Published in final edited form as:

J Chromatogr A. 2018 March 02; 1539: 62–70. doi:10.1016/j.chroma.2018.01.049.

Surface Fitting for Calculating the Second Dimension Retention Index in Comprehensive Two-dimensional Gas Chromatography Mass Spectrometry

Md Aminul Islam Prodhan^{1,2,3,4}, Xinmin Yi^{1,2,3,4}, Seongho Kim⁸, Craig McClain^{2,3,5,6,7}, and Xiang Zhang^{1,2,3,4,5,*}

¹Department of Chemistry, University of Louisville, Louisville, KY 40208, USA

²University of Louisville Alcohol Research Center, University of Louisville, Louisville, KY 40208, USA

³University of Louisville Hepatobiology & Toxicology Program, University of Louisville, Louisville, KY 40208, USA

⁴Center for Regulatory and Environmental Analytical Metabolomics, University of Louisville, Louisville, KY 40208, USA

⁵Department of Pharmacology & Toxicology, University of Louisville, Louisville, KY 40208, USA

⁶Department of Medicine, University of Louisville, Louisville, KY 40208, USA

⁷Robley Rex Louisville VAMC, Louisville, Kentucky 40292, USA

⁸Biostatistics Core, Karmanos Cancer Institute, Department of Oncology, School of Medicine, Wayne State University, Detroit, MI 48201, USA

Abstract

Comprehensive two-dimensional gas chromatography mass spectrometry (GC×GC-MS) has been widely used for analysis of volatile compounds. However, the second dimension retention index (*I*) of each compound is not widely used to aid compound identification owing to the limited accuracy of *I* calculation. We report a surface fitting approach to the calculation of *I* using *n*-alkanes (C₇ – C₃₀) as references, where the second dimension retention time (²*t_R*) and the second dimension column temperature (²*T_c*) formed the *X-Y* plane and the *I* was the *Z*-axis to form the *I* surface. Compared to the conventional approach for calculating *I* using isovolatility curves, the surface fitting approach eliminated the construction of isovolatility curves for the reference compounds and gives better reproducibility. The goodness of the proposed surface fitting achieved *R*² = 0.9999 and *RMSE* = 6.1 retention index units (iu). Ten-fold cross validation demonstrated the surface fitting approach had a good predictability with average *R*² = 0.9999 and *RMSE* = 6.6 iu. The developed method was also applied to calculate the second dimension retention indices of

*CORRESPONDING AUTHORS: Prof. Xiang Zhang, Department of Chemistry, University of Louisville, 2320 South Brook Street, Louisville, KY 40292, USA. Phone: +01 502 852 8878. Fax: +01 502 852 8149. xiang.zhang@louisville.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

compound standards in two commercial mixtures MegaMix A and MegaMix B. The mean standard deviation of the calculated I was only 1.6 iu for compounds in MegaMix A and 3.4 iu for compounds in MegaMix B. Compared with the literature results, the small value of standard deviation in the calculated retention index using surface fitting method shows that the surface fitting method has less measurement variability than the conventional isovolatility curve approach.

Keywords

Retention index calculation; second dimension retention index; surface fitting; GC×GC-MS

1. Introduction

While mass spectrometry (MS) coupled with a one-dimensional separation technique, such as gas chromatography (GC) or liquid chromatography (LC) has been widely used in metabolomics, the limited separation power of the one-dimensional GC or LC contributes to the poor performance of GC-MS and LC-MS in detecting low abundance compounds from complex samples such as human body fluids and tissues. Comprehensive two-dimensional gas chromatography mass spectrometry (GC×GC-MS) uses two GC columns with different stationary phases to increase the separation power. The first dimension column, i.e. the primary column, is usually a 10 – 60 m long column while the second dimension column is a short column about 1 – 2 m^(1–5). The two columns are connected by a thermal modulator. Compounds eluted from the first dimension column within a modulation period P_M are collected by the thermal modulator and subjected to the second dimension column for further separation at an increased temperature. Depending on the system configuration, the modulation period is usually set to 2 – 10 s and the temperature gradient on both GC columns is usually programmed to 5 – 10 °C/min. While the two columns are operated in linear temperature-programmed condition, the fast separation and slow temperature gradient on the second dimension column results in a small temperature change within a P_M , making a condition such that the second dimension column actually functions under a pseudo-isothermal condition^(6–12).

