

基于多元图形特征与云模型理论的质量评价方法研究*

李 昕, 洪文学, 宋艳东, 曹 静, 马 利

(燕山大学 电气工程学院, 河北 秦皇岛 066004)

摘 要: 针对综合质量评价中评价结果不能充分反映主、客观因素共同作用的问题, 提出了一种基于多元图表示原理与云模型理论相结合的综合质量评价方法。该方法基于多元图表示原理, 应用雷达图表示方法的图形面积融合多指标参数, 结合云模型理论定性定量转换强大功能, 实现综合质量评价, 得到符合人类思维方式的定性定量综合评价结果。以成都市土壤为研究对象, 实验结果表明, 该方法得到的定性定量相结合的综合评价结果符合实际数据的分布和人的思维方式, 是一种充分反映主、客观因素共同作用的评价方法。该方法适用于多指标参数的质量评价问题。

关键词: 质量评价; 复用雷达图; 云模型

中图分类号: X825

文献标志码: A

文章编号: 1001-3695(2010)08-3024-03

doi:10.3969/j.issn.1001-3695.2010.08.056

Research of soil evaluation based on cloud model and multi-parameter chart presentation

LI Xin, HONG Wen-xue, SONG Yan-dong, CAO Jing, MA Li

(Institute of Electrical Engineering, Yanshan University, Qinhuangdao Hebei 066004, China)

Abstract: In order to obtain the result of quality assessment with both quantitative and qualitative presentation, this paper researched a new assessable method based on the chart expression theory and the cloud mode theory. Used the radar chart theory to fuse the multi-parameter, so the basic assessment result was got ton and it was visual. Researched the quantitative and qualitative information transform based on the trait of cloudy model. The experiment was finished with the soil data in Chengdu. The result is being reacted the sense of human being with quantitative and qualitative information. It is an effective method.

Key words: quality assessment; multi-used radar chart; cloud model

综合质量评价是指对以多属性体系结构描述的对象系统作出全局性、整体性的评价, 是由分指标上的评价结果得到的指标上的评价结果的过程^[1,2]。

目前常用的评价方法分为专家群体打分评价法、经济分析法、运筹学方法、控制论、数理统计方法、信息论、灰色系统理论、多维标度分析、星座图法和雷达图法等。

其中运筹学方法又可分为多目标决策法、数据包络分析法和层次分析法等。数理统计方法又可分为聚类分析法、判别分析法、主成分分析法、因子分析法等。这些评价方法中, 有定性评价、定量评价及定性定量结合评价之分, 如专家评价法属定性评价、主成分分析法、因子分析法等数理统计方法属于定量评价, 层次分析法属定性定量相结合的评价^[3-6]。总之, 评价方法可以归纳为定量评价法、定性评价法和定性定量相结合的评价方法三种。

随着研究问题的深入, 研究的对象越来越复杂。一个复杂系统只用定量信息描述是不够的, 有时甚至是不可能的。复杂系统一定涉及定性定量信息的综合表示的问题, 一定会涉及

多维定性与定量信息的融合问题。

本文将可视化的图表示原理与定性定量转换模型——云模型理论相结合, 充分应用图表示原理的可视化的特性, 结合云模型理论的定性定量信息转换的强大功能^[7], 希望得到一种符合人类思维方式的定性定量综合的质量评价方法。

1 质量评价方法基本理论

1.1 图表示原理

多元数据的图表示和图分析是多元统计分析中重要的方法, 该方法不仅可以利用图形, 而且可以利用构图的几何图形的数学表达和图形形状来表达有用信息^[8]。本文采用复用雷达图表示理论。典型的复用雷达图如图 1 所示。

基于图形特征的雷达图表示理论, 将多变量映射到复用雷达图, 用图形的面积作为整体质量的表征函数, 图像的形状表征信息的细节特征。每一个输入变量用一条单位圆上的射线(数轴)表示。每个输入变量对输出整体质量变量贡献的权重数由该变量在单位圆上逆时针方向对应的弧度来表示。

