Journal of Chromatography, 295 (1984) 31–47
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 16,748

APPLICATION OF PRINCIPAL COMPONENTS ANALYSIS TO THE EVALUATION AND SELECTION OF ELUENT SYSTEMS FOR THE THIN-LAYER CHROMATOGRAPHY OF BASIC AND NEUTRAL DRUGS

GIUSEPPE MUSUMARRA*, GIUSEPPE SCARLATA and GIUSEPPE CIRMA
Istituto Dipartimentale di Chimica e Chimica Industriale, Università di Catania, Viale Andrea Doria 6, 95125 Catania (Italy)
GUIDO ROMANO and SILVANA PALAZZO
Istituto di Medicina Legale e delle Assicurazioni, Università di Catania, Via Biblioteca 4, 95124 Catania (Italy)
and
SERGIO CLEMENTI and GIANFRANCO GIULIETTI
Dipartimento di Chimica, Università di Perugia, Via Elce di Sotto 10, 06100 Perugia (Italy)
(First received December 30th, 1983; revised manuscript received March 14th, 1984)

SUMMARY

The R_F values of 55 drugs in 40 eluent mixtures are reported. Principal components analysis of these data provides a four-significant-components model, which explains 92% of the total variance. This analysis, showing that the eluent mixtures cluster into different groups according to their information content, provides a reliable criterion for the choice of optimal eluents.

Four eluent mixtures [ethyl acetate–methanol–30% ammonia (85:10:5), cyclohexane–toluene–diethylamine (65:25:10), ethyl acetate–chloroform (50:50) and acetone with the plate dipped in potassium hydroxide solution], chosen on the basis of the above criterion and of the R_F reproducibility, provide a two significant principal components model that can be used for the identification of unknown samples.

INTRODUCTION

The advantages of thin-layer chromatography (TLC) as a sensitive, simple and rapid method for the identification of organic compounds are well known. However, the applications of TLC in the identification of drugs in forensic toxicology and related fields have been severely limited by the problems related to (a) the choice of an objective criterion (*i.e.*, an appropriate statistical approach) that utilizes the information provided by the R_F values in different eluent systems to achieve the identification of unknowns; (b) the selection of the minimum number of suitable eluent systems (each providing a different piece of information); and (c) the poor reproducibility of R_F data in some eluent mixtures.

Extensive work in this area has been carried out by Stead *et al.*[1], who reported

0021-9673/84/$03.00

the R_F values of almost 800 basic, neutral and acidic drugs in eight carefully standardized TLC systems, ordering the drugs according to their increasing R_F values in each eluent, in order to facilitate the identification of unknown samples. The same group[1–4] has also investigated the individual information for each eluent system and the correlation between the systems using the discrimination power.

The possibility of using information theory for characterizing TLC separations has been investigated by Massart[5] and the merits of different solvents used for the separation of the same group of compounds have been compared. Application of numerical taxonomy techniques to the choice of optimal sets of solvents in TLC has also been reported[6,7] and paper and thin-layer chromatographic separations of phenolic compounds were classified into clusters according to their selectivities[8].

Following previous applications of multivariate analysis to gas chromatographic[9,10] and TLC[11] data, we have recently pointed out the potential of principal components analysis (PCA) as a suitable statistical approach both for the identification of unknown samples and for the evaluation of the information content of the eluent systems[12,13]. PCA is able to reduce the number of measurements to 'object' scores (θ parameters) that characterize the compounds in a two- or three-dimensional space, allowing a graphical representation that makes the identification of unknowns easier with respect to earlier approaches based on R_F values reported in tables[1]. Moreover, in contrast with previous procedures defining the information content of each single eluent mixture as if it were to be used alone[5–7] or correlating two systems at a time[1–4], PCA gives a direct measure of the spanning properties of each system in combination with the others, thus directly providing information on the minimum number of systems that are needed and the criterion for their selection. As the interdependence of TLC data is well known[14], the superior ability of PCA over regression methods in detecting multivariate patterns is expected (for a comparison of PCA with other approaches adopted in TLC, see ref. 12).

The application of PCA to R_F data for 54 drugs in eight eluent mixtures[12] and 596 basic and neutral drugs in four eluent mixtures[13] allowed the characterization of the drug on a plane by two principal component parameters (θ values), leading to a drastic restriction of the range of inquiry to a few candidates and, in many instances, to the unambiguous identification of the drug. In both of the examined cases, however, the principal component parameters characteristic of the eluent mixtures (β values) indicated grouping of the eluents, with the eluents in each group providing approximately the same information.

Following these studies, aimed at the development of the applications of TLC as a cheap, rapid and reliable method for the identification of organic compounds, we report here the PCA of the R_F values of 55 basic and neutral drugs in 40 solvent mixtures with the purpose of selecting the minimum number of eluent systems having the maximum information content.

The drugs examined, which belong to various classes of compounds (tranquillizers, analgesics, natural and synthetic opiates, alkaloids, anthistamines, local anaesthetics, etc.) differing in their structural and biological properties, can all be detected with Dragendorff reagent. The eluent mixtures were chosen from those available in the literature and include those already analysed by PCA[12,13]. In order to achieve a rapid determination and to improve sensivity and reproducibility, silica gel HPTLC plates were used.

TABLE I

CHARACTERISTICS OF ELUENT MIXTURES 1–40

No.	*Eluent mixture (v/v)*	*Plate**	*Reproducibility***
1	Toluene–acetone–ethanol–30% ammonia (45:45:7:3)	a	+ +
2	Ethyl acetate–benzene–methanol–30% ammonia (60:35:6:2.5)	a	+
3	Benzene–dioxane–ethanol–30% ammonia (50:40:7.5:2.5)	a	+
4	Methanol–30% ammonia (100:1.5)	a	+
5	Benzene–isopropanol–methanol–30% ammonia (70:30:20:5)	a	+ +
6	Ethyl acetate–methanol–30% ammonia (85:10:5)	a	+ + +
7	Acetone–7.5% ammonia (90:10)	a	+ + +
8	Cyclohexane–toluene–diethylamine (65:25:10)	a	+ + +
9	Cyclohexane–toluene–diethylamine (75:15:10)	a	+ + +
10	Cyclohexane–benzene–methanol–diethylamine (70:20:10:5)	a	+ + +
11	Chloroform–acetone–diethylamine (50:40:10)	a	+ +
12	Cyclohexane–chloroform–diethylamine (50:40:10)	a	+ + +
13	Benzene–ethyl acetate–diethylamine (50:40:10)	a	+ +
14	Xylene–methyl ethyl ketone–methanol–diethylamine (40:40:6:2)	a	+ +
15	Diethyl ether–diethylamine (95:5)	a	+ +
16	Ethyl acetate–chloroform (50:50)	a	+ + +
17	Ethyl acetate–chloroform (50:50)	b	+ +
18	Butanol–methanol (40:60)	a	+ + +
19	Butanol–methanol (40:60)	b	+ + +
20	Chloroform–methanol (90:10)	a	+ +
21	Chloroform–methanol (90:10)	b	+
22	Acetone	a	+ + +
23	Acetone	b	+ + +
24	Acetone	c	+ +
25	Benzene–acetonitrile (70:30)	a	+ + +
26	Benzene–acetonitrile (70:30)	b	+ +
27	Benzene–tetrahydrofuran (80:20)	a	+ + +
28	Benzene–tetrahydrofuran (80:20)	b	+ + +
29	Chloroform–ethyl acetate–methanol (40:40:20)	a	+ + +
30	Chloroform–ethyl acetate–methanol (40:40:20)	b	+ +
31	Chloroform–*n*-hexane–methanol (65:25:10)	a	+ +
32	Chloroform–*n*-hexane–methanol (65:25:10)	b	+ +
33	Dichloromethane–methanol (95:5)	a	+ +
34	Dichloromethane–methanol (95:5)	b	+
35	Chloroform–methanol (75:25)	a	+ + +
36	Chloroform–methanol (75:25)	b	+ +
37	Acetic acid–ethanol–water (30:60:10)	a	+ + +
38	Ethyl acetate–dimethylformamide–ethanol (86.5:12.5:1)	a	+
39	Methanol–acetone–triethanolamine (50:40:1.5)	a	+
40	Chloroform–acetone–methanol–triethylamine (30:40:10:20)	a	+ +

