Journal of Chromatography, 328 (1985) 11–34
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 17 693

COMPARISON OF MATHEMATICAL METHODS FOR THE CALCULATION OF KOVÁTS INDICES

R. J. SMITH*, J. K. HAKEN, M. S. WAINWRIGHT and B. G. MADDEN

Departments of Industrial Chemistry and Polymer Science, University of New South Wales, P.O. Box 1, Kensington, N.S.W. 2033 (Australia)

(Received February 28th, 1985)

SUMMARY

Several statistical and iterative methods for the calculation of the retention indices of Kováts have been compared using retention data obtained on a wide range of columns. Using homologues ranging from C_1 to C_{22} and the linear relationship between the logarithm of adjusted retention times and carbon numbers, it has been shown that the iterative method of Guardino *et al.* [*J. Chromatogr.*, 118 (1976) 13] is of adequate accuracy in the majority of analyses. It has further been shown that while higher degree polynomial relationships between carbon numbers and the logarithm of adjusted retention times give improved accuracy, their interpretation is difficult.

INTRODUCTION

Since the introduction of the retention index system of Kováts[1–3] in the late 1950s, a large number of papers have discussed a variety of methods of calculating Kováts indices. These methods are invariably based on the empirical linear relationship between the logarithm of the adjusted retention time and the carbon number of a substance (eqn. 1). The majority of such methods involve either the simultaneous or sequential estimation of the system dead time and the coefficients of the linear equation. Recently methods have been introduced which do not rely on this linear relationship.

$$\ln(t_{Rz} - t_M) = bz + c \qquad I = 100z \tag{1}$$

where t_{Rz} is the retention time of an n-alkane of carbon number z; t_M is the dead time; I is the Kováts retention index; and b and c are the slope and intercept, respectively.

It is not the purpose of this paper to carry out a review of the various methods for estimating dead time and calculating Kováts indices as an extensive review of this area has been carried out by us[4]. Instead, this paper will compare a range of statistical and iterative methods which have appeared in the literature[5–12].

0021-9673/85/$03.30

METHODS

The following methods will be compared.

(i) The method of Grobler and Balizs[5] which relies on the use of two linear regressions to first calculate the slope *b*, and then the mathematical dead time t_M and intercept *c*.

The method has been modified by the technique of Van Tulder *et al.*[6] to allow the use of all possible combinations of three or more homologues which are separated by equal distances. The combination giving the smallest sum of squares of the differences between known and calculated Kováts indices is used in the final calculation of equation parameters. (Note that all homologues are used in the sum-of-squares calculation, irrespective of the actual number used to determine the equation parameters.)

(ii) The method of Guardino *et al.*[7], in which an iteration is carried out on t_M, with *b* and *c* being calculated using a least squares linear regression.

(iii) The iterative method of Garcia Dominguez *et al.*[8] in which the actual gross retention times are used in a least squares regression to calculate a set of parameters which are then used to estimate a new retention time for the homologue with the smallest retention time. The difference in retention times for the first homologue is then used to correct the retention times of all other homologues and the procedure is repeated until the corrected and actual retention times of the first homologue are identical.

(iv) The method of Ambrus[9] in which an initial linear regression gives both the slope *b*, and the dead time t_M. A second regression then gives the intercept *c*.

This method has also been modified by the technique of Van Tulder *et al.*[6] in order to check all combinations of three or more homologues (see method i).

(v) An iterative technique mentioned by Tóth and Zala[10] in a recent paper in which t_M is determined by expressing the coefficient of determination (R^2) as a function of t_M and solving a non-linear equation

$$\mathrm{d}R^2/\mathrm{d}t_M = 0 \qquad (2)$$

The method of Guardino *et al.*[7] has been modified to solve this equation using a five-point numerical differentiation method outlined by Dorn and McCracken[13].

An alternative solution which is not only faster, but also does not involve the problems associated with numerical differentiation, is the minimisation of $(1 - R^2)$ using an appropiate iterative technique.

(vi) Flexible simplex[11] which is a non-linear least-squares estimation of t_M, *b* and *c* simultaneously using an objective function of the sum of squares of the differences between known and calculated Kováts indices.

(vii) The method of Heeg *et al.*[12] in which a cubic polynomial is fitted to the data. In order to fit a cubic polynomial to the data, Gaussian elimination with partial pivotal condensation[13] has been used. As t_M must be estimated separately, the result from method ii is used.

(viii) The method of Guardino *et al.*[7] modified to fit a cubic instead of linear equation to the data while carrying out an iteration on t_M. Again Gaussian elimination with partial pivotal condensation has been used.

This method is included to evaluate the effect of calculating the mathematical dead time in conjunction with the curve fitting method, instead of using the result from a linear method as suggested by Heeg *et al.*[12]

Initially the method of flexible simplex was to be used. However continuing problems with local minimums *versus* global minimums and instabilities in the dead time estimation caused the abandonment of the method.

(ix) Method viii, extended to permit a degree 5 polynomial to be fitted to the data in order to determine the degree of improvement in fit that occurs when higher degree polynomials are used.

No classical methods were included because of the large inaccuracies inherent in all such methods. We fully discussed this problem in refs. 4 and 14. In addition several other authors have critised the accuracy of such methods[15–19].

DATA

Previously reported comparisons have, in many cases, suffered from the small range of homologues over which they have been made. In addition the extension of such results to other analyses has been difficult due to the small range of columns and conditions that have been reported in individual articles. In an attempt to rectify this problem, thus allowing more general conclusions to be drawn, data have been selected from several sources covering a wide range of columns and conditions.

The raw data are presented in Table I while columns and temperatures are listed in Table II.

ANALYSES

All analyses were generated on a Dual Disk 48K Tandy Model 3 microcomputer using double precision arithmetic. This computer, based on a Z-80 microprocessor running at 2.02752 MHz, uses a version of microsoft extended disk BASIC (interpretative).

Comparison with literature results[12] shows some differences in precision. We believe this is due to the problem of round-off and subtractive error involved in some of the methods plus the fact that with the polynomial methods, apparently significantly different coefficients can lead to almost identical Kováts indices. Therefore we have used up to 16 digits of precision throughout all the methods (double precision routines for ln and exp. were written in BASIC as the Model 3 only provides single precision routines for these functions) with only the results being quoted to a lower precision for convenience. This also explains why the Kováts indices for the higher polynomials cannot be obtained with the quoted accuracy of the coefficients.

Before beginning the comparison, it is necessary to outline the criteria that will be used for the comparison. While the calculation of mathematical dead time is an integral part of each of the methods, it will be approached on the basis of its suitability in the calculation of Kováts indices and not as an estimate of the system hold-up volume which is discussed in detail elsewhere[4].

We will also be concerned with the accuracy of fit of the regression over the full range of reported homologue retention times. Ettre[17], amongst others, has suggested using only the linear part of the data to calculate the dead time and system

TABLE I

RAW RETENTION DATA

Homologue	*Source*										
	21	*21*	*21*	*21*	*12*	*7*	*17*	*22*	*23*	*23*	*6*
	Unit										
	sec	*sec*	*sec*	*sec*	*sec*	*sec*	*sec*	*mm*	*sec*	*sec*	*min*
C_1	200.6	158.5	136.0	58.5			879.6				2.326
C_2	212.6	182.5	164.5	72.3			885.1				2.361
C_3	233.6	229.0	222.5	94.3			901.1				2.413
C_4	281.5	335.3	373.0	139.5			958.9				2.494
C_5	379.9	568.0	730.0	225.8			1146.1			124.5	2.606
C_6	582.9	1079.0	1576.5	389.8			1752.2		69.0	135.4	2.767
C_7	1003.0	2206.0	3556.5	703.3		254.0	3715.0		90.5	155.3	2.987
C_8						271.0			135.2	192.6	3.266
C_9						301.0		101	226.5	262.2	3.648
C_{10}						352.5		130	416.0	393.3	4.168
C_{11}						440.0		170	808.7	637.25	4.843
C_{12}						580.0		226	1625.9		5.756
C_{13}					168.1	839.0		303			6.979
C_{14}					187.2	1267.0		406			8.605
C_{15}					215.9	1990.0		548			10.799
C_{16}					259.0			744			
C_{17}					323.7			1013			
C_{18}					420.5			1384			
C_{19}					565.2			1894			
C_{20}					780.8			2598			
C_{21}					1101.5						
C_{22}					1577.1						

