

· 研究报告 ·

喻家湖水水质时空分布特征和影响因子分析

王召唤, 张延荣*

(华中科技大学环境科学研究所, 湖北 武汉 430074)

摘要: 通过设计合理的水质监测网, 采用多元统计分析, 并结合地理信息技术对武汉市喻家湖在 2011 年—2012 年期间 12 个监测点、13 个水质参数监测值进行水质时空分布特征研究。结果表明, 喻家湖 13 个水质指标概括为 4 个主成分: 第一主成分代表喻家湖的重金属污染, 第二主成分代表其富营养化水平, 第三主成分代表有机污染程度, 第四主成分间接指示富营养化程度; 在时间和空间变化上都可分为二组, 显著性指标的时空差异较明显, 水质污染程度从南至北逐渐减弱, 湖溪河是喻家湖的最主要污染源。并对水质参数、监测点位和频次进行了优化。

关键词: 水质; 时空分布; 多元统计分析; 地理信息系统; 喻家湖

中图分类号: X824 文献标识码: B 文章编号: 1006-2009(2013)02-0016-05

Identification of Spatio-temporal Distribution Characteristics and Its Influencing Factors of Water Quality on Yujia Lake

WANG Zhao-huan, ZHANG Yan-rong*

(Environmental Science Research Institute, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China)

Abstract: Combined with a rational water quality monitoring program, Multivariate Statistical Analysis and geographic information system (GIS) were used to analyze data sets of water quality for 11 parameters monitored at 12 different sites in Yujia Lake from 2011 to 2012 to determine temporal and spatial distribution characteristics in water quality. The results demonstrated that: the significance of the four principal components was verified, the first varifactor (VF1) represented the pollution of heavy metals in Yujia Lake, the second varifactor (VF2) indicated the level of eutrophication, the third varifactor (VF3) represented the organic pollution, the fourth varifactor (VF4) indicated indirectly the level of eutrophication; the monitoring periods were classified into 2 clusters and the monitoring sites were partitioned into 2 groups, the temporal and spatial variations were significant. The water pollution levels became gradually weaker from south to north, Huxi River was the most important source of pollution. The frequency and sites for the monitoring of water quality parameters were optimized properly.

Key words: Water quality; Spatio-temporal distribution; Multivariate statistical analysis; GIS; Yujia Lake

多元统计分析 (MSA) 中的主成分分析 (Principal Component Analysis, PCA)、聚类分析 (Cluster Analysis, CA)、判别分析 (Discriminant Analysis, DA) 等统计分析方法已被成功用于河流^[1]、湖泊^[2]、沿海^[3-4]、地下水^[5] 等水体水质评价和时空分布特征的研究。GIS (地理信息系统) 在空间分析的核心优势为时空变异特征的可视化提供了更好的技术支持^[6]。MSA 结合 GIS 技术已成为研究水质时空分布模式的有效工具, 在最近几年的研究

中被广泛应用^[3-4]。而对研究结论有重要影响的原始数据的获得和预处理过程却经常被研究者忽视^[3]。

收稿日期: 2012-09-03; 修订日期: 2013-02-28

基金项目: 华中科技大学自主创新基金资助项目 (2010 zz 004-05)

作者简介: 王召唤 (1988—), 男, 河南鹤壁人, 硕士研究生, 研究方向为水环境评价与监测技术。

* 通信作者: 张延荣 E-mail: yanrong_zhang@yahoo.cn

喻家湖毗邻武汉市马鞍山森林公园,是武汉东湖 10 大子湖之一,属于典型的长江中下游浅水型内陆湖泊。现通过设计合理的水质监测网,解决了以往研究中普遍存在的指标单一,代表性差,监测点位不合理,不能全面反映水质状况的问题^[7]。原始数据经预处理后,运用 PCA 方法对所选指标降维,以较小的维度对喻家湖的整体水质状况作出评价和分析。在此基础上采用 CA、DA 等方法,结合 GIS 技术对喻家湖的水质时空变异模式进行探索和模拟。此研究旨在为中小型湖泊水质的时空分布特征提供一种新的研究方法,同时为喻家湖水体污染控制和水质管理提供科学依据。

