

# 半干旱区内陆湖泊透明度高光谱估测模型研究 ——以松嫩平原查干湖为例

宋开山, 张 柏, 王宗明, 段洪涛, 张渊智, 李 方

(中国科学院东北地理与农业生态研究所, 吉林 长春 130012)

**摘要:** 通过实测查干湖高光谱数据, 建立透明度 (Secchi Disk Depth, SDD) 单波段估测模型、比值估测模型以及神经网络高光谱估测模型, 并以确定性系数  $R^2$  以及剩余残差 RMSE 为指标进行了验证。通过对单波段估测模型和比值估测模型进行比较发现, 单波段模型估测结果与比值模型相差无几, 而水体透明度经对数处理有利于模型精度提高, 但是神经网络模型是三者中最优的。查干湖透明度高光谱定量估测模型的建立, 有利于今后利用遥感影像, 对查干湖水体透明度进行全面估测, 对于研究和监测查干湖水体水质状况有重要意义。

**关 键 词:** 半干旱区; 内陆湖泊; 透明度; 高光谱; 查干湖

**中图分类号:** P343.3      **文献标识码:** A      **文章编号:** 1001-6791(2006)06-0790-07

常规的地表水透明度观测主要采用塞氏盘 (Secchi Disk) 法, 在研究区域内定点剖面监测。这种方法不仅费时费力, 而且只能了解监测断面上的水质透明度状况, 对于整个水体而言, 这些测点数据只具有局部的代表意义。遥感技术具有快速、大面积和周期性的特点, 可以有效地解决这种局限性。同时, 在水体遥感中, 由于其光谱信息取决于多种物质的复杂集合体, 它包括纯水、溶解有机物质、浮游植物、悬浮物质等。各种信息重叠干扰, 造成各种目标物成分的反射模型及其光谱曲线不同, 呈现随机性与非线性。因此如何确定水体光谱信息中不同成分的贡献率就显得尤其重要, 而相对于传统多光谱传感器较低的光谱分辨率, 高光谱传感器由于其达纳米级的光谱分辨率, 可以获得目标物的诊断性光谱特性, 实现遥感信息模型参数或条件约束的确定, 很好地解决了常规遥感中出现的问题<sup>[1]</sup>。

## 1 实验方法及数据获取

### 1.1 研究区概况

查干湖大部位于吉林省西北部的前郭县境内, 大安市和乾安县各占水域的很小一部分。地处嫩江与霍林河交汇处的水网地区, 属内陆半咸水湖泊, 由查干湖和新庙泡、辛甸泡及库里泡组成, 是吉林省重要的渔业基地, 也是松辽平原第一大湖。

查干湖地处大陆性季风型半干旱气候区, 年均气温  $4.5^{\circ}\text{C}$ ,  $14^{\circ}\text{C}$  的水温日数 145d; 日照时数 2880 h, 年均辐射量  $4429\text{ MJ}/\text{m}^2$ , 5~9 月份辐射量为  $2319\text{ MJ}/\text{m}^2$ ; 年均降水量为 450.8 mm, 蒸发量为 1063.5 mm; 春季多大风, 其中风速  $17\text{ m}/\text{s}$  的日数年均 29 d; 冰封期 130 d 左右, 结冰期日数 150 d 左右, 冰厚 0.8~1.1 m。该湖正常运行水位 130 m 时, 水域面积  $3.72 \times 10^4\text{ hm}^2$ , 平均水深 2.52 m, 水体容积  $5.98 \times 10^8\text{ m}^3$ ; 集水区内均为盐

收稿日期: 2005-07-20; 修订日期: 2005-11-30

基金项目: 中国科学院知识创新工程重大项目 (KZCX1-SW-19); 国家自然科学基金资助项目 (40371082)

作者简介: 宋开山 (1974-), 男, 吉林白山人, 副研究员, 博士, 主要从事地物特征高光谱反演研究。

E-mail: songks@neigae.ac.cn

通讯作者: 王宗明, E-mail: zongmingwang@neigae.ac.cn

碱化农田和牧场; 湖底平坦, 湖盆为粉沙质土壤, 周围土壤为白钙碱土, 湖泊水质为苏打型盐碱水, 叶绿素含量较低, 多泥沙悬浮, 属富营养型湖泊<sup>[2]</sup>。

