

区域水资源承载能力综合评价的 GPPIM

郦建强¹ 陆桂华¹ 杨晓华² 杨志峰²

(1. 河海大学水问题研究所, 江苏 南京 210098; 2. 北京师范大学环境科学研究所, 北京 100875)

摘要:为解决单项水资源承载能力评价指标评价结果的不相容问题, 提高水资源承载能力综合评价各层次的分辨力和评价模型的精度, 采用大样本数据, 利用投影寻踪、遗传算法、插值型曲线和水资源承载能力评价标准, 为水资源承载能力综合评价建立了一种新的数学模型——遗传投影寻踪插值模型(GPPIM), 给出了实施 GPPIM 的详细步骤, 作为示例, 对山西省水资源承载能力进行了综合评价, 结果表明, GPPIM 的评价结果可靠、直观、适用性强, 不仅可以对水资源承载能力进行综合评价, 还可以科学地确定评价指标的权重, GPPIM 可广泛应用于各种水资源问题的综合评价。

关键词: 水资源承载能力 综合评价 投影寻踪 遗传算法 模型精度

中图分类号: TV213 **文献标识码:** A **文章编号:** 1000-198X(2004)01-0001-04

水资源承载能力是指在某一历史发展阶段, 以可预见的技术、经济和社会发展水平为依据, 以可持续发展为原则, 以维护生态环境良性发展为条件, 在水资源得到合理开发利用的情况下某一研究区域人口增长与经济发展的最大容量。可见, 对水资源承载能力进行评价必然要涉及经济、环境、资源和人口等多种因素^[1, 2]。所谓水资源承载能力综合评价, 就是根据水资源承载能力评价指标, 使问题研究定量化、程序化, 通过所建立的数学模型, 对一个地区的水资源承载能力进行评价, 从而全面地分析出其水资源承载能力的状况, 为水资源可持续利用提供科学依据和决策支持。由于实际水资源承载能力各项指标的评价结果常常是不相容的, 所以直接利用水资源承载能力评价标准表进行水资源承载能力等级评判缺乏实用性。模糊综合评判方法^[3]、人工神经网络法^[4]等也难以精确地从各层次考察水资源承载能力的综合能力。20 世纪 70 年代, Friedman^[5]给出了多元数据分析的投影寻踪(Projection Pursuit, 简称 PP)方法, 该算法能在一定程度上解决多指标样本分类等非线性问题。最近, 投影寻踪方法已被用来对水资源问题进行综合评价^[6], 但由于采用的样本过少(每一等级取 5 个样点), 所建立的数学模型不够准确, 对于小样本问题容易产生模型误差, 而对多元数据, 这种复杂的拓扑结构一般也很难找到最优的投影方向。目前, 现代优化算法中的遗传算法由于提供了一种求解复杂系统优化问题的通用框架, 它不依赖于问题的具体领域, 具有较强的全局搜索能力、信息处理的隐并行性、鲁棒性和可规模化等优点^[7], 是处理这类优化问题的较好方法。基于上述原因, 本文利用投影寻踪方法把多指标问题转化为单一的投影指标问题, 采用大量的样本, 利用全局收敛的实编码混合加速遗传算法^[8]来优化投影方向, 并根据最佳投影值与其对应等级之间所呈现的关系, 建立新的数学模型, 由此提出了一种适合于对水资源承载能力进行综合评价的新模型——遗传投影寻踪插值模型(Genetic Projection Pursuit Interpolation Model, 简称 GPPIM), 并对山西省水资源承载能力进行综合评价。

1 GPPIM 的计算技术

投影寻踪是处理高维数据, 尤其是非正态数据的一类新兴统计方法, 它的基本方法是按照实际问题的需要, 把高维数据投影到低维子空间上, 使用者通过观察投影图像, 采用投影目标函数来衡量投影揭示数据某种结构的可能性大小。其中投影目标函数的构造、优化方法的选择、数学模型的建立是应用投影寻踪解决实际问题的关键。传统的投影寻踪计算技术较复杂, 在一定程度上限制了投影寻踪方法的深入研究和广泛应用。为此, 本文提出一套基于实编码混合加速遗传算法的插值型投影寻踪建模方法, 该方法可用于水资源承

载能力的综合评价. GPPIM 的计算包括以下 5 步:

步 1:建立投影数据. 根据水资源承载能力评价标准产生用于水资源承载能力评价的原始数据, 包括水资源承载能力指标 $x^*(i, j)$ 及对应的等级值 $y(i)$ ($i = 1 \dots, n; j = 1 \dots, n_p$). 其中 n, n_p 分别为样品的个数和评价水资源承载能力的指标数. 对 $x^*(i, j)$ 进行归一化处理为 $x(i, j)$, 即

$$x(i, j) = (x^*(i, j) - x_{\min}(j)) / (x_{\max}(j) - x_{\min}(j)) \quad i = 1 \dots, n \quad j = 1 \dots, n_p \tag{1}$$

其中 $x_{\min}(j), x_{\max}(j)$ 分别为第 j 个指标的最小值和最大值.

