

基于分段映射模型的水质参数遥感反演研究

陈 军^{1,2}, 温珍河^{1,2}, 付 军^{1,2}

1. 国土资源部海洋油气资源与环境地质重点实验室, 山东 青岛 266071

2. 青岛海洋地质研究所, 山东 青岛 266071

摘 要 以数据为驱动的遥感建模方法片面地强调模型模拟值与实测值总体偏差最小, 而忽视局部偏差状况, 进而可能导致模型局部模拟值与实测值偏差过大。针对这种建模方法的缺陷, 以水质参数反演为例, 提出了分段映射反演算法。该算法将水质参数浓度与遥感参数之间的映射关系分解为若干分段函数, 每一区段参数间的关系由该区段内的实测数据(简称为节点) 和插值函数决定。通过分析以 Newton 插值算法作为插值函数的分段映射反演模型可知, 该算法能保证每个节点处的模型计算值与实验值一致, 并且能较好地拟合分布趋势复杂的实验数据。此外, 分段映射反演模型对野外水质样本采集实验的规范化具有较大理论指导意义。最后利用太湖 TM 影像数据和同步实测水质数据为例, 论证了该算法的可行性。

关键词 遥感; 分段映射反演算法; Newton 插值算法; 水质

中图分类号: TP732.2 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2010)10-2784-05

引 言

随着社会经济的发展, 水资源短缺问题日益突出, 传统的水质管理与监测技术则越来越难以应付这种矛盾的挑战^[1]。遥感数据、实测数据和水质反演模型综合体的出现有助于缓和这种矛盾^[2]。通常, 水质参数定量模型可分为^[3]: 分析模型、半分析模型与经验模型。虽然水色遥感取得了长足的发展, 但是业务化反演算法的精度仍然不高, 主要原因可归结为: (1) 水体成分的复杂性导致反演机理难以把握^[4]; (2) 水体信号太弱导致水体信息提取困难^[5-7]; (3) 定量模型的不统一性及模型偏差的验证难^[8,9]。

通常情况下, 定量模型所涉及的经验参数是通过优化算法获取, 且定量模型至少为一阶光滑曲线。然而, 由优化算法解得的最佳代价函数不能保证模型的计算值与实测值的精确相等, 甚至在有些函数区间段内存在过大偏差, 进而导致模型在偏差较大的函数区间段内反演效果较差。本文基于上述问题的考虑, 提出了分段映射反演算法。该算法在保证节点处模型值与实测值精确相等的前提下, 将水质浓度参数与光学参数之间的映射关系分解为若干分段函数, 每一区段特征量之间的关系由实测数据(简称为节点) 和采用的插值函数决定。此外, 本文以 Newton 插值算法为例, 结合 2003 年 10 月 27 日的太湖水质参数实测数据和 TM 影像数据, 综合反

演太湖悬浮泥沙和叶绿素 a 浓度的分布状况, 论证了该算法的可行性。

1 研究区域与数据

根据水色光学测量和数据处理有关规范, 在 2003 年 10 月 27 日, 利用 ASD 野外光谱仪 FieldSpec Pro 对图 1 所示的站点分别进行了水体光谱的测量。站点的光谱曲线见

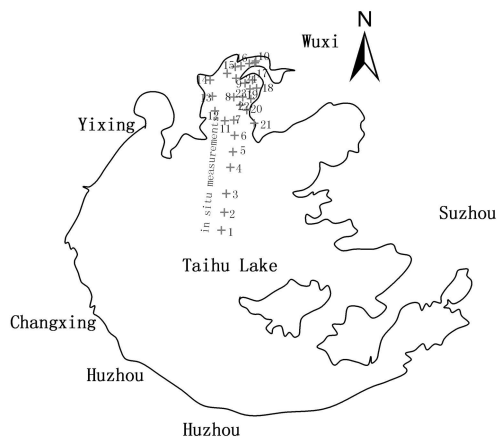


Fig 1 Distribution of the *in situ* measurements in Taihu Lake

收稿日期: 2009 11-30, 修订日期: 2010 03-06

基金项目: 国家“十一五”科技支撑项目(2008BAC34B03)和国家地质大调查专项(1212010511302)资助

作者简介: 陈 军, 1982 年生, 青岛海洋地质研究所实习研究员, e-mail: chenjun820711@163.com

图 2。在表观光学量观测的同时,使用尼克森采水器取表层大约从水面到水下 30 cm 处的水体样本,取样后立即放入有冰块的冷藏箱,当天带回实验室处理。根据湖泊测量操作流程^[19],测量了叶绿素 a 浓度、悬浮泥沙浓度和黄色物质浓度。

