

基于主成分和聚类分析的永泰县土壤养分地球化学综合评价

江晓龙 (福建省煤田地质勘查院, 福建福州 350000)

摘要 [目的]运用统计学方法对永泰县土壤养分地球化学综合等级进行评价,最大限度地减少人为主观因素的影响。[方法]对工作区内土壤进行系统采样,送检测试土壤养分大量元素(有机质、碱解氮、有效磷、速效钾、全氮、全磷、全钾)含量及pH,通过采样数据的一般性描述,初步判断工作区内土壤养分元素富集等情况,随后利用SPSS软件对数据进行标准化处理,建立相关系数矩阵,判断其进行统计学分析的可行性。通过主成分分析计算筛选出评价因子主成分,计算综合评价得分,并利用K-均值聚类分析法将综合得分划分为缺乏、较缺乏、中等、较富集、富集5类。[结果]工作区内土壤养分元素分布不均匀,变异性高,土壤总体呈酸性,富氮缺磷钾。土壤地球化学养分综合状况较差,综合得分属较缺乏、缺乏等级占比达54.68%。[结论]该研究为土壤养分评价提供科学依据。

关键词 主成分分析;聚类分析;土壤养分;地球化学;综合评价;永泰县

中图分类号 S151.9 文献标识码 A

文章编号 0517-6611(2023)01-0068-04

doi:10.3969/j.issn.0517-6611.2023.01.015

开放科学(资源服务)标识码(OSID):



Comprehensive Evaluation of Soil Nutrients Geochemistry in Yongtai County Based on Principal Component Analysis and Cluster Analysis

JIANG Xiao-long (Fujian Coalfield Geological Prospecting Institute, Fuzhou, Fujian 350000)

Abstract [Objective] To evaluate the comprehensive geochemical grade of soil nutrients in Yongtai County by using statistical methods, so as to minimize the impact of human subjective factors. [Method] Systematically sample the soil in the working area, the content of soil nutrient macroelements (organic matter, alkali hydrolyzable nitrogen, available phosphorus, available potassium, total nitrogen, total phosphorus, total potassium) and pH were submitted for testing, through the general description of the sampling data, the enrichment and grading of soil nutrient elements in the work area were preliminarily judged. Then, SPSS software was used to standardize the data and establish a correlation coefficient matrix to judge the feasibility of statistical analysis. The principal components of the evaluation factors were selected through principal component analysis, and the comprehensive evaluation scores were calculated. The comprehensive scores were divided into five categories: lack, relatively lack, medium, relatively enriched and enriched by using K-means cluster analysis. [Result] The soil nutrients elements in the working area was uneven distribution, high variability, and the soil was generally acidic, nitrogen enrichment, lack of potassium and phosphorus. The comprehensive nutrient status was poor, the lack and relatively lack of comprehensive level accounted for 54.68%. [Conclusion] This study provides scientific basis for soil nutrient evaluation.

Key words Principal component analysis; Cluster analysis; Soil nutrients; Geochemistry; Comprehensive evaluation; Yongtai County

土壤是物质生产最基本的载体,是作物吸收营养的主要渠道。农业生产中土壤的贡献率在50%~60%,土壤养分的丰缺是影响作物产量、品质的基本因素^[1-2],是地力评价的主要指标,因此,查明土壤养分状况对农业生产的合理化调配显得尤其重要。通过对工作区内的土壤养分地球化学情况进行综合评价,可为土壤资源开发管理提供一定的数据参考。土壤养分综合评价的方法主要有聚类分析法^[3]、模糊综合评判法^[4-5]、BP神经网络法^[6]、主成分分析法^[7-8]、层次分析法^[9]等。近年来,主成分分析法因其可通过降维方式找出具有代表性、综合性的影响因子而避免指标重叠的特性,得到学者的广泛应用,成为土壤定量研究中应用最为广泛的数理统计方法^[10-11]。该研究以永泰县土壤为研究目标,共采集土壤样品1955件,测试氮、磷、钾、有机质等大量养分指标,通过统计学方法,运用主成分分析、聚类分析筛选出影响土壤肥力的主成分及其影响权重大小,对土壤养分地球化学等级进行较为客观的综合评价,尽可能地减少人为主观因素干扰。