TABLE II

COLUMN AND OPERATING TEMPERATURES USED TO GENERATE DATA IN TABLE I

Column No.	*Details*
1	22 ft. × 0.25 in. O.D. aluminium packed with 10% SE-30 on Chromosorb W AW DMCS, operated at 30°C
2	6 ft. × 0.125 in. O.D. PTFE packed with Porasil C (75–125 μm), operated at 60°C
3	10 ft. × 0.125 in. O.D. PTFE packed with 10% squalane on Porasil C, operated at 50°C
4	6 ft. × 0.125 in. O.D. PTFE packed with Porapak S (80–100 μm), operated at 180°C
5	30 m glass capillary coated with SE-30, operated at 160°C
6	30 m × 0.3 mm I.D. stainless-steel capillary, coated with Apiezon L, operated at 120°C
7	100 m glass wall-coated open tubular coated with SF-96, operated at 25°C
8	3.0 m × 0.8 mm I.D. glass micro-packed with 1.5% OV-17 and 2.0% OV-210 on Kieselguhr (100–120 mesh), operated at 85°C
9	12 ft. × 0.25 in. O.D. aluminium packed with 10% OV-25 on Chromosorb W, operated at 25°C
10	12 ft. × 0.25 in. O.D. aluminium packed with 10% SILAR-7CP on Chromosorb W, operated at 60°C
11	HPLC data: 15 × 0.46 cm I.D. packed with Nucleosil C_{18}, (5 μm), operated at 25°C (thermostated). Methanol–water (9:1) at 1.0 ml/min. Data are for 2-ketoalkanes

parameters. While this approach may have some merit when one is estimating the system hold-up volume, it is inappropiate when calculating Kováts indices, as large errors may be introduced into those regions excluded from the calculation. While Ettre's arguement that with this approach "both the adjusted retention times and the gas hold-up time so obtained would be artificial, not corresponding to the true values" is essentially correct, it is in fact not a problem. The initial (and still most important) reason for introducing the relationship in eqn. 1 is to aid in the identification of substances by eliminating such interlaboratory variables as the effect of the equipment used. The introduction of the column hold-up volume, while based to some extent in theory, is still only a "means to an end". The real requirement is accurate and reproducible Kováts indices.

Finally it is helpful, when discussing the accuracy of these methods, to have a target in mind. Although the acceptable reproducibility of Kováts indices varies within the literature, we believe that an accuracy between laboratories of better than one unit is not, in general, achievable and will therefore accept the estimate of Schomburg and Dielmann[20] of one unit. Since the homologues have predefined index values, it will be assumed that the maximum desireable deviation for these index values in a particular method is one unit.

RESULTS AND DISCUSSION

Tables III–XXII compare the nine methods over a wide range of homologues (C_1–C_{22}), a variety of columns including wall-coated open-tubular, packed and capillary, as well as several stationary phases of low, moderate and high polarity including gas–solid chromatography columns. Tables XXIII and XXIV include a set of high-performance liquid chromatography data for comparison. The tables are divided between the linear methods (odd numbered tables) and the polynomial methods (even numbered tables). A study of the data presented in these tables leads to several conclusions.

Linear methods (i to vi)

(1) Using the sum of squares of the deviations between calculated and known Kováts indices divided by the number of homologues (S of S/N) as a general indication of the accuracy of fit, methods ii, v and vi give almost identical results and are also the most accurate of the linear methods. This is not surprising since the three are iterative methods which depend on the minimisation of the sum of squares to determine their parameters.

(2) In general all six linear methods give similar accuracy. However, in individual cases, method iv in particular (Tables VII and XV) and to a lesser extent methods i (Table IX) and iii (Table XXIII) give large variations in accuracy. It should also be noted that methods i and iv give even larger variations when the full range of homologues are used instead of determining the best combination using the technique of Van Tulder *et al.*[6]

(3) Method iii, by definition, sets the Kováts index of the homologue with the least retention time to its known value. This is not necessarily a desireable practice when dealing with the lower homologues.

TABLE III

COMPARISON OF METHODS i–vi USING DATA FROM REF. 21 (FIRST COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	190.38	190.31	189.93	190.40	190.31	190.31
Slope	0.72937	0.72576	0.72057	0.72909	0.72575	0.72588
Intercept	1.59465	1.61614	1.64678	1.61034	1.61623	1.61585
Alkanes used	C_1, C_4, C_7	C_1–C_7	C_1–C_7	C_1–C_6	C_1–C_7	C_1–C_7
S of S/N	6.812	5.842	6.530	7.986	5.842	5.841
I_1	100.00	98.56	100.00	97.73	98.54	98.59
I_2	206.51	205.04	204.59	204.36	205.03	205.06
I_3	297.73	296.49	295.58	295.66	296.49	296.50
I_4	400.00	399.14	398.34	397.99	399.14	399.13
I_5	500.41	499.99	499.61	498.44	499.99	499.97
I_6	600.23	600.28	600.49	598.31	600.28	600.25
I_7	700.00	700.53	701.39	698.12	700.54	700.49
S_1 (196 sec)	17.98	16.96	21.71	15.54	16.93	16.99
S_2 (320 sec)	448.32	447.67	447.04	446.33	447.67	447.66
S_3 (2500 sec)	843.22	844.45	846.31	841.39	844.46	844.39

TABLE IV

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 21 (FIRST COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	190.31*	182.87	96.70
Constant	-2.07572	-7.93405	$-8.33624 \cdot 10^3$
Coefficient X	1.25431	4.34647	$7.25056 \cdot 10^3$
Coefficient X^2	$3.11180 \cdot 10^{-2}$	$-5.18281 \cdot 10^{-1}$	$-2.51862 \cdot 10^3$
Coefficient X^3	$-2.43056 \cdot 10^{-3}$	$3.01507 \cdot 10^{-2}$	$4.36684 \cdot 10^2$
Coefficient X^4	*N/A*	*N/A*	-37.7702
Coefficient X^5	*N/A*	*N/A*	1.30342
S of S/N	5.458	3.015	0.000
I_1	98.66	99.51	100.01
I_2	204.50	202.29	200.01
I_3	296.22	296.70	300.00
I_4	399.53	401.47	400.00
I_5	500.83	500.93	500.00
I_6	600.82	598.75	600.00
I_7	699.45	700.37	700.00
S_1 (196 sec)	18.65	33.60	47.51
S_2 (320 sec)	448.33	449.70	445.41
S_3 (2500 sec)	837.62	865.32	3859.74

* Calculated by the method of Guardino *et al.*[7].

TABLE V

COMPARISON OF METHODS i–vi USING DATA FROM REF. 21 (SECOND COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	140.53	139.64	139.02	140.17	139.64	139.64
Slope	0.78761	0.77969	0.77511	0.78580	0.77968	0.77975
Intercept	2.11984	2.16735	2.19405	2.12955	2.16744	2.16708
Alkanes used	C_3, C_5, C_7	C_1–C_7	C_1–C_7	C_3–C_6	C_1–C_7	C_1–C_7
S of S/N	4.859	3.435	3.950	4.693	3.435	3.435
I_1	97.64	98.73	100.00	99.16	98.72	98.75
I_2	205.33	204.01	203.60	205.65	204.00	204.02
I_3	300.00	298.24	297.44	299.98	298.23	298.25
I_4	400.20	398.75	398.07	400.12	398.75	398.76
I_5	500.00	499.25	498.94	500.03	499.25	499.25
I_6	599.84	599.96	600.15	600.04	599.96	599.95
I_7	700.00	701.07	701.81	70.40	701.08	701.05
S_1 (149 sec)	2.16	8.89	13.69	6.23	8.88	8.90
S_2 (450 sec)	458.99	457.92	457.44	458.96	457.92	457.92
S_3 (4500 sec)	794.84	796.85	798.13	795.45	796.85	796.82

TABLE VI

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 21 (SECOND COLUMN TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	139.64*	128.62	63.13
Constant	-2.43909	-7.70614	$-7.23407 \cdot 10^2$
Coefficient X	1.05415	3.50416	$5.66226 \cdot 10^2$
Coefficient X^2	$4.78689 \cdot 10^{-2}$	$-3.33975 \cdot 10^{-1}$	$-1.77392 \cdot 10^2$
Coefficient X^3	$-3.15232 \cdot 10^{-3}$	$1.66662 \cdot 10^{-2}$	27.8365
Coefficient X^4	*N/A*	*N/A*	-2.18032
Coefficient X^5	*N/A*	*N/A*	$6.81362 \cdot 10^{-2}$
S of S/N	2.525	0.946	0.000
I_1	99.01	99.75	100.00
I_2	203.11	201.20	200.00
I_3	297.72	298.17	300.00
I_4	399.26	400.84	400.00
I_5	500.54	500.61	500.00
I_6	600.87	599.20	600.00
I_7	699.50	700.24	700.00
S_1 (149 sec)	12.26	27.93	40.22
S_2 (450 sec)	458.98	459.89	458.20
S_3 (4500 sec)	790.16	801.73	925.05

* Calculated by the method of Guardino *et al.*[7].

