

引用格式: Wu Huanhuan, Guo Qiaozhen, Zang Jinlong, *et al.* Study on Water Quality Parameter Inversion based on Landsat 8 and Measured Data[J]. Remote Sensing Technology and Application, 2021, 36(4): 898-907. [吴欢欢, 国巧真, 臧金龙, 等. 基于 Landsat 8 与实测数据的水质参数反演研究[J]. 遥感技术与应用, 2021, 36(4): 898-907.]
doi:10.11873/j.issn.1004-0323.2021.4.0898

基于 Landsat 8 与实测数据的水质参数反演研究

吴欢欢, 国巧真, 臧金龙, 乔悦, 朱丽, 何云海

(天津城建大学地质与测绘学院, 天津 300384)

摘要: 目前遥感技术已成为监测水质参数的重要手段, 精度更高的水质参数反演模型是当前水质监测的重点。但由于水环境的复杂性、遥感数据的局限性等多重原因, 水质参数遥感反演精度有限, 且多集中于水色水质参数反演。为了得到精度更高的水质参数反演模型, 以天津市海河下游段为研究区, 对 Landsat 8 OLI 遥感影像进行大气校正、辐射定标等预处理, 通过实验室理化分析测定水体的总磷、氨氮、总氮浓度及电导率, 建立实测水质参数与 Landsat 8 OLI 遥感影像数据的统计回归模型及神经网络模型, 采用决定系数(R^2)、平均绝对误差(MAE)、均方根误差(RMSE)进行精度检验, 神经网络模型反演结果 R^2 均大于 0.85, MAE 分别为 0.019、0.09、0.242、0.411, RMSE 分别为 0.024、0.118、0.286、0.562, 反演精度较好。结果表明: 基于神经网络建立的水质参数反演模型精度较高。

关键词: Landsat 8 OLI 遥感影像; 神经网络模型; 水质参数; 反演

中图分类号: TP79 **文献标志码:** A **文章编号:** 1004-0323(2021)04-0898-10

1 引言

遥感技术以其在时间、空间上的优势, 在一定程度上克服了传统水质监测中存在的问题。目前国内外学者在水质参数遥感反演方面已取得了诸多成就, 黄李童^[1]以杭州湾区域为研究区, 研究了总悬浮物和有色容溶解有机物的反演方法; 浦玲伟^[2]以太湖为研究区, 研究了叶绿素 a 的反演方法; 马方凯等^[3]以武汉汤逊湖为研究区, 研究了氨氮、总磷和化学需氧量三种水质参数的反演模型; Song 等^[4]以美国印第安纳州中部和南澳大利亚州的供水水源为研究区, 研究了饮用水源中的藻蓝蛋白的含量; 殷子瑶等^[5]以天津市于桥水库为研究区, 基于高光谱遥感数据研究了水体透明度和水体中悬浮物的含量。

水质参数反演研究主要集中于悬浮物、黄色物

质、叶绿素 a 等水色水质参数的反演, 总磷、总氮、氨氮和电导率等非水色水质参数同样是衡量水质的重要指标。水体中氮磷物质浓度过高将直接导致水体富营养化, 而富营养化是目前大部分水环境的主要问题, 总磷、氨氮含量是评价水体污染情况的主要标准。国家排放标准中一般会规定氨氮含量的指标, 但在一些实际应用中, 总氮含量是影响处理的关键, 所以测得水体的总氮含量对实际问题的解决有一定的指导意义。水的电导率能有效地反映出水中存在的电解质的程度, 能一定程度上反映其他水质参数含量, 也是评价水质优劣的重要指标参考。孙骄阳^[6]以密云水库为研究区, 研究了总氮、总磷、氨氮和 COD 4 个水质参数在多种算法下的反演模型; 王珊^[7]以潘家口水库和大黑汀水库为研究区, 研究了浊度、总悬浮物浓度、叶绿素 a 浓度、DO、COD 和 TP 等多种水质参数的反演方法。马方凯

收稿日期: 2020-04-25; 修订日期: 2021-06-14

基金项目: 天津市自然科学基金项目“天津滨海新区地表水环境遥感监测与生态风险评价”(18JCYBJC90900), 天津市教委科研项目“遥感技术视角下的天津市地表温度研究”(2018KJ164)资助。

作者简介: 吴欢欢(1994—), 女, 山西临汾人, 硕士研究生, 主要从事遥感反演方面的研究。E-mail: 18822085914@163.com

通讯作者: 国巧真(1979—), 女, 河北武邑人, 副教授, 博士, 主要从事资源与环境遥感监测研究。E-mail: gqiaozhen@tcu.edu.cn

等^[8]以武汉汤逊湖为研究区,基于高分辨率遥感影像数据研究了氨氮、总磷和化学需氧量的反演方法,为当地生态环境保护研究和相关决策的制定提供了理论依据。李怡静等^[9]研究了鄱阳湖的总氮、总磷等6种水质参数的遥感反演方法。虽然已有越来越多关于非水色的水质参数的研究,但由于水环境的复杂性、地域性等特点,使得水质参数反演方法一直难以克服其在普适性上的问题。

水质参数反演方法主要包括经验法^[10]、半经验法^[11]和分析法^[12]。目前已有许多算法应用于水质参数的遥感反演研究^[13-14]。基于遥感数据与实测数据的地表水环境参数反演的相关研究,建立了统计回归、神经网络等水质参数反演模型。人工神经网络^[15-19]具有分布式联想能力、自学习能力及自组织能力,常用来做数据关系的挖掘、构建预测模型等工作,近年来在水质参数反演及水质评价领域被广泛使用。国内外学者在神经网络应用于水环境方面的研究已有一定进展:杨国范等^[20]建立了铁岭市清河水库的叶绿素a浓度的比值线性回归模型与最小二乘支持向量机两种反演模型,结果显示最小二乘支持向量机模型精度更高。潘俊等^[21]建立了西鞍山铁矿地下水多项水质参数的BP神经网络模型,结果显示人工神经网络模型有效降低了主观误差。郭庆春等^[22]建立了长江的多项水质参数的人工神经网络改进算法模型,结果显示经合理的算法

