

基于知识和遥感图像的神经网络水质反演模型

石爱业¹⁾ 徐立中¹⁾ 杨先一^{1),2)} 黄凤辰¹⁾

¹⁾(河海大学计算机及信息工程学院, 南京 210098) ²⁾(Guelph 大学工程学院, Guelph, 加拿大)

摘要 为进一步提高遥感图像水质反演的精度,提出了一种基于知识和遥感图像相结合的神经网络水质反演模型。该模型利用遥感图像数据以及与水质相关的知识数据作为BP神经网络的输入,经训练后,确定神经网络的结构,在训练好的BP神经网络基础之上对水质进行反演。以中国太湖为例进行实证研究,实验中,使用的知识数据包括太湖的地理信息知识和先对太湖TM图像上的水域解译进而对水质进行分类的知识。实验结果表明,本文提出的水质反演模型较常规的线性回归模型和传统的神经网络模型有更高的反演精度。

关键词 遥感图像 水质反演 BP神经网络 知识

中图分类号: TP751.1 X824 **文献标识码**: A **文章编号**: 1006-8961(2006)04-0521-08

A Neural Network Model for Water Quality Retrievals Using Knowledge and Remote-sensed Image

SHI Ai-ye¹⁾, XU Li-zhong¹⁾, YANG Xian-yi^{1),2)}, HUANG Feng-chen¹⁾

¹⁾(College of Computer and Information Engineering, Hohai University, Nanjing 210098)

²⁾(School of Engineering, University of Guelph, Guelph, Canada)

Abstract In order to improve water quality retrievals of remotely sensed image accurately, this paper puts forward a neural network model for water quality retrievals using knowledge and remotely sensed image. The model uses remotely sensed image data and water quality related knowledge as input of BP neural network, then trains neural network, after that water quality is retrieved by the trained neural network. The proposed model is applied to the water quality retrievals of Tai Lake in China. In experiment, knowledge used includes Tai Lake geography information knowledge and classification knowledge of water quality by interpretation of TM image. The result of experiment shows that the developed model has more accuracy than the routine linear regression model and traditional neural network model.

Keywords remotely sensed image, water quality retrievals, BP neural network, knowledge

1 引言

针对不同地域、不同特点的水质参数反演,如何建立较合适的反演模型是一个研究难点。常规的遥感水质反演的方法主要有经验方法、分析方法和半分析方法^[1],这些方法本质上都是通过建立线性回归模型来实现水质反演的,虽然简单,但是也存在明显不足,因为遥感图像数据与水质参数之间是非线

性关系,利用线性回归来估计,则精度差,而且数据之间的互补性比较小,反演的结果不准确。针对线性回归方法的不足,国内外许多学者采用非线性回归方法来进行水质反演。非线性回归方法既可以是显式,也可以是隐式的,其中显式的模型参数常常难以确定,而隐式的模型参数则避开了上述问题。常用的隐式回归模型主要有神经网络回归模型^[2,3]和遗传算法回归模型^[4]。另外,也有学者利用其他方法与神经网络方法结合起来而建立模型进行水质反

基金项目:国家自然科学基金项目(60374033);江苏省自然科学基金项目(BK2002064)

收稿日期:2005-02-23; 改回日期:2005-05-19

第一作者简介:石爱业(1969~),男,讲师。2002年获南京理工大学光学工程专业硕士学位。目前从事信息处理、信息融合等方面的研究。E-mail:shiaiye@sina.com

演的研究,试图提高水质的反演精度^[5,6]。

已有文献都是从利用水质的光谱特性和遥感图像数据之间内在的关系入手,来构造反演模型。这些方法都有如下的不足:(1)仅仅考虑遥感图像数据与水质参数之间的关系,而没有考虑到与水质关联的其他相关知识;(2)已有的反演模型,都仅仅是针对监测点本身的反演精度进行分析,而对其相邻的非监测点的水质反演的精度没有给予分析。