TABLE VII

COMPARISON OF METHODS i–vi USING DATA FROM REF. 21 (THIRD COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	116.94	116.82	116.25	118.21	116.82	116.82
Slope	0.86589	0.86318	0.85908	0.87703	0.86316	0.86326
Intercept	2.08185	2.10050	2.12409	2.02908	2.10058	2.10020
Alkanes used	C_1, C_4, C_7	C_1–C_7	C_1–C_7	C_1–C_5	C_1–C_7	C_1–C_7
S of S/N	5.487	4.777	5.174	8.016	4.777	4.776
I_1	100.00	98.88	100.00	96.85	98.87	98.90
I_2	205.59	204.37	203.98	205.90	204.36	204.38
I_3	297.67	296.57	295.87	298.52	296.57	296.58
I_4	400.00	399.15	398.57	400.37	399.15	399.15
I_5	500.83	500.26	500.01	500.25	500.26	500.25
I_6	601.00	600.74	600.91	599.29	600.74	600.72
I_7	700.00	700.05	700.66	697.08	700.05	700.01
S_1 (130 sec)	56.35	55.42	57.85	49.94	55.41	55.45
S_2 (500 sec)	446.51	445.80	445.35	446.79	445.48	445.79
S_3 (7000 sec)	780.12	780.41	781.40	776.20	780.42	780.37

TABLE VIII

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 21 (THIRD COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	116.82*	104.80	22.81
Constant	-2.38655	-7.28199	$-8.05144 \cdot 10^{2}$
Coefficient X	1.12386	3.33037	$6.05156 \cdot 10^{2}$
Coefficient X^2	$7.74692 \cdot 10^{-3}$	-32.3497	$-1.81632 \cdot 10^{2}$
Coefficient X^3	$-5.27100 \cdot 10^{-4}$	$1.59594 \cdot 10^{-2}$	27.2467
Coefficient X^4	*N/A*	*N/A*	-2.03704
Coefficient X^5	*N/A*	*N/A*	$6.06834 \cdot 10^{-2}$
S of S/N	4.671	2.164	0.000
I_1	98.72	99.63	100.00
I_2	204.19	201.80	200.00
I_3	296.60	297.20	300.00
I_4	399.46	401.44	400.00
I_5	500.69	500.70	500.00
I_6	600.92	598.88	600.00
I_7	699.42	700.35	700.00
S_1 (130 sec)	55.40	63.20	69.87
S_2 (500 sec)	446.19	447.55	444.76
S_3 (7000 sec)	778.60	789.54	923.13

* Calculated by the method of Guardino *et al.*[7].

TABLE IX

COMPARISON OF METHODS i–vi USING DATA FROM REF. 21 (FOURTH COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	44.911	44.60	43.93	45.19	44.60	44.60
Slope	0.64675	0.63977	0.63297	0.64657	0.63975	0.63985
Intercept	1.96254	2.00470	2.04583	1.96714	2.00483	2.00444
Alkanes used	C_1, C_4, C_7	C_1–C_7	C_1–C_7	C_1–C_6	C_1–C_7	C_1–C_7
S of S/N	10.041	7.167	8.417	8.384	7.167	7.165
I_1	100.00	98.08	100.00	96.15	98.06	98.12
I_2	208.37	205.84	205.29	206.15	205.82	205.86
I_3	299.53	297.20	295.99	298.03	297.19	297.21
I_4	400.00	398.29	397.18	398.95	398.29	398.29
I_5	500.25	499.39	498.83	499.44	499.38	499.37
I_6	600.03	600.13	600.38	599.36	600.13	600.10
I_7	700.00	701.12	702.32	699.42	701.13	701.08
S_1 (55 sec)	53.95	52.75	56.59	48.98	52.73	52.80
S_2 (450 sec)	624.90	625.25	625.74	624.26	625.25	625.22
S_3 (1000 sec)	757.52	759.25	761.02	756.97	759.25	759.20

TABLE X

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 21 (FOURTH COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	44.60*	31.31	−171.11**
Constant	-2.62115	-14.7303	$-1.78291 \cdot 10^5$
Coefficient X	1.17458	7.48229	$1.46243 \cdot 10^5$
Coefficient X^2	$9.28896 \cdot 10^{-2}$	-1.01689	$-4.79492 \cdot 10^4$
Coefficient X^3	$-7.03144 \cdot 10^{-3}$	$5.84255 \cdot 10^{-2}$	$7.85470 \cdot 10^3$
Coefficient X^4	*N/A*	*N/A*	$-6.42814 \cdot 10^2$
Coefficient X^5	*N/A*	*N/A*	21.0238
S of S/N	5.851	2.758	0.090
I_1	98.54	99.44	100.22
I_2	204.72	202.36	199.89
I_3	296.49	296.92	299.59
I_4	398.84	401.03	400.56
I_5	500.91	501.14	499.73
I_6	601.23	598.80	600.07
I_7	699.27	700.33	700.00
S_1 (55 sec)	54.86	61.70	68.25
S_2 (450 sec)	625.90	623.45	608.71
S_3 (1000 sec)	754.10	763.37	1627.21

* Calculated by the method of Guardino *et al.*[7]

** Note that the mathematical dead time is negative.

TABLE XI

COMPARISON OF METHODS i–vi USING DATA FROM REF. 12 (FIFTH COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	129.18	129.07	129.38	129.17	129.07	129.07
Slope	0.40235	0.40217	0.40292	0.40350	0.40217	0.40216
Intercept	−1.56880	−1.56604	−1.58155	−1.58899	−1.56599	−1.56596
Alkanes used	C_{14}–C_{21}	C_{13}–C_{22}	C_{13}–C_{22}	C_{14}–C_{19}	C_{13}–C_{22}	C_{13}–C_{22}
S of S/N	0.310	0.289	0.356	0.983	0.289	0.289
I_{13}	1300.15	1300.55	1300.00	1301.33	1300.54	1300.54
I_{14}	1399.39	1399.60	1399.51	1400.25	1399.59	1399.59
I_{15}	1499.28	1499.37	1499.54	1499.83	1499.37	1499.37
I_{16}	1599.56	1599.59	1599.86	1599.81	1599.58	1599.58
I_{17}	1700.07	1700.07	1700.35	1700.02	1700.07	1700.07
I_{18}	1800.45	1800.45	1800.67	1800.11	1800.45	1800.45
I_{19}	1900.68	1900.69	1900.81	1900.04	1900.69	1900.69
I_{20}	2000.54	2000.57	2000.56	1999.61	2000.57	2000.57
I_{21}	2100.01	2100.07	2099.92	2098.80	2100.07	2100.07
I_{22}	2198.98	2199.08	2198.76	2197.48	2199.08	2199.08
S_1 (135 sec)	827.76	832.05	821.12	830.93	832.00	832.02
S_2 (375 sec)	1758.25	1758.24	1758.49	1753.02	1758.24	1758.24
S_3 (2500 sec)	2321.54	2321.68	2321.16	2319.69	2321.68	2321.69