改进,神经网络模型对于水质参数的反演具有更强的适用性。Li等^[23]研究了BP神经网络在水质评价中的应用,反演结果精度高,为水质评价提供了更好的方法。郑炎辉等^[24]研究了多光谱遥感影像水质参数参数定量反演模型,为水库的水质和富营养化监测提供了参考。

实验以海河下游段为研究区,利用Landsat 8 OLI遥感数据对水体的总磷、氨氮、总氮含量和电导率进行定量反演研究,分析波段组合和水质参数浓度的相关性,建立统计回归模型和神经网络模型,选择精度较高的反演模型进行反演,并对反演结果进行分析,为海河下游段的水质参数反演提供理论基础。

2 研究区与数据

2.1 研究区概况

研究区为海河流域下游段,位居天津市东部,如图1所示。所属华北地区最大的海河水系,全长2 458.4 km,流域总面积32.06万km²,占全国总面积的3.3%。海河自上游发起,汇聚北运河、永定河、大清河、子牙河、南运河五大支流,于天津大沽口注入渤海湾。海河流域下游位于海河平原,属于温带季风型大陆性气候。由于流域内越来越多的人类活动,导致中下游地区水量锐减,地表水资源匮乏,并伴随有严重的水质污染问题。

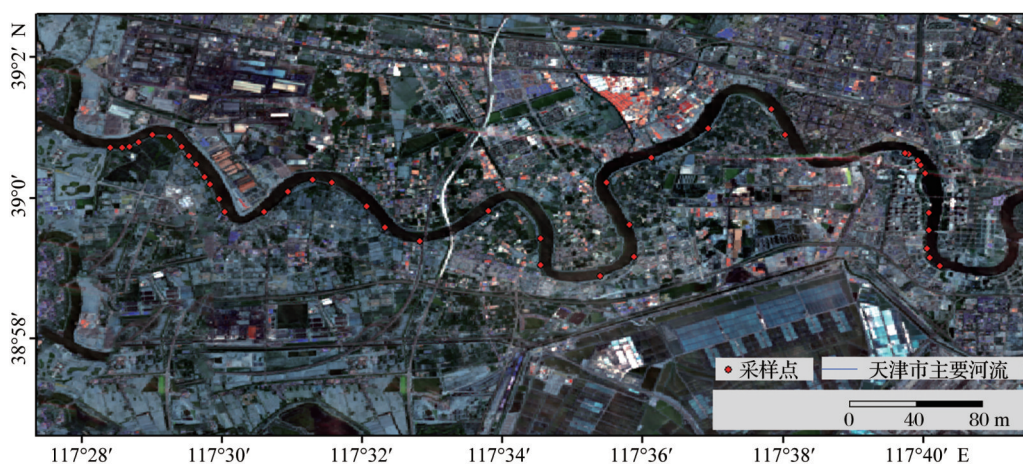


图1 采样点地理位置

Fig.1 Geographic location of sampling points

2.2 实测数据的采集与处理

选取的遥感影像为2019年6月23日的Landsat 8卫星影像数据,空间分辨率为30 m,传感器为OLI陆地成像仪,影像成像质量较好且覆盖海河水域的研究区。经对遥感影像进行裁剪、辐射定标和大气

校正等处理后,分别提取每个采样点对应的Landsat 8 OLI数据的前7个波段的反射率。同时进行了野外同步实验,采集了41个水样,采样点如图1所示。采样时用手持GPS记录每个样点的经纬度,将水样密封保存后当天送到实验室进行水质参数

理化分析。本实验使用的现场实测数据及其对应的遥感影像反射率数据共 41 组,去除一组异常点,选取其中 30 组数据用于建立反演模型,另外 10 组用于模型精度的检验。

2.3 Landsat 8 OLI 数据预处理

获取与实测数据同步的 2019 年 6 月 23 日的 Landsat 8 OLI 遥感数据,研究区影像质量良好。对该影像分别进行了影像裁剪、辐射定标、大气校正的预处理。辐射定标根据公式(1)将遥感影像的灰度值转换成辐射亮度,采用 FLAASH 大气校正工具对经过辐射定标后的遥感影像进行大气校正,自动获取中心点经纬度,选择对应传感器获取传感器高度及影像数据的分辨率,研究区地面高程数据采用天津市平均高程 3 m,影像生成时的飞行过境时间通过元文件获取,根据成像时间和纬度信息大气模型选择 Mid-Latitude Summer,气溶胶模型选择 Urban,气溶胶反演方法选择 2-Band(K-T),多光谱参数设置中波谱响应函数指向 Landsat 8 oli.sli,高级参数中内存大小设置为 100,其余参数默认,完成影像数据的大气校正。经校正的影像去除了大气散射和吸收引起的大气衰减的影响,绿地、水体等地物的波谱曲线趋于正常。

$$\text{Radiance} = \text{gain} * \text{DN} + \text{offset} \quad (1)$$

其中:gain 为各波段增益,offset 为偏差,均可从头文件中获取,DN 为各波段的像元值。

2.4 精度评价

为了验证水质参数反演结果的精度,使用平均绝对误差(MAE)、均方根误差(RMSE)对反演结果进行验证,计算方法如下:

$$\text{MAE} = \frac{\sum_{i=1}^n |x - x'|}{n} \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x - x')^2}{n}} \quad (3)$$

其中:x 为实测水质参数值;x'为模型反演所得水质参数值;n 为样本点数。

3 水质参数反演

水体中水质参数的种类及含量差别,表现在遥感影像上会有不同的反射率,且在各波段上的影响也不尽相同,所以说选择对水体中水质参数浓度变化敏感的波段或者波段组合是遥感技术应用于水质参数浓度反演的主要原理。分析水质参数浓度与遥感反射率之间的相关性,可以得到对于水质参

