

自動観測システムの利用-II

— 対数変換手法を用いた塩分の評価 —

吉本 宗央

Utilization of Automatic Observation System in the Ariake Sea -II Valuation of Salinity by Logarithmic Transformation

Muneo YOSHIMOTO

まえがき

内湾である有明海は、河川水による塩分の変動が大きく、生物に与える影響も少なくない。ノリ養殖における病害との関係¹⁾も深いし、貝類においても斃死の発生や稚貝の発生との関係²⁾が旧くから指摘されている。また、村上³⁾や平野⁴⁾の報告にみられるように、河川水の動向を把握する指標としても多くの解析例がある。当水試では、ノリ養殖漁期を中心に自動観測装置を稼働させ連続観測を行っており、この塩分データは養殖管理の指標として利用され、また、病害や作柄あるいは各種の調査・試験結果との因果関係^{1,5)}などについても調べられてきた。

ところで、有明海の生物は一般に広範囲の塩分に適応できる⁶⁾が、それには多くの(生理的な)労力が必要であること⁷⁾は知られており、塩分の変動により常に大小の影響を受けているわけである。塩分の影響度を判断する目安の一つは、それが通

常の生息条件から逸脱している程度を知ることであろう。そのためには“通常”の限度、つまり、例年と比較してどの程度以上が特異値であるのかを知っておく必要がある。水温に関しては、その変化は正規分布⁸⁾の範囲内にあるといわれ、所定の方法で比較的容易にその程度を判断できる。しかし、塩分に関しては、一般に正規分布から逸脱しているといわれており、確立された方法が示されているわけではない。その手法について吟味してみることは、塩分データを利用する上で基本的なことだけに、あながち無駄なことではないであろう。

本報では、早津江川自動観測塔の観測値を用い、塩分データを対数変換することにより、どの程度画一的な高低判断が可能になるか、また、生データを使用する場合と比較してどの程度評価精度が向上するかについて検討した。

方 法

塩分は、1971~1991年の9~12月に早津江川自動観測塔で観測された毎日の昼間満潮時観測値を用いた。昼間満潮時の決定は別報⁹⁾記載の満潮時水温の場合と同様である。観測装置のセンサー(鶴見精機製)はフロートで上下し、水面下約40cmの塩分を測定する。測定範囲は10~35‰、精度は±

0.1‰である。

1. 数的処理

塩分(S)の対数変換式としては、 $-\text{LOG}(33-S)$ と $-\text{LOG}(33/S-1)$ の2つを用い、変換データはそれぞれ10倍して表わした。いずれも常用対数である。

2. 生データと変換データの分布形の比較

次の3指標を導入して比較した。

1) 正規性の検定⁸⁾

統計数値の評価には、正規分布(Normal distribution)の使用が一般的であるので、変換後データがどの程度正規分布に近づいたかを判断する指標として正規性の χ^2 検定を行い、その数値を示した。数値が小さいほど正規性は高い。

2) 平均値の順位

高塩分側から数えた平均値の順位を、(順位/全データ数) $\times 100$ として%で表わした。

3) 歪度⁸⁾

分布の左右の不均衡性(アンバランス)を示す指標として示した。数値が大きいほど分布の左右が不均等であることを示し、符号は偏りの方向を示している。

結果及び考察

1. 塩分の分布

図1に1971~1991年まで、21年間の9~12月における塩分(生データ:以下無変換という)の経時変化を示した。図2には月別の分布を示し、表1にはこの分布形に関する3指標を示した。日塩分を月別に図2のように集計(一括)して同一母集団からのものとして取り扱うことには問題があるが、ここでは全体の分布傾向を大まかに把握する手段として使用することにした。

9、10月には日塩分の低下することが多く、変動幅は大きい。11、12月には30%前後でほぼ安定する。図2、表1に明らかなように塩分は、全般に右側(プラス側)に偏在し左側(マイナス)に延びた出現頻度分布を示すが、左への延びはこの时期的な変動傾向を反映して9、10月で長く、11、12月には短縮する。さらに、累年平均値(図1)

