

基于信息熵与 Theil 不等系数的 水资源可再生能力综合评价

龚艳冰, 房道伟, 张继国

(河海大学商学院, 江苏 常州 213022)

摘要: 为解决实际水资源可再生能力各项评价指标的不相容问题, 提出一种基于信息熵的 Theil 不等系数多属性评价方法。用该方法对黄河流域 9 个行政分区的水资源可再生能力进行综合评价。结果表明, 四川省在黄河流域 9 个行政区中的水资源再生能力最强; 内蒙古、宁夏地区的水资源可再生能力较弱。这一结果与用加权物元法和投影寻踪法得到的结果基本一致。

关键词: 信息熵; Theil 不等系数; 水资源可再生能力

中图分类号: C934 文献标识码: A 文章编号: 1003-9511(2009)03-0009-03

水资源具有可再生性, 这决定了水资源具有可再生能力。水资源可再生能力是指某一流域或区域环境, 在现有或近期科学技术和社会经济能力支撑下, 通过水自然循环与社会营造, 循环利用水资源的能力。对水资源可再生能力进行研究, 将为水资源可持续开发利用提供理论基础和决策依据^[1-3]。为了科学地评价黄河流域水资源可再生能力, 沈珍瑶等^[4]建立了黄河流域水资源可再生性评价指标体系, 并利用灰关联分析方法与模糊综合评判法对其进行了评价。杨晓华等^[5-7]分别提出了水资源可再生能力综合评价的遗传投影寻踪方法、遗传加权物元模型; 陈守煜等^[8]提出基于可变模糊集理论的水资源可再生能力评价模型。

由于影响水资源可再生能力的因素既有天然的, 又有人工的, 因此, 对水资源可再生能力进行综合评价就是一个多指标决策问题^[9]。由于实际多指标决策问题各项评价指标常常是不相容的, 因此, 各指标权重的确定较为困难。为此, 本文提出一种新的基于信息熵的 Theil 不等系数决策方法。该方法具有如下特点: ①由于评价水资源可再生能力的指标很多, 各指标的权重由熵权确定, 可避免人为因素的影响; ②通过计算各个方案与正、负理想方案之间的 Theil 不等系数, 可以充分利用决策矩阵提供的信息。通过对黄河流域 9 个行政分区的水资源可再生能力的客观、科学的评价, 评价结果将为该地区今后的水资源利用和保护的可持续发展起到一定的指导

作用。

1 基于熵权的 Theil 不等系数决策方法

设多属性决策问题的方案集为 $A = \{A_1, A_2, \dots, A_n\}$, 属性集为 $R = \{R_1, R_2, \dots, R_m\}$, 方案 A_i 相对于属性 R_j 的属性值用矩阵 $Y = (y_{ij})_{n \times m}$ 表示。通常属性有“效益型”、“成本型”、“固定型”和“区间型”几种类型。所谓效益型指标是指属性值愈大愈好的指标; 所谓成本型指标是指属性值愈小愈好的指标; 所谓固定型指标是指属性值既不能太大又不能太小, 而以稳定在某个固定值为最佳的一类指标; 所谓区间型指标是指属性值以落在某个固定区间内为最佳的一类指标。为了消除不同物理量纲对决策结果的影响, 决策时可按下列公式对决策矩阵进行规范化处理。

对于效益型和成本型指标, 一般可分别令

$$z_{ij} = \frac{y_{ij}}{\max_i(y_{ij})}$$

$$(i = 1, 2, \dots, n; j = 1, 2, \dots, m) \quad (1)$$

$$z_{ij} = \frac{\min_i(y_{ij})}{y_{ij}}$$

$$(i = 1, 2, \dots, n; j = 1, 2, \dots, m) \quad (2)$$

由于固定型指标集和区间型指标集在本文中并没有应用, 故不予讨论。记无量纲化处理后的决策矩阵为 $Z = (z_{ij})_{n \times m}$ 。显然, z_{ij} 总是愈大愈好, 设规范化决策矩阵 $\tilde{Z} = (\tilde{z}_{ij})_{n \times s}$ 的正、负理想方案的属性值

作者简介: 龚艳冰(1979—), 男, 江苏靖江人, 讲师, 博士, 主要从事水资源评价。

分别为 $U = (u_1, u_2, \dots, u_m)$ 和 $V = (v_1, v_2, \dots, v_m)$, 其中

$$u_j = \max_{1 \leq i \leq n} z_{ij} \quad v_j = \min_{1 \leq i \leq n} z_{ij} \quad (j = 1, 2, \dots, m) \quad (3)$$