数浓度变化敏感的波段或波段组合,实验分析了各个水质参数与 OLI 数据的单波段及波段组合之间的相关性,其中波段组合涉及到一些水体指数的运用,具体公式如表 1 所示。选出与水质参数浓度相关性较高的波段或波段组合建立相应的统计回归模型和神经网络模型。

表 1 水体指数公式^[25]

Table 1 Water body index formulae

水体指数	公式	水体指数	公式
归一化 水体指数	$\text{NDWI} = \frac{B_3 - B_5}{B_3 + B_5}$	增强型 水体指数	$\text{EWI} = \frac{B_3 - B_5 - B_6}{B_3 + B_5 + B_6}$
改进归一化 水体指数	$\text{MNDWI} = \frac{B_3 - B_6}{B_3 + B_6}$	新型水 体指数	$\text{NWI} = \frac{B_1 - (B_5 + B_6 + B_7)}{B_1 + B_5 + B_6 + B_7}$

3.1 统计回归模型

分析实测水质参数与 Landsat 8 OLI 遥感影像单波段及波段组合的相关性,选取相关性较好的数据组构建统计回归模型。对比分析线性、二次多项式、三次多项式、对数等算法的拟合效果,列出不同因子的模型中所有 R^2 较大的函数关系式,选取 R^2 较高的单波段或波段组合的拟合模型,利用检验用的 10 个样本数据对这些模型进行检验,比较模型间的平均绝对误差和相对误差,最终根据 R 、 R^2 以及误差大小,选择最好的统计回归模型。

分析总磷、氨氮、总氮、电导率与波段及波段组合间的相关性,选取相关性较高的波段组合。表 2 所示与水体中总磷浓度相关性较高的波段组合为 B_1/B_7 、 B_3/B_6 、 B_3/B_7 、EWI,表 3 所示与水体中氨氮浓度相关性较高的波段组合为 B_6 、 B_7 、 $B_5 + B_6$ 、 $B_6 + B_7$ 、 $B_5 + B_7$,表 4 所示与水体中总氮浓度相关性较高的波段组合为 $B_1 - B_5$ 、 $B_3 - B_5$ 、 $B_4 - B_5$ 、 $B_4 / (B_5 + B_6)$ 、EWI,表 5 所示与水体中电导率相关性较高的波段组合为与 NDWI、EWI、NWI、 B_5/B_3 ,这些与水质参数相关性较高的遥感影像单波段及波段组合可作为反演各水质参数建立统计回归模型的因子。

根据选定作为构建模型的波段组合因子,建立其与水质参数之间的统计回归模型。如表 6 所示

表 2 总磷与波段组合间的相关性指数

Table 2 Correlation index between total phosphorus and band

因子	相关性	因子	相关性
B_1/B_7	0.672	B_3/B_5	-0.581
B_2/B_7	0.649	B_3/B_4	0.645
B_3/B_7	0.705	NDWI	0.590
B_3/B_6	0.691	EWI	0.663

表 3 氨氮与波段组合间的相关性指数

Table 3 Correlation index between ammonia nitrogen and band combinations

因子	相关性	因子	相关性
B ₆	0.773	B ₆ +B ₇	0.781
B ₇	0.789	B ₅ +B ₇	0.779
B ₅ +B ₆	0.774		

表 4 总氮与波段组合间的相关性指数

Table 4 Correlation index between total nitrogen and band combinations

因子	相关性	因子	相关性
B ₁ -B ₅	0.654	B ₄ /(B ₅ +B ₆)	0.672
B ₄ -B ₅	0.661	EWI	0.621
B ₃ -B ₅	0.705		

表 5 电导率与波段组合间的相关性指数

Table 5 Correlation index between conductivity and band combinations

因子	相关性	因子	相关性
NDWI	0.684	NWI	-0.634
EWI	0.751	B ₅ /B ₃	-0.677

B₃/B₆作为反演总磷浓度的因子的三次多项式模型的 R^2 最好,达到了0.703,用10个检验站点数据对各因子的最好模型进行检验,比较模型的平均绝对误差和相对误差,最终根据 R 、 R^2 以及误差大小,选择B₃/B₆为因子的三次多项式模型作为统计回归模型。如表7所示,B₇作为反演氨氮浓度的因子的线性模型的 R^2 最好,达到0.622,经检验后进行对比分析,选择B₇为因子的线性模型作为统计回归模型。如表8所示,B₃-B₅作为反演总氮浓度的因子的三次多项式模型的 R^2 最好,达到0.542,经检验后进行对比分析,选择B₃-B₅为因子的三次多项式模型作为统计回归模型。如表9所示,EWI作为反演电导率的因子的三次多项式模型的 R^2 最好,达到0.592,经检验后进行对比分析,选择EWI为因子的三次多项式模型作为统计回归模型。

3.2 神经网络模型

利用实测水质参数与 Landsat 8 OLI 遥感影像数据,在 Matlab 上构建包含输入层、隐含层、输出层的 3 层 BP 神经网络模型。输入层为根据遥感影像提取的遥感反射率的单波段及波段组合数据,输出层为水质参数浓度数据,通过实验比较分析设置隐含层的神经元个数,不同的隐含层神经元个数将直接影响模型的精度。选择经过训练后,各组训练结果趋于稳定且满足误差要求,误差性能变化、训练情

表 6 总磷浓度的拟合模型

Table 6 Fitting model of total phosphorus concentration

因子	模型	R^2	函数关系式	MAD
EWI	三次	0.595	$Y=0.232+0.604X-0.141X^2-6.431X^3$	0.047
B ₁ /B ₇	三次	0.588	$Y=0.695-0.660X+0.254X^2-0.028X^3$	0.049
B ₃ /B ₆	三次	0.703	$Y=0.958-1.508X+0.885X^2-0.154X^3$	0.066
B ₃ /B ₇	三次	0.681	$Y=1.084-1.571X+0.838X^2-0.134X^3$	0.167

