

# 海洋悬浮泥沙二元特征参数 MODIS 遥感反演模型研究

王 芳<sup>1,2</sup>, 李国胜<sup>2,\*</sup>

(1 加拿大卡尔加里大学, 卡尔加里 加拿大 T2N 1N4;

2 中国科学院地理科学与资源研究所, 北京 100101)

**摘要:** 本文提出了一种采用海面离水辐射率和泥沙粒径二元特征参数来反演研究海区(渤海)海洋表层悬浮泥沙浓度的新的遥感反演算法, 以此为基础分别建立了基于 MODIS 遥感数据和泥沙粒径二元特征参数的主成分和神经网络两种泥沙浓度反演模型, 并对比分析了两类模型的反演精度以及泥沙粒径因子对模型的影响。分析结果表明, 新建立的二元特征参数反演算法在采用主成分模型和神经网络模型时的检验误差分别为 0.256 和 0.244, 而忽略泥沙粒径因子贡献的主成分模型和神经网络模型的检验误差分别为 0.384 和 0.390, 因此可以认为, 在泥沙浓度反演模型中加入粒径因子时, 模型的预测精度和模型稳定性均比只考虑浓度对反射率贡献的模型有显著改善。

**关 键 词:** 渤海; 悬浮泥沙浓度; MODIS; 遥感反演; 二元模型

**文章编号:** 1000-0585(2007)06-1186-11

## 1 引言

海洋悬浮泥沙浓度是海洋环境研究的重要组成部分。悬浮泥沙浓度在近岸海洋区域的分布格局不仅是海洋环境变化的重要原因, 同时也是港口选址、沿岸水利、近岸工程建设等所要考虑的一个重要因素。海洋悬浮泥沙浓度的测量过去一直采用水样调查法, 不但耗费大量的人力、物力、财力, 且获得的资料呈点状分布, 在时间和空间上都不连续, 无法建立数据的面源和长时间序列分布。遥感方法已被证实可有效地应用于海洋悬浮泥沙浓度的反演中, 它可以获取大面积、实时、同步的观测数据。对此, 前人已经进行了大量研究<sup>[1~7]</sup>。同时, 在过去近 20 年中, 国内学者对我国近岸河口地区的海洋悬浮泥沙分布开展了多项遥感反演模式的研究工作, 并取得了丰硕的研究成果<sup>[8~15]</sup>。

然而, 截止目前, 绝大部分已开展的海洋悬浮泥沙遥感反演研究, 都是基于由遥感影像获取的光谱值与泥沙浓度之间的直接或间接关系。实际上, 在二类水体区域, 传感器接收到的来自水体部分的光谱反射能量是水体中所有物质反射的综合结果, 其中包括悬浮泥沙、叶绿素和黄色物质等, 而不仅仅是悬浮泥沙浓度的函数。即便在泥沙浓度相同的情况下, 也会由于泥沙颗粒的大小、形状、矿物组成等不同而产生不同的光谱反射特征。此

收稿日期: 2007-03-23; 修订日期: 2007-08-14

基金项目: 国家自然科学基金资助项目 (40771030、40571020); 中国科学院知识创新工程领域前沿项目。

作者简介: 王芳 (1979-), 女, 江苏灌云人, 加拿大卡尔加里大学研究生。主要从事环境遥感研究。

\* 通讯作者: 李国胜 (1963-), 男, 江苏常州人, 研究员。主要从事陆海相互作用过程遥感与 GIS 模拟研究。

E-mail: ligs@igsnr.ac.cn

外，自然条件（如波浪、风场等）的不同也会对传感器接收到的后向散射能量有所影响。因此从理论上讲，要精确反演海洋悬浮泥沙浓度，所建立的模型必须包含以上所有影响因素。但是，由于实际操作中存在着种种难度，过去研究所建立的海洋悬沙遥感反演模式很少考虑泥沙颗粒大小、形状、矿物组成等要素的作用。

悬浮泥沙颗粒粒径因素是除浓度因素以外影响泥沙光谱反射贡献的最重要因素之一。恽才兴等实验研究表明，不同粒径的泥沙具有明显可区分的反射率曲线，其差异主要在于反射能量的大小和光谱敏感区段上<sup>[16]</sup>。黄海军等也曾提出增加泥沙的粒度参数对于提高该类模型的精度、增加其应用范围及进行不同区域的对比等方面作用较大，并建议在泥沙定量反演模型中加入粒径特征参数，但具体模式并未实现<sup>[9]</sup>。因此，考虑粒径参数对泥沙反射光谱的贡献是必需的也是可行的，基于这一点，本研究在不考虑后向散射的其他影响因素（矿物成分、泥沙颗粒形状等）的条件下，研究悬浮泥沙浓度和粒径两者对传感器接收的离水辐射率的贡献，将反射光谱和泥沙粒径两个参数同时纳入反演模型中，从而建立起渤海海域悬浮泥沙的二元特征参数遥感反演模型，并与不考虑粒径影响的情况作对比分析。