は、出現の仕方に偏りがあるため、常に低塩分側に偏って計算され、9、10月にその傾向が著しいことがわかる。

塩分のこのような特徴は、湾奥部混合域では塩分の高い方の出現が沖合水の塩分によって制約され、無降雨日が続いてもある限度を超えることがないのに対し、低い方へは降水によって容易に低下(ゼロまで可能)するために生じると考えられる。

さて、統計値を評価する場合、それが正規分布⁸⁾するならば別に問題はない。この分布では平均値が中央値に等しく、また左右の出現確率が均等で、平均値からの距離により測定値の高低(大小)評価が容易に行なえるからである。しかし、正規性の検定値の大きさや平均値の偏在など表1の指標は塩分が正規分布しないことを示している。特に

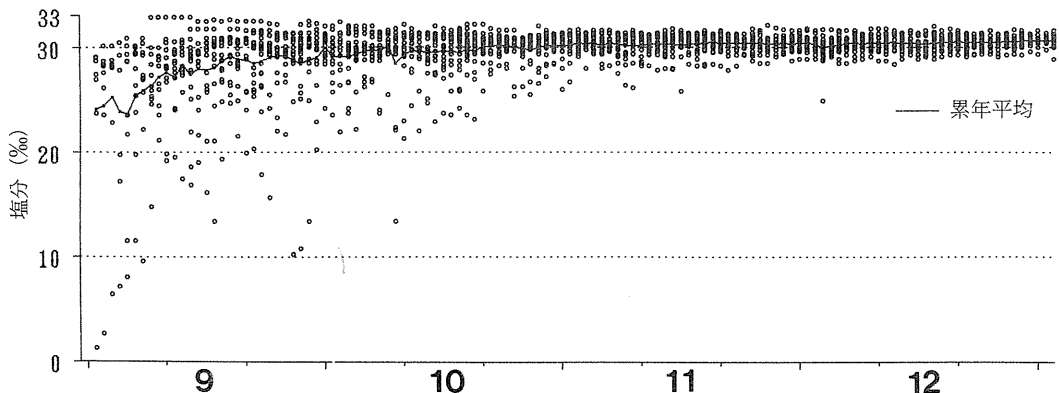


図1 昼間満潮時塩分の経日変化(1971~1991年の9~12月)

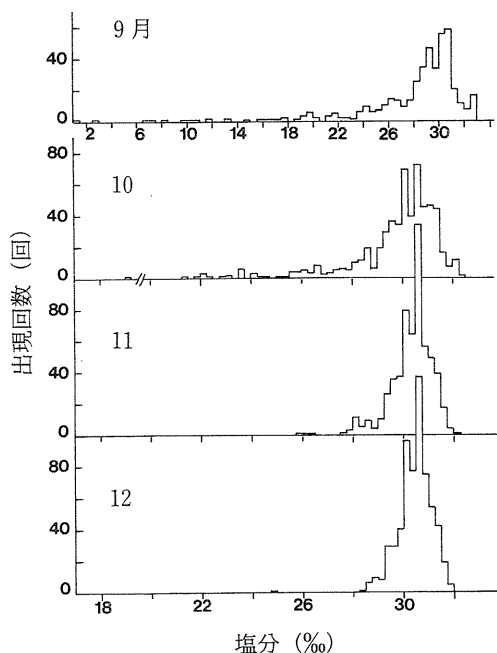


図2 月別にみた塩分の分布形

9～10月の塩分に、無条件にこの分布を適用してしまうと、評価の誤差は相当大きくなると考えられる。

2. 対数変換の導入

対数変換によってデータの集中した部分を展開し、ある程度正規分布に近づけるのであるが、通常、集中分布では偏りが左側にあることが多く、これを変換する方法⁹⁾は多数紹介されている。しかし、塩分では左右が通常とは全く反対の分布になっているのである。そこで、原点を右側にとって高塩分側を展開することを考える。原点としては、過去の観測記録(図1)から高塩分側の極限值(≒前述の沖合水の塩分)と思われる33‰を採用した。