1 监测网设计与方法

1.1 研究区域与点位设计

喻家湖位于湖北省武汉市,地处 114°25'32"—114°25'38"E、30°31'13"—30°31'56"N,总面积约为 0.65 km²,平均水深 2.18 m,位于喻家湖南侧的湖溪河河水通过涵洞流入湖内,而北侧与东湖相连。

监测网布点依据:(1)湖库面积和湖盆形态;(2)污染源位置;(3)污染物在水体中的迁移和扩散;(4)根据补给条件、入水及出水位置等原则^[6]确定设置 12 个采样点,分别为 S1~S12,覆盖了喻家湖全部区域,保证了所测数据的代表性和有效性,见图 1。

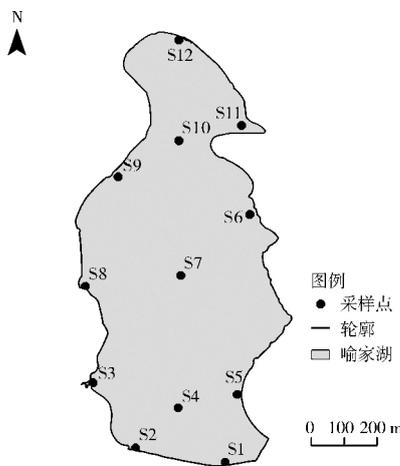


图 1 喻家湖水质监测点位分布

Fig. 1 Location of sampling sites in Yujia Lake

1.2 监测参数与频率

我国湖泊污染具有严重、复杂和全面化 3 个特

点,而富营养化和有毒有害物质形成的复合污染是我国湖泊污染的显著特征^[8]。研究表明,营养物浓度对湖泊富营养化起决定作用^[9]。霍守亮等^[9]将 P、N、Chla、SD 作为 4 项基本的营养物基准指标候选变量,适当加入其他水质指标如 pH 值、DO 等,根据我国的地表水环境质量和武汉市湖泊污染状况将 Cu、Zn、Pb、Cd 等重金属纳入水质监测指标。

现选用 SD、pH 值、DO、BOD₅、TN、NH₃-N、TP、I_{Mn}、Chla、Cu、Zn、Pb、Cd 共 13 个水质指标,pH 值和 DO 采用便携式仪器法测定,SD 采用塞氏盘法,BOD₅ 采用稀释与接种五日培养法,TN 采用碱性过硫酸钾消解紫外分光光度法测定,NH₃-N 采用纳氏试剂分光光度法测定,TP 采用钼酸铵分光光度法测定,I_{Mn}采用高锰酸钾滴定法测定,Chla 采用超声-乙醇萃取分光光度法测定,Zn 采用火焰原子吸收分光光度法测定,Cu、Pb、Cd 采用螯合萃取火焰原子吸收分光光度法测定。测定过程中均严格按照相关标准进行。采样点监测频率为每月 2 次,采用的数据为 2011 年 4 月—2012 年 3 月共计一整年喻家湖 12 个采样点的 13 个水质因子的监测值,见表 1。

表 1 喻家湖水质统计结果 mg/L

Table 1 Statistical description of Yujia lake's water quality mg/L

指标	最小值	最大值	平均值	偏度
pH 值	6.9	9.5	8.3	0.304
DO	0.4	18.0	9.0	0.113
SD ^①	0.28	0.65	0.45	0.133
BOD ₅	7.5	55.5	21.8	0.835
TN	0.65	17.76	5.45	0.983
NH ₃ -N	0.03	26.11	4.39	1.781
TP	0.03	0.66	0.22	1.735
I _{Mn}	5.0	26.0	13.4	0.455
Chla	0.0	321.0	82.0	1.138
Cu	0.001	0.046	0.008	2.252
Zn	0.05	0.30	0.08	2.127
Pb	0.01	0.15	0.03	1.620
Cd	0.000 1	0.011	0.002	1.926