## 2 数据

本研究选用了2000年8月至2004年6月间26对（天）实测悬浮泥沙浓度，以及对应日期的准同步MODIS遥感数据作为试验研究数据。数据的时间分布见表1。

表1 研究所采用的实测及对应的遥感数据的日期

Tab 1 Dates of data used in the study				
年份	2000	2001	2003	2004
日期	2000 08 29	2001 01 10	2003 07 22	2004 06 21
	2000 08 30	2001 01 11	2003 07 29	2004 06 22
	2000 08 31	2001 01 12	2003 08 13	2004 06 25
	2000 09 01	2001 01 13	2003 08 14	2004 06 26
	2000 09 02	2001 01 14	2003 08 17	2004 06 27
	2000 09 03	2001 01 15	2003 08 18	
	2000 09 04		2003 08 21	
			2003 08 22	

其中，MODIS遥感离水辐射率数据由1B级数据经辐射校正、几何校正后获得。实测泥沙浓度数据中的2004年度数据来自于国家卫星海洋应用中心的实测结果，2000、2001、2003年度数据中的一部分是通过渤海海域调查采集的浊度数据经浊度与浓度间的关系式转换获得。所有采样均在晴朗天气下进行，且时间与MODIS上午星和下午星过境时间相差在半小时之内，从而基本保证了数据的同步性。此外，所有实测泥沙浓度在0.3~418mg/l，采样站位相对均匀地分布于渤海海区。

粒径数据中2004年度数据（共42对）来自同步实测结果。2000、2001和2003年数据（共42对）中同步实测悬浮泥沙粒径资料缺乏，但由于渤海是一个相对封闭的海域，现代底质沉积主要是由表层悬浮颗粒物在海底经过长时间沉积而形成，虽然表层悬浮物颗粒粒径与现代底质沉积粒径之间难以存在瞬时性的严格对应，但从长期演化来看，底质沉积物颗粒粒度的区域分布，与表层悬浮颗粒物的粒径具有密切和稳定的对应关系。基于这一点，本研究中数字化了《渤海、黄海、东海海洋图集》中的底质粒径分布图（图1）中

粒径大小的等级分布, 作为这一时段中表层悬浮颗粒物粒径参数的替代数据, 为反演研究提供数据补充。

将得到的离水辐射率数据与实测泥沙浓度及粒径数据对应以后共获得 84 组数据, 所有数据几乎均匀地分布于整个渤海海域 (图 1)。为了建立模型并对模型进行评价, 需要将总体样本分为试验样本和检验样本 (10% 以上) 两部分。在实际试验研究中, 根据样本总量、空间分布以及泥沙浓度大小, 随机地将样本分成了实验样本 (70 组) 和检验样本 (14 组)。样本分组遵循了模式平衡和随机原则, 此外, 数据分组时还尽可能地考虑了样本模式间的平衡。



图 1 实测数据站点及渤海海域数字化底质粒径分布

Fig 1 Field data sites and sediment grain size distribution in Bohai Sea

### 3 模型建立

Bhargava 等对粒径范围为 0.0032~0.1253mm 的五类不同悬浮物质的反射光谱曲线进行了比较, 研究发现, 悬沙水体具有如下光谱特征: 在悬沙浓度一定的情况下, 悬沙水体的反射率和粒径成负相关关系; 在相同粒径条件下, 随着浓度的增加反射率增大<sup>[17]</sup>。通过对获取的 MODIS 遥感反射率与粒径直接的相关分析, 并参考文献 [18], 发现 MODIS 8-16 波段反射率与粒径之间同样存在负相关关系, 即同样浓度条件下, 随着粒径的增加, 反射率下降。因此, MODIS 8-16 波段的反射率( $R$ )与粒径( $D$ )之间存在  $R = m + n \times (1/D)$  的关系, 即  $(R - m) \times D = n$ , 其中  $m$ ,  $n$  为常数。根据研究需要, 本文将该式简化为  $R \times D = nX$  的格式, 其中  $X$  为常数, 即假设  $R$  与  $D$  成反比关系, 并将  $R \times D$  整体作为二元特征参数模型的参数, 以下简称二元参数。为进行包含粒径的模型和不包含粒径的模型性能对比, 本研究分别建立主成分模型和神经网络模型。

#### 3.1 主成分模型的建立

MODIS 波段众多, 但各波段间的信息量又存在很大的相关性。主成分分析是指用少量互不相关的指标代替原来较多的指标, 且尽可能多地反映原来指标体系的信息。本研究首先用主成分方法分别建立不包含粒径的一元特征参数主成分模型和包含粒径的二元特征参数主成分模型。

##### (1) 一元特征参数主成分模型

本研究采用的原始指标波段有 9 个 (MODIS 8-16 波段), 各波段之间存在很大的相关性, 信息重复。将原始数据进行主成分分析迭代运算 (25 次), 去除波段间相关性, 得出各主成分以及对应的方差解释量, 如表 2 所示。

第一主成分可以解释方差的 92.826%, 前三个主成分的方差解释量共达 96.548%, 基本可以满足要求。对所获取的前三个主成分进行多元回归分析, 笔者发现泥沙浓度与主

