

鄱阳湖区资源综合利用评价指标的形成方法

傅 春¹ 林永钦²

(1.南昌大学中国中部经济发展研究中心,江西 南昌 330047;2.南昌大学管理科学与工程系,江西 南昌 330047)

摘要:根据区域可持续发展理论,结合鄱阳湖区资源状况和社会经济发展的实际水平,综合运用物元理论、专家咨询法、频度统计法和 Rough 集理论,探索了一种构建鄱阳湖区资源综合利用的评价指标体系的方法,并引用信息熵的概念对其有效性作了初步分析。结果表明采用该方法形成的评价指标体系具有较强的完整性和有效性。

关键词:鄱阳湖区;资源开发;物元理论;Rough 集;评价指标

中图分类号: N94 文献标识码: A 文章编号: 1006-7647(2006)01-0037-03

Method for generation of evaluation index for comprehensive development and utilization of resources in Poyang Lake Region//FU Chun¹, LIN Yong-qin² 1. Research Center for Central China Economic Development, Nanchang University, Nanchang 330047, China; 2. Management Science and Engineering Department, Nanchang University, Nanchang 330047, China)

Abstract: In consideration of the current situation of natural resources and the reality of social and economic development in the Poyang Lake Region, a method for generation of the evaluation index system for comprehensive development and utilization of resources in the Poyang Lake Region was developed by use of the matter-element theory, the expert-consultation approach, the frequency statistics method, and the Rough set theory based on the regional sustainable development theory. With the concept of information entropy, the validity of the method was analyzed. The result shows that the evaluation index system formulated is of high integrity and effectiveness.

Key words: Poyang Lake Region; resource development; matter-element theory; Rough set; evaluation index

鄱阳湖流域面积 16.22 万 km²。区内水质优良,自然条件优越,资源丰富,昌九走廊、京九铁路、南昌、九江等江西省重要的经济增长带都坐落于该区,同时它还是国家重要的商品粮、棉、油基地,是国际最重要的湿地之一^[1]。然而,鄱阳湖滨湖 11 县市(包括南昌、新建、进贤、余干、波阳、都昌、湖口、星子、德安、永修 10 县和九江市,以下简称鄱阳湖区)的经济社会发展水平目前仍很落后。随着区内人口急增,大规模退田还湖后,由于土地资源的匮乏,水、土等自然资源被过度开发利用导致生态环境恶化,严重制约了该地区经济社会的可持续发展。为此,建立一套综合反映本区域资源开发利用状态和水平的、科学恰当的评价指标体系来指导鄱阳湖区资源的开发和利用,协调资源、环境和经济的发展,对鄱阳湖区的可持续发展有着重要的现实意义。

本文将尝试采用物元可拓方法和 Rough 集的属性约简方法,并同时结合其他方法,对鄱阳湖区资源综合利用评价指标的设置和筛选过程进行研究。

1 指标体系设计思路

运用物元理论^[2]中的发散-收敛的菱形思维过程,采用定性和定量相结合的方法提出一种新的指标体系设计思路,以增强指标体系设计的逻辑性和科学性。首先,利用物元理论进行分析,开拓出多个指标,并根据评价目的,结合采用频度统计法、专家咨询法,以形成能够比较全面地反映评价对象特征的、信息覆盖尽可能全的初始指标集;然后运用 Rough 集理论^[3]的属性约简方法进行筛选,形成能够满足科学性、独立性和可操作性原则的、包含较少指标且具有关键特征属性的核心指标体系。

2 初始指标集的形成

区域资源综合开发评价,不仅要考虑资源量的变化,也要考虑质的变化,还要考虑资源开发与区域环境、人口条件的协调一致。物元可拓性理论^[2]把对量变的研究和对质变的研究结合起来,为区域资

源综合开发评价指标体系的研究提供了一种新的思路。根据区域可持续发展的内涵^[4 5]以及鄱阳湖区资源、经济、社会等具体特点,结合专家咨询和频度统计法设置初始指标集。

