



## 基于 TM 数据渤海湾叶绿素浓度反演算法研究

郑小慎, 林培根

(天津市海洋资源与化学重点实验室, 天津科技大学海洋科学与工程学院, 天津 300457)

**摘要:** 海洋叶绿素浓度已成为衡量浮游植物的生物量和富营养化程度的最基本的指标. 利用 2005 年 4 月 13 日 Landsat-5/TM 数据和同步实测数据反演了渤海湾叶绿素浓度. 对 TM 数据进行几何校正和大气校正预处理后, 建立 9 种波段组合模型, 分别对 TM 图像进行模型运算. 结合实测数据, 建立相应的回归方程, 在其中选取  $R^2$  (拟合度) 和  $R$  (相关性) 最大的  $(tm_4 - tm_3)/(tm_4 + tm_3)$  波段组合进行反演, 并将反演结果与实测数据作相对误差分析, 实验结果表明, 4 和 3 波段组合建立的回归模型适合于渤海湾叶绿素浓度的反演.

**关键词:** 遥感; TM 数据; 叶绿素浓度

中图分类号: TP751 文献标志码: A 文章编号: 1672-6510(2010)06-0051-03

## Study on Algorithm of Chlorophyll Concentration Retrieval in Bohai Bay Based on TM Data

ZHENG Xiao-shen, LIN Pei-gen

(Tianjin Key Laboratory of Marine Resource and Chemistry, College of Marine Science and Engineering, Tianjin University of Science & Technology, Tianjin 300457, China)

**Abstract:** Ocean chlorophyll concentration has become the most basic indicators of measure of phytoplankton biomass and eutrophication. This paper established the inverse model of the Bohai bay chlorophyll concentration using the remote sensing data of Landsat-5/TM and the measured data on April 13, 2005. Firstly the TM data were preprocessed by geometric correction and atmospheric correction, the 9 kinds combined band models of TM image were built up to calculate respectively, then using the measured data and the result of model calculation, the regression equations was created, the band combination of  $(tm_4 - tm_3)/(tm_4 + tm_3)$  was selected, which  $R^2$  (fitting) and  $R$  (relevance) is the largest. Last the chlorophyll concentration was retrieved using the above band combination. The relative error of the inversion results and measured data shows that 4 and 3-band combination is appropriate to establish the regression model for the inversion of the Bohai bay chlorophyll concentrations.

**Keywords:** remote sensing; TM data; chlorophyll concentration

海洋叶绿素是浮游植物现存量的表征, 叶绿素浓度直接反映了浮游植物数量, 即海域初级生产者的现存生物量<sup>[1-2]</sup>. 叶绿素浓度的常规测定需要先采集水样, 对水样进行一系列处理, 在实验室对水样进行分析测定, 获得水样的叶绿素浓度. 该法速度慢、费用高, 难以实现大范围海域的同步采样测量. 随着 SeaSTAR 等海洋水色卫星的发展, 测定叶绿素浓度已经有多种成熟的遥感反演模型算法<sup>[3-4]</sup>. Landsat 陆地

卫星虽然没有海洋水色卫星在研究大范围海洋叶绿素浓度方面的优势, 但由于其高分辨率, 在近岸小范围研究中仍可以作为水色卫星数据的有效补充<sup>[5-7]</sup>.

### 1 研究区域及遥感数据预处理

#### 1.1 研究区域

近年来环渤海地区经济活动给渤海水域生态系

收稿日期: 2010-06-16; 修回日期: 2010-09-25

基金项目: 天津市自然科学基金资助项目(08JCYBJC10500); 天津市重点自然科学基金资助项目(09JCZDJC25400)

作者简介: 郑小慎(1973—), 女, 河北献县人, 副教授, zxs@tust.edu.cn.

统带来了重大的影响,加之其半封闭的地理环境,造成了渤海海水循环更新能力低于我国其他海域,渤海海洋环境问题日益突出<sup>[2]</sup>. 渤海湾位于渤海西部,湾口以大清河口至山东的旧黄河口一线为界. 它是一个向西凹、呈弧状的浅水湾,海底地势从湾顶向渤海中央倾斜,湾内水很浅,一般均在 20 m 以内.

### 1.2 遥感及实测数据

卫星影像数据为2005年4月13日 Landsat-5/TM 获取的遥感图像,包含 1~7 波段<sup>[8]</sup>. 由于一景 TM 图像范围在 37.96~39.82° N, 116.82~118.75° E 之间,只取到渤海湾部分,故本研究以渤海湾为例. 利用 ERDAS 软件对 TM 遥感图像进行几何精纠正,总误差控制在 0.5 个像素点之内,图像重采样采用最近邻近点法,以避免光谱信息的丢失,投影选择 UTM 投影,椭球体选择 WGS84,这样可以使采样点定位坐标和遥感图像投影坐标精确匹配. 实测数据为在 2005年4月11日至14日测量得到的渤海 24 个站点叶绿素质量浓度数据. 这4天中天气情况及海面情况比较稳定,没有出现异常,因此可以用这4天的实测数据与 Landsat5/TM 卫星遥感数据配合进行实验研究.

