

广东典型坡地土壤养分的聚类分析*

蔡昆争 段舜山 陈荣均

(华南农业大学热带亚热带生态研究所, 广州, 510642)

摘要 采用组平均聚类法、模糊聚类法、主成分分析法对广东不同坡地利用方式对土壤养分含量的影响进行聚类分析。综合3种分类方法, 可分为5大类: 第1类包括鹤山茶园和鹤山坡地花生; 第2类包括鹤山茶园空地、林地有覆盖和无覆盖、芒果地、草地、荒坡地和丰顺坡地青榄; 第3类包括鹤山坡地龙眼(新植)、番薯, 丰顺坡地龙眼和和高明坡地芒果; 第4类包括鹤山坡地龙眼, 丰顺坡地茶园, 曲江坡地花生和沙田柚; 第5类包括高州坡地甘蔗, 南雄坡地黄烟、银杏、三华李。结果表明3种分类方法均可用来进行分类, 但主成分分析法损失了较多的信息, 分类效果受到一定的影响。

关键词 坡地; 土壤养分; 聚类分析

中图分类号 S155.55

广东经济的高速发展, 城市化进程和工业用地的不断增加, 使得现有耕地面积在不断减少。据统计, 从1988~1996年, 广东平均每年耕地减少2.7%。在广东省的土地面积中, 山地丘陵占77%。开发山坡地, 发展有特色的山地生态农业模式, 已成为农业经济发展的新一轮增长点。聚类分析作为一种分析具有多个数量属性实体间的相似性的方法, 在农业生产的各个领域广为应用(吴乐民等, 1991; 王学荫等, 1994; 乔鑫山, 1995; 陶方玲等, 1995)。本研究通过在不同种植方式的坡地上取样, 采用不同的方法进行聚类分析, 分析和评价其对土壤养分的影响, 从而为坡地的进一步开发利用提供参考依据。

1 研究地点与调查测定项目

1.1 研究地点

选择广东省的典型坡地开展研究, 包括鹤山、高明、高州、南雄、丰顺、化州、曲江等, 这些县市坡地土壤以赤红壤为主, 有较大面积的坡地急待开发利用。在1996~1997年, 确定农户或农场作为定点分析研究, 采集土壤样品进行养分化验分析。

1.2 测定项目及聚类分析方法

在不同地区、不同坡地种植类型的样点(共21个)取0~40 cm土层的土样进行化验分析。土壤测试项目包括pH值、有机质、全氮、全磷、全钾、速效氮、速效磷、速效钾, 按常规方法进行(李西开, 1978)。

选择有机质、全氮、全磷、全钾、速效氮、速效磷、速效钾7个指标作为分类依据, 采用组平均聚类法、模糊聚类法和主元分析法(骆世明等, 1996)对样点进行聚类分析。

1998-10-06 收稿 蔡昆争, 男, 28岁, 助理研究员, 现为在职博士研究生

* 国家教委留学人员基金(1254)和广东省自然科学基金(950389)资助项目

2 聚类结果

2.1 组平均聚类法

(1) 数据标准化:由于各比较项目的数据属不同的量纲,因此采用标准差标准化:

$$SD_j = \sqrt{\frac{\sum_{i=1}^m a_{ij}^2 - (\sum_{i=1}^m a_{ij})^2 / M}{M-1}}$$

i 代表各样点 ($i=1, 2, \dots, n$); j 代表分类指标; SD_j 代表同一比较项目 (j) 在各样点 ($i=1, 2, \dots, m$) 数据的标准差. 原始数据 a_{ij} 经标准化后的数据为 a'_{ij} :

$$a'_{ij} = [(a_{ij} - \sum_{i=1}^m a_{ij}) / M] / SD_j$$

a_{ij} 代表第 i 个样点的第 j 项指标, a'_{ij} 代表标准化后的数据.

(2) 选择相异性指标, 并计算样点间相异性, 本文用欧氏距离作为相异性指标:

$$d_{ij} = \sqrt{\sum_{k=1}^m (a'_{ik} - a'_{jk})^2}$$

(3) 聚类组合: 根据矩阵中各样点所取的距离值, 以取值最小的样点首先聚合, 聚合而成的类群再与其它样点(类群)间的按最小距离聚合. 经过多次聚合与新距离矩阵计算, 逐步把不同样点聚合成新的类群. 当取聚合距离为 3.2 左右时, 21 个样点可聚合为 5 大类, 即第 I 类: [2, 11], 第 II 类: [1, 3, 4, 5, 7, 8, 13, 16, 20], 第 III 类: [9, 10, 14, 15], 第 IV 类: [6, 12, 17, 18], 第 V 类: [19, 21]. 见图 1.

2.2 模糊聚类分析

(1) 数据标准化: 同 2.1(1), 用标准差标准化.

(2) 计算相似性指数: 计算任 2 个样点之间的相似指数:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

x_i, y_i 分别为任意两个样点对应的数据. 由于 r 出现负值 ($-1 < r < 1$), 因此采用 $r' = (r+1)/2$ 式将 r 值变为 r' ($0 < r' < 1$) 以方便模糊合成.