* a, Not treated; b, dipped in 0.1 *M* potassium hydroxide methanolic solution and dried; c, after application of the drugs, the plate was kept for 30 min in a tank saturated with 30% ammonia solution and then transferred into the elution tank.

** + + +, All measurements for all compounds deviating less than 7% from the average; + +, some individual measurements for some of the compounds deviating between 7 and 14% from the average; +, some individual measurements for some of the compounds deviating more than 14% from the average.

TABLE II

R_F × 100 VALUES FOR COMPOUNDS 1–55 IN ELUENTS 1–40

No.	*Compound*	*Eluent*																	
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
1	Amidopyrine	65	57	66	73	83	66	81	28	19	18	76	61	64	46	46	13	18	72
2	Amitriptyline	77	77	81	55	88	81	83	58	57	45	79	73	77	34	69	2	9	22
3	Atropine	20	16	29	23	62	33	32	4	2	13	47	25	42	18	12	0	0	5
4	Benzphetamine	91	90	87	77	92	87	92	73	73	72	85	81	87	81	94	36	68	58
5	Benzydamine	65	63	72	48	85	75	76	42	37	31	71	62	69	24	54	0	2	15
6	Biperiden	91	90	87	68	92	87	92	73	72	64	85	81	86	68	94	11	40	40
7	Brompheniramine base	52	41	68	43	85	61	70	38	33	33	66	56	58	24	49	2	4	11
8	Brompheniramine maleate	58	43	66	42	83	66	70	37	34	28	65	57	61	22	47	0	0	8
9	Bupivacaine	83	84	84	80	86	83	90	45	41	32	84	74	81	69	73	22	56	69
10	Caffeine	55	42	52	68	77	54	73	8	5	13	60	30	51	41	20	13	12	54
11	Chlorpheniramine	57	42	65	43	82	66	70	37	34	25	66	58	62	24	47	0	0	9
12	Chlorpromazine	77	76	81	52	87	82	84	52	51	39	75	70	76	37	62	1	7	22
13	Cimetidine	25	10	13	67	61	30	69	0	0	6	15	0	2	14	0	0	0	51
14	Clemastine	67	55	80	48	88	72	79	51	51	45	77	69	72	26	67	0	6	14
15	Cocaine	81	82	81	71	87	82	89	46	41	38	81	72	80	52	72	6	24	30
16	Codeine	38	31	44	39	71	43	59	12	9	16	49	29	36	22	14	0	0	15
17	Cyclizine	71	72	80	64	85	80	82	49	47	40	75	71	73	39	59	2	9	34
18	Desipramine	50	42	57	26	78	57	55	27	25	24	61	52	52	23	32	0	16	7
19	Desmethyldiazepam	64	64	67	77	82	73	86	7	5	18	66	20	44	65	27	39	34	85
20	Diamorphine	55	51	62	46	82	63	75	20	13	21	67	50	59	26	29	0	2	19
21	Diazepam	76	79	80	78	85	80	88	28	21	29	80	61	75	72	54	54	50	85
22	Diphenhydramine	70	72	79	58	85	80	82	49	44	38	76	69	74	32	59	0	7	22
23	Flunitrazepam	74	78	79	73	85	80	89	17	11	22	80	50	72	72	43	48	43	83
24	Flurazepam	70	72	78	71	86	80	88	33	25	26	78	61	74	50	47	2	9	45
25	Haloperidol	70	74	80	75	85	83	90	18	14	20	74	42	66	45	43	1	6	47
26	Hydroxyzine	58	45	66	76	82	62	83	12	8	22	65	40	48	39	24	1	6	50
27	Imipramine	67	67	80	47	84	81	82	54	53	38	76	69	74	30	62	1	6	16
28	Isoxsuprine base	62	57	64	78	78	70	89	6	3	15	62	13	41	47	19	5	12	57
29	Ketamine	77	79	80	76	86	79	89	41	33	32	81	66	76	67	66	27	37	66
30	Lignocaine	77	79	80	73	86	80	89	35	30	28	84	73	77	66	64	25	54	68
31	Lignocaine base	78	82	83	76	87	78	88	40	29	27	87	70	79	67	66	29	63	69
32	Lorazepam	47	34	53	77	73	46	81	2	0	12	52	7	22	47	8	28	15	85
33	Mebeverine	85	90	90	65	90	85	90	43	33	38	88	70	87	62	76	5	29	29
34	Methadone	85	84	88	48	89	83	88	63	64	48	85	73	86	38	86	1	10	13
35	Methylamphetamine	45	33	49	31	80	46	46	33	31	29	61	54	51	23	41	0	1	7
36	Methylphenidate	70	66	80	64	88	73	83	43	36	37	77	65	71	34	59	2	12	31
37	6-Monoacetylmorphine	46	39	52	46	76	57	71	11	7	13	55	28	42	25	17	0	0	18
38	Morphine	18	9	15	39	56	20	42	2	0	5	20	2	8	13	3	0	1	16
39	Naloxone	48	40	62	75	79	48	81	15	11	22	52	26	40	60	21	18	21	67
40	Orphenadrine	74	68	83	58	88	76	83	50	49	43	79	69	75	31	66	1	11	20
41	Papaverine	68	66	76	79	88	71	84	12	7	18	78	56	62	54	30	28	41	76
42	Pentazocine	72	66	81	65	87	76	87	22	18	26	69	41	54	39	44	2	11	32
43	Pericyazine	57	34	59	68	85	54	81	7	4	16	59	21	40	27	17	1	2	36
44	Pethidine	64	54	76	57	88	69	78	41	37	35	73	62	64	27	56	1	6	25
45	Phenacetin	64	58	62	79	87	66	84	4	1	13	68	18	41	58	24	41	40	86
46	Phenazocine	76	75	83	74	88	79	89	23	17	27	70	42	55	54	43	11	22	51
47	Phenazone	66	53	70	73	86	65	80	30	22	24	77	60	66	45	54	15	21	68
48	Phendimetrazine base	65	56	75	64	86	67	78	40	37	38	72	63	62	36	56	12	19	45