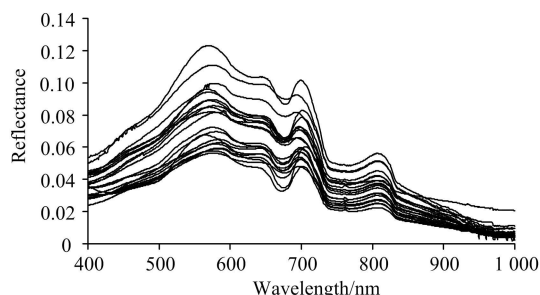


Fig 2 In situ measurements of spectral curves

2 研究方法

目前,用于描述节点与节点之间函数分布趋势的插值算法很多,主要包括:Newton 插值算法、Gauss 插值算法和样条插值算法等。本文以 Newton 插值算法为例,介绍该插值算法在构建反演模型中的应用,旨在阐述分段映射反演算法的基本思想。

2.1 算法描述

设一组实测数据如下

$$R = \{R_1, R_2, \dots, R_n\} \quad (1)$$

$$X = \{X_1, X_2, \dots, X_n\} \quad (2)$$

式中, R_i 和 X_i 表示第 i 个观测站点光学参数与水质参数的实测值。为了公式推导的方便,假设 R_i 的下标 i 表示 R_i 值从小到大排列的序号。

由 Newton 插值原理可得两个节点的插值公式如下

$$f(R) = f(R_i) + f[R_i, R_{i+1}](R - R_i) + f[R, R_i, R_{i+1}](R - R_i)(R - R_{i+1}) \quad (3)$$

$$f[R, R_i, R_{i+1}] = \frac{f[R, R_i] - f[R_i, R_{i+1}]}{R - R_{i+1}} \quad (4)$$

$$f[R_i, R_{i+1}] = \frac{f(R_i) - f(R_{i+1})}{R_i - R_{i+1}} \quad (5)$$

$$f[R, R_i] = \frac{f(R) - f(R_i)}{R - R_i} \quad (6)$$

式中, $f(\cdot)$ 为 $R_i \rightarrow X_i$ 的映射函数, $R \in [R_i, R_{i+1}]$ 。

变换(3)式可得

$$f(R) = N(R) + \xi(R) \quad (7)$$

其中,

$$N(R) = f(R_i) + f[R_i, R_{i+1}](R - R_i) \quad (8)$$

$$\xi(R) = f[R, R_i, R_{i+1}](R - R_i)(R - R_{i+1}) \quad (9)$$

式中, $N(\cdot)$ 为 Newton 插值多项式; $\xi(R)$ 为插值误差。

设 $\Delta R = R_{i+1} - R_i$, 当 $\Delta R \rightarrow 0$ 时, 则

$$|f(R)| \leq |f(R_i)| + |f[R, R_i]| \Delta R + |O(\Delta R)| \quad (10)$$

$$\lim_{\Delta R \rightarrow 0} f(R) = f(R_i) \quad (11)$$

式中, $||$ 为取绝对值算子; $O(\Delta R)$ 表示 ΔR 的高阶无穷小。该式说明, 当观测样本数量无限大, 且节点在定义域上的分

布够密集时, (8) 式就是水质参数与光学参数(如反射率)之间的精确关系。该性质对于水色遥感实验中的水质样本采集规范化具有较大的理论指导意义, 体现在以下几个方面:

(1) 合理的水质样本应分布在水质参数与光学参数关系突变的特殊点附近。这里的特殊点主要包括: 曲线的极值点、拐点和斜率变化较明显的曲线段上的点等。采用上述特殊点作为节点能有效地控制 Newton 算法的插值误差。