こうして、 $-\text{LOG}(33-S)$ と $-\text{LOG}(33/S-1)$ の2式を用いて、塩分(S)の対数変換を行なった(以下、順に変換-1, 2という)。対数の()内の式はそれぞれ次のような意味を持っている。

33-S : 塩分がS‰の海水において、1量の沖合水から希釈水(河川水)に移行した塩分量

表1 月別にみた塩分の分布¹⁾形に関する指標

指標	データ	月/9	10	11	12
正規性の検定 ²⁾	生データ	240.1	288.1	134.2	101.7
	変換-1	72.4	48.0	32.4	63.9
	変換-2	100.1	59.6	32.4	63.9
平均値の順位 ³⁾	生データ	70.1	67.0	62.0	62.4
	変換-1	53.4	53.1	51.1	44.3
	変換-2	55.9	54.5	51.1	44.3
歪度 ⁴⁾	生データ	-2.66	-2.67	-1.35	-1.23
	変換-1	0.98	-0.45	-0.11	0.14
	変換-2	-0.19	-0.74	-0.22	0.05

1) 1971～1991年の毎日の満潮時塩分を月別に累計したもの

2) χ^2 検定

3) 高塩分側から数えた平均値の順位/全データ数

4) 左右の均衡度

33/S-1 : 塩分がS‰の海水において、沖合水(=1)に対する希釈水(河川水)の割合
村上が報告³⁾した河川水の沖合水による希釈率(α)とは $33/S-1 = 1/(\alpha-1)$ という関係がある。

つまり、変換-1, 2は沖合水に対する河川水の影響度を対数で表わすという意味を持っているのである。

図3に、変換-1, 2の月別出現頻度分布を示した。また、その分布形に関する指標を表1に示す。無変換データに比べて正規分布に近くなり、平均値は中央値に近づいたことがわかる。以上から、対数変換が有効な手法であると推測できる。

3. 対数変換を用いた評価

日塩分の評価は、毎年の同じ月の同じ日に観測されたデータ(以下、同月日データという)間の比較でなされるべきであろう。それは、同一母集団に属する観測値を得るという観点から、同じ気候条件下において比較し、降水の確率を公平なものにしておく必要があるからである。

