水質項目の特性

主成分分析を行う際に、各項目のデータの分布形状が正規分布に近い方が情報の主成分への集約度が高くなると言われている⁴⁾。そこで、各々の水質項目の分布特性についての検討を行った。表-1に、各測定項目の最大値、最小値、平均値、および標準偏差を示す。また、各々のデータに対して、ルート交換および対数変換を施して、それぞれのヒストグラムを求め、その歪度と尖度から最も正規分布に近い形態を同表の最右欄に示した。この表から、水温、塩分、pHの値の変動幅が小さいこと、および $\text{PO}_4\text{-P}$ 、 $\text{NH}_4\text{-N}$ 、chl-aなどの変動幅が大きいことがわかる。また、水質項目の分布特性としては、自然的因子に関する項目は原データで、また、プランクトンの増殖あるいは赤潮状態の結果を示すような項目に関しては対数正規分布になっているものと考えられる。また、赤潮の原因物質と考えられている栄養塩や有機物に関する項目は、中間的なルート変換により正規分布に近づく分布形状を示している。これは、自然的な因子とプランクトンの増殖のような突発性の因子の両方から影響を受けているためと考えられる。

(2) データ相互間の相関と主成分分析

7月14日～16日の測定記録の一部を図-3に示す。赤潮の指標となる chl-a の濃度が高い時には COD 濃度も高いこと、また、 $\text{NH}_4\text{-N}$ 、 $\text{PO}_4\text{-P}$ の濃度の低いことがわかる。このような水質データ項目間の相互の関連を調べるために、各々の項目間の相関を求めた。その結果を表-2に示す。解析は、観測期の前半と後半に分けて行った。ここでは前半の結果を示す。この表において、右上半分については表層の、また左下半分については4m層のデータの相関である。chl-a との相関をみると、DO と COD と正の相関が強く、また、 $\text{NH}_4\text{-N}$ 、 $\text{PO}_4\text{-P}$ 、 SiO_2 、 $\text{NO}_3\text{-N}$ などの栄養塩と負の相関が強い

表-1 各水質項目の特性、および分布形状

測定項目・記号	単位	最大	最小	平均	標準偏差	分布
水温 TEMP	°C	28.7	22.7	26.27	1.15	○
pH	pH	9.02	7.32	8.36	0.24	○
塩分 SAL	‰	28.8	5.0	20.65	6.59	○
溶存酸素 DO	mg/l	15.7	3.9	9.32	2.39	L_n
濁度 SS	mg/l	16.9	1.9	6.19	2.66	L_n
アンモニア NH ₄ -N	µg/l	1128	1	148.6	152.2	R
遊離素 化学的酸素要求量 COD	mg/l	8.8	1.3	3.88	1.41	R
リン酸リン PO ₄ -P	µg/l	79.0	7.0	27.98	11.64	○
総リン TP	µg/l	193	15	79.8	30.5	L_n
ケイ酸 SiO ₂	mg/l	3.74	0.01	0.69	0.85	R
クロロフィル a chl-a	µg/l	142	12	49.0	27.5	L_n
硝酸態窒素 NO ₃ -N	µg/l	631	80	316.8	131.3	R

○: 原データ, R: ルート変換, L_n : 対数変換

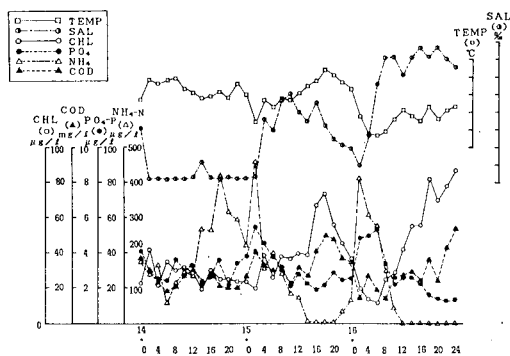


図-3 観測結果例(60年7月14日～7月16日)

