Published in final edited form as:

J Chromatogr A. 2012 October 19; 1260: 224–231. doi:10.1016/j.chroma.2012.08.071.

A new accurate quadratic equation model for isothermal gas chromatography and its comparison with the linear model

Liejun Wu^a, Maoxue Chen^b, Yongli Chen^c, and Qing X. Li^{a,*}

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

^b College of Information Science and Engineering, Shandong Agricultural University, Taian 271018, China

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

Abstract

The gas holdup time (t_M) is a dominant parameter in gas chromatographic retention models. The difference equation (DE) model proposed by Wu et al. (*J. Chromatogr. A* 2012, <http://dx.doi.org/10.1016/j.chroma.2012.07.077>) excluded t_M . In the present paper, we propose that the relationship between the adjusted retention time t'_{RZ} and carbon number z of n -alkanes follows a quadratic equation (QE) when an accurate t_M is obtained. This QE model is the same as or better than the DE model for an accurate expression of the retention behavior of n -alkanes and model applications. The QE model covers a larger range of n -alkanes with better curve fittings than the linear model. The accuracy of the QE model was approximately 2–6 times better than the DE model and 18–540 times better than the LE model. Standard deviations of the QE model were approximately 2–3 times smaller than those of the DE model.

Keywords

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

1. Introduction

The difference equation (DE) model that expresses the relationship between methylene retention increment with z carbon number (Δt_z) and the carbon number (z) of n -alkanes has been proposed by Wu et al. [1].

$$\log \Delta t_z = Az^2 + Bz + C \quad (1)$$

The slope of Eq. (1) is a linear equation (LE), and the slope of LE is a constant:

© 2012 Elsevier B.V. All rights reserved.

* Corresponding author at: Department of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, 1955 East-West Road, HI 96822, USA. Tel.: +1 808 956 2011; fax: +1 808 956 3542. qingl@hawaii.edu (Q.X. Li).

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.2012.08.071>.

$$\Delta \log \Delta t_z = \log \left(\frac{\Delta t_{z+1}}{\Delta t_z} \right) = \beta z + \gamma \quad (2)$$

$$\Delta^2 \log \Delta t_z = 2A = \text{constant} \quad (3)$$

where $\Delta t_z = t_{Rz+1} - t_{Rz}$ are the forward differences of retention times between two consecutive n -alkanes and A, B, C, β and γ are the regression coefficients that depend upon the experimental conditions.

The DE model excludes the influence of gas holdup time (t_M) and expresses the accurate retention behavior of homolog of n -alkanes. The adjusted retention time of an n -alkane with z carbon number (t'_{Rz}) is the total retention time (t_{Rz}) minus the t_M value. Therefore, if the t_M with a high degree of accuracy is obtained, the relationship between $\log t'_{Rz}$ and z of n -alkanes should also follow a quadratic equation (QE), which is called a QE model, and the adjusted retention time of an n -alkane with z carbon number is expressed as t'_{RT} hereafter.

$$\log t'_{RT} = Uz^2 + Vz + W \quad (4)$$

Being the same as Eq. (1), the slope of Eq. (4) is a LE, and the slope of LE is a constant:

$$\Delta \log t'_{RT} = \beta' z + \gamma' \quad (5)$$

$$\Delta^2 \log t'_{RT} = 2U = \text{constant} \quad (6)$$

where $t'_{RT} = t_{Rz} - t_{MT}$, and t_{MT} is an accurate gas holdup time which can be calculated using multiple linear regression analysis from the present work, and U, V, W, β' and γ' are regression coefficients which depend upon the experimental conditions.

For comparison with the DE and QE models, the linear Eqs. (7) and (9) and their corresponding slopes (8) and (10) are shown below:

$$\log \Delta t_z = z + \quad (7)$$

$$\Delta \log \Delta t_z = \text{constant} \quad (8)$$

$$\log t'_{Rz} = bz + c \quad (9)$$

$$\Delta \log t'_{Rz} = b = \text{constant} \quad (10)$$

where $t'_{Rz} = t_{Rz} - t_M$, B, C, b and c are regression coefficients which depend upon the experimental conditions. Eq. (9) (hereafter, LE model) under an isothermal GC condition has been used widely [2–7]. t_M is obtained via two types of common methods: one using air, methane (t_{MCH_4}) or neon (t_{MNe}) as a marker to measure t_M experimentally; the other using

calculation methods, such as the LE iterative method [8], which is called t_{MM} . Because the dependence of t'_{Rz} on z is not linear, many nonlinear models have been proposed [9–18].

It has been shown that the accuracy of curve fittings of LE was worse than that of the DE model that excludes t_M [1]. In the present work, we compare the QE model [Eq. (4)] with the DE [Eq. (1)] and LE models [Eqs. (7) and (9)]. Discussions also include the relationship between the development of retention models and the advances of the GC technology, the ratio of the differences of retention times between successive nalkanes ($R\Delta t_z$) and lack of a general acceptance of the nonlinear models.

2. Experimental

2.1. Chemicals, instruments and data collection

Chemicals, instruments and data collection were the same as previously described [1].