1 材料与与方法

1.1 研究区概况 永泰县隶属福州市管辖,地理位置为

118°07'08"~119°12'30"E、25°39'00"~26°04'21"N,土地面积共2229.86 km²。全县常年平均气温14.6~26.0℃,年无霜期近300 d,日照时数1445~2193 h,年降水量1400~2000 mm。凭借其得天独厚的自然环境,孕育出芙蓉李、绿茶、白云槟榔芋等一系列名特优农产品,获得国家地理标志证明商标、农产品地理标志产品等。

成土母质以沉积岩、中酸性侵入岩、火山岩为主,在多种因素的综合作用下,形成了诸多土壤类型。根据第二次土壤普查工作,查明县域地表土壤有8个土类,17个亚类,包括红壤、黄壤、黄棕壤、草甸土、紫色土、石灰(岩)土、潮土、水稻土。

1.2 评价方法

1.2.1 土壤样品采集。以《福建省农业地质调查评价工作手册》^[12]为依据,对永泰县域内耕地现状图斑进行评价单元划分,在评价单元中面积较大的地块内通过GPS定位采集土壤样品,根据评价单元面积大小采集3~5个子样品,再以等量混合的形式组合成1件样品,共采集耕层土壤1955件。经室内风干、过筛后按四分法缩分留样备用。

1.2.2 样品测试与方法。样品的检测根据《土地质量地球化学评价规范》(DZ/T 0295—2016)、《福建省农业地质调查评价工作手册》等相关规范规程由福建省121地质大队化验测试中心完成。土壤氮测定采用凯氏定氮法;土壤磷、钾测定采用X-荧光光谱法(XRF);有机质测定采用化学容量法

基金项目 福建省财政项目(2018001)。

作者简介 江晓龙(1992—),男,福建龙岩人,工程师,硕士,从事农业地质调查工作。

收稿日期 2022-02-27

(VOL), 碱解氮测定采用碱解-扩散法; 速效磷测定采用氟化铵-盐酸提取(碳酸氢钠溶液浸提)-钼锑抗比色法; 速效钾测定采用乙酸铵提取-等离子体发射光谱法。

1.3 数据处理 利用 SPSS 统计软件对样品测试结果进行统计学分析, 在标准化处理后构建相关系数矩阵并进行 KMO、Bartlett 球形度检验, 以判断进行因子分析的可行性。累计贡献率大于 80% 作为选择主成分的前提条件, 随后建立主成分的因子荷载矩阵, 通过换算得出主成分得分系数, 以计算主成分分值, 进一步通过综合得分公式求取各评价单元土壤养分综合得分值 (F), 随即根据综合得分值对土壤样品进行 K -均值聚类统计分析, 评价工作区内土壤养分地球化学综合等级。

主成分得分计算公式^[7]:

$$F_{ni} = \sum_{j=1}^z (\alpha_{ij} \times Z_{nj}) \quad (1)$$

式中, F_{ni} 代表第 n 个采样点第 i 项主成分的得分; α_{ij} 代表第 i 项主成分第 j 项原始指标的载荷; Z_{nj} 代表第 n 个采样点第 j

项原始指标的标准化数据。

综合得分计算公式^[13]:

$$F = \sum_{i=1}^3 (\gamma_i \times F_{ni}) \quad (2)$$

式中, F 表示综合得分, γ_i 表示第 i 项主成分贡献率。

2 结果与分析

2.1 土壤养分指标一般性描述 根据参数统计结果(表 1) 可见, 工作区内土壤养分含量分布不均匀, 从偏度、峰度可看出, 养分含量的分布形态多呈高尖峰右偏态分布, 即养分含量总体处于相对低值区, 但存在少量的异常值, 其中有效磷变化幅度最大, 极差达 1 018.47 mg/kg。从样品变异系数可看出, 有机质、碱解氮、全氮、全钾变异系数为 34%~39%, 属于中等变异; 有效磷、速效钾、全磷变异系数不小于 69%, 呈高变异性, 异常高值的存在一方面可能是受人为活动影响, 另一方面可能与成土母质不同有关。pH 中位数为 5.02, 众数为 4.79, 均值 5.04, 变异系数仅为 8%, 属于弱变异, 表明工作区内土壤酸碱度差异性较小, 区内土壤总体呈酸性。