由于资源具有环境的特征,因而把资源、环境合并为资源环境领域,评价指标体系归类分成经济、社会、资源环境和人口四个领域。另外,为利于比较分析,要求指标通用性较强,评价对象具有同质性。显然,作为地级城市的九江市与其他 10 个以农业为主的县域不同质,应不予包括在指标体系中。指标体系的基本结构确定为矩阵式,矩阵的行表示评价领域,列表示评价对象,即鄱阳湖区所包含的各个县域。在资源环境领域,选择人均土地面积、年平均降水量、森林覆盖率、年平均日照时数、年平均气温、人均水资源量、每公顷水产品量反映资源存量和质量,选择人均耕地面积、人均淡水养殖面积、人均耗水量反映开发利用现状。

在人口领域,人口与环境资源的协调要求在量上把握相对于资源环境的合理度,在质上具有相对于技术与生产力发展的适应和超前,在结构上具有相对于综合开发水平的合理与稳定。为此,选择人口自然增长率、人口密度反映人口规模,选择非农业人口占总人口比例、从业人员占总人口比例、人均科学事业费支出来反映综合开发的人口条件及其制约程度。在经济领域,以人均 GDP 反映综合开发所产生的经济效益,以第二、三产业占 GDP 比重及 GDP 增长率反映资源开发对经济的制约及适应程度。在社会领域,以职工年人均工资、农民年人均纯收入和公路密度来反映综合开发所产生的社会效益、社会条件及其制约程度。鉴于篇幅所限,仅列出口人领域的指标数据,如表 1 所示。

表 1 人口领域指标数据

区 域	人口自然 增长率 $f_1/\%$	人口密度 $f_2/$ (人·km ⁻²)	非农业 人口比例 $f_3/\%$	从业人员 比例 $f_4/\%$	人均科学 事业费支出 $f_5/(元·年^{-1})$
南昌县 u_1	1.478	560	14.9	44.0	0.25
新建县 u_2	0.697	280	17.6	47.1	0.69
进贤县 u_3	0.966	369	16.5	46.8	0.04
余干县 u_4	0.566	367	7.3	45.3	0.12
波阳县 u_5	0.910	321	9.0	42.9	3.61
都昌县 u_6	0.917	357	10.0	47.5	0.13
湖口县 u_7	0.475	385	15.5	43.9	0.30
星子县 u_8	0.980	325	13.4	40.7	0.13
德安县 u_9	0.864	170	34.1	42.5	2.09
永修县 u_{10}	0.587	174	22.8	40.7	0.79

注:数据来源于江西省和有关地市的统计年鉴以及江西省江湖委办公室、省统计局、省林业厅等单位的研究资料(2001 年数据)。

3 核心指标体系的形成

依据物元可拓方法形成的初始指标集,无法满足

实用性、独立性原则,必须作进一步的定量筛选。采用 Rough 集的属性约简方法,可提取出核心指标体系。

3.1 实值属性离散化

Rough 集理论处理数据用的是符号化的分析方法,其属性值“0”;“1”;“2”等都是离散的符号。但是前述指标数据(如表 1)都是连续性的实数,因而必须先离散化。对连续型属性进行离散化(即对取值区间进行划分,然后进行编码)的主要问题是确定区间的个数。现有的分组方法大都是根据经验或实际问题的需要事先选定区间的个数,带有一定的主观性,如等距离散法、等频离散法都需事先确定区间(或分类)个数。

本文采用一种基于 Ward 法的系统聚类^[6 7]的离散化方法,因为本文对信息系统表(表 1)的指标属性是逐一处理的,所以样本(各区域)都是一维变量,它们之间的距离定义为

$$d(X,Y)=\frac{|X-Y|^2}{2} \tag{1}$$

若以 $\bar{X}_K$ 和 $\bar{X}_L$ 分别表示类 C_K 和 C_L 中的均值,以 D_{KL} 表示第 K 类 C_K 与第 L 类 C_L 之间的距离, N_K 和 N_L 分别表示类 C_K 和 C_L 中的样本数,则

$$D_{KL}=\frac{|\bar{X}_K-\bar{X}_L|^2}{1/N_K+1/N_L} \tag{2}$$

类与类之间距离计算的递推公式如下:

$$D_{JM}=\frac{(N_J+N_K)D_{JK}+(N_J+N_L)D_{JL}-N_JD_{KL}}{N_J+N_M} \tag{3}$$

式中: D_{JM} 表示由类 C_K 和 C_L 合并的新类 C_M 与类 C_J 的距离; D_{JK} 表示类 C_J 与类 C_K 的距离; D_{KL} 表示类 C_K 与类 C_L 的距离; N_M 表示类 C_M 中的样本个数,则 $N_M=N_K+N_J$ 。