①单位为 m。

1.3 数据分析方法

1.3.1 方法选择

PCA 是将多个变量通过线性变换以选出较少个数重要变量的多元统计分析方法^[10]。数据经过

预处理之后,采用 IBM SPSS Statistics 19 的因子分析模块处理。在进行主成分分析之前,先作 Kaiser-Meyer-Olkin (KMO) 和 Bartlett 的球形度检验,验证其是否适用于主成分分析。通过 SPSS 得到 KMO 检验结果为 0.692, Bartlett 的球形度检验结果为 1490.302 (显著性水平 $p < 0.05$),说明主成分分析能够较好地降低原始变量的维度^[3,7]。经过几次试验后发现,将 13 个水质指标数据概括为 4 个主成分可以说明原始数据至少 2/3 的变化。

层次聚类分析 (hierarchical CA, HCA) 是 CA 中应用最为广泛的探索性方法。其实质是根据观察值或变量之间的亲疏程序,以逐次聚合的方法,将最相似的对象结合在一起,直到聚成一类^[3]。利用 HCA 进行时空相似性分析,采用的计算方法是欧氏距离平方和离差平方法。

DA 方法可通过判别函数判别 DA 分析结果和识别显著性的污染指标^[7,11]。其基本原理是按照一定的判别准则,建立一个或多个判别函数,用研究对象的大量资料确定判别函数中的待定系数,并计算判别指标,据此即可确定某一样本属于何类^[3]。根据上述分析结果,采用 ESRI 公司的 ArcGIS 9.3 中的地统计分析模块的普通 Kriging 法实现时空分布的可视化。

1.3.2 数据预处理

研究中涉及多元统计分析和多指标的综合评价,依据相应的原则进行数据预处理^[3,12]。

(1) 缺失数据采用剩余数据平均值填补,例如 Pb 和 Zn 检测值中 4 月份和 6 月份数据缺失;

(2) 低于检测限数据值采用最低检测限值的一半来代替;

(3) 在综合评价中要求指标同趋势化^[12],SD 值以《地表水资源质量评价技术规程》(SL 395-2007)中湖泊(水库)营养状态评价标准及分级方法中规定的中营养水平(EI=50)SD 值 1.0 m 为基准值,以基准值减去测量值得到新的变量,该变量值越大代表水质污染越严重;DO 值以 20 mg/L 作为基准值,以基准值减去测量值得到新的变量;

(4) 原始数据进行偏斜度检验水质指标分布特征,结果为 0.113 ~ 2.252,其中 $\text{NH}_3 - \text{N}$ 、TP、Chla、Cu、Zn、Pb、Cd 的 skewness 值均 > 1,明显正偏,采用平方根转换后进行 skewness 检验,结果为 -0.443 ~ 1.347,可认为各指数符合或接近正态分布;

(5) 在进行主成分分析和聚类分析时,考虑到水质指标数量级的差异和量纲的不同,需要对数据进行标准化(均值为 0,方差为 1)。

2 结果与讨论

2.1 主成分分析

前 4 个主成分的特征值都 > 1,累计方差贡献率为 68.861%,能够反映原始变量的绝大部分信息,选取前 4 个主成分作为分析对象,分别命名为 VF1、VF2、VF3、VF4。采用最大方差法进行正交旋转后得到主成分和各个参数之间的相关系数 r 。一般认为,相关系数的绝对值 > 0.6 的指标和该主成分显著相关^[7],根据各个监测点位的因子得分平均值进行空间插值,结果见图 2(a)(b)(c)(d)。

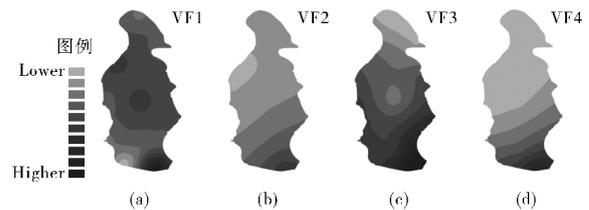


图 2 喻家湖因子得分空间插值

Fig. 2 Spatial interpolation of factor scores Yujia lake

第一主成分 (VF1) 旋转平方和方差贡献率为 19.129%,与 Cu ($r = 0.728$)、Pb ($r = 0.720$)、Cd ($r = -0.857$) 具有显著相关,可认为第一主成分代表喻家湖的重金属污染。