2.2. Mathematical treatment

The QE model was considered as the results from computational simulation. Computational simulation is a computer program that attempts to simulate an abstract model of a particular system. The DE model was considered as an abstract model, then using the multiple linear regression (SAS Institute Inc., Cary, NC, USA) to adjust the t_M value for curve fittings of Eq. (4). The range of the iteration and step length were set in SAS. After iteration, when the maximum correlation (R^2) was reached, the t_{MT} value, the holdup time with a high degree of accuracy in Eq. (4) and the QE model equation was acquired.

Reference points of retention indexes (I) were calculated according to the respective inverse forms [17,18] of Eqs. (1), (4), (7) and (9):

$$I_{cal} = \left[A' (\log \Delta t_z)^2 + B' (\log \Delta t_z) + C' \right] / 100 \quad (11)$$

$$I_{cal} = \left[U' (\log t'_{RT})^2 + V' (\log t'_{RT}) + W' \right] / 100 \quad (12)$$

$$I_{cal} = \left[(\log \Delta t_z) + =' \right] / 100 \quad (13)$$

$$I_{cal} = \left[b' (\log t'_R)^2 + c' \right] / 100 \quad (14)$$

The difference (ΔI) between the calculated and defined values [19,20] of reference points of I values is

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

The t_{MM} values were calculated by the same method as t_{MT} for curve fittings using Eq. (9). The t_{MCH_4} and t_{MNe} values were determined experimentally. The other data and statistical analysis were previously described [1].

3. Results and discussion

3.1. The QE and DE models

The DE and QE models that accurately describe dependence of retention times on carbon number of *n*-alkanes can be statistically and experimentally verified as discussed below.

3.1.1. Three coefficients of A, B and C in Eq. (1) and of U, V and W in Eq. (4)—

Hypothesis test uses statistics to determine the probability that a given hypothesis is true. It is also called confirmatory data analysis. In the present study, it was used to verify that the dependence of retention times on carbon number is the same for both models [Eqs. (1) and (4)].

The three coefficients (*A*, *B*, *C*) in the DE and (*U*, *V*, *W*) in the QE model on the ZB-1 column were served as three treatments (*t*) of the factor being tested (Table S1). The data were obtained at 10 temperatures (70–260 °C) as 10 random blocks (*r*). Both retention times (Δt_z and t'_{RT}) are control variables and all 60 coefficients (*A*, *B*, *C*, *U*, *V* and *W* at 10 temperatures) are observed variables or dependent variables. The overall one-way ANOVA was used to test the null hypothesis that all pair-observed variables of treatment groups have identical mean values, i.e., pair difference is equal to zero ($A - U = 0$, $B - V = 0$ and $C - W = 0$) in the entire experimental range. The ANOVA results of $F = 24.24$ and $P = < 0.0001$ indicate a significant difference between some pair-observed variables (Table S2). The results of $F = 1.00$ and $P = 0.4761$ indicate no significance of the effect of column temperatures on the pair difference between Eqs. (1) and (4).

To identify specific differences between coefficients (pairs of treatments), the multiple comparisons of least significant difference (LSD), *t* tests were applied. The $LSD_{0.05}$ value, the means of $A - U$ and $B - V$ were 0.08564, 0.00000 and 0.00249, respectively, which indicates no significant difference between the pairs of treatments (*A* and *U*, and *B* and *V*) for models [Eqs. (1) and (4)] (Table S3). Therefore, the QE model can express the retention behavior of homologs of *n*-alkanes as accurately as the DE model although the holdup time has been included in the QE model. In other words, the t_{MT} values are indeed the holdup time with a high degree of accuracy, because different results will be obtained from Eqs. (1) and (4) if the t_{MT} used in Eq. (4) is not accurate and reliable. It is noteworthy that the $LSD_{0.05}$ and the mean of $C - W$ were 0.08564 and 0.24704, respectively (Table S3), which indicates a significant difference of intercepts between Eqs. (1) and (4). This can be explained that the Δt_z values can characteristically reflect the molecular interactions between the solutes *n*-alkenes and the stationary phase [1], and the t'_{RT} is the sum of Δt_z fragments starting from propane [$t'_{RT} = (t_{R3} - t_{MT}) + \Delta t_4 + \Delta t_5 + \cdots + \Delta t_z$]. Therefore, both Δt_z and t'_{RT} values represent the retention behavior of the solutes. However, the Δt_z decreased faster than t'_{RT} values as the column temperature increased (Fig. 1a). This phenomenon can be explained by intermolecular forces between the solute *n*-alkanes and the stationary phase, which are presumably primarily dispersion forces in the present study. A first approximation of the interaction energy of dispersion forces is proportional to the square of the polarizability of the molecule (α^2) and inversely proportional to 6th degree of the distance between the molecules ($1/r^6$) [21]. At lower column temperatures (*T_c*), the thermodynamic energy of solutes is small, and *r* values are small, and consequentially the dispersion forces exponentially increase (Table S4). Therefore, at 70 °C, the Δt_{13} and corresponding t'_{RT} (*C*₁₃ *n*-alkane) values were 110.962 and 107.384 min, respectively. As the column temperature increases, the thermodynamic energy of the solute becomes large and results in large *r* values; consequentially, dispersion forces that are dominated by *r* values decrease

quickly. In such a case, the dispersion forces are primarily controlled by the α^2 values which are proportional to the molar volume of the solutes. Therefore, the Δt_{13} and $t'_{RT}(C_{13})$ values at 160 °C were 1.832 and 3.351 min, respectively, which the former was 1.8 times smaller than the latter (Table S4). It causes the DE curves of $\log \Delta t_z$ versus z to shift down relative to the QE curves (Fig. 1a); however, the dependence of retention times on the carbon number is the same for both models.

