Journal of Chromatography, 115 (1975) 1–7

CHROM. 8597

EVALUATION AND SELECTION OF OPTIMAL SOLVENTS AND SOLVENT COMBINATIONS IN THIN-LAYER CHROMATOGRAPHY

APPLICATION OF THE METHOD TO BASIC DRUGS

H. DE CLERCQ and D. L. MASSART

Farmaceutisch Instituut, Vrije Universiteit Brussel, Paardenstraat 67, B-1640 Sint-Genesius-Rode (Belgium)

(Received June 18th, 1975)

SUMMARY

A series of simple mathematical techniques for the evaluation of solvents and solvent combinations in thin-layer chromatography have been investigated. A strategy for the rapid selection of the optimum combination is proposed. It uses classification procedures based on calculation of the similarity between systems. The classification is carried out using a simple graph-theoretical procedure (Kruskal's algorithm) or numerical taxonomy. The selection of optimal sets from the clusters which appear in the classification is based on the information content as derived from Shannon's equation. The method has been applied to an R_F data set for basic drugs. It is concluded that these methods indeed allow the selection of optimal systems or combination of systems.

INTRODUCTION

Recently, attention has been paid to objective criteria for the evaluation of the separating ability of thin-layer chromatography (TLC) and paper chromatography (PC) systems. In this respect, the use of the information content[1] and the discriminating power[2] have been proposed for the evaluation of single systems or combinations of systems. This article critically investigates and compares these existing mathematical selection procedures. At the same time a strategy for optimizing combinations of more than one TLC system is discussed. A data library for the TLC separation of 100 basic drugs in eight systems published by Moffat and Smalldon[2] is used as an illustration.

THEORY

The discriminating power (*D.P.*) was introduced by Moffat and his co-workers[2–5] and has been extensively covered in a series of papers. We only recall to mind that two compounds are considered to be unresolved in a particular chromato-

graphic system if the difference between their R_F values does not exceed a certain value, the error factor E. To compute the discriminating power of a system in which N compounds are investigated, the total number, M, of matching pairs (within the limits of E) is counted. The $D.P.$ is then given by eqn. 1.

$$D.P. = 1 - [2M/N(N-1)] \tag{1}$$

Similarly, for a series of k systems in combination, and the chromatographic values for N compounds in each system, the discriminating power is given by an analogous formula, in which M is the number of pairs unresolved in all the systems.

Massart[1] used a concept of information theory, the information content, I, to obtain a numerical value representative of the separating ability of a chromatographic system. The R_F range is divided into m classes of a given class-width, for instance 0.05 R_F units. It should be noted that in the original publication[1] a slightly different dividing procedure was used, though this difference is unimportant. For each of the m classes the probability that an unknown compound will appear to have an R_F value within the limit of this class equals r_k/n for a group containing r_k members of the n comprising the total class. The information content, expressed in bit, is thus

$$I = -\sum_{k=1}^{m} \frac{r_k}{n} \cdot \log_2 \left(\frac{r_k}{n}\right) \tag{2}$$

It is well known that, in the ideal separation, TLC R_F values show a rectangular distribution over the plate. It can be shown that both $D.P.$ and I reach their maximum value for such a distribution[4,6]. In general, one might expect that every test allowing a measure of the "rectangularity" of the distribution could be used to evaluate a separation. One such test, the χ^2 test for goodness of fit, allows the calculation of the correspondence between an ideal, completely rectangular, distribution of R_F values and the experimentally observed one, and can thus provide a numerical value to describe each of the separating systems. The χ^2 statistic for goodness of fit is given by eqn. 3.

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{3}$$

For the selection of a preferred combination of two or more systems, two approaches are possible. The first is to consider every possible combination of two or more chromatographic systems and to calculate the information content or discriminating power for each combination. Such a procedure was carried out by Moffat and Smalldon[2], who calculated the discriminating power for k systems in series by means of a computer search program (to compute the number of similar pairs). It seems improbable, however, that the practising chromatographer will go to the trouble of writing a computer program for this purpose. A simpler method is therefore necessary. This second approach consists of the classification of the chromatographic systems into dissimilar groups and to select from each group (with similar separation characteristics) the best system according to, for instance, the information content.