第二主成分 (VF2) 旋转平方和方差贡献率为 18.268%,与 TN ($r = 0.695$)、 $\text{NH}_3 - \text{N}$ ($r = 0.817$)、Chla ($r = -0.639$) 存在显著相关,与 TP ($r = 0.572$) 中度相关, VF2 代表喻家湖的富营养化,结合 VF2 的空间分布图(图 2),可以认为 VF2 主要来自湖溪河的废水。

研究期间喻家湖的 TN 平均值为 5.45 mg/L, TP 平均值为 0.22 mg/L, $\rho(\text{TN}) / \rho(\text{TP}) \approx 24.77$ (> 7 ,藻类正常代谢所需的 $\rho(\text{N}) / \rho(\text{P}) = 7$),仅控制氮磷污染的外源输入,在短期并不能改变喻家湖的富营养化状态,只有在减少营养物质外源输入的同时,控制喻家湖水体中氮磷的内源释放,才能有效地控制喻家湖水体富营养化。

第三主成分 (VF3) 旋转平方和方差贡献率为 17.664%,与 SD ($r = -0.808$)、 BOD_5 ($r = 0.677$)、 I_{Mn} ($r = 0.768$) 存在显著相关,可以认为 VF3 代表

喻家湖的有机污染。由图 2 可见, VF3 和 VF2 的空间分布总体相同, 可以认为湖溪河废水是喻家湖有机污染的主要来源。

第四主成分(VF4) 旋转平方和方差贡献率为 13.800% , 与 pH 值($r = -0.842$) 、DO($r = 0.918$) 存在显著相关。研究期间 DO 平均值为 9.01 mg/L, 处于过饱和状态, 指示过高的光合作用活性和富营养状态^[9]。通过相关分析发现 pH 值和 DO 存在显著相关($r = 0.754$, $p < 0.05$) , 可能与水体中氧化还原电位和水生植物光合作用伴随的代谢活动相关。因此 VF4 间接反映喻家湖的富营养化。由图 2 可见, VF4 和 VF2 的分布基本相同, 从而可以认为 VF4 间接指示喻家湖的富营养化程度。

2.2 喻家湖水质年度变化特征

依据监测点位每个月的平均值, 对 2011 年 4 月—2012 年 3 月共计 12 个月进行聚类分析, 结果见图 3。

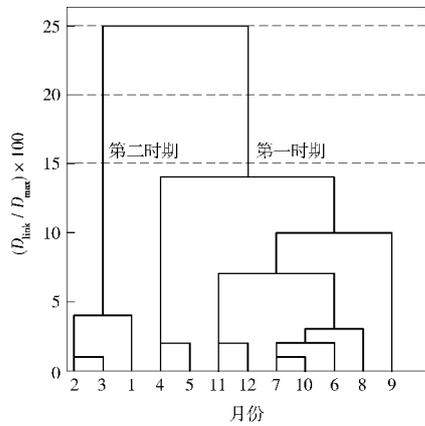


图 3 基于 Ward's 方法的时间聚类分析
Fig. 3 Temporal cluster analysis of monitoring periods based on Ward's methods

在(D_{link}/D_{max}) $\times 100 < 15$ 和 11 处(D_{link}/D_{max} 表示个案连锁距离与最大距离之比) 可分为 2 组和 3 组。按照 2 组可以分为: 第一时期为 4—12 月, 第二时期为 1—3 月, 按照 3 组可以分为: 第一时期为 4—5 月, 第二时期为 6—12 月, 第三时期为 1—3 月。分类结果表明, 简单地按照枯水期/丰水期或是季节变化来进行喻家湖的水质监测和评价是有偏差的, 还要考虑到水体污染特征^[3]。

为了验证上述时间相似性聚类分析结果和进一步识别显著性指标, 应用 DA 方法进行时间差异性判别分析, 采用 2 组分类的判别正确率为 95.8% , 采用 3 组分类的判别正确率为 93.4% , 因此在时间尺度上将喻家湖分为 2 个时期, 第一时期: 4—12 月份; 第二时期: 1—3 月份。显著性指标有 8 个, 即 SD、BOD₅、NH₃-N、I_{Mn}、Chla、Cu、Zn、Pb, 这些参数是造成喻家湖水质时间差异的主要因素。