(2) 样本的光学参数在一维空间上的展布间隔由插值精度决定。由(9)式可知, Newton 插值算法的误差与节点间隔成正比, 因此, 控制节点间隔将有助于模型误差的改善。这要求实验的有效站点数和光学参数在一维空间上的展布间隔应与插值精度相匹配。

(3) 在不考虑影像数据和环境因素影响的背景下, 节点间的插值精度应控制在一定的精度范围之内。

2.2 分段映射反演模型

最小二乘法是一种常用的数据分析方法, 该方法有助于从整体上把握现象变化的客观规律, 但是难以保证局部的计算精度, 以致现象局部的变换规律被弱化甚至淹没; 此外, 并不是所有的物理量与物理量之间的关系都可以用常规的数学函数(如指数函数、对数函数, 多项式等)表述。另外, 利用遥感模型与算法进行反演的一个基本假设为实测数据的精度是可靠的, 然而由最小二乘法等求得的最佳代价函数不能保证实测光学参数与水质组分浓度之间的关系完全精确成立, 而只能在整体上满足误差最小, 这显然不利于反演精度的提高。模型只有在整体和局部均与实测数据相符合才更具有应用价值, 这就是分段映射反演算法的核心思想。由本文提出的分段映射反演算法步骤如下:

(1) 函数定义域的分段。将实际测量而得的物理量(如光学参数)在定义域上从小到大排序, 相邻的两物理量构成一个分段区块。为了所构建的模型保证定义域和值域之间是一对一的关系, 当实测结果是一对多关系时, 不妨将这种一对多关系视为同一对象的多次观测, 并取平均值作为映射结果。

(2) 大气校正。大气校正是定量遥感的一个重要的环节, 当有地面同步观测数据时, 可以根据地表实测光学参数与遥感影像对应点的光学参数建立分段映射关系, 嵌套一次分段映射反演算法进行大气校正, 否则采用其他大气校正方法。

(3) 定量参数计算。(8) 式描述了分段中两物理量之间的映射关系, 对于一个给定的像元, 先检查物理量所属的分段区域, 然后根据该区段的两端点实测数据计算(8)式中的各个参数, 最后根据这些参数计算出该像元的反演结果。

3 数据分析

3.1 基于分段映射的大气校正方法

本次太湖实验总共采集了 25 个站点的实测数据。以往研究表明^[11, 12]: TM2/TM3 算法较适合于叶绿素 a 浓度的遥感反演; TM2 波段较适合于悬浮物浓度的反演。因此, 本研究将以 TM2/TM3 和 TM2 作为遥感参数, 以实测的 25 个站点的实测数据作为插值节点, 分别建立分段映射大气校正模

型,如图 3 与图 4 所示。其中,图 3(a)和图 4(a)是利用最小二乘法拟合得到的优化模型,图 3(b)和图 4(b)为基于 Newton 插值算法的分段映射大气校正模型。图 3 和图 4 表明:从整体趋势上看,因变量随着自变量的增大而增大;从局部来看,因变量随着自变量增大而增大的现象不是恒成立的,而是存在一定的波动(图 3 与图 4)。这将导致图 3(a)和图 4(a)两种优化模型的相关系数不高,分别为 0.710 6 与 0.537 4。这样的拟合结果显然不太理想,且在参与优化计算的实测数据的节点处,计算值与实测值不能精确成立。与图 3(a)和图

4(a)相比较,图 3(b)和图 4(b)将反演模型的定义域以两个节点为区间进行分段,实测遥感参数与 TM 影像观测值在分段区间上的变化用插值算法加以定义,这种做法至少保证了节点处的模型计算值是精确。对于以数据为驱动的反演模型,模型计算结果与实测数据的接近程度是衡量模型优劣的尺度。分段映射大气校正算法对于每个实测数据是精确的,因此,在该衡量尺度下,分段映射校正算法是具有更高模型精度的。