さて、同月日データの標準化を一般に用いられている次式で定義する。

標準化データ = (観測値 - 累年平均) / 標本標準偏差。

但し、累年平均と標本標準偏差は1971～1991年の同月日データからの計算値

図4は、無変換データ、変換-1, 2について、

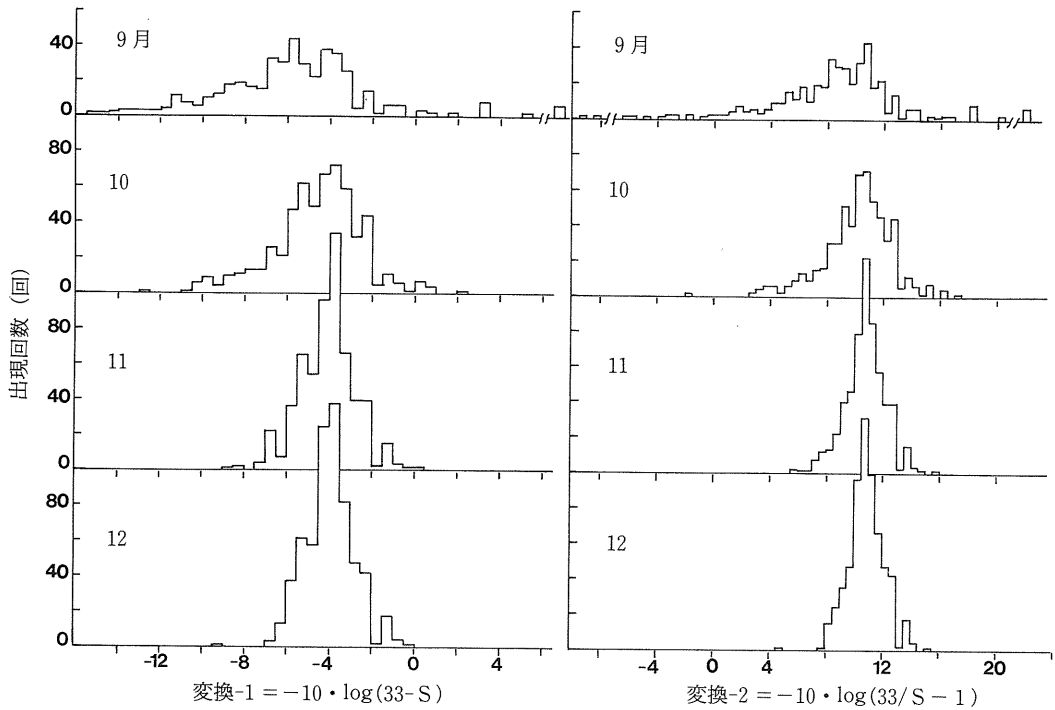


図3 月別にみた対数変換した塩分の分布形

同月日データ(1971~1991年まで21か年のデータ)を4か月分全て標準化し、月別に集計して頻度分布で示したものである。表2にはこの分布に関する正規性の検定、平均値の順位と歪度を示した。同月日データが正規分布するなら、その標準化データは全て平均値0、標準偏差1の正規分布に近いはずである。したがって、月別に集計した同月日データも平均値0、標準偏差1の正規分布で近似できると思われる。つまり、図4の分布について正規性を検定し、平均値の順位、歪度を調べる(表2)が、同月日データがどの程度正規分布に近いかを判定することになる。

無変換の分布に関する3指標は、いずれも変換-1, 2に比べて相当大きい。9, 10月には平均値の中央値からの偏りと左右のアンバランスは大きい(3指標は大きい)。11, 12月には正規分布にやや近づくものの、全期間を通じては χ^2 検定による正規性が認められない。

一方、変換-1, 2の分布でも、9, 10月には χ^2 検定による正規性は認められない。しかし、平均値は中央値に近く、左右もほぼ均等である。11, 12月になると変換-1, 2の分布には χ^2 検定による正規性が認められる。10月以前と比べると平均値はほぼ中央値に等しくなり、左右は均等と考えてもよい。