TABLE XII

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 12 (FIFTH COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	129.07*	130.22	126.20
Constant	3.25719	3.97749	−2.53604
Coefficient X	2.86327	2.54043	7.07279
Coefficient X^2	$-7.21528 \cdot 10^{-2}$	$-2.31109 \cdot 10^{-2}$	−1.32495
Coefficient X^3	$4.47872 \cdot 10^{-3}$	$1.97045 \cdot 10^{-3}$	$1.93128 \cdot 10^{-1}$
Coefficient X^4	*N/A*	*N/A*	$-1.42474 \cdot 10^{-2}$
Coefficient X^5	*N/A*	*N/A*	$4.28936 \cdot 10^{-4}$
S of S/N	0.005	0.000	0.000
I_{13}	1300.07	1300.01	1300.01
I_{14}	1399.92	1400.03	1400.00
I_{15}	1499.93	1500.00	1500.01
I_{16}	1599.99	1599.97	1599.99
I_{17}	1700.10	1700.02	1700.02
I_{18}	1800.06	1800.00	1799.98
I_{19}	1900.02	1900.00	1900.01
I_{20}	1999.92	1999.99	2000.00
I_{21}	2099.92	2099.99	2100.00
I_{22}	2209.07	2200.01	2200.00
S_1 (135 sec)	815.05	790.34	826.74
S_2 (375 sec)	1758.02	1757.95	1757.93
S_3 (2500 sec)	2325.24	2324.71	2324.74

* Calculated by the method of Guardino *et al.*[7].

TABLE XIII

COMPARISON OF METHODS i–vi USING DATA FROM REF. 7 (SIXTH COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	229.53	229.17	229.45	229.44	229.17	229.17
Slope	0.53532	0.53338	0.53463	0.53513	0.53338	0.53338
Intercept	−0.54996	−0.52523	−0.54207	−0.54453	−0.52518	−0.52519
Alkanes used	C_7–C_{14}	C_7–C_{15}	C_7–C_{15}	C_7–C_{14}	C_7–C_{15}	C_7–C_{15}
S of S/N	2.025	1.800	1.912	2.252	1.800	1.800
I_7	700.00	700.69	700.00	699.98	700.68	700.69
I_8	798.53	798.47	798.45	798.27	798.46	798.47
I_9	900.20	899.83	900.13	899.80	899.83	899.83
I_{10}	1001.57	1001.18	1001.55	1001.11	1001.17	1001.17
I_{11}	1101.96	1101.70	1102.03	1101.47	1101.70	1101.69
I_{12}	1197.21	1197.18	1197.38	1196.73	1197.18	1197.17
I_{13}	1300.57	1300.83	1300.86	1300.11	1300.83	1300.82
I_{14}	1399.94	1400.52	1400.35	1399.50	1400.52	1400.51
I_{15}	1498.72	1499.63	1499.26	1498.31	1499.63	1499.62
S_1 (235 sec)	420.24	429.07	421.64	422.50	429.03	429.12
S_2 (500 sec)	1148.81	1148.66	1148.93	1148.32	1148.65	1148.65
S_3 (3500 sec)	1614.42	1615.73	1615.10	1614.04	1615.73	1615.72

TABLE XIV

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 7 (SIXTH COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	229.17*	233.91	224.49
Constant	$9.45389 \cdot 10^{-1}$	3.29983	−28.1170
Coefficient X	1.90109	$8.29188 \cdot 10^{-1}$	28.8772
Coefficient X^2	$-5.53027 \cdot 10^{-3}$	$1.59435 \cdot 10^{-1}$	−10.0320
Coefficient X^3	$3.68399 \cdot 10^{-4}$	$-8.15049 \cdot 10^{-3}$	1.85080
Coefficient X^4	*N/A*	*N/A*	$-1.68891 \cdot 10^{-1}$
Coefficient X^5	*N/A*	*N/A*	$6.08758 \cdot 10^{-3}$
S of S/N	1.791	1.614	0.908
I_7	700.69	700.24	699.99
I_8	798.54	799.30	800.16
I_9	899.89	900.14	899.73
I_{10}	1001.18	1000.82	999.88
I_{11}	1101.63	1101.15	1101.24
I_{12}	1197.05	1196.90	1198.04
I_{13}	1300.71	1301.05	1301.51
I_{14}	1400.49	1400.96	1399.37
I_{15}	1499.81	1499.47	1500.11
S_1 (235 sec)	428.19	337.01	365.83
S_2 (500 sec)	1148.55	1148.19	1148.92
S_3 (3500 sec)	1616.36	1612.96	1640.66

* Calculated by the method of Guardino *et al.*[7].

TABLE XV

COMPARISON OF METHODS i–vi USING DATA FROM REF. 17 (SEVENTH COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	877.25	877.30	877.27	877.15	877.30	877.30
Slope	1.18278	1.18669	1.18478	1.19339	1.18658	1.18667
Intercept	−0.32869	−0.35054	−0.33988	−0.31486	−0.34990	−0.35030
Alkanes used	C_1, C_4, C_7	C_1–C_7	C_1–C_7	C_2–C_5	C_1–C_7	C_1–C_7
S of S/N	3.027	2.817	2.882	2.755	2.817	2.816
I_1	99.99	99.58	100.00	101.55	99.54	99.55
I_2	201.99	202.59	202.36	200.13	202.56	202.57
I_3	295.96	296.63	296.32	292.52	296.60	296.61
I_4	400.01	400.47	400.25	395.39	400.45	400.45
I_5	500.76	500.93	500.85	495.18	500.92	500.92
I_6	600.52	600.38	600.45	594.03	600.38	600.37
I_7	700.01	699.53	699.76	692.62	699.55	699.52

TABLE XVI

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 17 (SEVENTH COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	877.30*	876.58	871.07
Constant	$2.77480 \cdot 10^{-1}$	$-1.83319 \cdot 10^{-1}$	−12.5764
Coefficient X	$8.61391 \cdot 10^{-1}$	1.12532	12.5930
Coefficient X^2	$-4.9414 \cdot 10^{-3}$	$5.34703 \cdot 10^{-2}$	−4.28517
Coefficient X^3	$3.67458 \cdot 10^{-4}$	$3.22313 \cdot 10^{-3}$	$7.73107 \cdot 10^{-1}$
Coefficient X^4	*N/A*	*N/A*	$-6.87321 \cdot 10^{-2}$
Coefficient X^5	*N/A*	*N/A*	$2.40552 \cdot 10^{-3}$
S of S/N	2.697	1.764	0.000
I_1	99.17	99.85	100.03
I_2	202.92	201.35	200.01
I_3	296.99	297.52	300.00
I_4	400.48	401.64	400.00
I_5	500.58	500.49	500.00
I_6	600.01	598.87	600.00
I_7	699.85	700.39	700.00

* Calculated by the method of Guardino *et al.*[7].

(4) An examination of the homologues actually chosen by method i to give the lowest S of S/N shows that, in general, a subjet of three or four homologues separated by as large a distance as possible gives an improvement in accuracy. However it is also obvious (Table XIII for example) that the technique does not guarantee the best result for a particular set of data.

TABLE XVII

COMPARISON OF METHODS i–vi USING DATA FROM REF. 22 (EIGHTH COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	22.16	22.39	22.75	23.05	22.39	22.40
Slope	0.31693	0.31713	0.31749	0.31800	0.31713	0.31714
Intercept	1.51233	1.50908	1.50257	1.49316	1.50917	1.50891
Alkanes used	C_{12}–C_{19}	C_9–C_{20}	C_9–C_{20}	C_{11},C_{14}, C_{17},C_{20}	C_9–C_{20}	C_9–C_{20}
S of S/N	0.495	0.482	0.514	0.697	0.482	0.482
I_9	900.69	900.40	900.00	900.27	900.40	900.38
I_{10}	999.52	999.42	999.30	999.74	999.41	999.41
I_{11}	1099.07	1099.08	1099.13	1099.65	1099.07	1099.08
I_{12}	1200.42	1200.50	1200.65	1201.18	1200.49	1200.50
I_{13}	1301.53	1301.64	1301.83	1302.33	1301.64	1301.64
I_{14}	1400.12	1400.23	1400.42	1400.85	1400.23	1400.23
I_{15}	1499.44	1499.54	1499.70	1500.03	1499.53	1499.54
I_{16}	1599.40	1599.47	1599.58	1599.80	1599.47	1599.47
I_{17}	1699.34	1699.38	1699.42	1699.51	1699.38	1699.38
I_{18}	1799.70	1799.69	1799.64	1799.60	1799.69	1799.68
I_{19}	1900.06	1900.00	1899.87	1899.68	1900.00	1900.00
I_{20}	2000.80	2000.68	2000.45	2000.12	2000.69	2000.67