收稿日期: 2009-12-02; **修回日期:** 2010-03-09 **基金项目:** 国家自然科学基金资助项目(60304009); 河北省人事厅优秀专家出国培训资助项目

作者简介: 李昕(1970-), 女, 教授, 博士, 主要研究方向为信息融合与质量评价、智能传感器等(yddylixin@ysu.edu.cn); 洪文学(1953-), 男, 教授, 博导, 主要研究方向为模式识别、智能传感器。

1.2 基于图表示原理的多参数融合

1.2.1 基础指标的标准化

由于土壤环境质量综合评价指标体系中的基础指标有不同的单位,往往造成各指标之间数值上存在差别,需要对指标作标准化处理。

首先将各基础指标归一化,如式(1)所示,使各指标均变换为均值为0、方差为1的量^[9]。

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (1)$$

其中: x_{ij} 是某区域 j 指标 i 的测量值; \bar{x}_j 为 j 指标的均值; σ_j 为 j 指标的标准差; x'_{ij} 为标准化处理后的值。

对归一化后的基础指标 x'_{ij} 作如下非线性变换:

$$y_{ij} = \frac{2}{\pi} \tan^{-1}(x'_{ij}) + 1 \quad (2)$$

通过该变换,使得在均值附近变换具有较好的线性,而偏离均值越远变换的压缩性越强^[10],将无限区间变换为有限区间 $[0,2]$,使均值由0变为1。

1.2.2 基于三角形面积的变量融合方法

各输入信息根据权值系数 w_i 的不同映射到雷达图的射线上。输入信息变量的个数为 p ,雷达图的圆周据权值系数 w_i 分成 p 部分。

$$r_i = x_i / x_{i \max}; i = 1, 2, \dots, p \quad (3)$$

其中: x_i 是第 i 个信息变量, $x_{i \max}$ 是第 i 个信息变量的最大值。

则融合的结果为

$$S = (1/2) \times \sum_{i=1}^p r_i \times r_{i+1} \times \sin \phi_i \quad (4)$$

其中: ϕ_i 为半径 r_i 与半径 r_{i+1} 之间夹角的弧度值,且 $\phi_i = w_i \times 2\pi$ 。

1.3 云模型

云是用语言值描述的某个定性概念与其数值表示之间的不确定性转换模型。云由云滴组成,其整体形状反映了定性概念的重要特性。云滴则是对定性概念的定量描述,云滴的产生过程,表示定性概念与定量值之间的不确定性映射。云用期望值 Ex 、熵 En 、超熵 He 三个数字特征来整体表征一个概念,如图2所示。

云模型的数字特征把模糊性(定性概念的亦此亦彼性)和随机性(隶属度的随机性)完全集到一起,构成定性和定量相互间的映射,反映了定性知识的定量特性^[11,12]。

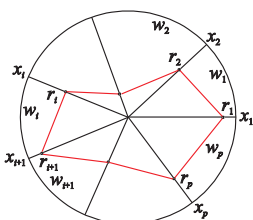


图1 复用雷达图

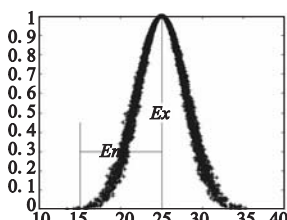


图2 隶属云及其数字特征

1.4 基于多元图形特征的云模型原理

基于多元图形特征的云模型原理如图3所示。

运用图表示原理,首先对图形特征进行提取,对多元数据融合以及降维。以提取的特征量作为基础变量,根据云模型理论的基本特征,构造描述系统评价所用的语言值。采用正向云进行系统的模糊评价,应用逆向云进行基础指标的分析,从而得到符合人类思维方式的集模糊性与随机性于一体的定性定

