This article was downloaded by: On: 24 January 2011 Access details: Access Details: Free Access Publisher Taylor & Francis Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Liquid Chromatography & Related Technologies

Publication details, including instructions for authors and subscription information: http://www.informaworld.com/smpp/title~content=t713597273

On the Use of the Knox Equation. I. the Fit Problem

Alain Berthod^a ^a Laboratoire des Sciences Analytiques U.A. CNRS 434 Universite de Lyon I, Villeurbanne, cedex, France

To cite this Article Berthod, Alain(1989) 'On the Use of the Knox Equation. I. the Fit Problem', Journal of Liquid Chromatography & Related Technologies, 12: 7, 1169 — 1185 To link to this Article: DOI: 10.1080/01483918908049499 URL: http://dx.doi.org/10.1080/01483918908049499

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

ON THE USE OF THE KNOX EQUATION. I. THE FIT PROBLEM

ALAIN BERTHOD

Laboratoire des Sciences Analytiques U.A. CNRS 434 Universite de Lyon I 69622 Villeurbanne, cedex France

ABSTRACT

To obtain significant A, B, and C parameters of the Knox equation, which corresponds to the plot of the reduced plate height, h, versus the reduced linear velocity, $\underline{\mathbf{v}}$, actual results must be fitted correctly. The Knox equation is analyzed, the rôle of each individual parameter is shown. The equation producing the coordinates of the minimum plate height (maximum efficiency) is fully derived from the derivative of the Knox equation. Two tables giving the minimum coordinates for usual A, B, and C ranges are listed. The classical fit method is derived and analyzed. A graphical fit method which uses the remarkable graphical capabilities of modern spreadsheet packages is described. A synthetical set of data was fitted. It is demonstrated that each Knox parameter, obtained through fit procedures, must be given with the fitting confidence interval which is often in the 20% range.

In Liquid Chromatography (LC), column testing is daily done for various important purposes: studies of selectivity and retention variation, and/or efficiency evaluation. It was recently shown (1) that 25 ODS phases had different selectivity comportment toward polycyclic aromatic hydrocarbons. This is a general trend (2): lot to lot variability of commercial columns is well-known. In evaluation of LC column performances, the silanol and metal activity, the bonding type (monomeric or polymeric) and bonding density, the pore size and volume are the important points of the stationary phase to be known (3). When assessing the total column efficiency, plots of h, the reduced plate height, versus $\underline{\mathbf{v}}$, the reduced mobile phase velocity, are commonly used.

The most widely accepted plate height equation is the Knox equation (4-6) that is:

$$h = A \underline{\mathbf{v}}^{1/3} + B/\underline{\mathbf{v}} + C\underline{\mathbf{v}} \qquad \text{Eq. 1}$$

in which: A, B and C are the constants of the Knox equation, h is the reduced plate height with $h=H/d_p$ (H is the plate height and d_p , the particle diameter), $\underline{\mathbf{v}}$ is the reduced velocity with $\underline{\mathbf{v}}=ud_p/D_m$ (u is the mobile phase linear velocity (cm/s), and D_m is the solute diffusion coefficient in the mobile phase (cm²/s)).

Measurements of efficiencies at different flow rates allows one to obtain a Knox plot which led to the A, B and C terms. The A term is related to the packing quality. A well-packed column has a A value around unity or less. Stationary phases squeezing or lissolution produces discontinuity inside the stationary phase bed. Efficiency is greatly reduced inducing a A term increase (7).

The B term is related to solute longitudinal diffusion. It is responsible for the decrease in efficiency at very low flow rates. The C term represents the mass-transfer contribution. It depends on the solute retention (k') on the pore size and volume, and on the solute diffusion coefficient in the mobile phase and in the stationary phase (7-8).

The A, B and C values allow to obtain an insight in the solute-stationary phase exchange. They are very important in fundamental chromatographic studies (9). To obtain significant A, B, and C values, h versus $\underline{\mathbf{v}}$ plots of actual results must be correctly fitted, that is the topic of this paper. Of course the plate heights

must be correctly measured, which is the matter of Part II of this work.

