

# 海洋生物—光学算法研究

## ——南黄海叶绿素浓度的估算\*

傅克忖 曾宪模 李宝华

(国家海洋局第一海洋研究所, 青岛 266003)

曾晓起

荒川久幸

(青岛海洋大学水产系, 青岛 266003) (东京水产大学, 东京 108)

**提要** 根据1995年9月中日合作南黄海现场观测资料,对南黄海生物—光学算法进行研究。首先对南黄海现场多光谱反射比数据进行因子分析,解反射比协方差矩阵的特征问题;解出特征值和对应的特征向量,求出主因子,然后建立叶绿素 $a$ 浓度对主因子的多元线性回归方程,并对算法进行回报和预报检验。计算结果表明,该算法具有较高精度,同时也显示出天气条件对估算精度产生较大影响。

**关键词** 因子分析 多光谱反射比 叶绿素 $a$ 浓度

海洋浮游植物光合作用的速率用初级生产力表示,叶绿素 $a$ 是计算初级生产力的主要因素之一。卫星海洋水色遥感的主要目的是估算全球海洋初级生产力。通过观测海水上涌辐射光谱来推测海中叶绿素浓度的生物—光学算法,Johnson (1978), Gordon 等 (1980), Morel 等 (1980), Smith 等 (1981) 等学者的研究都证明了浮游生物色素浓度与水体不同波长辐射比的指数相关性,即  $C = a(R_i / R_j)^b$ 。算法中主要差别在于使用不同的系数  $a$  和  $b$ 。在大洋 I 类水体中,算法具有相当精度,但在复杂水质的沿岸及海湾中,谱段和统计参数要根据具体水域和季节所采集的实测数据来确定(傅克忖等, 1995)。本文利用多光谱反射比数据,采用因子分析,建立叶绿素浓度对主因子的多元线性回归方程,用以估算南黄海叶绿素浓度。

### 1 方法

#### 1.1 观测时间及海上观测站的布设

于1995年9月1—5日,中日双方研究人员乘东京水产大学实习船“神鹰丸”,在南黄海实施了“中日黄海可见光遥感基础研究”的海上现场观测。设19个观测站,见图1。

#### 1.2 观测仪器及观测方法

\* 国家自然科学基金资助项目, 495.46004号; 东京水产大学创立百周年纪念学术研究奖励基金。傅克忖, 男, 出生于1941年1月, 副研究员。

收稿日期: 1996年6月3日, 接受日期: 1997年1月15日。

海洋辐射光谱由 PRR-600 型剖面辐射计(Bio-spherical Instrument 公司制)测定。剖面辐射计测定海面及水中下方向辐照度及上方向辐亮度,下方向辐照度通道为 412, 443, 490, 510, 555, 665nm, 另带有光子传感器,上方向辐亮度通道为 412, 443, 490, 510, 555, 665 及 683nm; 上下通道宽为  $\pm 10\text{nm}$ 。测量过程中,同时用 PRR-610 型天空照度计(通道与 PRR-600 相同)监测辐照度的变化。观测由海面以上开始,测量海面以上下方向辐照度及海面上方向辐亮度;其后,仪器潜入水中随深度连续测量至海底附近,并做返程连续测定。辐射计本身带有深度传感器,数据由计算机记录。

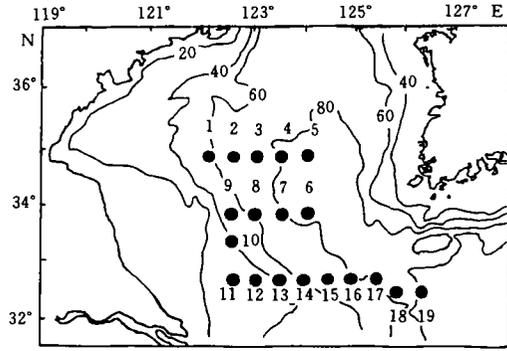


图1 南黄海观测站分布图

Fig.1 Station locations in the South Yellow Sea

叶绿素 *a* 由现场水中荧光计测定(荧光计附属在 OCTOPUS 装置上),从 0 m 水深开始按深度连续(间隔 1 m 取样)测定,数据由计算机记录。同时抽水样,用萃取荧光法测定叶绿素 *a* 浓度。