量集成综合质量评价结果。

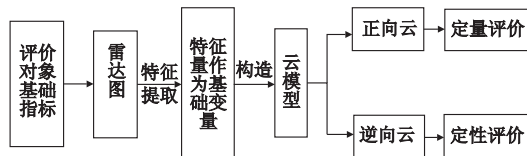


图3 基于多元图形特征的云模型原理框图

针对评价问题中的定性变量与定量变量具有的不同特征,采用不同方向、不同种类的云模型进行处理^[7]。

1) 变量云化

变量云化的实质是定性变量与定量变量的归一化过程,使其在统一标尺下形成可比性,是进行定性定量转换的前提。

(1) 对于定量变量,如其具有上下界,形如 $[B \min, B \max]$,云化为正向云模型。采用约束条件的中值作为期望值,用主要作用区域为双边约束区域的云来近似这个定量变量,如式(5)所示:

$$\begin{cases} Ex = (B \min + B \max) / 2 \\ En = (B \max - B \min) / 6 \\ He = k \end{cases} \quad (5)$$

其中: k 为常数,可根据变量本身的模糊程度来具体调整。

对于只有单边的定量变量,形如 $[B \min, \infty]$ 或 $[-\infty, B \max]$,可先根据测试数据的最大上限或下限确定其缺省边界参数或期望值,再参照式(5)计算云参数。

(2) 定性变量的云化,定性变量往往通过专家采用自然语言描述的评语来进行赋值。每个专家都可以根据自己的理解,为每个定性变量设定经验上的数字特征,形成评语云模型。 n 个专家提出的 n 个云模型表示的定性变量可以采用一个综合云来表征,数字特征为

$$\begin{cases} Ex = \frac{Ex_1 \times En_1 + Ex_2 \times En_2 + \dots + Ex_n \times En_n}{(En_1 + En_2 + \dots + En_n)} \\ En = En_1 + En_2 + \dots + En_n \\ He = \frac{He_1 \times En_1 + He_2 \times En_2 + \dots + He_n \times En_n}{En_1 + En_2 + \dots + En_n} \end{cases} \quad (6)$$

2) 构造云标尺

将定性变量和定量变量的云化结果置于统一坐标系上,形成云标尺。云标尺中的云簇形成定性或定量的激活区间。对任意时刻 T ,定性或定量的输入通过激活相应区间的云模型选转换云模型。

3) 构造规则发生器

根据不同的 X 条件组成,构造定性规则发生器,进行定性输入的处理。当输入平面的某一特定的输入值 $X = X_j$ 刺激各单规则生成器时,每个发生器随机产生一个 Y_j 值,这些值反映了各个单规则发生器中输入 X_j 对定性规则的激活强度(即隶属度),可以理解为对云标尺相应区间的激活强度。选择 Y_j 中最大值 Y_{\max} ,该最大值表明该条定性规则具有最高的支持率,可选中其为优先规则, Y_{\max} 控制输出平面对应的 CGY_j 产生云滴。

对于定量输入的处理是通过构造逆向云发生器生成云模型的过程,形成对大量数值总体特征的直观描述,反馈给专家。

2 实验结果与分析

以成都市土壤为评价对象。土壤中的 Hg、Cd、Pb、As、Cr、

Ni、Cu、Zn 重金属元素是评价土壤质量的基础指标。参照国家土壤环境质量标准,如表 1 所示,可将土壤污染程度分为四个等级:一级(清洁)、二级(警戒)、三级(轻度污染)、四级(中度以上污染)。

表 1 成都地区土壤重金属元素污染评价标准(元素含量:mg·kg⁻¹)

评价标准	Hg	Cd	Pb	As	Cr	Ni	Cu	Zn
一级(清洁/自然背景)	≤0.15	≤0.2	≤35	≤15	≤90	≤40	≤35	≤100
二级(警戒)	0.1~0.50.2	~0.3 35	~300	15~25 90	~300	40~50 35	~100	100~250
三级(轻度污染)	0.5~1.5 0.3	~1.0 300	~500	25~30 300	~400	50~200	100~400	250~500
四级(中度以上污染)	>1.5	>1.0	>500	>30	>400	>200	>400	>500