TABLE XVIII

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 22 (EIGHTH COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	22.39*	37.71	29.31
Constant	-4.17598	3.00640	-33.2383
Coefficient X	2.85237	$-4.91795 \cdot 10^{-2}$	29.2161
Coefficient X^2	$5.07560 \cdot 10^{-2}$	$4.48018 \cdot 10^{-1}$	-9.12539
Coefficient X^3	$-2.79787 \cdot 10^{-3}$	$-2.11370 \cdot 10^{-2}$	1.55032
Coefficient X^4	*N/A*	*N/A*	$-1.28742 \cdot 10^{-1}$
Coefficient X^5	*N/A*	*N/A*	$4.19830 \cdot 10^{-3}$
S of S/N	0.402	0.247	0.128
I_9	900.74	900.17	899.94
I_{10}	999.33	999.98	1000.35
I_{11}	1098.79	1099.33	1099.43
I_{12}	1200.19	1200.32	1200.00
I_{13}	1301.44	1301.17	1300.90
I_{13}	1400.20	1399.75	1399.65
I_{14}	1499.69	1499.32	1499.58
I_{16}	1599.78	1599.67	1600.11
I_{17}	1699.74	1699.98	1700.21
I_{18}	1799.97	1800.41	1800.14
I_{19}	1900.00	1900.31	1899.73
I_{20}	2000.14	1999.71	2000.09

* Calculated by the method of Guardino *et al.*[7].

TABLE XIX

COMPARISON OF METHODS i–vi USING DATA FROM REF. 23 (NINETH COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	49.03	48.90	48.86	48.98	48.90	48.90
Slope	0.72813	0.72660	0.72633	0.72576	0.72659	0.72660
Intercept	−1.37438	−1.35829	−1.35529	−1.35403	−1.35811	−1.35820
Alkanes used	C_6,C_9,C_{12}	C_5–C_{12}	C_6–C_{12}	C_7–C_{10}	C_6–C_{12}	C_6–C_{12}
S of *S*/*N*	0.124	0.069	0.070	0.328	0.069	0.069
I_6	599.99	599.94	600.00	599.47	599.93	599.93
I_7	700.34	700.03	700.00	699.98	700.02	700.02
I_8	800.78	800.45	800.40	800.66	800.45	800.45
I_9	900.00	899.77	899.73	900.17	899.77	899.77
I_{10}	999.78	999.70	999.68	1000.25	999.70	999.70
I_{11}	1099.70	1099.82	1099.83	1100.49	1099.82	1099.81
I_{12}	1200.00	1200.31	1200.36	1201.11	1200.31	1200.31

TABLE XX

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 23 (NINETH COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	48.90*	49.33	53.95
Constant	1.97752	2.16304	3.94557
Coefficient *X*	1.30509	1.21820	$5.30036 \cdot 10^{-1}$
Coefficient X^2	$1.47794 \cdot 10^{-2}$	$2.85073 \cdot 10^{-2}$	$3.39101 \cdot 10^{-2}$
Coefficient X^3	$-9.73557 \cdot 10^{-4}$	$-1.69971 \cdot 10^{-3}$	$3.11016 \cdot 10^{-2}$
Coefficient X^4	*N*/*A*	*N*/*A*	$-5.36840 \cdot 10^{-3}$
Coefficient X^5	*N*/*A*	*N*/*A*	$2.62510 \cdot 10^{-4}$
S of *S*/*N*	0.021	0.020	0.009
I_6	600.05	600.04	600.02
I_7	699.80	699.87	699.97
I_8	800.28	800.27	800.12
I_9	899.85	899.80	899.84
I_{10}	999.98	999.98	1000.13
I_{11}	1100.04	1100.09	1099.95
I_{12}	1199.99	1199.97	1200.01

* Calculated by the method of Guardino *et al.*[7].

(5) Mathematical dead time shows similar trends with methods ii, v and vi giving identical results, and the other methods showing slight variations. In fact, the variation between the six methods is, in most cases, less than 1%.

(6) In general, the variation is larger for the lower homologues (Tables III, V, VII, IX, XV and XXIII) than for the higher homologues (Tables XI, XVII, XIX and

TABLE XXI

COMPARISON OF METHODS i–vi USING DATA FROM REF. 23 (TENTH COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v*	*vi*
Dead time	112.18	112.12	112.02	112.16	112.13	112.13
Slope	0.62544	0.62405	0.62275	0.62528	0.62403	0.62405
Intercept	−0.61631	−0.60239	−0.58936	−0.61170	−0.60227	−0.60243
Alkanes used	C_5,C_8,C_{11}	C_5–C_{11}	C_5–C_{11}	C_7–C_{11}	C_5–C_{11}	C_5–C_{11}
S of S/N	0.296	0.187	0.230	0.311	0.187	0.187
I_5	500.00	499.67	500.00	499.68	499.64	499.64
I_6	601.36	600.89	600.78	600.93	600.87	600.87
I_7	700.34	699.90	699.66	699.86	699.89	699.89
I_8	800.00	799.68	799.46	799.51	799.67	799.67
I_9	899.70	899.54	899.43	899.21	899.54	899.54
I_{10}	1000.11	1000.15	1000.19	999.64	1000.15	1000.14
I_{11}	1100.00	1100.24	1100.47	1099.55	1100.25	1100.24

TABLE XXII

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 23 (TENTH COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	112.12*	109.65	97.08
Constant	1.08760	$-8.12517 \cdot 10^{-1}$	−67.8340
Coefficient X	1.50524	2.51497	64.4043
Coefficient X^2	$2.41895 \cdot 10^{-2}$	$-1.58528 \cdot 10^{-1}$	−23.4548
Coefficient X^3	$-1.89933 \cdot 10^{-3}$	$9.24411 \cdot 10^{-3}$	4.44027
Coefficient X^4	*N/A*	*N/A*	$-4.23596 \cdot 10^{-1}$
Coefficient X^5	*N/A*	*N/A*	$1.62262 \cdot 10^{-2}$
S of S/N	0.108	0.012	0.000
I_5	499.78	500.01	500.02
I_6	600.59	600.10	600.01
I_7	699.70	699.83	700.01
I_8	799.79	800.18	800.00
I_9	899.90	899.92	900.00
I_{10}	1000.43	1000.01	1000.00
I_{11}	1099.82	1100.00	1100.00

* Calculated by the method of Guardino *et al.*[7].

XXI) as would be expected if the lower homologues showed some non-linear behaviour.

(7) The degree of variation appears to be independent of the stationary phase or solid support. For example Tables III, V, VII, IX and XV all give a high variation over a range of phases as well as porous polymers. On the other hand Tables XI,

TABLE XXIII

COMPARISON OF METHODS i–vi USING DATA FROM REF. 6 (ELEVENTH COLUMN, TABLE I)

	Method					
	i	*ii*	*iii*	*iv*	*v**	*vi*
Dead time	2.204	2.195	2.214	2.196	2.195	2.195
Slope	0.31034	0.30461	0.31412	0.30653	0.30452	0.30467
Intercept	−2.44093	−2.39098	−2.50736	−2.39496	−2.38983	−2.39108
2-Ketoalkanes used	C_2,C_5,C_8,C_{11}	C_1–C_{15}	C_1–C_{15}	C_1,C_3,C_5,C_7,C_{11}	C_1–C_{15}	C_1–C_{15}
S of S/N	89.627	74.386	121.936	95.876	74.384	74.354
I_1	112.06	117.16	100.04	115.61	116.82	117.06
I_2	193.21	195.00	186.91	193.41	194.68	194.91
I_3	285.28	284.56	283.57	282.81	284.27	284.47
I_4	340.73	388.36	392.48	386.31	388.10	388.27
I_5	495.89	492.86	499.74	490.41	492.64	492.77
I_6	604.38	601.42	609.38	598.48	601.23	601.32
I_7	710.63	708.28	716.08	704.81	708.12	708.17
I_8	808.82	807.38	814.23	803.37	807.25	807.25
I_9	907.81	907.53	912.88	902.97	907.43	907.39
I_{10}	1006.90	1007.98	1011.41	1002.83	1007.91	1007.82
I_{11}	1102.08	1104.58	1105.89	1098.87	1104.54	1104.41
I_{12}	1197.81	1201.84	1200.80	1195.55	1201.83	1201.65
I_{13}	1293.15	1298.76	1295.23	1291.89	1298.78	1298.56
I_{14}	1383.58	1394.82	1388.71	1387.36	1394.87	1394.60
I_{15}	1482.55	1491.46	1482.66	1483.40	1491.54	1491.22
Time (min)	28:25	2:04	1:12	21:14	6:00**	20:14

* The method did not converge when using eqn. 2. The quoted results were obtained using the alternative method (see text).