(3) 利用模糊关系合成, 求出模糊隶属矩阵对应的模糊关系矩阵 R . 当取截距 $\lambda = 0.77$ 时, 21 个样点可聚合成 5 个类群, 见图 2, 即第 I 类: [2, 11], 第 II 类: [1, 3, 4, 5, 6, 7, 8, 12, 13, 17, 18] 第 III 类: [9, 10, 14, 15, 16], 第 IV 类: [20], 第 V 类: [19, 21].

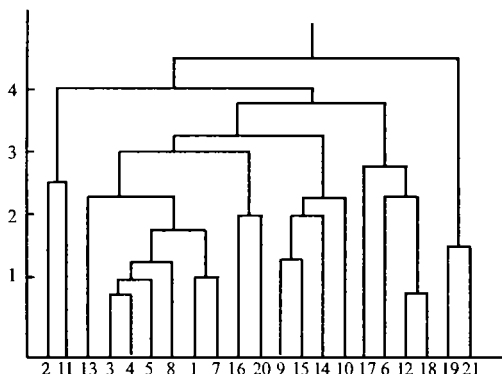


图1 组平均聚类分析结果

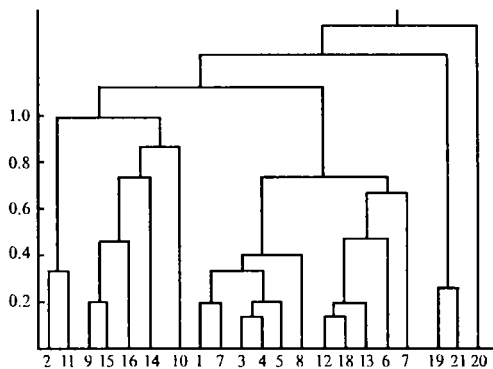


图2 模糊聚类分析结果

2.3 主成分分析

- (1) 原始数据标准化: 得到 X 矩阵
- (2) 求内积矩阵 S : $S = X \times X^T$, X^T 为 X 矩阵的转置矩阵.
- (3) 求内积矩阵的特征值和特征向量. 对于矩阵 S , 如果存在一组非零向量 U , 同时存在一组数 A , 使式 $US = AU$ 成立. 要使该式有非零解, 必须使其系数行列式为零, 即必须满足下式: $|S - \lambda I| = 0$. 解行列式, 则得到相应的特征根矩阵 $A(\lambda_1, \lambda_2, \cdots \lambda_n)$ 和特征向量 (U). 利用 $Y = UX$ 可得到坐标轴刚性转动为新坐标系后, 各样点在新轴上的值(即排序坐标), Y 为新坐标矩阵.
- (4) 求新坐标中各主分量轴集中的信息量比例:

$$P_i = \lambda_i / \sum_{i=1}^n \lambda_j, \text{ 前 } k \text{ 个分量轴反映的信息量合计 } B_k = \sum_{i=1}^k \lambda_i / \sum_{i=1}^n \lambda_j.$$

经计算得到相应的特征值与所含的信息量. 取信息量最大的两条新轴为排序轴, 代表了总信息量的 60.67%. 排序坐标见图 3, 从图中可看出, 各土壤样点可划分为五类, 即第Ⅰ类: [1, 3, 4, 7], 第Ⅱ类: [2, 5, 6, 8, 11], 第Ⅲ类: [16, 19, 20, 21], 第Ⅳ类: [9, 10, 13, 14, 15], 第Ⅴ类: [12, 17, 18].

3 种聚类分析方法的结果比较见表 1.

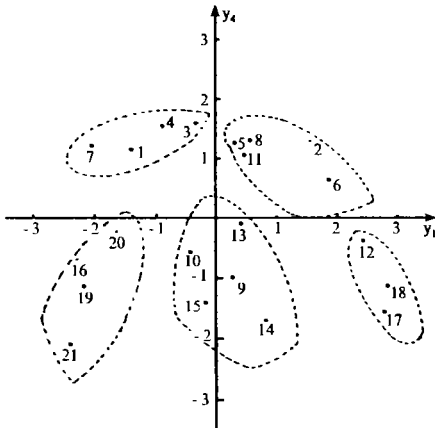


图3 主分量分析聚类结果

表 1 不同聚类分析法的分析结果比较

组平均聚类法	模糊聚类分析法	主成分分析法
2, 11(Ⅰ类)	2, 11(Ⅰ类)	2, 5, 6, 8, 11(Ⅰ类)
1, 3, 4, 5, 7, 8, 13, 16, 20(Ⅱ类)	1, 3, 4, 5, 6, 7, 8, 12, 13, 17, 18(Ⅱ类)	1, 3, 4, 7(Ⅱ类)
9, 10, 14, 15(Ⅲ类)	9, 10, 14, 15, 16(Ⅲ类)	9, 10, 13, 14, 15(Ⅲ类)
6, 12, 17, 18(Ⅳ类)	20(Ⅳ类)	12, 17, 18(Ⅳ类)
19, 21(Ⅴ类)	19, 21(Ⅴ类)	16, 19, 20, 21(Ⅴ类)