本文提出的水质反演模型中,把与水质相关的知识引入该模型,是考虑到如下原因:(1)待反演的水质参数的分布不仅与遥感图像数据有关,而且还与所要反演的水域的地理结构以及遥感成像当时的气候条件等有关;(2)依据遥感水域的目视解译知识对水域进行分类,对于水质参数的分布可以有总体上的了解;(3)尽管这些知识本身与水质的光谱特性没有直接的联系,但是如果能够在建立反演模型中考虑这些与水质相关的知识,那么对水质反演是一个有效的补充。另外,水质反演模型中之所以采用神经网络方法,是因为神经网络是一种有效的非线性逼近方法,而这正与待反演的水质参数与遥感图像数据及水质相关的知识数据之间具有的非线性关系相符合。

2 基于知识和遥感图像的神经网络水质反演模型

基于知识和遥感图像的神经网络水质反演模型框图如图 1 所示,该模型中神经网络的训练集是由与水质相关的知识数据、地面监测的水质数据和与地面监测点对应的遥感图像数据组成(图 1(a)),其中,水质相关的知识数据与地面监测点对应的遥感图像数据作为学习样本,地面监测的水质数据作为目标样本。在神经网络已训练好的基础之上,把部分或者全部遥感图像数据和相应的知识数据输入到

训练好的神经网络,得到部分或者整个水域最终的水质反演结果,如图 1(b)所示。

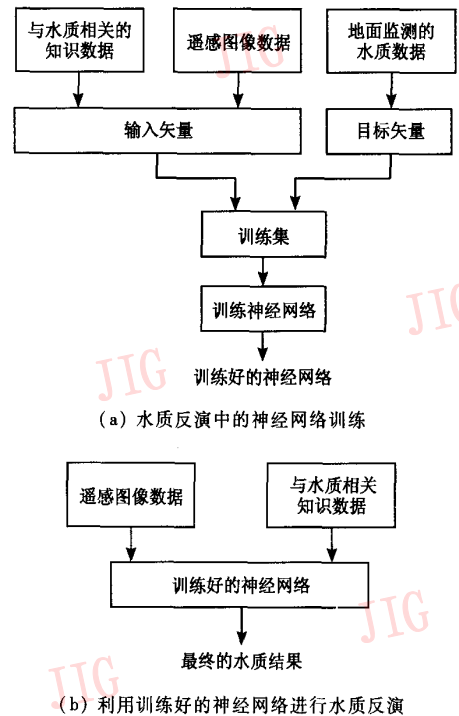


图 1 基于知识和神经网络结合的水质反演模型
Fig. 1 The proposed neural network model for water quality retrievals

不失一般性,设神经网络具有 3 层结构,即输入层、隐层和输出层,如图 2 所示。记输入向量为 p (由与水质相关的知识数据、遥感图像数据组成), W_1 为输入层与隐层的神经元之间的权值矩阵, W_2 为隐层与输出层的神经元之间的权值矩阵, b_1 、 b_2 为隐层、输出层的阈值矩阵, f_1 、 f_2 为隐层、输出层的传递函数, O_1 、 O_2 为隐层、输出层的输出,则

$$O_1 = f_1(W_1 \times p + b_1) \tag{1}$$

$$O_2 = f_2(W_2 \times O_1 + b_2) \tag{2}$$

O_2 再经过尺度转换就是待反演的水质。

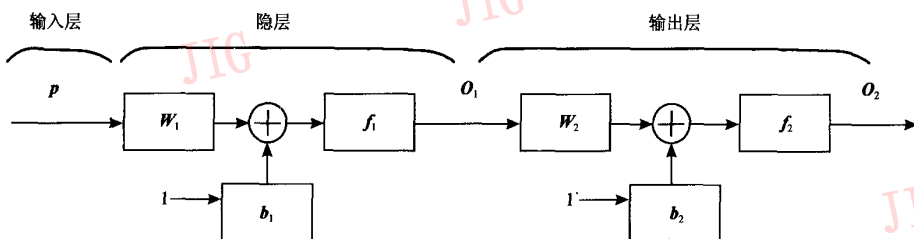


图 2 3 层的神经网络结构
Fig. 2 Structure of three layers neural networks

2.1 水质反演模型中神经网络的选择

在水质反演模型中采用BP神经网络进行水质的反演。因为水质反演本质上是函数拟合问题,尽管BP神经网络有一些固有的缺点,比如训练速度慢、学习过程容易陷入局部极小等缺点,但是由于BP神经网络是全局逼近网络,每个输入都和输出有关,加上BP网络每一层的传递函数通常是Sigmoid型可微函数,所以可以实现输入与输出之间的任意非线性关系,能够较好地实现函数拟合。又BP神经网络在经过适度的训练后,具有较好的推广能力,这一点在水质反演中尤显重要,因为在水质反演中,建立的模型不仅要能够对监测点本身给出输入到输出的正确映射,而且在非监测点也能够给出输入到输出较正确的映射。综上所述,在水质反演模型中,采用BP神经网络对水质反演是合适的。

