

成县土壤数值分类的初步研究

李文学 刘东顺 冯学智
(甘肃省土地勘测设计院) (中国科学院兰州冰川研究所)

摘 要

本文应用标准化数据—斜交距离—误差平方和聚类分析法对甘肃成县 34 个土壤剖面进行了数值分析研究。研究表明,供试 34 个剖面可分为五个土壤类型。对五类土壤的 pH 值、有机质、代换量、 SiO_2 、 Al_2O_3 、 Fe_2O_3 、 CaCO_3 及粘粒含量八项主要指标进行方差分析的结果表明,除 pH、代换量和 SiO_2 含量未达类型间 F 测验显著水平外,其他各项指标均达显著或极显著水平。各类土壤在地理分布、植被类型、田间石灰反应等特性上也有明显的差异,并且每一类土壤都具有一定的发生学意义。并分别对应于棕壤、淋溶褐土、褐土、碳酸盐褐土和潮土五种发生学类型。

土壤数值分类是近十多年来随着电子计算机技术日益广泛应用而发展起来的一种土壤分类方法。本文采用标准化数据—斜交距离—误差平方和聚类分析方法^[4],对甘肃成县 34 个剖面的 pH 值、有机质、全氮、全磷、全钾、速效磷、速效钾、阳离子代换量、 SiO_2 、 Al_2O_3 、 Fe_2O_3 、 CaCO_3 及物理性粘粒含量共 13 项指标的分析数据,经电子计算机处理进行了土壤数值分类的初步探讨,以完善成县的土壤分类。

一、数据处理与计算方法

(一) 数据整理

供试土壤剖面共 34 个。以剖面各发生层土层厚度为权,求得各指标的剖面加权平均值列于表 1。

将 N 个指标值作为变量 x_1, x_2, \dots, x_N 。令 x_{ij} 记第 i 个剖面第 j 个变量的计量, $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{iN}]$, ($i = 1, 2, \dots, M$), 记第 i 个剖面的计量向量,并看作是 N 维空间的一个点。然后根据 \mathbf{x}_i 与 \mathbf{x}_k ($i = 1, 2, \dots, M-1; k = i+1, \dots, M$) 之间的相似性对它们进行逐步归并,最后形成一个可反映土壤剖面间亲疏关系的谱系图,可以较客观地描述它们之间的差异性。

(二) 计算方法

1. 数据的标准化: 为了消除各指标的量纲差异,并平衡数据量级,本文对数据首先进行了标准化处理。设原始数据为 x_{ij} , 则标准化后的数据为 z_{ij} 。即

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (i = 1, 2, \dots, M; j = 1, 2, \dots, N)$$

其中

$$\bar{x}_i = \frac{1}{M} \sum_{j=1}^M x_{ij}$$

$$\sigma_i = \sqrt{\frac{1}{M} \sum_{j=1}^M (x_{ij} - \bar{x}_i)^2}$$

上式中 \bar{x}_i 为第 i 个指标的平均值, σ_i 为第 i 个指标的标准差。经过标准化后的变量均为 0, 标准差为 1。

表 1 供试 34 个剖面理化性状的剖面加权平均值

Table 1 Weighted averages of the physical and chemical properties of 34 profiles used in experiment