成分的自然指数之间存在最大相关系数和最小标准差。回归方程为：

$$S' = e^{2.223 + 1.193F1 - 0.292F2 + 0.043F3} \quad (R = 0.835, R^2 = 0.697)$$

其中，F1 为第一主成分，F2 为第二主成分，F3 为第三主成分。

表 2 一元主成分方差解释量

Tab 2 Variance explained of unary parameters

成分	初始特征值			选取方差量		
	总方差	方差解释量(%)	累积方差解释量(%)	总方差	方差解释量(%)	累积方差解释量(%)
1	8.354	92.826	92.826	8.354	92.826	92.826
2	0.218	2.417	95.243	0.218	2.417	95.243
3	0.117	1.305	96.548	0.117	1.305	96.548
4	0.096	1.063	97.611			
5	0.078	0.872	98.483			
6	0.048	0.537	99.020			
7	0.035	0.394	99.414			
8	0.031	0.340	99.754			
9	0.022	0.246	100.000			

(2) 二元特征参数主成分模型

对各波段二元参数 ( $R \times D$ ) 值进行相关分析，发现参数间也具有极高的相关度。因此，将 9 个二元参数进行主成分迭代运算 (25 次)，去除波段间相关性，得出新的主成分和方差解释量，如表3所示。其中，第一因子可以解释总方差的 93.311%，前三个因子的方差解释量已高达 96.854%。

所以可只取前三个主成分代替原来的 9 个参数，将前三个因子进入回归模型，得回归方程如下：

$$S' = e^{2.223 + 1.230F1 + 0.161F2 - 0.113F3} \quad (R = 0.846, R^2 = 0.716)$$

其中，F1 为第一主成分，F2 为第二主成分，F3 为第三主成分。

表 3 二元主成分方差解释量

Tab 3 Variance explained of binary parameters

成分	初始特征值			选取方差量		
	总方差	方差解释量(%)	累积方差解释量(%)	总方差	方差解释量(%)	累积方差解释量(%)
1	8.398	93.311	93.311	8.398	93.311	93.311
2	0.205	2.274	95.585	0.205	2.274	95.585
3	0.114	1.270	96.854	0.114	1.270	96.854
4	0.088	0.974	97.828			
5	0.073	0.808	98.636			
6	0.042	0.467	99.103			
7	0.035	0.390	99.493			
8	0.026	0.294	99.788			
9	0.019	0.212	100.000			

### 3.2 神经网络模型的建立

为监控训练过程使之不出现“过拟合”现象,并评价所模型的性能和泛化能力,在模型建立之前必须将输入样本进行归一化处理,保证数据值在 0~1 之间。

#### (1) 一元特征参数神经网络模型

Hornik 等证明了仅含一个隐层的前向 BP 网络,能以任意精度逼近定义在实数集上的任意非线性函数<sup>[19]</sup>。据此本文采用三层 BP 网络(即有 1 个隐层)。由于模型的输入参数为 MODIS 9 个波段的离水辐射率,输出参数为悬浮泥沙浓度,所以确定网络的输入层神经元为 9 个,输出层神经元为 1 个。在训练过程中逐步调整隐层的神经元数目,得到最佳隐层神经元为 4 个。

模型中的函数主要有传递函数、学习函数、训练函数和性能函数。通过综合分析,该模型在隐含层选用 S 型正切函数,  $\text{tansig}$ , 为输入层至隐含层的传递函数;选用线性函数,  $\text{purelin}$ , 作为隐含层与输出层间的传递函数;选择  $\text{learn\_gdm}$  作为学习函数,该函数为梯度下降动量学习函数,它利用神经元的输入和误差、权值或阈值的学习速率和动量常数,来计算权值或阈值的变化率,在本模型中具有很好的学习性能;选择  $\text{trainbfg}$  函数作为训练函数;  $\text{mse}$  函数作为性能函数。

模型中的参数包括学习率  $\eta$ 、动量因子  $\alpha$ 、初始权值及收敛误差边界值  $E$ 。这些参数对训练速度和精度有关键影响。为保证算法的收敛性,学习率  $\eta$  必须小于某一上限,一般取  $0 < \eta < 1$ 。学习率影响系统学习过程的稳定性,大的学习率可使权值每次的修正量过大,甚至导致权值在修正过程中超出误差极小值呈不规则跳跃而不收敛;小的学习率导致学习时间过长,但可保证网络收敛。所以,一般选取较小的学习率以保证学习过程的收敛性,通常在  $0.01 \sim 0.8$  之间。动量项是为了避免网络训练陷于局部极小点,一般取常量,通常在  $0 \sim 1$  之间,且一般比学习率大。在上述范围内通过对不同的  $\eta$  和  $\alpha$  的取值进行试验,得出网络的最佳学习率和动量因子  $\eta = 0.7$ ,  $\alpha = 0.9$ 。初始权值的选择对于防止局部极小点和提高网络收敛速度均有一定影响。由于 S 型传递函数的特性,一般要求初始权值在  $-0.5 \sim 0.5$  之间。本文考察了不同初始权、阈值的赋值范围对网络收敛速度的影响,在该范围内,对模型随机赋予初始权值。误差边界值的选择决定于对模型的收敛速度和样本的学习精度的要求。当  $E$  值较小时,学习效果好,但收敛速度慢,训练次数多。当  $E$  值较大时则相反。本文采用  $0.0002$  模型作为误差边界值,即当误差值小于  $0.0002$  时,认为学习完成,停止计算,输出结果。

对以上所构建的一元特征参数神经网络模型进行训练,选择训练步数为 100。训练过程的误差曲线如图 2 (左):网络收敛速度较快,在经过 80 步训练后网络性能已趋稳定,且误差满足要求。