表 2 综合分类各组别的土壤养分平均值

类群	$w/(g \cdot kg^{-1})$				$w/(mg \cdot kg^{-1})$		
	有机质	全氮	全磷	全钾	碱解氮	速效磷	速效钾
2, 11(Ⅰ类)	23.90	1.091	0.547	52.49	77.28	45.25	35.0
1, 3, 4, 5, 7, 8, 13(Ⅱ类)	16.82	0.765	0.235	50.72	73.19	5.23	24.9
9, 10, 14, 15(Ⅲ类)	14.76	0.748	0.291	17.38	80.37	16.52	102.4
6, 12, 17, 18(Ⅳ类)	25.83	1.245	0.433	18.21	118.76	4.80	76.3
16, 19, 20, 21(Ⅴ类)	11.71	0.345	0.368	10.23	43.68	8.91	40.3

3 讨论

3.1 综合 3 种分类结果, 可把不同坡地作物利用对土壤养分含量影响的聚类分析分为 5 大类(见表 1 和表 2): 第Ⅰ类包括样点 2 和 11, 即鹤山茶园和鹤山坡地花生; 第Ⅱ类包括 1, 3,

4,5,7,8,13, 即鹤山茶园空地、林地有覆盖和无覆盖、芒果地、草地、荒坡地和丰顺坡地青榄;第Ⅲ类包括9,10,14,15, 即鹤山坡地龙眼(新植)、番薯, 丰顺坡地龙眼和和高明坡地芒果;第Ⅳ类包括6,12,17,18, 即鹤山坡地龙眼, 丰顺坡地茶园, 曲江坡地花生和沙田柚;第Ⅴ类16,19,20,21, 即高州坡地甘蔗, 南雄坡地黄烟、银杏、三华李。其中第Ⅳ类由于经营时间较长,且作物均采用复合种植模式,植被覆盖率较高,因此土壤养分含量在5类中最高。第Ⅴ类大都属于新开垦的坡地,作物种植单一,土壤的养分含量很低。第Ⅰ类鹤山坡地茶园经营时间也较长,同时坡顶有成片树林生长,对于茶园保持水土有积极的作用,而坡地花生样点由于花生为肥地作物,因此这两个样点养分含量也较高。其它样点类群的养分含量则介于上述举例的样点类群之间。

3.2 不同坡地作物种植方式对土壤养分的影响均可用3种聚类方法进行分类,分类结果大体相似(见表2),尤以组平均聚类分析和模糊聚类分析较为接近,而主成分分析两条最大信息量的轴加起来才反映信息总量的60%左右,可见损失了一部分信息,分类效果受到影响。因此,选择适当的分类方法有助于对客观实体的准确分类。

3.3 通过这种分类,可为坡地的开发利用提供一定依据。对于新开垦的坡地,应注意采取复合种植方式,同时在幼龄果园利用空隙地间套种如大豆、花生、牧草、甘薯、旱生蔬菜等作物,以增加植被覆盖率,减少水土流失,有利于土壤养分的积累和肥力的保持。

参 考 文 献

- 王学荫,聂宏声,郭常莲.1994.山西省生态农业区域划分的研究.生态学报,14(1):16~23
李西开主编.1978.土壤农业化学常规分析方法.北京:农业出版社,67~286
乔鑫山.1995.黄淮海地区生态农业模式分类.生态学杂志,14(4):41~47
吴乐民,温演望.1991.模糊聚类分析在旱作农业分区上的应用.华南农业大学学报,12(4):37~44
骆世明,彭少麟.1996.农业生态系统分析.广州:广东科技出版社,529~570
陶方玲,梁广文,庞雄飞.1995.模糊聚类分析法及其在群落聚类分析中的应用.生态科学,(2):60~65

Cluster Analysis of Soil Nutrient on Typical Hillsides in Guangdong

Cai Kunzheng Duan Shunshan Chen Rongjun

(Tropical and Subtropical Ecology Research Institute, South China Agric. Univ., Guangzhou, 510642)

Abstract Group mean cluster analysis(GCA), fuzzy cluster analysis(FCA) and principal component analysis(PCA) were used to classify soil nutrient on hillsides in Guangdong Province. The soil nutrient on hillsides could be classified into five types by integration of the above three methods: Type I. soil with tea and groundnut in Heshan; Type II. bare soil of the tea plantation, forest land with or without mulch, mango plantation, grass land, bare soil in Heshan and olive plantation in Fengshun; Type III. soil planted to longan or sweet potato in Heshan, longan plantation in Fengshun and mango in Gaoming; Type IV. soil planted to longan in Heshan, tea in Fengshun, groundnut and pummelo in Qujiang; Type V. soil planted to sugarcane in Gaozhou, tobacco, ginkgo and plum in Nanxiong. The results showed that the three methods could be used, but the PCA method lost some information and the classification effect had limitations.

Key words hillside; soil nutrient; cluster analysis

【责任编辑 张 砾】