剖面编号 Profile No.	pH	有机质 (%) O.M.	全氮 (%) Total N	全磷 (%) Total P ₂ O ₅	全钾 (%) Total K ₂ O	速效磷 (ppm) Avai- lable P ₂ O ₅	速效钾 (ppm) Avai- lable K ₂ O	代换量 (meq /100g) CEC	SiO ₂ (%)	Al ₂ O ₃ (%)	Fe ₂ O ₃ (%)	CaCO ₃ (%)	物理性粘粒 (%) Clay (<0.01 mm)
1	8.30	0.28	0.042	0.054	1.62	5.47	86.12	9.63	47.04	10.96	4.02	25.45	51.25
2	8.12	0.41	0.040	0.055	2.12	9.90	133.01	15.12	61.61	13.80	5.24	5.73	48.99
3	8.19	0.76	0.071	0.084	2.44	3.99	98.42	11.64	61.22	13.74	5.12	5.72	42.75
4	8.12	0.91	0.085	0.057	2.12	2.74	112.07	18.08	60.17	13.96	5.53	5.84	50.87
5	8.15	0.71	0.13	0.057	2.43	6.14	150.03	16.84	61.50	14.79	6.06	4.05	46.54
6	7.69	0.89	0.091	0.056	2.25	3.10	120.00	20.94	66.04	15.16	5.94	0	54.20
7	8.11	0.62	0.061	0.053	2.16	6.20	104.30	17.39	65.73	14.98	5.54	0.49	43.63
8	7.98	1.06	0.083	0.081	2.31	59.64	201.55	10.31	65.15	13.37	5.00	4.84	25.92
9	8.10	1.13	0.10	0.089	2.32	7.96	271.00	13.96	62.61	14.04	5.03	4.53	32.46
10	7.92	1.70	0.13	0.065	2.24	32.65	107.00	15.50	62.23	14.87	5.50	4.43	42.17
11	7.55	0.36	0.057	0.059	2.37	8.08	117.97	17.47	64.45	15.68	6.11	0	54.02
12	7.99	2.97	0.18	0.037	2.12	5.22	121.95	26.35	61.61	16.23	5.60	0.42	46.26
13	7.90	1.92	0.12	0.046	1.97	3.87	82.13	17.59	66.66	14.34	5.54	0	37.69
14	8.02	0.67	0.069	0.059	2.31	8.48	116.94	18.17	65.40	14.94	5.64	0.095	50.44
15	7.51	0.44	0.056	0.048	2.24	5.11	102.07	20.42	64.43	15.60	6.01	0	56.68
16	8.26	0.37	0.044	0.049	2.11	7.84	114.93	15.65	61.32	14.13	5.38	6.88	50.22
17	8.13	0.63	0.064	0.076	2.20	10.78	127.89	15.78	68.94	13.99	5.16	0.12	39.91
18	7.78	1.93	0.12	0.049	2.04	7.57	97.18	22.55	65.27	14.43	5.57	0.069	49.42
19	8.03	0.61	0.071	0.044	2.29	5.04	123.49	21.10	65.00	14.92	5.85	0.41	52.09
20	7.51	0.56	0.058	0.058	2.33	7.37	107.67	17.26	64.65	15.36	6.04	0	54.30
21	7.43	1.41	0.10	0.061	2.23	11.97	135.41	18.72	63.29	14.92	5.91	0	48.28
22	8.09	0.53	0.051	0.067	1.83	13.63	94.67	11.67	55.99	12.25	4.70	17.03	50.97
23	7.68	0.34	0.038	0.051	2.04	2.74	94.97	28.56	56.33	12.83	4.39	13.61	37.89
24	7.95	0.58	0.047	0.064	2.01	26.25	76.67	9.88	66.69	11.83	4.49	7.02	24.79
25	7.99	0.47	0.054	0.072	2.23	5.20	126.47	16.07	61.59	13.82	5.30	5.63	45.10
26	8.07	0.49	0.051	0.057	2.23	7.02	119.91	16.14	64.78	13.98	5.29	1.82	45.48
27	8.18	0.15	0.035	0.067	2.54	5.38	135.06	21.92	64.62	14.96	5.53	0.27	48.99
28	8.01	0.60	0.066	0.058	2.29	3.63	125.70	20.24	60.13	13.62	5.07	7.06	47.94
29	7.93	1.30	0.11	0.061	2.21	2.34	94.07	17.68	60.61	15.25	6.16	2.77	49.34
30	7.12	0.40	0.054	0.032	2.59	4.45	88.66	15.95	65.11	14.40	5.49	0	58.87
31	7.76	0.68	0.061	0.066	2.32	2.54	131.08	17.05	62.98	15.00	5.84	1.85	47.35
32	6.66	1.66	0.12	0.044	1.93	3.06	69.36	17.86	65.48	14.38	5.28	0	50.70
33	7.73	2.82	0.18	0.058	2.55	4.80	111.34	22.57	62.42	15.62	5.82	0	51.29
34	5.91	3.18	0.18	0.026	1.71	2.97	121.56	18.09	67.38	13.56	4.70	0	49.53