After an analytical description of the Knox equation, the importance of the fit in the obtention of the A, B, and C terms, and their subsequent significance is exposed. The classical least square method is described and analyzed. A different approach of the fit problem, using the tremendous graphical capabilities of modern software packages is presented.

1- Analytical properties of the Knox equation

Derivative and minimum of h

It is mathematically well-known that the minimum value of h occurs when the dh/d<u>w</u> derivative is nil. The particular point, h_0 -<u>w</u>₀, corresponds to the maximum efficiency of the column that's why it is an interesting point to locate. The derivative of Eq. 1 is:

$$dh/d\underline{v} = A/(3\underline{v}^{2/3}) - B/\underline{v}^{2} + C$$
 Eq. 2

With the variable change $X=\underline{v}^{-2/3}$, Equation 2 becomes:

$$-B X^{3} + AX/3 + C = 0$$

This last equation has always at least one real solution:

with
$$\blacktriangle = (C^2/B^2)(1 - [4A^3/(729BC^2)])$$

and $\phi = 4A^3 / (729BC^2)$

the solutions are:

-when $\blacktriangle > 0$ (or $\phi < 1$) $X_{o} = (C/2B)^{1/3} [(1 + (1-\phi)^{1/2})^{1/3} + (1 - (1-\phi)^{1/2})^{1/3}]$ -when $\blacktriangle < 0$ (or $\phi > 1$) $X_{o} = 2/3 (A/B)^{1/2} \cos(1/3 \arccos(1/\phi))$

in both cases: $\underline{\mathbf{v}}_{0} = X_{0}^{-3/2}$

Table I lists the reduced velocity corresponding to the minimum h value for a well-packed column (A=0.5) and for a poorly packed column (A=2), with the B and C terms in the usual ranges. It can be seen that the optimal flow rate increases markedly when both B and C decrease. This means that fast and efficient separations can be achieved only with well-packed columns (low A values) and fast mass-transfer between phases (low C values).

A effect

Figure 1 shows the effect of A. While the B and C values were kept constant, the value of A was varied from 0.5 to 10, corresponding to a very well-packed column (A=0.5) to a poorly packed column or a column with a squeezed stationary phase (A=10). One can see that a poorly packed column presents a sharp minimum, at very low flow rates, in the h vs \underline{v} plot. A well-packed column presents a diffuse minimum. It is possible to work at high flow rates with a good efficiency. If needed, Figure 1 demonstrates the extreme importance of a very good stationary phase packing.

B effect

Figure 2 shows the effect of B. At high flow rates, the influence of B is insignificant: the plate height contribution of the B/\underline{v} term decreases rapidly when \underline{v} increases. The most important effect of a B increase is an increase of the optimum flow rate.

C effect

Figure 3 shows the effect of C. As shown by the derivative (Eq. 2), the C term has the main importance on slope at high flow rates. If the solute mass-transfer is slow (elevated C values), it will be necessary to work at reduced flow rates to obtain an adequate efficiency.

Figures 1-3 illustrate the fit problem: with different sets of A, B, and C values, it is possible to obtain rather similar overall plots as demonstrated below.

в	2	3	4	5	6	8	10
C=10 ⁻⁶	6.45	8.74	10.84	12.82	14.69	18.24	21.56
0.01	5.98	7.84	9.43	10.84	12.10	14.50	16.80
0.05	4.18	5.36	6.37	7.28	8.10	9.59	10.92
0.10	3.39	4.29	5.05	5.73	6.35	7.46	8.44
0.15	2.94	3.69	4.34	4.90	5.42	6.34	7.16
0.20	2.64	3.30	3.87	4.36	4.82	5.63	6.34
0.30	2.24	2.79	3.26	3.68	4.05	4.72	5.31