### 1.3 海水分光反射比特征

南黄海各观测站分光反射比分布见图 2。1—5 观测站海面上分光反射比分布见图 2a,表明各测站各通道反射比峰值不明显。水下 1 m 深度处 1—5 各测站各通道海水反射比分布见图 2b,峰值及极小值明显,各通道峰值位于 490nm 通道处,极小值明显共同位于 665nm 通道处。海面以上观测到的反射比中,包含了海面反射光及海面杂散光成分(图 2a)。由于各站观测时间不同,太阳高度不同,海况不同,海面反射光及杂散光情况不同,数据内含较复杂,不能单一地反映出海洋水体对光的反射特性。因此本文采用能反映出水质特性,不包含海面反射光及杂散光的水下反射比数据来做分析研究。第 8, 16—19 测站的水下 1 m 处分光反射比的分布见图 2c。图中第 16 站反射比曲线峰值位于 555 nm,显示出不同类型海水反射光谱。

16 站水下 1 m 处叶绿素 *a* 浓度为  $0.84 \text{ mg/m}^3$ ,海水中悬浮物质较多,海水透明度为 5.5 m。其它各测站叶绿素浓度仅为  $0.13 - 0.300 \text{ mg/m}^3$ ,透明度为 11.0—17.5 m,海水较清澈。

### 1.4 多光谱反射比数据的因子分析

试图用少数几个变量来代表多个通道反射比场的主要特征,对多通道反射比数据进行因子分析。设有  $m$  个变量  $x_1, x_2, \dots, x_m$  代表  $m$  个通道的海水反射比,每个变量有  $n$  个观测值,则反射比数据可表示为  $x_{it}, i=1, 2, \dots, m, t=1, 2, \dots, n$ , 写成矩阵形式:  $X=(x_{it})$ 。

将原变量矩阵  $X$  进行正交线性变换,得到由  $m$  个新变量  $z_1, z_2, \dots, z_m$  组成的矩阵  $Z=(z_{it})$

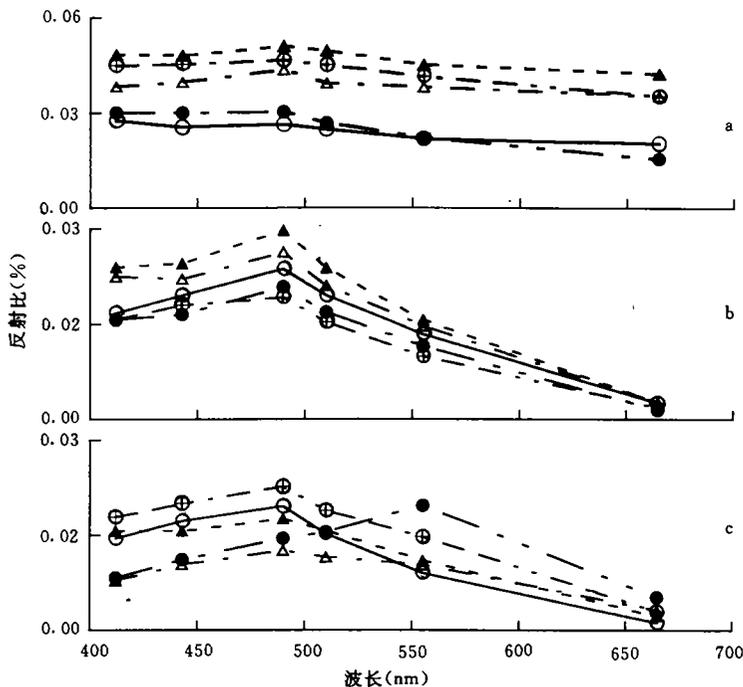


图 2 南黄海各测站分光反射比分布

Fig.2 Spectral reflectance distribution at various stations in the South Yellow Sea

- a. 1—5 测站海面上分光反射比: ○ 1 测站; ● 2 测站; △ 3 测站; ▲ 4 测站; ⊕ 5 测站.
- b. 1—5 测站水下 1m 处分光反射比: ○ 1 测站; ● 2 测站; △ 3 测站; ▲ 4 测站; ⊕ 5 测站.
- c. 8,16—19 测站水下 1m 处分光反射比: ○ 8 测站; ● 16 测站; △ 17 测站; ▲ 18 测站; ⊕ 19 测站.