ことがわかる。

この相関行列を用いて主成分分析を実施した。主成分分析とは、互いに相関のある多種類の特性値のもつ情報を、互いに無相関な少数個の総合特性値に要約するものである。したがって、多くの水質指標から水質の総合特性値を得ようとするものである。表-3に、前半のデータに対する固有値とその累積寄与率を示す。この表において、その主成分の持つ情報の寄与率が平均以上、すなわち1以上の固有値を有する主成分についてのみ示す。また、これらの主成分と各々の水質項目との相関を示す因子負荷量の結果も同表に示す。表-3から、表層においては第1主成分が57%の寄与率をもち、第3主成分

までで全体の 83 % の情報を集約している。この第 1 主成分は、chl-a、DO、pH、および COD との正の相関が高く、また、PO₄-P、NO₃-N、SiO₂ と負の相関が高いことから、植物プランクトンによる赤潮に関する総合特性値を示しているものと思われる。また第 2 主成分は、濁度と正の相関が高く、塩分と負の相関が高いことから、河川起源の水塊を示す特性値であると思われる。また、4 m 層についても同様の解析を行っており、第 1 主成分に対しては表層と同様に植物プランクトンに関する特性値を示し、第 2 主成分に対しては濁りを示す特性値となっているものと思われる。しかし、情報の集約度は表層ほど高くない。

(3) 重回帰分析

赤潮を示す指標として chl-a の濃度がよく用いられる。これは、chl-a が植物プランクトンの光合成活動に関与する物質を含み、故に植物プランクトンの現存量を表す最もよい物質であることによる。したがって、ここでは chl-a の量を赤潮の水質指標と仮定して、重回帰分析による chl-a の量の算定を行った。重回帰分析とは、

ある目的変数 y_i を、いくつかの説明変数 x_{pi} で表そうとするもので、次式によって表される。

$$y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi} + \varepsilon_i \quad \cdots \cdots \cdots (1)$$

ここに、 β_n は x_n の y に対する偏回帰係数であり、 ε は残差である。

重回帰分析を行う際に、求める重回帰式が目的変数の特性をよく表すためには、数多くの説明変数の中から適切な変数を選択する必要がある。ここでは、chl-a の目的変数に対して、最も一般的に用いられている変数増減法による変数選択を行った。表-4 に、60 年度データについての重回帰分析の結果を、観測の前半、後半の表層および 4 m 層の別に示す。選択変数の取込み基準 F_{in} および追出し基準 F_{out} はそれぞれ 2.0 と設定した。各々のデータに対して選択された変数は若干異なっているが、全体的に共通して選ばれているものは DO、SiO₂、NH₄-N、PO₄-P などである。前半・表層のデータについて、表-4 の偏回帰係数より求めた重回帰式と実測値の chl-a との比較を図-4 に示す。かなり良い近似を示していると思われる。実測値と算定値との相関係数は

表-2 相 関 係 数 表 (右上半分: 表層, 左下半分: 4 m 層)

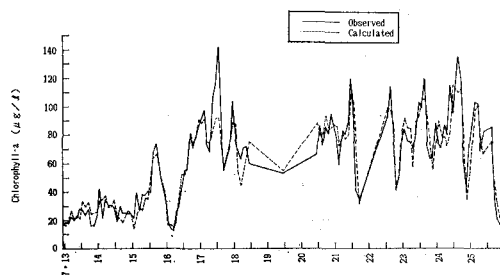
	TEMP	pH	SAL	DO	SS	NH ₄ -N	COD	PO ₄ -P	TP	SiO ₂	chl-a	NO ₃ -N
TEMP	*	0.51	0.16	0.37	0.42	-0.47	0.75	-0.24	-0.10	-0.62	0.54	-0.56
pH	0.08	*	0.44	0.61	0.24	-0.67	0.52	-0.64	-0.44	-0.48	0.64	-0.56
SAL	-0.83	-0.14	*	0.38	-0.15	-0.71	0.44	-0.55	-0.62	-0.77	0.60	-0.77
DO	0.30	0.33	-0.46	*	0.46	-0.70	0.43	-0.67	-0.38	-0.37	0.75	-0.50
SS	0.00	0.16	0.04	0.02	*	-0.22	0.33	-0.13	0.13	-0.17	0.46	-0.27
NH ₄ -N	-0.27	-0.02	0.33	-0.61	-0.05	*	-0.56	0.81	0.64	0.75	-0.89	0.74
COD	0.55	0.16	-0.38	0.15	0.05	-0.02	*	-0.25	-0.24	-0.72	0.58	-0.71
PO ₄ -P	-0.28	-0.29	0.52	-0.75	-0.01	0.62	0.05	*	0.70	0.46	-0.76	0.52
TP	-0.38	-0.20	0.30	-0.26	0.11	0.09	-0.41	0.10	*	0.53	-0.53	0.56
SiO ₂	-0.75	-0.12	0.54	-0.29	0.11	0.22	-0.61	0.13	0.38	*	-0.73	0.89
chl-a	0.44	0.07	-0.60	0.72	-0.11	-0.57	0.15	-0.61	-0.20	-0.31	*	-0.73
NO ₃ -N	-0.71	0.00	0.39	-0.09	-0.11	0.12	-0.70	-0.13	0.45	0.81	-0.13	*