** Estimated.

XVII, XIX and XXI which involve stationary phases ranging all the way from low to high polarity, show very similar (low) variation.

(8) The three most accurate methods (ii, v and vi) give a variation within the required one unit for almost all data involving only homologues above C_4 (the exception is the data from Guardino *et al.*[7]; Table XIII), while all results involving the lower homologues include at least one index value outside the one unit requirement. The other methods fail to meet the requirement for most sets of data.

(9) The calculated Kováts indices of three theoretical compounds (S_1–S_3) have been added to the initial Tables (III–XIV). All methods with the exception of method iv give very similar results when interpolation is involved (in almost all cases within one unit). However the variation is significantly greater when extrapolation is involved.

(10) Finally, Table XXIII gives examples of typical times for each of the methods. It can be seen that, of the three most accurate methods, the method of Guardino *et al.*[7] is the fastest. It should however be recognised that the timing of this method is highly dependent not only on the initial estimate, but also on the actual dead time. This results from the method's stepwise approach in which a dead

TABLE XXIV

COMPARISON OF METHODS vii–ix USING DATA FROM REF. 6 (ELEVENTH COLUMN, TABLE I)

	Method		
	vii	*viii*	*ix*
Dead time	2.195*	2.287	2.222
Constant	7.82714	8.07764	7.85166
Coefficient X	3.17692	2.91948	3.15038
Coefficient X^2	$9.78213 \cdot 10^{-3}$	$1.85079 \cdot 10^{-1}$	$1.62172 \cdot 10^{-1}$
Coefficient X^3	$3.49483 \cdot 10^{-2}$	$-1.42877 \cdot 10^{-2}$	$-3.00569 \cdot 10^{-2}$
Coefficient X^4	*N/A*	*N/A*	$-1.82949 \cdot 10^{-2}$
Coefficient X^5	*N/A*	*N/A*	$6.98252 \cdot 10^{-3}$
S of S/N	35.762	9.239	0.814
I_1	111.68	103.88	100.20
I_2	195.13	198.08	200.88
I_3	288.70	294.96	299.13
I_4	394.44	399.32	401.15
I_5	498.55	500.41	498.96
I_6	604.94	603.96	600.23
I_7	708.64	706.01	701.94
I_8	804.51	801.55	798.73
I_9	901.73	899.44	898.82
I_{10}	1000.15	999.23	1000.93
I_{11}	1096.24	1096.85	1100.08
I_{12}	1194.93	1196.78	1200.26
I_{13}	1295.79	1298.02	1300.16
I_{14}	1398.73	1399.88	1399.39
I_{15}	1505.83	1503.80	1500.26
Time (min)	0:41	13:59	14:13

* Calculated by the method of Guardino *et al.*[7].

time such as 50.11 will be found much faster than 49.99. A factor of 2 to 3 in timing is not unusual.

The general conclusion from these analyses is that, of the linear methods, the method of Guardino *et al.*[7] is not only the most accurate, but also the fastest. This conclusion differs markedly from our previous papers. This results from the newly recognised instability of the method of Grobler and Balizs[5], resulting in the need to check all homologue combinations to find the best fit. This greatly increases both the complexity of the program and its running time.

The accuracy that can be expected using the method of Guardino *et al.*[7] depends mainly on the region of the homologue curve that is involved. An accuracy of one unit or better can be expected in most cases when using homologues above C_4.

Polynomial methods (vii to ix)

(11) Comments 6 to 9 are equally applicable to the polynomial methods. In particular the problems introduced by extrapolation are increased. This can be seen in Tables IV and X where the variation introduced by fitting a degree 5 polynomial

TABLE XXV

COMPARISON OF POLYNOMIALS OF INCREASING DEGREE USING DATA FROM REF. 21 (FOURTH COLUMN, TABLE I)

	Polynomial					
	Degree 2	*Degree 3*	*Degree 4*	*Degree 5*	*Degree 6*	*Degree 7*
Dead time	42.000	31.312	3.141	−171.111*	41.889	40.990
Constant	−4.17526	−14.7303	$-1.79500 \cdot 10^{2}$	$-1.78291 \cdot 10^{5}$	$-1.65570 \cdot 10^{2}$	$1.20649 \cdot 10^{2}$
Coefficient X	1.93584	7.48229	$1.23557 \cdot 10^{2}$	$1.46243 \cdot 10^{5}$	$-2.27509 \cdot 10^{2}$	$-1.47247 \cdot 10^{2}$
Coefficient X^2	$-3.33179 \cdot 10^{-2}$	−1.01689	−31.8332	$-4.79492 \cdot 10^{4}$	$1.26710 \cdot 10^{2}$	66.8810
Coefficient X^3	*N/A*	$5.84255 \cdot 10^{-2}$	3.70294	$7.85470 \cdot 10^{3}$	−36.6818	−12.5451
Coefficient X^4	*N/A*	*N/A*	$-1.616370 \cdot 10^{-1}$	$-6.42814 \cdot 10^{2}$	5.87569	$1.71373 \cdot 10^{-1}$
Coefficient X^5	*N/A*	*N/A*	*N/A*	21.0238	$-4.94596 \cdot 10^{-1}$	$2.96994 \cdot 10^{-1}$
Coefficient X^6	*N/A*	*N/A*	*N/A*	*N/A*	$1.71148 \cdot 10^{-2}$	$-4.27321 \cdot 10^{-2}$
Coefficient X^7	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	$1.90494 \cdot 10^{-3}$
S of S/N	5.043	2.758	1.251	0.090	0.0000	0.000
I_1	98.97	99.44	99.66	100.22	100.00	100.00
I_2	204.05	202.36	201.33	199.89	200.00	200.00
I_3	296.31	296.92	298.07	299.59	300.00	300.00
I_4	399.17	401.03	401.55	400.56	400.00	400.00
I_5	501.22	501.14	499.19	499.73	500.00	500.00
I_6	601.17	598.80	600.25	600.07	600.00	600.00
I_7	699.13	700.33	699.96	700.00	700.00	700.00
S_2 (450 sec)	625.76	623.45	626.22	608.71	623.96	623.74

* Note that the mathematical dead time is negative.