### 1.3 采样数据有效点的选择

对于实测数据,采用格拉布斯(GRUBBS)检验方法进行异常值的检验. 设  $n$  个测试数据为  $x_1, x_2, \dots, x_n$ , 检验步骤如下:首先计算平均值  $\bar{x}$  和标准差  $s$ ,然后将  $n$  个数据按从小到大的顺序排列成  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ . 接着计算统计量  $G_n$ .

$$G_n = \max \left( \frac{x_{(n)} - \bar{x}}{s}, \frac{\bar{x} - x_{(1)}}{s} \right)$$

最后,查显著性水平  $\alpha = 0.05$  和  $\alpha = 0.01$  时与  $n$  对应的格拉布斯检验法临界值  $G_{(0.05)}$  和  $G_{(0.01)}$ , 并进行判断.

按照格拉布斯检验方法的检验步骤进行,与 GB 4883—1985《数据的统计处理和解释·正态样本异

常值的判断和处理》中格拉布斯检验法临界值表作比较,  $G_n \leq G_{(0.05)}$ , 说明其中24个数据可使用,在 24 个实测数据中抽取14个用来建立非线性回归模型,另外 10 个作为检验数据.

## 2 基于 TM 数据非线性回归模型建立与分析

### 2.1 在 ERDAS 中建立波段组合模型

根据 TM 各波段的特点和各个研究中使用情况选择波段组合,由于  $tm_1, tm_2, tm_3, tm_4$  这 4 个波段对叶绿素差异较敏感,所以本次研究选择以下几个波段组合模型做比较:  $(tm_4 - tm_3)/(tm_4 + tm_3)$ 、 $(tm_2 - tm_4)/(tm_2 + tm_4)$ 、 $(tm_1 - tm_4)/(tm_1 + tm_4)$ 、 $tm_4/tm_3$ 、 $tm_3/(tm_1 + tm_4)$ 、 $(tm_1 + tm_4)/(tm_2 + tm_3)$ 、 $(tm_2 + tm_3)/(tm_1 + tm_4)$ 、 $(tm_2 - tm_3)/(tm_2 + tm_3)$ 、 $(tm_1 - tm_3)/(tm_1 + tm_3)$ .

利用 ERDAS 遥感软件的图像处理中的条件判断功能,对影像各运算波段中的值采取逻辑判断处理,即对波段运算公式中分母值为零的情况特殊处理,以避免运算结果产生溢出的情况.

### 2.3 建立回归方程

#### 2.3.1 记录 ERDAS 处理后的对应经纬度点数据

把选出建立方程的 14 个数据经纬度逐个输入,分别记录每个经纬度得到的“FILE PIXEL”值,即该经纬度点的  $(tm_4 - tm_3)/(tm_4 + tm_3)$  值. 在图形模块中选择建立散点图,在弹出对话框中选择叶绿素质量浓度的自然对数  $\ln \rho$  为因变量,依次选择各个波段组合为变量,生成散点图. 并进行线性估计,建立非线性回归方程.

9 个波段组合模型生成了二次回归方程和参数估计表,把参数估计值转写成方程式的形式,总结成表,以便进行对比,结果见表 1. 选取  $R^2$ (拟合度)或者  $R$ (相关性)较大的一组回归方程进行反演.

表 1 波段组合回归方程

Tab.1 Band combinations regression equation

波段组合	替代编号	回归方程	$R^2$	$R$
$(tm_4 - tm_3)/(tm_4 + tm_3)$	$a$	$\ln \rho = 3.948 + 11.621a + 20.993a^2$	0.373	0.610
$(tm_2 - tm_4)/(tm_2 + tm_4)$	$b$	$\ln \rho = 2.645 - 3.096b + 9.227b^2$	0.214	0.463
$(tm_1 - tm_4)/(tm_1 + tm_4)$	$c$	$\ln \rho = 6.001 - 15.099c + 15.733c^2$	0.210	0.428
$tm_4/tm_3$	$d$	$\ln \rho = 4.027 - 4.990d + 3.905d^2$	0.073	0.270
$tm_3/(tm_1 + tm_4)$	$e$	$\ln \rho = -16.403 + 85.771e - 97.012e^2$	0.038	0.195
$(tm_1 + tm_4)/(tm_2 + tm_3)$	$f$	$\ln \rho = 2.690 - 0.842f + 0.142f^2$	0.030	0.173
$(tm_2 + tm_3)/(tm_1 + tm_4)$	$g$	$\ln \rho = -0.859 + 7.452g - 4.043g^2$	0.032	0.179
$(tm_2 - tm_3)/(tm_2 + tm_3)$	$h$	$\ln \rho = 3.693 + 21.662h + 94.126h^2$	0.104	0.323
$(tm_1 - tm_3)/(tm_1 + tm_3)$	$i$	$\ln \rho = 5.200 - 21.162i + 41.421i^2$	0.013	0.114

### 3 回归方程的检验与分析

#### 3.1 检验回归方程

从表1中可以看出,拟合度  $R^2$  或者相关性  $R$  最大的一组回归方程是关于  $(tm_4 - tm_3)/(tm_4 + tm_3)$

波段组合的回归方程. 所以,选择这个回归方程进行反演检验.