### 1.2 数据获取及预处理

为研究查干湖水质, 2004年5月到10月, 每月一次, 共计6次, 对研究区内代表点位分别用GPS进行定位导向, 同步测量各点水体透明度和反射光谱(采样点位置见图1), 并采集水样, 用以实验室分析其他水质参数。6次采样共计77点, 随机选取其中40个作为建模数据, 其余进行模型验证。水体透明度采用传统赛氏盘法; 反射光谱测量采用美国分析光谱仪器公司(Analytical Spectral Devices, ASD公司)生产的Fieldspec-VNIR(r)野外光谱仪, 可在350~1050 nm波长范围内进行连续测量, 采样间隔1.4 nm, 光谱分辨率3 nm, 视场角25°, 共有512个采样值。光谱测量时间为北京时10:00~14:00(太阳高度角大于45°), 天气晴朗, 微风或无风, 湖面基本平静。光谱测量在距水面上方1.5 m处进行, 光谱仪传感器探头基本垂直于水面, 同时与太阳入射与反射光构成的主平面正交, 避免镜面反射对光谱反射率的影响。在每个点位, 进行20次反射光谱测量, 最后取其平均值作为该点的光谱反射值, 每个样点测定前都进行暗电流与白板定标。对每条反射光谱反射率进行归一化处理, 选取400~1000 nm波谱数据作为高光谱模型研究光谱范围。

表1、表2是历次采样点水体透明度以及悬浮物、总有机碳含量(TOC)的基本统计信息。从中可以发现, 5月、9月以及10月份查干湖水体透明度较差, 这主要是由于松嫩平原中西部地区在这3个月份风力较大、水体浅、地势低平、湖面开阔, 因此当有风吹动湖面时, 很容易扰动底部碱土与粉沙, 致使查干湖透明度显著降低; 在6~8月份, 降水相对丰沛, 风力小, 因此水体透明度相对较大。同期样本的悬浮物则表现出相反的趋势, TOC的变化则出现一定的波动性。

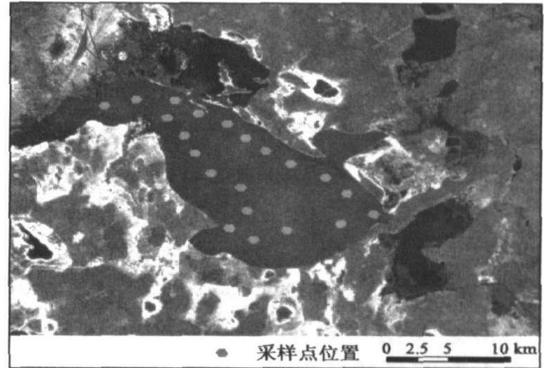


图1 查干湖及采样点空间位置

Fig. 1 Sampling positions on the Chagan lake

表1 各期水体采样点样本数及平均透明度

Table 1 Statistic informations of sampling in different months

| 采样时间 | 样本数/n | 平均值/cm | 标准差/cm | 最大值/cm | 最小值/cm |
|------|-------|--------|--------|--------|--------|
| 5月   | 11    | 14.4   | 6.20   | 30.0   | 9.0    |
| 6月   | 8     | 25.3   | 11.7   | 27.3   | 10.0   |
| 7月   | 8     | 22.5   | 0.96   | 24.0   | 22.0   |
| 8月   | 9     | 24.8   | 3.66   | 32.0   | 18.0   |
| 9月   | 20    | 14.1   | 1.43   | 17.0   | 12.0   |
| 10月  | 21    | 14.16  | 7.27   | 44.5   | 10.0   |

表2 各期水体样本悬浮物以及总有机碳含量平均值

Table 2 Average suspended matter and total organic carbon of water samples in different months

| 平均值 | 悬浮物/(mg L <sup>-1</sup> ) | TOC/(mg g <sup>-1</sup> ) |
|-----|---------------------------|---------------------------|
| 5月  | 47.04                     | 15.75                     |
| 6月  | 27.14                     | 10.30                     |
| 7月  | 19.5                      | 16.69                     |
| 8月  | 22.8                      | 12.42                     |
| 9月  | 38.1                      | 14.96                     |
| 10月 | 27.6                      | 19.73                     |