TABLE I

DISTANCES BETWEEN THE NODES IN FIG. 1

City	*A*	*B*	*C*	*D*	*E*	*F*	*G*
A	0						
B	13	0					
C	27	25	0				
D	15	26	40	0			
E	8	21	29	11	0		
F	23	26	9	35	24	0	
G	38	37	12	50	39	15	0

We have recently introduced a simple mathematical technique based on this idea, namely numerical taxonomy[7], in order to to carry out this classification. A still simpler method makes use of a graph-theoretical algorithm, namely Kruskal's[8] algorithm for the calculation of a minimum spanning tree in a network. This algorithm can be introduced by the following example, in which seven cities are to be connected to each other by highways. One wants to know in which way to connect the cities so that the highways are of minimum length. All cities must be linked directly or indirectly to each other. There must evidently be no cycle. This is called a tree, and the tree for which the sum of the values of the edges is a minimum is called the minimum spanning tree. Kruskal's[8] algorithm for finding the minimum spanning tree can be stated as follows: "add to the tree the edge with the smallest value that does not build up a cycle with the edges already chosen". Applying this algorithm to the distances between the cities (A, B, C, D, E, F and G) as given in Table I, one starts by selecting edge AE (value 8), followed by CF (9), DE (11), CG (12) and then AB (13). In the next step, neither edge AD nor FG (value 15) are chosen, because they form a cycle with the already chosen edges AE and ED, or FC and CG. Edge EB (21) is not chosen because it also forms a cycle, and AF (23) is eventually chosen.

The optimum distribution network is therefore as given in Fig. 1. By careful inspection of this figure, which is drawn to scale, two clusters can be distinguished, namely (A, B, D and E) and (C, F and G). These clusters can be obtained formally by breaking up the longest edge (AF) in the tree. This distribution problem can be used in TLC for the classification of systems according to similarities in their chromatographic behaviour. The resemblance of the chromatographic systems can be pictured as a distance. The correlation coefficient calculated for two chromatographic systems using all the R_F values of each substance in both systems is a very appropriate

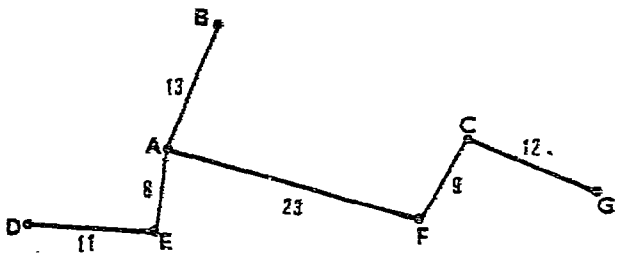

Fig. 1. Minimum spanning tree for seven cities.

expression of this resemblance. Therefore the distances between two systems for the identification of basic drugs can be defined as $(1 - \varrho)$ where ϱ is the correlation coefficient, and the chromatographic systems can be considered as constituting nodes in a graph. The values of the edges are given by the distance between the nodes: the smaller the distance, the larger the similarity between the nodes.

APPLICATIONS AND CONCLUSIONS

A data set of R_F values for the separation of 100 basic drugs in eight chromatographic systems published by Moffat and Smalldon[2] is now used. Moffat calculated the *D.P.* and the correlation between the systems in order to select the most effective series of chromatographic systems. Moffat's observations are compared here with a classification by numerical taxonomy of the same eight systems and selection of the best system by information content. As similarity coefficient, either the taxonomic distance defined as

$$\Delta_{kl} = \left[\sum_{i=1}^{n} (x_{ik} - x_{il})^2/n\right]^{\frac{1}{2}} \tag{4}$$

x_{ik} and x_{il} being the R_F values of compound i in systems k and l, or ϱ are used. For Kruskal's calculation of a minimum spanning tree, a distance defined as $(1 - \varrho) \cdot 1000$ is used as measure for the resemblance of two chromatographic systems.

In Table II the results obtained for the data set of Moffat are given. The discriminating power for each system, computed at an error factor of 0.10, was taken from his publication. The information content and the χ^2 values were calculated for R_F classes of 0.10 and 0.05 R_F units. Taking into account that the best system should display the highest *D.P.* and *I* and the lowest χ^2 value, it is clear that either system 6 or 7 should be considered as the optimum. According to the *D.P.* value, system 7 is best, and according to χ^2, system 6. When the separation criterion is reduced, *i.e.*, when 0.05 R_F units are considered to allow a distinction between two substances, then system 7 is best. When the systems are ranked according to decreasing efficiency, in general, the agreement between the three criteria is good. It is nearly complete between χ^2 and *I*. It should be noted that the *D.P.* is computed by considering the

TABLE II

COMPARISON OF DISCRIMINATING POWER, *D.P.*, INFORMATION CONTENT, *I*, AND χ^2 VALUES FOR BASIC DRUGS

The data are taken from ref. 2.