2. 计算相关系数矩阵 $R(\gamma_{ij})$ (表 2):

$$\gamma_{ij} = \frac{1}{M} \sum_{l=1}^M z_{li} \cdot z_{lj} \quad (i = 1, 2, \dots, N-1; j = i+1, \dots, N; l = 1, 2, \dots, M)$$

3. R-式聚类分析采用相关系数^[4]: 首先以 $R(\gamma_{ij})$ 为初始相关矩阵 $R^{(1)}$, 在 $R^{(1)}$ 中找出相关系数最大的一对指标合并。由公式

$$w_i = \frac{m_p w_p + m_q w_q}{m_p + m_q}$$

计算已合并指标的重心 w_i 。式中 w_p 和 w_q 分别为被合并指标的重心, m_p 和 m_q 分别为两个被合并指标内所包括的个体数。将合并以后的指标标准化并计算与其他指标间的相关系数得 $R^{(2)}$ 。重复以上步骤直到所有指标归并完毕, 然后根据各指标的连接顺序作出 R-式分析的二维谱系图(图 1)。

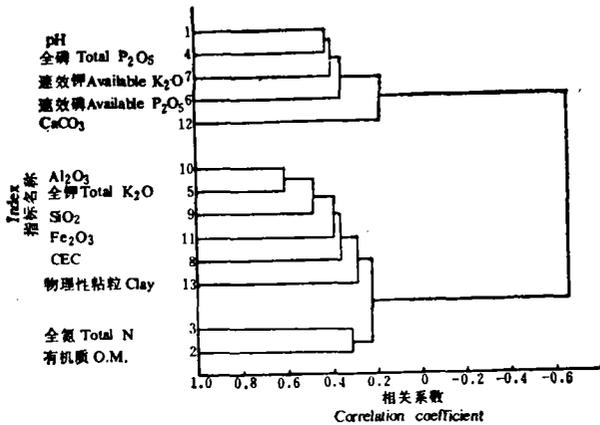


图 1 成县土壤 13 项指标的 R 聚类谱系图

Fig. 1 Dendrogram of R-cluster of 13 indexes of the soil in Cheng County

4. Q-式聚类分析采用标准化数据—斜交距离—误差平方和法: (1) 计算初始误差平方和增量矩阵 $W^{(1)}$ 。矩阵元素为

$$w_{ij} = \frac{1}{2} \sum_{k=1}^N \sum_{l=1}^N (x_{ki} - x_{kj})(x_{li} - x_{lj}) \gamma_{kl}$$

$$(i = 1, 2, \dots, M-1; j = i+1, \dots, M; k, l = 1, 2, \dots, N)$$

式中 γ_{kl} 为变量 k 与 l 间的相关系数。此时, 每群中只包含一个剖面, 群内误差平方和为 0。(2) 在 $w^{(1)}$ 中挑选误差平方和增量最小的一对点群归并, 并根据误差平方和增量的刷新公式

$$w_{ir} = \frac{m_p}{m_p + m_q} w_{pr} + \frac{m_q}{m_p + m_q} w_{qr} - \frac{m_p m_q}{(m_p + m_q)} w_{pq}$$