19	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*	*33*	*34*	*35*	*36*	*37*	*38*	*39*	*40*
73	58	62	59	67	62	13	23	11	10	59	63	50	54	44	34	85	82	43	47	61	76
39	37	48	12	26	59	2	15	3	7	24	52	35	54	16	27	64	75	67	38	36	88
8	2	4	0	1	5	0	0	1	1	1	4	2	10	0	2	8	21	41	3	11	48
83	66	74	69	88	78	31	77	39	70	69	85	62	84	43	76	85	87	81	59	58	92
26	28	35	7	16	50	0	4	0	3	13	37	23	43	8	16	48	67	52	25	29	79
65	48	64	44	81	76	9	53	15	48	45	76	44	72	25	47	73	85	80	49	46	93
19	28	35	8	10	44	0	0	0	1	11	28	19	36	6	11	32	62	30	14	28	63
19	14	31	2	7	29	0	2	0	1	6	22	15	36	4	10	36	59	27	10	21	73
87	65	71	65	85	76	25	64	20	48	64	81	56	77	45	68	85	87	74	49	62	88
57	59	63	53	54	55	16	20	9	8	53	56	46	50	41	40	82	80	70	48	58	62
20	15	32	3	7	29	0	2	0	0	6	23	16	36	5	12	35	59	28	10	22	73
38	36	47	13	32	60	2	18	2	7	23	50	30	48	17	29	61	77	63	41	34	86
55	9	12	7	14	42	0	1	1	1	17	25	7	12	3	2	42	48	54	18	55	35
27	30	42	6	16	62	0	4	0	4	11	44	24	46	5	15	50	72	65	23	35	74
57	42	60	35	73	71	5	34	8	30	29	67	34	63	21	44	67	79	40	49	44	86
21	24	32	4	8	21	0	1	0	1	9	21	18	34	7	12	46	59	35	12	27	57
54	44	50	15	32	59	4	16	5	9	36	57	38	57	20	30	73	80	64	45	42	82
13	8	28	1	4	15	1	2	0	1	8	16	13	32	4	9	42	54	78	3	11	65
87	58	55	80	81	75	38	40	28	23	73	71	44	47	38	33	83	81	90	65	72	75
30	43	46	13	25	13	0	6	0	4	23	41	32	46	16	23	70	78	41	38	36	72
87	78	78	81	83	73	52	57	45	44	78	78	67	70	66	60	89	88	89	67	75	82
40	31	44	13	29	60	1	11	3	7	20	48	31	50	13	24	54	72	62	31	39	82
85	78	80	82	83	74	55	58	41	40	78	77	66	68	74	68	89	88	89	67	72	82
66	47	53	36	67	70	5	18	5	11	43	62	39	53	21	29	77	83	53	49	51	83
74	31	44	27	68	74	2	14	2	7	34	58	27	46	12	20	65	76	89	48	49	83
77	45	52	20	59	68	1	13	1	5	42	60	40	54	19	28	78	82	72	48	55	74
32	32	46	10	22	57	0	9	2	6	19	44	31	51	12	23	59	71	63	30	32	84
85	24	39	44	79	71	4	19	5	11	44	63	20	42	9	18	57	70	90	48	56	71
79	65	64	74	80	72	37	56	25	33	64	74	53	66	45	55	84	85	67	49	63	83
84	71	74	67	83	73	29	56	16	37	68	79	60	76	53	67	87	87	52	49	62	83
85	71	82	70	84	80	24	50	16	41	67	81	66	75	49	62	87	90	51	47	72	78
79	47	42	76	69	74	23	20	20	11	64	58	41	39	20	21	76	73	89	57	81	61
53	50	72	35	81	83	2	22	7	35	41	71	45	68	20	32	76	89	74	47	48	84
29	21	34	13	51	75	0	7	2	13	17	48	17	43	6	14	52	62	68	37	51	81
16	13	18	0	7	13	0	0	0	0	5	20	9	25	3	7	37	39	68	5	19	50
52	36	54	17	47	69	0	10	2	10	29	60	31	55	13	25	58	78	72	34	48	73
30	23	34	10	21	8	0	3	0	2	13	31	15	36	7	11	47	63	35	17	34	68
18	7	13	2	4	11	0	0	0	1	5	14	5	13	2	3	21	39	37	6	28	24
77	67	68	74	80	76	16	32	17	28	68	71	53	57	38	41	85	85	62	49	75	64
41	34	49	11	34	66	0	7	2	9	20	56	30	53	11	24	53	77	62	31	45	75
80	79	82	66	72	74	23	34	19	29	72	78	64	72	48	50	92	92	62	47	79	70
59	23	32	16	57	75	0	5	3	13	32	58	16	38	5	9	52	66	84	36	46	73
57	26	36	17	41	66	0	5	1	3	22	50	20	38	6	13	55	72	70	38	54	70
40	38	49	10	23	57	0	5	2	5	23	52	33	51	12	27	61	78	54	34	46	71
84	57	56	78	82	75	39	31	24	26	70	73	43	48	40	39	82	79	88	56	86	67
78	38	48	42	75	79	4	15	8	31	51	72	28	46	13	20	70	78	88	48	57	75
74	60	64	60	65	68	12	17	13	15	59	68	51	57	31	31	84	85	43	49	75	68
58	55	63	31	43	65	11	17	9	15	46	61	48	61	26	32	77	85	47	47	57	68

(*Continued on p. 36*)

TABLE II *(continued)*

No.	*Compound*	*Eluent*																	
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
49	Phendimetrazine bitartrate	64	54	75	62	86	67	78	39	36	38	71	63	62	34	55	6	14	40
50	Pirenzepine	15	5	15	39	59	18	31	0	0	5	29	3	8	12	1	0	1	8
51	Prazepam	81	83	86	83	88	81	89	41	31	35	83	66	82	75	74	65	67	86
52	Procaine	64	60	70	65	82	73	85	8	5	16	66	24	54	37	50	1	11	29
53	Promazine	64	53	79	47	87	70	79	43	41	42	73	63	66	26	56	0	6	12
54	Strychnine	47	33	50	28	78	47	49	13	9	18	56	42	40	22	13	36	0	7
55	Thenyldiamine	63	57	80	53	86	72	80	44	41	42	74	65	66	29	59	1	6	19

The R_F determinations were carried out independently in two laboratories where the eluent mixtures were freshly prepared.

EXPERIMENTAL

R_F Measurements

The eluent mixture compositions are reported in Table I, together with an estimate of the reproducibility of the R_F measurements for all 55 compounds.

Each drug (10 mg) was dissolved as the hydrochloride (except where stated otherwise) in methanol (10 ml). All drug solutions were freshly prepared and aliquots (4 μl, containing 4 μg of drug) were applied approximately 1 cm apart to 10 × 20 cm silica gel 60 F_{254} HPTLC plates (Merck). In some instances (b in Table I) the plates were dipped in 0.1 *M* potassium hydroxide solution and dried before application of the drugs; in other instances (c in Table I) the plates, after application of the drugs, were placed in a tank saturated with 30% ammonia solution and kept there for 30 min before being quickly transferred into the elution tank.

The solvents (100 ml) were placed in TLC tanks, which were sealed and allowed to equilibrate for at least 30 min before use. The systems were run for 5 cm from the baseline. The solvent front was marked and the plates were air-dried. The drugs were detected using Dragendorff spray reagent, after spraying with 10% sulphuric acid.

The R_F values were measured independently in two laboratories where the eluent mixtures were freshly prepared using commercial solvents often provided by different companies. The R_F values are uncorrected. The R_F × 100 data for compounds 1–55 in eluent mixtures 1–40 reported in Table II are averages of four determinations (two in each laboratory). The reproducibilities reported in Table I include both intra- and inter-laboratory errors.