2.2 BP神经网络结构的设计及训练方法的选择

由于3层结构的BP神经网络能够实现任意的非线性逼近,所以本文选择3层结构的BP神经网络,即输入层、隐层以及输出层。输入层的神经元数目由所使用的遥感图像波段数目和所采用的知识种类所确定;而输出层的数目则由待反演的具体水质参数确定。输入层和输出层的神经元数目较容易确定,关键是隐层的神经元数目的确定,由于目前还没有一个统一的指导原则来确定神经元数目,在实际应用中往往都是利用“试错法”来进行选择。另外,如何尽可能地提高BP神经网络的推广能力是一个关键问题。BP神经网络要具有较好的推广性能,必须经过大量样本的学习,然而在实际应用中,受到人力、物力等因素限制,地面水质监测点数量的设置总是有限的,这样导致用来训练的神经网络的训练集规模偏小,如按常规的训练方法训练BP神经网络,就会使得BP神经网络的推广能力不强,从而使得反演的水质准确性不高。

故在这种情况下对神经网络的训练要进行特殊的处理。在此,选择适合小样本训练的交叉训练方式^[2,7],即先把训练集(包括学习集和目标集)随机地分为若干组,其中每一组包括训练集和验证集,但是不包括测试集(受样本数的限制)。然后利用上述若干组的划分,分别对神经网络进行提前停止的训练,最后找出对于若干组的平均训练误差和平均验证误差最小的网络连接权值作为最后的神经网络结构。

2.3 遥感图像数据及与水质相关的知识数据

输入到BP神经网络的遥感图像波段要与水质

的光谱特性密切相关。待反演水域的具体地理信息知识,包括水域水质的地理分布的先验知识、水底的地理结构知识等。另外一类知识是对遥感图像上(通常是假彩色合成图像)的水域进行目视解译后再进行分类而获取的,在解译过程中,主要是利用判读标志,它是依据判读特征以及成像时间、季节、图像种类、比例尺等因素制作而成的^[8]。

在湖泊水质反演中,与水质相关的知识规则的制定要依据所采用的具体知识,另外为了降低计算量,规则的编码要尽量简单。

3 实验与分析

以中国的太湖为例,对实验的结果进行分析。实验所采用的地面监测数据为1997-05-04~1997-05-07的12个地面监测点的水质参数,这12个地面监测点分别分布在北部湖岸区、西部湖岸区、湖心区、东部湖岸区等区域,地面监测点的分布如图3所示,其中,★1~★12为地面监测点。



图3 太湖地面监测点的分布图

Fig. 3 Location of sampling stations of Tai Lake

由于太湖的污染主要体现在水体的富营养化方面,而叶绿素 α (Chl- α)浓度是反映湖泊水质富营养化的一个重要指标,又Chl- α 浓度影响水色、水质,与水域中初级生产力之间存在密切关系,所以这里选用的待反演水质参数为Chl- α 浓度。实验中使用的遥感图像数据为与水质监测同期的Landsat5卫星的TM图像数据,共7个波段。由于Chl- α 具有特定

的吸收和反射光谱,在 440nm 附近有一吸收峰,在 550nm 附近有一反射峰,在 680nm 附近有较明显的荧光峰,而这些特征的吸收峰和反射峰与 TM 图像的蓝波段、绿波段、红波段——对应^[9]。又通过对由 TM 图像的绿色波段、红色波段和近红外波段组成的假彩色合成图像的目视解译可以获取与水质相关的知识,所以本实验使用 TM 图像的蓝波段、绿波段、红波段和近红外波段,并把这 4 个波段分别记作 TM1、TM2、TM3、TM4。

