

# 基于 GIS 的区域水环境压力分区研究

张 培<sup>1,2</sup>, 田富姣<sup>3</sup>

(1. 郑州大学 环境政策规划评价研究中心, 郑州 450002;

2. 郑州大学 环境技术咨询工程公司, 郑州 450002; 3. 郑州大学 水利与环境学院, 郑州 450002)

**摘 要:**在参阅大量文献的基础上,界定了水环境压力分区的内涵,并从人口压力、社会经济压力、水资源压力、水环境污染压力和生态环境压力方面初步建立了水环境压力分区的指标体系;以漯河和平顶山为研究对象,运用 GIS 技术在提取研究区域子流域单元的基础上建立了子流域单元表征指标数据库,并通过相关分析方法选择出人口密度、单位面积水资源量和产业结构作为研究区域的分区指标;结合 GIS 技术,建立了研究区域 9 种聚类方案,并运用 PFS 方法选择出了最优聚类个数;在保证湖库节点完整性的基础上对最优聚类结果进行调整,最终将研究区域分为 6 类 11 个区,并分析了各分区的特点,为区域水环境管理目标与措施制定提供了科学依据和决策支持。

**关键词:**水环境压力; 分区; GIS; 指标体系; 聚类检验

中图分类号:F323.213; TP79

文献标识码:A

文章编号:1005-3409(2011)05-0014-05

## Regional Water Environmental Pressure Regionalization Based on GIS

ZHANG Pei<sup>1,2</sup>, TIAN Fu-jiao<sup>3</sup>

(1. Research Center for Environmental Policy & Planning & Assessment of Zhengzhou University, Zhengzhou 450002, China;

2. Zhengzhou University Environment Technology and Consulting Company, Zhengzhou 450002, China;

3. College of Water Conservancy & Environmental Engineering, Zhengzhou University, Zhengzhou 450002, China)

**Abstract:** The concept of water environmental pressure regionalization is defined on the basis of a large of references and the indicator system is set up basically from population pressure, social economy pressure, water resource pressure, water environmental pollution pressure, ecological environment pressure. The expression indicator database of the sub-watershed is established on the basis of sub-watershed extracted using GIS technology in Luohe and Pingdingshan as the research site, and then, population density, unit area water quantity and industrial structure are selected for regionalization indicators through the correlation analysis method. Nine kinds of clustering schemes of the study areas are built up with GIS technology, and the optimal cluster number is selected using PFS method. The optimal cluster scheme is adjusted in guarantee lake integrity basis, and finally, the study area is divided into six types of eleven regions, and the characteristic of each region is analyzed, which provides the scientific basis and decision support for making regional water environmental management objectives and measures.

**Key words:** water environmental pressure; regionalization; GIS; indicator system; clustering test

水环境压力是一个作用因素繁多、内涵十分复杂的概念,目前国内外对其鲜有专门的研究。袁洪锋认为水环境压力是水环境影响于人口、社会与经济持续发展的阻力,形成水环境压力的因素有自然因素和非自然因素<sup>[1]</sup>。顾晓薇等人定义环境压力为人类活动引起的能够造成环境服务功能退化的对环境状态的

扰动力<sup>[2]</sup>;王恒认为环境压力是人类的各种活动对其周围的环境所施加的影响力,这种影响力介入、扰乱或改变了宏观和微观的自然过程,致使环境状态发生“非自然”变化,引起环境功能的退化,对人类的生存、生活条件 and 健康带来不利影响,进而制约社会、经济的发展<sup>[3]</sup>。结合环境压力的内涵,笔者认为水环境压

力是人类社会经济活动对水环境所施加的影响,这种影响包括水资源的过度开发、水体的污染、水环境系统功能的退化等。

GIS 技术作为一种可视化的数据采集、存储、分析、管理工具,已经被广泛地应用到分区规划中,为分区规划进一步科学化、精确化和合理化提供了可能,同时 GIS 技术强大的空间分析功能,极大地增强了分区规划的科学性、目的性和现实意义。水环境压力分区是基于人类活动对水环境所产生的影响而提出的,目的是识别出不同区域水环境压力特征。本研究以漯河和平顶山地区(以下简称研究区域)为研究对象,运用 GIS 技术建立研究区域子流域单元表征指标数据库,并应用分区指标对研究区域进行水环境压力分区,识别不同区域水环境压力状况,为区域水环境