$$z_{it} = v_{1i}x_{1t} + v_{2i}x_{2t} + \dots + v_{mi}x_{mt} = \sum_{k=1}^m v_{ik} \cdot x_{kt} \tag{1}$$

$$i = 1, 2, \dots, m. \quad t = 1, 2, \dots, n$$

写成矩阵形式

$$Z = V' \cdot X \tag{2}$$

式中,

$$Z = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1n} \\ z_{21} & z_{22} & \dots & z_{2n} \\ \vdots & \vdots & \dots & \vdots \\ z_{m1} & z_{m2} & \dots & z_{mn} \end{bmatrix} \quad V = \begin{bmatrix} v_{11} & v_{21} & \dots & v_{m1} \\ v_{12} & v_{22} & \dots & v_{m2} \\ \vdots & \vdots & \dots & \vdots \\ v_{1m} & v_{2m} & \dots & v_{mm} \end{bmatrix} \quad X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \dots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}$$

$V'$  为矩阵  $V$  的转置矩阵, 变换矩阵  $V$  是待定的。要求最大限度地使方差的贡献集中于前面少数几个新变量上, 并要求新变量彼此相互独立, 即新变量彼此正交。这样我们

可以用较少的新变量描述原变量  $X$  反射比场的主要特征。本研究称新变量  $Z$  为主因子。理论分析表明, 求待定的线性变换矩阵  $V$ , 归结为求解样本原变量  $X$  的协方差矩阵  $S$  的特征问题。即化解为求解方程:

$$(S - \lambda I)V = 0 \quad (3)$$

式中协方差矩阵  $S$  的元素  $s_{ii} = \frac{1}{n} \sum_{i=1}^n (x_{ii} - \bar{x}_i)(x_{ii} - \bar{x}_i)$   $i, j=1, 2 \dots m$  其中,  $x_{ii}, x_{ji}$  为原变量  $X$  矩阵的元素;  $\bar{x}_i$  为  $X$  的第  $i$  个变量各观测值的均值;  $\bar{x}_j$  为  $X$  的第  $j$  个变量各观测值的均值。

式(3)欲使  $V$  有解, 则必须满足

$$|S - \lambda I| = 0 \quad (4)$$

式(4)为原变量  $X$  协方差矩阵  $S$  的特征方程。

本文利用 jacobi 方法, 求  $S$  的特征值  $\lambda$  和对应的特征向量  $V$ 。  $m$  阶矩阵有  $m$  个特征值  $\lambda_1 \geq \lambda_2 \geq \lambda_3 \dots \geq \lambda_m$ 。对应的特征向量为  $v_1, v_2, \dots, v_m$ 。根据(2)式求出主因子  $z_1, z_2, \dots, z_m$ 。

## 2 结果与讨论

本文设定所选取的主因子对方差的贡献率为 99% 来确定选取主因子的个数, 选取了有阳光照射的 10 个测站(1—5, 8, 16—19 站)的数据进行分析。运算结果选出了 3 个主因子  $z_1, z_2, z_3$ , 第 1、第 2、第 3 主因子对反射比方差的贡献率分别为 66.37%, 31.85% 和 0.97%。3 个主因子反映了全部信息的 99.19%。所得的因子荷载表明, 第一主因子主要由 412, 443, 490, 510nm 通道的反射比线性组合而成; 第二主因子主要由 555, 665nm 通道的反射比线性组合而成。因子荷载随波长的分布见图 3。

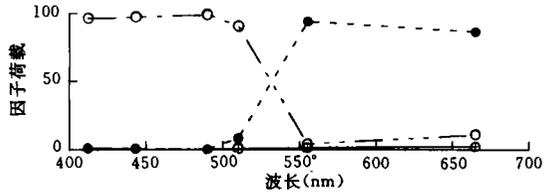


图 3 因子荷载按波长的分布

Fig.3 Distribution of factor loading with wavelength

○ 66.37%; ● 31.85%; ⊕ 0.97%.

### 2.1 叶绿素浓度模式建立

由于  $z_1, z_2, z_3$  前 3 个主因子能反映的信息量为 99.19%, 后 3 个主因子  $z_4, z_5, z_6$  的作用可忽略。试建立叶绿素  $a$  的多元线性回归方程, 探讨叶绿素  $a$  与主因子之间的数值关系, 方程如下:

$$\ln(C) = B_0 + B_1 z_1 + B_2 z_2 + B_3 z_3 \quad (5)$$

式中,  $C$  为各测站叶绿素浓度 ( $\text{mg} / \text{m}^3$ )。系数  $B_0, B_1, B_2, B_3$  由多元线性回归运算求出。