表 1 土壤养分指标描述统计

Table 1 The descriptive statistics of nutrients content index

项目 Item	有机质 OM g/kg	碱解氮 Alkalin-N mg/kg	有效磷 Olsen-P mg/kg	速效钾 AK//mg/kg	全氮 TN//g/kg	全磷 TP//g/kg	全钾 TK//g/kg	pH
极小值 Minimum	2.46	4.65	0.05	7.37	0.09	0.06	1.10	3.45
极大值 Maximum	99.31	331.98	1 018.52	511.60	3.30	4.13	26.09	7.54
均值 Mean	24.46	111.61	53.77	79.32	1.21	0.50	12.11	5.04
众数 Mode	16.62	102.91	0.05	27.57	1.08	0.06	1.10	4.79
变异系数 CV//%	39	36	166	69	37	75	34	8
偏度 Skewness	1.41	0.88	4.04	2.42	0.94	2.14	0.22	0.45
峰度 Kurtosis	4.57	1.74	24.82	9.43	1.22	8.35	-0.34	1.66

2.2 土壤养分指标等级 根据全国第二次土壤普查养分分级标准, 从表 2 可以看出, 有机质含量 3 等(中等)以上评价单元比例达 64.50%; 约 68.29% 的评价单元内碱解氮含量水平在 3 等(中等)以上; 全氮含量 3 等(中等)以上比例为

63.68%; 有效磷含量在 33.50% 的评价单元中超过了 40 mg/kg, 表现为 1 等(丰富); 全磷含量水平较低, 72.08% 的评价单元全磷等级在 4 等(较缺乏)以下; 速效钾、全钾等级较低, 以 4 等(较缺乏)为主。

表 2 土壤养分含量等级分布比例

Table 2 Distribution proportion of soil nutrient content grades

等级 Grade	有机质 OM		碱解氮 Alkalin-N		有效磷 Olsen-P		速效钾 AK	
	含量 Content g/kg	比例 Ratio//%	含量 Content mg/kg	比例 Ratio//%	含量 Content mg/kg	比例 Ratio//%	含量 Content mg/kg	比例 Ratio//%
1 等(丰富) Level 1 (rich)	>40	6.39	>150	15.40	>40	33.50	>200	3.73
2 等(较丰富) Level 2 (relatively enriched)	>30~40	16.42	>120~150	21.38	>20~40	15.55	>150~200	5.42
3 等(中等) Level 3 (medium)	>20~30	41.69	>90~120	31.51	>10~20	16.93	>100~150	14.32
4 等(较缺乏) Level 4 (relatively lack)	>10~20	34.22	>60~90	24.91	>5~10	17.60	>50~100	43.17
5 等(缺乏) Level 5 (lack)	≤10	1.28	≤60	6.80	≤5	16.42	≤50	33.36
等级 Grade	全氮 TN		全磷 TP		全钾 TK		pH	
	含量 Content g/kg	比例 Ratio//%	含量 Content g/kg	比例 Ratio//%	含量 Content g/kg	比例 Ratio//%	数值 Value	比例 Ratio//%
1 等(丰富) Level 1 (rich)	>2.00	5.78	>1.0	9.46	>25	0.10	≤5.0	46.96
2 等(较丰富) Level 2 (relatively enriched)	>1.50~2.00	16.93	>0.8~1.0	7.21	>20~25	3.58	>5.0~6.5	52.69
3 等(中等) Level 3 (medium)	>1.00~1.50	40.97	>0.6~0.8	11.25	>15~20	21.84	>6.5~7.5	0.31
4 等(较缺乏) Level 4 (relatively lack)	>0.75~1.00	24.14	>0.4~0.6	20.26	>10~15	41.13	>7.5~8.5	0.04
5 等(缺乏) Level 5 (lack)	≤0.75	12.18	≤0.4	51.82	≤10	33.35	>8.5	0