表 2 数据是不同土壤环境质量区重金属元素含量。根据每个地区的基础指标绘制雷达图,如图 4 所示。图中第一条实线的部分是土壤环境质量的一级标准,介于第一条实线与第二条实线之间属于二级,第二条与第三条实线之间属于三级,超出第三条实线外的属于中度以上污染;虚线表示成都市各个地区的重金属含量。以雷达图的面积作为特征量完成特征提取,运用云模型定性定量转换思想完成转换。评语云标尺如图 5 所示。图中从左到右依次为一级、二级、三级、四级的云模型。

表 2 不同土壤环境质量区重金属元素含量对比(元素含量:mg·kg⁻¹)

区域	Hg	Cd	Pb	As	Cr	Ni	Cu	Zn
成都市区浅层	0.50	0.23	52.72	16.81	81.63	38.42	67.64	148.91
成都市区深层	0.05	0.13	23.23	13.20	81.39	36.71	30.04	75.52
彭州—青白江	0.37	0.19	43.86	12.42	161.87	53.65	40.62	101.81
龙门山	0.11	0.35	48.31	13.94	141.64	51.20	53.58	147.63
峨山东坡	0.09	0.49	39.43	15.53	128.80	46.95	48.27	125.00
成都地区浅层	0.11	0.21	32.58	8.95	76.01	33.87	29.87	85.50
成都地区深层	0.05	0.15	25.51	9.15	77.25	34.12	26.22	77.85

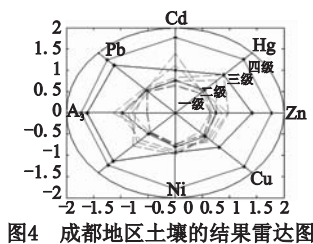


图 4 成都地区土壤的结果雷达图

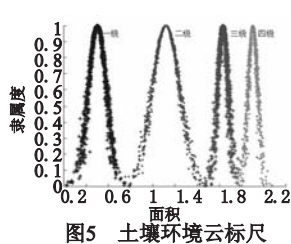


图 5 土壤环境云标尺

以成都市区土壤为例^[12],由图表示原理将成都市各个区域土壤数据进行融合,得到面积为 0.766 2。采用逆向云发生器实现定量→定性转换,即(Ex, En, He)分别为(0.76, 0.154 5, 0.01),再运用正向云发生器生成云滴,从而形成云模型,置于云标尺中,即可以确定整个成都地区的土壤等级介于一级与二级之间,如图 6 所示。