表 7 氨氮浓度的拟合模型

Table 7 Fitting model of ammonia nitrogen concentration

因子	模型	R^2	函数关系式	MAD
B ₆	线性	0.598	$Y=0.289+0.002X$	0.339
	对数	0.588	$Y=-5.349+1.079\ln(X)$	0.360
B ₇	线性	0.622	$Y=0.139+0.003X$	0.246
	对数	0.603	$Y=-5.992+1.2\ln(X)$	0.336
B ₅ +B ₆	线性	0.600	$Y=0.251+0.001X$	0.410
	对数	0.597	$Y=-6.237+1.116\ln(X)$	0.343
B ₅ +B ₇	线性	0.606	$Y=0.180+0.001X$	0.491
	对数	0.597	$Y=-6.587+1.176\ln(X)$	0.330
B ₆ +B ₇	线性	0.611	$Y=0.219+0.001X$	0.404
	对数	0.597	$Y=-6.450+1.137\ln(X)$	0.345

表 8 总氮浓度的拟合模型

Table 8 Fitting model of total nitrogen concentration

因子	模型	R^2	函数关系式	MAD
EWI	三次	0.519	$Y=3.425+5.440X-11.517X^2-66.953X^3$	0.409
B ₃ -B ₅	三次	0.542	$Y=4.254-0.018X+5.425E-5X^2-4.228E-8X^3$	0.331

表 9 电导率的拟合模型

Table 9 Fitting model of conductivity

因子	模型	R^2	函数关系式	MAD
EWI	三次	0.592	$Y=11.753+14.843X-10.176X^2-95.526X^3$	1.077
	线性	0.540	$Y=12.016+10.157X$	1.156
NWI	三次	0.572	$Y=12.182+12.532X-13.980X^2-60.690X^3$	0.986

形及线性回归拟合情况较好。总磷、氨氮、总氮和电导率反演浓度各部分样本的拟合情况如图2所示。

将若干组相关性较高的数据作为各个水质参数构建神经网络的输入层,实测水质参数值作为输出层,选择合适的隐含层神经元个数的BP神经网络模型。经分析单波段和波段组合与实测水质参

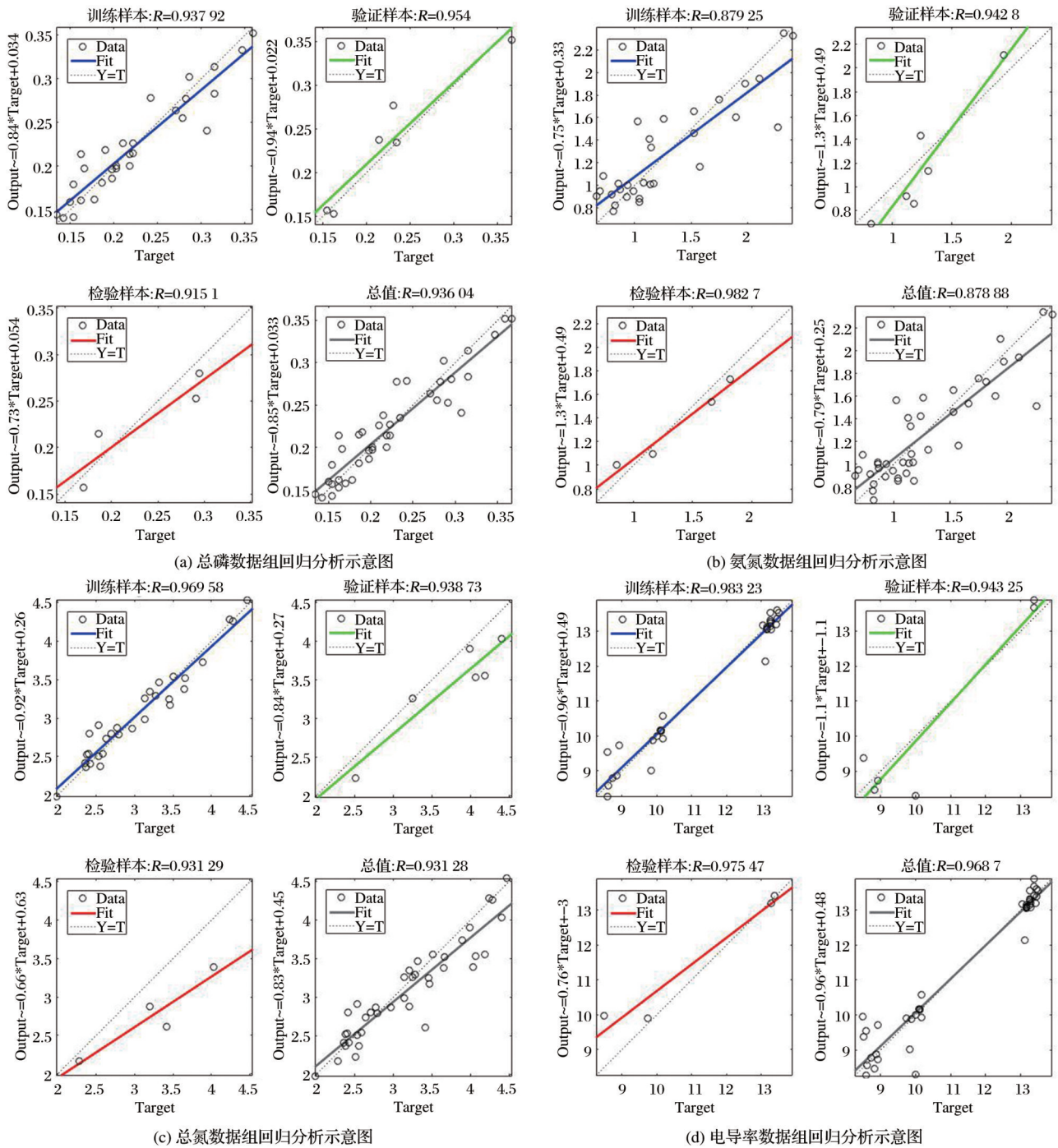


图2 各水质参数网络回归分析示意图

Fig.2 Schematic diagram of network regression analysis of various water quality parameters