工作区内土壤具有较高的氮肥力,速效钾、全钾、全磷含量较低,说明工作区内土壤钾、磷的缺乏,但较高的有效磷含量表明土壤中水溶性磷、部分吸附态磷及有机态磷等可被植物吸收的磷组分含量较多。土壤 pH 总体偏低,呈酸性,过酸的土壤会影响土壤和肥料中养分的溶解、沉淀和微生物的活动,进而影响养分的有效率,从而影响作物正常生长,并极易导致病虫害发生,应适时进行酸土改良。

2.3 土壤养分综合评价

2.3.1 相关系数矩阵及统计学检验。各评价指标间存在相关性是进行因子分析的前提,利用 SPSS 软件对各指标进行

标准化处理并构建相关系数矩阵(表 3)。由表 3 可见,各因子间均存在不同程度的相关性,表明各因子对土壤养分的作用效果存在着重叠关系且可相互影响,其中有机质与全氮的相关系数为 0.918,相关性最高,其次是碱解氮与全氮以及有机质与碱解氮间的相关性。

对各因子相关性进行 KMO 和 Bartlett 检验,结果显示,KMO 值为 0.74,表明各因子有较强的相关性,相关矩阵可行,适合进行因子分析^[14];Bartlett 值(P)为 0.00,进一步说明各因子间存在显著的相关性,即各因子间存在着可信赖较高的线性相关,可进行因子分析^[14]。

表 3 相关系数矩阵
Table 3 Correlation matrix

因子 Factor	有机质 OM	碱解氮 Alkalin-N	有效磷 Olsen-P	速效钾 AK	全氮 TN	全磷 TP	全钾 TK
有机质 OM	1						
碱解氮 Alkalin-N	0.861	1					
有效磷 Olsen-P	0.089	0.145	1				
速效钾 AK	-0.006	0.062	0.228	1			
全氮 TN	0.918	0.885	0.123	-0.003	1		
全磷 TP	0.377	0.438	0.642	0.183	0.459	1	
全钾 TK	-0.308	-0.234	0.028	0.152	-0.262	-0.060	1

2.3.2 主成分分析。以累计贡献率大于 80%为提取条件,可提取 3 个主成分,贡献率分别为 45.966%、22.540%、13.204%,累计贡献率达 81.710%(表 4),说明 3 个主成分可体现土壤养分 7 个因子中 81.710%的信息。

表 4 主成分贡献率

Table 4 The contribution rate of principal component

主成分 Principal component	特征值 Eigenvalue	贡献率 Contribution rate//%	累计贡献率 Cumulative contribution rate//%
1	3.218	45.966	45.966
2	1.578	22.540	68.506
3	0.924	13.204	81.710
4	0.773	11.037	92.747
5	0.295	4.212	96.959
6	0.140	1.995	98.954
7	0.073	1.046	100

进一步通过因子荷载系数判断各因子在不同主成分上的权重比例^[15],由表 5 可见,主成分 1 可有效反映有机质、碱解氮、全氮和全磷的权重影响,主成分 2 可反映有效磷的权

重影响,主成分 3 反映速效钾、全钾的权重影响。

根据公式(1)与表 5 得出各主成分得分方程分别为:

$$F_{n1} = 0.51Z_{n1} + 0.51Z_{n2} + 0.19Z_{n3} + 0.05Z_{n4} + 0.52Z_{n5} + 0.36Z_{n6} - 0.19Z_{n7} \quad (3)$$

$$F_{n2} = -0.20Z_{n1} - 0.12Z_{n2} + 0.62Z_{n3} + 0.47Z_{n4} - 0.15Z_{n5} + 0.46Z_{n6} + 0.33Z_{n7} \quad (4)$$

$$F_{n3} = 0.16Z_{n1} + 0.20Z_{n2} - 0.41Z_{n3} + 0.56Z_{n4} + 0.15Z_{n5} - 0.27Z_{n6} + 0.59Z_{n7} \quad (5)$$