#### (2) 二元特征参数神经网络模型

该网络的构建过程与不包含粒径的网络基本相同,

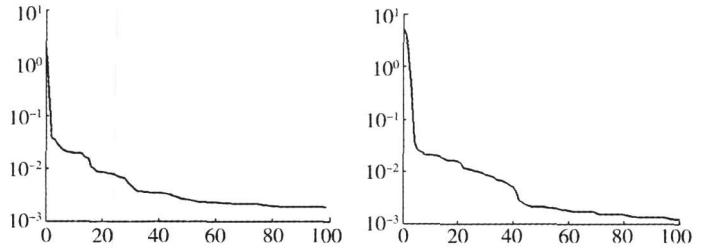


图 2 神经网络模型训练过程  
(左:一元特征参数模型;右:二元特征参数模型)

Fig 2 Training processes of the NN-based model  
(left: unary-parameter NN-based model;  
right: binary-parameter NN-based model)

采用只包含一个隐层的三层网络作为基础模型。输入向量比一元特征参数的网络多出一个粒径因子,将其单独作为一个输入元素,故该网络的输入层共10个神经元。经反复试验,考虑粒径情况的网络隐含层采用4个神经元时收敛效果最好。网络的输出同样只有悬浮泥沙浓度,因此输出层也只有一个神经元。

二元特征参数神经网络模型各函数为:输入层至隐含层选择tansig函数作为传递函数,隐含层至输出层采用purelin函数作为传递函数;learnngdm作为学习函数;traingdm函数作为训练函数,仍然采用mse函数作为性能函数。对于模型中的参数, $\eta=0.5$ , $\alpha=0.6$ ;初始权值在 $-0.5\sim+0.5$ 之间随机赋予;误差边界值 $E$ 仍采用0.0002。

选择训练步数为100,训练过程的误差曲线如图2(右):网络的收敛速度比不考虑粒径的网络更快,在经50次训练以后,网络的训练误差已基本满足要求。

4 结果分析

4.1 主成分模型

基于MODIS 8-16波段离水辐射率而建立的一元特征参数主成分模型和二元特征参数主成分模型的基本统计量如表4所示:

表4 主成分模型基本统计量  
Tab 4 Statistics of the principal components models

主成分模型	相关系数 ( $R$ )	决定系数 ( $R^2$ )	标准偏差 ( $Std. dev$ )	$F$ 检验值
一元特征参数主成分模型	0.835	0.697	0.828	34.143
二元特征参数主成分模型	0.851	0.724	0.802	55.399

一般在实际建模过程中采用决定系数 $R^2$ 和标准偏差 $Std. dev$ 作为评价标准, $R^2$ 值越大且 $Std. dev$ 越小,理论上模型的预测效果就越好。由表4可以看出,二元特征参数主成分模型的相关系数和决定系数明显比一元特征参数主成分模型高,而标准偏差却明显比一元特征参数模型低,这在一定程度上说明了二元特征参数模型性能比一元特征参数模型更为有效和可靠。

本文采用的置信水平 $\alpha=0.01$ ,由 $F$ 分布表得 $F_{0.01}(3,66)=4.13$ , $F_{0.01}(4,65)=3.65$ 。各反演模型的 $F$ 值远远大于对应的 $F_{\alpha}(m,n-m-1)$ 值,可见方程回归效果显著。

此外,对14组验证数据提取前三个主成分,分别代入前面所建立的一元特征参数和二元特征参数主成分回归模型中。

表5 主成分模型预测值及相对误差  
Tab 5 Predicted values and relative errors of the principal components models

站位	实测值 LnS	一元		二元	
		估算值	残差	估算值	残差
1	0.763677	1.70	0.936817	1.64	0.878508
2	0.869199	1.73	0.861115	1.35	0.485436
3	0.982078	1.25	0.272811	0.89	0.093759
4	1.039980	2.02	0.942345	1.43	0.375027
5	1.258461	1.42	0.171115	1.00	0.205379
6	1.762159	1.37	0.222545	1.22	0.307667
7	2.147454	2.03	0.054695	2.28	0.061722
8	2.434322	2.48	0.018764	3.10	0.273455
9	2.767349	1.88	0.32065	3.42	0.23584
10	3.058712	1.99	0.349399	2.69	0.120545
11	3.423659	4.18	0.220916	3.41	0.00399
12	4.018467	4.96	0.234302	3.28	0.183768
13	4.246639	1.74	0.590264	3.04	0.28414
14	5.179632	4.26	0.177548	4.74	0.084877
平均			0.383809		0.256722

得到预测值与实测值的相对误差 (表 5, 图 3)。

二元特征参数主成分模型的预测误差均值为 38.4%, 而二元特征参数主成分模型的预测误差均值为 25.6%, 可以看出在模型中加入粒径因子时, 平均相对误差比只考虑浓度对反射率贡献时有显著改善, 且误差除个别值外较之不考虑粒径贡献的情况下更为稳定 (图 3)。因此进行主成分模型构建时, 在模型中考虑粒径因子的贡献是非常必要的。