3.1 TM 图像的预处理

预处理包括辐射校正、几何校正和大气校正,其中,主要是大气校正。一般可通过 3 种途径,即辐射传递方程式计算法、野外波谱测试回归法及多波段图像对比分析法进行大气校正^[10]。本实验采用多波段图像对比分析法中的直方图法,经过实验确定将 TM 图像的 TM1、TM2、TM3 这 3 个波段的灰度值(因为灰度值是地物辐射的平均值的度量)分别减去 17、11、8,得到经过大气校正的 TM 图像数据。另外为了消除噪声的影响,用 3×3 中值滤波器对经过大气校正过的 TM 图像进行滤波。把经过上述处理过后的 TM 图像的 TM1、TM2、TM3 这 3 个波段的像素值分别记作 B_1, B_2, B_3 。最后利用 MapInfo 7.0 来实现地面监测点与对应的 TM 图像间的坐标变换。

3.2 与水质相关的知识数据的预处理

采用两类知识,其中,第 1 类知识是地理信息知识,第 2 类知识是通过 TM 图像假彩色合成图像的解译而对太湖水质进行分类而获取的。依据这两类知识数据进行编码规则的制订。

(1)地理信息知识的编码规则的制订 考虑到太湖在每年的 5、6 月份盛行东南风,太湖的重污染水域主要是太湖的西北区域,比如梅梁湖区。又在湖岸区,由于人类活动如污水排放、围网养殖等会加剧富营养化的发展,所以湖岸区污染较重。另外,从湖泊形态方面考虑,由于太湖湖体的地理结构近似为一个“碟形”,湖心在整个太湖流域污染最轻。根据上述分析,依据水质的污染轻重将太湖的各个湖区划分成如图 4 所示的区域,并针对各个湖区的地理位置进行相应的污染程度编码,如表 1 所示。

(2)水质分类的编码规则的制订 利用 TM 图像的 TM2、TM3、TM4 波段的假彩色合成图像对太湖流域进行目视解译,依据先验知识,在假彩色图像上,水质污染较重的区域往往呈现黑色,而污染较轻的区域则色调偏浅,采用监督分类得到关于水质类



图 4 太湖的各个湖区
Fig.4 District of Tai Lake

表 1 根据太湖的地理信息对各个湖区的污染程度编码
Tab.1 Code of pollution degree based on geography information of Tai Lake

| 区域 | 东部湖岸区 | 南部湖岸区 | 北部湖岸区 | 西部湖岸区 | 湖心区 | 东太湖 |
|--------|-------|-------|-------|-------|-----|-----|
| 污染程度编码 | 1 | 1 | 3 | 2 | 0 | 2 |

别的分布如图 5 所示。其中,黑色表示污染最严重,浅黑色的区域表示污染较重,灰色的区域表示污染中等,浅灰色的区域表示污染最轻。



图 5 太湖(部分)的水质类别划分
Fig.5 Partition of water quality classification of Tai Lake

根据上述分析,对太湖水质类别进行编码,如表 2 所示。

表 2 依据假彩色图像对太湖水质类别编码

Tab.2 Code of water quality classification of Tai Lake based on pseudo color image

| 区域类别(颜色) | 黑色 | 浅黑色 | 灰色 | 浅灰色 |
|----------|----|-----|----|-----|
| 水质分类编码 | 6 | 5 | 4 | 3 |

3.3 BP 神经网络的训练

由上述分析可知,用于水质反演的神经网络的输入共有 5 个,这些输入又分为 3 种类型,即:(1)经过预处理的遥感图像数据 B1, B2, B3;(2)由地理信息知识获取的水质编码;(3)依据假彩色图像获取的水质编码。从 12 个地面监测点中任意选取 9 个监测点实测的 Chl- α 浓度值,再加上与监测点对应的遥感图像数据和水质相关的知识数据共同组成训练样本,余下的 3 个监测点的实测 Chl- α 浓度值及对应的遥感图像数据和水质相关的知识数据作为验证样本,这种划分总共可以有 $C_{12}^3 = 220$ 种,如图 6 所示。