本方法首先将样品各自分成一类,然后每次缩小一类,每次合并时(例如把 C_K 和 C_L 合并为 C_M 时)选择这样两个类,使得合并(为 C_M)后,类内离差平方和的增量为最小。

为了客观确定聚类个数(即区间个数),选用 R^2 统计量和伪 t^2 统计量作为判断标准,两个统计量的定义如下:

$$R^2=1-(P_G/T) \tag{4}$$

$$\text{伪 } t^2=\frac{B_{KL}}{(W_K+W_L)(N_K+N_L-2)} \tag{5}$$

式中: G 表示类的个数; P_G 表示 G 个类的类内离差平方和的总和; T 表示所有样本的总离差平方和; B_{KL} 表示合并类 C_K 和 C_L 后类内离差平方和的增量; W_K 和 W_L 分别表示类 C_K 和类 C_L 中样本的类内离差平方和。

R^2 值越大,表示类内离差平方和在总离差平方和中占的比例越小,即 G 个类越分开,故聚类效果好,伪 t^2 值大说明合并 C_K 和 C_L 后,离差平方和的增加相对于 C_K 和 C_L 的类内离差平方和要大,表明该次合并的两个类 C_K 和 C_L 是很分开的,也就是上一次聚类的效果是好的。

实际应用时常寻找两种统计量之间的一致性,即 R^2 统计量的较大下降与伪 t^2 统计量的较大增加相吻合,以此确定合适的聚类个数。

以人口自然增长率 f_1 为例(其他类同),系统聚类结果如表 2 所示。

表 2 指标 f_1 的系统聚类结果

类	合并的类	类中样本个数	R^2	伪 t^2
C_9	波阳县,都昌县	2	1.000	0.0
C_8	进贤县,星子县	2	1.000	0.0
C_7	余干县,永修县	2	1.000	0.0
C_6	C_9 ,德安县	3	0.997	66.7
C_5	C_7 ,湖口县	3	0.988	31.1
C_4	C_8 , C_6	5	0.979	11.8
C_3	新建县, C_5	4	0.955	5.0
C_2	南昌县, C_4	6	0.617	116.0
C_1	C_2 , C_3	10	0.000	12.9

由表 2 后两栏可知,当聚类个数 G 由 3 到 2 时, R^2 统计量下降较大(由 0.955 下降为 0.617),而且伪 t^2 统计量增加很多并在 $G=2$ 时达最大值(由 5.0 增加为 116.0),这说明了 $G=3$ 的聚类效果很好,因而应把样本分为 3 类。由表 2 前两栏可看出:南昌单独成为一类,其离散值用符号“0”表示;新建与类 C_5 合并,即新建、湖口、余干、永修归为一类,用符号“1”表示;类 C_8 与类 C_6 合并,即进贤、星子、德安、波阳、都昌归为一类,用符号“2”表示。类似地,可以求得表 1 其他指标的离散化结果(见表 3)。

表 3 人口领域离散化结果

区 域	f_1	f_2	f_3	f_4	f_5
南昌县 u_1	0	0	0	0	0
新建县 u_2	1	1	0	1	1
进贤县 u_3	2	1	0	1	0
余干县 u_4	1	1	1	0	0
波阳县 u_5	2	1	1	0	2
都昌县 u_6	2	1	1	1	0
湖口县 u_7	1	1	0	0	0
星子县 u_8	2	1	0	2	0
德安县 u_9	2	2	2	0	3
永修县 u_{10}	1	2	3	2	1