式中下标 p, q 为被归并的一对点群, m_p, m_q 分别为 p, q 点群所包含的个体数目, r 为归并后形成的新点群, τ 为其他点群。

刷新 $W^{(1)}$ 得 $W^{(2)}$ 。(3) 重复(1)和(2)步骤, 直到所有点群归并完毕, 形成 Q 式聚类的

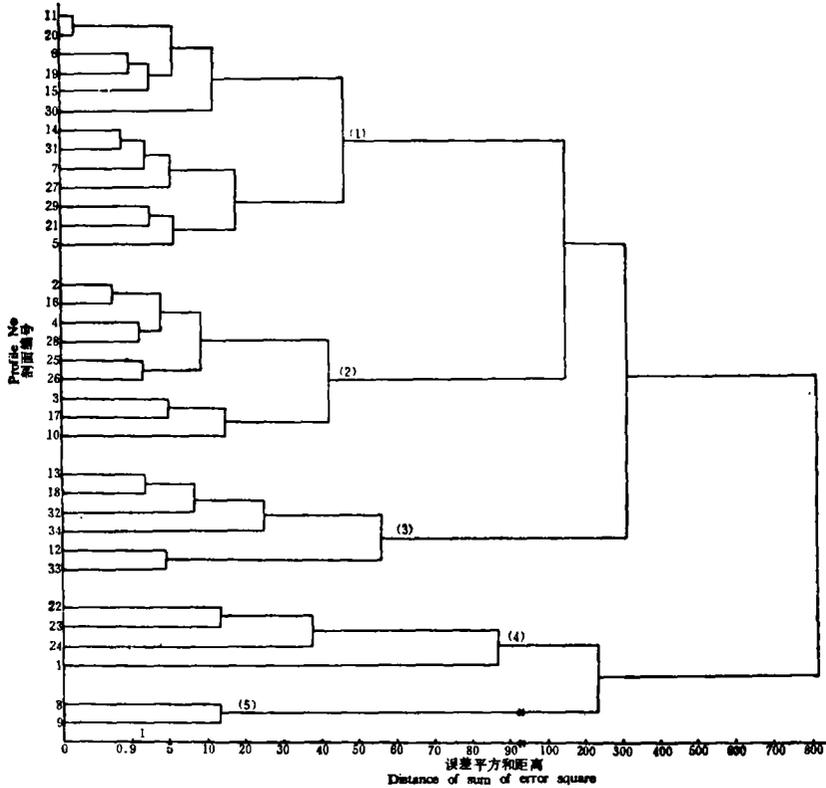


图2 成县土壤数值分类谱系图

Fig. 2 Dendrogram of numerical classification of 34 soil profiles in Cheng County

二维谱系图(图2)。

二、结果与分析

图1是13项指标间R聚类谱系图。由图可见,所选13项指标虽然根据-0.65的相关距离可将13项指标分为两大类,一类与pH呈正相关,另一类与pH呈负相关,但总的看来,各指标基本上是离散的,相互间并无较显著关系。通过对表2相关系数的显著性分析也表明,这13项指标间相关系数均未达到 $r_{0.05} = 0.553$ 的显著性水平。这证明了本文选择这13项指标进行Q式聚类分析是正确的。

根据方开泰所介绍^[1]由Demirmen(1976)提出的聚类图分类的四条准则,我们将图2从类内误差平方和86.70处划一条垂直于横轴的直线作为土壤类型划分界限,则34个剖面大致可分为五个类型,其类内误差平方和分别为46.93, 42.51, 55.47, 86.56, 13.73。对五类土壤的pH,有机质、代换量、 SiO_2 、 Al_2O_3 、 Fe_2O_3 、 CaCO_3 和粘粒含量八项主要指标类型间差异进行方差分析的结果(表3)表明,除pH、代换量和 SiO_2 含量未达类型间F测验显著水平外,粘粒含量达 $F_{0.05}$ 显著水平,其他四项指标均达 $F_{0.01}$ 的极显著水平。这表明五类土壤理化性状间存在着明显可辨的差异。