TABLE XXVI

COMPARISON OF POLYNOMIALS OF INCREASING DEGREE USING DATA FROM REF. 17 (SEVENTH COLUMN, TABLE I)

	Polynomial					
	Degree 2	*Degree 3*	*Degree 4*	*Degree 5*	*Degree 6*	*Degree 7*
Dead time	877.296	876.583	874.385	871.070	870.890	871.990
Constant	$2.91372 \cdot 10^{-1}$	$-1.83319 \cdot 10^{-1}$	-2.68483	-12.5764	-13.7271	-7.76144
Coefficient X	$8.44268 \cdot 10^{-1}$	1.12532	3.04498	12.5930	13.9253	7.18596
Coefficient X^2	$-1.46080 \cdot 10^{-4}$	$-5.34703 \cdot 10^{-2}$	$-6.02552 \cdot 10^{-1}$	-4.28517	-4.93052	-1.76426
Coefficient X^3	*N/A*	$3.22313 \cdot 10^{-3}$	$7.15720 \cdot 10^{-2}$	$7.73107 \cdot 10^{-1}$	$9.39076 \cdot 10^{-1}$	$1.60315 \cdot 10^{-1}$
Coefficient X^4	*N/A*	*N/A*	$-3.10528 \cdot 10^{-3}$	$-6.87301 \cdot 10^{-2}$	$-9.25059 \cdot 10^{-2}$	$1.07513 \cdot 10^{-2}$
Coefficient X^5	*N/A*	*N/A*	*N/A*	$2.40552 \cdot 10^{-3}$	$4.19825 \cdot 10^{-3}$	$-2.33916 \cdot 10^{-3}$
Coefficient X^6	*N/A*	*N/A*	*N/A*	*N/A*	$-5.54936 \cdot 10^{-5}$	$3.03142 \cdot 10^{-5}$
Coefficient X^7	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	$6.11239 \cdot 10^{-6}$
S of S/N	2.815	1.764	0.483	0.000	0.001	0.001
I_1	99.59	99.85	99.99	100.03	100.06	100.06
I_2	202.54	201.35	200.41	200.01	200.02	200.02
I_3	296.61	297.52	299.14	300.00	300.01	300.01
I_4	400.49	401.64	401.16	400.00	400.00	400.00
I_5	500.96	500.49	499.04	500.00	500.00	500.00
I_6	600.38	598.87	600.43	600.00	600.00	600.00
I_7	699.47	700.39	699.43	700.00	700.00	700.00

TABLE XXVII

COMPARISON OF POLYNOMIALS OF INCREASING DEGREE USING DATA FROM REF. 6 (ELEVENTH COLUMN, TABLE I)

	Polynomial					
	Degree 2	*Degree 3*	*Degree 4*	*Degree 5*	*Degree 6*	*Degree 7*
Dead time	2.2567	2.2871	2.2768	2.2221	2.2000	2.2000
Constant	7.99171	8.07764	8.02309	7.85166	7.78298	7.78263
Coefficient X	2.98581	2.91948	2.98025	3.15038	3.21908	3.22759
Coefficient X^2	0.143359	0.185079	$2.08878 \cdot 10^{-1}$	$1.62172 \cdot 10^{-1}$	$1.27088 \cdot 10^{-1}$	$1.28838 \cdot 10^{-1}$
Coefficient X^3	*N/A*	$-1.42877 \cdot 10^{-2}$	$-2.32421 \cdot 10^{-2}$	$-3.00569 \cdot 10^{-2}$	$-3.00271 \cdot 10^{-2}$	$-4.60557 \cdot 10^{-2}$
Coefficient X^4	*N/A*	*N/A*	$-6.95261 \cdot 10^{-3}$	$-1.82949 \cdot 10^{-2}$	$-1.55690 \cdot 10^{-2}$	$-1.68716 \cdot 10^{-2}$
Coefficient X^5	*N/A*	*N/A*	*N/A*	$6.98252 \cdot 10^{-3}$	$1.21051 \cdot 10^{-2}$	$1.94619 \cdot 10^{-2}$
Coefficient X^6	*N/A*	*N/A*	*N/A*	*N/A*	$-2.23870 \cdot 10^{-3}$	$-2.00149 \cdot 10^{-3}$
Coefficient X^7	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	$-9.48216 \cdot 10^{-4}$
S of S/N	13.843	9.239	1.486	0.814	0.739	0.698
I_1	104.31	103.88	100.47	100.20	100.17	100.31
I_2	197.49	198.08	201.88	200.88	200.84	200.51
I_3	294.39	294.96	298.61	299.13	299.31	299.38
I_4	399.34	399.32	400.44	401.15	401.20	401.48
I_5	500.98	500.41	499.09	498.96	498.79	498.84
I_6	604.79	603.96	601.11	600.23	600.05	599.83
I_7	706.74	706.01	702.85	701.94	701.96	701.75
I_8	801.94	801.55	799.08	798.73	798.92	798.94
I_9	899.34	899.44	898.39	898.82	899.03	899.26
I_{10}	998.60	999.23	999.98	1000.93	1000.95	1001.15
I_{11}	1095.84	1096.85	1099.23	1100.08	1099.88	1099.83
I_{12}	1195.65	1196.78	1200.17	1200.26	1200.02	1199.75
I_{13}	1297.20	1298.02	1301.10	1300.16	1300.17	1300.05
I_{14}	1399.97	1399.88	1400.62	1399.39	1399.74	1400.07
I_{15}	1505.60	1503.80	1499.26	1500.26	1500.10	1499.98
S (2.85 min)	647.10	646.28	643.16	642.22	642.10	641.85

TABLE XXVIII

COMPARISON OF POLYNOMIALS OF INCREASING DEGREE USING DATA FROM REF. 7 (SIXTH COLUMN, TABLE I)

	Polynomial					
	Degree 2	*Degree 3*	*Degree 4*	*Degree 5*	*Degree 6*	*Degree 7*
Dead time	229.617	233.913	237.109	224.494	156.542	182.151
Constant	1.11737	3.29983	5.24385	−28.1170	$-6.05712 \cdot 10^3$	$1.55928 \cdot 10^3$
Coefficient X	1.83300	$8.29188 \cdot 10^{-1}$	$-2.97637 \cdot 10^{-1}$	28.8772	$6.00307 \cdot 10^3$	$-2.41505 \cdot 10^3$
Coefficient X^2	$3.27995 \cdot 10^{-3}$	$1.59435 \cdot 10^{-1}$	$4.15346 \cdot 10^{-1}$	−10.0320	$-2.47211 \cdot 10^3$	$1.50848 \cdot 10^3$
Coefficient X^3	*N/A*	$-8.15049 \cdot 10^{-3}$	$-3.48156 \cdot 10^{-2}$	1.85080	$5.41555 \cdot 10^2$	$-5.02054 \cdot 10^2$
Coefficient X^4	*N/A*	*N/A*	$1.06387 \cdot 10^{-3}$	$-1.68891 \cdot 10^{-1}$	−66.5025	97.3184
Coefficient X^5	*N/A*	*N/A*	*N/A*	$6.08758 \cdot 10^{-3}$	4.33884	$-1.10584 \cdot 10^{-1}$
Coefficient X^6	*N/A*	*N/A*	*N/A*	*N/A*	$-1.17473 \cdot 10^{-1}$	$6.84866 \cdot 10^{-1}$
Coefficient X^7	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	$-1.78836 \cdot 10^{-2}$
S of S/N	1.767	1.614	1.572	0.908	0.298	0.281
I_7	700.52	700.24	700.24	699.99	700.13	700.10
I_8	798.68	799.30	799.21	800.16	799.54	799.64
I_9	900.05	900.14	900.10	899.73	900.84	900.74
I_{10}	1001.23	1000.82	1000.87	999.88	999.04	999.07
I_{11}	1101.58	1101.15	1101.25	1101.24	1100.78	1100.83
I_{12}	1196.96	1196.90	1196.95	1198.04	1199.56	1199.51
I_{13}	1300.62	1301.05	1300.98	1301.51	1300.16	1300.18
I_{14}	1400.47	1400.96	1400.82	1399.37	1399.96	1399.96
I_{15}	1499.91	1499.47	1499.56	1500.11	1500.00	1500.01
S (500 sec)	1148.47	1148.19	1148.28	1148.92	1149.72	1149.70

is so large that it results in a Kováts index for the theoretical unknown that is 2 to 5 times the one calculated by the method of Heeg *et al.*[12]. However, it is also apparent that with these methods care must be taken when interpolation is used. For example Table X gives a result for a degree 5 fit which is obviously too low [this method also gives a negative dead time (see the next point)].

(12) The variation in the mathematical dead time is so large with these methods, that it can no longer be equated with the system hold-up volume. In one case (Table X), the result is negative! Therefore the term should be considered in the same way as the equation coefficients, a constant which does not have a direct physical analogue. There is, however, no reason that this should effect the calculation of Kováts indices.

(13) In general, the method of Heeg *et al.*[12] when used in conjunction with the dead time calculated by the method of Guardino *et al.*[7] produces only a slight improvement over the best linear method. This contrast with the results quoted in the paper by Heeg *et al.*[12]. The difference occurs because of the limited data tested by Heeg *et al.*[12]. This highlights the reason for our use of such an extensive set of data.