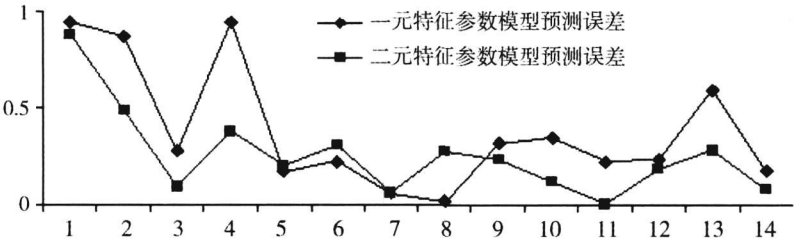


图 3 主成分模型预测误差曲线

Fig. 3 Testing error curves of the PCA-based models

4 2 神经网络模型

判断模型是否可有效逼近样本所蕴含的规律最直接和客观的指标是考察检验样本的误差。当检验样本误差比训练样本误差小或者相当, 一般可认为建立的模型已有效逼近训练样本所蕴含的规律; 否则, 则说明模型并未有效逼近训练样本所蕴含的规律, 而只是在训练样本点上逼近, 这时所建立的网络模型是对训练样本所蕴含规律的错误反映。将本研究中的检验样本输入上述所建立的一元特征参数模型和二元特征参数模型中, 得到模型的预测误差曲线, 并与模型的训练误差曲线进行比较分析, 结果如图 4。

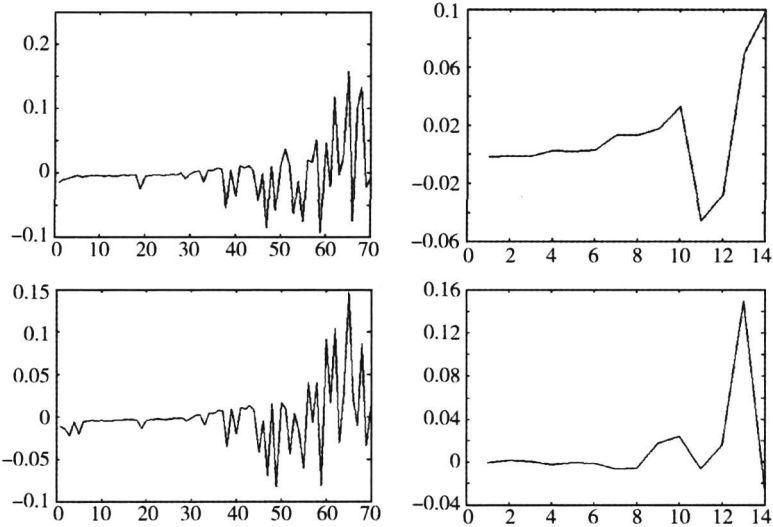


图 4 神经网络模型训练及预测误差曲线

(上左: 一元特征参数神经网络训练误差曲线 上右: 一元特征参数神经网络预测误差曲线  
下左: 二元特征参数神经网络训练误差曲线 下右: 二元特征参数神经网络预测误差曲线)

Fig. 4 Training and testing error curves of the NN-based models

(Upper left and upper right: training errors and testing errors of unary-parameter model)

(Down left and down right: training errors and testing errors of binary-parameter model)

由图 4 可以看出，在一元特征参数神经网络模型中，检验样本的误差非常接近训练样本的误差，因此，模型具有良好的泛化能力，可以很好地进行悬浮泥沙的预测。

在二元特征参数神经网络模型中，预测值与真实值之间的误差非常小。除第 10 次和第 13 次出现了两个比较大的误差之外，其余误差都在 0 左右，所有点误差均在 0.15 以内，此外，检验样本的误差与训练样本的误差非常接近，说明模型具有较好的泛化能力。因此，训练得到的模型可以很好地进行悬浮泥沙的预测。两模型得到的预测误差如表 6 及图 5 所示。

模型的预测误差结果显示，二元特征参数神经网络模型比一元特征参数神经网络模型的预测误差方面更优，前者的总体平均预测相对误差达到了 39.0%，而后者却只有 24.4%的。但由于训练样本有限，使得模型在个别点的误差较大，因此，为进一步提高网络模型的反演精度不但要在模型中加入粒径因子，还要增加模型的训练样本数量，并使训练样本尽可能覆盖区域的所有浓度范围。

表 6 神经网络模型预测结果及相对误差

Tab 6 Predicted values and relative errors of the artificial neural network models

序号	悬浮泥沙浓度 S (归一化结果)			相对误差 (%)	
	实测值	预测值 (一元)	预测值 (二元)	一元	二元
1	0.004616	0.003013	0.003882	0.347172	0.158983
2	0.005187	0.003904	0.006989	0.247345	0.347195
3	0.00587	0.004444	0.006549	0.24296	0.115702
4	0.006251	0.008696	0.00419	0.391188	0.329698
5	0.007904	0.009747	0.007174	0.233215	0.092307
6	0.013421	0.016372	0.012496	0.2199	0.06893
7	0.019974	0.033375	0.013548	0.670925	0.321716
8	0.026783	0.040093	0.020826	0.496957	0.222436
9	0.037573	0.055126	0.055095	0.467157	0.466329
10	0.050458	0.083502	0.074483	0.654884	0.476131
11	0.072911	0.027176	0.079012	0.627275	0.083676
12	0.132588	0.104616	0.148539	0.210971	0.120304
13	0.166704	0.237118	0.258999	0.422387	0.553641
14	0.424583	0.52347	0.397016	0.232903	0.064927
平均值				0.244427	0.390374

