

粤西海域水质评价方法的研究及应用^{*}

张莹¹, 谢仕义^{1*}, 杨锋²

(1. 广东海洋大学 海洋遥感与信息技术实验室, 广东 湛江 524088; 2. 湛江市海洋与渔业环境监测站, 广东 湛江 524039)

摘要:针对海洋水质评价问题,建立了一套先利用主成分分析法明确各水质指标权重关系,继而参考权重关系依照海水水质分类标准建立分类样本,最终通过判别分析方法建立判别函数实现水体自动分类的方案。该方案在粤西海域,水体分类回判符合率达到91.4%。依据自动分类结果绘制了粤西海域2008—2010年水体分布专题图,为粤西海域环境污染防治提供了决策依据。

关键词:粤西海域;水体分类;主成分分析;判别分析

中图分类号: X824

文献标志码: A

文章编号: 1671-6647(2012)02-0198-07

海洋水质评价是海洋环境管理和决策的重要组成部分,其评价内容是根据海水水体的主要物质成分及其含量,按照《国家海水水质标准》方案,分析水质的时空分布状况,从而为海洋养殖业、临海工业及地方经济的发展规划及管理决策提供依据。

从1965年Jacobs提出水质指数概念以及相应公式^[1]以来,有关各类水体评价方法的研究成果不断涌现。早期主要有单因子指数法^[2-4]、综合指数法^[5-6]、分级评价法^[7]等,近期主要有模糊理论、灰色系统理论、投影寻踪及神经网络等方法^[8-11]。但由于水质系统是一个随时空变化的多维因子组成的复杂系统,依据不同区域,不同时期水体情况建立起来的评价模型很难通用,因而至今国内外还没有统一的水质评价模型与方法。

我们针对南海粤西海域水质实测资料,为完成海水水质评价工作,设计下述方案:首先,利用主成分分析法明确海水水质各指标的权重关系;其次,依据权重关系按照《国家海水水质标准》建立水体分类样本;最后,利用判别分析方法分析样本数据,建立水质自动分类模型。利用上述方案对粤西海域2010年35个站位水质实测资料进行了处理,其数点回判正确率达到91.4%。同时利用建立的自动分类模型对另外10个站位的水质实测资料进行了自动分类,分类结果全部正确。又将上述方案应用到2008和2009年粤西海域水体分类中,通过绘制水体分类专题图,对比分析了粤西海域2008—2010年水体类型分布的变化情况,为粤西海域的生态治理与资源规划提供了依据。

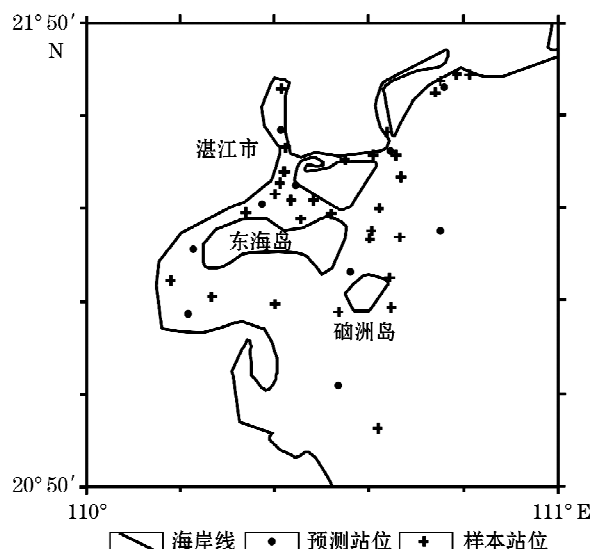


图1 粤西海域站位分布图

Fig. 1 Distribution of monitoring stations in seawaters off western Guangdong

^{*} 收稿日期:2011-02-14

资助项目:广东省科技项目——粤西海域卫星遥感水质监测速报与预警系统(2009B030600007)

作者简介:张莹(1982-),女,吉林四平人,讲师,博士,主要从事海洋探测信息处理与解释方面研究。

E-mail: zhangying198206@126.com

^{*} 通讯作者,男,四川巴中人,教授,硕士. E-mail: shiyxie@126.com

(陈靖 编辑)

1 研究区及水质资料概况

粤西海域重点研究范围为 $20^{\circ}20' \sim 21^{\circ}40'N, 110^{\circ}10' \sim 111^{\circ}00'E$ 。沿海地区包括茂名和湛江两市,是广东海洋经济圈的重要组成部分。该海域水质采样站位共计 45 个(图 1),其中国定站位 8 个,省定站位 37 个。