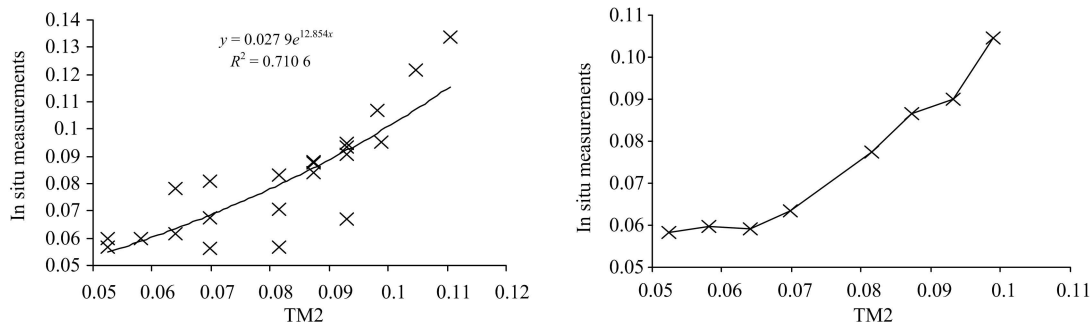


Fig 3 Atmospheric correction for TM2

(a): Exponential atmospheric correction for TM2; (b): Subsection mapping algorithm for TM2 atmospheric correction

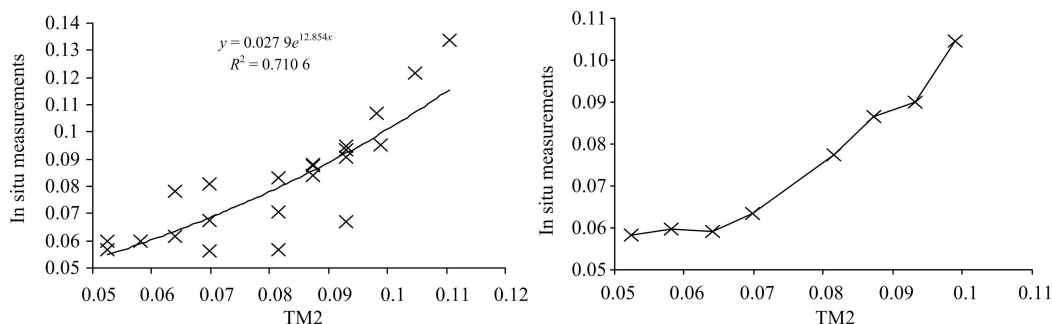


Fig 4 Atmospheric correction for TM2/TM3

(a): Exponential atmospheric correction for TM2/TM3; (b): Subsection mapping algorithm for TM2/TM3 atmospheric correction

3.2 基于分段映射算法反演太湖悬浮泥沙浓度与叶绿素 a 浓度

在 IDL 开发环境下,提取影像中每一像元值,并确定像元值归属的分段区间。若像元值落在实测物理量最大值与最小值范围之内,则选择离该像元值归属的分段子区间的两个节点,利用 Newton 插值算法进行内插得到函数值;若像元值落在实测物理量最大值与最小值范围之外,则选择离该像元值最近的区段,利用 Newton 插值算法构建该区段的反演模型,然后进行外推得到反演值。图 5(a)和图 5(b)分别为悬浮泥沙浓度与叶绿素 a 浓度分段映射反演模型。经过图 3(b)和图 4(b)所示模型的大气校正和图 5(a)和图 5(b)所示模型的定量反演可得太湖区域 2003 年 10 月 27 日泥沙浓度和叶绿素 a 浓度的分布状况如图 6 所示。

从图 6 可以看出:在太湖区域,叶绿素浓度呈湖心,湖西与湖东高,而湖南与湖北低的格局,部分边缘区域叶绿素浓度很高(太湖沿岸健康植被对边缘像元叶绿素浓度的贡献,即所谓边缘效应),平均叶绿素 a 浓度为 $43.57 \mu\text{g} \cdot \text{L}^{-1}$,已

达富含藻类,并存在水华类水质的指标^[13];悬浮泥沙浓度分布为湖心和湖南高,其他区域低的特点。

综合分析图 6(a)与图 6(b)可知:太湖水体叶绿素浓度较高区域恰好是悬浮物浓度较低的区域,而悬浮浓度较低区域则恰好是叶绿素浓度较高的区域。其原因可解释为:悬浮物浓度与水动力强度密切相关^[14],当水动力强度过强时,水体中的营养成分难以被浮游植物吸收利用,进而不利于浮游植物的生长;与之相反,在水动力条件较弱的区域,上游被水体携带下来的有机质在该区域逐渐沉淀,为浮游植物的生长提供了物质条件。由此可知,太湖水体的叶绿素浓度与悬浮浓度的空间分布呈负相关的特性。