数的相关性,得出 B_1/B_7 、 B_3/B_7 、 B_3/B_4 、 B_3/B_6 、 B_5/B_3 、NDWI、EWI 7组数据与实测总磷浓度相关性较高,且隐含层神经元个数为8时模型拟合效果最好,如图2(a)所示,各组样本拟合度均在0.9以上,模型拟合度很好。 B_6 、 B_7 、 B_1/B_5 、 B_1/B_6 、 B_1/B_7 、 B_5+B_6 、 B_5+B_7 、 B_6+B_7 、NWI 9组数据与实测氨氮浓度相关性较高,且在神经元个数为9时模型拟合程度最好,如

图2(b)所示,各组样本拟合度均在0.87以上,模型拟合度良好。 $B_4/(B_5+B_6)$ 、 $B_4/(B_5+B_7)$ 、 B_3/B_5 、 B_1-B_5 、 B_2-B_5 、 B_3-B_5 、 B_4-B_5 、NDWI、EWI 9组数据与实测总氮浓度相关性较高,且在神经元个数为14时模型拟合程度最好,如图2(c)所示,各组样本拟合度均在0.9以上,模型拟合度很好。NDWI、EWI、NWI、 B_3/B_5 4组数据与实测总氮浓度相关性较高且

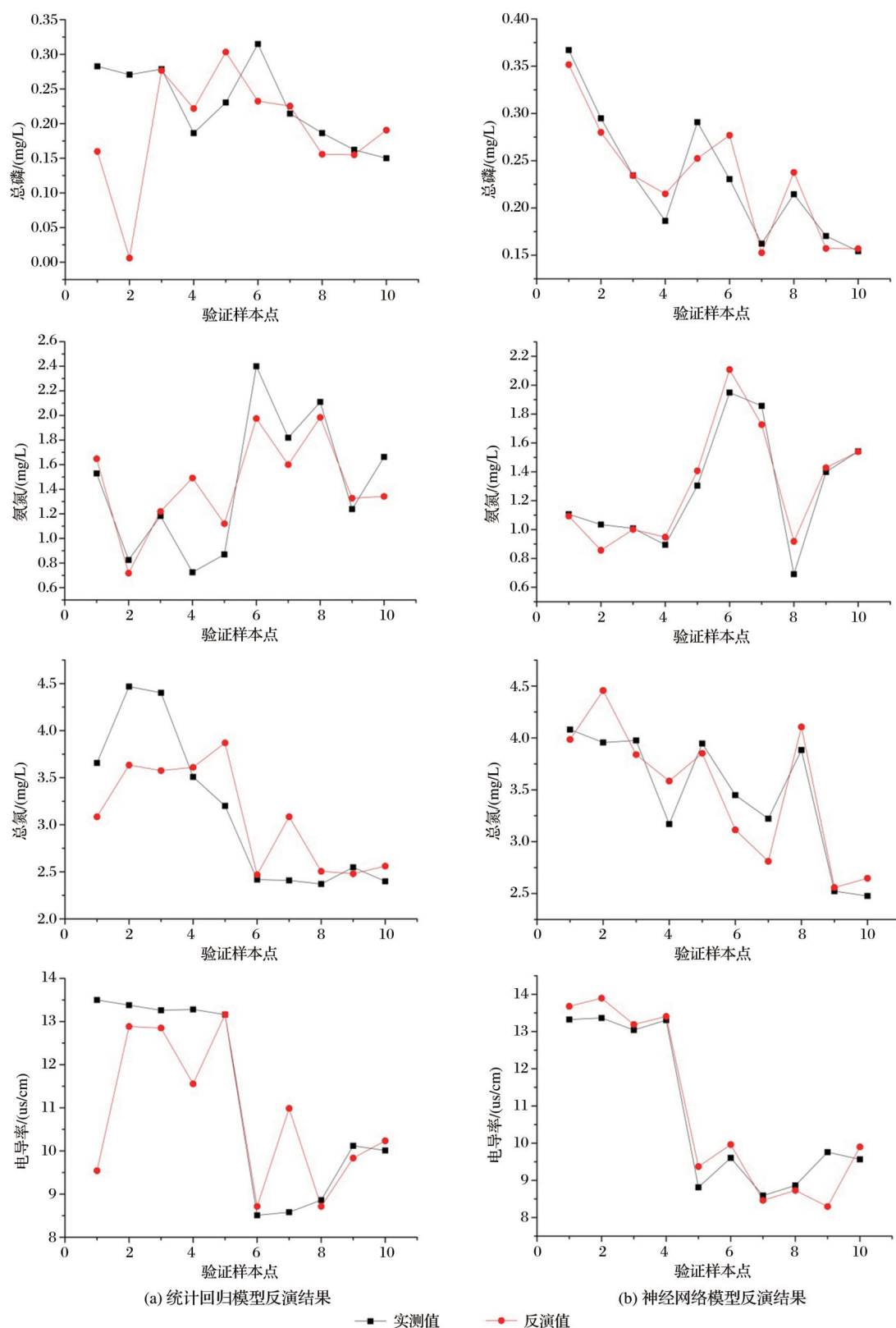


图3 水体总磷、氨氮、总氮浓度和电导率统计回归模型与神经网络反演值与实测值对比图

Fig.3 Comparison of inversion and measured values of statistical regression model and neural network of total phosphorus, ammonia nitrogen, total nitrogen concentration and conductivity in water

在神经元个数为8时模型拟合程度最好,如图2(d)所示,各组样本拟合度均在0.9以上,模型拟合度很

好。整体来说,4个水质参数的3组样本与整体数据的拟合值均为0.85以上,大部分集中在0.9以上,拟

合效果较好。

4 结果与分析

4.1 模型对比

实验分别采用统计回归和BP人工神经网络两种方法构建了总磷、氨氮、总氮浓度和电导率的遥感反演模型。利用检验样本点对所建立的关于总磷、氨氮、总氮浓度和电导率的统计回归和神经网络遥感反演模型进行精度检验与比较。