以上のように、塩分は無変換データのままで偏りのある(集中)分布をするが、対数変換によって正規分布、あるいはそれに近い分布に訂正することができる。

4. 誤差の範囲

塩分を無変換データのまま使用した場合の過誤の程度及び変換-1, 2によってどの程度まで理論値に近づいたかについて調べてみることにする。

普通、標準化データの評価は、それが正規分布(理論値)に従うことを前提に7ポイントを配置して行なうことが多い。そのポイントは中央値

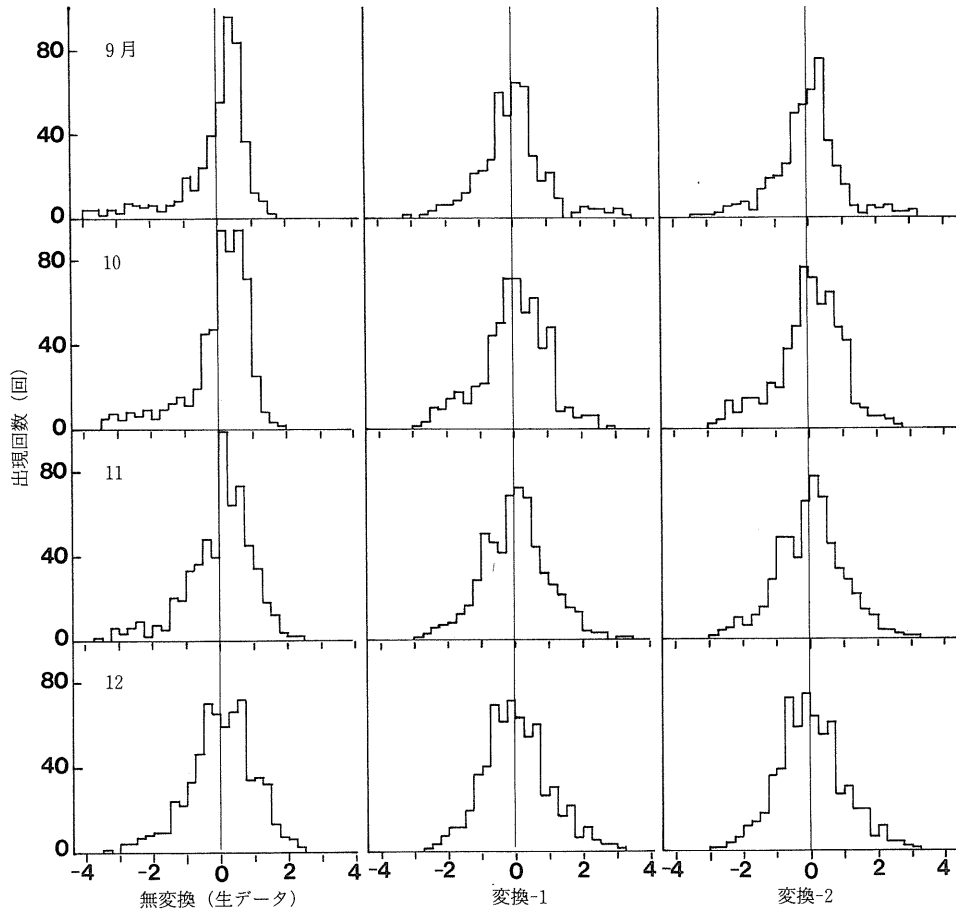


図4 標準化データの月別分布 (同月日データを標準化し、月別に累計したもの)

(0)と ± 0.6 , ± 1.3 , ± 2.0 で、理論値では中央値と平均値は等しく、中央値から左右に測った出現頻度は、 ± 0.6 が22.6%, ± 1.3 が40.3%, ± 2.0 が47.7%である。これらは、一般に $-0.6 \sim 0.6$ を『平年並』、 $0.6 \sim 1.3$ ($-0.6 \sim -1.3$)を『やや高(低)』、 $1.3 \sim 2.0$ ($-1.3 \sim -2.0$)を『かなり高(低)』、 $2.0 \sim (-2.0 \sim)$ を『甚だ高(低)』と統一した呼称で評価されている。評価がその頻度に基づいていることはいうまでもない(以下、平年並をA、分布の外側に向かってB~D ($-B \sim -D$)ランク、 $22.6\% \times 2$, $22.6 \sim 40.3\%$, $40.3 \sim 47.7\%$, $47.7\% \sim$ をランクの期待頻度という)。

表3には、図4の分布について平均値の中央値

表2 同月日標準化データの分布¹⁾形に関する指標

指標	データ 月/	9	10	11	12
正規性の検定 ²⁾	生データ	212.4	257.7	78.7	24.6
	変換-1	68.6	31.7	*14.4	*14.0
	変換-2	77.2	49.0	20.7	*12.1
平均値の順位 ³⁾	生データ	66.4	65.3	60.1	51.9
	変換-1	51.1	53.3	51.3	47.6
	変換-2	54.1	54.3	52.5	47.6
歪度 ⁴⁾	生データ	-1.71	-1.53	-0.84	-0.43
	変換-1	0.39	-0.36	-0.01	0.27
	変換-2	-0.01	-0.51	-0.08	0.22