管理目标与措施制定提供科学依据和决策支持<sup>[4-5]</sup>。

1 水环境压力分区指标体系

指标体系的确定应该尽可能体现水环境压力分区的目的与特色,并反映出不同区域水环境压力的空间差异性规律。水环境压力分区应该遵循流域完整性、综合性与主导性、相似性与差异性以及地域发生学与共轭性原则<sup>[6]</sup>,指标在选取的过程中应遵循科学性、全面性、独立性、代表性、适用性、简明性和可操作性的原则,采用理论推导法、专家选取法和文献频数法从人口压力要素、社会经济压力要素、水资源压力要素、水环境污染压力要素和生态环境压力要素等方面进行分析,选取表征各要素层的指标建立区域水环境压力分区指标体系,如表 1 所示。

表 1 水环境压力分区指标体系

要素层	表征指标层	计算方法
人口压力	人口密度	区域人口/区域面积
	人口增长率	区域年增长人口/区域总人口
社会经济压力	人均 GDP	区域 GDP/区域总人口
	产业结构	区域第一产业生产总值/区域第二产业生产总值
	城市化率	城镇人口/区域总人口
水资源压力	人均水资源量	区域水资源总量/区域总人口
	单位面积水资源量	区域水资源总量/区域面积
	单位 GDP 耗水量	区域用水量/区域 GDP
	用水负荷	区域用水量/区域水资源量
水环境污染压力	废水排放强度 <sup>[7]</sup>	区域废水排放量/区域面积
	单位水资源 COD 排放量	区域 COD 排放量/区域水资源量
	化肥施用强度	区域化肥施用量/区域面积
生态环境压力	土地垦殖率	区域耕地面积/区域面积
		$L_a = 100 \times \sum A_i \times C_i, L_a \in (100, 400)$
	土地利用程度综合指数 <sup>[8]</sup>	式中: $L_a$ ——土地利用程度综合指数; $A_i$ ——第 $i$ 级的土地利用程度分级指数; $C_i$ ——第 $i$ 级土地利用程度面积百分比

2 研究区域与研究方法

2.1 研究区概况

研究区域位于河南省中南部,北纬 33°08′—33°59′,东经 112°14′—114°16′,总面积 10 499 km<sup>2</sup>。东部和中部以平原为主,西部和西北部以山地为主,自西北向东南方向倾斜,全境最高海拔 2 153 m,最低 50.1 m,地形地貌多样。2009 年研究区域总人口 761.77 万人,人口密度为 725.56 人/km<sup>2</sup>,城镇化率为 40.93%,人均 GDP 为 22 555.36 元。研究区域全年水资源总量 22.087 亿 m<sup>3</sup>,总用水量 14.060 3 亿

m<sup>3</sup>,人均用水量 184.7 m<sup>3</sup>,单位 GDP 用水量 81.83 m<sup>3</sup>/万元。2009 年研究区域全年废水排放量 30 584.3 万 t,COD 排放量 49 926.02 t。

2.2 研究方法及技术路线

在大量资料收集、文献阅读和现场调研的基础上,以 SRTM 30 m 分辨率的高程图为基础,在 ARC-GIS 软件的支撑下对研究区域进行流域提取,获取研究区域子流域单元;通过数字化研究区域行政区划图将收集到的数据分配到各子流域单元,建立子流域单元表征指标数据库,形成表征指标专题图;对表征指标进行相关分析选取具有代表性的指标参与分区,将

分区指标专题图进行叠加,运用 ERDAS 软件将叠加结果聚为 2—10 类,运用伪  $F$  检验方法选出最佳聚类数,并参考研究区域湖库节点完整性对聚类结果进行调整,最终形成研究区域水环境压力分区结果。研究区域水环境压力分区技术路线如图 1 所示。