3.1.2. The experimental results in both DE and QE models [Eqs. (1) and (4)]—

Fig. 1 shows the experimental results from both models Eqs. (1) and (4) on the ZB-1 column at 70, 100, 130 and 160 °C, which confirms the results of hypothesis test in statistics. Fig. 1a shows that both ($\log \Delta t_z$ versus z) and ($\log t'_{RT}$ versus z) curves at 70 °C were almost overlapped, which indicates that the three coefficients in Eq. (1) were almost the same as the corresponding coefficients in Eq. (4). The differences of intercept between both models gradually increased with increasing column temperatures. The two curves started to separate at 100 °C and became more obviously apart at 160 °C. Both curves at each temperature were completely parallel because A value equaled U value, and B equaled V , which followed Eqs. (1) and (4), respectively. Fig. 1b shows that $\Delta \log \Delta t_z$ and $\Delta \log t'_{RT}$ values followed Eqs. (2) and (5), respectively, which decreased with increasing z and the pair curves were always parallel under each temperature. Fig. 1c shows similar results of which the raw retention times were cited from the literature [12,13]. Fig. 1d shows that the mean $\Delta^2 \log \Delta t_z$ was a constant which equaled $2A$ [Eq. (3)] and the mean $\Delta^2 \log t'_{RT}$ equaled $2U$ [Eq. (6)]. Because the coefficient A equaled U , the mean $\Delta^2 \log \Delta t_z$ equaled mean $\Delta^2 \log t'_{RT}$, both were -0.0014 at 70 °C and -0.0011 at 130 °C, respectively (Table S5). Therefore, Eqs. (4), (5) and (6) correspond to Eqs. (1), (2) and (3), describing the dependence of retention times on the carbon number, except the different types of the retention parameters: one is t'_{RT} in Eq. (4) and the other is Δt_z in Eq. (1).

3.2. Accuracy and precision of curve fitting of Eq. (4)

3.2.1. Accuracy—To compare the accuracy of a curve fitting of Eq. (4) with Eqs. (1), (7) and (9), the differences (ΔI) between calculated (I_{cal}) and defined (100 z) reference points of retention indices are considered as an indicator of the evaluation. Table 1 shows the calculated (t_{MT} , t_{MM}) and experimental (t_{MNe} , t_{MCH_4}) holdup times. The t_{MT} values were calculated by the QE model [Eq. (4)] and were compared with t_{MM} values and the experimental t_{MNe} and t_{MCH_4} values. The t'_{RMM} , t'_{RNe} and t'_{RCH_4} are referred to as adjusted retention times using t_{MM} , t_{MNe} and t_{MCH_4} , respectively. Using the raw retention data at 70 °C as an example, Fig. 2a shows that the ΔI values calculated from Eqs. (11) and (12) (symbols Δt_{zDE} , t'_{RT}) of the DE and QE models were almost always close to zero at the range of C₇–C₁₄ n -alkanes. However, three sets of ΔI values calculated from Eqs. (13) and (14) (symbols Δt_{zLE} , t'_{RMM} and t'_{RCH_4}) for the linear equation spread to both sides of zero, especially the ΔI values from t'_{RCH_4} . The mean squared errors (MSE) of the ΔI values were 0.0001, 0.0007, 0.5574, 0.1606 and 3.1943 for t'_{RT} , Δt_{zDE} , Δt_{zLE} , t'_{RMM} and t'_{RCH_4} , respectively (Table 2). The accuracy of the QE model (t'_{RT}) was approximately 2–6 times better than the DE model (Δt_{zDE}) and more than 18–540 times better than the LE model (Δt_{zLE} , t'_{RMM} and t'_{RCH_4}) (Table 2). The t_{MM} value is an artificial t_M value that is useful only for comparison of I values among inter-laboratory. There is no real meaning in spite of its deviation of ΔI values being less than that of Δt_{zLE} . The large deviations of ΔI values from Δt_{zLE} resulted from a systematical error of the LE method, because the plot of $\log \Delta t_z$ versus z is a slightly convex

curve. The LE regression resulted in negative deviations at both ends and positive deviations in the middle for z values. The more number of individual n -alkanes were evaluated, the larger the deviations of ΔI values became. Fig. 2b shows similar results calculated using raw retention times reported by Lebrón-Aguilar et al. [12]. Many publications considered that neon shows the smallest retention time under any GC conditions. Fig. 2b shows clearly that the deviations of ΔI values from t'_{RNe} were much larger than those from the Δt_{LE} , although t'_{MNe} was less than t_{MCH_4} . The MSE values of the ΔI values under various conditions: different column temperatures, ranges of carbon, and numbers of individual n -alkanes evaluated are detailed in Table 2. The QE model looks apparently perfect under any condition tested. The MSE values of the LE model, not only were larger than those of QE, but also increased quickly with increasing the number of the n -alkanes evaluated, especially those from t'_{RCH_4} , which is similar with the results previously described [1].