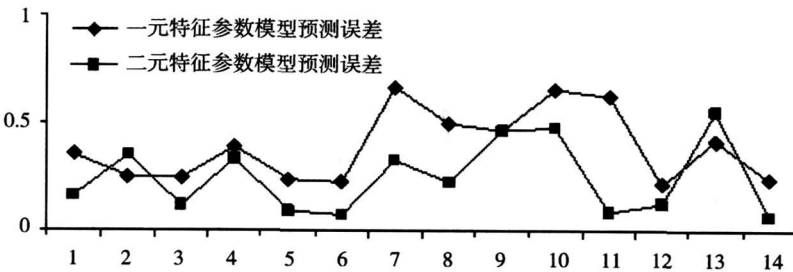


图 5 神经网络模型的预测误差曲线

Fig 5 Testing error curves of the NN-based models



总体而言,无论是主成分模型还是神经网络模型,当模型中加入粒径因子时,模型的预测误差和稳定性均比不加入粒径因子时有很大的提高,且总体来看神经网络模型的性能优于主成分模型。

## 5 讨论

由于神经网络模型的性能较优,本研究选取神经网络模型进行悬浮泥沙浓度的反演示范,并对反演结果进行分析讨论。

流入渤海的大部分河流处于大陆性气候地带,冬季干燥寒冷,夏季暴雨集中,年内有明显的枯水、丰水交替特征<sup>[20]</sup>。因此,试验研究中分枯水期和丰水期对泥沙浓度反演结果分别进行分析。试验研究选取了2003年3月18日和8月12日作为枯水期和丰水期的代表。将两天的遥感离水辐射率和对应的粒径数据应用于上述所建立的神经网络模型中,获得了悬浮泥沙浓度分布图(图版4图6)。总体来看,模型反演结果基本反映了渤海海域悬浮泥沙浓度的分布时间差异特征:丰水期的悬浮泥沙浓度普遍低于枯水期;并由近岸向远海海域逐渐降低。

枯水期悬浮泥沙浓度的空间分布特征为:整个海区的悬浮泥沙浓度普遍较高,渤海南部的高浓度区基本与等深线一致(近岸浅水区浓度较高,离岸深水区较低)。其中渤海湾,莱州湾以及辽东湾大部分地区的悬沙浓度都高于 $50\text{mg/l}$ ,且范围基本与岸线平行。辽东湾近岸出现高于 $450\text{mg/l}$ 的超高浓度区域。丰水期悬浮泥沙浓度的空间分布特征为:总体上悬沙浓度在大部分海域都低于枯水期,且小于 $5\text{mg/l}$ 的低浓度区占主导地位。高浓度区主要出现在渤海湾,莱州湾以及辽东湾,在渤海湾和莱州湾,部分地区甚至高于 $450\text{mg/l}$ ,主要是受河流输沙的强烈影响。

为进一步验证本研究所建立的模型及应用结果的正确性,将其与前人对渤海海域悬浮泥沙浓度的研究结果作了对比分析。秦蕴珊等曾分别于1958年和1962年两次按不同季节,对渤海海水中的悬浮体状况进行了实际调查和分析,并根据调查结果,对渤海悬浮体的分布状况等问题进行了较为详细的讨论,认为4月份枯水期海水中悬浮体含量较7月份丰水期整体较高,且高浓度区的范围也较7月份大很多。而4月和7月的高浓度区均位于渤海湾、莱州湾以及辽东湾海域的浅水区<sup>[21]</sup>。恽才兴等也曾应用TM影像对1998年渤海海域悬浮泥沙浓度及其分布进行反演,结果表明,冬季枯水期的渤海表层悬浮泥沙浓度较夏季丰水期整体较大,高浓度区均分布在渤海湾,莱州湾以及辽东湾部分海域,且高浓度区的分布基本与岸线平行<sup>[22]</sup>。这些研究结论与本文的反演结果基本一致。

## 6 结论

在总结前人研究成果的基础上,本文对渤海海域悬浮泥沙的光谱特性、MODIS数据特性及其与悬浮泥沙浓度和对应粒径的关系进行了分析与研究,并建立了渤海海域的一元特征参数和二元特征参数的主成分模型以及神经网络模型。通过对模型的对比分析及误差稳定性比较,发现在主成分模型中加入粒径因子贡献时,模型的决定系数由0.697提高到0.724,而模型预测误差由0.383809降低到0.256722。在神经网络模型中加入粒径因子时,模型的预测精度由0.390374降低至0.244427。由此说明,在悬浮泥沙浓度遥感反演模型中加入粒径因子的贡献,可有效地提高模型的反演精度和稳定性。此外,通过模型应用,选取典型日期的遥感影像对渤海海域的悬浮泥沙浓度分布的反演结果显示,模型反演

的趋势与前人实测调查结果及遥感反演趋势基本一致, 因此, 本项研究所建立的二元模型可以作为渤海海域悬浮泥沙浓度反演的有效手段。

诚然, 本项研究还是刚刚起步, 仍然有许多不足之处。例如, 为弥补数据不足, 本研究采用的一部分泥沙粒径数据是由海洋调查所获得底质粒径作为替代, 因此, 必然与悬浮泥沙粒径间会存在一定的偏差。要进一步提高模型精度, 定量考察粒径对后向散射的贡献程度, 还必须广泛获取同步的表层悬浮泥沙粒径数据。此外, 悬浮泥沙浓度的遥感反演除受泥沙浓度和粒径因素制约外, 还受矿物成分、形状等特性的影响, 在今后的研究中, 必须将这些因素逐步纳入反演模型中, 以进一步提高反演模型的精度和可靠性。