表-3 主成分分析における固有値と因子負荷量

主成分	表 層			4 m 層			
	1	2	3	1	2	3	4
TEMP	0.621	0.558	-0.253	0.844	0.298	-0.115	0.150
pH	0.757	0.063	0.206	0.252	-0.151	0.758	-0.433
SAL	0.739	-0.474	-0.301	-0.825	0.049	0.127	-0.055
DO	0.728	0.137	0.508	0.635	-0.567	0.134	-0.087
SS	0.327	0.736	0.359	-0.016	0.011	0.701	0.657
NH ₄ -N	-0.928	0.141	-0.123	-0.544	0.528	0.103	-0.281
COD	0.721	0.404	-0.389	0.586	0.567	0.162	-0.030
PO ₄ -P	-0.765	0.337	-0.435	-0.565	0.717	-0.071	0.024
TP	-0.650	0.572	-0.065	-0.512	-0.274	-0.122	0.494
SiO ₂	-0.853	0.009	0.450	-0.768	-0.435	0.087	-0.010
chl-a	0.917	0.077	0.201	0.698	-0.469	-0.226	0.017
NO ₃ -N	-0.881	-0.004	0.310	-0.640	-0.667	-0.021	-0.185
固有値	6.873	1.765	1.281	4.633	2.480	1.22	1.011
寄与率	57.3%	14.7	10.7	38.6%	20.7	10.2	8.4
累積寄与率	57.3%	72.0	82.7	38.6%	59.3	69.5	77.9

表-4 変数増減法による重回帰分析の偏回帰係数

	60 年度 表層		60 年度 4 m 層	
	前 半	後 半	前 半	後 半
TEMP	0	0	0	-0.171
pH	0	-0.779	-0.370	0
SAL	0	-0.075	-0.125	0
DO	0.372	0.675	1.207	0
SS	0.457	1.115	0	0.324
NH ₄ -N	-0.033	0	-0.027	-0.058
COD	0	0.262	0	0.410
PO ₄ -P	-0.013	-0.021	0	-0.007
TP	0	0.162	0	0
SiO ₂	-0.464	0.372	0	-1.221
NO ₃ -N	0.020	0	0	0
相関	0.94	0.90	0.77	0.92



図—4 変数増減法による重回帰式と実測値との比較

0.94 とかなり高い値を示した。

4. 赤潮の予測

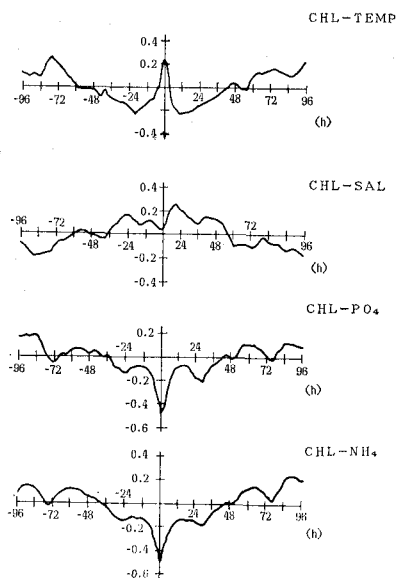
重回帰分析によって、かなりの精度で chl-a 量の算定ができた。しかし、図—4 に用いられた説明変数には DO と濁度が含まれている。これらは chl-a 量が増加したことによる DO と濁度の上昇の結果を示し、赤潮発生の結果を説明する指標と成り得ても、これを chl-a 量の予測として用いることはできない。そこで、ここでは赤潮の予測として、予め説明変数を定めておき、説明変数の時間レベルを目的変数よりも前に設定した形での重回帰分析を実施した。