水质资料包括全氮(TN)、全磷(TP)、溶解氧(DO)、化学需氧量(COD)、酸碱度(pH)、石油类(Oils)共计 6 个水质指标,时间为 2008—2010 年每年 8 月。我们选用 2010 年粤西海域水质资料为研究对象,选取其中平面分布均匀的 35 个站位水质资料作为样本,其余 10 个站位水质资料作为预测(图 1)。在 2010 年水体分类方案应用的基础上,又利用该方案得出了 2008 和 2009 年粤西海域水体类型的分布情况。

2 水质指标关系

海洋水质评价是 1 种多指标的复杂评价问题,因此明确影响水体分类的主导指标,对于后续的水质评价研究尤为重要。主成分分析是研究如何通过少数几个主成分来解释多个变量的方差—协方差结构的 1 种数学变换方法,它能够将多个实测变量转换为少数几个不相关的综合指标,并描述原始实测变量与后期综合指标间的关系。由此我们可以通过主成分分析方法来描述海洋水质各指标间的权重关系,为后续的水体分类研究奠定基础。

将粤西海域 2010 年 45 个采样站位的 6 个水质指标(TN, TP, DO, COD, pH 和 Oils)利用主成分分析方法进行处理。第 1 个主成分的特征根为 3.05,解释 6 个原始变量总方差的 50.831%;第 2 个主成分的特征根为 1.078,解释 6 个原始变量总方差的 17.968%;第 3 个主成分的特征根为 0.874,解释 6 个原始变量总方差的 14.571%。截止到第 3 个主成分,累计方差贡献率达到 83.37%,说明前 3 个主成分已经能够很好的解释原来 6 个水质指标的信息了。从主成分数与特征值所绘制的碎石图(图 2)也可以看出,第 1,2 主成分的特征值连线较之第 2,3 主成分特征值连线的斜率有较大差别,而第 2,3 主成分特征值连线与其后的各主成分特征值连线斜率基本一致。

利用回归法计算得出各水质指标与各主成分的相关系数(表 1)。其中第一主成分中 TN 的权重最高,继而为 TP。说明粤西海域 2010 年水体分类着重考虑的水质指标应为 TN 和 TP。

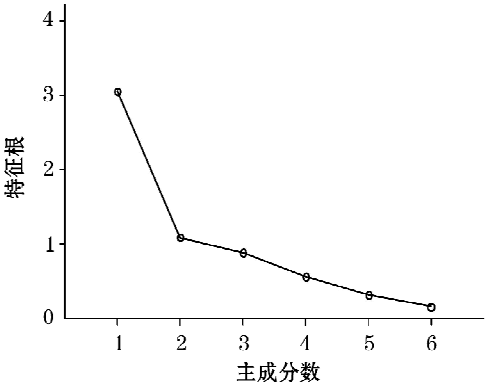


图 2 2010 年主成分分析碎石图
Fig.2 Scree plot of principal component analysis in 2010

表 1 2010 年粤西海域水质指标与主成分关系
Table 1 Relationship between the principal components and the water quality indexes in 2010 in the sea area off western Guangdong

主成分	TN/mg · L ⁻¹	TP/mg · L ⁻¹	DO/mg · L ⁻¹	COD/mg · L ⁻¹	pH	石油类/mg · L ⁻¹
第一主成分	0.927	0.832	0.683	-0.661	-0.586	-0.501
第二主成分	0.062	-0.328	0.437	0.212	0.496	-0.696

3 水体分类样本的建立

拟采用判别分析方法建立粤西海域水质自动分类模型,判别分析是对已知分类的数据建立判别函数,然后把这样的规则应用到未知分类的数据中去实现其分类。因此首先应建立粤西海域已知分类的水质资料样本。建立水体分类样本的步骤如下:

1) 建立粤西海域水质指标标准

按照 GB3097—1997《国家海水水质标准》,将粤西海域水体分为 4 类(表 2)。

表 2 粤西海域水质指标标准

Table 2 National standard for the sea water quality

评价指标	GB3097—1997《国家海水水质标准》			
	第一类	第二类	第三类	第四类
无机氮/ $\text{mg} \cdot \text{L}^{-1} \leq$ (以 N 计)	0.2	0.3	0.4	0.5
活性磷酸盐/ $\text{mg} \cdot \text{L}^{-1} \leq$ (以 P 计)	0.015	0.015	0.03	0.045
溶解氧/ $\text{mg} \cdot \text{L}^{-1} \geq$	6	5	4	3
化学需氧量/ $\text{mg} \cdot \text{L}^{-1} \leq$	2	3	4	5
pH	7.8~8.5 同时不超出该海域 正常变动范围的 0.2pH 单位		6.8~8.8 同时不超出该海域 正常变动范围的 0.5pH 单位	
石油类/ $\text{mg} \cdot \text{L}^{-1} \leq$	0.05		0.3	0.5