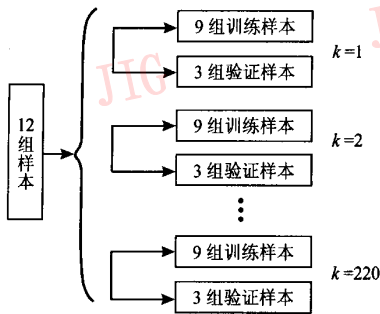


图 6 训练样本和验证样本的划分

Fig. 6 Partition of train samples and validation samples

另外为了加快神经网络的收敛速度以及使神经网络具有更好的拟合效果,需对神经网络的输入、输出矢量进行归一化处理。

针对 220 种划分中的某一种划分,先要确定 BP 神经网络的隐层神经元个数。记输入层神经元个数为 m 、隐层神经元个数为 n 、输出层神经元个数为 p 的 BP 神经网络为 $m \rightarrow n \rightarrow p$, 这里, $m = 5, p = 1$ 。为了不使神经网络结构过分复杂,又能使得神经网络具有较好的推广性能,先选择神经网络的隐层神经元个数 n 分别为 3、4、5, 然后通过对神经网络的训练性能进行比较,找出最优的神经网络结构。在隐层的神经元数目固定时,针对某一组划分,先选定若

干组任意的初始权值,分别对该组划分进行提前停止的训练。如果网络的训练误差较大,表明训练的次数不够;如果训练误差较小,而验证误差较大,表明训练过度。训练适度的神经网络应该是训练误差和验证误差接近,而且都较小。按照上述要求,找出若干个 BP 神经网络,都满足训练误差和验证误差接近而且较小,然后把 220 种划分全部代入这若干个 BP 神经网络,找出平均训练误差和平均验证误差均最小的网络作为最优的神经网络结构。当隐层神经元数目改变时,训练方式和上述类似。当隐层数目取不同值时,本模型的最优 BP 神经网络结构对样本的 220 种划分,得到的平均训练误差 (E_{tr})、平均验证误差 (E_{val})、网络的输出与实测的 Chl- α 浓度之间的相关系数 (R) 如表 3 所示。

表 3 隐层数目由 3 到 6 时的本文模型的最优 BP 神经网络对样本的 220 种划分的 E_{tr}, E_{val}, R

Tab.3 E_{tr}, E_{val} , and R of optimum BP neural networks which hidden layers is from 3 to 6 for 220 partition of samples in the developed model

| BP 神经网络结构 | E_{tr} | E_{val} | R |
|-----------|----------|-----------|---------|
| 5-3-1 | 14.523 4 | 14.646 2 | 0.997 1 |
| 5-4-1 | 15.001 0 | 15.540 5 | 0.996 9 |
| 5-5-1 | 14.733 1 | 15.377 4 | 0.997 1 |

最后,通过比较隐层神经元数目不同的神经网络,选取 E_{tr}, E_{val} 均最小而且 R 最大的 BP 神经网络结构。根据表 3,选取隐层神经元个数为 3 的 BP 神经网络。

3.4 太湖水质反演结果及分析

在 MATLAB6.5 软件环境下,通过试验确定最后的神经网络结构,也即得到 BP 神经网络各层之间的连接权值及阈值。然后,在遥感图像处理软件 ERDAS8.4 的环境下,利用该软件的空间建模工具,来实现神经网络的反演算法。之所以这样做是因为,空间建模工具是一个面向目标的模型语言环境,其中的空间建模语言是 Erdas 中所有 GIS 分析和图像处理功能的基础,功能非常强大。据此得到太湖(部分)的 Chl- α 浓度的反演结果如图 7 所示。

由图 7 可知,太湖流域的 Chl- α 浓度分布呈如下特点:

(1)西北湖岸区、北部湖岸区、东北湖岸区和东太湖 Chl- α 浓度最高,是严重的富营养化区域;

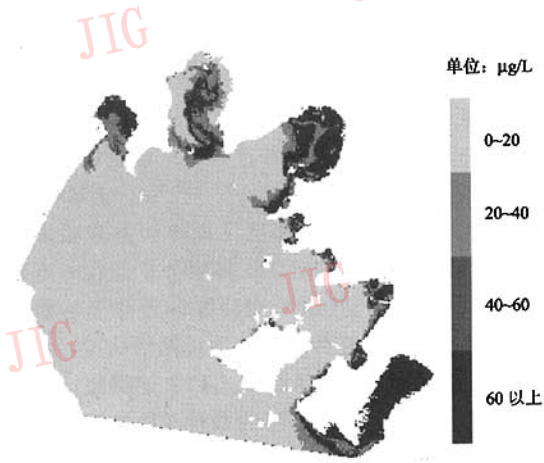


图 7 本文模型反演的太湖(部分)Chl- α 浓度分布图
Fig. 7 Chl- α concentration distribution of Tai Lake using the proposed model