(14) The shortcoming of the method of Heeg *et al.*[12] is in its use of a dead time calculated by a linear method. This is shown by the results of method viii in which an iteration is carried out on the dead time while fitting the cubic equation, thus allowing the mathematical dead time to be optimised for a cubic fit, resulting in an improvement in results.

(15) Even with the method viii, the individual Kováts indices are not always within 1 unit of their known values. Therefore method ix shows the result of fitting a degree 5 polynomial to the data. This produces an excellent fit with all the data except the data of Guardino *et al.*[7] (Table XIV) and the HPLC data of Tulder *et al.*[6] (Table XIV).

(16) Tables XXV–XXVIII show the effect of fitting polynomials of degree 2 up to 7 to the data. As would be expected, the accuracy of fit improves as the degree of the polynomial increases. Three points emerge from these tables. First, it is necessary to carry out all calculations to a very high precision (in our case up to 16 digits were carried throughout the calculations with the final results quote to a lower accuracy for convenience) with increasing precision required as the degree of the polynomial increases. Second, great care must be taken in the interpretation of the results as discussed in points 11 and 12. Finally, some data will need a very complex polynomial to be fitted to reduce the individual variation below one unit. For example, while a degree 5 polynomial is sufficient for data by Ettre[17] (Table XXVI), both data by Wainwright *et al.*[21] (Table XXV) and data by Guardino *et al.*[7] (Table XXVIII) require a degree 6 polynomial, while the HPLC data of van Tulder *et al.*[6] would need a polynomial of degree higher than 7.

(17) Even with these complex methods, the largest difference in the retention index of the theoretical unknown between the best of the linear methods and the most suitable polynomial method is approximately two and a half units for the lower homologues and one unit for the higher homologues.

The general conclusion from these analyses is that the polynomial methods offer very little improvement for the higher homologues and only a small improvement for the lower homologues. Therefore they should only be necessary when extreme accuracy is required or where highly variable data is involved. In the latter

TABLE XXIX

VARIATION OF KOVÁTS RETENTION INDICES DUE TO RANDOM DEVIATIONS

	t_{Rz}*	t_{Rz} (rounded)	t_{Rz} (+ deviation)
t_m	100.00	100.01	99.83
b	0.620	0.620	0.616
c	−0.610	−0.609	−0.582
R^2	1.000000	0.999999	0.999939
z			
4	106.49	106.5	106.4
5	112.06	112.1	112.0
6	122.42	122.4	122.3
7	141.68	141.7	141.8
8	177.48	177.5	177.6
9	244.03	244.0	239.9
10	367.74	367.7	376.6
Kováts retention indices			
4	400.03	399.95	399.99
5	499.99	500.33	500.05
6	600.01	599.75	599.58
7	700.02	700.04	700.99
8	800.01	800.04	801.09
9	900.00	899.99	896.59
10	1000.00	1000.02	1001.76

* Calculated from $\ln (t_{Rz} - 100) = 0.62z - 0.61$.

case the experimental method should be checked for problems and the analyses repeated. When used, great care must be taken in the interpretation of results and under no conditions should results be extrapolated outside the range of homologues used in the analysis.

A perspective can be put on the above discussion by reference to Table XXIX. The first column shows calculated retention times based on eqn. 3.

$$\ln (t_{Rz} - 100) = 0.62z - 0.61 \qquad (3)$$

The second column in Table XXIX shows that rounding the data to the nearest 0.1 of a second introduces a maximum error of 0.33 in the retention indices. However, the third column which involves a deviation of ±0.1 sec to each retention time increases this error to a maximum of 3.41 units. To reduce this below 1 unit requires a degree 4 polynomial to be fitted to the data. This emphasises the need to ensure high accuracy in the raw data if accurate Kováts indices are to be achieved.

RECOMMENDATIONS

Any recommendation must involve not only the actual method to be used to achieve the desired accuracy, but also an outline of a suitable experimental procedure.

We believe the following is a suitable procedure (it is assumed that the equipment has been checked for correct operation, in particular the stability and accuracy of oven temperature and carrier gas flow-rate):

(a) Determine the range over which subsequent measurements will be carried out and choose a series of suitable homologues to completely cover this range, ensuring that extrapolation will not be required.

(b) Carry out a series of runs using the homologues under the appropiate conditions. Retention times should be measured to at least the nearest 0.1 of a second unless extremely long times are involved, in which case a lower accurary may be acceptable. The average retention times should be used in all subsequent analyses and standard deviations calculated to ensure an acceptable variation.

(c) The method chosen to analyse the data will depend on the required accuracy. For the majority of analyses the method of Guardino *et al.*[7] will give the required accuracy (especially if only higher homologues are involved). As a routine part of all such analyses, the retention indices of the standard homologues should be determined. If their deviations are not acceptable, a polynomial may be fitted to the data. The required degree of the polynomial will have to be determined by trial and error or by experience. As previously mentioned the results of such a procedure should be carefully evaluated before they are accepted.

(d) Three or four homologues (two for simple analyses) should be added to the sample which is to be chromatographed. The homologues should bracket the sample at both ends where possible and must have been included in the standard runs. The sample should then be run several times and the means of the retention times calculated. The homologues can now be compared with the standard runs to ensure that an acceptable accuracy is being maintained.

(e) If satisfied with the accuracy of the retention times, Kováts indices can now be calculated and should be quoted only to the nearest whole number unless they are being used for a particular purpose which requires higher accuracy.

The procedure outlined above should allow an interlaboratory repeatibility of 1 unit for the higher retention indices and 2 units for the lower retention indices.

REFERENCES

1 E. Kováts, *Helv. Chim. Acta*, 41 (1958) 1915.
2 A. Wehrli and E. Kováts, *Helv. Chim. Acta*, 42 (1959) 2709.
3 L. S. Ettre, *Anal. Chem.*, 36(8) (1964) 31A.
4 R. J. Smith, J. K. Haken and M. S. Wainwright, *J. Chromatogr.* (1985) in press.
5 A. Grobler and G. Balizs, *J. Chromatogr. Sci.*, 12 (1974) 57.
6 P. J. M. van Tulder, J. P. Franke and R. A. de Zeeuw, presented at the VIIth International Symposium on Column Liquid Chromatography, Baden-Baden, May 3–6. 1983.
7 X. Guardino, J. Albaigés, G. Firpo, R. Rodríguez-Viñals and M. Gassiot, *J. Chromatogr.*, 118 (1976) 13.
8 J. A. Garcia Dominguez, J. Garcia Munoz, E. Fernandez Sanchez and M. J. Molera, *J. Chromatogr. Sci.*, 15 (1977) 520.
9 L. Ambrus, *J. Chromatogr.*, 294 (1984) 328.
10 Á. Tóth and E. Zala, *J. Chromatogr.*, 284 (1984) 53.
11 J. A. Nelder and R. Mead, *Comput. J.*, 7 (1965) 308.
12 F. J. Heeg, R. Zinburg, H. J. Neu and K. Ballschmiter, *Chromatographia*, 12 (1979) 451.
13 W. S. Dorn and D. D. McCracken, *Numerical Methods with Fortran IV Case Studies*, Wiley, New York, 1972.
14 M. S. Wainwright and J. K. Haken, *J. Chromatogr.*, 184 (1980) 1.
15 R. J. Smith, J. K. Haken and M. S. Wainwright, *J. Chromatogr.*, 147 (1978) 65.
16 H. Groenendijk and A. W. C. van Kemenade, *Chromatographia*, 1 (1968) 472.
17 L. S. Ettre, *Chromatographia*, 13 (1980) 73.
18 J. Guberska, *Chem. Anal. (Warsaw)*, 19 (1974) 161.
19 W. E. Sharples and F. Vernon, *J. Chromatogr.*, 161 (1978) 83.
20 G. Schomburg and G. Dielmann, *J. Chromatogr. Sci.*, 11 (1973) 151.
21 M. S. Wainwright and J. K. Haken, *J. Chromatogr.*, 256 (1983) 193.
22 J. Ševčík and M. S. H. Löwentap, *J. Chromatogr.*, 147 (1978) 75.
23 M. S. Wainwright, J. K. Haken and D. Srisukh, *J. Chromatogr.*, 188 (1980) 246.