注:表2中的样本数量与表1对应月份相同。

### 1.3 透明度遥感原理

对水体来说,水的光谱特征主要是由水本身的物质组成决定,同时又受到各种水状态的影响。湖泊水体中影响光的强度及光谱分布的物质大致可分为3类<sup>[3,4]</sup>:浮游藻类色素,黄色物质和悬浮物质。除了悬浮颗粒物在其自然浓度条件下对光不发生明显吸收外,其余两种物质分别选择吸收一定波长范围的光,形成各自的吸收波谱。因此,水体组分含量的差别造成一定波长范围反射率的显著不同,成为用遥感模型定量监测水质参数的基础<sup>[5]</sup>。

## 2 模型建立及分析

### 2.1 不同季节水体光谱反射率的变化趋势

在本研究中共对查干湖进行了6次取样,基本代表了温带内陆湖体的季节变化趋势。由表1、表2和图2中的光谱曲线变化趋势的相互认证表明,查干湖较为浑浊,透明度低,而且透明度与光谱反射率的季节变化较大。从图2可以发现查干湖水体波长在850 nm之前的可见光波段,不同季节各个采样点的光谱反射率差异明显,为水体透明度的遥感反演奠定了物理基础;在550 nm、705 nm和810 nm左右有明显的反射峰,而在430 nm、670 nm以及900 nm左右则存在3个谷区,这些反射峰与吸收谷是水体悬浮物、黄色物质以及叶绿素共同作用下形成的,是构建光谱比值法,建立反演模型的诊断性波段。

### 2.2 单波段数据与水体透明度相关分析

在对反射光谱数据进行归一化处理,在400~1000 nm之间逐波段分析光谱反射率、导数光谱与采样点水体透明度之间的相关关系,并在此基础上建立各个波段与水体透明度的线性回归关系。经以上分析后,相关系数、线性回归确定性系数随波长的变化趋势见图3。由图3(a)可以看出,在蓝光区域(400~500 nm),其相关系数平均在-0.51上下,相关度比较好,且随着波长增加,相关系数略有增加;绿光区域(500~600 nm),相关系数平均在-0.52左右,与蓝光相比偏高,在绿光的长波边缘区域相对较高,这是由于绿光、蓝光具有较深的穿透能力,与Koponen等的研究结果相似<sup>[6]</sup>。红光区域(600~760 nm),相关系数平均在-0.70左右,有较大增高。近红外区域(760~1000 nm),相关系数可以分为两个区域,760~890 nm,相关系数都在-0.78以上,是我们研究波段区域中相关度最高的区域;从890 nm以后,相关度急剧降低,到950 nm时,相关系数变为-0.2。红光与近红外部分波段与透明度相关系数较高,其原因可能是由于查干湖水体混浊度较高,透明度受水中非生物物质对光吸收和散射的影响大于藻类生物量的影响,这种情况在我国泥沙含量高的大型浅水湖泊中较为普遍,而在500~800 nm波段可以反映悬浮物含量。更为重要的原因可能是,蓝绿光对水体叶绿素含量比较敏感,而查干湖水体叶绿素含量相对较低,同时悬浮物对红光、近红外波段光谱反射率影响较大,因而这一区域反射率与透明度有较高相关度,Harna在研究芬兰湖泊时也发现近红外光谱与透明度有一定相关<sup>[10]</sup>。导数光谱在红光与近红外波段与透明度的相关系数较大,而在蓝、绿光波段相关系数较小。通过对比可以发现,导数光谱只有在少数波段的相关系数大于光谱反射率与透明度的相关系数,而在绿光、红光以及近红外短波方向的相关系数相对较低。由图2可以发现在蓝、绿光波段各个样本的光谱曲线有一定的交叉,而在红光与近红外波段,各个样本光谱曲线差异明显。因此各个光谱曲线的变化趋势进一步说明了为什么在红光与近红外波段的相关系数较高,而在蓝、绿光波段的相关系数较低。