Principal components analysis

PCA using the SIMCA method[15–18] and its application to the identification of drugs by TLC in different eluent systems[12,13] have been reported in detail. In the present instance, the matrix Y with the elements y_{ik}, contains R_F values where subscript i is used for the eluent mixtures (variables) and k for the compounds (objects). From this data matrix, the number of significant product terms A and the parameters

19	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*	*33*	*34*	*35*	*36*	*37*	*38*	*39*	*40*
57	51	61	30	39	65	6	13	6	11	42	60	46	60	24	30	76	82	46	46	54	68
15	4	12	0	3	7	0	0	0	0	3	8	4	11	0	1	18	43	18	5	27	29
88	89	82	84	85	80	55	59	58	60	30	84	73	74	62	58	92	91	91	73	89	79
53	23	40	25	57	70	1	7	2	9	20	55	16	40	7	21	41	70	48	33	54	68
26	29	43	8	22	57	0	6	1	4	13	43	26	47	8	21	46	72	50	18	35	72
11	27	36	3	5	14	0	2	0	0	7	19	19	38	10	15	48	66	33	8	15	61
32	41	56	9	21	60	0	6	1	6	20	51	38	57	14	26	60	81	53	24	43	72

α_i, β_{ia} and θ_{ak} in eqn. 1 are estimated by minimizing the sum of the cross-validated squared residuals ε_{ik}.

$$y_{ik} = \alpha_i + \sum_{a=1}^{A} \beta_{ia}\,\theta_{ak} + \varepsilon_{ik} \qquad (1)$$

In this model, α_i and β_{ia} are constants which are dependent only on the eluent mixtures and θ_{ak} are the compound-dependent parameters. The deviations from the model are expressed by the residuals ε_{ik}.

Before the PCA computation, the eluent values were autoscaled (see, *e.g.*, ref. 18), *i.e.*, the variables were given the same variance (fixed to unity). With this scaling, all variables were given the same importance in the PCA.

After a model has been determined with autoscaling, it can be refined by a reweighing of the variables, in this instance by multiplying each variable with its modelling power ψ_i, defined as

$$\psi_i = (1 - s_i/S_{iy}) \qquad (2)$$

Here s_i and S_{iy} are the residual standard deviations for variable i with A significant components and with $A = 0$ respectively. This means that variables for which the $\beta\,\theta$ terms contain no or little information will have modelling powers close to zero. Thus with this type of reweighing, such variables are given small weights.

RESULTS AND DISCUSSION

PCA with forty variables (autoscaled model)

The R_F values were arranged into a matrix (see Table II) with the compounds as 'objects' and the eluent mixtures as 'variables'. Each of the 2200 elements of the matrix is indicated in eqn. 1 as y_{ik}.

The variables (R_F values for each eluent mixture) were first autoscaled[18]. Each element was multiplied by the weight typical of the eluent (the reciprocal of the standard deviation of the variable) in order to give unitary variance to each eluent

TABLE III

WEIGHTS, α, β_1, β_2, β_3 AND RESIDUAL VARIANCES AFTER THREE PRINCIPAL COMPONENTS, $s_i^2(3)$ FOR VARIABLES (ELUENT MIXTURES) 1–40 (AUTOSCALED MODEL)

Eluent mixture	*Variable*					
	Weight	α	β_1	β_2	β_3	$s_i^2(3)$
1	0.0586	3.6948	−0.1664	−0.1703	0.0931	0.0341
2	0.0466	2.6958	−0.1677	−0.1536	0.0532	0.0842
3	0.0547	3.7615	−0.1512	−0.2010	0.1180	0.0521
4	0.0626	3.7899	−0.1619	0.1208	0.2468	0.0795
5	0.1278	10.5633	−0.1427	−0.1970	0.1269	0.1280
6	0.0605	4.0743	−0.1519	−0.1910	0.1260	0.0760
7	0.0691	5.3803	−0.1659	−0.0707	0.2959	0.0675
8	0.0519	1.5738	−0.0801	−0.2911	−0.1537	0.0832
9	0.0511	1.3541	−0.0664	−0.2940	−0.1513	0.1239
10	0.0738	2.0917	−0.0899	−0.2682	−0.1245	0.1696
11	0.0653	4.4863	−0.1535	−0.1913	0.0138	0.0983
12	0.0450	2.2829	−0.1037	−0.2679	−0.1691	0.0683
13	0.0500	2.9952	−0.1343	−0.2369	−0.0755	0.0629
14	0.0548	2.2420	−0.1856	0.0668	−0.0513	0.0816
15	0.0423	1.9423	−0.1224	−0.2514	−0.1002	0.0731
16	0.0620	0.6935	−0.1347	0.1699	−0.2526	0.1651
17	0.0525	0.9240	−0.1698	0.0730	−0.2311	0.1085
18	0.0383	1.5007	−0.1502	0.2126	0.0666	0.0396
19	0.0390	2.0400	−0.1655	0.1492	0.1744	0.0516
20	0.0469	1.8990	−0.1803	0.0822	−0.0898	0.1013
21	0.0525	2.5682	−0.1854	0.0109	−0.0504	0.1225
22	0.0354	1.1192	−0.1649	0.1759	−0.0408	0.0469
23	0.0337	1.5601	−0.1752	0.0994	0.1377	0.0974
24	0.0432	2.4662	−0.1674	−0.0136	0.2775	0.1207
25	0.0648	0.6454	−0.1460	0.1701	−0.2289	0.1057
26	0.0492	0.9545	−0.1737	0.0721	−0.2166	0.0883
27	0.0774	0.7033	−0.1527	0.1406	−0.2389	0.1216
28	0.0589	0.9151	−0.1702	0.0447	−0.2047	0.1603
29	0.0411	1.4651	−0.1739	0.1559	0.0063	0.0266
30	0.0466	2.4610	−0.1907	0.0296	0.1337	0.0266
31	0.0550	1.8779	−0.1825	0.0582	−0.1131	0.0989
32	0.0590	2.9119	−0.1855	−0.0665	−0.0639	0.0796
33	0.0540	1.1545	−0.1671	0.1327	−0.1870	0.0674
34	0.0531	1.5119	−0.1809	0.0354	−0.1949	0.0752
35	0.0485	3.0458	−0.1798	0.0676	0.0679	0.1334
36	0.0678	4.9634	−0.1772	−0.0189	0.1270	0.1692
37	0.0512	3.1144	−0.1162	0.0528	0.2079	0.5690
38	0.0546	2.0098	−0.1787	0.0836	0.1518	0.0804
39	0.0518	2.5075	−0.1532	0.1749	0.1492	0.1028
40	0.0712	5.1073	−0.1493	−0.1850	0.0960	0.1331

mixture. Weights for individual variables 1–40 are recorded in Table III*.

The PCA of the data matrix (Table II) gave a four significant principal components model, according to the cross-validation technique[17]. The first component

* It should be noted that the variable weights depend only on the range covered by the 55 average R_F values in each solvent, and they have nothing to do with the reproducibility of individual eluents, which is used only for the selection of the best eluent within groups of similar ones, and is not used in the statistical analysis.

explains 62% of the total variance, the second one a further 21%, the third one a further 6% and the fourth one a further 3%; the four components model then accounts for 92% of the total variance. However, the residual variances after three components reported in Table III are low for all variables (eluents) except 37, and do not vary significantly by adding a fourth principal component, showing that the three-component model, accounting for 89% of the total variance, describes satisfactorily the systematic behaviour of all eluents except 37. The latter eluent mixture, containing acetic acid, is peculiar and can be better modelled by adding a fourth component [$s_i^2(4)$ for 37 = 0.2367; *cf.* $s_i^2(3)$ = 0.5690 in Table III], which explains only a further 3% of variance. In other words, the fourth component, which has a very low information content, is required only to describe the peculiar behaviour of eluent mixture 37.