Table I - Optimal reduced velocity

A = 0.5

Α	=	2
---	---	---

В	2	3	4	5	6	8	10
C=10 ⁻⁶	1.68	2.28	2.83	3.34	3.83	4.76	5.62
0.01	1.68	2.28	2.83	3.34	3.83	4.75	5.61
0.05	1.67	2.25	2.78	3.27	3.73	4.58	5.37
0.10	1.62	2.16	2.63	3.07	3.46	4.17	4.80
0.15	1.55	2.04	2.45	2.81	3.14	3.77	4.37
0.20	1.47	1.90	2.26	2.62	2.96	3.57	4.12
0.30	1.33	1.74	2,10	2.42	2.71	3.25	3.74

 $\underline{\mathbf{v}}$ values (dimensionless) for which dh/d $\underline{\mathbf{v}}$ (Eq. 2) is zero.







Figure 2: The effect of B.



Figure 3: The effect of C.

2- Fitting procedures

Classical fit method

The classical method to fit a set of N data, $h_i - \underline{v}_i$, is to use the Knox equation minimizing the deviation between the calculated and the actual values (least square method). This method is briefly described below:

Defining the S(A) term as the sum of the N experimental plate heights, h_i , multiplied by the reverse of the cubic root of the corresponding reduced velocity, $\underline{\mathbf{v}}_i$:

$$S(A) = \sum h_i \underline{v}_i^{-1/3} \qquad \text{Eq. 3}$$

also, from Eq. 1:

$$S(A) = NA + B \Sigma \underline{v}_{i}^{-4/3} + C \Sigma \underline{v}_{i}^{2/3}$$
 Eq. 4

then:

A =
$$(1/N) * [S(A) - B \Sigma \underline{v}_i^{-4/3} - C \Sigma \underline{v}_i^{2/3}]$$
 Eq. 5

Defining the same way S(B) and S(C):

$$S(B) = \Sigma h_i \underline{\mathbf{v}}_i = NB + A \Sigma \underline{\mathbf{v}}_i^{4/3} + C \Sigma \underline{\mathbf{v}}_i^2 \qquad \text{Eq. 6}$$

$$S(C) = \Sigma h_i / \underline{\mathbf{v}}_i = NC + A \Sigma \underline{\mathbf{v}}_i^{-2/3} + B \Sigma \underline{\mathbf{v}}_i^{-2} \qquad \text{Eq. 7}$$

it is possible to derive an expression for A, B, and C using a linear combination of the sums of the linear velocities raised to some ith power. Using the convenient notation:

$$S(x) = \Sigma \underline{v}_i^x$$
 Eq. 8

A can be expressed as:

$$A = (1/D) \bullet \{ [NS(A) - S(C)S(2/3)] * [N^2 - S(2)S(-2)] - [NS(B) - S(C)S(2)] * [NS(-4/3) - S(2/3)S(-2)] \} Eq. 9$$

in which D is the common denominator whose full derivation is exposed in Anner along with the B and C equations.

This widely used method can be easily automatized on microcomputers to give the A, B, and C values and the error of fit usually taken as the mean of the squared deviations between the measured and calculated h, values. We want to point out that the results, always computed in seconds, must be carefully considered. Even with a very low error of fit, the A, B, and C terms obtained may be highly questionable. Often one out of the three terms can be predetermined. For example, the C term can be estimated at high flow rates, neglecting the B term (7). On the opposite, the B term can be obtained at very low flow rates (9) or by the arrested elution method (10). Once a term is known, it can be put in Eqs 4-7 to simplify the derivation of the two remaining terms. For example, when B is known, A is estimated using:

$$A = (1/D') * [NS'(A) - S'(C)S(2/3)]$$
 Eq. 10

USE OF THE KNOX EQUATION. I

The complete derivation of the relations can be found in Annex. Commonly, the error of fit is higher when one of the Knox parameters was measured apart.

Visual fit method

Recently introduced software packages are powerful tools that must be known and used by Analytical Chemists. Spreadsheets are electronic worksheets arranged horizontally in rows and vertically in columns. This defines a grid of cells. Each cell can contain a number, a string of characters, a formula calling other cells, or a macro-command. Any cell in the worksheet can be linked to any other by a user-defined relationship. Most modern software packages, such as Lotus 123 (Lotus Development Corp., Cambridge, MA), Excel (Microsoft Corp.) or Quattro (Borland International Inc.), have graphic capabilities (11).