表 3 用 8 项主要指标进行的土壤类型间差异方差分析¹⁾

Table 3 Analysis of variance between soil types by 8 important indexes

主要指标 Important indexes	变 异 来 源 ²⁾ Variant source				F	差异显著性 ³⁾ Differential significance	
	类 型 间 Between soil types		误 差 Error			F _{0.05}	F _{0.01}
	SS ³⁾	MS ⁴⁾	SS	MS			
pH	2.37	0.59	5.20	0.18	3.31	—	—
有机质(%)	15.43	3.86	5.03	0.17	22.24	—	**
CEC (mcq/100g)	184.02	46.01	406.57	14.02	3.28	—	—
SiO ₂ (%)	208.70	52.17	317.47	10.95	4.77	—	—
Al ₂ O ₃ (%)	32.38	8.10	9.30	0.32	25.24	—	**
Fe ₂ O ₃ (%)	7.03	1.76	1.94	0.067	26.21	—	**
CaCO ₃ (%)	786.96	196.74	240.93	8.31	23.68	—	**
物理性粘粒 (%) (<0.01 mm)	997.43	249.36	973.31	33.56	7.43	*	—

注: 1) 根据表 1 数据计算而来;

2) 供试样品个数 $n = 34$, $DF_{总} = 33$, $DF_{类型} = 4$, $DF_{误差} = 29$;

3) SS 表示平方和;

4) MS 表示均方;

5) $F_{0.05}$ 和 $F_{0.01}$ 的显著性水准分别为 5.765 和 13.83; *显著, **极显著。

表 4 是五类土壤各指标的类平均值和标准差。根据表中各类土壤 CaCO₃ 和粘粒含量的多少将五类土壤谓之低钙粘质土、中钙粘壤土、无(极少)钙粘壤土、高钙壤质土及中钙砂壤土。现将按图 2 划分的五类土壤分述如下。

1. 低钙粘质土: 这类土壤包括 13 个剖面。剖面物质的平均含量表现为 <0.01 毫米的粘粒含量最高 ($\bar{x} = 51.25\%$, 变幅 43.63—58.87%), Al₂O₃ 和 Fe₂O₃ 含量也居五类土壤之首。CaCO₃ 含量仅 0.76% (变幅 0—4.05%), pH 值 7.77, 呈中性至微碱性反应。类内除 5 号剖面外, 绝大多数无田间石灰反应。从图 3 可以看出, CaCO₃ 和粘粒的分布在第二层有明显的聚积现象, 说明剖面有一定的淋溶作用。此外, 这类土壤绝大多数分布在海拔 1000—1600 米的丘陵山地, 土壤分布区年平均降水量约 700—750 毫米, 自然植被主

表 4 五类土壤理化性状的

Table 4 Means (\bar{x}) and standard deviations (S) of 13

土壤类型 Soil type	剖面数 Profile No.	pH	有机质 (%) O. M.	全氮 (%) Total N	全磷 (%) Total P ₂ O ₅	全钾 (%) Total K ₂ O	速效磷 (ppm) Available P ₂ O ₅
低钙粘质土	13	7.77±0.33	0.68±0.35	0.073±0.027	0.056±0.0095	2.33±0.13	5.86±2.68
中钙粘壤土	9	8.09±0.105	0.704±0.41	0.068±0.027	0.064±0.0115	2.22±0.104	9.31±9.19
无(极少)钙粘壤土	6	7.33±0.85	2.41±0.65	0.15±0.033	0.043±0.011	2.05±0.28	4.58±1.72
高钙壤质土	4	8.005±0.20	0.43±0.15	0.045±0.006	0.059±0.0077	1.88±0.19	12.02±10.55
中钙砂壤土	2	8.04±0.085	1.01±0.05	0.092±0.012	0.085±0.0057	2.32±0.0071	33.8±36.54