3.2.2. Precision—The mean relative dispersion of the raw retention times was $1.04 \pm 0.21 \times 10^{-3}$, which is roughly the same level as the data published [14] and was discussed by Wu et al. [1]. Here, the discussion focuses on comparison of the dispersion of the data points between t'_{RT} and Δt_z , for which $\Delta^2 \log t'_{\text{RT}}$ and $\Delta^2 \log \Delta t_z$ values were compared (Tables S5 and S6). Both mean $\Delta^2 \log t'_{\text{RT}}$ and $\Delta^2 \log \Delta t_z$ values were approximately equal at the same column temperature according to Eqs. (6) and (3), whereas SDs of $\Delta^2 \log t'_{\text{RT}}$ were approximately 2–3 times smaller than those of $\Delta^2 \log \Delta t_z$ (Tables S5 and S6 and Fig. 1d). Each t'_{RT} value was larger than the corresponding Δt_z for each individual n -alkane, which caused the smaller relative dispersion of each t'_{RT} value than that of Δt_z for the same set of experimental data. Therefore, use of the t'_{RT} values in Eq. (4) can minimize the dispersion.

A disadvantage of the DE model is that the number ($z - 1$) of calculable retention values is smaller than the number (z) of raw retention times. It is avoided in the QE model. In summary, the QE model kept the essence of the DE model and overcame its weakness. The QE model is arguably perfect for practical applications.

3.3. The relationship between retention models and the advancement of GC technology

The development of a retention model depends upon the accuracy and precision of raw retention data which depend upon the advancement of GC technology. Many factors are involved here; discussion of them all is beyond the scope of this paper. Here we discuss the type of recorders and their corresponding deviations for raw retention data only. Table 3 shows a brief account of improvement of measured deviations of raw retention data with GC advancement from the 1950s until now [10–18,22–28]. The differences (ΔI) between calculated (I_{cal}) and defined (100 z) reference points of retention indices still can be considered as an indicator for evaluating the accuracy of DE model curve fittings. Table 4 shows the effects of three different deviations of raw retention data on ΔI values for eight examples [1,12–14,17,22].

When a potentiometer was used to record a solute profile, the retention distance of each solute was measured manually. The absolute error of one's ability to measure the distance on the chart paper was 0.1–1 mm [22–25], which corresponds 0.6–6.0 s when the chart speed was 10 mm/min. The MSE of the ΔI values was as high as 6.38 (Table 4). After the HP 3370 series integrator was launched in 1968 [28], the integrator with the deviation of the retention time of 0.01 min, i.e. 0.6 s has been used widely [13,26–28], which lead the MSE values decreased to 0.06–0.97. In the recent years, the deviation was improved to 0.001 min, i.e. 0.06 s [10–12,14–18] for commonly used integrators or computers and, therefore, the

MSE values were about less than 0.03. It is evident that the MSE values are about 100 times less than those of about five decades ago. Therefore, the improvement of raw retention data acquisition inspired the developments of the DE and QE models. The precision of retention times can now reach up to 0.0008 min, i.e. 0.048 s, for example, using Agilent 5989N GC [www.agilent.com/chem].

3.4. $R\Delta t_z$ values, the reliability of the raw retention data and retention models

Eq. (2) shows that the $\log(\Delta t_{z+1}/\Delta t_z)$ of n -alkanes is a linear function with the carbon number, while the ratio $\Delta t_{z+1}/\Delta t_z$ itself (hereafter, $R\Delta t_z$) should be also linear, as expressed.

$$R\Delta t_z = \frac{\Delta t_{z+1}}{\Delta t_z} = \beta''z + \gamma'' \quad (16)$$