2) 依据水质指标权重关系建立样本

针对 2010 年粤西海域 35 个站位水质资料,按照粤西海域水质指标分类标准,重点考虑 TN 和 TP 质量浓度的分类界限,建立粤西海域 2010 年由 6 个水质指标(TN, TP, DO, COD, pH 和 Oils)组成的对应 4 种水体类型的水体分类样本。其中第一类水体 16 个样本点,第二类水体 2 个样本点,第三类水体 5 个样本点,第四类水体 12 个样本点。

4 水质自动分类方法

4.1 方法原理

判别分析在解决多元分析问题中应用较广,其基本思想是在已知若干个体的几组属性(或特征)观测数据及个体所属类别的前提下,根据某些准则建立判别公式,然后对未知类型的个体进行判别分析。判别分析法按判别准则的不同可分为距离判别法、Fisher 判别法、Bayes 判别法和逐步判别法等。

设有 n 项水质指标,建立判别函数:

$$Y_i = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \quad (1)$$

式中, Y_i 为第 i 个判别函数; $a_1, a_2, \cdots, a_n$ 为各水质指标的系数; $x_1, x_2, \cdots, x_n$ 为样本的各水质指标。本次研究的海水水质指标只有 6 项,则应将所有指标都进行判别分析,故选取了应用较为广泛的 Fisher 判别法。

4.2 实际应用

利用 Fisher 判别法对 2010 年粤西海域水质样本数据进行了分析,从中共可获得 3 个典型判别函数(表 3),其特征值为该函数组间平方和与组内平方和的比值。其中第 1 个判别函数对于分类的贡献率最大,达到 90.1%,也就是说利用第 1 个判别函数就已经能够较好的区分 4 类水体了。

表 3 判别函数说明表

Table 3 Necessary explanation for discriminant functions

判别函数	特征值	贡献率%	累计贡献率 %	典型相关系数
函数 1	8.379	90.1	90.1	0.945
函数 2	0.790	8.5	98.6	0.664
函数 3	0.134	1.4	100	0.343

用 $x_1, x_2, \dots, x_6$ 表示 TN, TP, DO, COD, pH, Oils 这 6 个水质指标,得到 3 个判别函数表达式:

$F_1 = 7.597 + 4.315x_1 + 28.043x_2 - 0.468x_3 - 0.138x_4 - 0.863x_5 - 1.140x_6$ (2)

$F_2 = -33.084 - 1.223x_1 + 51.095x_2 + 1.145x_3 - 0.224x_4 + 3.236x_5 - 12.627x_6$ (3)

$F_3 = -16.408 + 2.927x_1 - 6.713x_2 + 0.095x_3 - 0.405x_4 + 1.794x_5 + 29.094x_6$ (4)

由于前 2 个判别函数的贡献率已经达到 98.6%,故选取前 2 个判别函数进行分类。将全部 35 个水质数据的第一判别函数值与第二判别函数值在平面上投点,再根据样本中各点与类中心位置距离远近划分各点归属哪类(图 3)。由 Fisher 判别法分析得出的 2010 年粤西海域水质样本分类问题的回判正确率情况:第一类水体共计 16 个样本点,有 1 个样本点分类错误,对于一类水体判断符合率为 93.8%;第二类水体共计 2 个样本点,全部分类正确,对于二类水体判断符合率为 100%;第三类水体共计 5 个样本点,全部分类正确,对于三类水体判断符合率为 100%;第四类水体共计 12 个样本点,有 2 个样本点分类错误,对于四类水体判断符合率为 83.3%(表 4)。全部 35 个样本点判断符合率为 91.4%。说明利用 Fisher 判别法所建立的判别函数,能够有效的解决粤西海域水质自动分类问题。为了进一步验证上述方案的有效性,将粤西海域其余 10 个站点的水质数据代入式(2)和(3)中,计算 F_1 和 F_2 的值,投点判断其归属哪类(图 4),其中 5,6,7 号站点为第一类水体,判断正确;8 号站点属第二类水体,判断正确;1,2 号站点属第三类水体,判断正确;3,4,9,10 号站点属第四类水体,判断正确,则全部 10 个站点的水质预测分类正确率为 100%。

