

# 渤海湾海水光谱特性的数值分析

## ——复相关分析\*

平仲良

(中国科学院海洋研究所)

### 一、序 言

从遥感信息中提取海洋信息是海洋遥感应用的核心问题。

从遥感数据提取海洋数据,现在已经有多种方法与模型。各种方法都有它们的长处和短处。从目前发表的遥感文章来看,大多还是使用数理统计模型。它的优点是方法比较简单,且比较精确。缺点是有很大的局限性,它的结果只有在特定的海区、特定的大气条件和飞行高度等实验条件下才是正确的。作者在《渤海湾海水光谱特性的数值分析》<sup>[1]</sup>一文中使用的正是 R. W. Johnson (1978, 1979) 使用的数理统计方法和线性回归方程。

为了提高由遥感数据得到的计算值与海洋实测数据之间的相关系数,即提高由遥感数据计算海洋数据的置信度,本文提出一种新的方法——复相关分析方法,建立回归方程。计算表明,复相关方法建立的数学模型比 R. W. Johnson 的方法——简相关方法(有的数理统计书称单相关方法)建立的模型有更好的结果。

R. W. Johnson 用逐次回归分析法寻求与自变量(叶绿素 a 浓度或悬浮泥沙含量)相关系数较大的依变量(某些波长的海水反射率),并建立数学关系。而依变量(海水的光谱反射率)不单与一个自变量(叶绿素 a 浓度或悬浮泥沙含量等)有关,而是与二个或二个以上的自变量(叶绿素、泥沙和其他溶解的、悬浮的物质)都有关系,因此我们用复相关的方法来分析。

本文应用了《渤海湾海水光谱特性的数值分析》一文中的原始数据<sup>[1]</sup>。对文中未叙述或未充分叙述的问题,本文作了某些必要的说明和补充,因此本文是那篇文章的继续、补充和发展。

### 二、方 法

为了说明复相关的方法及便于与简相关方法比较,需要将简相关的方法作简要的回顾。

首先是将在渤海湾 8 个站位航空测得的 8 条海水光谱反射曲线数字化,即将每条曲线以 20 毫微米为间隔,分割为若干波段,测得每一波段的反射值,除以该波长的白板的反射值,再乘上白板的反射率,即得每一站位、各波段的反射率数据(见表 1)。

然后逐个波段计算 8 个站位的海水反射率与各个海洋参数(如叶绿素 a 浓度、悬浮泥沙含量和水色)的相关系数。对每一个海洋参数,选择与它相关系数最大的一个或较大的几个波段的反射率与它建立回归方程,于是就得到每一个海洋参数的回归方程。

复相关分析是逐个波段计算海水反射率与二个或二个以上的海洋参数(但主要是叶绿素和悬浮泥沙二个海洋参数)的复相关系数,选择二个复相关系数较大的反射率与每一个海洋参数建立回归方程<sup>[2]</sup>,具体步骤如下。