由图3验证样本点的实测值与统计回归模型、神经网络模型的反演结果的对比,可清楚地比较出两种反演模型对于研究区水质参数的反演效果。由图3可看出,统计回归建立的反演模型精度较低,存在较多与实测值相差很大的反演结果,表现在图像上为实测值与反演值两条折线的拟合度不高,出现折点的跳跃,而神经网络建立的反演模型精度较高,实测值折线与反演值折线拟合度较高,说明神经网络所建立的水质反演模型精度更高。

表10列出4种水质参数的统计回归模型和BP神经网络模型的反演精度检验结果。总磷选择波段组合 B_3/B_6 建立统计回归模型,选择7个波段组合作为BP神经网络的输入层,BP神经网络的均方根误差为0.024。氨氮选择波段组合 B_7 建立统计回归模型,选择9个波段组合作为BP神经网络的输入层,BP神经网络的均方根误差为0.118。总氮选择波段组合 B_3-B_5 建立统计回归模型,选择9个波段组合作为BP神经网络的输入层,BP神经网络的均方根误差为0.286。电导率选择指数EWI建立统计回归模型,选择4个波段组合作为BP神经网络的输入层,BP神经网络的均方根误差为0.562。从表10中可以清楚地看出,BP神经网络构建的水质参数反演模型精度明显高于统计回归反演模型。

表 10 统计回归模型与BP神经网络模型反演结果对比
Table 10 Comparison of inversion results between statistical regression model and BP neural network model

水质参数	统计回归模型		BP神经网络模型	
	MAD	RMSE	MAD	RMSE
总磷	0.066	0.101	0.019	0.024
氨氮	0.246	0.321	0.09	0.118
总氮	0.409	0.517	0.242	0.286
电导率	0.986	1.583	0.411	0.562

4.2 水质参数分布图

基于Landsat 8 OLI数据与实测水质参数构建关于水质参数总磷、氨氮、总氮浓度和电导率的统

计回归和神经网络反演模型,经试验,其反演结果精度高于统计回归模型。截取面积为 2.7 km^2 的研究区遥感影像,分别反演区域内水体的4个水质参数,反演结果如图4所示。可以看出所选水域总磷浓度在 0.5 mg/L 左右,其中后半段水域总磷浓度有小幅升高,氨氮浓度普遍在 $3.5\sim 4\text{ mg/L}$,前三分之一段水域的总氮浓度约 5 mg/L ,随后浓度增加到 $6\sim 7\text{ mg/L}$,水域内电导率为 $11\text{ }\mu\text{s/cm}$ 左右。参考《地面水环境质量标准》(GB3838-83),所选区域的氨氮和总氮含量很高,总磷含量虽相对较低,电导率趋于正常范围,但总的来说该段水质较差,存在一定的问题。

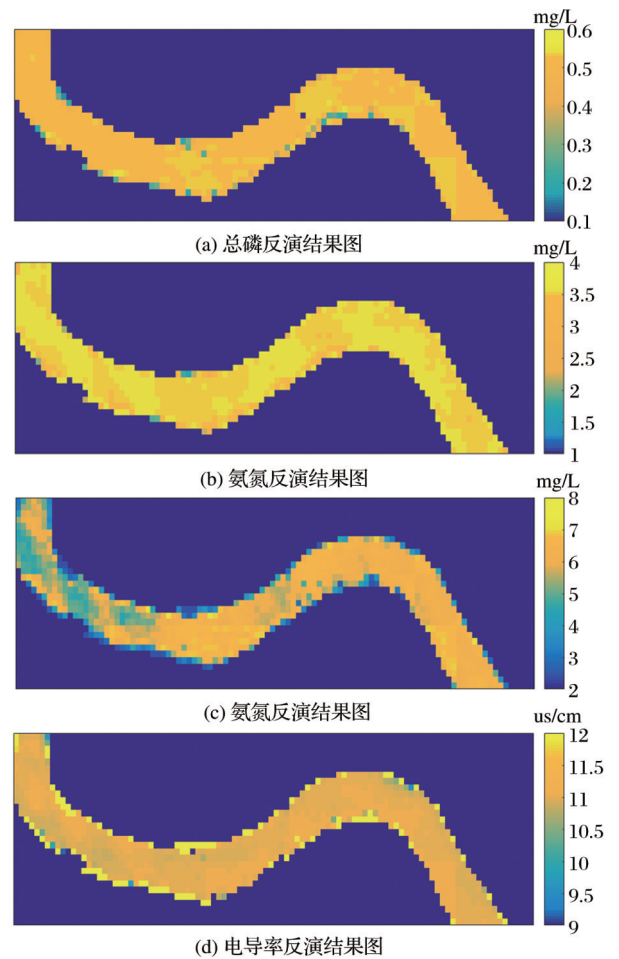


图 4 总磷、氨氮、总氮浓度及电导率神经网络反演结果
Fig.4 Neural network inversion results of total phosphorus, ammonia nitrogen, total nitrogen concentration and conductivity

5 结 论

实验以海河下游段为研究区,分析了总磷、氨氮、总氮浓度及电导率4个水质参数与Landsat 8

OLI数据的单波段及波段组合的相关性,建立了统计回归模型和BP神经网络模型,并对比了两种方法构建的反演水质参数模型的结果。研究区水体的4个水质参数与Landsat 8 OLI数据的单波段反射率,与波段反射率加减法及比值,与相关水体指数NDWI、MNDWI、EWI、NWI密切相关,相关性均大于0.6。基于相关性分析的结果,建立水质参数与遥感数据的统计回归及BP神经网络反演模型。所建统计回归模型的拟合度最高为0.703,拟合效果相对较差,尤其体现在总氮的反演过程中,拟合度为0.542。所建的BP神经网络模型的拟合度明显提高,4个水质参数的拟合值均大于0.93。经检验,所建的4个水质参数的BP神经网络模型相较于传统统计回归模型,在MAD、RMSE上均得到不同程度的提高。在此基础上,截取一定范围的水域,运用所建立的神经网络反演模型反演了该水域的总磷、氨氮、总氮浓度及电导率。结果表明,遥感数据反演水质参数是非线性的,统计回归的方法并不适用于水质参数的反演,神经网络模型具有解决复杂非线性映射的能力,使用BP神经网络模型较好地拟合了水质参数与光谱反射率之间的非线性函数,证明利用BP神经网络模型来进行水质遥感反演研究是适用的,且反演结果精度较高。