步 2: 计算投影值. 设 $a = (a(1), a(2), \dots, a(n_p))$ 为投影方向, 把 $x(i, j)$ 投影到 a 上, 得到一维投影值 $z(i)$,

$$z(i) = \sum_{j=1}^{n_p} a(j)x(i, j) \quad i = 1 \dots, n \tag{2}$$

步 3: 建立投影目标函数. 在综合投影值时, 要求投影值 $z(i)$ 应尽可能大地提取 $x(i, j)$ 中的变异信息, 即 $z(i)$ 的标准差 S_z 尽可能大, 同时要求 $z(i)$ 与 $y(i)$ 的相关系数 R_{ZY} 的绝对值 $|R_{ZY}|$ 尽可能大^[6]. 本文取投影目标函数 $f(a)$ 为

$$f(a) = S_z |R_{ZY}| \tag{3}$$

步 4: 用实编码混合加速遗传算法来优化投影方向. 投影目标函数 $f(a)$ 随着投影方向 a 的变化而变化, 可通过求解投影目标函数最大值来估计最佳投影方向 a , 即

$$\max f(a) = S_z |R_{ZY}| \tag{4}$$

$$\text{s.t.} \quad \sum_{j=1}^{n_p} |a(j)| = 1 \quad -1.0 \leq a(j) \leq 1.0$$

这是一个以 $a = (a(1), a(2), \dots, a(n_p))$ 为优化变量的非线性优化问题, 用常规的方法处理较困难, 为了增强 GPPIM 的实际应用能力和计算效率, 本文用实编码混合加速遗传算法给出最佳投影方向. 其基本思想是, 在单位超球面中随机抽取若干个初始投影方向作为初始群体, 用实编码混合加速遗传算法来优化投影目标函数, 并建立与投影目标函数的大小相应的适应度函数, 适应度函数大的个体被保留的机会大, 通过选择、杂交、变异等遗传操作, 最后最大适应度函数值对应的个体与最大投影目标函数对应的最佳投影方向 a^* 相对应.

步 5: 建立水资源承载能力综合评价数学模型. 由步 4 求得的最佳投影方向 a^* 代入式 (2) 后, 得到第 i 个样品的投影值 $z^*(i)$. 根据 $z^*(i)$ 与 $y(i)$ 的散点图可建立水资源承载能力综合评价的插值型投影寻踪数学模型. 详见第 2 部分.

2 水资源承载能力综合评价

鉴于影响水资源承载能力的因素很多, 根据区域水资源特征, 笔者选取的评价指标为^[1] (a) 人均水资源可利用量 U_1 ($\text{m}^3/\text{人}$) (b) 水资源利用率 U_2 (%) (c) 人均供水量 U_3 ($\text{m}^3/\text{人}$) (d) 供水量模数 U_4 ($\text{万 m}^3/\text{km}^2$) (e) 工业用水重复利用率 U_5 (%) (f) 耕地灌溉率 U_6 (%) (g) 生活用水定额 U_7 ($\text{L}/(\text{人} \cdot \text{d})$); (h) 需水量模数 U_8 ($\text{万 m}^3/\text{km}^2$) (i) 生态用水率 U_9 (%). 本文将水资源承载能力分为 3 个等级: 等级 1 表示水资源承载能力较强, 等级 2 表示水资源承载能力中等, 等级 3 表示水资源承载能力较弱.

表 1 评价标准^[1]
Table 1 Assessment standard^[1]

等级	U_1 ($\text{m}^3 \cdot \text{人}^{-1}$)	U_2 /%	U_3 ($\text{m}^3 \cdot \text{人}^{-1}$)	U_4 ($\text{万 m}^3 \cdot \text{km}^{-2}$)	U_5 /%	U_6 /%	U_7 ($\text{L} \cdot \text{人}^{-1} \cdot \text{d}^{-1}$)	U_8 ($\text{万 m}^3 \cdot \text{km}^{-2}$)	U_9 /%
1	> 550	< 30	> 330	< 1.5	< 25	< 20	< 70	< 1.5	> 5
2	400 ~ 550	30 ~ 60	240 ~ 330	1.5 ~ 3.9	25 ~ 75	20 ~ 30	70 ~ 130	1.5 ~ 4.7	2 ~ 5
3	< 400	> 60	< 240	> 3.9	> 75	> 30	> 130	> 4.7	< 2