1. 用一个二元一次方程表示二个自变量

\* 中国科学院海洋研究所调查研究报告第 1142 号。本文承方国洪副研究员指教,谨此致谢。

表 1 各站海水的光谱反射率和海洋实测数据

| 站 位                       | 11 A         | 11 B  | 11 C  | 11 D  | 12 A  | 12 B  | 12 C  | 12 D  |
|---------------------------|--------------|-------|-------|-------|-------|-------|-------|-------|
| 叶绿素a (毫克/米 <sup>3</sup> ) | 2.287        | 1.987 | 0.927 | 1.613 | 2.627 | 1.687 | 0.976 | 2.056 |
| 悬浮泥沙 (毫克/升)               | 160          | 40    | 80    | 40    | 110   | 140   | 160   | 60    |
| 水色 (牌号)                   | 21           | 15    | 15    | 20    | 21    | 19    | 14    | 19    |
| 波长 (毫微米)                  | 海水的光谱反射率 (%) |       |       |       |       |       |       |       |
| 460                       | 15.13        | 14.81 | 14.19 | 14.83 | 15.83 | 16.26 | 15.27 | 14.48 |
| 480                       | 15.62        | 15.26 | 14.79 | 15.57 | 16.71 | 17.03 | 17.45 | 14.96 |
| 500                       | 19.17        | 16.26 | 16.72 | 17.07 | 18.53 | 18.47 | 19.77 | 15.99 |
| 520                       | 18.12        | 17.52 | 17.64 | 18.38 | 19.37 | 19.50 | 20.03 | 17.39 |
| 540                       | 20.32        | 19.46 | 19.01 | 20.30 | 21.92 | 21.88 | 23.48 | 17.87 |
| 550                       | 22.16        | 20.52 | 20.16 | 21.90 | 22.53 | 22.96 | 24.42 | 19.47 |
| 560                       | 23.03        | 22.07 | 20.75 | 23.87 | 24.40 | 24.72 | 25.85 | 20.96 |
| 580                       | 24.34        | 23.53 | 23.00 | 25.67 | 27.43 | 27.35 | 28.91 | 23.58 |
| 600                       | 25.78        | 23.67 | 23.11 | 25.67 | 28.16 | 28.35 | 28.59 | 23.83 |
| 620                       | 25.38        | 21.55 | 21.25 | 23.59 | 27.01 | 27.18 | 24.87 | 21.73 |
| 640                       | 24.83        | 20.58 | 20.43 | 22.81 | 27.10 | 27.27 | 23.80 | 20.81 |
| 660                       | 24.06        | 19.08 | 19.04 | 22.80 | 26.74 | 27.32 | 22.02 | 18.87 |
| 680                       | 24.12        | 18.81 | 18.19 | 23.36 | 26.29 | 27.22 | 20.16 | 18.20 |
| 700                       | 21.80        | 16.57 | 16.75 | 21.35 | 24.29 | 24.60 | 17.56 | 16.79 |
| 720                       | 15.92        | 11.69 | 12.15 | 13.17 | 19.63 | 20.13 | 13.03 | 12.71 |
| 740                       | 13.57        | 6.58  | 7.21  | 7.50  | 11.47 | 12.23 | 6.62  | 6.75  |
| 760                       | 11.11        | 5.20  | 6.87  | 5.82  | 8.96  | 9.41  | 5.61  | 5.50  |
| 780                       | 10.82        | 5.09  | 6.80  | 6.63  | 8.58  | 8.96  | 5.31  | 5.10  |

(叶绿素a浓度和悬浮泥沙含量) 与依变量(海水光谱反射率) 之间的关系。

$$Y_{\lambda} = A_0 + A_1 x_1 + A_2 x_2 \quad (1)$$

$Y_{\lambda}$  为波长为 $\lambda$ 的反射率;

$x_1$  为叶绿素a浓度;

$x_2$  为悬浮泥沙含量;

$A_0, A_1, A_2$ 为波长为 $\lambda$ 的回归系数。

2. 用最小二乘法计算波长为 $\lambda$ 的系数

$A_0, A_1, A_2$ 的估计值 $a_0, a_1, a_2$  (见表2)。

估计值 $a_1, a_2$ 反映叶绿素a浓度、悬浮泥沙含量与海水光谱反射率之间的关系, 称为依变量 $y$ 依自变量 $x_1, x_2$ 的回归系数。

3. 用海洋实测数据 (叶绿素a浓度、悬浮泥沙含量) 和波长为 $\lambda$ 的系数 $A_0, A_1, A_2$ 的估计值 $a_0, a_1, a_2$ 代回回归方程:

$$Y_{\lambda} = a_0 + a_1 x_1 + a_2 x_2 \quad (2)$$

得到每个波段每个站位海水反射率的计算值 $R_{\lambda}$ 计。

4. 逐个波段计算海水反射率实测值与计算值之间的相关系数 $r$ , 可以证明<sup>[4]</sup>所得 $r$ 为二个自变量与依变量之间的复相关系数。

5. 选择复相关系数 $r$ 最大的波长的实测反射率 $R_{500}, R_{740}$ 与海洋实测数据进行复回归, 得到以 $R_{500}, R_{740}$ 为二个自变量海洋参数为依变量的复回归方程:

$$\begin{aligned} \text{叶绿素a (毫克/米}^3\text{)} = & 3.871 - 20.377 \\ & R_{500} + 16.849 R_{740} \end{aligned} \quad (3)$$

$$\begin{aligned} \text{悬浮泥沙 (毫克/升)} = & -377.082 + \\ & 2429.82 R_{500} + \end{aligned}$$

$$455.006 R_{740} \quad (4)$$

$$\begin{aligned} \text{水色牌号} = & 24.096 - 80.792 R_{500} \\ & + 91.656 R_{740} \end{aligned} \quad (5)$$

式中 $R_{500}, R_{740}$ 为波长为500毫微米、740毫微米海水的反射率。

### 三、相关性的比较

用每一站位的遥感实测数据 $R_{500}, R_{740}$ 代入上面叶绿素a浓度、悬浮泥沙含量、水色的回归方程, 就可得到每一站位各个海洋参数的计算值。再计算每一海洋参数的计算值与实测值之间的相关系数。计算结果表明, 用复相关方法得到的结果比简相关方法得到的结果好。为便于比较, 表3列举了两种方法得到的相关系数数据。

表 2 各波长的回归系数

| 波 长<br>(毫微米) | $a_0$     | $a_1$      | $a_2$      |
|--------------|-----------|------------|------------|
| 460          | 13.597715 | 0.398665   | 0.0083856  |
| 480          | 14.783084 | -0.0661823 | 0.0132406  |
| 500          | 15.285051 | -0.0214335 | 0.0263197  |
| 520          | 18.032524 | -0.5166667 | 0.0157970  |
| 540          | 18.988276 | -0.3971048 | 0.0236277  |
| 550          | 20.735323 | -0.6033012 | 0.0220790  |
| 560          | 21.337911 | -0.0367382 | 0.0204695  |
| 580          | 24.596179 | 0.0466005  | -0.0047624 |
| 600          | 22.340951 | 0.3648049  | 0.0306403  |
| 620          | 18.299123 | 1.422945   | 0.0347081  |
| 640          | 23.533553 | -2.156379  | 0.0393367  |
| 660          | 14.477046 | 2.240043   | 0.0426245  |
| 680          | 13.359734 | 2.839128   | 0.0385133  |
| 700          | 11.829807 | 2.837489   | 0.0327535  |
| 720          | 6.0308299 | 2.792556   | 0.0403168  |
| 740          | 0.1436131 | 2.706003   | 0.0427159  |
| 760          | 1.0198433 | 1.702360   | 0.0344945  |
| 780          | 1.0643589 | 1.601279   | 0.0330277  |

表 3 用简相关与复相关方法得到的计算  
值与实测值之间的相关系数

| 海 洋 参 数  | 简 相 关 | 复 相 关 |
|----------|-------|-------|
| 叶绿素 a 浓度 | 0.67  | 0.68  |
| 悬浮泥沙含量   | 0.73  | 0.91  |
| 水 色      | 0.70  | 0.78  |

#### 四、置信度的比较

用数学模型计算得到的海洋参数,即用遥感数据预报的海洋参数到底有多大的可信水平,我们把相关系数代入公式  $F = \frac{r^2}{1-r^2}(n-2)$ , 得到统计量 F 值,再用查表的方法得到各个海洋参数计算值的置信度<sup>[3]</sup>,见表 4。

由表 4 可以看出,使用复相关方法,悬浮泥沙含量和水色的计算值的置信度有明显提高。叶绿素 a 浓度的计算值的置信度也有提高,

但不甚明显。

表 4 用简相关与复相关方法得  
到的计算值的置信度

| 海 洋 参 数  | 简 相 关 | 复 相 关 |
|----------|-------|-------|
| 叶绿素 a 浓度 | 0.90  | 0.90  |
| 悬浮泥沙含量   | 0.95  | 0.99  |
| 水 色      | 0.90  | 0.95  |