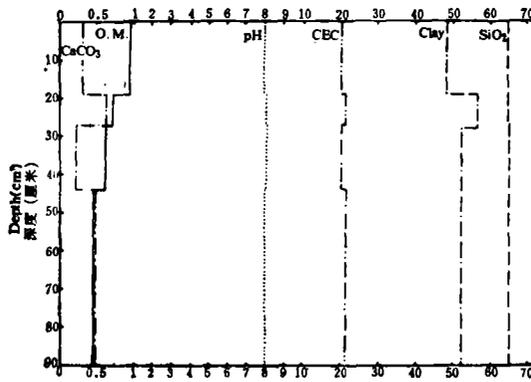


图 3 19 号剖面的物质剖面分布

Fig. 3 Distribution of some materials in No. 19 soil profile

要有华山松 (*Pinus armandii* Franch) 等木本植物及马兰 (*Kalimeris* Cass) 和苔草 (*Cyperus* L.) 属一些草本植物。植被覆盖度较高。由此可见, 这类土壤属森林草原淋溶土壤。

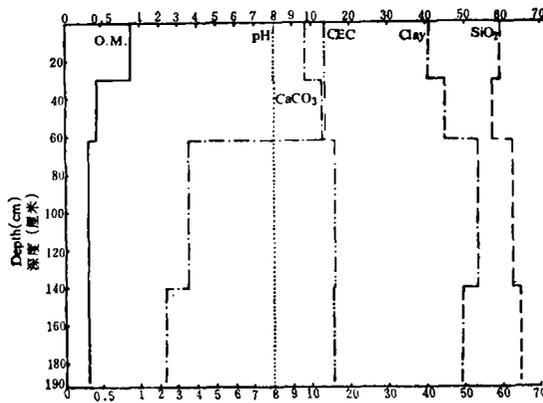


图 4 2 号剖面的物质剖面分布

Fig. 4 Distribution of some materials in No. 2 soil profile

类平均值(\bar{x})和标准差(S)

physical and chemical properties of five soil types

速效钾 (ppm) Available K ₂ O	代换量 (meq/100g) CEC	SiO ₂ (%)	Al ₂ O ₃ (%)	Fe ₂ O ₃ (%)	CaCO ₃ (%)	物理性粘粒 (%) Clay (<0.01mm)
117.44±17.81	18.53±1.92	64.14±1.62	15.07±3.41	5.85±0.25	0.76±1.3	51.13±4.35
118.38±11.21	16.02±2.31	62.44±2.79	13.99±0.36	5.29±0.16	4.80±2.34	45.9±3.83
99.59±23.43	20.84±3.56	64.8±2.31	14.76±0.98	5.42±0.39	0.082±0.17	47.48±5.10
88.11±8.66	14.94±9.13	56.51±8.03	11.97±0.79	4.40±0.28	15.78±7.67	41.23±12.61
236.28±49.11	12.14±2.58	63.88±1.80	13.71±0.47	5.02±0.021	4.69±0.22	29.19±4.63

2. 中钙质粘壤土: 这类土壤包括 9 个剖面, 其平均粘粒含量 45.9% (变幅 39.91—50.87%), CaCO_3 含量 4.8% (变幅 0.12—7.06%), pH 8.09 (变幅 7.92—8.26), 呈弱碱性到碱性反应。从 2 号剖面各土层物质分布(图 4) 看, 这类土壤有轻度的物质淋溶现象。类内剖面多分布在海拔 940—1200 米的低山丘陵地区, 该区年平均降水量在 700 毫米以下, 自然植被由菊科一些植物如苦蕒菜 (*Lxeris cass*) 等和禾本科小草共同组成。由此看, 这类土壤属弱淋溶草原土壤。

