Published in final edited form as:

J Chromatogr A. 2012 October 19; 1260: 215–223. doi:10.1016/j.chroma.2012.07.077.

Difference equation model for isothermal gas chromatography expresses retention behavior of homologues of *n*-alkanes excluding the influence of holdup time

Liejun Wu^a, Yongli Chen^b, Sarah A.L. Caccamise^{a,1}, and Qing X. Li^{a,*}

^aDepartment of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, HI 96822, USA

^bCollege of Natural and Computational Sciences, Hawaii Pacific University, 45-045 Kamehameha Highway, Kaneohe, HI 96744, USA

Abstract

A difference equation (DE) model is developed using the methylene retention increment (Δt_z) of *n*-alkanes to avoid the influence of gas holdup time (t_M). The effects of the equation orders (1st–5th) on the accuracy of a curve fitting show that a linear equation (LE) is less satisfactory and it is not necessary to use a complicated cubic or higher order equation. The relationship between the logarithm of Δt_z and the carbon number (z) of the *n*-alkanes under isothermal conditions closely follows the quadratic equation for C₃–C₃₀ *n*-alkanes at column temperatures of 24–260 °C. The first and second order forward differences of the expression ($\Delta \log \Delta t_z$ and $\Delta^2 \log \Delta t_z$, respectively) are linear and constant, respectively, which validates the DE model. This DE model lays a necessary foundation for further developing a retention model to accurately describe the relationship between the adjusted retention time and z of *n*-alkanes.

Keywords

Gas chromatography; Retention model; Retention time; Gas holdup time; *n*-Alkanes

1. Introduction

An isothermal retention model is a mathematical description of retention behavior of solutes during the separation process in isothermal gas chromatography (GC). Development of an accurate retention model is essential for identification of unknown compounds, prediction of retention times and calculations of chromatographic and thermodynamic parameters. It has been a fundamental and extensively studied topic [1–21]. However, a main challenge is that the gas holdup time (t_M) is hidden in the total retention time (t_{Rz}) [22,23].

$$t_{Rz} = t'_{Rz} + t_M \quad (1)$$

The adjusted retention time (t'_{Rz}) is an important retention parameter for GC theory and application, reflecting interactions between solutes and the stationary phase, while t_M is

independent of solutes. t_M depends on the volume from the point of injection to the point of detection in a GC system. The effect of t_M on a retention model should be eliminated completely. Methods used to determine t_M values in the literature include using air [24], inert gas [25], methane [26] and an iterative method using a least squares linear regression [27]; and all of those methods cannot separate t_M and t'_{Rz} accurately.

The linear relationship (i.e., LE model) between logarithm of t'_{Rz} and carbon number (z) of the homologous n -alkanes under an isothermal GC condition has been widely used since the 1950s [1,2,22,23,28–30].

$$\log(t_{Rz} - t_M) = bz + c \quad (2a)$$

or

$$\log t'_{Rz} = bz + c \quad (2b)$$

The slope of Eq. (2b) is a constant:

$$\Delta \log t'_{Rz} = b \quad (3)$$

where coefficients b and c are the slope and intercept, respectively, and the t_M is obtained via experimental or calculated methods. The t_M values vary largely upon the determination methods. Small variations (in an order of 2 s) in t_{Rz} could lead to large differences (up to 10 s) in t_M values, and incorrect t_M values have a remarkable influence on the retention index (I) values [3]. As the GC technology has been advanced to generate highly improved precision of retention times, many studies have found that the LE model is not a straight line and actually is a curve [3–14], especially for the low n -alkane homologues [17–21]. Therefore, many nonlinear equations have been proposed to describe the relationship between retention time and carbon number of n -alkanes, which includes the quadratic [5,10], cubic [5–7], degree 4 [7], degrees 5, 6, 7 polynomials [8], hyperbolic, and second order hyperbolic models [5]. González [9] proposed a nonlinear equation from statistical thermodynamics. Lebrón-Aguilar et al. [10,11] selected two nonlinear equations from 25 mathematical expressions and questioned the linear relationship based on a large set of data [12]. None of the above nonlinear equations, however, has been commonly recognized and the linear equation [Eq. (2a) or (2b)] is still commonly used.

The difference of retention times (Δt_z) between two consecutive n -alkanes or the methylene retention increments in a homologous series has been employed to calculate t'_{Rz} values in the literature [31,32]. This method has yet been commonly accepted for applications. The difference equation is a mathematical equality involving the differences between successive values of a function of a discrete variable. In the present study, we found that the dependence of logarithm of the Δt_z on z closely obeys a quadratic equation when the raw retention times are accurate and precise enough. Eq. (4) is a zero order difference equation, and thus, this model is called the difference equation (DE) model.

$$\log \Delta t_z = Az^2 + Bz + C \quad (4)$$

where Δt_z values can be calculated as the forward differences ($\Delta t_z = t_{Rz+1} - t_{Rz}$) or backward differences ($\nabla t_z = t_{Rz} - t_{Rz-1}$) when $\Delta z = 1$, A , B , and C are the regression

coefficients that depend on the experimental conditions. It is noteworthy that both forward difference and backward difference methods give the same results. In the present study, the forward difference method is used. Only Δt_z rather than t'_{Rz} , which is governed by the interactions between the stationary phase and *n*-alkanes, is involved in the DE model. For example, Δt_{10} is the difference of retention times of *n*-undecane and *n*-decane. The dependence of $\log \Delta t_z$ on *z* reflects the retention behavior of the *n*-alkanes excluding the influence of t_M . The first order forward differences of Eq. (4) as expressed by Eq. (5) follow a linear equation:

$$\Delta \log \Delta t_z = \beta z + \gamma \quad (5)$$

where $\beta = 2A$; $\gamma = A + B$. Similarly, the second order forward differences expressed by Eq. (6) are a constant (for details see supplementary information):

$$\Delta^2 \log \Delta t_z = 2A \quad (6)$$

The relationship between $\Delta \log \Delta t_z$, $\Delta^2 \log \Delta t_z$ and *z* confirms the retention behavior of the *n*-alkanes excluding the influence of t_M and thus, validates the DE model. Kováts's retention index (*I*) [23,29,30,33,34] is the most reliable method of expressing inter-laboratory chromatographic retention data. In order to compare the effect of equation order for regression of $\log \Delta t_z$ versus *z* on the accuracy of a curve fitting, the mean squares errors (MSE) of the differences (ΔI) between calculated (I_{cal}) and defined ($100z$) *I* values from LE [Eq. (7)], DE [Eq. (4)], and 3rd, 4th and 5th order polynomial [Eq. (8)] are also discussed.