由图3(b)可知,光谱反射率、导数光谱与透明度的线性回归的确定性系数与相关系数随波长的变化趋势密切相关,同样是在红光与近红外波段取得较大值或最大值,在红光的长波方向以及近红外的短波方向,确定性系数 $R^2$ 在0.61左右,在置信度 $P < 0.001$ 水平上,说明水体高光谱反射率与透明度之间具有统计分析意义。

### 2.3 单波段估测模型的构建

选取单波段与透明度相关系数最大的特征波长730 nm,与透明度做一元回归分析,建立的一元回归线性方程见图3;由于6次采样水体透明度、悬浮物含量的变化较大(表1),因此这里对水体透明度数据进行对数处理,然后建立一元回归方程(图4)。对比图3(b)与图4,可以发现非线性回归对提高模型的反演精度略有帮助,但确定性系数 $R^2$ 只提高了0.014。但是对透明度进行对数处理后,再建立模型则确定性系数会有较为显著的改

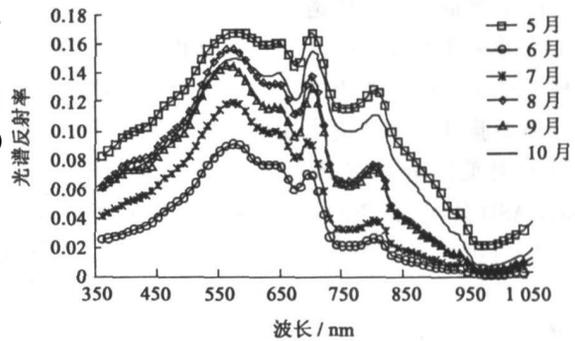


图2 采样点典型水体光谱反射率随波长的变化趋势

Fig. 2 Spectral reflectance characteristics of all samples

善， $R^2$  提高了 0.1(图 5)。对其他波段做如上处理，结果发现非线性回归对模型的精度提高能力是有限的，都在百分之几级别上。

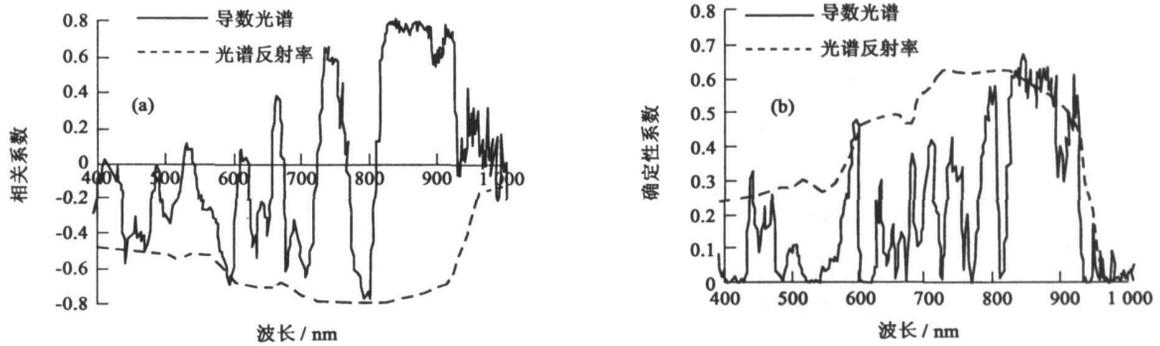


图 3 水体光谱反射率、微分光谱与查干湖水体透明度的相关分析(a)、线性回归关系(b)

Fig.3 Relationship between water reflectance, derivative reflectance and Secchi Disk depth (SDD) in Chagan Lake