(2)湖心区、南部湖区和西部湖区 Chl- α 浓度最低,但是这种最低是相对于整个太湖流域的;

(3)其他湖区的 Chl- α 浓度介于上述两种情况之间,属于较重的富营养化区域。

3.5 线性回归模型、传统的神经网络模型与新模型反演结果的比较

为了分析比较本文提出的新模型与线性回归模型及传统的神经网络模型之间的性能,从监测点的反演精度和非监测点的反演精度两个方面加以说明。

根据实验,得到 Chl- α 浓度的线性回归模型为

$$\text{Chl-}\alpha \text{ 浓度} = 85.4B_1/B_3 - 148 \quad (3)$$

为了便于比较,传统的神经网络模型中隐层神经元个数也选为 3,训练方法和本文的模型一样。

(1)监测点的反演精度

对于本文选择的 BP 神经网络,反映监测点的反演精度可以用输出向量和目标向量之间平方和误差(SSE)来表示。为了提高监测点的反演精度,在对 BP 神经网络训练时,就要使 SSE 最小。又 SSE 与误差向量及神经网络的权值和阈值参数有如下的关系^[11]:

$$SSE = s(E, X) \quad (4)$$

式中, s 是平方和误差函数, E 是误差向量, X 是由网络权值和阈值参数构成的矩阵。

可见 SSE 与神经网络的结构及训练方式有很大关系。所以为了保证反演的精度,选择适合的网络结构和采用适合的训练方式是很重要的。实验中,结合提前停止的训练方法,通过不断地权值、阈

值调整过程,可以使得监测点的反演精度得到提高。为了更加准确地描述监测点的反演精度,还引入了相关系数指标。表 5 给出了线性回归模型、神经网络模型与本文提出的新模型对 12 个地面监测点的 Chl- α 浓度的反演的比较。由表 5 可知,新模型反演的 Chl- α 浓度和实际测定的 Chl- α 浓度之间的相关系数最大(新模型为 0.997),而传统的神经网络模型反演的 Chl- α 浓度和实际测定的 Chl- α 浓度之间相关系数介于新模型和线性回归模型之间(神经网络模型为 0.996),线性回归模型反演的 Chl- α 浓度与实际测定的 Chl- α 浓度之间的相关系数最小(线性回归模型为 0.8600)。又新模型反演的 Chl- α 浓度与实测值的平方和误差(SSE)与其他二种模型相比最小,新模型的 SSE 为 44.8810,传统的神经网络模型为 62.2429,线性回归模型为 2002.00。从以上分析表明,从监测点本身的 Chl- α 浓度反演来说,新模型反演的精度最高。

表 5 3 种模型的 Chl- α 浓度反演数值的比较

Tab.5 Comparison of retrieval of Chl- α concentration for three models 单位: $\mu\text{g/L}$

| 地面监测点 编号 | Chl- α 浓度实测值 | 线性回归 模型 | 神经网络 模型 | 新模型 |
|-----------------------|------------------------|------------|------------|---------|
| 1 | 39 | 51.2 | 39 | 38.9 |
| 2 | 22 | 7.7 | 15.5 | 16.8 |
| 3 | 16 | 33 | 17 | 17.1 |
| 4 | 17 | 24.7 | 16.3 | 17.1 |
| 5 | 13 | 7.5 | 15.8 | 16.8 |
| 6 | 16 | 19.2 | 16.1 | 16.9 |
| 7 | 17 | 22.8 | 16.1 | 16.9 |
| 8 | 8 | 6.4 | 10.6 | 8.7 |
| 9 | 39 | 57 | 39 | 39 |
| 10 | 47 | 41.3 | 45.2 | 46.5 |
| 11 | 16 | 7.2 | 16.1 | 16.2 |
| 12 | 104 | 75.9 | 104 | 103.9 |
| 各模型反演的数值与 实测值的相关系数 | | 0.8600 | 0.996 | 0.997 |
| 各模型反演的数值与 实测值的 SSE | | 2002.00 | 62.2429 | 44.8810 |