相应的判别函数 TDF1 和 TDF2 为式(1)和式(2) (式中参数均代表标准值)。在今后的监测系统中, 只需选择上述 8 个指标在两个时期的测定值来反映喻家湖水质在时间尺度上的差异性, 这样既保证了监测方案的合理及结果的准确, 并在最大程度上节约了监测成本。

$$TDF1 = -2.009 SD - 0.722 BOD_5 + 1.815 NH_3 - N + 0.868 I_{Mn} - 1.714 Chla - 1.165 Cu - 0.719 Zn - 1.735 Pb - 4.977 \quad (1)$$

$$TDF2 = 0.667 SD + 0.236 BOD_5 - 0.625 NH_3 - N - 0.285 I_{Mn} + 0.547 Chla + 0.82 Cu + 0.236 Zn + 0.562 Pb - 1.159 \quad (2)$$

根据判别函数, 选择 SD、NH₃-N、Chla、Pb 4 种显著性指标进行差异性空间模拟, 结果见图 4 (a) (b) (c) (d), 其中(1) (2) 分别为第一时期、第二时期。

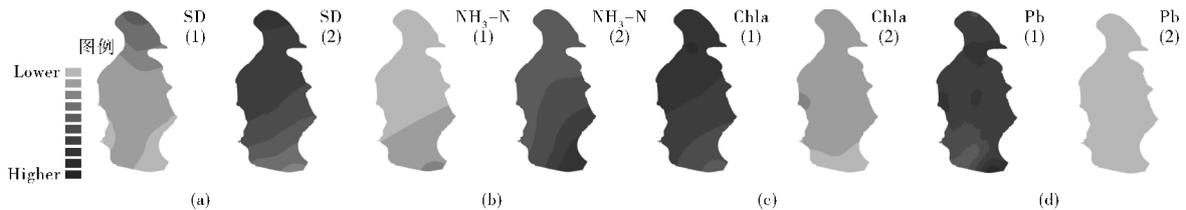


图 4 时间尺度显著性指标的插值
Fig. 4 Spatial interpolation of significant parameters of temporal scale

从时间上来看,这 4 个显著性指标在 2 个时期存在显著的差异,进一步证明了判别分析的正确性。SD 在第二时期(1—3 月)的值大于第一时期(4—12 月),并且在 2 个时期的空间分布特征都呈现由南至北逐渐增加的趋势,Chla 第一时期的浓度明显高于第二时期,但 Chla 和 SD 在空间上表现相同的分布,因此 SD 在时间尺度上与 Chla 负相关,在空间尺度上正相关。 $\text{NH}_3 - \text{N}$ 在第一时期的浓度明显高于第二时期,空间分布上 2 个时期都表现出由南至北逐渐减小的趋势,靠近湖溪河部分水域的污染程度最高。

2.3 水质空间相似性和差异性分析

依据监测点位在整个时期的各个水质指标的监测平均值,对 S1 ~ S12 共计 12 个采样点进行聚类分析,结果见图 5。

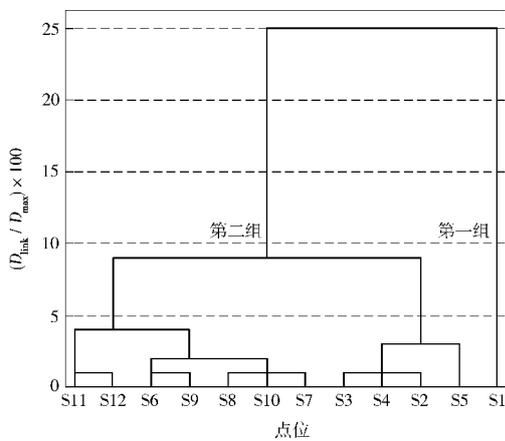


图 5 基于 Ward's 方法的空间聚类分析

Fig. 5 Spatial cluster analysis of monitoring periods based on Ward's methods

在 $(D_{link} / D_{max}) \times 100 < 10$ 和 5 处明显可分为 2 组和 3 组。如果按 2 组来分,第一组为 S1,第二组

为 S2 ~ S12; 如果按 3 组来分则上述第 2 组进一步分为 S2 ~ S5 和 S6 ~ S12。S1 点位于湖溪河入口处,湖溪河是喻家湖的最主要污染源,主要接纳生活污水; S2 ~ S5 位于 S1 周围; S6 ~ S12 靠近东湖。