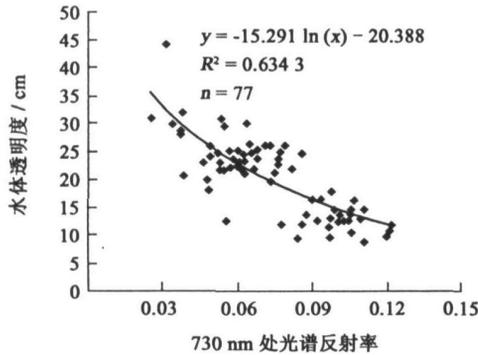


图 4 730 nm 反射率与透明度的回归关系

Fig.4 Regression between spectral reflectance of 730 nm and SDD

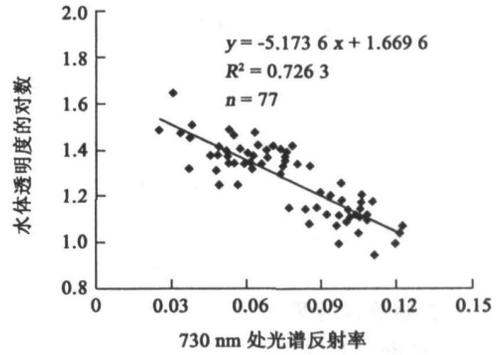


图 5 730 nm 反射率与透明度对数关系

Fig.5 Relation between spectral reflectance of 730 nm and logarithm of SDD

对 730 nm 处的微分光谱与透明度做非线性回归分析,结果为

$$y = - 15.723\ln(x) - 21.019 \tag{1}$$

式中  $y$  为水体透明度;  $x$  为 730 nm 处微分光谱; 回归的确定性系数  $R^2 = 0.6326$ 。对透明度进行对数处理后, 回归结果为

$$y = - 5.2297x + 1.6838 \tag{2}$$

式中  $y$  为水体透明度的对数值,  $x$  含义同上。回归的确定性系数  $R^2 = 0.7023$ 。当采用 830 nm 左右的微分光谱作为自变量时, 以上两种回归方式所得到的确定性系数都有明显提高。

### 2.4 比值估测模型的构建

使用两波段反射率比值法对水质参数进行研究, 有利于校正大气、光照条件等环境背景对光谱反射率的影响<sup>[6]</sup>。通过前人的研究成果以及对图 2 不同样点光谱数据随波长变化趋势的分析, 对部分可能的波段进行了组合运算, 最后发现 780 nm 和 684 nm 组合效果最好。使用其建立透明度线性估测模型(图 6):

$$y = - 41.59x + 47.333 \tag{3}$$

式中  $y$  为水体透明度, cm;  $x$  为 780/684 组合波段反射率的比值。

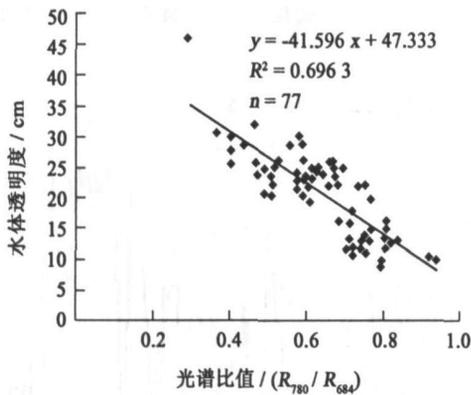


图6 光谱反射率比值和透明度的回归关系

Fig.6 Correlation between spectral reflectance ratio and SDD

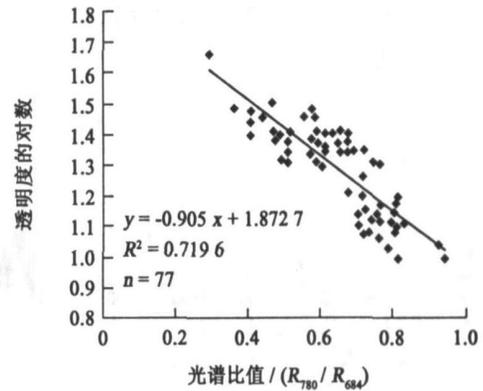


图7 光谱反射率比值和透明度对数的关系

Fig.7 Correlation between spectral reflectance ratio and logarithmic of SDD

在本研究中仍然对水体透明度进行了对数处理, 来实现对采样数据的归一化, 然后再建立模型(图7), 通过对比图5、图6的数据散点的分布可以发现, 水体透明度经对数处理后, 回归模型的拟合度有了略微提高, 但是对回归精度并没有太大的改善。对比图3、图4与图5、图6, 可以发现比值估算模型的精度并没有得到改善, 甚至还没有单波段构建的模型精度高, 这种现象在地物高光谱定量遥感中经常出现<sup>[6]</sup>, 这是由于高光谱数据能够直接发现某些地物物质组成的诊断波段, 通过这些波段构建的单波段模型一样可以提高模型的反演精度。作者还与张霄宇、吴敏等研究结果进行了对比<sup>[9,10]</sup>, 发现本研究的波段比值模型精度略低于同类研究。分析认为, 一是由于各类水体组成成分不同; 二是由于本研究的时间跨度大, 不同季节的数据综合在一起分析, 不确定环境因素更多, 影响了模型精度。