6 讨 论

地表水环境复杂,且各水质参数相互影响,所以水质参数反演模型的构建需要考虑多因素、信息处理不精确等问题。针对这种情况,神经网络的强大非线性拟合能力可以很大程度上提高水质参数遥感反演模型的精度。目前神经网络算法在水质参数反演研究中应用逐渐广泛,但也有两个问题在研究中不容忽视。一方面,神经网络模型训练到一定程度可能会出现过拟合,如何有效地利用其他算法对其进行优化。另一方面,今后的研究中考虑在已有条件的基础上加大样本量进行进一步研究。

参考文献(References):

- [1] Huang Litong. Retrieval and application of remote sensing water quality of complicated water body in coastal zone: a case study of the Hangzhou Bay area[D]. Hangzhou: Zhejiang University, 2019.[黄李童. 海岸带复杂水体水质遥感反演与应用:以杭州湾区域为例[D]. 杭州:浙江大学, 2019.]
- [2] Pu Lingwei. Remote sensing inversion of Chlorophyll-a concentration using spectral smoothing method to improve the validation precision of the Chlorophyll-a estimation models in turbidity waters[D]. Suzhou: Suzhou University of Science and Technology, 2019.[浦玲伟. 基于高分影像的太湖东岸带叶绿素a浓度遥感反演研究[D]. 苏州: 苏州科技大学, 2019.]
- [3] Zhao Song. Inversion of water quality parameters of Fuyang River in Handan City based on multi-source remote sensing data[D]. Handan: Hebei University of Engineering, 2019.[赵松. 基于多源遥感数据的邯郸市滏阳河水水质参数反演[D]. 邯郸: 河北工程大学, 2021.]
- [4] Song K S, Li L, Lenore P Tedesco, *et al.* Remote quantification of phycocyanin in potable water sources through an adaptive model[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 95: 68-80. DOI: 10.1016/j.isprsjprs.2014.06.008.
- [5] Yin Ziyao, Li Junsheng, Fan Haisheng, *et al.* Preliminary study on water quality parameter inversion for the Yuqiao Reservoir based on Zhuhai-1 hyperspectral satellite data[J]. Spectroscopy and Spectral Analysis, 2021, 41(2): 494-498.[殷子瑶, 李俊生, 范海生, 等. 珠海一号高光谱卫星的于桥水库水质参数反演初步研究[J]. 光谱学与光谱分析, 2021, 41(2): 494-498.]
- [6] Sun Siyang. Inversion of water quality parameters of Miyun Reservoir based on multi-source remote sensing and machine learning[D]. Beijing: Beijing Forestry University, 2019.[孙驷阳. 基于多源遥感与机器学习的密云水库水质参数反演研究[D]. 北京:北京林业大学, 2019.]
- [7] Wang Shan. Monitoring and analyzing of Panjiakou-Daheiting Reservoir water quality change based on multi-source remote sensing images[D]. Lanzhou: Lanzhou Jiaotong University, 2019.[王珊. 基于多源遥感影像的潘家口—大黑汀水库水质变化监测与分析[D]. 兰州: 兰州交通大学, 2019.]
- [8] Ma Fangkai, Gao Zhaobo, Ye Bangling. Remote sensing inversion of Tangxun Lake water quality based on high score remote sensing satellite image[J]. Water resources development and management, 2021(5): 69-75.[马方凯, 高兆波, 叶帮玲. 基于高分遥感卫星影像的汤逊湖水水质遥感反演[J]. 水资源开发与管理, 2021(5): 69-75.]
- [9] Li Yijing, Sun Xiaomin, Guo Yuyin, *et al.* Remote sensing retrieval of water quality parameters in Poyang Lake based on the gradient boosting decision tree algorithm[J]. Space Return and Remote Sensing, 2020, 41(6): 90-102.[李怡静, 孙晓敏, 郭玉银, 等. 基于梯度提升决策树算法的鄱阳湖水环境参数遥感反演[J]. 航天返回与遥感, 2020, 41(6): 90-102.]
- [10] Park Y, Cho K H, Park J, *et al.* Development of early-warning protocol for predicting Chlorophyll-a concentration using machine learning models in freshwater and estuarine reservoirs, Korea[J]. Science of The Total Environment, 2015,