System	*D.P.*	*I(0.10)*	$\chi^2(0.10)$	*I(0.05)*	$\chi^2(0.05)$
1	0.730	2.76	71.6	3.73	78.0
2	0.688	2.83	55.2	3.71	81.2
3	0.749	2.99	35.8	3.91	50.8
4	0.657	2.68	76.8	3.60	82.8
5	0.672	2.56	80.2	3.57	90.4
6	0.742	3.08	32.6	3.95	46.8
7	0.753	3.08	35.4	4.06	43.2
8	0.549	2.10	276.4	2.97	334.8

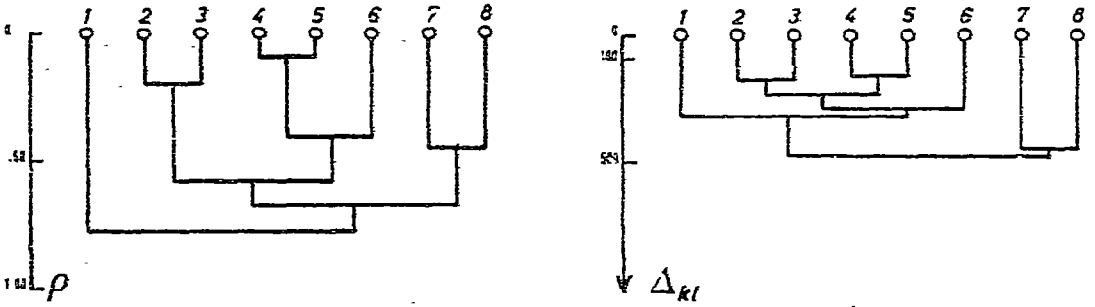

Fig. 2. Classification obtained by weighted pair average linkage numerical taxonomy with correlation coefficient, ϱ, as the similarity parameter.

Fig. 3. Classification obtained by weighted pair average linkage numerical taxonomy with taxonomic distance, Δ_{kl}, as the similarity parameter.

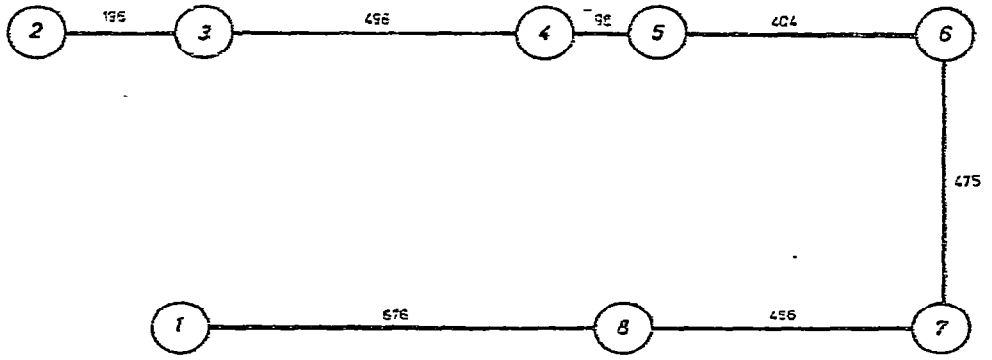

Fig. 4. Minimum spanning tree between chromatographic systems for the identification of basic drugs. The similarity parameter used was $(1 - \varrho) \cdot 1000$.

actual distances in R_F units, while for I and χ^2 the substances are grouped into R_F classes. Small differences between *D.P.* on the one hand and I and χ^2 on the other are therefore to be expected. Each of the proposed evaluation criteria (*D.P.*, χ^2 and I) takes into account only the distance between spots and not the spot size. It is evident that this can lead to erroneous conclusions when spot sizes are very different from one system to another. However, the purpose of using these criteria is to evaluate published data sets or at least to compare newly developed separations with alternative possibilities. Since spot sizes or shapes are hardly ever listed in R_F data sets, a systematic consideration of these factors is impossible. It should also be noted that the information content may be calculated in such a way that the size of individual spots is taken into account. However, this leads to complex mathematical formulae, thereby defeating our object of keeping the method simple and within reach of the possibilities of a small pocket calculator.