### 2.5 水体透明度神经网络模拟

人工神经网络是20世纪80年代中后期世界范围内迅速发展起来的一个前沿领域, 因其良好的预测性和实用性被广泛应用于各个领域, 尤其在遥感影像自动分类与定量分析方面得到了广泛应用<sup>[11~13]</sup>。由于BP神经网络具有并行处理、非线性、容错性、自适应和自学习的特点, 在数据拟合与模拟中有着无比的优越性。因此本文将尝试采用BP神经网络, 把多个波段的反射率作为输入矢量来预测查干湖水体透明度随光谱反射率的变化, 改善以往单波段或比值法建立模型参与波段少、反演精度低的缺陷。

本研究采用的BP神经网络模型是由Matlab的Neural Network Toolbox提供, 网络共有3层, 依次为输入层、隐藏层和输出层。输入层为不同诊断波段的光谱反射率, 中间层的神经元个数为2~10之间的可变因子; 输出层神经元为1。在77个水体透明度数据中随机抽取40个作为训练样本的学习目标 $T$ , 将450、650、720、800 nm处的反射率作为输入矢量 $P$ , 设置网络学习的迭代次数为300。

当神经网络训练完毕后, 把剩余样本的光谱数据作为输入矢量 $P$ , 进行模拟所得预测值与实测值进行拟合, 结果如图8(a)所示。当隐藏层的节点数为9时, 确定性系数 $R^2$ 与剩余残差 $RMSE$ 分别为0.9806、0.8710 cm。可以看出, 神经网络的模拟精度远远高于单波段与比值法构建的回归模型。

由图3(b)微分光谱与水体透明度的线性回归的趋势可以发现, 水体透明度与微分光谱在部分波段关系密切, 因此在本研究中将480、590、709、810 nm波长处的微分光谱作为输入变量, 进行了神经网络模拟, 得到的结果见图8(b)。从图8(b)的结果可以看出, 以微分光谱作为输入变量的神经网络模型同样可以收到良好的反演精度。笔者在本研究过程中, 曾尝试用不同波段的反射率以及微分光谱值作为输入变量进行神经网络模拟, 都收到了良好的模拟结果, 但是当输入变量与水体透明度相关性较低时, 神经网络的模拟精度就会有所下降, 因此在神经网络模拟构成中, 所选择的波段一定要与模拟目标有一定的相关性。

以上分析了隐藏节点数为9时的模拟精度, 当隐藏层的节点数为2~8时,  $R^2$ 与 $RMSE$ 的变化情况以及二

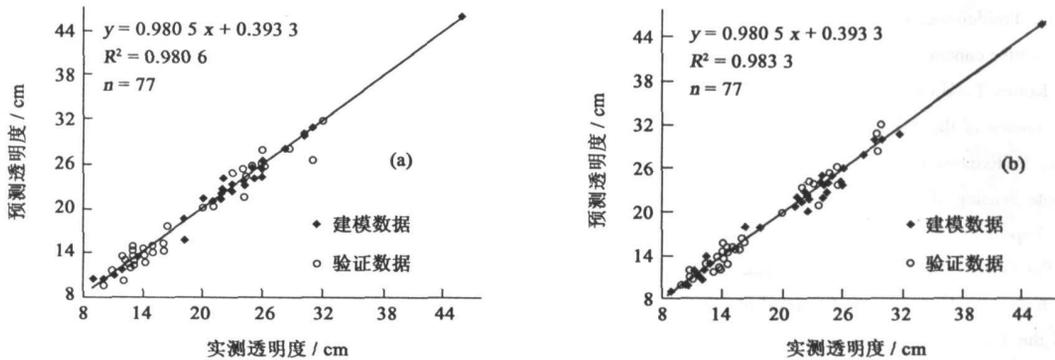


图 8 神经网络预测查干湖水体透明度与实测值之间的关系以及模型验证

Fig. 8 Relation between predicted by ANN-BP and measured SDD, and calibration of ANN-BP