- 502: 31-41. DOI: 10.1016/j.scitotenv.2014.09.005.
- [11] Cheng Chunmei, Wei Yuchun, Wang Guoxiang, *et al.* Using spectral smoothing method to improve the validation precision of the Chlorophyll-a estimation models in turbidity waters[J]. Remote Sensing Technology and Application, 2013, 28(6): 941-948.[程春梅, 韦玉春, 王国祥, 等. 使用光谱平滑提高浑浊水体叶绿素a浓度估算模型的应用精度[J]. 遥感技术与应用, 2013, 28(6): 941-948.]
- [12] Huang Changchun, Li Yunmei, Xu Liangjiang, *et al.* Study on influencing factors and universality of Chlorophyll-a retrieval model in inland water body [J]. Environmental Science, 2013, 34(2): 525-531.[黄昌春, 李云梅, 徐良将, 等. 内陆水体叶绿素反演模型普适性及其影响因素研究[J]. 环境科学, 2013, 34(2): 525-531.]
- [13] Cui, T W, Zhang J, Wang K, *et al.* Remote sensing of chlorophyll a concentration in turbid coastal waters based on a global optical water classification system [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 163: 187-201. DOI: 10.1016/j.isprsjprs.2020.02.017.
- [14] Tavora J E, Boss S, Doxaran D, *et al.* An algorithm to estimate suspended particulate matter concentrations and associated uncertainties from remote sensing reflectance in coastal environments[J]. Remote Sensing, 2020, 12(13): 2172. DOI: 10.3390/rs12132172.
- [15] Chen Jun, Wen Zhenhe, Fu Jun. The application of the subsection mapping retrieval model to water qualities quantitative analysis [J]. Spectroscopy and Spectral Analysis, 2010, 30(10): 2784-2788.[陈军, 温珍河, 付军. 基于分段映射模型的水质参数遥感反演研究[J]. 光谱学与光谱分析, 2010, 30(10): 2784-2788.]
- [16] Xue Pengsong, Feng Minquan, Xing Xiaopeng. Water quality prediction model based on Markov chain improved grey neural network[J]. Engineering Journal of Wuhan University, 2012, 45(3): 19-24.[薛鹏松, 冯民权, 邢肖鹏. 基于马尔科夫链改进灰色神经网络的水质预测模型[J]. 武汉大学学报工学版, 2012, 45(3): 19-24.]
- [17] Liu S Y, Tai H J, Ding Q S, *et al.* A hybrid approach of support vector regression with genetica gorithm optimization for aquaculture water quality prediction [J]. Mathematical and Computer Modelling, 2013, 58(1): 458-465. DOI: 10.1016/j.mcm.2011.11.021.
- [18] Yang Liu, Han Yu, Wang Zumao, *et al.* Study on retrieval model of water quality parameter in Wenyu River based on BP neural network[J]. Journal of Water Resources and Water Engineering, 2013, 24(6): 25-28.[杨柳, 韩瑜, 汪祖茂, 等. 基于BP神经网络的温榆河水质参数反演模型研究[J]. 水资源与水工程学报, 2013, 24(6): 25-28.]
- [19] Zhang Sisi, Gao Xuguang, Hua Wenqiang. Remote sensing image information extraction method based on clustering and artificial neural network [J]. Electronic Design Engineering, 2020, 28(15): 106-109.[张思思, 高旭光, 滑文强. 基于聚类与人工神经网络的遥感图像信息提取方法[J]. 电子设计工程, 2020, 28(15): 106-109.]
- [20] Yang Guofan, Yan Mengdong, Yin Fei. Inversion modeling of Chlorophyll-a in Qinghe Reservoir[J]. Remote Sensing Information, 2016, 31(5): 74-78.[杨国范, 阎孟冬, 殷飞. 清河水库叶绿素a浓度反演模型研究[J]. 遥感信息, 2016, 31(5): 74-78.]
- [21] Pan Jun, Liang Haitao, Yue Dandan, *et al.* Application of BP artificial neural network model in the evaluation of groundwater quality[J]. Water Supply Technology, 2015, 9(6): 6-11.[潘俊, 梁海涛, 岳丹丹, 等. BP人工神经网络模型在地下水水质评价中的应用[J]. 供水技术, 2015, 9(6): 6-11.]
- [22] Guo Qingchun, He Zhenfang, Li Li. Water quality assessment model of Yangtze River based on artificial neural networks[J]. Information Technology, 2013(8): 54-56, 60.[郭庆春, 何振芳, 李力. 基于人工神经网络的长江水质评价模型[J]. 信息技术, 2013(8): 54-56, 60.]
- [23] Li M S, Wu W, Chen B S, *et al.* Water quality evaluation using back propagation artificial neural network based on self-adaptive particle swarm optimization algorithm and chaos theory [J]. Computational Water Energy & Environmental Engineering, 2017, 6(3): 229-242. DOI: 10.4236/cweee.2017.63016.
- [24] Zheng Yanhui, Zhang Yuanbo, He Yanhu. Study on the multi-parameter inversion of reservoir water quality based on GF-1 WFV images and neural network model [J]. People's Pearl River, 2020, 41(7): 57-62, 84.[郑炎辉, 张园波, 何艳虎. 基于GF-1 WFV影像与神经网络模型的水库水质多参数反演研究[J]. 人民珠江, 2020, 41(7): 57-62, 84.]
- [25] Xiao Qian, Yang Kun, Hong Liang. Comparative study of extraction methods of lake water in Yunnan-Guizhou plateau based on multi-source images [J]. Journal of Yunnan Normal University (Natural Science Edition), 2018, 38(2): 73-78.[肖茜, 杨昆, 洪亮. 基于多源数据的云贵高原湖泊水体提取方法对比研究[J]. 云南师范大学学报(自然科学版), 2018, 38(2): 73-78.]

Study on Water Quality Parameter Inversion based on Landsat 8 and Measured Data

Wu Huanhuan, Guo Qiaozhen, Zang Jinlong, Qiao Yue, Zhu Li, He Yunhai

(School of Geology and Geomatics, Tianjin Chengjian University, Tianjin 300384, China)

Abstract: At present, remote sensing technology has become an important method for monitoring water quality parameters, and a more accurate water quality parameter inversion model is the focus of current water quality monitoring. However, due to multiple reasons such as the complexity of the water environment and the limitations of remote sensing data, the accuracy of water quality parameter remote sensing inversion is limited, and most of them focus on the inversion of water color water quality parameters. In order to obtain a better accurate water quality parameter inversion model, taking the lower reaches of the Haihe River in Tianjin as the research area, Landsat 8 OLI remote sensing images were subjected to atmospheric correction, radiometric calibration and other pretreatments, and the total phosphorus, ammonia nitrogen, total nitrogen concentration and conductivity of the water body were determined by laboratory physical and chemical analysis. The statistical regression model and neural network model of measured water quality parameters and Landsat 8 OLI remote sensing image data are established. Coefficient of determination (R^2), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used to test the accuracy, and the neural network model inversion results R^2 is greater than 0.85, MAE is 0.019, 0.09, 0.242, 0.411, RMSE is 0.024, 0.118, 0.286, 0.562, and the inversion accuracy is better. The results show that the water quality parameter inversion model based on neural network has high accuracy.

Key words: Landsat 8 OLI remote sensing images; Neural network model; Water quality parameters; Inversion