3.2 约简过程

采用简化分明矩阵的算法^[3]对离散化后的人口领域(表 3)指标进行约简,运算过程为:

a. u_1 与 $u_2 \sim u_{10}$ 关于属性值是分明的属性直接从表 3 中提取并构成分明合取范式 F_1 ,再约简得

$$F_1=(f_1 \vee f_2 \vee f_4 \vee f_5) \wedge (f_1 \vee f_2 \vee f_4) \wedge \dots \wedge (f_1 \vee f_2 \vee f_3 \vee f_4 \vee f_5)=f_1 \vee f_2 \quad (6)$$

b. u_2 与 $u_3 \sim u_{10}$ 关于属性值是分明的属性直接从表 3 中提取并构成分明合取范式,然后加上式(6)得 F_2 并约简得

$$F_2=(f_1 \vee f_5) \wedge (f_3 \vee f_4 \vee f_5) \wedge \dots \wedge (f_2 \vee f_3 \vee f_4) \wedge (f_1 \vee f_2)=f_1 \wedge f_4 \quad (7)$$

c. 同理,可以分别提取 $u_3 \sim u_9$ 关于后续属性值是分明的属性构成分明合取范式并加上前一步的结果,再约简得

$$f_1 \wedge f_3 \wedge f_4 \quad (8)$$

d. 对式(8)施行“ $\wedge$ ”对“ $\vee$ ”的分配律运算(这里无需做此运算),则得到表 3 的约简结果: $R=\{f_1, f_3, f_4\}$,即把人口领域的指标由原来的 5 个约简为 3 个,分别是人口自然增长率、非农业人口占总人口比例、从业人员占总人口比例。

类似过程,可求得评价核心指标体系如图 1 所示。

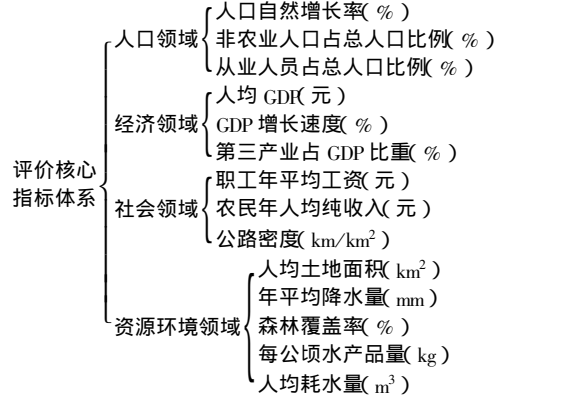


图 1 评价核心指标体系

4 进一步讨论

前面的约简过程由离散化和属性约简组成。运用 Rough 集的属性约简方法所删除的属性(即指标)对原有系统的分类不构成影响^[3],不会造成必要信息量的损失。但是在离散化这一步骤上有可能会造成必要信息量的损失。本文采用的离散化方法保证了信息的有效性和完整性,具体为:引用信息论中的“熵”来测度信息量的大小。设系统有 m 个对象, x_i 为指标 f_j 原始数据,令 x_j^0 为某个最优常数(值 $x_j^- \leq x_j^0 \leq x_j^+$),则离散化前熵的计算步骤是:①对指标进行标准化

$$\text{转换(极差法)}:y_j=\frac{\max_{1 \leq i \leq m} |x_i-x_j^0|}{\max_{1 \leq i \leq m} |x_i-x_j^0|} \quad \text{②进行归一化处理,得 } p_i=y_i/\sum y_i \quad \text{③计算熵值:} e=-\left(\sum p_i \ln p_i\right)/\ln m。$$

(下转第 53 页)

3520 m³/d;下水库渗漏库段集中于主副坝段,渗漏
量小,处理措施较简单。水库防渗处理措施应根据
库周具体工程地质及水文地质条件,针对库周不同
地段的渗漏特点确定。

a. 坡积土及全风化带具弱至中等透水性,上部
部分为中等透水性,中下部具弱透水性,为相对隔水
层,具有一定的防渗作用。因此,应充分利用其作为
水平防渗铺盖,以减少防渗处理措施。

b. 主副坝址宜采用截渗墙结合帷幕灌浆做垂
直防渗体,垂直防渗体深度应深入弱透水性岩层中
透水率小于或等于 3 Lu 以下 5m;主坝址上游侧还需
作水平防渗铺盖。

c. 低矮垭口和单薄分水岭可采用帷幕灌浆做
垂直防渗体处理,并与相邻主副坝防渗体连成一体
或延伸至相邻地下分水岭高于正常蓄水位处。

d. 对于地下水位略低于正常蓄水位的低矮垭
口,可设置地下水位长期观测孔进行观测,暂不进行
防渗处理,根据地下水位长期观测结果再采取针对
性处理措施。

e. 上、下水库库盆底地层岩性以花岗岩和混合
岩为主,为非可溶性不透水岩石,且不存在集中渗漏
管道,因此无需对库盆底进行水平防渗处理。

(上接第 39 页)