若设评价指标间的加权向量为 $W = (w_1, w_2, \dots, w_m)^T$ 。目前, 大多采用专家权重法确定评价指标的权重, 为了克服专家权重法因各个专家经验和审查问题角度不同而造成权重的不确定性, 采用熵权法来确定权重向量。由熵理论可知属性 R_j 输出的信息熵为

$$H_j = -\frac{1}{\ln n} \sum_{i=1}^n f_{ij} \ln f_{ij} \quad (j = 1, 2, \dots, m) \quad (4)$$

式中 $f_{ij} = \frac{z_{ij}}{\sum_{i=1}^n z_{ij}}$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$), 进而第 j 个评价指标的熵权 w_j 定义为

$$w_j = \frac{1 - H_j}{\sum_{k=1}^m (1 - H_k)} \quad (5)$$

如果将每个决策方案看成一个行向量, 则每个决策方案 A_i 与正、负理想方案的 U, V 之间的加权 Theil 不等系数可分别定义为^[10]

$$\theta_U(i) = \frac{\sqrt{\sum_{j=1}^m (z_{ij} - u_j)^2 w_j}}{\sqrt{\sum_{j=1}^m z_{ij}^2} + \sqrt{\sum_{j=1}^m u_j^2}} \quad (i = 1, 2, \dots, n) \quad (6)$$

$$\theta_V(i) = \frac{\sqrt{\sum_{j=1}^m (z_{ij} - v_j)^2 w_j}}{\sqrt{\sum_{j=1}^m z_{ij}^2} + \sqrt{\sum_{j=1}^m v_j^2}} \quad (i = 1, 2, \dots, n) \quad (7)$$

显然 $\theta_U(i) \geq 0, \theta_V(i) \geq 0$, 且加权 Theil 不等系数 $\theta_U(i)$ 愈小, 方案 A_i 越接近正理想方案的 U , 即方案 A_i 越好; 同理加权 Theil 不等系数 $\theta_V(i)$ 愈小, 方案 A_i 越接近负理想方案的 V , 也即方案 A_i 越差。

因此, 可以通过比较各个方案之间的加权 Theil 不等系数的大小对方案进行排序。为了综合考虑方案 A_i 与正、负理想方案的 U, V 加权 Theil 不等系数的大小, 笔者用每个方案与理想方案的相对贴近度的大小来对方案进行排序, 其计算公式为

$$D_i = \frac{\theta_U(i)}{\theta_U(i) + \theta_V(i)} \quad (i = 1, 2, \dots, n) \quad (8)$$

$D_j^* = \max_i \{D_i\}$ 所对应的方案即为最优方案。

2 应用实例

2.1 评价指标的确定

评价指标是综合评价的基础, 指标选取的合理与否, 直接关系评价结果的合理与否, 因此, 建立区域水资源可再生能力评价指标体系, 应该根据区域水资源的区域特点, 全面、合理地考虑各个因素。一般的原则如下^[11]: ①功效性原则; ②系统性原则; ③可度量性原则; ④代表性原则; ⑤层次性原则。文献[4]利用主成分赋权方法, 建立了黄河流域水资源可再生能力的评价指标, 主要包括: 单位面积的水资源量 U_1 ; 单位面积地表水资源量 U_2 ; 单位面积地下水资源量 U_3 ; 丰水年单位面积水资源量 U_4 ; 枯水年单位面积水资源量 U_5 ; 干旱指数 U_6 (倍比); 降水量 U_7 ; GDP 年增长率 U_8 ; 农业总产值增长率 U_9 ; 万元产值农业耗水率 U_{10} ; 牲畜用水定额 U_{11} 。具体数据见表 1。

2.2 评价计算结果与分析

以全国水资源可再生能力数据为评价标准, 可知 $U_1, U_2, U_3, U_4, U_5, U_7, U_8, U_9$ 可按效益型指标的原则无量纲化; U_6, U_{10}, U_{11} 3 个评定指标按成本型指标的原则无量纲化。由式(1)、式(2)和表 1 的数据, 计算可得规范化决策矩阵 Z 。