図—3 に示された水質データの変化曲線から、chl-a 量の増加する 12~16 時間前ごろに $\text{NH}_4\text{-N}$ と $\text{PO}_4\text{-P}$ の濃度が上昇していることがわかる。この結果から、これらの栄養塩濃度の上昇が赤潮発生の前兆となることが予想される。そこで、これらの水質項目の時系列データを用いて赤潮の予測を試みた。

まず、代表的な水質項目についての周期特性と相互相関について若干の検討を行った。周期特性に関しては、各々の水質項目の変動のエネルギースペクトルを求めた。スペクトル形状はここでは示せないが、全般的に 10 時間、24 時間、32 時間前後にピークを持った分布特性を示したものとなった。また、図—5 には chl-a とその他の項目との相互相関係数を示す。この図において、横軸の正の値が chl-a の時間レベルよりも前の時間の値との相関を示している。図—5 より、相関関数には約 24 時間の周期性があることがわかる。

以上の検討を踏まえて、重回帰分析による chl-a の予測を行った。用いた説明変数は、前記の相関係数、重回帰分析の他に、クラスター分析などの結果を考慮して、水温、塩分、 $\text{NH}_4\text{-N}$ 、 $\text{PO}_4\text{-P}$ の 4 項目に選定した⁵⁾。

これら 4 項目の 12 時間前、24 時間前、32 時間前、48 時間前のデータを用いて重回帰分析による chl-a の算定を行った。その結果の一例を図—6 に示す。また表—5 には、それらの偏回帰係数の値と実測値と予測値の相関係



図—5 chl-a と他の項目との相互相関関数

数を示す。図—6 より、12 時間前のデータを用いた場合に、chl-a のデータを、急激な現象までは再現できていないが、かなり良く追隨しているものと思われる。データのずらし時間を大きくすると、増加・減少の傾向はあっているが、その近似度はかなり悪くなっていく様子がみられる。12 時間前のケースの偏回帰係数の値をみると、 $\text{NH}_4\text{-N}$ の値が時間差無しのケースと比べて負から正に転じ、他については同符号であることがわかる。このことから、12 時間前のデータを用いての赤潮の予測は、 $\text{NH}_4\text{-N}$ の濃度が前兆となりうるものと考えられる。したがって、本調査時での植物プランクトンによる赤潮発生の過程において、窒素が制限因子であったものと考えられる。

次に、時系列データの予測手法の一つとして、カルマン・フィルターによる手法があり、水質予測においても Beck⁶⁾ の DO-BOD の予測例がある。本研究においても、chl-a 量の時系列データを用いてカルマン・フィルターによる赤潮予測を試みた。その結果を図—7 に示す。全体的な傾向は若干あっていると言えるが、予測値の揺れ幅が大きく、まだ満足できるものではない。赤潮のような突発性の現象を、時系列データから予測することは、かなり難しいものと思われる。

5. おわりに

昭和 60 年 7 月から 8 月にかけて、大阪湾の西宮防波堤脇で実施された水質自動観測装置の開発調査の際に得られたデータを用いて、水質データの処理および水質の予測手法の検討を行った。本研究において明らかになっ