GC×GC-MS provides three pieces of information for each compound, the first dimension retention 1t_R , the second dimension retention time 2t_R , and an electron ionization (EI) mass spectrum, if an EI source is used^(13, 14). To reduce the dependence of retention time on GC×GC-MS experimental conditions, 1t_R and 2t_R of a compound u are respectively converted into linear retention index I^T and Kováts retention index I as follows^(15, 16):

$$I_u^T = 100 \times \left[n + \frac{{}^1t_{R(u)} - {}^1t_{R(n)}}{{}^1t_{R(n+1)} - {}^1t_{R(n)}} \right] \quad (1)$$

$$I_u = 100 \times \left[n + \frac{\log({}^2t'_{R(u)}) - \log({}^2t'_{R(n)})}{\log({}^2t'_{R(n+1)}) - \log({}^2t'_{R(n)})} \right] \quad (2)$$

where n is the number of carbon atoms in a n -alkane. $^1t_{R(u)}$, $^1t_{R(n)}$, and $^1t_{R(n+1)}$ are the first dimension retention times of the compound u , the n -alkane with n carbon atoms, and the n -alkane with $n+1$ carbon atoms, respectively. $^2t'_R$ is the adjusted second dimension retention time, i.e. the second dimension retention time of a chromatographic peak minus hold-up time (2t_M) that is the retention time of a unretained chromatographic peak.

During GC×GC-MS analysis, compounds eluted within the same P_M from the first dimension column are collected and subjected to the second dimension column for further separation at a different temperature, as well as on a different stationary phase. While the separation of the compounds on the second dimension column within one P_M can be considered as in pseudo-isothermal condition, the overall temperature of the second dimension column 2T_e changes during the entire GC×GC-MS analysis. The retention time 2t_R of a compound in the second dimension column decreases with the increase of column temperature 2T_e .

The relation $^2t_R \sim ^2T_e$ of a compound can be captured by an isovolatility curve. Several methods have been proposed to construct the isovolatility curves for the reference compounds such as n -alkanes, including Beens et al.^(17, 18), Marriot et al.^(19–22), and Zhao et al.⁽²³⁾. In these reported methods, the hold-up time 2t_M at different temperatures of the second dimension column must be first estimated to get the adjusted retention time $^2t'_R$. The $^2t'_R$ values of the same reference compound at different column temperatures were then used to construct the isovolatility curve of a reference compound. The isovolatility curves of all reference compounds formed the second dimension retention index map, and the I value of a compound was finally calculated from the two adjacent isovolatility curves of two reference compounds.

The objective of this study was to develop a method to calculate I value of each compound without the construction of isovolatility curves for the reference compounds. We developed a surface fitting method for the calculation of I values for all compounds detected by GC×GC-MS. The developed surface fitting method used n -alkanes as reference compounds, where the second dimension retention time 2t_R and the second dimension column temperature 2T_e formed a X - Y plane and the I was the Z -axis to form the I surface. The robustness of the surface fitting method was validated using 10-fold cross validation. The method was then applied to calculate the retention indices of 152 compound standards in two commercially available mixtures, MegaMix A and MegaMix B.

2. Experimental

2.1. Chemicals

A mixture containing $C_7 - C_{30}$ n -alkanes was purchased from Sigma-Aldrich Corp., St. Louis, MO. MegaMix A (Cat. No. 8270) and MegaMix B (Cat. No. 8260) were purchased from Restek Corp., Bellefonte, PA. All of the compounds were present at a purity of between 95–99%. Two solutions were prepared from these three commercial mixtures. One solution contained only $C_7 - C_{30}$ n -alkanes, while the other contained MegaMix A and

MegaMix B. These two solutions were prepared in dichloromethane at a concentration of 25 µg/mL per compound.