Table IV lists the θ_1, θ_2 and θ_3 values for each of the compounds 1–55, together with their residual standard deviations s_k from the three principal components model. The θ values indicate the position of each point along each new dimension defining the three-dimensional model, whereas the s_k values give the distance of each compound from the model. The critical distance for the compounds to be considered belonging to the model is evaluated by the appropriate statistical F-test[15]. In the present instance, according to the F-test at the 99% confidence level, compounds with residual standard deviations lower than 0.42 lie within the confidence interval around the three-dimensional model.

Table IV shows that benzphetamine (4), biperiden (6), cimetidine (13), desipramine (18), diamorphine (20), strychnine (54) and probably atropine (3) and methadone (34) are 'outliers'. These compounds might be omitted from further PCA, in order to obtain improved models. However, as our main objective is to utilize PCA for the identification of unknown drugs, we decided not to exclude any of the compounds in the four-variables model (see below).

Figs. 1 and 2, showing plots of θ_2 *vs.* θ_1 of θ_3 *vs.* θ_2 (*i.e.*, the projections of the points identifying each compound into the θ_1–θ_2 and θ_2–θ_3 planes) indicate the non-homogeneity in the examined set of drugs, which is a direct consequence of the analytical purpose of this study. In fact, the drug set was formed by selecting a few representatives from each group of drugs exhibiting the same pharmacological properties (analgesics, antidepressants, antihistamines, local anaesthetics, tranquillizers, antiulcer agents, parasympatholytics, anorectics, antispasmodics, central stimulants, neurological psycotonics, vasodilators) choosing, within each group, compounds of very similar chemical structure in order to find eluents able to differentiate between them.

In this context, Figs. 1 and 2 are not expected to provide much information for purposes of classification of drugs 1–55. However, we note that in both Figs. 1 and 2, the anthistamines 14, 17, 22 and 55 are grouped and very close to a compact group of anthistamines with almost identical chemical formulae (brompheniramine base 7, brompheniramine maleate 8 and chlorpheniramine 11).

Table III reports the 'loading' parameters, *i.e.*, the α, β_1, β_2, β_3 values for eluent mixtures 1–40, which define the model.

Fig. 3, where the β_2 values are plotted against β_1, shows that all eluents lie on the left-hand side with respect to the origin (*i.e.*, all eluents contain the same major information) and shows the existence of three groups of eluents: (A) including mix-

TABLE IV

θ_1, θ_2 AND θ_3 VALUES IN THE AUTOSCALED PRINCIPAL COMPONENTS MODEL USING 40 VARIABLES, RESIDUALS AFTER THREE PRINCIPAL COMPONENTS, $s_k(3)$ AND θ_1 AND θ_2 VALUES IN THE REFINED MODEL USING FOUR VARIABLES FOR OBJECTS (COMPOUNDS) 1–55

Object (compounds)		*Forty variables (autoscaled model)*			$s_k(3)$	*Four variables (refined model)*	
No.	*Name*	θ_1	θ_2	θ_3		θ_1	θ_2
1	Amidopyrine	− 2.3693	1.4677	0.2968	0.4166	−0.1229	0.3270
2	Amitriptyline	− 0.5093	−4.0292	0.1179	0.2017	−0.6354	−1.0528
3	Atropine	10.6440	1.8551	−2.5929	0.4314	1.9923	−0.0281
4	Benzphetamine	− 9.5790	−1.9289	−3.2061	0.5318	−1.9327	0.0166
5	Benzydamine	2.6264	−2.9522	0.1191	0.1633	−0.0640	−0.8519
6	Biperiden	− 5.9234	−3.9491	−0.9465	0.4485	−1.6587	−0.5777
7	Brompheniramine base	4.4553	−2.1312	−0.8103	0.2587	0.4587	−0.7227
8	Brompheniramine maleate	4.9836	−2.5432	−0.8661	0.2502	0.3732	−0.8047
9	Bupivacaine	− 7.2282	−0.0025	−1.1792	0.3211	−1.1688	0.2818
10	Caffeine	− 0.2783	3.5501	0.3311	0.3564	0.7077	0.6481
11	Chlorpheniramine	4.9171	−2.4060	−0.8724	0.2541	0.3732	−0.8047
12	Chlorpromazine	− 0.1869	−3.5479	0.3102	0.1997	−0.6040	−0.8892
13	Cimetidine	8.2743	5.4141	0.9945	0.5190	2.0337	0.2167
14	Clemastine	1.7668	−3.8825	0.1321	0.2285	−0.1365	−1.0177
15	Cocaine	− 2.8895	−2.4862	0.0256	0.3523	−0.9131	−0.1989
16	Codeine	6.7421	1.2736	−0.9447	0.2019	1.4901	−0.1641
17	Cyclizine	− 1.0372	−2.5300	0.4357	0.2384	−0.4997	−0.7972
18	Desipramine	6.0884	−1.2738	−1.4273	0.4435	0.8451	−0.5882
19	Desmethyldiazepam	− 4.0432	4.6477	0.8121	0.3349	−0.3002	1.4083
20	Diamorphine	2.7420	−0.3553	−0.0701	0.4643	0.6015	−0.2402
21	Diazepam	− 8.1973	3.0524	−1.9231	0.2915	−1.0268	1.2776
22	Diphenydramine	0.5246	−3.1781	0.4156	0.1396	−0.4555	−0.8725
23	Flunitrazepam	− 7.7974	3.7079	−1.5753	0.3783	−0.7796	1.3765