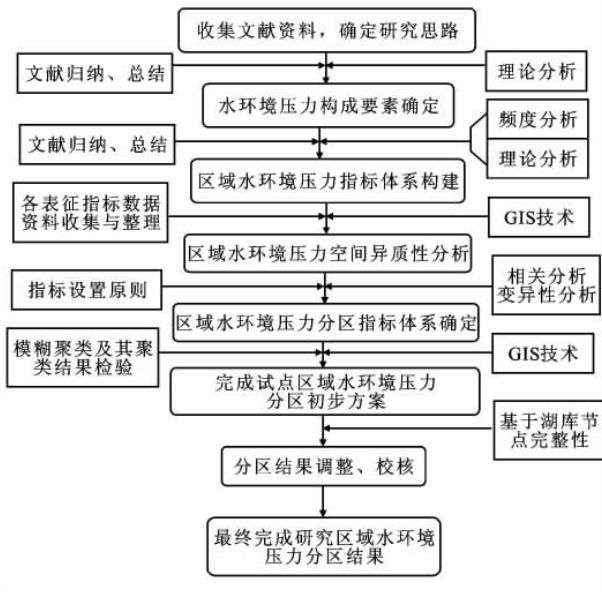


图 1 研究区域水环境压力分区技术路线

表 2 表征指标的相关性分析

因子	$I_1$	$I_2$	$I_3$	$I_4$	$I_5$	$I_6$	$I_7$	$I_8$	$I_9$	$I_{10}$	$I_{11}$	$I_{12}$	$I_{13}$	$I_{14}$
$I_1$	1.000													
$I_2$	0.008	1.000												
$I_3$	0.119	-0.246	1.000											
$I_4$	-0.170	0.378	-0.834*	1.000										
$I_5$	0.593*	0.179	0.323	-0.306	1.000									
$I_6$	-0.801*	0.261	0.100	-0.102	-0.068	1.000								
$I_7$	0.027	0.443*	0.542*	-0.314	0.495*	0.429	1.000							
$I_8$	-0.297	0.239	-0.723*	0.448*	-0.460*	0.116	-0.461*	1.000						
$I_9$	0.695*	-0.160	0.115	-0.208	0.172	-0.788*	-0.289	0.168	1.000					
$I_{10}$	0.773*	0.024	0.179	-0.120	0.724*	-0.353	0.388	-0.358	0.362	1.000				
$I_{11}$	0.824*	-0.021	0.123	-0.034	0.596*	-0.494*	0.270	-0.338	0.411	0.965*	1.000			
$I_{12}$	0.725*	-0.354	0.196	-0.001	0.280	-0.773*	-0.135	-0.319	0.564*	0.520*	0.640*	1.000		
$I_{13}$	0.774*	0.132	0.122	0.196	0.111	-0.849*	-0.046	-0.227	0.612*	0.363	0.498*	0.703*	1.000	
$I_{14}$	0.670*	-0.006	0.238	-0.010	0.213	-0.685*	0.068	-0.380	0.479*	0.384	0.439	0.536*	0.763*	1.000

注: \* 表示 0.01 水平显著。 $I_1$ — $I_{14}$  分别为人口密度、人口增长率、人均 GDP、产业结构、城市化率、人均水资源量、单位面积水资源量、单位 GDP 耗水量、用水负荷、废水排放强度、单位水资源 COD 排放量、化肥施用强度、土地垦殖率、土地利用程度综合指数。

从表 2 中可以看出,人口密度指标综合性最强,富含信息量最大,其与城市化率、人均水资源量、用水负荷、废水排放强度、单位水资源 COD 排放量、化肥施用强度、土地垦殖率、土地利用程度综合指数等 8

### 3 结果与分析

#### 3.1 相关分析

由于各表征指标的单位不同,不同指标之间没有可比性,因此在进行相关分析和聚类分析前需要对不同量纲指标的初始数据进行标准化处理,把所有指标数据转换成统一的含义。表征指标标准化量化公式为<sup>[9-10]</sup>

$$C_i = (X_i - \min X_i) / (\max X_i - \min X_i)$$

式中: $C_i$ ——第  $i$  个指标标准化后的值; $X_i$ ——第  $i$  个指标的原始值; $\max X_i$ ——第  $i$  个指标原始值中的最大值; $\min X_i$ ——第  $i$  个指标原始值中的最小值。

选用多个表征指标对研究区域进行水环境压力分区能够尽可能全面地表达出区域所受到的各种压力影响,但并非指标越多越好。由于一些表征指标之间存在一定的重复性,会造成水环境压力分区过程中的冗余,因此本研究利用相关分析对表征指标进行筛选。相关系数  $r$  表示两个指标之间的相关性, $-1 \leq r \leq 1$ , $r$  的绝对值越接近 1,说明两个指标之间的相关性越强。本研究运用 SPSS 软件对量化后的子流域单元表征指标数据进行相关分析,结果如表 2 所示。