1) 1971~1991年の満潮時塩分を同月日ごとに標準化し、月別に累計したもの

2) χ^2 検定: *正規性が認められるもの

3) 高塩分側から数えた平均値の順位/全データ数

4) 左右の均衡度

表3 同月日標準化データの分布における各標準化距離までの出現頻度

月 / データ	標準化距離 ¹⁾	-2.0	-1.3	-0.6	0.6	1.3	2.0
	理論頻度% ²⁾	(47.7)	(40.3)	(22.6)	(22.6)	(40.3)	(47.7)
9 生データ	生データ	43.3	40.1	31.7	26.1	48.4	50.0
	変換-1	47.4	42.7	27.4	31.1	43.6	45.5
	変換-2	46.3	41.8	29.2	30.6	44.1	46.4
10 生データ	生データ	43.3	38.8	32.7	22.7	48.6	50.0
	変換-1	46.1	39.0	27.4	22.3	44.5	47.7
	変換-2	45.5	39.0	27.7	22.2	44.7	48.3
11 生データ	生データ	45.4	40.8	26.3	24.0	44.8	49.4
	変換-1	47.1	42.0	23.8	26.0	40.4	47.9
	変換-2	46.6	42.0	23.8	25.6	40.8	47.9
12 生データ	生データ	46.2	40.9	26.7	21.5	41.3	48.7
	変換-1	48.2	43.0	23.7	24.4	40.1	46.3
	変換-2	48.2	42.7	24.2	23.9	40.1	46.5

1) 平均値からの標準化距離, 2) 正規分布における理論出現頻度

表4 正規分布なみの出現頻度を得るための標準化距離 (補正值)

月 / データ	出現頻度の実現点 ¹⁾						
	-C	-B	-A	中央値	A	B	C
9 変換-1	-2.06	-1.17	-0.44	0.03	0.40	1.07	2.53
	-2.20	-1.21	-0.37	0.07	0.42	1.00	2.30
10 変換-1	-2.23	-1.39	-0.44	0.06	0.61	1.13	1.99
	-2.35	-1.42	-0.40	0.08	0.61	1.09	1.88
11 変換-1	-2.11	-1.20	-0.57	0.02	0.51	1.29	1.97
	-2.15	-1.20	-0.57	0.03	0.52	1.27	1.98
12 変換-1	-1.88	-1.17	-0.58	-0.07	0.55	1.32	2.17
	-1.92	-1.18	-0.57	-0.06	0.55	1.32	2.15

1) |A|, |B|, |C|は平均値からの出現頻度が, それぞれ22.6, 40.3, 47.7%となる点 (標準化距離) で, 正規分布の理論値では0.6, 1.3, 2.0 (表3) である。

からのズレ (確率) と中央値をゼロとした場合の各ポイントまでの出現頻度を示した。

無変換の場合, 9月においては, 期待した頻度よりも約14%も多い観測値をAランクと誤認し, -Dは多めに, 逆にDはゼロと判定してしまう。10月もほぼ同じ。11月にはAランクの過誤(期待頻度からのズレ)が約5%, B, Cが7%, -Bが4%で他は2%前後である。12月にはAの過誤は同じであるが, ±C, ±Dはほぼ期待頻度に近い。

一方, 変換-1, 2でも9月にはAランクの過誤は無変換と変わらない。しかし, 左右はほぼ均等

で±C, ±Dは期待頻度にかかなり近い。10月にはAの過誤は約5%に低下する。11, 12月には全般に期待頻度に近いと考えてよい程度まで過誤は小さくなっている。

表4に変換-1, 2について計算した中央値の位置と期待頻度 (正規分布のA~D評価) どおりとなる各ポイントを標準化距離 (標本平均からの距離/標準偏差) で示しておく。9, 10月のA~Dランクの評価にはこの数値を使用すべきである。