以“非农业人口占总人口比例”指标为例,离散
化前熵值 $e_1 = 0.9584$,离散化后熵值 $e_2 = 0.9599$,熵
值增加(即信息量减少)了 0.16%。这证明了所采
用的离散化方法的有效性。

至于整个系统经约简后总的信息量损失的测度
问题,还有待进一步的探讨。

参考文献:

[1] 甘筱青.鄱阳湖区资源综合利用与社会可持续发展[J].
南昌大学学报 理科版 2002 26(4) 328-330.
[2] 蔡文 杨春燕 林伟初.可拓工程方法[M].北京:科学出
版社,1997:16-129.
[3] 刘清. Rough 集及 Rough 推理[M].北京:科学出版社,
2001:1-75.
[4] 李金华.中国可持续发展核算体系[M].北京:社会科学
文献出版社,2000:1-28 234-248.
[5] 朱启贵.可持续发展评估[M].上海:上海财经大学出版
社,1999:1-48.
[6] 涂汉生,何平,赵联文.应用统计[M].重庆:西南交通大
学出版社,1994:177-186.
[7] 高惠璇,耿直,李贵斌.SAS 系统 SAS/STAT 软件使用手册
[M].北京:中国统计出版社,1997:630-658.

(收稿日期 2005-04-12 编辑 熊水斌)

4 结 语

抽水蓄能电站水库蓄水后库水反复循环使用,
为保证电站正常运行,库区不允许库水发生永久性
渗漏,因此对库区进行库水外渗条件的专题调查和
分析显得非常重要。惠州抽水蓄能电站上、下水库
库周以雄厚山体为主,仅主副坝址、单薄分水岭及低
矮垭口地形低矮,存在渗漏问题。库水外渗主要沿
通往库外的断裂构造与强风化带连通形成的渗漏通
道渗漏,而库周坡积土及全风化带具相对隔水作用,
充分利用其作为水平防渗铺盖可以降低库水外渗
量,减少防渗处理措施,对渗漏库段采用以垂直防渗
体为主的防渗措施,主坝址上游侧结合水平防渗铺
盖是能够满足防渗要求的。总体上看,惠州抽水蓄
能电站上、下水库库容大,防渗处理工程量不太大,
是工程地质条件较好的天然库盆。

参考文献:

[1] 陆兆溱.工程地质学[M].北京:水利电力出版社,1989:
246-379.
[2] 王贵军,刘国玉.回龙抽水蓄能电站上库渗流计算及防
渗方案选择[J].小水电,2003(6):7-10.

(收稿日期 2004-10-09 编辑 熊水斌)

(上接第 45 页)

3 大型渠道衬砌设备在我国的发展趋势

a. 滚筒衬砌设备 逐渐向履带式滚筒衬砌设备
发展,皮带布料机将被螺旋布料机代替,衬砌机的衬
砌小车行走由链传动牵引改为自行机构。

b. 滑模式衬砌设备 振捣、输送布料部分的动
力采用电驱动,行走、升降采用液压传动。这样可充
分降低设备的体积和重量,使用起来更为灵活。

c. 无论是滑模衬砌设备还是滚筒衬砌设备均朝
着智能化和模块化方向发展。改变其装配形式可满足
不同坡长和坡度的渠道衬砌及平面混凝土的衬砌要求。

参考文献:

[1] 商青.南水北调工程所需主要机械设备简介(二)[N].中
国水利报,2003-4-24(5).
[2] 傅智.水泥混凝土路面滑模施工技术[M].北京:人民交
通出版社,2000:24-28.
[3] 霍兵.渠道机械化衬砌设备及应用前景[J].中国机电工
业,2004(1):31-32.
[4] 乔世珊,胡胜利.南水北调中线关键机械设备的探讨
[J].中国水利,2003(9):58-59.
[5] 聂生勇.山东济平干渠积极进行机械化衬砌试验[N].中
国水利报,2003-12-11(5).

(收稿日期 2004-11-29 编辑 熊水斌)