者的对比情况见表 3。从表中可以看出当人工神经网络的隐藏层节点数由 2 增加的 8 时，对应  $R^2$ 、 $RMSE$  的变化情况，基本上是随着节点数量的增加，网络的模拟精度在不断增加，但也有个别情况下，节点数增大反而预测精度略有下降。一般而言，当节点数为 7~8 时，网络的模拟精度已经趋于稳定，节点数的增加对模拟的精度不会有太大的提高，而且节点过多可能出现过度拟合，因此在以神经网络模型进行水体透明度反演时，一定要在大样本事件的前提下进行，并且网络的训练以及隐藏层节点数都要合理控制，避免过度训练与过度拟合现象的发生。

表 3 隐藏层含有不同节点时，神经网络的模拟精度  $R^2$ 、 $RMSE$  的变化情况

Table 3  $R^2$  by ANN BP prediction accuracy and  $RMSE$  with different nodes in the hidden layers

| 所采用光谱反射率的波段 |           | 450 nm | 650 nm | 720 nm | 800 nm |        |        |        |
|-------------|-----------|--------|--------|--------|--------|--------|--------|--------|
| 精度评价指标      |           | 隐藏层节点数 |        |        |        |        |        |        |
|             |           | 2      | 3      | 4      | 5      | 6      | 7      | 8      |
| 透明度         | $R^2$     | 0.7731 | 0.8877 | 0.9062 | 0.8751 | 0.8959 | 0.9425 | 0.9653 |
|             | $RMSE/cm$ | 3.3240 | 2.3383 | 2.1365 | 2.4662 | 2.2516 | 1.6738 | 1.2992 |
| 所采用微分光谱的波段  |           | 480 nm | 590 nm | 709 nm | 810 nm |        |        |        |
| 透明度         | $R^2$     | 0.8223 | 0.8371 | 0.8500 | 0.8668 | 0.9430 | 0.9673 | 0.9785 |
|             | $RMSE/cm$ | 2.9411 | 2.8160 | 2.7022 | 2.5468 | 1.6657 | 1.2612 | 1.0781 |

### 3 结 论

利用高光谱遥感监测模型对查干湖透明度进行估算和监测，能够获取较为准确的评价结果，相对于传统监测方法具有省时省力的特点。通过对单波段估测模型和比值估测模型进行比较发现，单波段模型估测结果好于比值模型，这与他人的研究成果有些不同，可能与查干湖的水质特征有很大关系。而对数比值模型又强于单纯的比值模型，这是由于求对数可以对光谱数据进行简单归一化处理，弱化了外界因素对数据的干扰，提高了数据精度。在本研究中，神经网络模型的反演精度是令人满意的，当神经网络的隐藏层节点数为 2 时，其预测精度已经明显优于单波段以及比值法建立的回归模型，随着节点数的增加，网络的预测精度得到了进一步提高。在今后的研究中，如何更为有效的控制采样时的环境干扰，以及如何纠正镜面反射以及随季节变化而引起的太阳高度的变化对采集的光谱数据的影响将要引起特别的关注。查干湖高光谱估测模型的建立，有利于今后结合遥感影像，对查干湖水体水质状况进行全面研究和监测，并对进一步应用到其他内陆湖泊具有参考意义。

#### 参考文献：

[1] 韩震, 陈西庆, 恽才兴. 海洋高光谱遥感研究进展[J]. 海洋科学, 2003, 27(1): 22 - 25.  
 [2] 杨富亿. 查干湖的综合开发与利用[J]. 资源开发与市场, 1998, 14(6): 247 - 249.  
 [3] 张运林, 秦伯强, 陈伟民. 湖泊光学研究动态及其应用[J]. 水科学进展, 2003, 14(5): 653 - 659.