$$\log \Delta t_z = Bz + C \quad (7)$$

$$\log \Delta t_z = o + pz + qz^2 + rz^3 + \dots \quad (8)$$

where coefficients B , C , o , p , q , and r are the regression coefficients of the corresponding Eqs. (7) and (8). The objective of the present work is to develop a retention model that reflects retention behavior of *n*-alkanes excluding the influence of t_M . This DE model lays a necessary foundation for further developing a retention model to accurately describe the relationship between t'_{Rz} and *z* of *n*-alkanes.

2. Experimental

2.1. Chemicals, instruments, and data collections

C₇–C₃₀ *n*-alkane standards were purchased from Sigma–Aldrich Co. (St. Louis, MO, USA). The standard mixtures were prepared at a concentration of 0.2 mg/mL of each *n*-alkane. Retention times were collected with a Hewlett–Packard (HP) 5890 II gas chromatograph equipped with a flame ionization detector at (1) a head pressure of 31 psi at isothermal 70, 100, 130, 140, 150, 160 and 170 °C and (2) a head pressure of 29 psi at isothermal 240, 250 and 260 °C. The column was a ZB-1 column with dimensions of 60 m × 0.25 mm × 0.25 μm (Phenomenex, Torrance, CA, USA). Helium was the carrier gas. The injector and detector temperatures were set at 290 °C. After GC conditions such as temperatures and head pressure were optimized, the GC conditions were stabilized at least 14 h and 4 h for the 31 psi and 29 psi head pressures, respectively. Samples (1.0 μL) were injected in 3–5

replicates using an HP 7673A autosampler in split mode at a ratio of 1:30. The retention times recorded had a precision of 0.001 min using an HP 3396A integrator. To verify the DE model and concept, data from the literature were also cited and used in the calculations.

2.2. Data treatment

The relationships between the Δt_z and z were regressed using the least squares method with linear, quadratic and polynomial fitting [Eqs. (4)–(8)] in Microsoft Excel. The logarithm of Δt_z , first order differences and second order differences are calculated using Eqs. (9)–(11), respectively.

$$\log \Delta t_z = \log(t_{Rz+1} - t_{Rz}) \quad (9)$$

$$\Delta \log \Delta t_z = \log \Delta t_{z+1} - \log \Delta t_z \quad (10)$$

$$\Delta^2 \log \Delta t_z = \Delta \log \Delta t_{z+1} - \Delta \log \Delta t_z \quad (11)$$

The reference points of retention index (I) can be calculated according to their inverse form of Eqs. (4), (7) and (8) [6,7]:

$$I_{\text{cal}} = [A'(\log \Delta t_z)^2 + B'(\log \Delta t_z) + C']100 \quad (12)$$

$$I_{\text{cal}} = [B'(\log \Delta t_z) + C']100$$

(13)

$$I_{\text{cal}} = [o' + p'(\log \Delta t_z) + q'(\log \Delta t_z)^2 + r'(\log \Delta t_z)^3 + \dots]100 \quad (14)$$

where coefficients A' , B' , C' , B' , C' , o' , p' , q' and r' are the regression coefficients of the corresponding Eqs. (12), (13) and (14). The differences (ΔI) between the calculated (I_{cal}) and defined values $100z$ [33,34] of retention indices are

$$\Delta I = I_{\text{cal}} - 100z \quad (15)$$

The DE, LE and polynomial models in the present work were compared by results of calculations with Eqs. (4), (7) and (8), respectively.

3. Results and discussion

3.1. Accuracy and precision of raw retention data

The t_{Rz} values for methane and C₇–C₃₀ *n*-alkanes on a ZB-1 column at a column temperature range of 70–260 °C were accurate at a confidence level of 95% (Table 1). The mean relative dispersion of the retention times at a head pressure of 31 psi and column temperature of 130 °C was $(1.04 \pm 0.21) \times 10^{-3}$ and was similar to that $(1.63 \pm 0.83) \times 10^{-3}$ reported by Lebrón-Aguilar et al. [35] (Table S1).

3.2. Difference equation (DE) model under isothermal GC and its validity

The quadratic equation was selected after regressions of the 1st, 2nd, 3rd, 4th and 5th order equations were statistically tested for curving fittings. The quadratic equation was then used as a zero order difference equation in the DE model that was systematically evaluated.

3.2.1. Effect of equation order (1st, 2nd, 3rd, 4th and 5th) for regression of $\log(\Delta t_z)$ versus z on accuracy of curve fitting—A good retention model should meet three basic requirements: (a) define the relationship between retention time and carbon number of n -alkanes excluding the influence of t_M , (b) have a high accuracy and precision, and (c) be simple.

The MSE of differences (ΔI) between calculated (I_{cal}) and defined ($100z$) reference points of retention indices were used as an indicator to compare the effect of the equation order on the accuracy of curve fitting. Fig. 1a shows that the MSE of the 1st order equation were greater than those of the 2nd to 5th order polynomial equations which were close to zero. The results were obtained with different numbers of n -alkanes evaluated (N), ranges of carbon, raw data sources, and different column temperatures (Table S2). The MSE values of the present study agreed well with those of published data [6,11,12]. It is noteworthy that the MSE (1.11) from the published data according to the 1st order equation [Eq. (13)] was slightly smaller than that from the present study (2.52), which is most likely due to differences in the n -alkanes evaluated between the present and previous studies (Table S2).