24	Flurezepam	− 1.8121	−0.8310	1.5522	0.2552	−0.5331	−0.0759
25	Haloperidol	− 0.7078	−0.2304	3.1188	0.2791	−0.3540	0.2019
26	Hydroxyzine	− 0.0334	1.5731	2.3635	0.2936	0.4553	0.3340
27	Imipramine	0.9033	−3.6976	−0.0207	0.1987	−0.5190	−1.0380
28	Isoxsuprine	0.3767	2.6586	3.2298	0.3435	0.1102	0.7219
29	Ketamine	− 6.0332	0.7190	−1.1066	0.2105	−0.9755	0.4303
30	Lignocaine	− 6.5615	0.7455	−1.3902	0.3337	−0.9083	0.5391
31	Lignocaine base	− 6.9310	0.7929	−1.2443	0.3738	−0.9815	0.5425
32	Lorazepam	− 0.3749	6.1158	1.6163	0.3328	0.7881	1.2918
33	Mebeverine	− 4.0739	−2.6375	0.8856	0.3400	−1.0126	−0.0840
34	Methadone	− 0.0577	−5.0801	0.6821	0.4340	−1.0038	−0.9010
35	Methylamphetamine	6.9624	−1.0981	−1.7991	0.4061	1.0351	−0.6161
36	Methylphenidate	− 0.3643	−2.2174	1.0822	0.1756	−0.3227	−0.4677
37	6-Monoacethylmorphine	5.4868	0.8367	−0.0443	0.3365	0.9761	−0.0715
38	Morphyne	10.8080	4.4970	−1.7291	0.3281	2.3844	0.1164
39	Naloxone	− 2.3986	4.4579	0.3613	0.3953	0.4802	0.9375
40	Orphenadrine	0.0980	−3.4659	0.5361	0.1630	−0.4094	−0.7938
41	Papaverine	− 5.1876	2.9837	0.0488	0.4073	−0.1560	−0.9892
42	Pentazocine	0.9640	−0.7217	2.5925	0.3432	−0.1270	0.0543
43	Pericyazine	2.8261	1.6035	2.5909	0.2303	0.9452	0.2735
44	Pethydine	1.0057	−2.1364	0.4522	0.2327	0.0586	−0.6992
45	Phenacetin	− 3.7577	5.1240	0.8193	0.3250	−0.0659	1.5592
46	Phenazocine	− 2.0328	0.4394	2.4230	0.3361	−0.4734	0.4079
47	Phenazone	− 2.6512	1.2234	0.4409	0.4199	−0.1277	0.3104
48	Phendimetrazine base	− 1.5666	−0.5424	−0.2288	0.3487	−0.1395	−0.2144
49	Phendimetrazine bitartrate	− 0.9180	−0.8465	0.0938	0.3590	−0.0347	−0.3632
50	Pirenzepine	11.2419	4.1889	−2.1254	0.2912	2.4883	0.1562
51	Prazepam	− 9.9708	2.6880	−2.6798	0.4208	−1.3996	1.2587
52	Procaine	1.8229	0.3931	2.0765	0.3886	0.2204	0.3329
53	Promazine	2.1692	−3.1209	−0.0900	0.1902	0.0109	−0.7770
54	Strychnine	6.2262	0.5811	−2.1316	0.4972	1.0789	0.5012
55	Thenyldiamine	0.8169	−2.7681	0.0476	0.2923	−0.0653	−0.7987

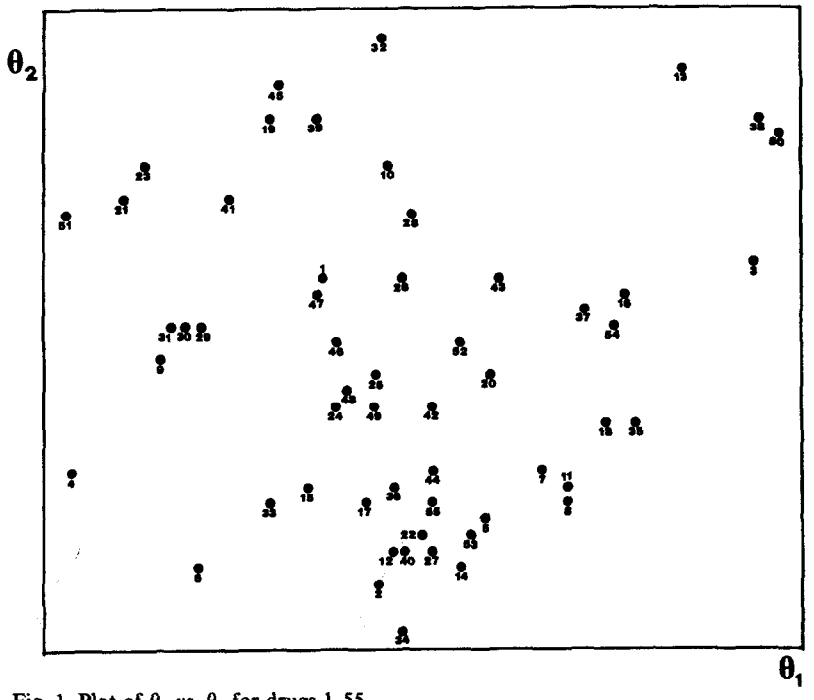

Fig. 1. Plot of θ_2 *vs.* θ_1 for drugs 1–55.

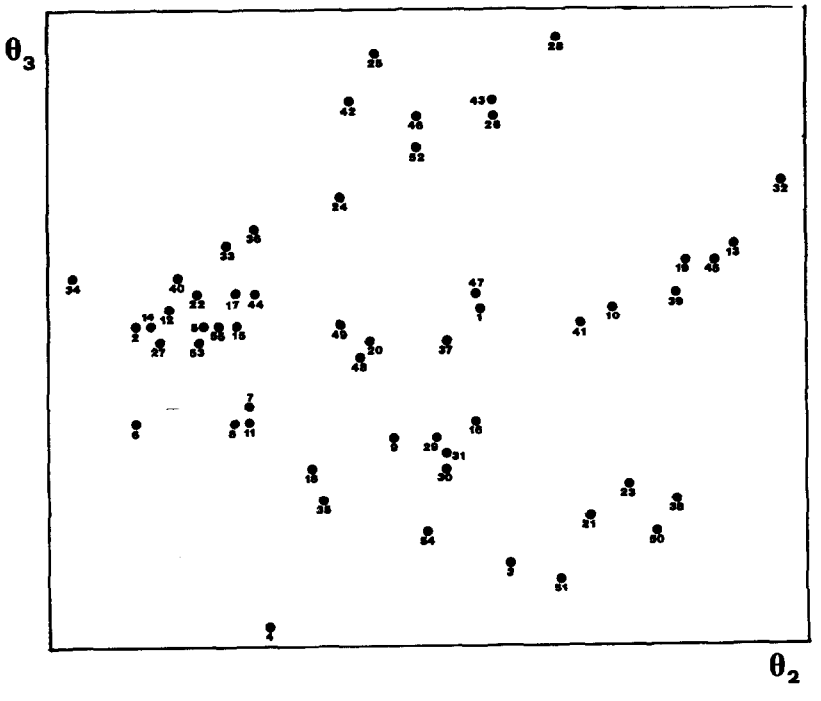

Fig. 2. Plot of θ_3 *vs.* θ_2 for drugs 1–55.

tures containing diethylamine (8–10, 12, 13 and 15); (B) eluents with 30% ammonia solution (1–3, 5 and 6), one with diethylamine (11) and that with triethylamine (40); and (C) all the others, which are mixtures of neutral organic solvents, plus 4 and 14, which contain low concentrations of 30% ammonia solution and diethylamine, respectively. Eluent mixture 7, containing more dilute ammonia solution (7.5%), lies between groups B and C. Fig. 3 also confirms the peculiar behaviour of eluent mixture 37, already indicated to contain its own individual information.

Treatment of the plate with potassium hydroxide solution results in a decrease in both β_1 and β_2 values (*cf.*, 22–23 and analogous eluent pairs), and a drastic decrease in β_2 is observed when, before elution, the plate is placed in a tank saturated with ammonia solution (eluent 24).

Fig. 4, where β_3 is plotted against β_2, shows a uniform distribution of the eluents around the origin (0,0), which lies approximately in the middle of the plot. Groups A and B, already evidenced by the β_2 *vs.* β_1 plot (Fig. 3), are present in this plot also, while eluents in former group C (Fig. 3) are now better differentiated by their β_3 values and a new group, D, can be identified in the lower right-hand corner of the plot.

The information provided by Fig. 3 and 4 is suitable for the evaluation and selection of eluent mixtures.