$R\Delta t_z$ values of the n -alkanes have been defined as the ratio A that was used to assess the linearity of n -alkanes in many publications in the 1970s–1980s [29–31]. We found that the relationship between $R\Delta t_z$ and z [Eq. (16)] is also a simple and accurate method to assess the reliability of the raw retention data and consequently the type of the retention model. When the accuracy and precision of the raw retention data are low, the $R\Delta t_z$ becomes a constant [Eq. (8)], indicating the LE type. When the deviation of raw retention data is within 0.001 min, i.e. 0.06 s, $R\Delta t_z$ decreases with increasing z , indicating the QE type. Fig. 3a shows that when the t_{Rz} value 379.5 s for n -hexane [12] was used to calculate the $R\Delta t_z$ value, the “incorrect” curve was generated, the R^2 of the regression of the $R\Delta t_z$ versus z was close to zero ($R^2 = 0.0027$), indicating the LE type. After the error of 395.5 s was corrected to 279.5 s, the “correct” curve indicated that the $R\Delta t_z$ value decreased with increasing carbon number ($R^2 = 0.9926$), indicating the QE type. This phenomenon suggests that 379.5 s was a typo which has been confirmed by Dr. Rosa Lebrón-Aguilar, one of the authors. In the later publications [34,35], the typo data was still cited, while the authors acknowledged and discussed the deviations. Fig. 3b shows the $R\Delta t_z$ values were constant ($R^2 = 0.0024$), indicating the LE type due to the deviation of the raw retention distance of 1 mm [22]. The true dependence of $R\Delta t_z$ on z [Eq. (16)] was concealed by the larger random error of raw retention data under this condition. The LE model was adequately used under the technology level of the early period of GC. Fig. 3c shows that the $R\Delta t_z$ values decreased with increasing carbon number [follows Eq. (16)], indicating the LE type due to the decreased deviations from 6 s to 0.6 s, i.e. 0.01 min. However, the R^2 values decreased with increasing column temperatures, which were 0.8175, 0.7610 and 0.3751 for 150, 160 and 200°C, respectively. It is because that the dispersion of the data points increased with increasing column temperature. Therefore, obtaining a correct retention behavior requires both lower deviations of raw retention data and correct GC condition settings. The regression lines 1 and 4 in Fig. 3d show that the $R\Delta t_z$ values decreased with increasing carbon number, which follows Eq. (16) well due to the decreased deviation from 0.6 s to 0.06 s (i.e. 0.001 min). Equation of line 1 was $R\Delta t_z = -0.0034z + 1.6148$ and R^2 was 0.9913. To further confirm the effect of different levels of deviation of retention times on the type of the retention model, 0.001 min and 0.01 min were added to the retention time of n -undecane, lines 2 and 3 in Fig. 3d were obtained, respectively. The regression equation of line 2 became $R\Delta t_z = -0.0034z + 1.6149$ and R^2 was 0.9876. The facts that the slope and intercept were almost the same for both lines 1 and 2 and the small decrease of R^2 for line 2 indicate the deviation of 0.001 min does not affect the expression of the true dependence of $R\Delta t_z$ on z . However, the regression equation of line 3 became $R\Delta t_z = -0.0014z + 1.5846$, and R^2 was 0.0221, which is different from line 1. It indicates that the dependence of $R\Delta t_z$ on z changes from the QE to the LE model, when the deviation of only one retention time in C_8 – C_{20} n -alkanes reached to 0.6 s (0.01 min). Therefore, the validity of the QE model is based on the good accuracy and

precision of raw retention data, which has been made possible by the advancement of GC technology.

3.5. Thoughts for lack of general acceptance of the nonlinear model for wide applications

The linear relationship between the logarithm of retention data (t'_{RZ} , corrected retention times, net retention times or corresponding volumes) in homologs of n -alkanes and their carbon number was established about 50 years ago [6,19,20,36,37]. Smith et al. [33] stated two reasons that nonlinearity models have not been uniformly recognized: first, the nonlinearity is minimized for large n -alkanes at high temperatures and on non-polar stationary phases; second, nonlinearity can be obscured (if not large) by mathematical treatments of the data. We would question that the calculation complexity may hinder applications of the complicated nonlinear models. Another major reason is lack of a model being simple, but more accurate and reliable than the LE model.

Many experimental and theoretical evidences have been well presented for nonlinearity of the plots of $\log t'_{RZ}$ versus the carbon number not only for small n -alkanes [29–33,38], but also all homolog of n -alkanes [10–18]. Our previous [1] and the present work have shown that the DE and QE models have significantly improved the accuracy of curve fittings in comparison with the LE model for C_3 – C_{30} n -alkanes at column temperatures of 24–260°C. Of course, the nonlinearity for n -alkanes with large carbon numbers of at high temperature can be treated as in LE approximately for simplicity. However, nowadays quadratic calculations can be carried out as easily as that of LE. The QE model has two main advantages over the LE model. (1) The QE model has much better accuracy and a larger range of curve fittings than the LE model. (2) The holdup time can be accurately and reliably calculated by the QE model itself; while for the LE model, the t_M value depends upon experimental or calculation methods. Therefore, the QE model can be a plausible substitute of the LE model.

4. Conclusion

The QE model was developed based on the DE model. Both models have been developed via two approaches. The DE model is obtained from searching a mathematic expression of retention behavior of n -alkanes excluding the influence of t_M . On the contrary, the QE model is obtained from searching a t_M value for the optimal fitting of the quadratic equation. Therefore, the QE model can be considered as the results obtained from simulating the DE model. The relationship between retention times and the carbon number can be quantitatively described at three levels [Eqs. (4, 5 and 6)]. For an accurate expression of the retention behavior of n -alkanes and model applications, the QE model [Eq. (4)] is the same as or better than DE [Eq. (1)]. The plot ($\log t'_{RT}$ versus z or $\log \Delta t_z$ versus z) is slightly convex and, therefore, the curve best fits with a quadratic equation. Because of this, it is no surprise to notice the large deviations from the fitted curve in the LE model, no matter what t_M (t_{MM} , t_{MNe} or t_{MCH_4}) is used. The errors come from two sources: the system error of LE and the incorrect t_M value. The QE model has significant improvements over the LE model and, thus, is recommended for applications.