11, 12月においても厳密にはそうであるが, 前述の正規分布の理論値 (0.6, 1.3, 2.0) をそのまま

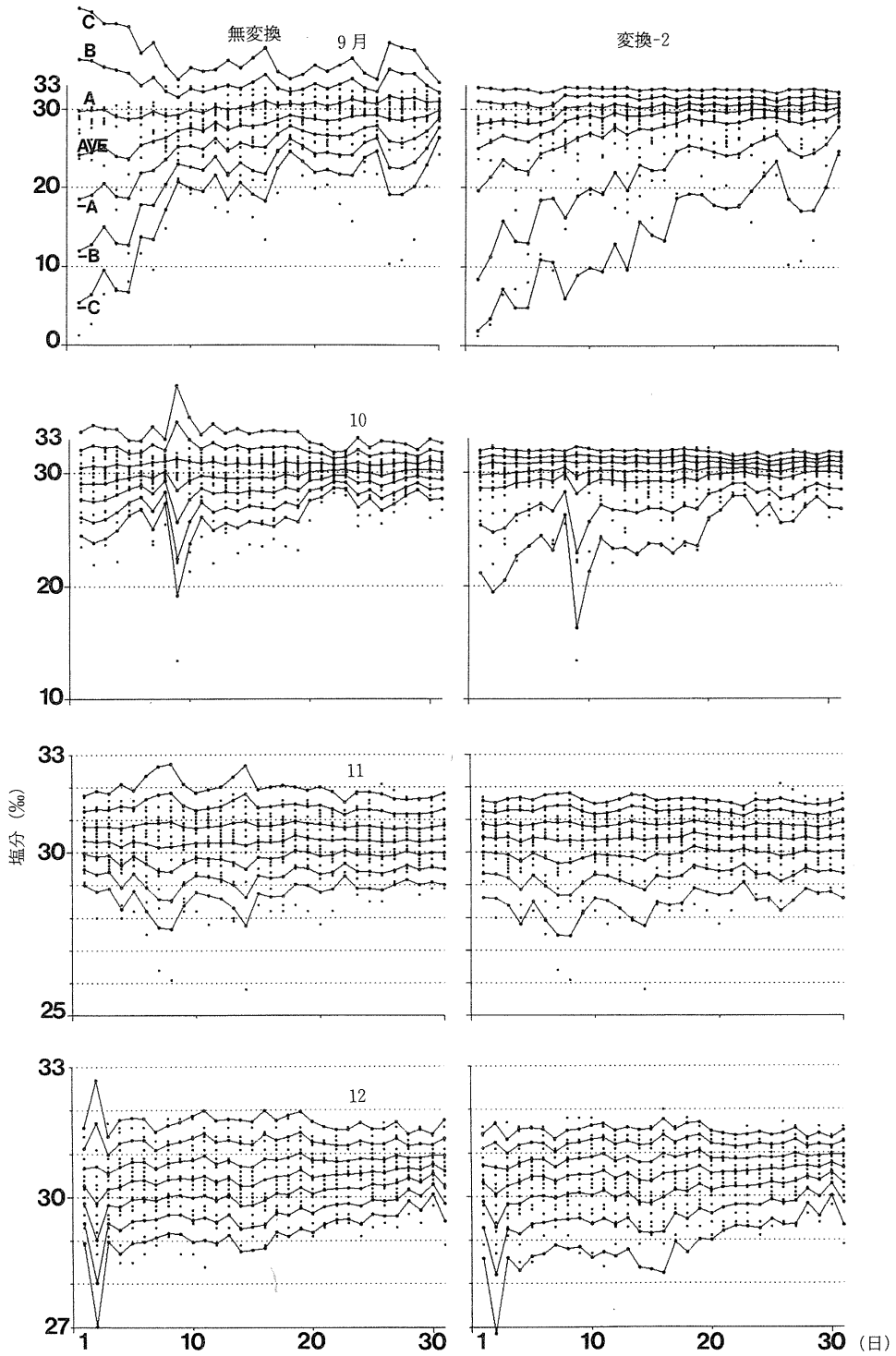


図5 対数変換-2と生データ使用のランク評価区分点*の比較

(* 0~|A|の出現確率 $P(A)=22.6$, $P(B)=40.3$, $P(C)=47.7\%$ となるように計算した点)

用いても大きな誤差はないであろう。

5. 対数の再変換

新しい観測値をA～Dランクで評価するような場合、前年度までの同月日データから変換-1, 2を用いて各ランクの区分点を計算し、その値を対数から塩分値に再変換しておくといふ。高低の評価が即座にでき、変換値をそのまま用いるよりも観測値の位置を実感的に捉え易い利点がある。

変換-1, 2の計算値を再び塩分に変換する式(再変換式)を示す。

但し、 α は標準化データのランク区分点で、正規分布のA～Dにおける $\pm 0.6, \pm 1.3, \pm 2.0$ に相当する。また、 $S(\alpha)$ は区分点の塩分であり、標本平均は、対数値の前年までの累年平均、標準偏差も対数値の偏差のことである。