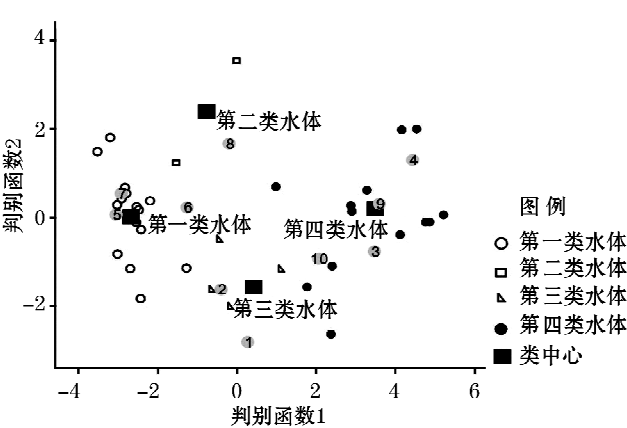


图 3 判别分析分类图

Fig. 3 Classification figure resulting from the discriminant analysis

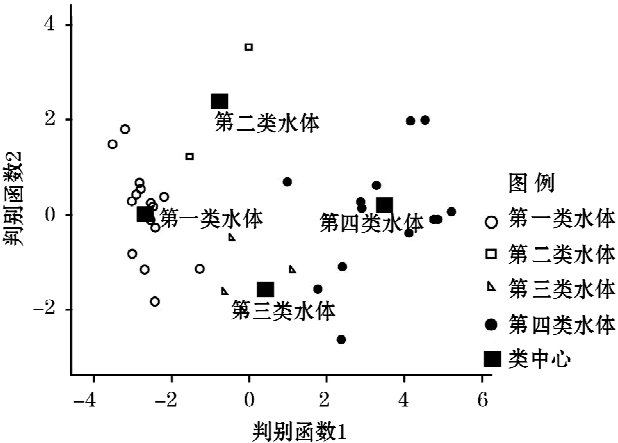


图 4 预测站点判别分析分类图

Fig. 4 Classification figure predicted from the discriminant analysis

表 4 逐步判别分析正确率
Table 5 Accuracy of stepwise discriminant analysis

统计量	原属类型	预测类型				合 计
		第一类水体	第二类水体	第三类水体	第四类水体	
个 数	第一类水体	15	0	1	0	16
	第二类水体	0	2	0	0	2
	第三类水体	0	0	5	0	5
	第四类水体	0	0	2	10	12
百分比/%	第一类水体	93.8	0	6.3	0	100
	第二类水体	0	100	0	0	100
	第三类水体	0	0	100	0	100
	第四类水体	0	0	16.7	83.3	100

在粤西海域 2010 年水质资料的分析中该套自动分类方案同样适用于研究不同年限同一采样站位群所获得的水质资料,因此,按照先明确水体分类主导指标,再依照海水水质标准建立样本,继而利用判别分析方法建立判别函数进行水质自动分类的思路,对粤西海域 2008—2010 年 45 个站位的水质资料进行了自动分类,并绘制了水体分类专题图(图 5)。