3. 无(极少)钙粘壤土: 类内包括 6 个土壤剖面。这类土壤 pH 值及 CaCO_3 含量是五类中最低的, 而有机质和代换量居五类之首。除 12 号剖面在 61 厘米土层以下有较强的石灰反应外, 其他剖面通层均无石灰反应。从图 5 中可看到粘粒和 SiO_2 在土层下部发生淀积。证明这类土壤经受较强的淋溶作用。从土壤地理分布看, 这类土壤分布于海拔 1400—1920 米的山地, 该区年平均降水量约 750—800 毫米, 植被属森林和森林草原植被, 主要植物有华山松, 箭竹 (*Sinarundinaria nitida* (Mitf) Nakai), 杜鹃花 (*Rhododendron* L.), 秦岭小檗 (*Berberis circumerrata* Schneid), 虎榛子 (*Ostryopsis davidiana* Decne), 蔷薇 (*Rosa* L.), 唐松草 (*Thalictrum aquilegifolium* var. *sibiricum* Regel et Tiling) 以及苔草属 (*Carex* L.) 和莎草属 (*Cyperus* L.) 植物。本区植被覆盖度最高, 植物种类丰富。土

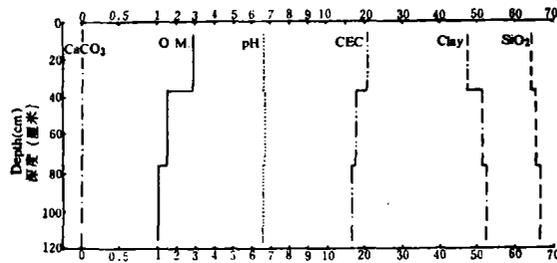


图 5 32 号剖面的物质剖面分布

Fig. 5 Distribution of some materials in No. 32 soil profile

壤呈微酸性至中性反应, 属森林或森林草原淋溶土壤。

4. 高钙壤质土: 这类土壤共 4 个剖面, 其 CaCO_3 含量最高 ($\bar{x} = 15.78\%$, 变幅 7.02—25.45%), 有机质、全氮等养分含量最低。 CaCO_3 和粘粒含量有明显的表聚现象(图 6), 类内所有剖面通层都有强的石灰反应。从剖面的地理分布看, 这类剖面多分布于 1000—

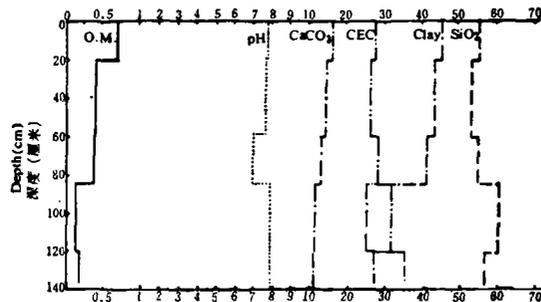


图 6 23 号剖面的物质剖面分布

Fig. 6 Distribution of some materials in No. 23 soil profile

1400 米以下的丘陵地区,年平均降水量 600—700 毫米,自然植被主要有菊科蒿属 (*Artemisia* L.)植物和禾本科小草。土壤无淋溶作用,属草原非淋溶性土壤。

5. 中钙砂壤土: 仅两个剖面,分布于海拔 1000 米左右的河谷川坎区,系河流冲积物形成的土壤。这类土壤粘粒含量最低(平均为 29.19%),土壤为砂壤质。这两个剖面养分含量,特别是速效养分的含量显著高于其他类型土壤,表现出熟化土壤的特征。整个土壤剖面中物质的分布(图 7)表明其没有淋溶作用。剖面分布区降水量 600—700 毫米,自然植被以菊科,禾本科植物为主。根据这类土壤所处地理位置及剖面特性,应属于河谷川坎潮土类型。

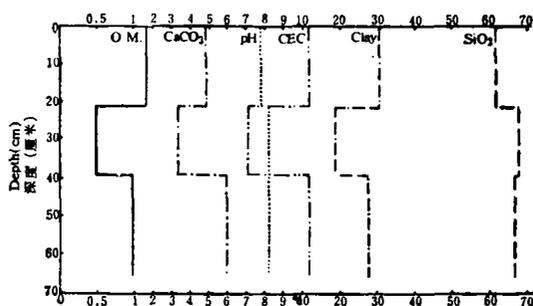


图 7 8 号剖面的物质剖面分布

Fig. 7 Distribution of some materials in No. 8 soil profile

综上所述,除少数剖面(如 5 号和 17 号剖面)外,绝大多数剖面的分类结果基本符合发生分类原则。因此,本文数值分类的五类土壤在发生分类上可相当于棕壤、淋溶褐土、典型褐土、碳酸盐褐土和潮土。