为了验证空间相似性聚类分析结果,利用上述 12 个采样点全年监测值作 DA 分析。采用 2 组分类的判别正确率为 92.0%,采用 3 组分类的判别正确率为 67.4%,因此采用 2 组分类是比较合适的。

对于空间尺度而言,显著性指标有 6 个,分别为 pH 值、SD、 $\text{NH}_3 - \text{N}$ 、TN、Cu、Pb,相应的判别函数为式(3)和式(4)(式中各参数均指标准值),采用普通 Kriging 插值方法对这 6 个显著性指标进行差异性模拟,结果见图 6(a) (b) (c) (d) (e) (f)。

根据上述分析结果以及图 6,可以进一步优化监测网络,第一组仅设置 S1,第二组设置 S4、S7、S12; 选择上述 6 种水质参数就可以较好地表征喻家湖水质的空间分布特征。

$$\text{SDF1} = -1.937 \text{pH} + 1.411 \text{SD} - 1.311 \text{TN} + 2.388 \text{NH}_3 - \text{N} + 1.841 \text{Cu} + 0.611 \text{Pb} - 4.425 \quad (3)$$

$$\text{SDF2} = 0.177 \text{pH} - 0.132 \text{SD} + 0.130 \text{TN} - 0.238 \text{NH}_3 - \text{N} - 0.177 \text{Cu} - 0.068 \text{Pb} - 0.727 \quad (4)$$

TN、 $\text{NH}_3 - \text{N}$ 质量浓度呈现由南至北逐渐减少态势,与 Chla 的趋势相反,可能是因为过多的营养盐输入和有毒物质对藻类的生长起到抑制作用。Cu 质量浓度在空间表现由南至北逐渐升高的趋势,Pb 则表现复杂的不规则分布,可能是因为底泥中重金属的内源释放是喻家湖表面水体重金属的重要来源。

通过时空差异性模拟,得知喻家湖最主要的污染源是湖溪河。通过重点削减或截断湖溪河入湖污水,可以减小或切断外源污染的输入,并适当控制喻家湖底泥的内源释放,从而达到喻家湖的水污染控制和管理的目的。

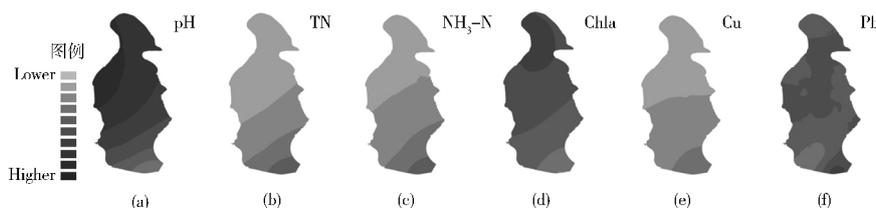


图 6 空间尺度显著性指标的插值

Fig. 6 Spatial interpolation of significant parameters of spatial scale

(下转第 66 页)

期基本上都为V类水,只有徐家大塘为IV类水。

2.5 评价结果比较

分别采用单因子指数评价、综合指数评价、模糊综合评价法对漕桥河各断面不同水期进行水质评价。通过上述的评价结果比较可以看出:单因子指数评价方法简单明了,但缺乏对环境及其综合质量的整体反映,仅表示了水质类别。综合污染指数是相对标准值计算的结果,综合反映了河流污染状态,能全面平衡污染因子对水质的贡献度,可以比较各监测断面污染程度,但不能说明水质类别。模糊综合评价法将模糊数学应用到水质综合评价,采用隶属函数来描述水质的分级,因而能够较好地反映水体真实状况。

(上接第20页)

3 结论

(1) 通过设计合理的水质监测方案,所得数据用于主成分分析,将喻家湖13个水质指标统计分类为4个主成分,可以反映至少2/3的原始信息:VF1与Cu、Pb、Cd显著相关,代表喻家湖的重金属污染;VF2与NH₃-N、TN、Chla显著相关,与TP中度相关,代表喻家湖富营养化水平;VF3与SD、BOD₅、I_{Mn}显著相关,代表喻家湖有机污染程度;VF4与pH值、DO显著相关,间接指示其富营养化程度;