### 参考文献:

- [ 1 ] Ruhl C A. Combined use of remote sensing and continuous monitoring to analyze the variability of suspended sediment concentration in San Francisco bay California Estuarine, Coastal and Shelf Science, 2001, 53: 801~ 812
- [ 2 ] Meters L K, Smith M O, Adams J B Estimating suspended sediment concentration in surface waters of the Amazon River Wetlands from Landsat images Remote Sensing of Environment, 1993, 43: 281~ 301
- [ 3 ] Tassan S Local algorithms using SeaWiFS data for the retrieval of phytoplankton, pigments, suspended sediment, and yellow substance in coastal waters Applied Optics, 1994, 33(12): 2369~ 2378
- [ 4 ] 潘德炉, 李淑菁. 卫星海洋水色遥感信息特征量的研究. 遥感学报, 1998, 2(1): 26~ 31
- [ 5 ] 傅克付, 荒川久幸, 曾宪模. 悬沙水体不同波段反射比的分布特征及悬沙量估算实验研究. 海洋学报, 1999, 21(3): 134~ 140
- [ 6 ] 唐军武, 田国良. 水色光谱分析与多成份反演算法. 遥感学报, 1997, 1(4): 252~ 256
- [ 7 ] 万新宁, 李九发, 沈焕庭. 长江口外海滨悬沙分布及扩散特征. 地理研究, 2006, 25(2): 294~ 302
- [ 8 ] 恽才兴, 胡嘉敏. 遥感技术在河口海岸研究中的应用. 海洋通报, 1982, 11(2): 61~ 70
- [ 9 ] 黄海军, 李成治, 郭建军. 黄河口海域悬沙光谱特征的研究. 海洋科学, 1994, (5): 40~ 45
- [ 10 ] 陈涛, 李武, 吴曙初. 悬浮泥沙浓度与光谱反射率峰值波长红移的相关关系. 海洋学报, 1994, 16(1): 38~ 43
- [ 11 ] Li Yan, Huang W, Fang Ming An algorithm for the retrieval of suspended sediment in coastal waters of China from AVHRR data Continental Shelf Research, 1998, 18(5): 487~ 500
- [ 12 ] 李京. 利用 NOAA 卫星的 AVHRR 数据测杭州湾海域的悬浮泥沙含量. 海洋学报, 1987, 9(1): 132~ 136
- [ 13 ] 李炎, 李京. 基于海面- 遥感器光谱反射率斜率传递现象的悬浮泥沙遥感算法. 科学通报, 1999, 44(17): 1892~ 1897.
- [ 14 ] 李四海, 唐军武, 恽才兴. 河口悬浮泥沙浓度 SeaWiFS 遥感定量模式研究. 海洋学报, 2002, 24(2): 51~ 58
- [ 15 ] 李四海, 恽才兴. 河口表层悬浮泥沙气象卫星遥感定量模式研究. 遥感学报, 2001, 5(2): 154~ 160
- [ 16 ] 恽才兴, 等. 遥感文集. 北京: 科学出版社, 1981
- [ 17 ] Bhargava D S, Mariam Dejene W. Effects of suspended particle size and concentration on reflectance measurements Photogrammetric Engineering and Remote Sensing, 1991, 57(5): 519~ 529
- [ 18 ] Qin Y, Li F. Study of the suspended matter of the sea water of the Bohai Gulf Acta Oceanologica Sinica, 1982, 4(1): 191~ 200
- [ 19 ] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators Neural Networks, 1989, 2: 359~ 366
- [ 20 ] 李宁, 刘珍, 顾卫. 渤海与环渤海地区年降水量的统计分析. 地理研究, 2006, 25(6): 1022~ 1030
- [ 21 ] 秦蕴珊, 李凡. 渤海海水中悬浮体的研究. 海洋学报, 1982, 4(2): 191~ 200
- [ 22 ] 恽才兴. 海岸带及近海卫星遥感综合应用技术. 北京: 海洋出版社, 2005

## Two parameters retrieval models of suspended sediment concentration of Bohai Sea based on MODIS data

WANG Fang<sup>1, 2</sup>, LI Guo-sheng<sup>2</sup>

(1 University of Calgary, Calgary, Alberta T2N1N4, Canada;

2 Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China)