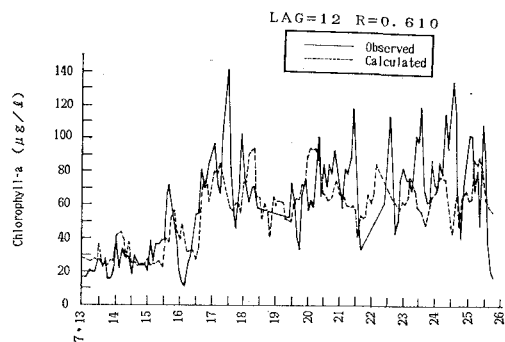


図-6(1) 重回帰分析による chl-a の予測 (12 h)

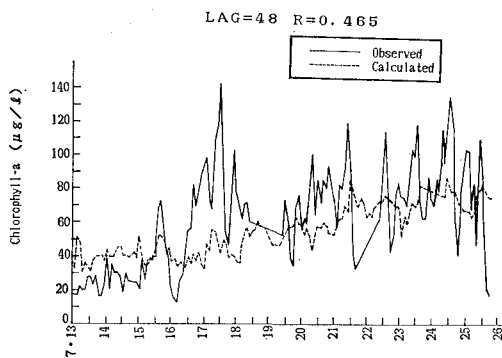


図-6(2) 重回帰分析による chl-a の予測 (48 h)

表-5 chl-a の予測における偏回帰係数

	TEMP	SAL	NH ₄ -N	PO ₄ -P	相 関
0 時間	0.147	0.011	-0.036	-0.020	0.84
12	0.034	0.086	0.017	-0.007	0.61
24	0.065	0.042	-0.003	-0.012	0.55
32	0.025	0.047	-0.023	-0.002	0.49
48	0.148	0.059	0.027	-0.003	0.46

たこととして、①水質を表す総合特性値としての第1主成分は、植物プランクトンによる赤潮に関する特性を表している。②赤潮の指標として chl-a を使い、変数増減法による重回帰分析により、比較的良好な精度で chl-a 量の算定ができた。③また、赤潮予測の試みとして、重回帰分析による chl-a 量の予測を行い、12 時間前のデータを用いた場合、ある程度の予測ができた。

しかし、赤潮という現象はプランクトンの異常繁殖という突発性の現象であって、水質データの統計処理あるいは時系列データ処理という、赤潮発生メカニズムがブラック・ボックスでの解析法だけでは、現象を解明するには困難が伴う。今後は更に多くの水質データを収集するとともに、赤潮発生メカニズム、メカニズムを考慮した生態系モデル等の研究、および本研究のような時系列データ解析等を考慮した研究が必要であると思われる。

最後に、本データは水質自動観測装置の開発調査委員会（委員長：藤永奈良教育大学学長）での調査結果を提

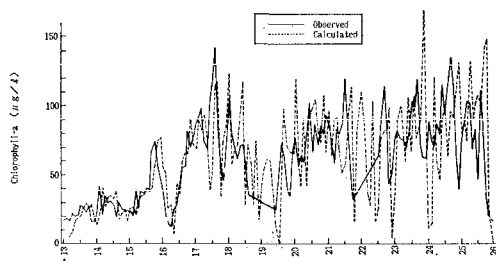


図-7 カルマン・フィルタによる chl-a の予測

供していただいたものである。同調査委員会の委員の方々、および関係者の方々に深く感謝致します。

参 考 文 献

- 1) 例えば、大内 晟：広島湾奥部における赤潮予報研究の現状，沿岸海洋研究ノート，20 巻 1 号，1982。
- 2) 昭和60年度水質自動観測装置開発に関する調査研究報告書，運輸省第三港湾建設局神戸機械整備事務所，1986。
- 3) 奥野忠一・久米 均・芳賀敏郎・吉澤 正：多変量解析法一改訂版一，日科技連，430 p.，1981。
- 4) 吉見 洋・岡 敬一・井口 潔・関根廣子：相模川水系の水質解析について，水質汚濁研究，5 巻 4 号，pp.193~199，1982。
- 5) 村上和男：水質自動観測データの解析と赤潮予測への応用，港研報告，26 巻 2 号，(印刷中)。
- 6) Beck, M.B.: Modelling of dissolved oxygen in a non-tidal stream, Mathematical Models in Water Pollution Control, John Wiley & Sons, pp.137~166, 1978。