4 结 论

分段映射反演模型以实测物理量作为分段节点,将物理量的定义域分割为若干子区间,然后对各个子区间进行拟合与插值运算。在很多情况下,随观测数据点数增加,传统的

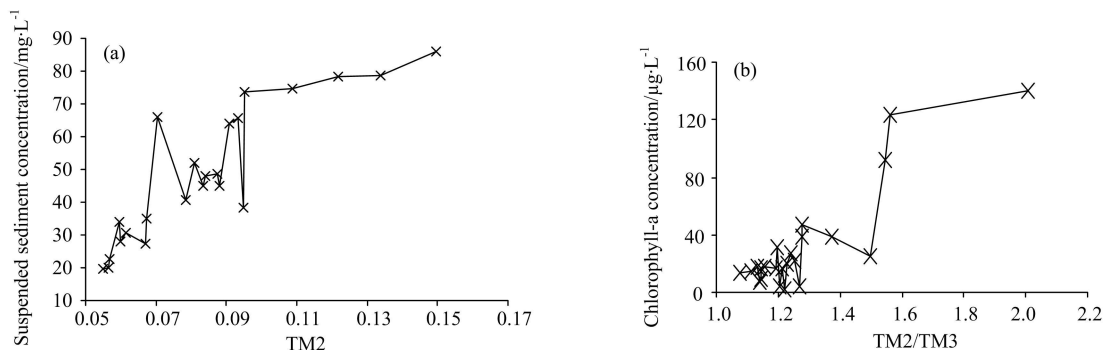


Fig 5 Subsection mapping algorithm for water qualities inversion

(a): Subsection mapping algorithm for suspended sediment concentration retrieval;

(b): Subsection mapping algorithm for chlorophyll-a concentration retrieval

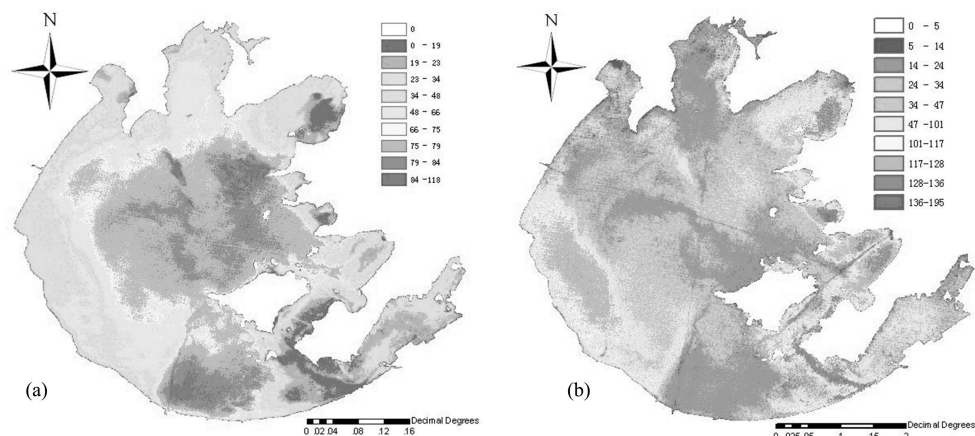


Fig 6 Inversion results of subsection algorithm

(a): Subsection algorithm retrieval results for suspended sediment concentration;

(b): Subsection algorithm retrieval results for chlorophyll a concentration

经验算法, 如基于最小二乘法的反演模型, 其回归系数或均方根误差将急剧增加, 甚至需要重新尝试新的数学表达式进行拟合。这是因为水质参数与光学特性之间的映射关系难以用一个简单的数学函数很好地描述, 拟合结果只能是粗略地表达样本数据的一个分布趋势, 而不能照顾到细节。与传统算法相比较, 分段映射反演模型不因实测数据量的增加或改变而调整反演函数的形式, 并且能够保证分段节点处的反演值与实测值的吻合; 此外, 随着实测数据量的增加, 分段映射反演模型是对水质参数与光学特性之间映射关系的一个无限逼近, 当样本数趋向于无穷大时, 分段映射模型就是水质