由式(3)可得正、负理想方案的属性值分别为

$$U = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$$

$$V = (0.0683, 0.0540, 0.2520, 0.0743, 0.0628, 0.0714, 0.4014, 0.7831, 0.3619, 0.0280, 0.5442)$$

表 1 黄河流域水资源可再生能力评价指标

行政区	$U_1/$ ($\text{m}^3 \cdot \text{m}^{-2} \cdot \text{a}^{-1}$)	$U_2/$ ($\text{m}^3 \cdot \text{m}^{-2} \cdot \text{a}^{-1}$)	$U_3/$ ($\text{m}^3 \cdot \text{m}^{-2} \cdot \text{a}^{-1}$)	$U_4/$ ($\text{m}^3 \cdot \text{m}^{-2} \cdot \text{a}^{-1}$)	$U_5/$ ($\text{m}^3 \cdot \text{m}^{-2} \cdot \text{a}^{-1}$)	U_6	$U_7/$ mm	$U_8/\%$	$U_9/\%$	$U_{10}/$ ($\text{m}^3 \cdot \text{万元}^{-1}$)	$U_{11}/$ ($\text{m}^3 \cdot \text{头}^{-1}$)
青海	0.137	0.137	0.061	0.158	0.117	2.3	443.3	8.3	3.98	1906	4.90
四川	0.278	0.278	0.127	0.323	0.239	0.5	712.6	8.0	6.75	154	3.94
甘肃	0.091	0.091	0.036	0.111	0.071	2.1	496.7	8.1	6.75	904	4.33
宁夏	0.019	0.017	0.032	0.024	0.015	5.0	313.2	7.9	6.75	5499	4.98
内蒙古	0.034	0.015	0.032	0.037	0.029	7.0	286.9	6.5	10.13	4341	4.90
山西	0.085	0.069	0.054	0.103	0.067	1.9	549.0	7.3	3.88	943	6.09
陕西	0.096	0.080	0.056	0.114	0.078	1.9	549.9	7.4	6.75	1007	7.22
河南	0.168	0.132	0.095	0.210	0.125	2.0	660.5	6.8	6.75	1341	7.24
山东	0.185	0.172	0.061	0.193	0.177	1.8	714.7	7.4	10.72	404	5.77

$$Z = \begin{pmatrix} 0.4928 & 0.4928 & 0.4803 & 0.4892 & 0.4895 & 0.2174 & 0.6203 & 1.0000 & 0.3713 & 0.0808 & 0.8041 \\ 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.9971 & 0.9639 & 0.6297 & 1.0000 & 1.0000 \\ 0.3273 & 0.3273 & 0.2835 & 0.3437 & 0.2971 & 0.2381 & 0.6950 & 0.9759 & 0.6297 & 0.1704 & 0.9099 \\ 0.0683 & 0.0612 & 0.2520 & 0.0743 & 0.0628 & 0.1000 & 0.4382 & 0.9518 & 0.6297 & 0.0280 & 0.7912 \\ 0.1223 & 0.0540 & 0.2520 & 0.1146 & 0.1213 & 0.0714 & 0.4014 & 0.7831 & 0.9450 & 0.0355 & 0.8041 \\ 0.3058 & 0.2482 & 0.4252 & 0.3189 & 0.2803 & 0.2632 & 0.7682 & 0.8795 & 0.3619 & 0.1633 & 0.6470 \\ 0.3453 & 0.2878 & 0.4409 & 0.3529 & 0.3264 & 0.2632 & 0.7694 & 0.8916 & 0.6297 & 0.1529 & 0.5457 \\ 0.6043 & 0.4748 & 0.7480 & 0.6502 & 0.5230 & 0.2500 & 0.9242 & 0.8193 & 0.6297 & 0.1148 & 0.5442 \\ 0.6655 & 0.6187 & 0.4803 & 0.5975 & 0.7406 & 0.2778 & 1.0000 & 0.8916 & 1.0000 & 0.3812 & 0.6828 \end{pmatrix}$$