个指标都有非常强的相关性,其相关系数基本上都在 0.6 以上,说明人口密度指标能较好地综合反映各相关指标所表达的信息量,具有很强的代表性,故选取人口密度作为分区指标,则与其相关性强的指标不再

考虑;剩余的指标有人口增长率、人均GDP、产业结构、单位面积水资源量和单位GDP耗水量,这5个指标中,单位面积水资源量和单位GDP耗水量这两个指标综合性比较强,富含信息量比较大,单位面积水资源量与人口增长率、人均GDP、城市化率、单位GDP耗水量等4个指标具有很强的相关性,单位GDP耗水量与人均GDP、产业结构、城市化率、单位面积水资源量等4个指标具有很强的相关性,同时这两个指标之间又具有很强的相关性,因此需要从这两个指标中选取其一。

从表2中可以看出,这两个指标的综合性相似,且与其相关指标的相关程度也大体一致,本研究运用变异性分析来筛选确定,变异系数越大,说明空间差异特征越明显,因此选择变异系数大的指标。变异系数=标准差/均值,通过计算得知单位面积水资源量的变异系数为0.88,单位GDP耗水量的变异系数为0.76,故选取单位面积水资源量指标作为分区指标,则与其相关的指标不再选取;最后剩余产业结构一个指标,将其纳入到分区指标中。综上所述,最终确定的分区指标是:人口密度、单位面积水资源量和产业结构。

### 3.2 聚类分析

运用ERDAS软件将研究区域子流域单元分区指标专题图进行叠加,基于非监督分类方法(ISODATA)将研究区域分区2~10类,如附图1所示。

尽管运用ERDAS软件将研究区域划分为2~10类,但是无法确定哪一种分区结果为最好的分类方案,因此需要选择一种检验方法以从众多的分类方案中选择最好的一种。很多文献对最佳聚类个数的确定,提出了相应的检验指标以确定聚类个数在局部或整体最优,用于检验判断最佳聚类个数的指标有 $R^2$ 统计量、半偏相关统计量、伪F统计量(PFS)、伪 $t^2$ 统计量等<sup>[11]</sup>。本研究应用伪F统计量(PFS)来确定研究区域水环境压力分区最佳聚类方案。伪F统计量计算公式为

$$PFS = \frac{(T - P_G) / (G - 1)}{P_G / (n - G)}$$

式中: $G$ ——聚类个数; $n$ ——聚类单元数; $P_G$ ——分类数为 $G$ 个类时的总类内离差平方和; $T$ ——所有变量的总离差平方和。

伪F统计量评价分为 $G$ 个类的效果,如果分为 $G$ 个类合理,则类内离差平方和(分母)应该较小,类间平方和(分子)相对较大。考虑PFS值随着分类 $G$ 值的增大, $P_G$ 总是趋于下降,而 $(T - P_G)$ 总是趋于上

升,但这并不使PFS值始终上升,因为随着 $G$ 值的增大,比值 $(n - G) / (G - 1)$ 却不断下降。因此可以预知,在规定的聚类数范围内,PFS可能在某一 $G$ 值处达到最大值,而这个最大值正是我们所要寻找的最佳分类数。

根据伪F统计量的计算公式分别对这9种聚类方案的三个分区指标计算其F值,结果如图2所示。

对一个样本点数为 $n$ 的数据对象可能有的最佳聚类个数上限,很多研究者使用经验规则为 $\sqrt{n}$ 或 $2\ln n$ 下取整<sup>[12]</sup>。虽然此结论无明显的理论指导,但很多专家已在使用,也有间接旁证文献和实例应用<sup>[13]</sup>。根据此最佳聚类数取值经验规则,本研究子流域单元个数为33个,其最佳分类数目的取值范围应在2~6类,从图2中可以看出,随着聚类个数的增多,分区指标PFS值加权平均后的值逐渐增大至8类后开始降低,8类处为一个最大值点,同时6类处为一个拐点。结合经验公式及图2,确定本研究区域最佳聚类个数为6,因此将研究区域聚为6类。