结合公式(2)及表 5 得出主成分综合得分方程为:

$$F = 0.4597F_{n1} + 0.2254F_{n2} + 0.1320F_{n3} \quad (6)$$

综合得分越高则表明土壤养分的综合程度越好,土壤越肥沃,反之则表示土壤相对贫瘠^[7]。对各采样点综合得分进行描述性统计,结果发现,工作区内土壤养分综合得分在 -2.08~4.10,均值为 0,标准差为 0.88,不同评价单元土壤综合养分等级差异较大;偏度为 0.891,峰度为 1.065,呈低尖峰右偏态分布,表明工作区内土壤肥力离群值较少,总体肥力较低,需完善配方肥料供应体系,根据实际需求测土配方施肥,提高施肥效率,以改善土壤肥力。

表 5 因子荷载与得分系数

Table 5 Factor load and score coefficient

项目 Items	因子荷载 Factor load			得分系数 Score coefficient		
	主成分 1	主成分 2	主成分 3	主成分 1	主成分 2	主成分 3
Z_{n1} (有机质 OM)	0.915	-0.254	0.149	0.51	-0.20	0.16
Z_{n2} (碱解氮 Alkalin-N)	0.918	-0.146	0.195	0.51	-0.12	0.20
Z_{n3} (有效磷 Olsen-P)	0.340	0.778	-0.395	0.19	0.62	-0.41
Z_{n4} (速效钾 AK)	0.087	0.587	0.540	0.05	0.47	0.56
Z_{n5} (全氮 TN)	0.938	-0.193	0.146	0.52	-0.15	0.15
Z_{n6} (全磷 TP)	0.646	0.578	-0.263	0.36	0.46	-0.27
Z_{n7} (全钾 TK)	-0.342	0.414	0.571	-0.19	0.33	0.59

2.3.3 聚类分析。采用 K -均值聚类将土壤养分综合得分分为 5 类,最终聚类中心见表 6。第 1 聚类小组占比最高(30.69%),土壤养分综合得分均值为 -0.35 ,属于土壤肥力较缺乏区;第 2 小组(肥力中等区)占比 25.47%;肥力较富集以上区域累计占比仅为 19.85%;工作区内土壤综合养分状况较差,较缺乏、缺乏占比达 54.68%,进一步表明工作区内土壤肥力的不足。

表 6 综合得分分类

Table 6 Classification of comprehensive scores

聚类分组 Cluster group	聚类中心(F) Clustering center	占比 Proportion %
1(较缺乏 Relatively lack)	-0.35	30.69
2(中等 Medium)	0.29	25.47
3(缺乏 Lack)	-0.99	23.99
4(较富集 Relatively enriched)	1.09	15.45
5(富集 Enriched)	2.33	4.40

3 讨论

该研究以统计学为基础,通过主成分分析、聚类分析对土壤地球化学综合养分进行评价,结果显示,永泰县有机质含量水平主要为 3 等,土壤普遍富氮缺磷、钾。通常土壤中可被植物吸收的氮、磷、钾多以碱解氮、有效磷、速效钾的形式出现,但县域内土壤普遍呈现出速效钾缺乏、有效磷较富集的现象,碱解氮水平则以中等为主。 pH 平均值为 5.04,变异系数小,表明区内土壤总体呈酸性,且不同土壤类型、成土母质下的差异性较小,推测区内土壤酸化应与农业生产活动有关,如过磷酸钙等酸性肥料的使用。主成分分析显示,提取 3 个主成分可呈现出土壤综合养分 81.710% 的信息,能取得较好的分析效果。从土壤养分地球化学综合等级来看,永泰县内土壤综合得分普遍不高,应加强土壤及耕作方式的科学化管理,加强人为的正向干预,改良土壤肥力及质量。