変換-1 : $S(\alpha) = 33 - 10^{-Y}$, $Y = \bar{X} + \alpha\sigma$

(\bar{X} : 標本平均, σ : 標本標準偏差)

変換-2 : $S(\alpha) = 33 / (1 + 10^{-Y})$, $Y = \bar{X} + \alpha\sigma$

(\bar{X} : 標本平均, σ : 標本標準偏差)

上式の $\alpha = 0$ の時が平均値である。

図5に再変換した対数-2の区分点を示し、無変換の場合と比較しておく。但し、対数-2の9, 10月の α は表4に示した訂正後の値を用いた。11, 12月の α 値は理論値どおり($\pm 0.6, \pm 1.3, \pm 2.0$)である。

なお、変換-1と変換-2のどちらを用いても変換性能(変換後の正規分布への近似度、平均値の中央値化、均等性など)はほとんど変わらない。ただ変換-1を用いた場合、 $(33 - S)$ のSが理論上、 $-\infty$ まで可能であるため、実用にあたって、塩分変化の激しい時期(9月上旬のように)にはDランクがマイナスに計算されてしまうことがある。このような時期には変換-2を用いる必要がある。

要 約

1. 塩分を生データから対数変換して正規分布(均等分布: Normal distribution)に近づけることを試みた。

対数変換式としては $-\text{LOG}(33 - S)$ と $-\text{LOG}(33/S - 1)$ の2つを用いた。

2. 9～12月の塩分は全般に右側(プラス方向: 高塩分側)に偏し左側(マイナス)に延びた出現頻度分布を示す。9, 10月には低塩分になることが多く、変動幅は大きい。11, 12月には30%前後でほぼ安定する。左への延びはこの時期的な変動傾向を反映して9, 10月で長く、11, 12月には短縮する。

3. このため、無条件に正規分布を適用してしまうと、評価(平年値に比べての高低評価)の誤差

は相当大きく、特に9, 10月の塩分では顕著であった。

4. 対数変換によって9, 10月の塩分の出現頻度分布はほぼ均等になり、11, 12月には正規分布と認められるようになった。

5. 9, 10月については現在行なわれている正規分布のランク評価の出現頻度に等しくなる標準化距離を求めた。

6. 新規観測値の評価は、9, 10月は補正後の標準化距離を用い、11, 12月は従来の正規分布の基準で各ランクの基準値(区分点: 対数値)を計算し、さらにこの値を塩分に再変換することで画一的に行なえる様になった。

文 献

1) 馬場裕文・山下康夫 1985: 佐賀県有明海のノリ養殖における漁海況情報の利用。佐有水試報, (9), 39-44.

2) 佐賀県水産試験場 1916: 有明海鯉研究会報告書。

3) 村上彰男 1968: 有明海湾奥部における塩素量分布。筑後川調査報告附属資料, 環境-3, 水産庁,

- 1-15.
- 4) 平野敏行 1968: 有明海湾奥部における筑後川水の稀釈混合について. 筑後川調査報告附属資料, 環境-5, 水産庁, 11-21.
 - 5) 川村嘉応・山下康夫・異儀田和弘・古賀秀昭 1985: プイロボットの運用技術 秋芽網期の漁場環境と養殖ノリの生長. マリンランディング計画昭和59年度委託事業報告書, 佐賀有明水試.
 - 6) 異儀田和弘 1986: 六角川感潮域における稚仔魚等の分布について. 佐有水試報, (10), 35-45.
 - 7) 内田清一郎 1970: 魚類生理. 川本信之編, 302-325, 恒星社厚生閣, 東京.
 - 8) George W. Snedecor and William G. Cochran 1967: STATISTICAL METHODS, 6th edition. (畑村又好・奥野忠一・津村善郎共訳 1972: 統計的方法. 岩波書店, 東京.)
 - 9) 吉本宗央 1992: 自動観測システムの利用-I 水温の日周変化と潮候. 本誌, 25-38.
 - 10) 岩井重久・住友 恒・松岡 譲 1980: 水質データの統計的解析. 森北出版, 東京.