(2)非监测点的 Chl- α 反演精度

线性回归模型和传统的神经网络模型都没有对非监测点的水质反演精度给予分析,而本文从“邻域判别原则”来对非监测点的水质反演精度进行分

析。所谓的“邻域判别原则”就是以某一监测点为中心的一个小区域中的每一个点反演的 Chl- α 浓度与该监测点的实测 Chl- α 浓度差的绝对值和 (SABS) 来判断模型的反演精度。判断的准则是:如果某一种模型的反演精度较高,那么以某一监测点为中心的一个小区域内的所有点反演的 Chl- α 浓度与该监测点实测的 Chl- α 浓度不应该相差太大,即 SABS 较小;如果模型的反演精度较小,则 SABS 较大。以 1 号监测点为例加以说明。在选择邻域的大小时,要注意的是邻域范围选择得不宜过大,因为根据经验,两点之间距离如果超过 1000m,那么这两点的水质就会有较大的变化。而对于 TM 图像,地面分辨率是 30m。综合以上因素,在实验时选择的邻域大小为以某一监测点为中心的 5×5 邻域。已知 1 号监测点实测的 Chl- α 浓度为 $39 \mu\text{g/L}$,线性回归模型、传统的神经网络模型和本文的新模型对以 1 号监测点为中心的 5×5 邻域进行反演而得到的 Chl- α 浓度(单位: $\mu\text{g/L}$)如图 8 所示。

| | | | | |
|------|------|------|------|------|
| 45.4 | 49.6 | 50.8 | 51.1 | 47.4 |
| 45.7 | 49.6 | 50.8 | 49.2 | 46.1 |
| 46.2 | 48.8 | 51.2 | 49.9 | 47.9 |
| 48.5 | 47.7 | 49.6 | 47.9 | 47.0 |
| 49.1 | 46.4 | 47.2 | 46.6 | 47.1 |

(a) 线性回归模型

| | | | | |
|------|------|------|------|------|
| 41.8 | 36.6 | 40.2 | 39.0 | 47.1 |
| 40.5 | 38.6 | 40.2 | 41.1 | 46.6 |
| 36.1 | 40.1 | 39.0 | 45.7 | 48.3 |
| 32.0 | 36.5 | 31.6 | 32.2 | 33.2 |
| 29.2 | 36.2 | 33.2 | 37.0 | 38.8 |

(b) 传统的神经网络模型

| | | | | |
|------|------|------|------|------|
| 38.6 | 38.6 | 38.7 | 38.6 | 38.7 |
| 38.6 | 38.5 | 38.7 | 38.9 | 39.0 |
| 38.6 | 38.7 | 38.9 | 39.0 | 39.0 |
| 38.9 | 38.9 | 38.8 | 38.8 | 38.7 |
| 39.1 | 39.3 | 39.2 | 39.0 | 38.7 |

(c) 新模型

图 8 以 1 号监测点为中心的 5×5 区域 3 种模型反演的 Chl- α 浓度(单位: $\mu\text{g/L}$)

Fig. 8 Chl- α Concentration of 5×5 neighbor of 1 sampling station by three models(unit: $\mu\text{g/L}$)

由图 8 可知,对于以 1#监测点为中心的 5×5 邻域中的 Chl- α 浓度反演,线性回归模型、传统的神经网络模型和本文的新模型的 SABS 分别为 $231.8 \mu\text{g/L}$ 、 $97.5 \mu\text{g/L}$ 、 $5.7 \mu\text{g/L}$ 。由此可见,新模型 SABS 最小(为 $5.7 \mu\text{g/L}$);其次是神经网络模型的 SABS ($97.5 \mu\text{g/L}$);最大的是线性回归模型的 SABS (为 $231.8 \mu\text{g/L}$)。根据太湖水质的实际情况和经验,1 号监测点周围小邻域内的水质反演值与实测值相差应该很小,而新模型反演结果的 SABS 最小,正符合这一实际情况,这表明,新模型在结合了地理信息知识和遥感图像解译知识以后,在非监测点也提高了水质反演的精度。