(2) 水质的时间变化可以分为2期,第一时期为4—12月,第二时期为1—3月,DA分析表明分为2组的正确率为95.8%;表征时间差异的显著性指标有8个,分别为SD、BOD₅、NH₃-N、I_{Mn}、Chla、Cu、Zn、Pb;

(3) 空间变化上也可分为2组,第一组为S1,第2组为S2~S12;DA分析表明分为2组的正确率为92.0%;表征空间差异的显著性指标是pH值、SD、NH₃-N、TN、Cu、Pb;主要污染物TN、NH₃-N的浓度从北到南逐渐减小,Chla的趋势与其相反。湖溪河是喻家湖最主要污染源,底泥的内源释放也是喻家湖表面水体重金属含量的重要来源。

[参考文献]

- [1] SU S, LI D, ZHANG Q, et al. Temporal trend and source apportionment of water pollution in different functional zones of Qiantang River, China [J]. *Water Research*, 2011, 45(4): 1781-1795.

[参考文献]

- [1] 彭文启,张祥伟.现代水环境质量评价理论和方法[M].北京:化学工业出版社,2005.
 [2] 张新钰,辛宝东,刘文臣,等.三种地下水水质评价方法的对比分析[J].*水资源与水工程学报*,2011,22(3):113-118.
 [3] 周林飞,高云彪.模糊数学在湿地水质评价中应用的研究[J].*水利水电技术*,2005,36(1):35-38.
 [4] 杨庆.地下水易污染性评价-DRASTIC指标体系[J].*水文地质工程地质*,1999(2):4-9.
 [5] 赵昌刚.福建省南部沿海地区地下水资源易污性评价[J].*矿产保护与利用*,2006(4):51-54.
 [6] 孙卫红,程炜,崔玉霞,等.太湖流域主要入湖河流环境综合整治[J].*中国资源综合利用*,2009,27(11):39-42.
 [7] 邢肖鹏,薛鹏松,冯民权.汾河运城段河流水质评价[J].*水土保持通报*,2011,31(2):142-145.

本栏目责任编辑 李文峻

- [2] ZHAO Y, XIA X H, YANG Z F, et al. Assessment of water quality in Baiyangdian Lake using multivariate statistical techniques [J]. *Procedia Environmental Sciences* 2012, 13: 1213-1226.
 [3] ZHOU F, HUANG G H, GUO H, et al. Spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong [J]. *Water Research*, 2007, 41(15): 3429-3439.
 [4] GU Y, WANG Z, LU S, et al. Multivariate statistical and GIS-based approach to identify source of anthropogenic impacts on metallic elements in sediments from the mid Guangdong coasts, China [J]. *Environmental Pollution* 2012(163): 248-255.
 [5] MATHES S, RASMUSSEN T. Combining multivariate statistical analysis with geographic information systems mapping: a tool for delineating groundwater contamination [J]. *Hydrogeology Journal* 2006, 14(8): 1493-1507.
 [6] MANTZAFLERI N, PSILOVIKOS A, BLANTA A. Water quality monitoring and modeling in lake kastoria, using GIS. Assessment and management of pollution sources [J]. *Water Resources Management* 2009, 23(15): 3221-3254.
 [7] 谢森,何连生,田学达,等.巢湖水质时空分布模式研究[J].*环境工程学报*,2010(3):531-539.
 [8] 吴丰昌,孟伟,宋永会,等.中国湖泊水环境基准的研究进展[J].*环境科学学报*,2008,28(12):2385-2393.
 [9] 霍守亮,陈奇,席北斗,等.湖泊营养物基准的候选变量和指标[J].*生态环境学报*,2010,26(6):1445-1451.
 [10] KAZI T G, ARAIN M B, JAMALI M K, et al. Assessment of water quality of polluted lake using multivariate statistical techniques: A case study [J]. *Ecotoxicology and Environmental Safety* 2009, 72(2): 301-309.
 [11] 何万谦,黄金良.澳门半岛近岸海域水质时空变异分析[J].*环境科学*,2010(3):606-611.
 [12] 叶宗裕.关于多指标综合评价中指标正向化和无量纲化方法的选择[J].*浙江统计*,2003(4):24-25.