由式(4)和规范化决策矩阵 Z , 可计算出信息熵为

$$H = (0.9179 \ 0.8971 \ 0.9573 \ 0.9198 \ 0.9110 \ 0.8876, \\ 0.9830 \ 0.9988 \ 0.9806 \ 0.7673 \ 0.9923)$$

再由 H_j 和式(5)可得指标属性的熵权为

$$W = (0.1043 \ 0.1307 \ 0.0542 \ 0.1019 \ 0.1130 \ 0.1428, \\ 0.0216 \ 0.0015 \ 0.0246 \ 0.2955 \ 0.0098)^T$$

最后由式(6)~(8)计算可得加权 Theil 不等系数值, 见表 2。

表 2 Theil 不等系数值和水资源可再生能力排序

行政区	θ_{ij}	θ_{ij}	D	水资源可再生能力排序		
				本文方法	物元模型	投影模型
青海	0.1224	0.0753	0.3809	4	4	7
四川	0.0088	0.1869	0.9548	1	1	1
甘肃	0.1296	0.0543	0.2953	6	7	5
宁夏	0.1758	0.0144	0.0759	9	9	9
内蒙古	0.1708	0.0279	0.1404	8	8	8
山西	0.1348	0.0510	0.2744	7	6	4
陕西	0.1296	0.0561	0.3021	5	5	6
河南	0.1102	0.0897	0.4489	3	2	3
山东	0.0843	0.1084	0.5624	2	3	2

由表 2 可见, 在黄河流域 9 个地区中, 四川省的水资源可再生能力排第 1 位, 因此, 该地区的水资源可再生能力是最强的。这是因为该地区除了农业总产值增长率这项指标外, 其他各项指标都高于其他地区, 山东省和河南省次之, 内蒙古和宁夏地区的水资源可再生能力较弱, 这与该地区的自然环境、社会环境和经济环境有关。本文结果和文献 [6-7] 的加权物元方法和投影寻踪方法得到的黄河流域地区的水资源可再生能力的实际情况基本一致。由于本文通过信息熵方法确定指标权重, 且计算各方案与理想方案的 Theil 不等系数, 因此, 本文方法与加权物元方法和投影寻踪方法相比, 具有计算简便、模型简单、实际操作性强的特点。

3 结 论

为了解决实际水资源可再生能力各项评价指标

的不相容问题, 笔者提出一种基于信息熵的 Theil 不等系数方法, 详细论述了该方法的计算步骤, 即采用信息熵来确定各项评价指标的权重, 并利用各个地区的指标值与正、负理想方案之间的 Theil 不等系数来判断各个地区的水资源可再生能力的强弱, 并应用该方法对黄河流域 9 个行政分区的水资源可再生能力进行了综合评价。评价结果表明, 四川省在黄河流域众多地区中的水资源可再生能力最强, 而内蒙古和宁夏地区的水资源可再生能力较弱。

参考文献:

- [1] 马文慧. 三江源区的生态保护与水资源可持续利用[J]. 水利经济, 2007, 25(3): 19-22.
- [2] 李朝霞. 区域水资源可持续利用指标体系研究[J]. 河海大学学报: 自然科学版, 2007, 35(1): 81-85.
- [3] 周长进, 董锁成, 张小军. 党河水环境特征与水资源可持续利用[J]. 水资源保护, 2008, 24(1): 42-44.
- [4] 沈珍瑶, 杨志峰. 黄河流域水资源可再生性评价指标体系与评价方法[J]. 自然资源学报, 2002, 17(2): 188-197.
- [5] 杨晓华, 杨志峰, 沈珍瑶. 水资源可再生能力综合评价的多目标决策理想区间法[J]. 中国科学, 2004, 34(S1): 34-41.
- [6] 杨晓华, 杨志峰, 沈珍瑶. 水资源可再生能力评价的遗传投影寻踪方法[J]. 水科学进展, 2004, 15(1): 73-76.
- [7] 杨晓华, 杨志峰, 沈珍瑶. 水资源可再生能力综合评价的遗传加权物元模型[J]. 数学的实践与认识, 2004, 34(11): 56-63.
- [8] 陈守煜, 李敏. 基于可变模糊集理论的水资源可再生能力评价模型[J]. 水利学报, 2006, 37(4): 431-435.
- [9] 彭少明, 黄强. 黄河流域水资源可持续利用多目标规划模型研究[J]. 河海大学学报: 自然科学版, 2007, 35(2): 153-158.
- [10] 陈华友. 基于 Theil 不等系数的组合预测模型的性质[J]. 电子科技大学学报, 2004, 33(1): 105-108.
- [11] 来海亮, 汪党献, 吴涤非. 水资源及其开发利用综合评价指标体系[J]. 水科学进展, 2006, 17(1): 95-101.

(收稿日期: 2008-08-16 编辑: 徐广生)