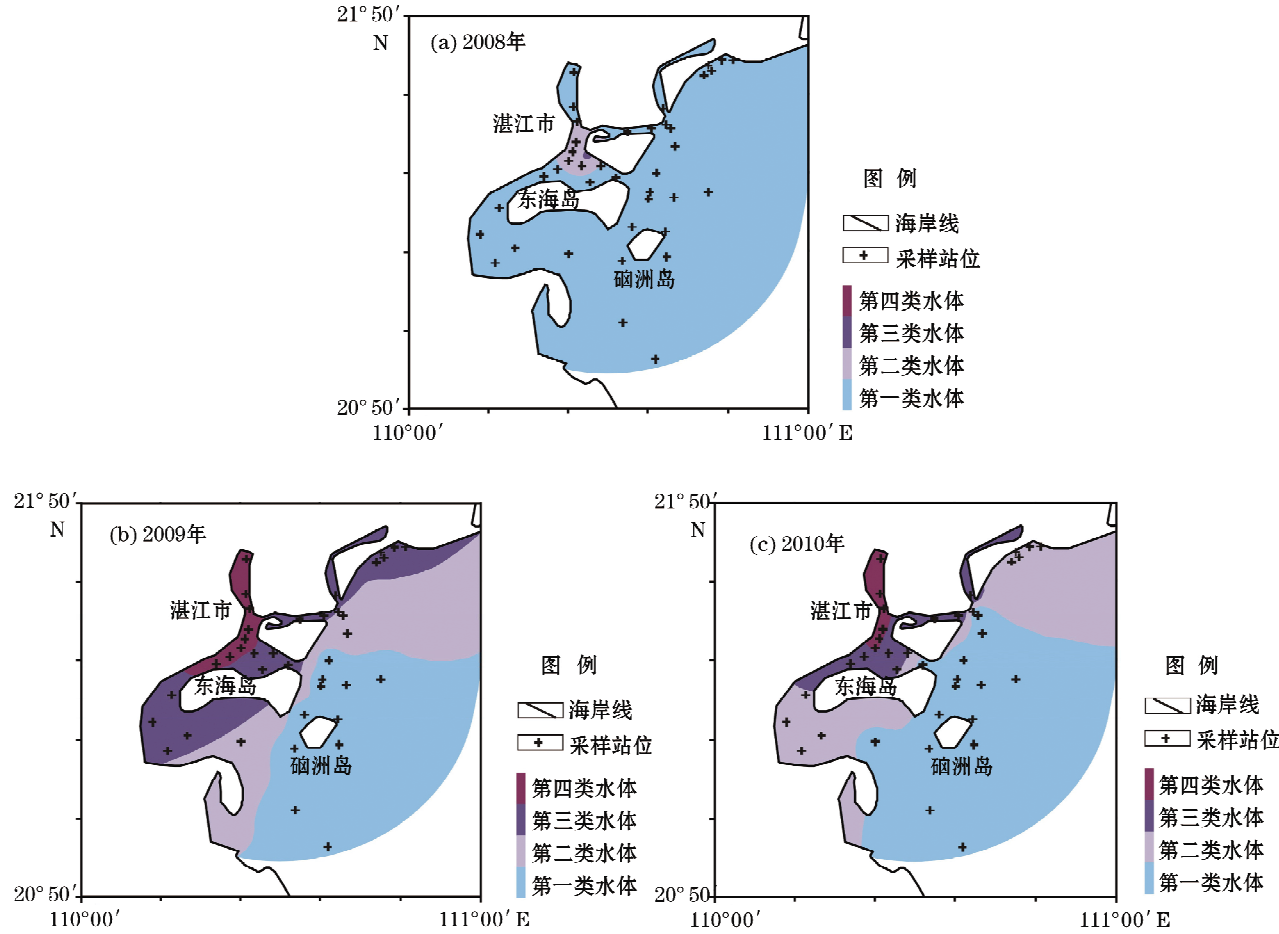


图 5 2008—2010 年水体分类图
Fig. 5 Water body classification from 2008 to 2010

2008—2010 年粤西海域水体分类图表明:2008 年时粤西海域大部分水体属第一类水体,仅在特呈岛附近海域出现了第二类水体和及小范围的第三类水体;2009 年时,粤西海域水质状况急剧变差,近岸海域出现了大面积的第三类及第四类水体;2010 年时,粤西海域的水质状况出现了好转,第二、三、四类水体的面积均变小,但是近岸仍然存在大面积的第三、四类水体,严重影响了海洋水产养殖业的发展。

5 结 论

依据粤西海域 2008—2010 年水体实测情况,建立了适用于该区域的水体分类方案,利用该分类方案得出了该海域水体分布的具体情况:

1) 粤西海域的水体分类方案:先通过主成分分析确定水体分类主导指标,再通过判别分析方法建立关于主导指标的 2 个判别函数,通过各类水体两函数交会区域范围的限定实现水体自动分类。该模型对粤西海域 2010 年 35 个站位的回判符合率达到 91.4%,10 个预测站位的正确率达到 100%。

2) 通过上述水质分类方案得出了粤西海域这 3 年中各类水体的分布状况,可概括为 2008 年整体水质状况较好,2009 年水质状况急剧变差,特别是近岸海域,到 2010 年时水质状况稍有好转,但近岸的海水水质状况仍不利于海洋水产养殖业的发展。

参考文献(References):