由 10 个测站的叶绿素浓度数据及前 3 个主因子做三元线性回归得到:

$$\ln(C) = -1.6275 - 0.174 \cdot z_1 + 0.270 \cdot z_2 - 0.340 \cdot z_3 \quad (6)$$

子样全相关系数  $R=0.9713$ 。(6)式即为利用主因子估算叶绿素  $a$  的计算模式。

### 2.2 模式检验

**2.2.1 F 检验** 取显著性水平  $\alpha=0.01$ 。查  $F$  分布表得临界值  $F_{0.01}(3, 6)=9.78$ ,  $F=33.36 > F_{0.01}(3, 6)$ , 线性显著。

**2.2.2 回报检验** 将  $n$  个样本的前三个主因子代入模式(6), 得到  $n$  个计算值  $C_e$ , 本文中  $n=10$ 。计算值  $C_e$  与实测值  $C_m$  列入表 1 中。10 个样品回报均方根误差为  $0.0262\text{mg}/\text{m}^3$ , 平均相对误差为 10.06%, 其中 St. 1 站回报检验相对误差达 30.66%。分析其原因, 此测站观测时间为早晨 06:00–07:00, 太阳高度还很低, 辐射计接收光能中天空反射光成分较多所致。

**表 1 南黄海叶绿素  $a$  浓度的计算值与实测值的比较<sup>1)</sup>**

Tab.1 Comparison between estimated and measured values for chlorophyll  $a$  concentration in the South Yellow Sea

站位	1	2	3	4	5	8	16	17	18	19
$C_m$	0.140	0.170	0.160	0.140	0.130	0.170	0.840	0.300	0.160	0.180
$C_e$	0.183	0.178	0.139	0.131	0.136	0.173	0.785	0.296	0.173	0.191
$C_m - C_e$	-0.043	-0.008	0.021	0.009	-0.006	0.033	0.055	0.004	-0.013	-0.011
$E(\%)$	30.66	4.97	12.82	6.21	4.65	19.45	6.56	1.20	7.86	6.22

1)  $C_m$  为实测值;  $C_e$  为计算值;  $C_m - C_e$  为残差;  $E(\%)$  为相对误差(表 2、表 3 同)。

**2.2.3 预报检验** 假设被预报某一样本也符合参与因子分析的那些样本所得出的统计规律。将预报样本各变量  $x_k$   $k=1, 2, \dots, 6$ , 代入下式:

$$z_i = \sum_{k=1}^6 v_{ik} x_k \quad i=1, 2, \dots, 6 \quad (7)$$

求出 6 个主因子  $z_1, z_2, z_3, z_4, z_5, z_6$ , 选出前三个主因子代入(6)式, 得到预报值  $C_f$ 。由于本次海上观测遇到阴天和雨天, 为此只选取了 10 个有阳光情况下的观测数据参与统计分析。采用有阳光的 10 个测站中的 9 个测站数据, 循环进行上述的统计分析, 求出估算模式, 将另一站作为预报对象, 轮流得到 10 个测站的预报值, 列入表 2。10 个样本预报的均方根误差为 0.1136, 平均相对误差为 17.79%, 其中 16 测站在预报中, 相对误差达 41.58%, 从图 2 中 b, c 图所见, 16 测站与其它 9 个站相比较属另一类型海水。用一种类型海水经统计分析所得估算模式, 用于估算另一类型海水的叶绿素

**表 2 叶绿素  $a$  浓度的预报值与实测值的比较<sup>1)</sup>**

Tab.2 Comparison between forecasting and measured values for chlorophyll  $a$  concentration in the South Yellow Sea

站位	1	2	3	4	5	8	16	17	18	19
$C_m$	0.140	0.170	0.160	0.140	0.130	0.170	0.840	0.300	0.160	0.180
$C_f$	0.188	0.154	0.132	0.124	0.140	0.121	0.491	0.317	0.181	0.194
$C_m - C_f$	-0.048	-0.016	0.028	0.016	-0.010	0.049	0.349	-0.017	-0.021	-0.014
$E(\%)$	34.37	9.62	17.68	11.38	7.86	28.77	41.58	5.79	12.96	7.89

1)  $C_m$  为实测值;  $C_f$  为预报值;  $C_m - C_f$  为残差(下同)。

表 3 叶绿素  $a$  浓度的预报值与实测值的比较Tab.3 Comparison between forecasting and measured values for chlorophyll  $a$  concentration in the South Yellow Sea