Fig. 1b shows the MSE values of the present study according to Eq. (13) (1st order, Wu1st) and Eq. (12) (2nd order, Wu2nd) in comparison with those (L.H.1st and L.H.2nd) of the published data [6,11,12]. The MSE values of the 1st order equation increased quickly with increasing the number of n -alkanes evaluated (N), following the equation of $MSE = 0.000012 \times N^{5.1189}$ (R^2 , 0.9995), whereas the MSE values of Wu2nd were close to zero. The MSE values calculated from the data in the literature [6,11] according to Eq. (13) (1st order) and Eq. (12) (2nd order) matched well correspondingly with those of the present study. The MSE values of “L.H.1st” were 0.4330, 0.8903, and 1.6329 when N was 8, 9 and 10, which were comparable with the corresponding Wu1st values of 0.4995, 0.8608 and 1.5115. The dependence of the MSE values of “L.H.1st” on N followed that of $MSE = 0.000002 \times N^{5.9518}$ (R^2 , 0.9997) although the regression was carried out with only 3 data points. The MSE values of L.H.2nd were close to zero.

The results of different ranges and number of n -alkanes evaluated under various column temperatures indicate that the 1st order equation can only be used in a narrow range of carbon number, whereas the 2nd order equation gives better accuracy of curve fitting for a much wider range of n -alkanes (Tables S3–S6). The 2nd order equation was selected instead of 3rd or higher order equations based on four reasons. (1) The decrease of MSE values between the 1st and 2nd order equations was 37–709-fold, which was much greater than 1.8–7-fold between the 2nd and 3rd order equations (Tables S3–S6). (2) The adjusted R^2 (R_{adj}^2) values of the 1st, 2nd and 3rd–5th order equations were approximately 0.9999, 0.9999993–0.9999999 and 0.9999999–1.0000000, respectively (Tables S3–S6). The increase of R_{adj}^2 between the 1st and 2nd were much greater than that between the 2nd and 3rd–5th order equations. (3) The residuals of curve fitting of $\log \Delta t_z$ versus z for the 1st, 2nd and 3rd order equations were examined (Fig. S1). The confidence interval ($\pm 2\hat{\sigma}$) of the residuals of the data determined in the present study were ± 0.0091 , ± 0.0005 and ± 0.0003 for the 1st, 2nd and 3rd order equations, respectively (Fig. S1A–C). The $\pm 2\hat{\sigma}$ values of the data cited from the literature were ± 0.0061 , ± 0.0021 and ± 0.0017 for the 1st, 2nd and 3rd order equations, respectively (Fig. S1D–F). The residuals varied regularly with the carbon number in a convex shape for the linear equation (Fig. S1A and D), which indicates that systematic

errors and it is necessary to use the 2nd order equation. The residuals randomly spread around zero, and most residuals fall within the $\pm 2\hat{\sigma}$ except a few data points for the 2nd and 3rd order equations (Fig. S1B, E and 1C, F). The results of the 2nd order regression were very similar to those of the 3rd order equation and much better than those of the 1st order equation. (4) Calculation simplicity of the 2nd order equation relative to the 3rd order equation significantly outweighs the little gain of curve fit-ting accuracy. Eq. (4) is, therefore, chosen as a zero order different equation of the DE model.

3.2.2. Validity of Eqs. (4)–(6) for curve fitting—Fig. 2a shows that the four curves of $\log t_{RZ}$ versus z on a ZB-1 column at 70, 100, 130 and 160 °C are very complicated and cannot be described with a simple equation. The four curves (bold) of $\log \Delta t_z$ versus z can be described well with the DE model [Eq. (4)] with a coefficient of determination (R^2) greater than 0.9999997.

Fig. 2b shows that all the quadratic fittings of $\log \Delta t_z$ versus z at 70–260 °C have high correlations. The coefficient A being a small negative value indicates that the curves were slightly convex. Table 2 shows that the lower was the column temperature, the more negative were the A values and the sharper was the convex curve. For example, the A values were -0.000715 , -0.000499 , -0.000439 and -0.000263 at 70, 140, 170 and 260 °C, respectively.

Fig. 2c shows that the $\Delta \log \Delta t_z$ decreased with increasing carbon number of n -alkanes, as shown in Eq. (5). For example, $\Delta \log \Delta t_7$ was 0.3140, while $\Delta \log \Delta t_{12}$ was 0.3068 at 70 °C. The slope $\Delta \log \Delta t_z$ is a good estimation of the thermodynamic free energy per methylene group partitioning between the mobile and stationary phases. The molecular interaction between the terminal methyl groups and stationary phases is stronger than that of the internal methylenes. The longer is the carbon chain, the larger is the methylene portion and the smaller becomes the difference of molecular interactions between the consecutive n -alkane and stationary phases. It coincides with the fact that the free energy of methyl is greater than methylene [13,14]. The mean $\Delta \log \Delta t_z$ values decreased as the column temperatures increased, which were 0.3104, 0.2068, 0.1749, and 0.1102 at 70, 140, 170, and 260 °C, respectively. These results agreed well with distribution constant (K), the capacity ratio (k) and the relative retention (r), which decreased with increasing column temperatures [13,14,23,29,30]. The slope equals $2A$ and the intercept equals coefficient A plus B . The experimental slope (β) and intercept (γ) obtained from Eq. (5) agreed very well with the calculated values ($2A$) and ($A+B$) obtained from Eq. (4). The slope of the plot for β (experimental versus calculated) and γ (experimental versus calculated) values at 70–260 °C were 0.9758 and 0.9982, respectively.