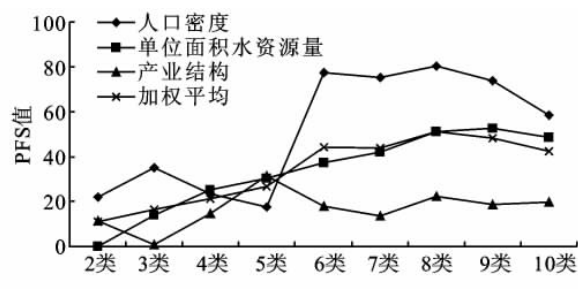


图2 各种聚类方案分区指标 PFS 值趋势

### 3.3 聚类结果调整

从附图1中可以看出,研究区域水环境压力6类分区方案中研究区域被分成11部分。同时,研究区域内包含三个较大的水库,分别为白龟山水库、昭平台水库和燕山水库,一个完整的湖库具有相似的生态系统特征,因此不能将其分裂开,在进行分区时应保证湖库节点的完整性,即要保证一个湖库完全位于一个分区内。在现有的6类分区方案中,白龟山水库跨越16号、29号和30号子流域单元,也即是横跨3个分区,因此需要对这3个分区进行调整,调整的原则是将白龟山水库划入其落入面积最大的区域内,也即是将其划到16号子流域单元中;昭平台水库跨越11号、25号和30号子流域单元,横跨两个分区,按照同样的原则,将昭平台水库划到11号子流域单元所在的分区;燕山水库完全落于27号子流域单元内,因此不需要对其进行调整。经过调整后的研究区域水环境压力分区结果如附图2所示。

### 3.4 结果分析

研究区域水环境压力各分区特征如表3所示。

表 3 研究区域水环境压力分区结果分析

分区名称	包含区域	分区特征
分区 I	位于平顶山市汝州市境内	以山地为主,水资源压力对水环境的影响最大,居于六区之首,其次为产业结构的影响
分区 II	位于平顶山市鲁山县境内	以山地为主,人口压力对水环境的影响很小,主要是水资源压力对水环境的影响较重
分区 III	包括平顶山市郏县、宝丰县和漯河市的临颍县、舞阳县及鄆城区	水资源匮乏,对水环境的影响较重,同时人口压力对水环境的影响较重
分区 IV	包括平顶山市鲁山县、石龙区和舞钢市	水资源充足,但是产业结构对水环境的影响较重
分区 V	包括平顶山市区,漯河市源汇区、召陵区、鄆城区部分区域	为城市所在地,水环境压力较重,尤其是人口压力对水环境产生的影响较大,其次为水资源量的影响
分区 VI	主要位于平顶山市叶县,包括漯河市舞阳县部分区域	人口压力对水环境的影响比较严重

4 结 论

本研究选择漯河、平顶山地区作为研究案例,主要是由于这两个地区无论从地形地貌、经济发展,还是从污染物排放、生态建设等方面都有着很大的差异性,有利于更好地研究人类活动对水环境产生的压力。