站 位	6	7	9	10	11	12	13	14	15
$C_m$	0.100	0.120	0.180	0.160	1.570	1.250	1.800	2.540	3.310
$C_f$	0.150	0.139	0.249	0.278	35.088	9.087	0.999	0.849	0.509
$C_m - C_f$	-0.050	-0.019	-0.069	-0.118	-33.518	-7.837	0.801	1.691	2.801
$E(\%)$	50.23	15.81	38.54	73.75	2134.92	626.94	44.50	66.57	84.64

浓度, 会带来相应的估算误差。

以参加因子分析的 10 个测站(1—5, 8, 16—19, 有阳光观测)所得的计算叶绿素  $a$  浓度的算式(6)为依据, 预报其余未参加因子分析的 9 个测站(多云、阴天、雨天)的叶绿素  $a$  浓度, 预报结果列于表 3。预报结果表明, 对无阳光天气(多云、阴天、雨天)条件下, 叶绿素  $a$  的预报效果较差, 特别是雨天, 预报相对误差达 2134.92%, 表明天气条件对预报值产生较大影响。

### 3 结语

在南黄海水域的生物光学算法中, 对多光谱反射比进行因子分析, 分析结果表明, 第一、二、三主因子对反射比方差的贡献率分别为 66.37%, 31.8% 和 0.97%, 前三个主因子已反映出全部信息量的 99.19%。对由前三个主因子估算叶绿素  $a$  浓度的回归方程进行回报和预报检验。回报检验叶绿素  $a$  浓度的均方根误差为  $0.0262\text{mg}/\text{m}^3$  (叶绿素  $a$  浓度范围:  $0.130\text{mg}/\text{m}^3 - 0.840\text{mg}/\text{m}^3$ ), 其平均相对误差为 10.06%。预报检验对于有阳光的测站, 均方根误差为  $0.1136\text{mg}/\text{m}^3$ , 其平均相对误差为 17.79%; 对于无阳光(多云、阴天、雨天)的测站, 叶绿素  $a$  预报效果较差, 表明天气条件对估算精度有较大影响。

### 参 考 文 献

- 傅克付、曾宪模, 1995, 黄渤海海洋, 13(1):55—63。  
 Gordon, H. R. et al., 1980, *Science*, 210:63—68。  
 Johnson, R. E., 1978, *Photogr. Engin. Rem. Sens.*, 44:617—624。  
 Morel, A. and Gordon, H. R., 1980, *Bound. Layer Meteo.*, 18:343—355。  
 Smith, R. C. and Wilson, W. H., 1981, *Oceanography From Space*, ed. by Gower, J. F. R., Plenum Press (New York), pp. 281—294。

# A STUDY ON MARINE BIO-OPTICAL ALGORITHM — ESTIMATION OF CHLOROPHYLL CONCENTRATION IN THE SOUTH YELLOW SEA

Fu Kecun, Zeng Xianmo, Li Baohua

(First Institute of Oceanography, SOA, Qingdao 266003)

Zeng Xiaoqi

(Department of Fisheries, Ocean University of Qingdao, Qingdao 266003)

Hisayuki Arakawa

(Tokyo University of Fisheries, Tokyo 108)

**Abstract** On the basis of the multi-spectral reflectance data in the South Yellow Sea and the chlorophyll *a* concentration data determined *in situ* with a fluorimeter during the “Sino-Japanese Joint Basic Research on visible light remote sensing in the Yellow Sea” in September, 1995, a marine bio-optical algorithm has been studied. Factor analysis has been undertaken for the multi-spectral reflectance data. The reflectance data covariant matrix is solved for eigenvalues and their corresponding eigenvectors; the principal factors are obtained. The contributions of the first three principal factors to the variance of reflectance are, respectively, 66.37%, 31.8% and 0.97%. From the first three principal factors 99.19% of the total information on the original variables can be retrieved. A multivariate linear regression equation has been established between the natural logarithm of the chlorophyll *a* concentration and the first three principal factors. The result shows that the overall correlation coefficient of the subsample reaches 0.9713 and the root-mean-square error of the hindcasting test estimate of chlorophyll *a* concentration is 0.0262 mg / m<sup>3</sup> (the range of the chlorophyll *a*: 0.130 — 0.840 mg / m<sup>3</sup>), with a relative error of 10.06%. The relative error for the forecasting test is 17.79%. Predicted values of chlorophyll *a* concentration are greatly affected in cloudy and rainy days.

**Key words** Factor analysis      Multi-spectral reflectance      Chlorophyll *a* concentration