Evaluation and selection of eluent mixtures

The factors considered to be most important for the selection of suitable TLC

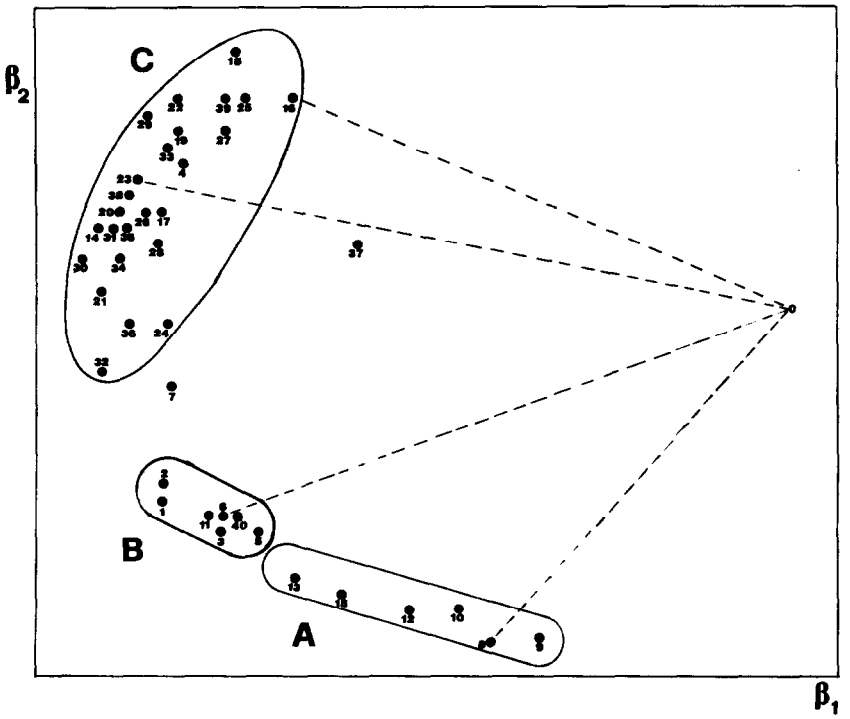

Fig. 3. Plot of β_2 *vs.* β_1 for variables 1–40; 0 indicates origin (0,0).

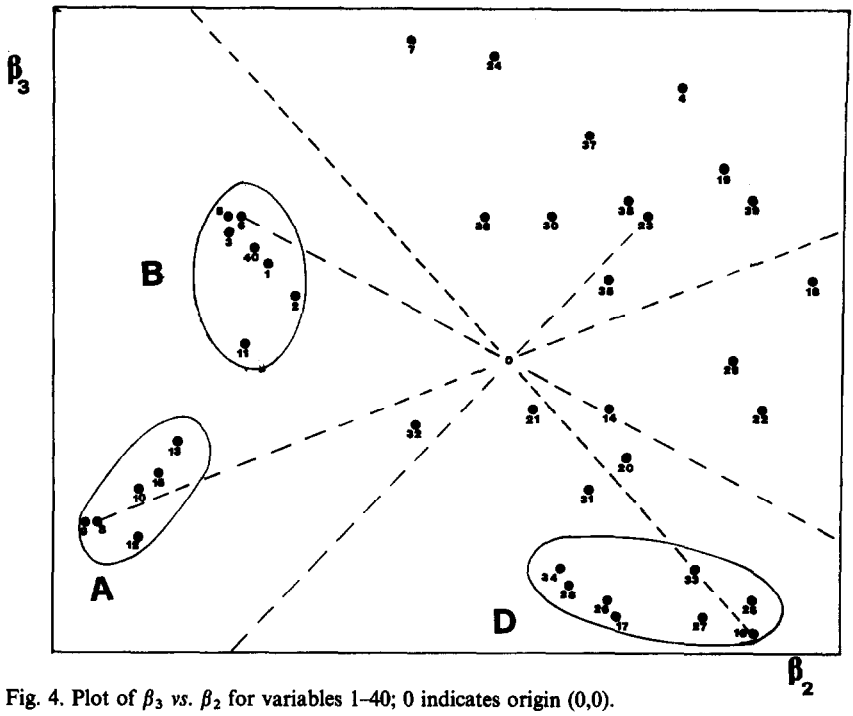

Fig. 4. Plot of β_3 *vs.* β_2 for variables 1–40; 0 indicates origin (0,0).

systems are as follows[4,19]: (1) distribution of chromatographic values over the useful range of the system; (2) correlation between systems when more than one is used; (3) speed; (4) reproducibility; and (5) sensitivity. In the present instance, the use of HPTLC plates improves the speed, the reproducibility and the sensitivity. The distribution of chromatographic values in the systems examined can be evaluated from Table II, while Table I reports an estimate of the R_F reproducibility in each eluent mixture for all the compounds examined. In this study, the evaluation of the different information contents of the eluent systems and the criteria for their selection are achieved by PCA.

We shall now discuss the information provided by PCA on how systems can be selected and how many of them are needed, and then propose a set of eluents on the basis of the above criteria and of the desired properties of 'optimum' eluents.

An important piece of information on which eluent mixtures are to be selected is provided by Fig. 4, which suggests that three eluents should be picked from groups A, B and D, respectively (each group providing different information). We note that group D includes eluents in which the R_F values for many compounds (basic drugs) are very close to zero, the R_F values not being distributed over a satisfactory range. However, we believe that a representative of this group should be included in the 'optimum' set as it is able to distinguish basic from neutral drugs.

The presence of the origin (0,0) in Figs. 3 and 4 gives a further criterion of selection: the lines joining the eventually selected eluents with the origin should not lie in the same or close directions. In fact eluents that lie on the same line are very

similar to each other (provide the same information): the closer they are to the origin the less systematic variation they have.

Another piece of information provided by PCA is the number of independent effects generating the data structure, *i.e.*, the number of significant principal components. As the number of significant groups is spanned by the same number of components, four eluents should be selected. In this respect, from a purely statistical point of view, as the fourth component is required to describe the behaviour of eluent 37, this system should be included in the 'optimum' set. However, owing to the peculiarity of this eluent mixture (containing a protic acid such as acetic acid), which has only a very small information content (only 3% of the total variance is explained by the fourth principal component), we decided not to include it in the 'optimum' set. As a consequence, the number of eluents selected should be limited to three. However, in consideration of the unsatisfactory R_F distribution in group D and of the fact that in Fig. 4 a significant region of the plot would remain uncovered by choosing only representatives of groups A, B and D, we decided to include four eluents in the 'optimum' set in order to improve the capacity for the identification of unknowns of the resulting principal component model.

On the basis of the above statistical considerations, of the R_F reproducibility (Table I), of the R_F distributions (Table II) and of the shape of the spots after elution, we chose eluents 6, 8 and 16 as representatives of groups B, A and D, respectively. As the fourth system, which should lie in a direction as different as possible from those of the three already selected, eluent 23 (acetone) was preferred to 19 (butanol–methanol, 40:60) for its simplicity (it is a pure solvent) and for the better shape of the spots. Eluents 4 and 38 were not taken into consideration owing to the poor R_F reproducibility (see Table I).

PCA with four variables (autoscaled model)

PCA was then repeated using only the selected variables (eluent mixtures 6, 8, 16 and 23) in order to check if this choice was really useful for our identification purposes, *i.e.*, if the information contained in these four eluents is representative enough of all possible systems. A two components model which explained 66% of the total variance was obtained by autoscaling. The modelling power ψ_2 for each variable, reported in Table V, permitted the calculation of new weights by which the refined model (see below) was worked out.