三、结 语

国内外一些研究者曾对土壤数值分类进行过大量研究^[2,3,5-8],但迄今还很少应用于土壤分类实践。本文对成县 34 个土壤剖面的数值分类虽然属于物质分类,但其结果与发生分类基本吻合,每一类型都体现土壤一定的发生学特点而自成一类,并且各类土壤都具有实际的分类意义。这表明,只要土壤指标选择适当,物质分类是可以揭示土壤发生学基本属性的。

我们的工作只是初步尝试。但我们相信,随着方法的不断完善,分类指标选择的更加合理,数值分类必定会推动土壤分类学的发展,使之更科学、更客观地反映土壤的本质。

参 考 文 献

- [1] 方开泰、潘恩沛, 1982: 聚类分析。地质出版社, 96—104 页。
- [2] 李天杰, 1979: 数学方法在土壤分类中的应用。《土壤分类及土壤地理论文集》, 浙江人民出版社, 241—247 页。
- [3] 刘多森、徐琪、陆彦椿, 1980: 聚类分析在太湖地区水稻土物质分类上的应用。土壤学报, 第 17 卷 4 期, 374—381 页。
- [4] 於崇文等, 1980: 数学地质的方法与应用, 地质与化探中的多元分析。78—114 页, 冶金出版社。

- [5] 熊国炎, 1980: 两组线性判别分析在土壤分类中的应用。土壤, 第5期, 177—181页。
- [6] Bidwell, O. W. and Hole, F. D., 1964: An experiment in the numerical classification of some Kansas soil. Soil Sci. Soc. Amer. Proc., 28: 263—268.
- [7] Sarkar, P. K. Bidwell, O. W. and Marcus, L. F., 1966: Selection of characteristic for numerical classification of soil. Soil Sci. Soc. Amer. Proc., 30: 269—272.
- [8] Webster, R., 1977: Quantitative and Numerical Methods in Soil Classification and Survey. Oxford University Press.

STUDY ON NUMERICAL CLASSIFICATION OF SOILS IN CHENGXIAN COUNTY OF GANSU PROVINCE

Li Wenxue and Liu Dongshun

(Institute of Land Survey and Design, Gansu Province)

Feng Xuezhi

(Institute of Glacier, Frozen Soil and Desert, Academia Sinica, Lanzhou)

Summary

Study on numerical classification for 34 soil profiles in Chengxian County of Gansu Province was carried out by means of a cluster analysis method of standardized variate-oblique distance-sum of square error. Indexes used in numerical classification include soil pH, organic matter (O. M.), total N, total P, total K, available P, available K, cation exchange capacity (CEC) SiO_2 , Al_2O_3 , Fe_2O_3 , CaCO_3 and physical clay (<0.01 mm). Results obtained from this study showed that 34 soil profiles used in the experiment may be divided into five soil groups. From variance analysis for eight main indexes (i.e. pH, O.M., CEC, SiO_2 , Al_2O_3 , Fe_2O_3 , CaCO_3 and clay) of the five soil groups shows that the differences of indexes between soil groups are all significant ($F_{0.05}$) or very significant ($F_{0.01}$) respectively except pH, CEC and SiO_2 content. In addition, there are also significant differences between soil groups in geographical distribution, vegetation types and calcareous reaction in fields etc.. Each soil group has its marked genetic characters. Compared with genetic classification system, the five soil groups belong to brown earth, leached drab soil, ortho-drab soil, carbonate drab soil and cultivated fluvio-aquic soil respectively.