Fig. 2d shows that all $\Delta^2 \log \Delta t_z$ values fluctuated around the coefficient $2A$ at each column temperature. The data fluctuations resulted from random errors of the GC measurements. The slopes of the plots of the $\Delta^2 \log \Delta t_z$ versus z (C_7 – C_{30}) at 70–260 °C were calculated. To verify the validity of Eq. (6), the hypothesis $H_0: \mu$ (slope) = 0 was applied for statistical analysis (Table S7). The results show that 9 out of the 10 test statistic (T) values, except that at 130 °C, fell in their corresponding regions of acceptance [$t_{(1-\alpha/2)}(n-1)$]. The $\Delta^2 \log \Delta t_z$ values are, therefore, constant ($2A$).

The $\Delta^2 \log \Delta t_z$ values increased with increasing column temperature. The mean $\Delta^2 \log \Delta t_z$ values were -0.0014 , -0.0009 , -0.0008 , and -0.0004 at column temperatures of 70, 140, 170, and 260 °C, respectively. The results further indicate that the K values of the n -alkanes decreased with increasing column temperature. Table 2 shows that the experimental mean $\Delta^2 \log \Delta t_z$ values according to Eq. (11) coincided well with the calculated values ($2A$) obtained from coefficient A in Eq. (4). The slope of the plot for experimental versus

calculated $\Delta^2 \log \Delta t_z$ values at 70–260 °C was 0.9342 (Table 2). In summary, the relationship between retention time and carbon number can be quantitatively described with Eq. (4), which is validated by Eqs. (5) and (6).

Raw retention times published in the literature [11] were used to verify the DE model for a wide range of GC conditions including different column types, stationary phase polarities and column temperatures. The six curves of $\log \Delta t_z$ versus z on two packed columns (QF-1 and PS-225) and two capillary columns (CPSIL-5CB and HP-5) at 90, 120 and 150 °C follow Eq. (4) well with R^2 values greater than 0.999999. All coefficient A values (−0.00034 to −0.00069) being small negative values indicate that the curves are slightly convex (data not shown). The $\Delta \log \Delta t_z$ values were not constant, which decreased with increasing carbon number of n -alkanes. The mean $\Delta \log \Delta t_z$ values decreased as the column temperatures increased (Table S8). The mean $\Delta \log \Delta t_z$ values were 0.2509 and 0.2533 for columns CPSIL-5CB (100% dimethyl polysiloxane) and HP-5 (5% phenyl methylpolysiloxane) at 120 °C, which agreed well with the results of 0.2507 and 0.2541 for SE-30 (100% dimethyl polysiloxane) and SE-52 (5% phenyl methylpolysiloxane), respectively, reported by Rohrschneider [15]. All $\Delta^2 \log \Delta t_z$ values were slightly negative values (−0.0006 to −0.0015) and fluctuated around the coefficient $2A$ (data not shown).

The linear dependence of the $\Delta \log \Delta t_z$ on z agrees well with published data. Rohrschneider [16] determined that the logarithm of the ratio of vapor pressures (P_n/P_{n+1}) and activity coefficients (γ_n/γ_{n+1}) of n -alkanes versus the carbon number was not a constant, and it decreased from ethane to n -decane. Golovnya and Grigoryeva [13] calculated the slope of the partial molar free energy of sorption (ΔG_n) versus the carbon number of n -alkanes and showed that the energy contribution per CH_2 -group is not a constant, but decreased with increasing z on both polar and non-polar stationary phases. Tudor and Oncescu [14] measured the contributions of the methylene increment to the standard free energy at different column temperatures on SE-30 and Carbowax-20M capillary columns and found it decreased with increasing column temperatures and the carbon number of n -alkanes.

3.3. Range of n -alkanes for the DE model

It is well known that the first few members of the n -alkanes deviate from the linearity of the traditional LE model. Different starting points of the linearity range reported in the literature began at ethane [17], propane [18,19], n -butane and greater [20], n -hexane [23], n -heptane [29], n -decane [16]. Heeg et al. [6,7] and Lebron-Aguilar et al. [10–12,35] found that the semilog regressions of the retention times of n -alkanes vs the carbon number were not strictly linear within the whole range. We found that the n -alkanes with $z \geq 3$ display a regular retention characteristic of the homologue, following the DE model.

3.3.1. Low homologue of C_3 – C_{12} n -alkanes also follows the DE model—The working range of the DE model starts at n -alkanes with $z \geq 3$. Eqs. (4)–(6) can quantitatively describe the relationship between retention time and carbon number of n -alkanes using published data [11,21] starting at propane. The relationships of $\log \Delta t_z$ versus z followed Eq. (4) with R^2 values of 0.9999976–0.9999999 when the Δt_z values were obtained from $t_{R,z}$ of C_3 – C_{12} n -alkanes on three different columns at 24, 50, 60, and 90 °C in publications (Table S9). The coefficient A values were all negative and increased with increasing column temperatures. For example, the coefficient A value of −0.00174 at 24 °C increased to −0.00065 at 90 °C. Regression of $\Delta \log \Delta t_z$ versus z was linear [Eq. (5)] and decreased with increasing z at a constant column temperature when $z \geq 3$. The mean of the $\Delta \log \Delta t_z$ also decreased with increasing column temperature (e.g., 0.4379 at 24 °C and 0.2994 at 90 °C) (Table S9). The slope of the plot for β (experimental versus calculated) and γ (experimental versus calculated) values at 24–90 °C were 0.9622 and 1.0168, respectively. All mean $\Delta^2 \log$

Δt_z values were negative constants and gradually increased with increasing column temperature. The slope of the plot for the mean experimental *versus* calculated $\Delta^2 \log \Delta t_z$ values was 0.8910 (Table S9).