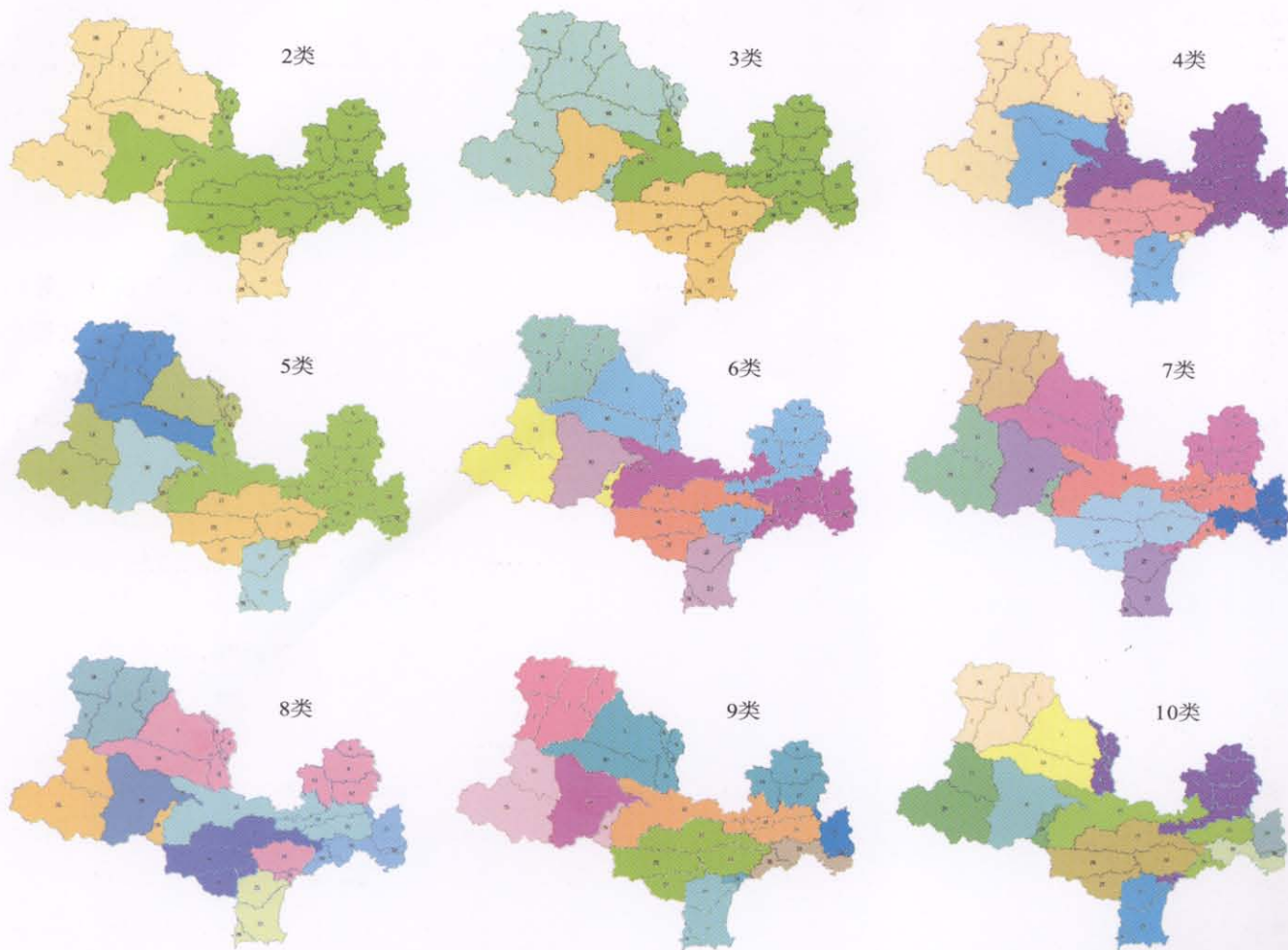
本研究通过参阅大量的文献,初步建立研究区域的指标体系,在运用地理信息系统提取研究区域子流域单元的基础上建立表征指标数据库,并运用相关分析方法结合变异分析选择出人口密度、单位面积水资源量和产业结构作为研究区域的分区指标;在 GIS 和 RS 的辅助下,建立研究区域的聚类方案,并运用 PFS 方法选择出了最优聚类数目,同时在保证湖库节点完整性的基础上对最优聚类结果进行了调整,最终将研究区域分区 6 类 11 个区,并阐述了各分区的特点。

本研究尝试运用 GIS 技术对研究区域进行水环境压力分区,并取得了较好的效果。但是,在分区过程中发现,数据获取和子流域单元提取是一个关键点和难点,如何保证流域提取结果与现状相符及数据可得性是未来要研究的重点。

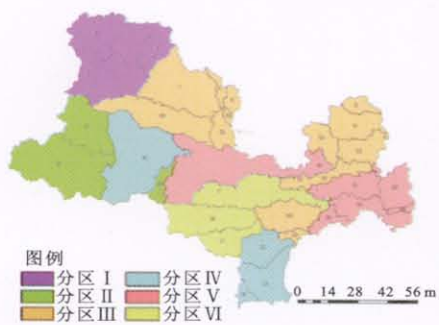
参考文献:

[1] 袁洪锋. 我国水环境压力及其特征分析[J]. 浙江大学学

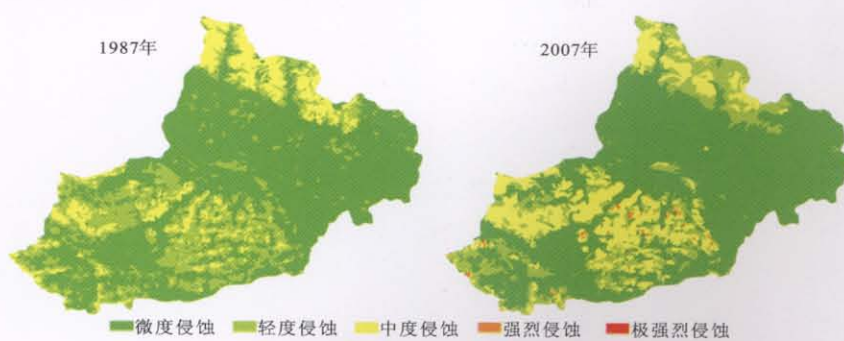
报:理学版,2002,29(2):219-224.  
[2] 顾晓薇,王青,刘敬智,等. 环境压力指标及应用[J]. 中国环境科学,2005,25(3):315-319.  
[3] 王恒. 四川省环境承载力与压力研究[D]. 成都:西南交通大学,2008.  
[4] 刘木生,林联盛,郭秋忠. 基于 GIS 的生态功能分区技术方法刍议[J]. 江西农业学报,2008,20(10):111-113.  
[5] 张海星,谭斌. 基于 GIS 方法的赣江流域水环境功能区划研究[J]. 江西科学,2003,21(1):26-29.  
[6] 尹民,杨志峰,崔保山. 中国河流生态水文分区初探[J]. 环境科学学报,2005,25(4):423-428.  
[7] 程道品,王金叶,胡新添,等. 广西生态环境质量的动态驱动机制[J]. 桂林工学院学报,2006,26(4):591-594.  
[8] 庄大方,刘纪远. 中国土地利用程度的区域分异模型研究[J]. 自然资源学报,1997,12(2):105-111.  
[9] 丁耕,吴伟. 基于 GIS 的东辽河流域生态安全空间差异评价研究[J]. 安全与环境学报,2006,6(3):119-122.  
[10] 王敏,熊丽君,黄沈发. 上海市主体功能区划分技术方法研究[J]. 环境科学研究,2008,21(4):205-209.  
[11] 李斌,郭剑毅. 一种带约束的最小离差平方和系统聚类法及应用[J]. 计算机应用,2005,25(1):45-48.  
[12] 吴艳文,胡学钢. 一种 K-means 算法的 k 值优化方案[J]. 巢湖学院学报,2007,9(6):21-24.  
[13] 于剑,程乾生. 模糊聚类方法中的最佳聚类数的搜索范围[J]. 中国科学:E 辑,2002,32(4):274-280.



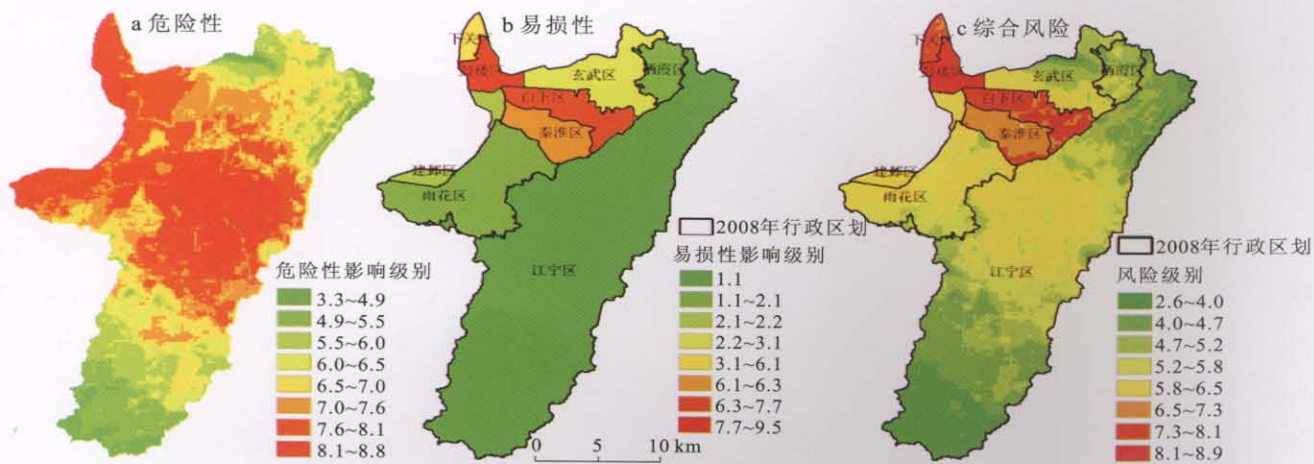
附图1 研究区域水环境压力分区方案 (每幅图同一种颜色代表一类)



附图2 研究区域水环境压力分区结果



附图3 山东省费县不同年份土壤侵蚀强度等级分布



附图4 2008年秦淮河中下游洪涝灾害危险性、易损性、综合风险区划图