参数与光学特性之间的映射关系。因此, 与传统经验算法相比较, 分段映射模型具有较真实描述水质参数与光学特性之间关系。

分段映射算法对水色遥感实验中的水质样本采集规范化具有较大的理论指导意义。合理的水质样本应分布在反映水质参数与光学参数关系突变的特殊点附近, 并且样本的水质参数或光学参数在一维空间上的展布间隔由插值精度决定。此外, 在不考虑影像数据和环境因素影响的背景下, 节点间的插值精度应控制在一定得國值之内。

参 考 文 献

- [1] Dekker A G, Vos R J, Peters S W. International Journal of Remote Sensing, 2002, 23(1): 15.
- [2] Curran P J, Novo E M. Journal of Coastal Research, 1988, 4: 351.
- [3] Gohin F, Druon J N, Lampert L. International Journal of Remote Sensing, 2002, 23(8): 1639.
- [4] ZHANG Ting-lu, SHI Ying-ni (张亭禄, 施英妮). Periodical of Ocean University of China(中国海洋大学学报), 2005, 35(5): 849.
- [5] Michael D K, Yoram J K, Didier T, et al. Bulletin of American Meteorological Society, 1999, 80(11): 2229.
- [6] Zhao W J, Masayuki T, Hidenori T. Remote Sensing of Environment, 2000, 76: 202.
- [7] Shi W, Wang M H. Remote Sensing of Environment, 2007, 110: 149.
- [8] Lira J, Morales A, Zamora F. Internal Journal of Remote Sensing, 1997, 18(1): 171.

- [9] Neillis M D, Harrington J A, Wu J P. Geomorphology, 1998, 21: 281.
- [10] Mueller J L, Austin R W. Ocean optics protocols for SeaWiFS validation; NASA Technical Memorandum 104566. Greenbelt, MD, NASA Goddard Space Flight Center, 1992. 5.
- [11] Gordon H R, Brown O B, Evans R H, et al. Journal of Geophysical Research, 1988, 93: 10909.
- [12] Dekker A G, Vos R J, Peters S W M. International Journal of Remote Sensing, 2005, 23(1): 15.
- [13] HE Junhua, CHENG Yongjin, ZHANG Hao(贺俊华, 程永进, 张昊). Optics and Optoelectronic Technology(光学与光电技术), 2007, 5(5): 16.
- [14] CHEN Jun, WEN Zhenhe, FENG Jiajia, et al(陈军, 温珍河, 丰佳佳, 等). Marine Geology & Quaternary Geology(海洋地质与第四纪地质), 2008, 28(2): 31.

The Application of the Subsection Mapping Retrieval Model to Water Qualities Quantitative Analysis

CHEN Jun^{1,2}, WEN Zhenhe^{1,2}, FU Jun^{1,2}

1. The Key Laboratory of Marine Hydrocarbon Resource and Geology, Qingdao 266071, China

2. Qingdao Marine Geosciences Institute, Qingdao 266071, China

Abstract Taking the *in situ* measurements as the driver, the traditional remote sensing models modeling approaches emphasize particularly on the least total errors between the modeled estimate and the measurements while ignore its local error status, which may lead to a large warp between the modeled prediction and the observation at some position. Considering such default of general approach, the present paper developed the subsection mapping retrieval algorithm (SMRA), which decomposes the mapping mechanism between the water qualities and its optical parameters into several subsection functions, and each subsection function is determined by the *in situ* measurements (named nodes as follows) and an interpolating function. The analysis results of subsection mapping retrieval algorithm based on Newton interposing algorithms indicate that the algorithm keeps the inversion results accuracy at nodes, and is preferably suitable for regression estimate of the complicated relationship between the parameters. Additionally, the method has a great theoretical meaning for the standardization of the sampling interval and sample number in water qualities experiments. Combined with the analysis results of the Landsat/TM imagery and the experimental data of Taihu Lake, it could be known that the SMRA is preferably suitable for describing the relationship between remotely sensed parameters and water qualities, especially for complicated case II water bodies such as Taihu Lake.

Keywords Remote sensing; Subsection mapping retrieval algorithm; Newton interposing algorithm; Water qualities

(Received Nov. 30, 2009; accepted Mar. 6, 2010)