#### 五、讨 论

1. 如前所述,海水的光谱反射率与浮游藻类(叶绿素的载体)和悬浮泥沙含量都有关系,可以用方程  $Y = A_0 + A_1x_1 + A_2x_2$  表达。反射率是叶绿素和悬浮泥沙两种作用(除水的作用外的叠加。某一波长的反射率如蓝绿波段 480, 520 毫微米的反射率与叶绿素 a 浓度的简相关系数较大,而 500 毫微米的反射率与叶绿素 a 浓度的简相关系数较小,乍看起来,480, 500 毫微米的反射率与叶绿素 a 浓度的相关性比 500 毫微米的反射率显著,其实不一定如此,因为没有考虑泥沙的作用,如果把泥沙对反射率的作用考虑进去,480, 520 毫微米的反射率与叶绿素 a 浓度的相关性就不比 500 毫微米的反射率显著。因此,只有在考虑叶绿素(或泥沙)与反射率关系的同时,也考虑泥沙(或叶绿素)的作用,即叶绿素与泥沙对反射率的共同作用,在这个基础上(即选取复相关系数最大的反射率)建立数学关系才更全面更精确地体现一个自变量(叶绿素或泥沙)与依变量(反射率)的关系。因此复相关系数不仅反映两个自变量与依变量综合关系的紧密程度,也较全面、较精确地反映一个自变量与依变量关系紧密的程度,这就是复相关方法比简相关方法优越的原因。

2. 如表 2 所示,在光谱的蓝绿区(480—560 毫微米),  $a_1$  值均为负值,在光谱的红区均为正值,即叶绿素 a 浓度增加,使蓝绿区反射率降低而红区反射率升高,这是叶绿素在像渤海湾这样悬浮泥沙含量很高的海域的光谱效应

(它与叶绿素在清澈海水中的光谱效应不同), 这就为在飞机上卫星上使用的各种多光谱扫描仪, 包括陆地卫星上使用的谱带较宽的多光谱扫描仪定量检测叶绿素浓度提供了理论依据。悬浮泥沙含量较高、信号较大, 检测就更容易。

3. 为了提高用遥感数据预报海洋数据的精度, 即提高遥感数据与海洋实测数据之间的相关系数, 一方面要提高遥感数据与海洋实测数据的测量精度, 增加站位(即增加样品数); 另一方面, 要通过各种方法建立更确切地反映遥感数据与海洋实测数据之间客观规律的数学模型。计算表明, 复相关方法就比简相关方法

优越, 但不一定就是最好的方法, 其他方法如偏相关(有的数理统计书称净相关)方法, 非线性回归等是否更好, 还需一一试验。

### 参 考 文 献

- (1) 平仲良, 1984. 渤海湾海水光谱特性的数值分析. 海洋科学集刊22: 307—312页。
- (2) 范福仁, 1980. 净回归与复回归及净相关与复相关(第八章). 生物统计学. 江苏科学技术出版社, 308—346页。
- (3) 阿姆斯特特, B. L., 1978 (彭兴文译, 1978). 附录B统计用表. 可靠性数学. 科学出版社, 353—371页。
- (4) 陈敦隆, 1982. 相关与回归(第八章). 海洋科学研究中的概率统计方法. 海洋出版社, 217—275页。

## DATA ANALYSIS OF SPECTRAL SIGNATURE OF SEA WATER IN BOHAI BAY

### —COMPLEX CORRELATION ANALYSIS

Ping Zhongliang

(Institute of Oceanology, Academia Sinica)

#### Abstract

Spectral reflectance of sea water depends on both chlorophyll-a concentration and suspended sediment, so the complex correlation analysis approach was adopted in data analysis of spectral reflectance of sea water to select the independent variables (reflectance) which have the highest complex correlation coefficient with the dependent variables (chlorophyll-a concentration and suspended sediment content) and to determine the complex regression equations for quantitatively relating sea truth measurements to remotely sensed data. The results obtained show that the complex correlation analysis approach is better than the single correlation analysis approach in increasing 1) the correlation coefficients of remotely sensed values calculated from regression equations to sea truth measurements and 2) the confidence.