Development of retention models depends on the accuracy and precision of measurements of raw retention data as well as the advancement of GC technology. The LE model established about 50 years ago was based on the data measured with deviations of 0.6–6 s. Consequently, the LE model describes the approximate retention behavior of n -alkanes. When the deviation of the measurement of retention times improved to 0.06 s, the DE and QE models can accurately describe the retention behavior of n -alkanes. The $R\Delta t_z$ value can be used to assess the reliability of the raw retention data and the type of the retention model.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by the U.S. Fish and Wildlife Service, US ONR HEET (N00014-09-1-0709), the National Center for Research Resources (5G12 RR003061-26; 5P20RR016467-11), the National Institute on Minority Health and Health Disparities (8G12MD007601-26) and the National Institute of General Medical Sciences (8P20GM103466-11) from the National Institutes of Health. We thank Prof. Maolin Li and Prof. Minglin Wang for helpful discussions.

References

1. Wu, L.; Chen, Y.; Caccamise, SAL.; Li, QX. J. Chromatogr. A. 2012. <http://dx.doi.org/10.1016/j.chroma.2012.07.077>
2. Kaiser R. Chromatographia. 1969; 2:125, 215.
3. Conder, JR.; Young, CL. Physicochemical Measurements by Gas Chromatography. John Wiley & Sons; New York: 1979.
4. Budahegyi MV, Lombosi ER, Lombosi TS, Mészáros SY, Nyiredy S, Tarján G, Timár G, Takács JM. J. Chromatogr. 1983; 271:213.
5. Pacakova, V.; Feltl, L. Chromatographic Retention Indices. An Aid to Identification of Organic Compounds; Ellis Horwood, Chichester, UK: 1992.
6. Ettre, LS.; Hinshaw, JV. Basic Relationships of Gas Chromatography. Advanstar Communications Inc; Cleveland, OH: 1993.
7. Poole CF, Poole SK. J. Chromatogr. A. 2009; 1216:1530. [PubMed: 19013576]
8. Guardino X, Albaigés J, Firpo G, Rodríguez-Vinãls R, Gassiot M. J. Chromatogr. 1976; 118:13.
9. Santiuste JM, Quintanilla-López JE, Takács JM, Lebrón-Aguilar R. J. Chromatogr. A. 2012; 1222:90. [PubMed: 22197020]
10. Touabet A, Maeck M, Badjah Hadj Ahmed AY, Meklati BY. Chromatographia. 1988; 25:389.
11. Lebrón-Aguilar R, Quintanilla-López JE, Garcia-Dominguez JA. J. Chromatogr. A. 1997; 760:219.
12. Quintanilla-López JE, Lebrón-Aguilar R, Garcia-Dominguez JA. J. Chromatogr. A. 1997; 767:127.
13. Garcia-Dominguez JA, Quintanilla-López JE, Lebrón-Aguilar R. J. Chromatogr. A. 1998; 803:197.
14. Lebrón-Aguilar R, Quintanilla-López JE, Garcia-Dominguez JA. J. Chromatogr. A. 2002; 945:185. [PubMed: 11860135]
15. González FR. J. Chromatogr. A. 2000; 873:209. [PubMed: 10757298]
16. González FR, Gagliardi LG. J. Chromatogr. A. 2000; 879:157. [PubMed: 10893032]
17. Heeg FJ, Zinburg R, Neu HJ, Ballschmiter K. Chromatographia. 1979; 12:451.
18. Ballschmiter K, Heeg FJ, Zinburg R, Fresenius Z. Anal. Chem. 1985; 321:426.
19. Sz Kováts E. Helv. Chim. Acta. 1958; 41:1915.
20. Ettre LS. Anal. Chem. 1964; 36:31A.
21. Scott, RPW. Introduction to Analytical Gas Chromatography. Marcel Dekker Inc; New York: 1998. p. 33
22. Šev ík J, Löwentap MSH. J. Chromatogr. 1978; 147:75.
23. Merritt C Jr. Walsh JT. Anal. Chem. 1962; 34:903.
24. Merritt C Jr. Walsh JT. Anal. Chem. 1962; 34:908.
25. Ettre LS. Chromatographia. 1973; 6:489.
26. Garcia-Dominguez JA, Garcia Munoz J, Fernandez Sanchez E, Molera MJ. J. Chromatogr. Sci. 1977; 15:520.
27. The Sadtler Research Laboratories. The Sadtler Standard Gas Chromatography Retention index Library. Vol. 4. Philadelphia, PA: 1986.
28. Dyson, N. Chromatographic Integration Methods. 2nd ed.. Dyson Instruments Ltd.; UK: 1998. p. 110