3.3.2. Methane and ethane do not obey the DE model—The R_{adj}^2 values of quadratic fitting according to Eq. (4) are 0.9999998 for the ranges of C₃–C₁₀, C₄–C₁₀, and C₅–C₁₀ at 60 °C. However, the R_{adj}^2 values suddenly dropped to 0.9995 and 0.9997 for the ranges of C₁–C₁₀ and C₂–C₁₀, respectively (Table 3). Similar results were also obtained for C₁–C₁₂ *n*-alkanes at 90 °C. The results indicate that the quadratic regression line perfectly fits the data when methane and/or ethane are excluded. Fig. 3a shows that the $\Delta t_{z+1}/\Delta t_z$ data points of methane and ethane at four column temperatures apparently deviated from the line of $z = 3$. Fig. 3b shows an excellent quadratic fit between the boiling (bp) [36] and z when $z = 3$, and melting points (mp) [37] of methane and ethane also deviated from the regular property patterns. Fig. 3c shows that the curve of the ratio of bp differences ($\Delta bp_{z+1}/\Delta bp_z$) *versus* z is similar to the curve of the $\Delta t_{z+1}/\Delta t_z$ *versus* z shown in Fig. 3a. The first two $\Delta bp_{z+1}/\Delta bp_z$ values deviated from the line of $z = 3$. This phenomenon can be explained by the different interaction between the molecules in the vapor state due to the difference of the molecular composition of *n*-alkanes. The C–C bond is longer than a C–H bond, and the molecular interaction between the terminal methyl groups and stationary phases is stronger than that of the internal methylenes in the vapor state. Methane contains four equal C–H bonds. The bp of methane is much lower than the calculated value according to the relationship between bp and the carbon number of *n*-alkanes with $z = 3$. Ethane composes of two methyl groups and has the strongest intermolecular interactions among *n*-alkanes in the vapor state. Fig. 3d shows further that the experimental mp differences of methane and ethane deviated from the regression mp of *n*-alkanes with $z = 3$. This can be well explained by the molecular packing of *n*-alkanes. In crystalline solid state, methane has the most effective packing; therefore, the mp is much higher than the theoretical value and even a little higher than that of ethane. The symmetry of the ethane molecule results in a tighter packing than that of propane and, therefore, a higher mp. All *n*-alkanes of $z = 3$ contain C–C bonds, C–H bonds, and methyl and methylene groups, and the difference between the *n*-alkanes lies only in the number of methylene groups. Therefore, the *n*-alkanes of $z = 3$ display regular property patterns.

3.4. Relationship between Δt_z and column temperature (T_C)

The relationship between K , k and r *versus* column temperature (T_C) can be described with Antoine equations [23]:

$$\log (K, k, \text{ or } r) = \frac{a}{T_C} + b \quad (16)$$

where the T_C is the absolute column temperature. In a relatively narrow temperature range, the above relationship can be approximated using a linear equation [23]:

$$\log (K, k, \text{ or } r) = a' T_C + b' \quad (17)$$

We found that the $\log \Delta t_z$ *versus* T_C can be well described with a quadratic equation in a wide temperature range and the $\log (K, k, r)$ *versus* T_C as well when the t'_{Rz} and t_M values are correct, which is discussed in details in the companion paper [38].

$$\log(\Delta t_z, K, k, \text{ or } r) = A' T_C^2 + B' T_C + C' \quad (18)$$

where the T_C is the absolute or centigrade column temperature. Fig. 4a and b show the dependence of Δt_z , K on T_C . The R^2 values are greater than 0.9999. The raw data are listed in Table 1. Fig. 4c and d shows similar results from raw data reported by Quintanilla-López et al. [11] and González and Gagliardi [39], respectively. Eq. (18) is useful for applications although it is empirical. The accuracy indicating the dependence of retention parameters on T_C is better than that of Eqs. (16) and (17). Eq. (18) is an easier calculation than the conventional thermodynamics equation [40,41].

4. Conclusions

The DE model reveals the relationship between retention parameter and carbon number excluding the influence of t_M using the difference of retention times (Δt_z). This difference method based on accurate raw retention times is a logical method because the relationship between retention time and carbon number can be quantitatively described at three levels [Eqs. (4)–(6)]. The validity of the DE model has been verified by vigorous statistical analysis and curve fitting of experimental data collected in the present study and those cited from the literature. The DE model applies well for n -alkanes larger than propane, except methane and ethane. The DE model is reliable and simple.

Under the same condition (excluding the influence of t_M), the MSE of LE [Eq. (7)] were larger than those of the 2nd to 5th order polynomial and increased quickly with increasing the number of n -alkanes evaluated. The LE model is less accurate than DE, whereas it is not necessary to use a complicated polynomial with 3rd or higher order. Therefore, the DE model is an optimal choice and lays a foundation to further develop retention models in real GC systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by grants from the U.S. Fish and Wildlife Service, US ONR HEET (N00014-09-1-0709), the National Center for Research Resources (5G12 RR003061-26; 5P20 RR016467-11), the National Institute on Minority Health and Health Disparities (8G12 MD007601-26) and the National Institute of General Medical Sciences (8P20GM103466-11) from the National Institutes of Health. We thank Dr. Ning-Shou Xu and Professor Maolin Li for helpful discussions and Mrs. Margret Baker for the literature search. We also thank Dr. R. Lebrón-Aguilar for communications and data confirmation.

References

1. Poole CF, Poole SK. *J Chromatogr A*. 2009; 1216:1530. [PubMed: 19013576]
2. Peng CT. *J Chromatogr A*. 2010; 1217:3683. [PubMed: 20227699]
3. Haken JK, Wainwright MS, Smith RJ. *J Chromatogr*. 1977; 133:1.
4. Reido F, Fritz D, Farjan G, Kováts Esz. *J Chromatogr*. 1976; 126:63.
5. Touabet A, Maeck M, Badjah H, Ahmed AY, Meklati BY. *Chromatographia*. 1988; 25:389.
6. Heeg FJ, Zinburg R, Neu HJ, Ballschmiter K. *Chromatographia*. 1979; 12:451.
7. Ballschmiter K, Heeg FJ, Neu HJ, Zinburg R. *Z Fresenius Anal Chem*. 1985; 321:426.
8. Smith RJ, Haken JK, Wainwright MS, Madden BG. *J Chromatogr*. 1985; 328:11.
9. González FR. *J Chromatogr A*. 1999; 832:165.
10. Lebrón-Aguilar R, Quintanilla-López JE, García-Domínguez JA. *J Chromatogr A*. 1997; 760:219.