**Abstract:** In recent decades, remote sensing has been proven to be an effective method to retrieve suspended sediment concentration. However, most retrieval methods developed today only based on the direct relationship between suspended sediment concentration and the remote sensing reflectance which is not reasonable when there is a big difference among the grain size of the study area. Bohai Sea is a relative closure region in which the grain size of the suspended sediment varies widely. This paper brought out a new method on the retrieval of suspended sediment concentration of Bohai Sea, which used both the reflectance from remote sensing data and the grain size of the suspended sediment. After analyzing the spectrum characteristics of suspended sediment, MODIS data characteristics and the relationship between suspended sediment concentration and grain size, the unary-parameter and binary-parameter PCA-based and NN-based models were constructed in Bohai Sea based on the data collected. The analytical results show that when introducing grain size parameter into the PCA-based models, model's correlation coefficient was increased from 0.697 to 0.724, while its predicting error was decreased from 0.383809 to 0.256722. This phenomenon also happened in the NN-based models where the predicting error was decreased from 0.390374 to 0.244427. The stability of the models with a grain size parameter being also better than the one without the grain size. It is proved that a model's retrieval precision and stability can be improved effectively by introducing grain size into the model. Therefore, it's necessary to add the grain size into the retrieval model in order to improve the precision of the prediction. Moreover, representative remotely sensed MODIS imagery was used to validate the model built in this paper which has the same conclusion with the previous researches. Generally speaking, in Bohai Sea, suspended sediment concentration in winter low water period is higher than that in summer abundant water period; and high concentration areas are located in Bohai Bay, Laizhou Bay and Liaodong Bay and their distributions are parallel to the shoreline.

**Key words:** Bohai Sea; suspended sediment concentration; remote sensing; two parameters model

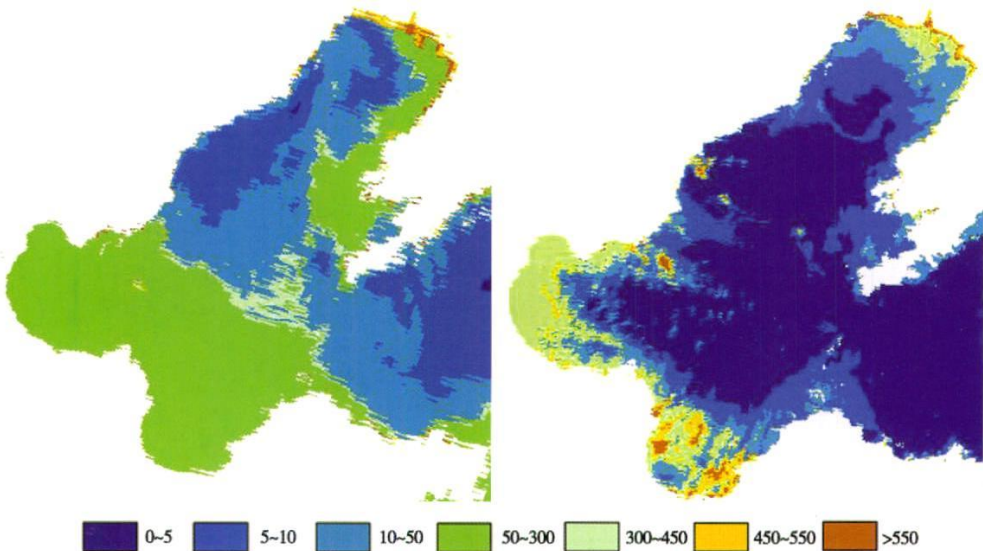


图 6 渤海悬浮泥沙浓度分布 ((左)2003 年 3 月,(右)2003 年 8 月)  
Fig.6 Suspended sediment concentration distribution results of Bohai Sea in March,2003  
(left) and in August,2003 (right)

曹小曙 等：火车站地区建设用地功能组合及空间结构——以广州站和广州东站为例

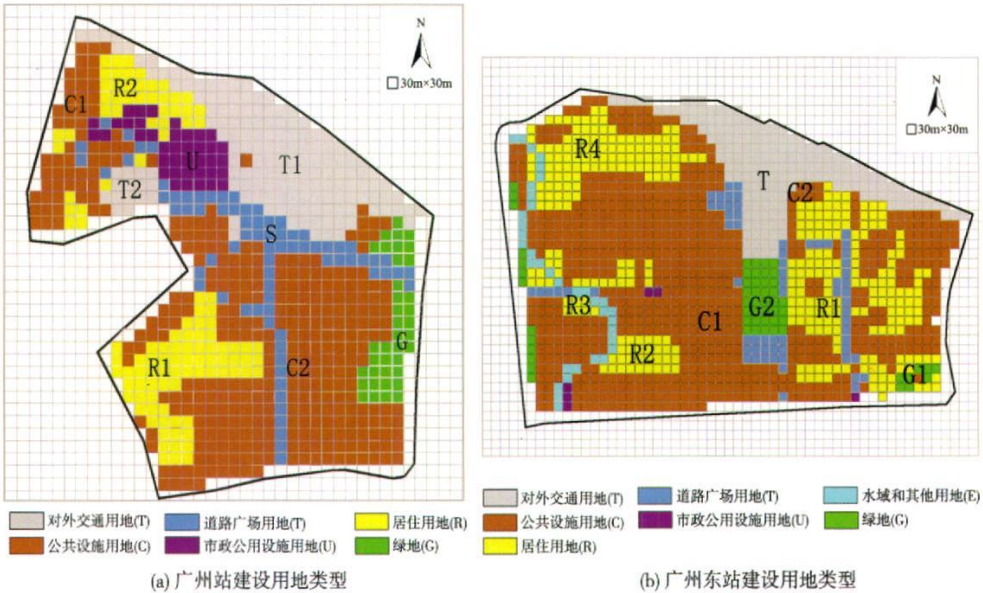


图 1 广州站和广州东站建设用地信息图  
Fig.1 Land use information map of Guangzhou railway station and railway eastern station areas