4 结 论

水质反演是一个不确定性问题,反演的结果在多大程度上是可靠的,一直是水质反演的研究难点。新模型的输入不仅仅有遥感图像的数据,而且有与水质有关的知识数据,即相当于在遥感图像的数据基础之上融合了与监测区域水质相关的知识,从而在一定程度上能够提高水质的反演精度,也就是说,水质反演的可靠性得到了提高。本文的工作仅仅是对 Chl- α 浓度进行反演,进一步的工作是将进行其他的水质参数反演研究,比如悬浮物、总磷、总氮等。另外,在水质反演中,还有几个问题有待进一步研究:(1)如何提高 TM 图像的大气校正精度,以使后续的遥感图像反演是在一个较可靠的基础之上进行;(2)如何更准确地实现地面实测点与遥感图像的坐标配准;(3)为了使反演的推广性能进一步提高,反演模型中应该尽量地增加地面监测点的数量,(并且监测点的布设要有一定的代表性),那么地面监测点的数量究竟是多少较合宜等。

参考文献 (References)

- 1 Lin Zhi-gui, Xu Li-zhong, Huang Feng-chen. The method of data fusion based on water quality remote sensing image and land monitoring[A]. In: The Chinese Conference 2004 of Electronic & Information Technology[C]. Beijing, 2004;47 ~ 55. [林志贵,徐立中,黄凤辰. 水质遥感图像与地面监测数据融合处理方法[A]. 见:中国电子与信息工程学术年会论文集[C],北京,2004;47 ~ 55.]
- 2 Louis K E, Yan X H. A neural network model for estimating sea surface chlorophyll and sediments from Thematic Mapper imagery [J]. Remote Sensing of Environment, 1998, 66(2):153 ~ 165.
- 3 Keiner L E, Brown C W. Estimating oceanic chlorophyll

- concentrations with neural networks [J]. International Journal of Remote Sensing, 1999, 20(1):183~188.
- 4 Chen L. A study of applying genetic programming to reservoir trophic state evaluation using remote sensor data[J]. International Journal of Remote Sensing, 2003, 24(11):2265~2275.
- 5 Zhang Y Z, Pulliainen J T, Koponen S S. Water quality retrievals from combined landsat TM data and RES-2 SAR data in the gulf of Finland[J]. IEEE Transactions on Geoscience and Remote Sensing, 2003, 41(3):622~629.
- 6 Zhan Hai-gang. Inversion of ocean color based on soft computing [D]. Qingdao:Qingdao Ocean University, 2001. [詹海刚. 基于软计算的海洋水色遥感反演[D]. 青岛:青岛海洋大学,2001.]
- 7 Yan A X, Zhang R S, Liu M C. Large artificial neural networks applied to the prediction of retention indicies of acyclic and cyclic alkanes, alkenes, alcoholos, esters, ketones, and ethers [J]. Computers and Chemistry, 1998, 22(5):405~412.
- 8 Zhu S L, Zhang Z M. Acquisition and Analysis of Remotely Sensed Image[M]. Beijing: Science Press, 2003:103. [朱述龙,张占睦著. 遥感图像获取与分析[M]. 北京:科学出版社,2003:57~60.]
- 9 Chen Chu-qun, Shi Ping, Mao Qing-wen. Study on modeling chlorophyll concentration of surface coastal water using TM data[J]. Remote Sensing of Environment China, 1996, 11(3):168~176. [陈楚群,施平,毛庆文. 应用 TM 数据估算沿岸海水表层叶绿素浓度模型研究[J]. 环境遥感,1996,11(3):168~176.]
- 10 Zhang Xiao-can, Huang Zhi-cai, Zhao Yuan-hong. Remote Sensing Digital Image Processing[M]. Hangzhou: Zhejiang University Press, 2003:57~60. [章孝灿,黄智才,赵元洪著. 遥感数字图像处理[M]. 杭州:浙江大学出版社,2003:57~60.]
- 11 Xu Dong, Wu Zheng. System Analysis and Design Based on Matlab6.x—Neural Networks [M]. Xi'an: Xidian University Press, 2002:23~24. [许东,吴铮著. 基于 MATLAB6.x 的系统分析与设计——神经网络[M]. 西安:西安电子科技大学出版社,2002:23~24.]