Placing the experimental $\underline{\mathbf{v}}_i$ values in column a, the corresponding \mathbf{h}_i values in column b, it is possible, with a keystroke, to visualize the experimental h vs $\underline{\mathbf{v}}$ plot. Placing the Knox equation (Eq. 1) in column c, referring to three named cells for the A, B, and C terms, it is possible to calculate a \mathbf{h}_i set corresponding to any A, B, and C values. A second plot, using column c (calculated \mathbf{h}_i values) and column a (experimental $\underline{\mathbf{v}}_i$) can be drawn with a different color and compared to the experimental plot. If the cell named A is changed, a keystroke (F10 key) allows to see the evolution of the calculated plot compared to the experimental plot. Placing in column d the squared deviation ($\mathbf{h}_i \exp_i - \mathbf{h}_i$ calc), the mean of this squared deviation can be used as a test to optimize the fit using macrocommands. This is described in Part II of the annex.

The main advantage of the visual method is to reveal erroneous experimental results. Points corresponding to inaccurate measurements will stand apart of the general trend of the plot. Such points can be removed to improve the goodness of fit (12).

Another feature put forward by the visual method is the absolute necessity to have experimental points well around the minimum of the h vs \underline{v} plot to be able to obtain a significant B term. As Figure 2 showed, if the minimum is not present in the experimental plot (too high \underline{v} values), any B value will fit the data.

BERTHOD

3- Goodness of fit

To illustrate the fit problem, an artificial set of data was generated using the values 1.01, 6.2, and 0.18 for A, B, and C, respectively. 15 couples, $h_i - \Psi_i$, were generated. The classical method gave the exact A, B, and C values with an error of fit of 10^{-5} due to computer truncating. When the exact B or C value was injected, the two other values were correctly found. An real set of data is much more difficult to handle. To generate a more realistic set of data, the previous computed set was convoluted to introduce a only 1% random deviation on each h_i value. Figure 4 shows the points and the fit obtained with the least square method. The error of fit was 5.1×10^{-4} , which is fairly acceptable. However, a only 1% error on the h data induced a 6%, 1.6%, and 5.6% error on the A, B, and C terms, respectively (Table II).

Figures 5 and 6 presents the same points with force-fitted plots using B=6.5 and B=5.7, respectively. The error of fit were 2.7×10^{-3} and 3.5×10^{-3} , respectively, which would be accepted by most people. The errors increased dramatically (Table II).

On Figures 4-6 (bottom), the contribution of each term is represented. This shows that the low A value of the Figure 5 fit is compensated by a high B and specially a high C value, and vice-versa with the Figure 6 fit. The sum of the three contributions produces acceptable fits on the reduce \underline{v} range studied.

Conclusion

Since the early derivation of fit procedures (13), use and abuse of fitting methods were done. For example, equations relating chromatographic retention times to solute refractive index or density were derived (14). It is very important to be aware that a "bug"free computer program will produce 8-digit parameters to fit an experimental set of results, when the first digit of such parameters could be not significant. This is the case in the fit of a h vs \underline{v} set of data using the Knox equation: if the set of experimental data does not display the minimum, the least square fitting method will



Figure 4: Top: synthetical set of data (open squares) fitted as A=0.95, B=6.3, and C=0.19 (full line). Bottom: the contribution of each separate term, the $A\underline{v}^{1/3}$ curve crosses the C \underline{v} line at \underline{v} =11.

Comment	A	error	В	error	С	error	error of fit
exact values	1.01	-	6.20	-	0.18	-	-
Figure 4	0.95	-6%	6.30	-1.6%	0.19	5.6%	0.00051
(fease square) Figure 5 (force fit B)	0.89	-12%	6.50	4.8%	0.20	11%	0.0027
(force fit B) (force fit B)	1.20	19%	5.70	-8.1%	0.15	-17%	0.0035

Table II - The goodness of fit

The result of such a fit must be given as: A = 1.04 ± 0.16 ; B = 6.10 ± 0.40 ; C = 0.17 ± 0.03 ; with an error of fit lower than 0.0035.