使用之前选出作为检验数据的10个实测数据,根据二次回归方程  $\ln \rho = 3.948 + 11.621a + 20.993a^2$  反演出该10个检验点的  $\ln \rho$  值(叶绿素质量浓度的自然对数值),记录在表2中.

表2 叶绿素浓度反演和误差分析

Tab.2 Chlorophyll concentration inversion and error analysis

$\rho_{\text{实测}}/(\text{mg}\cdot\text{m}^{-3})$	$\ln \rho_{\text{实测}}$	FILE PIXEL 值	$\ln \rho_{\text{反演}}$	$\rho_{\text{反演}}/(\text{mg}\cdot\text{m}^{-3})$	相对误差/%
12.25	2.505 5	-0.143	2.715 5	15.111 9	23.36
11.37	2.431 0	-0.271	2.340 5	10.386 0	8.65
21.32	3.059 6	-0.127	2.810 7	16.622 0	22.04
15.62	2.748 6	-0.224	2.398 2	11.003 8	29.55
15.65	2.750 5	-0.200	2.463 5	11.746 1	24.95
14.05	2.642 6	-0.413	2.729 3	15.321 9	9.05
9.03	2.200 6	-0.253	2.351 6	10.502 7	16.31
9.09	2.207 2	-0.262	2.344 3	10.426 4	14.70
11.51	2.443 2	-0.204	2.451 0	11.599 5	0.78
2.98	1.091 9	-0.351	2.455 4	11.650 9	290.97

#### 3.2 叶绿素浓度的反演

利用 ERDAS IMAGINE 中的 Modeler 模块,以 ndvi.img 为输入图像,根据得到的回归方程建立简单的数学计算模型:

$$\$n1\_ndvi * \$n1\_ndvi * 20.993 + \$n1\_ndvi * 11.621 + 3.948$$

得到灰度值的变化就是所得的  $\ln \rho$  变化,即代表叶绿素浓度变化. 运用上述回归方程和本次试验的 TM 图像进行叶绿素反演,得到图1.

合,建立非线性回归模型,并对叶绿素浓度进行反演. 得到的二次回归模型,平均相对误差为 18.5%,形式较简明,精度较高,回归模型适合于渤海湾叶绿素浓度的反演.

#### 参考文献:

- [1] 李健,舒晓波,陈水森. 基于 Landsat-TM 数据鄱阳湖湿地植被生物量遥感监测模型的建立[J]. 广州大学学报:自然科学版,2005,4(6):494-498.
- [2] 丛丕福. 海洋叶绿素遥感反演及海洋初级生产力估算研究[D]. 北京:中国科学院研究生院,2006.
- [3] Coskun H G, Tanik A, Alganci U, et al. Determination of environmental quality of a drinking water reservoir by remote sensing, GIS and regression analysis[J]. Water Air & Soil Pollution, 2008, 194(1/2/3/4):275-285.
- [4] Duan H T, Zhang Y Z, Zhang B, et al. Assessment of chlorophyll-a concentration and trophic state for Lake Chagan using landsat TM and field spectral data[J]. Environmental Monitoring and Assessment, 2007, 129(1/2/3):295-308.
- [5] 吕恒,江南,罗淑葱. 基于 TM 数据的太湖叶绿素 A 浓度定量反演[J]. 地理科学, 2006, 26(4):472-476.
- [6] 李云亮,张运林. 基于 TM 影像的太湖夏季悬浮物和叶绿素 a 浓度反演[J]. 遥感应用, 2008(6):22-27.
- [7] 杨一鹏,王桥,肖青,等. 基于 TM 数据的太湖叶绿素 a 浓度定量遥感反演方法研究[J]. 地理与地理信息科学, 2006, 22(2):5-8.
- [8] 梅安新,彭望球,秦其明,等. 遥感导论[M]. 北京:高等教育出版社,2005:156.

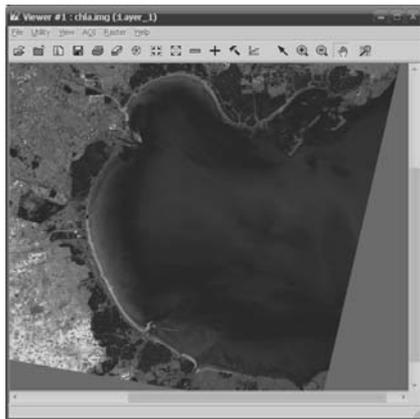


图1 叶绿素浓度分布灰度值图

Fig.1 Chlorophyll concentration gray value map

### 4 结论

以 TM 遥感数据和准同步的叶绿素浓度实测数据为信息源,基于 TM 遥感数据对 9 种波段组合方法进行比较,选择出相关性最高 ( $R^2$  为 0.373,  $R$  为 0.610) 的波段组合,即  $(tm_4 - tm_3)/(tm_4 + tm_3)$  波段组