29. Wainwright MS, Haken JK, Srisukh D. J. Chromatogr. 1980; 188:246.
30. Wainwright MS, Haken JK, Srisukh D. J. Chromatogr. 1980; 189:156.
31. Haken JK, Wainwright MS, Srisukh D. J. Chromatogr. 1983; 257:107.
32. Riedo F, Fritz D, Tarjan G, Sz Kovats E. J. Chromatogr. 1976; 126:63.
33. Smith RJ, Haken JK, Wainwright MS, Madden BG. J. Chromatogr. 1985; 334:95.
34. Watanachaiyong T, Jeyashoke N, Krisnangkura K. J. Chromatogr. Sci. 2000; 38:67. [PubMed: 10677835]
35. Wu NS, Wu GS, Wu MY. J. Chromatogr. Sci. 2006; 44:244. [PubMed: 16774708]
36. James AT, Martin AJP. Biochem. J. 1952; 50:679. [PubMed: 14934673]
37. Littlewood AB, Phillips CSG, Price DT. J. Chem. Sci. 1955:1480.
38. Rohrschneider L. Chromatographia. 1969; 2:437.

Fig. 1.

Comparison between the QE and DE models for Eqs. (4)–(6) and Eqs. (1)–(3), respectively. (a), (b) ZB-1 column (60 m × 0.25 mm × 0.25 μm) at 70, 100, 130 and 160°C (Table 1 in [1]). (c) CPSIL-5CB (50 m × 0.32 mm × 0.43 μm) at 60 and 90°C [12] and HP-5 (60 m × 0.25 mm × 0.25 μm) at 120°C [12] and 150°C [13]. (d) ZB-1 column at 70, 130°C (Table 1 in [1]).

Fig. 2.

The differences (ΔI) between calculated (I_{cal}) and defined ($100z$) of reference points of retention indices of the QE model in comparison with those of the DE and LE models. ΔI was calculated according to Eq. (15). The Δt_{ZDE} and Δt_{ZLE} data were the calculated I_{cal} values from Eqs. (11) and (13), respectively. t'_{RT} was the calculated I_{cal} value from Eq. (12). t'_{RMM} , t'_{RNe} and t'_{RCH_4} were the calculated I_{cal} values from Eq. (14) with the t_{MM} [8], t_{MNe} and t_{MCH_4} , respectively. (a) The raw data for C₇–C₁₄ *n*-alkanes on the ZB-1 column at 70°C (Table 1 in [1]). (b) The raw data for C₆–C₁₃ *n*-alkanes on the HP-5 column at 120°C [12].

Fig. 3.

Relationship between $R\Delta t_z$ and z at different GC development stages and assessing the reliability of raw retention data. (a) CPSIL-5CB at 60°C [12]. "Incorrect" and "correct" are referred to the t_{Rz} value of *n*-hexane as 379.5 s (typo) and 279.5 s, respectively, which was confirmed by Dr. Rosa Lebrón-Aguilar. (b) 1.5% OV-17 and 2.0% OV-210 micro-packed column (3 m × 0.8 mm I.D.) at 85°C; the measurement deviation of the raw retention distance was 1 mm [22]. (c) HP-5 at 150, 160 and 200°C with recording precision of 0.01 min [13]. (d) ZB-1 column at 150°C; the precision of retention time of line 1 was ±0.001 min [1]. Lines 2 and 3 data had an addition of 0.001 min and 0.01 min to only the retention time of C₁₁, respectively (the other data remained the same as line 1). Line 4 data from a 30-m glass capillary column with SE-30 at 160°C [17].

Table 1

Calculated (t_{MT} , t_{MM})^a and experimental (t_{MNe} , t_{MCH4}) holdup times cited from different sources of data.

Column	Column temperature (°C)	Holdup times		Sources of data		
				t_{MM}	t_{MNe}	t_{MCH4}
ZB-1	70	min	3.659	3.602	–	3.705
ZB-1	100	min	3.831	3.794	–	3.878
ZB-1	130	min	4.015	3.968	–	4.062
ZB-1	160	min	4.166	4.137	–	4.209
CPSIL-5CB	60	s	193.53	193.19	194.8	195.3
CPSIL-5CB	90	s	209.84	209.55	210.9	211.4
HP-5	120	s	360.20	359.50	361.3	361.8
HP-5	150	min	14.55	14.51	–	14.58

^aThe t_{MT} values were calculated by the QE model [Eq. (4)]; the t_{MM} values were calculated by using the multiple linear regression and curve fitting of Eq. (9).

Table 2

Mean squared errors (MSE) of the ΔI values from the QE, DE and LE models at different column temperatures ($^{\circ}\text{C}$) and different numbers of n -alkanes evaluated (N).