- [1] JACOBS H L. Water quality criteria[J]. Journal of Water Pollution Control Federation, 1965, 37(5): 292-300.
- [2] LIU C M, WEN W Y, HE Q X. Actuality survey and assessment of marine environment in Qiongzhou Strait[J]. Marine Environmental Science, 2002, 21(4): 24-28. 刘翠梅, 温伟英, 何清溪. 琼州海峡海洋环境现状调查与评价[J]. 海洋环境科学, 2002, 21(4): 24-28.
- [3] ZHANG Q F, ZHANG J L, DING J K. Assessment of marine environmental quality in Yanghe River Estuary[J]. Marine Environmental Science, 1997, 16(4): 38-42. 张秋丰, 张建乐, 丁家坤. 洋河口海域海洋环境质量现状评价[J]. 海洋环境科学, 1997, 16(4): 38-42.
- [4] SONG L, ZHOU Z C, WANG N B, et al. An appraisal of the environmental quality status and trend in Liaoning Coastal Area[J]. Fisheries Science, 2007, 26(11): 613-618. 宋伦, 周遵春, 王年斌, 等. 辽宁省近岸海洋环境质量状况与趋势评价[J]. 水产科学, 2007, 26(11): 613-618.
- [5] WALSKI T M, PARKER F L. Consumers water quality index[J]. Journal of Environmental Engineering Division, 1974, (EE3): 593-601.
- [6] SHANNON E E, BROEZONIK P L. Eutrophication analysis: A multivariate approach[J]. Journal of the Sanitary Engineering Division, 1972, 98(SA1): 37-57.
- [7] FU H, SUN Y L, SUN L, et al. The application of grey relational analysis in assessment of marine environmental quality[J]. Transactions of Oceanology and Limnology, 2007, 5(3): 127-131. 付会, 孙英兰, 孙磊, 等. 灰色关联分析法在海洋环境质量评价中的应用[J]. 海洋湖沼通报, 2007, 5(3): 127-131.
- [8] QIN C B, ZHENG B H, QIN Y W, et al. Gray correlation assessment of water environment quality for Tianjin coastal zone in Bohai Bay[J]. Research of Environmental Sciences, 2006, 19(6): 94-99. 秦昌波, 郑丙辉, 秦延文, 等. 渤海湾天津段海岸带水环境质量灰色关联度评价[J]. 环境科学研究, 2006, 19(6): 94-99.
- [9] LIU J Y, YANG T X, LI M, et al. A weight absolute gray correlation degree and it's application in evaluation of water quality of Miyun Reservoir [J]. Journal of Jilin University: Earth Science Edition, 2005, 35(1): 54-58. 刘金英, 杨天行, 李明, 等. 一种加权绝对灰色关联度及其在密云水库水质评价中的应用[J]. 吉林大学学报: 地球科学版, 2005, 35(1): 54-58.
- [10] WANG B X, YANG L F. Bioassessment of Qianhuai River using a river biological index[J]. Acta Ecologica Sinica, 2003, 23(10): 2082-2091. 王备新, 杨莲芳. 用河流生物指数评价秦淮河上游水质的研究[J]. 生态学报, 2003, 23(10): 2082-2091.
- [11] CAI R Y. Artificial neural network predicting model of noctiluca scientillans density and physical and chemical parameters[J]. Environmental Monitoring in China, 2001, 17(3): 52-55. 蔡如钰. 基于人工神经网络的夜光藻密度预测模型[J]. 中国环境监测, 2001, 17(3): 52-55.

Research and Application of the Method for Water Quality Evaluation in the Sea Area off Western Guangdong

ZHANG Ying¹, XIE Shi-yi¹, YANG Feng²

(1. *Lab of Ocean Remote Sensing and Information Technology, Guangdong Ocean University, Zhanjiang 524088, China;*

2. *Zhanjiang Oceanic and Fishery Environmental Monitoring Station, Zhanjiang 524039, China*)

Abstract: Aimed at the existing problems of seawater quality evaluation, a plan has been established, in which first is to determine the weight relationships among various water quality indexes by using principal component analysis, then to classify the samples according to the weight relationships and the water quality classification criteria, and last to establish a discriminant function by discriminant analysis and to realize the water body classification automatically. Good results have been obtained from the application of the plan for the seawaters off western Guangdong. The returned discriminant result of the water body classification reaches to a high compliance rate of 91.4%. Based on the automatic classification, thematic maps of water body distributions during the period from 2008 to 2010 have also been plotted for the seawaters off western Guangdong, which could provide a decision basis for the pollution control in this sea area.

Key words: sea area off western Guangdong; water quality classification; principal component analysis; discriminant analysis

Received: February 14, 2011