- [4] Doxarana D, Froidefonda J M, Lavender S, *et al.* Spectral signature of highly turbid waters: Application with SPOT data to quantify suspended particulate matter concentrations[J]. *Remote Sensing of Environment*, 2002, 81: 149 - 161.
- [5] Kallio K, Kutser T, Hannonen T, *et al.* Retrieval of water quality from airborne imaging spectrometry of various lake types in different seasons [J]. *The Science of the Total Environment*, 2001, 268: 59 - 77.
- [6] Koponen S, Pulliainen J, Kallio K, *et al.* Lake water quality classification with airborne hyperspectral spectrometer and simulated MERIS data [J]. *Remote Sensing of Environment*, 2002, 79: 51 - 59.
- [7] Harma P, Vepsäläinen J, Hannonen T, *et al.* Detection of water quality using simulated satellite data and semi-empirical algorithms in Finland[J]. *The Science of the Total Environment*, 2001, 268: 107 - 121.
- [8] Kallio K, Kutser T, Hannonen T. Retrieval of water quality from airborne imaging spectrometry of various lake types in different seasons[J]. *The Science of the Total Environment*, 2001, 268:59 - 77.
- [9] 张霄宇, 林以安, 唐仁友, 等. 遥感技术在河口颗粒态总磷分布及扩散研究中的应用初探[J]. *海洋学报*, 2005, 25(1): 51 - 56.
- [10] 吴 敏, 王学军. 应用 MODIS 遥感数据监测巢湖水质[J]. *湖泊科学*, 2005, 17(2): 110-113.
- [11] Keiner L E. A neural networks model for estimation sea surface chlorophyll and sediments from Thematic Mapper Imagery. *Remote Sensing of Environment*, 1998, 66: 153 - 165.
- [12] Danson F M, Rowland C S. Training a neural network with a canopy reflectance model to estimate crop leaf area index[J]. *International Journal of Remote Sensing*, 2003, 24(23): 4891 - 4905.
- [13] Zhang Y, Pulliainen S, Koponen, *et al.* Application of an empirical neural network to surface water quality estimation in the Gulf of Finland using combined optical data and microwave data[J]. *Remote Sens Environ*, 2002, 81: 327 - 336.

## Hyperspectral model for transparency of Chagan lake in semi-arid environment of Songnen plain<sup>\*</sup>

SONG Kai-shan, ZHANG Bai, WANG Zong-ming, DUAN Hong-tao, ZHANG Yuan-zhi, LI Fang

(*Northeast Institute of Geography and Agricultural Ecology, CAS, Changchun 130012, China*)

**Abstract:** The transparency, which can demonstrate the limpid or muddy degree of lake water also be used to evaluate water eutrophic state, is one of the water body visible degree indicators. The traditional transparency of surface water is observed with Secchi Disk. It is time-consuming and strenuous, and only represents some local information. The remote sensing technology, which can cover large area simultaneously and periodically, may deal with this kind of limitation effectively. In this paper, 77 spectral reflectance data of different water sites in Chagan lake were collected during 6 field works, and the water transparency data were acquired simultaneously. The linear regression constructed with every single band of reflectance and derivative against the water transparency data; the band ratio model with the reflectance of 780 nm and 654 nm also established. Finally, ANN - BP model is established with diagnostic band reflectance and derivative as the input vector. The result shows that water transparency has an intimate relation with water reflectance, and correlation coefficient is about 0.5 in blue and green spectral region, it is about 0.6 in red spectral region; while it obtains the maximum value in short wave of near infrared region. As for derivative reflectance, it also obtains higher value in near infrared spectral region. The regression model established with single band gets the similar performance as that with band ratio as regression dependent variable. By comparison, the ANN - BP model performed best with determination coefficient ( $R^2$ ) of 0.98. It indicates that hyperspectral remote sensing models established for estimation Chagan lake water transparency can obtain the comparatively accurate result. The construction of hyperspectral models in Chagan lake will to help estimate the water body transparency with satellite images, and also provide a method for remote sensing monitoring of other inner water body with similar water status as that in Chagan lake.

**Key words:** Semi-arid region; inland lake; Secchi Disk Depth; Hyperspectral; Chagan lake

\* The study is financially supported by the National Natural Science Foundation of China (No. 40371082).