11. Quintanilla-López JE, Lebrón-Aguilar R, Garcia-Dominguez JA. *J Chromatogr A*. 1997; 767:127.
12. Garcia-Dominguez JA, Quintanilla-López JE, Lebrón-Aguilar R. *J Chromatogr A*. 1998; 803:197.
13. Golovnya RV, Grigoryeva DN. *Chromatographia*. 1983; 17:613.
14. Tudor E, Oncescu T. *J Chromatogr A*. 1999; 844:201.
15. Rohrschneider L. *Chromatographia*. 1993; 37:250.
16. Rohrschneider L. *Chromatographia*. 1969; 2:437.
17. Parcher JF, Johnson DM. *J Chromatogr Sci*. 1980; 18:267.
18. Ettre LS. *Chromatographia*. 1980; 13:73.
19. Becerra MR, Fernandez-Sanchez E, Garcia-Dominguez JA, Garcia-Munoz J, Molera MJ. *J Chromatogr Sci*. 1982; 20:363.
20. Wainwright MS, Haken JK, Srisukh D. *J Chromatogr*. 1980; 188:246. Wainwright MS, Haken JK, Srisukh D. *J Chromatogr*. 1980; 198:156.
21. Haken JK, Wainwright MS, Srisukh D. *J Chromatogr*. 1983; 257:107.
22. Kaiser R. *Chromatographia*. 1969; 2:125.
23. Ettre, LS.; Hinshaw, JV. *Basic Relationships of Gas Chromatography*. Advanstar Communications Inc; Cleveland, OH: 1993.
24. Littlewood AB, Phillips CSG, Price DT. *J Chem Sci*. 1955:1480.
25. Ambrose D, Keulemans AIM, Purnell JH. *Anal Chem*. 1958; 30:1582.
26. McReynolds WO. *J Chromatogr Sci*. 1970; 8:685.
27. Guardino X, Albaigés J, Firpo G, Rodríguez-Vināls R, Gassiot M. *J Chromatogr*. 1976; 118:13.
28. Conder, JR.; Young, CL. *Physicochemical Measurements by Gas Chromatography*. John Wiley & Sons; New York: 1979.
29. Budahegyi MV, Lombosi ER, Lombosi TS, Mészáros SY, Nyiredy S, Tarján G, Timár G, Takács JM. *J Chromatogr*. 1983; 271:213.
30. Pacakova, V.; Feltl, L. *An Aid to Identification of Organic Compounds*. Ellis Horwood; Chichester, U.K: 1992. *Chromatographic Retention Indices*.
31. Sevcík J. *J Chromatogr A*. 1977; 135:183.
32. Sevcík J, Löwentap MSH. *J Chromatogr A*. 1978; 147:75.
33. Kováts, ESz. *Helv Chim Acta*. 1958; 41:1915.
34. Ettre LS. *Anal Chem*. 1964; 36:31A.
35. Lebrón-Aguilar R, Quintanilla-López JE, Garcia-Dominguez JA. *J Chromatogr A*. 2002; 945:185. [PubMed: 11860135]
36. Cao C, Yuan H. *J Chem Inf Comput Sci*. 2001; 41:867. [PubMed: 11500103]
37. Loudon, GM. *Organic Chemistry*. 4. Oxford University Press; New York: 2002. p. 45
38. Wu L, Chen M, Chen Y, Li QX. *J Chromatogr. A* submitted for publication.
39. González FR, Gagliardi LG. *J Chromatogr A*. 2000; 879:157. [PubMed: 10893032]
40. Castells RC, Arancibia EL, Nardillo AM. *J Chromatogr*. 1990; 504:45.
41. González FR. *J Chromatogr A*. 2002; 942:211. [PubMed: 11822386]

Appendix A Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.chroma.2012.07.077>.

Fig. 1.

Effects of equation orders (a) and number of n -alkanes evaluated (b) on the MSE of the differences (ΔI) between calculated (I_{cal}) and defined (100 z) reference points of retention indices. $\text{MSE} = (\sum(\Delta I)^2)/(\text{number of } n\text{-alkanes evaluated})$. I_{cal} (2nd order), I_{cal} (1st order), I_{cal} (3rd order) and ΔI were calculated with Eqs. (12)–(15), respectively. (a). The MSE values of “Wu” were calculated with mean values of the present study on the ZB-1 column at 70, 100, 130, 160 and 240 °C (raw data in Table 1). The MSE values of “L.H.” were calculated with mean values from references 11, 12 (L.) and 6 (H.) on the columns CPSIL-5CB and HP-5 [11], HP-5 [12] and SE-30 [6] at 60, 90, 120, 150 and 160 °C. (b) The MSE values of “Wu1st” and “Wu2nd” were calculated with mean values of the present study on the ZB-1 column at 130, 160, 240, 250 and 260 °C, but according to Eq. (13) (1st order) and Eq. (12) (2nd order), respectively. The raw retention times are listed in Table 1. The MSE values of “L.H.1st” and “L.H.2nd” were calculated with mean values from references [11] (L.) and [6] (H.), but according to Eq. (13) (1st order) and Eq. (12) (2nd order), respectively.

Fig. 2.

Relationships between different model parameters and carbon number at different column temperatures on the ZB-1 column: (a) nonlinear scatter graphs of $\log t_{Rz}$ and $\log \Delta t_z$ versus z at 70, 100, 130 and 160 °C, (b) quadratic fitting curves of $\log \Delta t_z$ versus z , (c) regression of $\Delta \log \Delta t_z$ versus z and (d) plots of $\Delta^2 \log \Delta t_z$ versus z at 70, 100, 130, 140, 150, 160, 170, 240, 250 and 260 °C. Raw retention times are listed in Table 1. The retention times of C_{27} and C_{29} n -alkanes at 240, 250, and 260 °C were calculated from interpolation according to data in Table 1.

Fig. 3.