运用统计学方法虽可最大程度地减少人为主观因素对评价结果的影响,但土壤是一个开放的载体,综合养分不仅由地球化学元素决定,土壤的类型、耕作方式、生态环境变化等因素时时刻刻都在影响着土壤的状态。此次研究仅从地球化学角度出发,选取土壤养分大量元素作为评价指标较为有限,如何对土壤养分进行更加全面的、多角度、多维度的科学评价仍需进一步探讨。

4 结论

(1)工作区土壤养分元素含量分布不均匀,变异性高,总体具有较高的氮肥力,但普遍缺钾缺磷。 pH 总体较低,酸性土占比 99.65%,土壤总体酸性较强,变异性低。

(2)土壤养分地球化学综合得分在 $-2.08\sim 4.10$,均值为 0,标准差 0.88,呈低尖峰右偏态分布,表明工作区内土壤综合养分普遍偏低,仅有少部分区域具有较好的土壤养分。

(3)通过 K -均值聚类分析,工作区内有 30.69% 的评价单元土壤综合养分较缺乏,23.99% 的评价单元综合养分呈缺乏状态,养分不足区域占总评价单元的 54.68%,进一步说明工作区内土壤肥力的不足。

(4)建议根据土壤养分分布情况,针对性地测土施肥,在生产过程中应注重培肥地力及土壤的酸性改良,建立科学的施肥、耕作方案。

参考文献

- [1] 耿维,袁漫漫,邬刚,等.安徽省作物养分供需分析及化肥减施潜力研究[J].中国生态农业学报(中英文),2020,28(2):221-235.
- [2] 汪学美.谈土壤肥料对农业生产的影响[J].农业开发与装备,2017(4):107.
- [3] 范向龙,吕新,张泽,等.基于距离聚类与 K -means 动态聚类的棉田土壤养分评价研究[J].干旱区研究,2021,38(4):980-989.
- [4] 王杰,张春燕,卢加文,等.广安区柑橘土壤养分状况及综合肥力评价[J].土壤通报,2021,52(6):1360-1367.
- [5] 武伟,唐明华,刘洪斌.土壤养分的模糊综合评价[J].西南农业大学学报,2000,22(3):270-272.
- [6] 周婧,白云龙,张小宝,等.基于 MEA-BP 神经网络的土壤养分评价模型[J].中国农机化学报,2020,41(2):231-236.
- [7] 李金涛,李守岭,王晓媛,等.校园间作土壤养分及土壤肥力的综合评价[J].江西农业学报,2020,32(9):73-79.
- [8] 王天娇,朱卫红,吴婷婷.延吉市农田土壤肥力综合评价[J].安徽农业科学,2017,45(26):116-118,125.
- [9] 董起广,韩雾昌,张卫华,等.陕西汉中土壤养分评价[J].西部林业科学,2015,44(6):125-128.
- [10] 马光跃,卢桂宾,陈红玉,等.临漪县冬枣园土壤肥力综合评价[J].北方园艺,2021(20):90-96.
- [11] 刘文民,田鹏,刘利昆,等.青藏高原典型沙化区草地土壤质量评价[J].水土保持研究,2022,29(2):118-124.
- [12] 福建省自然资源厅.福建省农业地质调查评价工作手册[M].福州:福建省自然资源厅,2016.
- [13] 张瑞莲,袁海波,尹军峰,等.主成分分析与聚类分析在茶饮料汤色稳定性评价中的应用[J].茶叶科学,2010,30(4):288-294.
- [14] 旷爱萍,胡超.因子分析与熵值法下广西低碳农业发展质量综合评价[J].云南农业大学学报(社会科学),2021,15(3):61-69.
- [15] 李小胜,陈珍珠.如何正确应用 SPSS 软件做主成分分析[J].统计研究,2010,27(8):105-108.