In Figs. 2 and 3 are given the classifications obtained by weighted pair average linkage numerical taxonomy with correlation coefficient and taxonomic distance respectively, as the similarity parameter. Fig. 4 shows the distribution tree obtained

with $(1 - \varrho) \cdot 1000$ as the similarity parameter. The resulting separations into classes are given in Table III. By selecting the best system in each class, according to I or $D.P.$ (Table II), one obtains the (supposedly) best combinations (Table III).

The agreement between the two classification methods based on correlation and the discrimination power is excellent, at least if one takes into account that the best pair according to Moffat (3 and 7, $D.P. = 0.929$) is only very slightly better than the pair chosen by our much simpler technique (1 and 7, $D.P. = 0.925$), and our best combinations, 1, 6 and 7 and 1, 3 and 7, are second and third best according to the $D.P.$ Also the two individually best systems, 6 and 7 or 3 and 7, are nearly equivalent to 1 and 7, although 1 is clearly of less efficiency than either 3 or 6 (Table II). This is explained by the fact that the correlation between systems 1 and 7 is much smaller than between 3 or 6 and 7 and shows, again, as was remarked by several of the authors cited in the present article, that two factors determine the best combination: *nl*, the individual efficiency of the systems, and their dissimilarity.

To conclude, the proposed criteria and strategies lead in the example to optimum or near optimum TLC systems. The numerical techniques presented here allow the rational classification and selection of separating systems in chromatography. We are aware that many practising TLC or PC specialists have more confidence in the experience of the analyst for finding optimal separation systems than in such techniques. Therefore it should be stated explicitly here that the proposed methods are not at all intended to replace this experience but only to aid the analyst in finding an easier and more rapid solution to his problem, and to provide him with objective methods for evaluating several alternative possibilities. This is especially clear for the numerical-taxonomy and Kruskal procedures. These classification procedures divide the systems into two or more clusters. From each of these clusters the analyst chooses the best system, according to a selection criterion in the way proposed here or by taking into account such important practical parameters as availability, reproducibility or cost.

TABLE III

SEPARATION INTO CLASSES AND THE BEST COMBINATIONS OF CHROMATOGRAPHIC SYSTEMS FOR THE IDENTIFICATION OF BASIC DRUGS

	Class	*Best combination*
Numerical taxonomy (correlations)	1, 2–8	1, 7
	1, 2–6, 7–8	1, 6, 7 or 1, 3, 7*
	1, 2–3, 4–6, 7–8	1, 3, 6, 7
Numerical taxonomy (distances)	1–6, 7–8	1, 7
	1–6, 7, 8	1, 7, 8
	1, 2–6, 7, 8	1, 6, 7, 8
Distribution tree	1, 2–8	1, 7
	1, 2–3, 4–8	1, 3, 7
	1, 2–3, 4–6, 7–8	1, 3, 6, 7
Discriminating power (from Moffat)		3, 7
		3, 6, 7
		1, 3, 6, 7

* Systems 3 and 6 have similar efficiencies: according to I, 6 is somewhat better; according to $D.P.$, 3 is best.

REFERENCES

1 D. L. Massart, *J. Chromatogr.*, 79 (1973) 157.
2 A. C. Moffat and K. W. Smalldon, *J. Chromatogr.*, 90 (1974) 9.
3 K. W. Smalldon and A. C. Moffat, *J. Forensic Sci. Soc.*, 13 (1973) 291.
4 A. C. Moffat, K. W. Smalldon and C. Brown, *J. Chromatogr.*, 90 (1974) 1.
5 A. C. Moffat, A. H. Stead and K. W. Smalldon, *J. Chromatogr.*, 90 (1974) 19.
6 H. N. Nahikian, *A Modern Algebra for Biologists*, 2nd Ed., Univ. of Chicago Press, Chicago and London, 1969.
7 D. L. Massart and H. De Clercq, *Anal. Chem.*, 46 (1974) 1988.
8 J. B. Kruskal, *Proc. Amer. Math Soc.*, 7 (1956) 48.