	Column temperature ($^{\circ}\text{C}$)	N	Range of carbon	MSE ^d of the ΔI values					
				QE t'_{RT}	DE Δt_{zDE}	LE Δt_{zLE}	LE t'_{RMM}	LE t'_{RNe}	LE t'_{RCH_4}
70 ^b		8	C ₇ –C ₁₄	0.0001	0.0007	0.557	0.161	–	3.194
100 ^b		10	C ₇ –C ₁₆	0.0008	0.0044	1.632	0.465	–	13.062
130 ^b		12	C ₈ –C ₁₉	0.0030	0.0150	4.214	1.136	–	23.992
160 ^b		12	C ₉ –C ₂₀	0.0016	0.0125	4.255	1.255	–	26.446
Mean		10.5	–	0.0014	0.0082	2.665	0.754	–	16.673
60 ^c		8	C ₃ –C ₁₀	0.0006	0.0016	0.765	0.272	31.933	59.300
90 ^c		10	C ₃ –C ₁₂	0.0120	0.0520	1.381	0.496	56.332	118.513
120 ^d		8	C ₆ –C ₁₃	0.0023	0.0064	0.416	0.090	3.807	6.250
150 ^e		9	C ₇ –C ₁₅	0.0450	0.0599	1.119	0.209	–	5.520
Mean		8.75	–	0.0150	0.0300	0.920	0.267	30.691	47.396

${}^a\text{MSE} = \left(\sum (\Delta I)^2 \right) / N$. $\Delta I = I_{\text{cal}} - 100z$ [Eq. (15)]; $I_{\text{cal}} (t'_{\text{RT}} \text{ and } \Delta t_{\text{zDE}})$ were calculated with Eqs. (12) and (11), i.e., $I_{\text{cal}} = \left[U' \left(\log t'_{\text{RT}} \right)^2 + V' \left(\log t'_{\text{RT}} \right) + W' \right] 100$ and $I_{\text{cal}} = (A' (\log \Delta t_{\text{z}})^2 + B' (\log \Delta t_{\text{z}} + C') 100)$; $I_{\text{cal}} (\Delta t_{\text{zLE}})$ was calculated with Eq. (13) $I_{\text{cal}} = [B' (\log \Delta t_{\text{z}} + C') 100]$; $I_{\text{cal}} (t'_{\text{RMM}}, t'_{\text{RNe}} \text{ and } t'_{\text{RCH}_4})$ were calculated with Eq. (14) $I_{\text{cal}} = (b' (\log t'_{\text{RZ}}) + c') 100$.

^bZB-1 [1].

^cCPSIL-5CB [12].

^dHP-5 [12].

^eHP-5 [13].

Table 3

Advances of raw data acquisition at different GC technology development stages.

	1950s–1970s	1970s–1980s	1980s–now
Recorders	Potentiometric with response time 0.5 s	Integrator	Digital integrator and workstation
Types of data processor	Manual measurement of distances by a ruler	Digital	Digital
Deviations of retention data	1–0.1 mm, i.e., 0.1–0.01 min (6–0.6 s)	0.01 min (0.6 s)	0.0008–0.001 min (0.048–0.06 s)
References	[22–25]	[13,26–28]	[10–12,14–18, www.agilent.com/chem]

Table 4

Effects of three different deviations (0.1, 0.01 and 0.001 min) of raw retention data on ΔI values at different column temperatures ($^{\circ}\text{C}$) for eight GC studies [I_{cal} and ΔI values were calculated according to Eqs. (11) and (15)].

z^a	ΔI	0.1 min							
		85 $^{\circ}\text{C}$ [22]				200 $^{\circ}\text{C}$ [13]			
		0.001 min	0.01 min	0.1 min	0.001 min	0.001 min	0.01 min	0.1 min	0.001 min
		120 $^{\circ}\text{C}$ [14]	160 $^{\circ}\text{C}$ [17]	60 $^{\circ}\text{C}$ [12]	140 $^{\circ}\text{C}$ [1]	150 $^{\circ}\text{C}$ [13]	160 $^{\circ}\text{C}$ [13]	200 $^{\circ}\text{C}$ [13]	160 $^{\circ}\text{C}$ [13]
3				-0.02					
4				0.02					
5				0.00					
6				0.07					
7		-0.25	0.37	-0.05		0.00	-0.94	-0.76	
8	0.32	0.06		-0.05		0.19	0.87	1.82	
9	-0.12	-0.16		0.04		-0.28	0.96	-1.71	-2.56
10	-0.24	-0.01				-0.04	-0.48	0.46	-1.2
11	-0.16	-0.06				0.17	-0.20	1.23	4.82
12	-0.05	-0.05				-0.28	-0.12	-1.20	5.13
13	-0.03	0.09	0.28			0.50	-0.09	0.45	-3.26
14	0.08	-0.19	-0.19			-0.25	-0.03	-0.34	-2.18
15	0.14	-0.30	-0.30				-0.47	-0.44	-0.76
16	0.14	0.00	0.00			0.50	0.50	0.49	-1.17
17	0.16	-0.03	-0.03					0.06	0.06
18	-0.07	0.21	0.21						-0.03
19	-0.17	0.13	0.13						1.29
20		0.10	0.10						
21		-0.21	-0.21						
MSE ^b		0.03	0.03	0.00	0.03	0.06	0.30	0.97	6.38
R^2 ^c		0.9606	0.9606	0.9926	0.9924	0.8175	0.7610	0.3751	0.0049

^a z , carbon number of n -alkanes.

^b MSE, mean of square error, its calculation is shown in footnote of Table 2.

^c R^2 of the regression of the $R\Delta I_z$ versus z [Eq. (16)].