在表 1 中各等级取值范围内均匀随机产生各 1000 个指标样本 $x^*(i, j)$, 与对应的 3 个等级

$$y(i) = \begin{cases} 1 & 1 \leq i \leq 1000 \\ 2 & 1001 \leq i \leq 2000 \\ 3 & 2001 \leq i \leq 3000 \end{cases}$$

一起组成样本系列,并对 $x^*(i,j)$ 进行归一化处理为 $x(i,j)$ ($i=1,\dots,3\,000$; $j=1,\dots,9$),用 GPPIM 中的式 (1)~(4) 得最佳投影方向:

$a^*=(0.082,-0.135,0.144,-0.089,-0.113,-0.010,-0.173,-0.091,0.163)$

其中 $0.082,-0.135,0.144,-0.089,-0.113,-0.010,-0.173,-0.091,0.163$ 可作为 1~9 个指标的权重.

最大投影目标函数值 $f(a^*)=0.75$.

最佳投影值与等级的散点关系见图 1. 其中 A,B,C,D,E,F 的坐标分别为 $(-1.511,3),(-0.824,3),(-0.502,2)(0.017,2)(0.309,1)(0.928,1)$.

从图 1 可以看出:

a. $z^*(i)$ 与 $y(i)$ 的图形为阶梯形、下降曲线. 用 A,B,C,D,E,F 6 个点进行分段线性插值,则该折线所对应的函数即式 (5) 建立水资源承载能力综合评价的数学模型. 用 GPPIM 计算的水资源承载能力等级评价模型为

$$y^*(i)=\begin{cases} 3 & z^*(i)\leq -0.824 \\ 3+\frac{2-3}{-0.502+0.824}(z^*(i)+0.824) & -0.824<z^*(i)<-0.502 \\ 2 & -0.502\leq z^*(i)\leq 0.017 \\ 2+\frac{1-2}{0.309+0.017}(z^*(i)+0.017) & 0.017<z^*(i)<0.309 \\ 1 & z^*(i)\geq 0.309 \end{cases} \quad (5)$$

b. 用 GPPIM 计算的各个样本的最佳投影值 $z^*(i)$ 所对应的等级值 $y(i)$ 误差为零,可见用 GPPIM 对水资源承载能力进行综合评价精度较高.

3 应 用

山西省水资源承载能力综合评价的各指标值详见文献 [1]. 用 GPPIM 和文献 [1] 的数据,对山西省水资源承载能力作出综合评价,评价结果详见表 2. GPPIM 计算结果表明:

a. 现状年 从等级值为 2.487~2.750 来看,山西省各流域水资源开发利用达到相当大的规模,已接近其开发容量,水资源供需矛盾突出,在现有的技术经济条件下水资源承载能力介于中等和较弱之间^[1].

b. 2010 年 (a) 方案 1(用水水平充分考虑未来节水措施,供水以本地水资源为主):从等级值为 3 来看,水资源承载能力较弱. 水资源供需矛盾更加突出,水资源承载能力比现状年进一步减弱. (b) 方案 2(用水水平充分考虑未来节水措施,供水除本地水资源外,还包括引、提黄及污水处理回用):从等级值为 3 来看,水资源承载能力较弱. 从投影值来看,水资源供需矛盾仍然突出. 水资源承载能力比现状年有所减弱,但较方案 1 有所提高. 纵观全局,在水供需战略上必须突出资源优化配置、需水管理和控制.

然而,用模糊综合评判方法计算的结果,只能模糊地断定山西省现状年、2010 年各流域水资源承载能力介于中等和较弱之间^[1].

可见, GPPIM 不仅可利用等级值来判定水资源承载能力所处的等级,而且可利用投影值的大小进一步判断处于同一等级的水资源承载能力的相对大小. 而模糊综合评判方法只能用等级值来评价水资源承载能力的大小,对处于同一等级的水资源承载能力的相对大小的判断较为模糊.