Figure 6: Top: synthetical set of data (open squares) fitted as A=1.20, B=5.7, and C=0.15 (full line). Bottom: the contribution of each separate term, the $A\underline{v}^{1/3}$ curve does not cross the C \underline{v} line in the \underline{v} range presented.

produce a B term (may be with 8 digits) which has no meaning (Figure 2).

When the Knox equation is used to fit h vs \underline{v} plots, the confidence interval of each fitted term A, B, and C should be given. Depending on the accuracy and reproducibility of the experimental results, the width of the confidence intervals is commonly higher than 20% of the term (9).

References

1-	Sander, L.C.; Wise, S.A., HRC & CC , <u>11</u> , 383 (1988).
2-	Scott, R.P.W.; Kucera, P., J. Chromatogr., <u>142</u> , 213 (1977).
3-	Sander, L.C., J. Chromatogr. Sci. , <u>26</u> , 380 (1988).
4	Knox, J.H.; Saleem, M., J. Chromatogr. Sci. , <u>10</u> , 80 (1972).
5-	Done, J.N.; Knox, J.H., J. Chromatogr. Sci. , <u>10</u> , 606 (1972).
6-	Kennedy, G.J.; Knox, J.H., J. Chromatogr. Sci. , <u>10</u> , 549 (1972).
7-	Knox, J.H., J. Chromatogr. Sci. , <u>15</u> , 352 (1977).
8-	Giddings, J.C., Dynamics of Chromatography, Marcel Dekker,
	New York, 1965.
9-	Berthod, A.; Chartier, F.; Rocca, J.L., J. Chromatogr., in
10-	Knox, J.H.; McLaren, L., Anal. Chem. , <u>36</u> , 1477 (1964).
11-	Allars, A., Int. Biotechnology Lab., <u>6, 2</u> , 38 (1988).
12-	Stout, R.W.; DeStephano, J.J.; Snyder, L.R., J. Chromatogr.,
	<u>282</u> . 263 (1983).
13-	Savitzky, A.; Golay, M.J.E., Anal. Chem. , <u>36</u> , 1627 (1964).
14-	Yixing, Z.; Keji, H., Daxua Xuebao , <u>1</u> , 177 (1986).

ANNEX

Mathematical derivation of A, B, and C

Using the notation exposed by Eqs. 5-8, it can be written:

$$S(A) = NA + BS(-4/3) + CS(2/3)$$
 Eq. A1

$$S(B) = NB + CS(2) + AS(4/3)$$
 Eq. A2

$$S(C) = NC + AS(-2/3) + BS(-2)$$
 Eq. A3

From Eqs. Al-A3, we derive A, B, and C:

$$A = (1/N)[S(A) - BS(-4/3) - CS(2/3)]$$
 Eq. A4

$$B = (1/N)[S(B) - CS(2) - AS(4/3)]$$
Eq. A5

$$C = (1/N)[S(C) - AS(-2/3) - BS(-2)]$$
 Eq. A6

Substituting Eq. A6 in Eq. A5, it comes:

$$B = (1/N^{2})[NS(B) - S(C)S(2) + A\{S(-2/3)S(2) - NS(4/3) + BS(-2)S(2)\}]$$

and

$$B = [1/(N^2 - S(2)S(-2))][NS(B) - S(C)S(2) + A\{S(-2/3)S(2) - NS(4/3)\}]$$
Eq. A7

Substituting Eq. A6 in Eq. A4, it comes also:

$$A = [1/(N^2 - S(2/3)S(-2/3))][NS(A) - S(C)S(2/3) - B\{NS(-4/3) - S(2/3)S(-2)\}]$$
Eq. A8

With Eqs A7 and A8, it can be derived:

$$A = (1/D) \{ [NS(A) - S(C)S(2/3)] [N^2 - S(2)S(-2)] - [NS(B) -S(C)S(2)] [NS(-4/3) - S(2/3)S(-2)] \}$$
Eq. A9

in which the denominator D is:

$$D = [N^{2} - S(2/3)S(-2/3)][N^{2} -S(2)S(-2)] + [S(-2/3)S(2) -NS(4/3)][NS(-4/3) - S(2/3)S(-2)]$$
Eq. A10

Similar derivations for B and C produce:

USE OF THE KNOX EQUATION. I

$$B = (1/D) \{ [NS(B) - S(A)S(4/3)] [N^2 - S(2/3)S(-2/3)] + [NS(C) - S(A)S(-2/3)] [S(2/3)S(4/3) - NS(2)] \}$$
Eq. All

$$C = (1/D) \{ [NS(C) - S(A)S(-2/3)] [N^2 - S(4/3)S(-4/3)] + \{ NS(B) - S(A)S(4/3)] [S(-2/3)S(-4/3) - NS(-2)] \}$$
Eq. A12

If one term is known, the calculus is very simplified. Say B is known; then, the h data set must be changed in h':

$$h' = h - B/\mathbf{v} = A\mathbf{v}^{1/3} + C\mathbf{v}$$

S(A) becomes S'(A):

$$S'(A) = \Sigma h_i' \underline{v}^{-1/3} = NA + S(+2/3)$$
 Eq. A13

similarly

$$S'(C) = \Sigma h_{1}/\Psi = NC + AS(-2/3)$$
 Eq. A14

Eqs A13 and A14 are easily combined to produce:

$$A = \{NS'(A) - S'(C)S(2/3)\}/\{N^2 - S(2/3)S(-2/3)\}$$
 Eq. A15

and

$$C = \{NS'(C) - S'(A)S(-2/3)\}/\{N^2 - S(2/3)S(-2/3)\} Eq. A16$$

We note that only four summations were needed: S'(A), S'(C), S(2/3) and S(-2/3) instead of nine in the general case. This decreases possible errors due to computer truncating.

If A or C is known, the two other terms can be easily obtained the same way.

Visual fit method, a macro-command for Lotus 123

Lotus 123 is controlled by a menu system and access to the main menu is achieved by pressing the "/" key. The menus offer a



Figure 7: Top: Flow chart of the macro-command to optimize the A term. "precedent" is the name of the cell in which the deviation before calculation is stored. The macro stops when the deviation increases. Bottom: actual lines written in the Lotus 123 language corresponding to the optimization of the A, B, and C terms.

USE OF THE KNOX EQUATION. I

series of options displayed across the top of the screen, each specified by a single word such as "print" or "file". Each option is selected by pressing the key corresponding to the first letter of the word describing the chosen option which either causes the command to be obeyed or displays the next level menu. Thus a sequence of keystrokes such as "/ppg" representing "print", "printer", "go" could be used to cause a printout of the worksheet.

Any of the cells in Lotus 123 can be given a label; a macrocommand is simply a sequence of keystrokes and/or keywords placed in one of these labeled worksheet cells. To activate a macro-command, simply press the ALT key and the appropriate label key and this causes all the keystrokes and keywords stored in the cell to be processed automatically (11).

Once the experimental h and \underline{v} values were entered in two columns, the Knox equation was entered in a third one, referring to three cells named A, B, and C. A fourth column will contain the squared deviation. A cell named "deviation" will contain the mean value of the deviation column (@AVG(first cell..last cell)). Another cell, named "precedent" will be prepared, it is needed by the program chart shown in Figure 7-top (for A search) and the Lotus macro-commands for the A, B, and C search (Figure 7-bottom).

To save time, A, B, and C values can be estimated by visual inspection. As the macros increase the studied term, underestimated A, B and C values must be present when invoking the macro. Drawing and data corresponding to Figures 4-6 were obtained with these macros. The computer must have an arithmetic coprocessor (8087 chip on IBM PCs or compatibles, 80287 chip on IBM ATs and compatibles, etc...) in order to recalculate (keyword {calc}, Figure 7) the sheet in a fraction of a second (instead of tenths of seconds or minutes).