本文分别对重庆市、漳州市和昆山市三个不同城市的土壤环境进行了综合评价,数据来源于文献[13~15],如图 7 所示,三个城市的土壤都处在清洁与警戒状态之间。

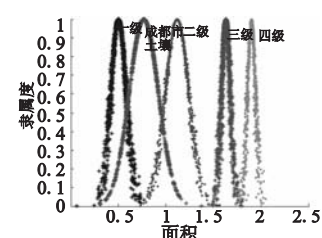


图 6 成都地区土壤的结果云模型

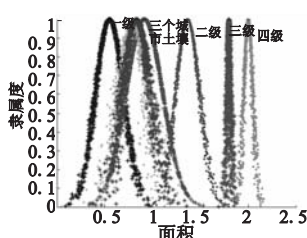


图 7 三个城市土壤的结果云模型

仿真结果表明,用于评价的云模型在概念所描述的期望值点上具有凝聚性,符合用模糊语言直接进行评价时的基本准则,能够明确地反映各语言值的覆盖程度,对于相同的面积值,

其隶属度具有随机性,能够反映模糊评价所体现的差异性。

3 结束语

基于多元图形特征云模型原理的质量评价方法,将评价对象的各项指标映射到二维复用雷达图中,首先可视化地得到被评价对象的基本状况,并对被评价对象哪些指标存在问题作出初步判断。进一步应用复用雷达图面积实现多参数的融合,得出关于质量的总体评价。应用云模型理论,用语言值定性描述定性概念与其数值表示之间的不确定转换关系,使评估结果除了保留传统定性划分所具有的优点以外,也更加符合实际的数据分布和人的思维方式。将这两种方法有效地结合并运用在土壤环境质量评价方法中,雷达图可以表示各单项指标结果:成都市区主要是 Hg 的污染;彭州—青白江主要是 Hg 和 Ni 的污染;峨山东坡和龙门山主要是 Cd、Cr、Ni、Cu 和 Zn 的污染。基于图形特征云模型结果给出关于系统的总体评价:整个成都地区的土壤环境质量介于清洁与警戒之间。

参考文献:

- [1] GOVAERTS B, SAYRE K D, DECKERS J. A minimum data set for soil quality assessment of wheat and maize cropping in the highlands of Mexico[J]. *Soil & Tillage Research*, 2006, 87(2): 163-174.
- [2] VELASQUEZ E, LAVALLE P, ANDRADE M. GISQ, a multifunctional indicator of soil quality[J]. *Soil Biology & Biochemistry*, 2007, 39(12): 3066-3080.
- [3] BOHANEK M, CORTET J, GRIFFITHS B. A qualitative multi-attribute model for assessing the impact of cropping systems on soil quality[J]. *Pedobiologia*, 2007, 51(3): 239-250.
- [4] ZORNOZA R, MATAIX-SOLERA J, GUERRERO C, et al. Evaluation of soil quality using multiple linear regressions based on physical, chemical and biochemical properties[J]. *Science of the Total Environment*, 2007, 378(1-2): 233-237.
- [5] HOFMAN J, BEZCHLEBOVO J, DUSEK L, et al. Novel approach to monitoring of the soil biological quality[J]. *Environment International*, 2003, 28(8): 771-778.
- [6] MUNOZ A, LOPEZ-PINEIRO A, RAMIREZ M. Soil quality attributes of conservation management regimes in a semi-arid region of south western Spain[J]. *Soil & Tillage Research*, 2007, 95(1-2): 255-265.
- [7] 李德毅, 孟海军, 史雪梅. 隶属云和隶属云发生器[J]. *计算机研究与发展*, 1995, 32(6): 16-21.
- [8] 李伟明. 多元描述统计方法[M]. 上海: 华东师范大学出版社, 2001: 16-58.
- [9] 郑惠莉, 刘陈, 翟丹妮. 基于雷达图的综合评价方法[J]. *南京邮电学院学报*, 2001, 21(2): 75-78.
- [10] 刘红樱, 谢志仁, 陈德友, 等. 成都地区土壤环境质量初步评价[J]. *环境科学学报*, 2004, 24(2): 1-7.
- [11] 宋远俊, 李德毅, 杨孝宗, 等. 电子产品可靠性的云模型评价方法[J]. *电子学报*, 2002, 12(12): 74-76.
- [12] 李德毅, 杜鹤. 不确定性人工智能[M]. 北京: 国防工业出版社, 2005: 137-182.
- [13] 李章平, 陈玉成, 杨学春, 等. 重庆市主城区土壤重金属的潜在生态危害评价[J]. *西南农业大学学报*, 2006, 28(2): 227-230.
- [14] 乔胜英, 李望成, 何方, 等. 漳州市城市土壤重金属含量特征及控制因素[J]. *地球化学*, 2005, 34(6): 635-642.
- [15] 陈凤, 濮励杰, 赵翠薇. 昆山市农用地土壤重金属污染现状及评价[J]. *淮阴工学院学报*, 2006, 15(1): 58-62.