Relationships (a) between $\Delta t_{z+1}/\Delta t_z$ and carbon number z , (b) between the boiling point (bp), melting point (mp) and z , (c) between $\Delta bp_{z+1}/\Delta bp_z$ and z , (d) between Δmp_z of experimental (exp.) mp and regression (reg.) mp and z . Δbp_z and Δmp_z are the forward differences of boiling and melting points of two consecutive n -alkanes, respectively, i.e., $\Delta bp_z = bp_{z+1} - bp_z$ and $\Delta mp_z = mp_{z+1} - mp_z$. Regression (reg.) mp values of even and odd carbon n -alkanes were separately calculated with quadratic equation regression. Raw retention times in Fig. 3a are listed in Table S9; data of boiling points and melting points are cited from [36] and [37].

Fig. 4.

Dependence of log Δt_z and log distribution constant (K) on column temperature (T_c). (a, b) C_9 – C_{14} n -alkanes at 70–170 °C on the ZB-1 column, K values were calculated using t_{MT} value as discussed in the companion paper [38]. Raw data see Table 1; (c) C_5 – C_9 n -alkanes at 60–150 °C on CPSIL-5CB column, raw data from [11]; (d) C_{10} – C_{18} n -alkanes at 80–200 °C on SPB-1, raw data from [39].

Table 1

Total retention times ($t_{R,Z}$, min)^a and the confidence interval^b of *n*-alkanes on a ZB-1 column at different column temperatures.

z^c	70 °C	100 °C	130 °C	140 °C	150 °C	160 °C	170 °C
1	3.705 ± 0.001	3.878 ± 0.001	4.062 ± 0.002	4.101 ± 0.001	4.147 ± 0.001	4.209 ± 0.001	4.262 ± 0.004
7	5.078 ± 0.002	4.509 ± 0.004					
8	6.601 ± 0.002	5.087 ± 0.004	4.654 ± 0.006	4.584 ± 0.002	4.543 ± 0.002		4.699 ± 0.006
9	9.740 ± 0.003	6.149 ± 0.006	5.092 ± 0.006	4.922 ± 0.002	4.805 ± 0.002	4.744 ± 0.001	
10	16.188 ± 0.008	8.100 ± 0.008	5.824 ± 0.007	5.470 ± 0.002	5.222 ± 0.002	5.066 ± 0.002	4.949 ± 0.006
11	29.376 ± 0.018	11.668 ± 0.013	7.045 ± 0.008	6.362 ± 0.003	5.881 ± 0.003	5.564 ± 0.001	5.328 ± 0.007
12	56.299 ± 0.036	18.182 ± 0.019	9.080 ± 0.010	7.807 ± 0.003	6.924 ± 0.003	6.333 ± 0.002	5.898 ± 0.008
13	111.043 ± 0.077	30.033 ± 0.035	12.461 ± 0.014	10.147 ± 0.004	8.569 ± 0.003	7.517 ± 0.002	6.758 ± 0.009
14	222.005 ± 0.114	51.557 ± 0.058	18.067 ± 0.019	13.923 ± 0.006	11.161 ± 0.004	9.340 ± 0.002	8.052 ± 0.011
15		90.486 ± 0.104	27.336 ± 0.029	20.007 ± 0.010	15.232 ± 0.006	12.137 ± 0.003	9.994 ± 0.014
16		160.720 ± 0.302	42.631 ± 0.050	29.783 ± 0.019	21.614 ± 0.005	16.422 ± 0.003	12.903 ± 0.018
17			67.785 ± 0.065	45.453 ± 0.030	31.596 ± 0.013	22.969 ± 0.004	17.254 ± 0.028
18		109.071 ± 0.102	109.071 ± 0.102	70.515 ± 0.056	47.172 ± 0.009	32.959 ± 0.006	23.747 ± 0.035
19			176.587 ± 0.144	110.467 ± 0.100	71.422 ± 0.009	48.157 ± 0.008	33.416 ± 0.050
20				174.058 ± 0.192	109.049 ± 0.025	71.225 ± 0.016	47.775 ± 0.070
21							69.031 ± 0.108
22							100.463 ± 0.161

	240 °C	250 °C	260 °C
14	5.677 ± 0.007	5.551 ± 0.020	5.497 ± 0.019
15	6.004 ± 0.009	5.807 ± 0.021	5.703 ± 0.019
16	6.442 ± 0.006	6.145 ± 0.023	5.970 ± 0.022
17	7.027 ± 0.007	6.590 ± 0.019	6.316 ± 0.021
18	7.808 ± 0.009	7.176 ± 0.018	6.764 ± 0.023
19	8.850 ± 0.011	7.944 ± 0.017	7.346 ± 0.027
20	10.236 ± 0.009	8.953 ± 0.017	8.100 ± 0.031
21	12.079 ± 0.011	10.277 ± 0.015	9.074 ± 0.036
22	14.525 ± 0.011	12.008 ± 0.014	10.333 ± 0.041
23	17.766 ± 0.009	14.271 ± 0.013	11.956 ± 0.050

	240 °C	250 °C	260 °C
24	22.052 ± 0.013	17.223 ± 0.012	14.047 ± 0.059
25	27.715 ± 0.010	21.070 ± 0.015	16.739 ± 0.071
26	35.183 ± 0.014	26.076 ± 0.018	20.195 ± 0.081
28	57.943 ± 0.051	41.022 ± 0.044	30.296 ± 0.103
30	97.155 ± 0.074	66.084 ± 0.094	46.840 ± 0.122

^aThe R_Z values were the means of five and three replicate GC runs at 70–170 °C and 240–260 °C, respectively.

^bThe confidence interval of 95% was calculated by [(Student factor (t) × standard deviation (SD))/(the squared root of number of replicates (n)), i.e. $tSD/\sqrt{n}$].

^c Z_c , carbon number of n -alkanes.