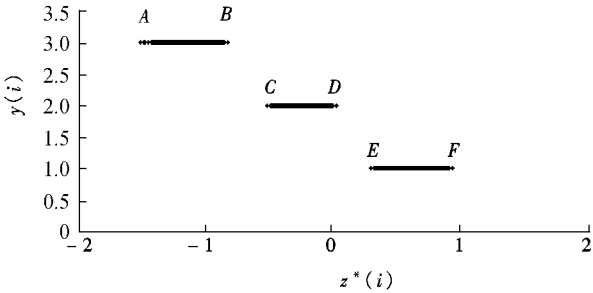


图 1 投影值 $z^*(i)$ 与经验等级 $y(i)$ 的散点关系

Fig.1 Scatter dots of projection values and water resources bearing capacity grades of samples

表 2 用 GPPIM 对山西省水资源承载能力评价的结果

Table 2 Assessment result of water resources bearing capacity of Shanxi Province by GPPIM

评价年	地区	投影值	等级值
现状年	黄河流域	-0.659	2.487
	海河流域	-0.744	2.750
	全 省	-0.715	2.661
2010 年 (方案 1)	黄河流域	-0.897	3.000
	海河流域	-0.905	3.000
	全 省	-0.901	3.000
2010 年 (方案 2)	黄河流域	-0.901	3.000
	海河流域	-0.916	3.000
	全 省	-0.906	3.000

4 结 论

a. GPPIM 解决了评价指标评价结果的不相容问题,评价结果直观、可靠。GPPIM 可把水资源承载能力多维评价指标综合成一维投影指标,根据等级值的大小就可以对水资源承载能力进行评价,解决了各单项水资源承载能力评价指标评价结果的不相容问题。由于采用了大样本数据,且建立的数学模型所计算的各样本的等级值误差为零,所以 GPPIM 对水资源承载能力进行综合评价具有较高的精度。

b. GPPIM 可以被用来科学地确定权重。GPPIM 不仅可以对水资源承载能力进行综合评价,还可利用最佳投影方向科学地确定水资源承载能力评价指标的权重,特别适合于被评价对象的评价指标难以确定的问题。

c. GPPIM 适用性强、分辨能力高。由于投影值与等级值的图形为阶梯形、下降曲线,可以得出投影值越小的区域其水资源承载能力越强的结论。对于多个区域水资源承载能力的评价问题, GPPIM 既可以利用等级值从整体上来判别水资源承载能力所属等级,又可以利用投影值对属于同一等级的各区域水资源承载能力进行细致评价。GPPIM 不仅具有较强的分类功能,而且具有较好的排序功能,提高了区域水资源承载能力综合评价问题各层次的分辨力。

参考文献:

- [1] 肖满意,董羽立. 山西省水资源承载能力评估[J]. 山西水利科技, 1998, 123(4): 5.
- [2] 秦莉云, 金忠青. 淮河流域水资源承载能力的评价分析[J]. 水文, 2001, 21(3): 14.
- [3] 杨纶标, 高英仪. 模糊数学原理及其应用[M]. 广州: 华南理工大学出版社, 2001. 129—146.
- [4] 薛建军, 姚桂基. 人工神经网络在水质评价中的应用[J]. 水文, 1997, 17(3): 37.
- [5] FRIEDMAN J H. A projection pursuit algorithm for exploratory data analysis[J]. IEEE Trans on Computer, 1974, 23(9): 881.
- [6] 金菊良, 魏一鸣, 丁晶. 水质综合评价的投影寻踪模型[J]. 环境科学学报, 2001, 21(4): 431.
- [7] 周明, 孙树栋. 遗传算法原理及应用[M]. 北京: 国防工业出版社, 2001. 15.
- [8] 杨晓华, 陆桂华, 酆建强. 混合加速遗传算法在流域模型参数优化中的应用[J]. 水科学进展, 2002, 13(3): 340—344.

DPPIM for comprehensive assessment of regional water resources bearing capacity

LI Jian-qiang¹, LU Gui-hua¹, YANG Xiao-hua², YANG Zhi-feng²

(1. College of Water Resources and Environment, Hohai Univ., Nanjing 210098, China;

2. Institute of Environmental Sciences, Beijing Normal Univ., Beijing 100875, China)

Abstract: To avoid the nonuniformity of different indexes-based assessment results of water resources bearing capacity and to raise the precision of assessment, a new mathematical model—Genetic Projection Pursuit Interpolation Model (GPPIM), is developed for comprehensive assessment of water resources bearing capacity based on a great deal of sampled data, projection pursuit, genetic algorithm, interpolation curve, and assessment standard for water resources bearing capacity, and the detailed steps for the GPPIM are also given. As an example, the model is applied to the assessment of water resources bearing capacity of Shanxi Province, and the result shows the reliability and high applicability of the GPPIM. With the model, not only can the bearing capacity of water resources be comprehensively assessed, but also the weight of assessment indexes can be scientifically determined, so the model can be widely used for comprehensive assessment of various water resources problems.

Key words: water resources bearing capacity; comprehensive assessment; projection pursuit; genetic algorithm; model precision