PCA with four variables (refined model)

The principal components model can now be refined by reweighing the variables. The new weights are obtained multiplying the weights for each variable listed in Table III by the modelling powers ψ_2 listed in Table V. The data analysis again provides a two components model (a plane), in which the first component explains 40% of the residual variance and the second one a further 35%; the model then shows that 75% of the residual variance after refinement is systematic. However, as part of the variation is now included in the weights, the fraction of variance remaining unexplained by the refined model is as low as 15%. Accordingly, the capability of the planar model with four eluents to describe the data set is comparable to that of the three-dimensional model with 40 systems and no significant loss of information is involved in the reduction of eluents from 40 to 4.

TABLE V

MODELLING POWER ψ_2 IN THE AUTOSCALED FOUR VARIABLES MODEL, AND REFINED WEIGHT, α, β_1 AND β_2 IN THE REFINED FOUR-VARIABLES MODEL FOR VARIABLES (ELUENT MIXTURES) 6, 8, 16 AND 23

Eluent mixture	*Variable*				
	ψ_2	*Refined weight*	α	β_1	β_2
6	0.6995	0.0423	2.8500	−0.6973	−0.1272
8	0.6745	0.0350	1.0615	−0.5114	−0.5779
16	0.5621	0.0349	0.3898	−0.2411	0.5906
23	0.6153	0.0207	0.9599	−0.4407	0.5487

The new weights together with the α, β_1 and β_2 values for the refined model are recorded in Table V.

In the β_1 *vs.* β_2 plot (Fig. 5), eluents 6, 8, 16 and 23 lie in the same position as in the plot obtained from the 40 variables model (Fig. 3), confirming that the selected systems have a different information content, as they are properly spread along the arch formed by the 40 eluents.

The principal components parameters reported in Table V could be used for calculating t_1 and t_2 values for unknowns from the equation

$$t_{ak} = \Sigma\,(100\,R_{F_i}\,w_i - \alpha_i)\,\beta_{ia} \tag{3}$$

where, in the present instance, $a = 1$ and 2. These t values could be compared with the θ_{ak} values of the samples used as training set in the four variables model (Table

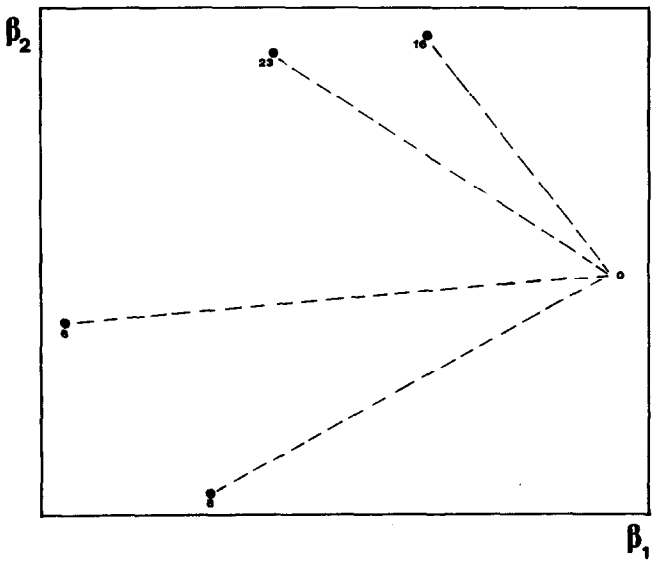

Fig. 5. Plot of β_2 *vs.* β_1 in the four variables refined model; 0 indicates origin (0,0).

IV) to identify unknowns, provided that they are included in the examined set (see refs. 12 and 13). The small number of compounds in this set, however, precludes the use of eqn. 3 for the general purpose of identifying any unknown, which is outside the scope of this study.

CONCLUSIONS

This work confirms that PCA is a suitable method for the evaluation and selection of eluent systems in TLC. On the basis of PCA and of other practical considerations (reproducibility, shape of the spot, etc.) we have proposed a minimum set of eluents that contains virtually all the information obtainable from a much larger set and could be conveniently used for the identification of unknowns.

We are aware that the selection of 'optimum' eluent mixtures represents a subjective choice. However, this decision can be made as 'objective' as possible by the use of an appropriate statistical procedure such as PCA, which gives a simple graphical representation of the relationships between the information provided by individual eluents.

ACKNOWLEDGEMENTS

We thank Professor Svante Wold (Umeå University, Sweden) for helpful discussions and for providing the SIMCA package; financial support (to S. C. and G. M.) by the Ministero della Pubblica Istruzione to a research project on chemometrics is also gratefully acknowledged.

REFERENCES

1 A. H. Stead, R. Gill, T. Wright, J. P. Gibbs and A. C. Moffat, *Analyst (London)*, 107 (1982) 1106.
2 A. C. Moffat and K. W. Smalldon, *J. Chromatogr.*, 90 (1974) 9.
3 P. Owen, A. Pendlebury and A. C. Moffat, *J. Chromatogr.*, 161 (1978) 187.
4 P. Owen, A. Pendlebury and A. C. Moffat, *J. Chromatogr.*, 161 (1978) 195.
5 D. L. Massart, *J. Chromatogr.*, 79 (1973) 157.
6 D. L. Massart and H. de Clerq, *Anal. Chem.*, 46 (1974) 1988.
7 H. de Clercq and D. L. Massart, *J. Chromatogr.*, 115 (1975) 1.
8 L. Nagels, *J. Chromatogr.*, 209 (1981) 377.
9 G. Blomquist, E. Johansson, B. Söderström and S. Wold, *J. Chromatogr.*, 173 (1979) 7.
10 B. Söderström, S. Wold and G. Blomquist, *J. Gen. Microbiol.*, 128 (1982) 1773.
11 D. Carbone, G. Musumarra, S. Occhipinti, G. Scarlata and S. Wold, *Ann. Chim. (Rome)*, 73 (1983) 183.
12 G. Musumarra, G. Scarlata, G. Romano and S. Clementi, *J. Anal. Toxicol.*, 7 (1983) 286.
13 G. Musumarra, G. Scarlata, G. Romano, S. Clementi and S. Wold, in preparation.
14 K. A. Connors, *Anal. Chem.*, 46 (1974) 53.
15 S. Wold and M. Sjöström, in B. Kowalski (Editor), *Chemometrics, Theory and Application (ACS Symposium Series*, No. 52), American Chemical Society, Washington, DC, 1977, p. 243.
16 C. Albano, G. Blomquist, D. Coomans, W. J. Dunn, III, U. Edlund, B. Eliasson, S. Hellberg, E. Johansson, B. Norden, M. Sjöström, B. Söderström, H. Wold and S. Wold, in A. Höskuldsson *et al.* (Editors), *Proceedings, Symposium i Anvent Statistik*, Copenhagen, NEUCC, RECAU, RECKU and the Danish Society of Theoretical Statistics, 1981, p. 183.
17 S. Wold, *Technometrics*, 20 (1978) 397.
18 S. Wold, in D. L. Massart, A. Dijkstra and L. Kaufman, *Evaluation and Optimization of Laboratory Methods and Analytical Procedures*, Elsevier, Amsterdam, 1978, pp. 385 and 423.
19 A. C. Moffat and B. Clare, *J. Pharm. Pharmacol.*, 26 (1974) 665.