Table 2

Regression parameters (A , B , C , B' , C') and coefficients of determination (R^2) for LE [Eq. (7)] and DE [Eq. (4)] regressions of $\log \Delta t_z$ versus z and comparison of regression and calculated parameters from $\Delta \log \Delta t_z$ and $\Delta^2 \log \Delta t_z$ values.^a

Column temperature (°C)	70	100	130	140	150	160	170	240	250	260
Range of <i>n</i> -alkanes	C ₇ –C ₁₄	C ₇ –C ₁₆	C ₈ –C ₁₉	C ₈ –C ₂₀	C ₈ –C ₂₀	C ₉ –C ₂₀	C ₉ –C ₂₂	C ₁₄ –C ₃₀	C ₁₄ –C ₃₀	C ₁₄ –C ₃₀
log Δ <i>t_z</i> vs. <i>z</i> regression with LE [Eq. (7)]										
<i>B'</i>	0.3104	0.2606	0.2189	0.2069	0.1961	0.1856	0.1749	0.1222	0.1162	0.1104
<i>C'</i>	–1.9865	–2.0571	–2.1022	–2.1186	–2.1413	–2.1557	–2.1670	–2.1846	–2.2075	–2.2237
<i>R</i> ²	0.99998	0.99997	0.99995	0.99995	0.99995	0.99995	0.99993	0.99988	0.99988	0.99990
Mean <i>R</i> ²	0.99993									
log Δ <i>t_z</i> vs. <i>z</i> regression with DE [Eq. (4)]										
<i>A</i>	–0.000715	–0.000599	–0.000531	–0.000499	–0.000472	–0.000452	–0.000439	–0.000333	–0.000307	–0.000263
<i>B</i>	0.3247	0.2738	0.2327	0.2204	0.2088	0.1982	0.1882	0.1365	0.1293	0.1217
<i>C</i>	–2.0551	–2.1256	–2.1866	–2.2035	–2.2217	–2.2399	–2.2597	–2.3313	–2.3426	–2.3398
<i>R</i> ²	1.000000	0.999999	0.999999	0.999999	0.999998	0.999999	0.999996	1.000000	0.999998	0.999983
Mean <i>R</i> ²	0.9999997									
Δlog Δ <i>t_z</i> vs. <i>z</i> regression with LE [Eq. (5)]										
Mean Δlog Δ <i>t_z</i>	0.3104	0.2606	0.2188	0.2068	0.1960	0.1855	0.1749	0.1221	0.1160	0.1102
<i>β</i> ^{exp} , <i>b</i>	–0.00144	–0.00119	–0.00106	–0.00098	–0.00094	–0.00090	–0.00086	–0.00066	–0.00062	–0.00049
<i>β</i> ^{cal} , (=2 <i>A</i>)	–0.00143	–0.00120	–0.00106	–0.00100	–0.00094	–0.00090	–0.00088	–0.00067	–0.00061	–0.00053
<i>Slope</i> ^{<i>c</i>}	0.9758									
<i>γ</i> ^{exp} , <i>b</i>	0.3241	0.2731	0.2321	0.2196	0.2083	0.1977	0.1874	0.1361	0.1291	0.1205
<i>γ</i> ^{cal} , (= <i>A</i> + <i>B</i>)	0.32398	0.27321	0.23216	0.21988	0.20837	0.19778	0.18771	0.13616	0.12904	0.12142
<i>Slope</i> ^{<i>c</i>}	0.9982									
Δ ² log Δ <i>t_z</i>										
exp. ^{<i>d</i>}	–0.00144	–0.00115	–0.00106	–0.00093	–0.00096	–0.00089	–0.00080	–0.00067	–0.00063	–0.00041
cal. (=2 <i>A</i>) ^{<i>d</i>}	–0.00143	–0.00120	–0.00106	–0.00100	–0.00094	–0.00090	–0.00088	–0.00067	–0.00061	–0.00053
<i>Slope</i> ^{<i>c</i>}	0.9342									

^aRaw data are from Table 1. Total retention times of C_{27} and C_{29} at 240, 250, and 260 °C were calculated by interpolation according to data in Table 1.

^b Experimental values (exp.) from Eq. (5) and calculated values (cal.) obtained from coefficient A and B values in Eq. (4).

^c The slopes of the plots of the experimental values *versus* the calculated values.

^d Experimental mean values (exp.) from Eq. (11) and calculated values (cal.) obtained from coefficient A value in Eq. (4).

Table 3

Comparison of the adjusted coefficient of determination (R_{adj}^2)^a values for different ranges of *n*-alkanes at 60 °C and 90 °C on CPSIL-5CB column using Eq. (4) curve fitting.^b

Column temperature	Range of <i>n</i> -alkanes	Coefficients of Eq. (4)			R_{adj}^2
		A	B	C	
60 °C	C ₁ –C ₁₀	–0.001476	0.377176	–0.165446	0.99946
	C ₂ –C ₁₀	–0.00328	0.399175	–0.224951	0.99968
	C ₃ –C ₁₀	–0.000964	0.369066	–0.134623	0.9999998
	C ₄ –C ₁₀	–0.000935	0.368670	–0.133326	0.9999998
	C ₅ –C ₁₀	–0.000905	0.368231	–0.131777	0.9999998
90 °C	C ₁ –C ₁₂	–0.001019	0.31407	–0.220855	0.99962
	C ₂ –C ₁₂	–0.001925	0.32730	–0.263261	0.99976
	C ₃ –C ₁₂	–0.000621	0.307215	–0.192836	0.9999984
	C ₄ –C ₁₂	–0.000715	0.308737	–0.198604	0.9999996
	C ₅ –C ₁₂	–0.000771	0.309696	–0.202498	0.9999994

^a $R_{adj}^2 = 1 - [(n-1)/(n-p-1)](1-R^2)$, where *n* is the number of observations; *p* is the number of independent variables used for regression: e.g., 1, 2, 3, 4 and 5 for 1st, 2nd, 3rd, 4th and 5th order regression, respectively.

^b Raw retention times are cited from Ref. [11] and listed in Table S9.