2.2. GC×GC-MS experiment

The LECO (St. Joseph, MI) Pegasus® 4D GC×GC-MS instrument was equipped with a Gerstel MPS2 autosampler and an Agilent 6890 gas chromatograph featuring a LECO two-stage cryogenic modulator and a secondary oven. The first dimension column was a non-polar DB-5ms (phenyl arylene polymer virtually equivalent to a (5%-phenyl)-methylpolysiloxane) 60.0 m × 0.25 mm 1d_c × 0.25 µm 1d_f column, and the second dimension column was a medium polar DB-17ms ((50%-phenyl)-methylpolysiloxane) 1.0 m × 0.1 mm 2d_c × 0.1 µm 2d_f column. Both columns were obtained from Agilent Technologies (Santa Clara, CA), and were connected by means of a press-fit connector before the thermal modulator. The flow rate of ultra-high purity helium carrier gas (99.999%) was set to 2.0 mL/min at a corrected constant flow via pressure ramps with an initial head pressure of 36.9 psi.

A 1 µL sample was injected into the inlet chamber of the GC×GC-MS system and split ratio was set to 20:1. The inlet temperature was 280 °C. The temperature of thermal modulator was set to +15 °C relative to the secondary oven temperature. The MS parameters were: MS range 45 – 800 m/z; data acquisition rate 200 spectra/s; temperature of ion source chamber 230 °C; temperature of MS transfer line 280 °C; detector voltage 1641 V; electron energy 70 eV. The acceleration voltage was turned on after a solvent delay of 220 s.

To test the variability of the GC×GC-MS system, the mixture of *n*-alkanes was analyzed 25 times in five days in temperature-programmed mode, where the primary column temperature was held at an initial temperature 60 °C for 1 min and then ramped at 5 °C/min to 285 °C, and kept at 285 °C for 13 min. The secondary column temperature program was identical to that of the first dimension column but temperature was set to +5 °C higher than that of the first dimension column. P_M was set as 10 s.

To construct the second dimension retention index surface, the mixture of C₇ – C₃₀ *n*-alkanes was analyzed by GC×GC-MS 46 times under isothermal conditions. Temperature of the first dimension column 1T_e was set to 60 °C in the first injection, and increased +5 °C in each of the following 45 injections with 1T_e = 285 °C in the last injection. The second dimension column was also set under isothermal conditions but at +5 °C with respect to the temperature of the first dimension column. P_M was set as 10 s.

The mixture of MegaMix A and MegaMix B was analyzed four times in temperature-programmed conditions with the first dimension column temperature initially set to 60, 80, 100, and 120 °C, respectively. The column temperature was held at the initial temperature for 1 min and then ramped at 5 °C/min to 285 °C, and kept at 285 °C for 13 min. The secondary column temperature program was identical to that of the first dimension column but temperature was set +5 °C higher than that of the first dimension column. P_M was set as 4 s.

To study the robustness of surface fitting method, the mixture of MegaMix A and MegaMix B was also analyzed under different instrumental conditions, by varying the temperature offset of the secondary oven from +5 to +10 °C, the ramp rate of temperature gradient from 4 to 6 °C/min, and modulation period from 4 to 5 s. The other instrumental parameters were kept the same as described above.

2.3. Spectrum deconvolution and compound identification

The LECO ChromaTOF software (version 4.51), equipped with the National Institute of Standards and Technology MS database (NIST 2014), was used for instrument control, spectrum deconvolution, and compound identification. Manufacturer recommended parameters for ChromaTOF were used to reduce the instrument data into a peak list of compounds. These parameters were: baseline offset 0.5; auto smoothing; peak width in first dimension 6 s; peak width in the second dimension 0.1 s; signal-to-noise ratio 100; match required to combine peaks 600; R.T. shift 0.08 s; minimum forward similarity match 600. The true peak spectrum was exported as part of the information for each chromatographic peak as absolute intensity values.

2.4. Hold-up time calculation

To calculate the hold-up time on the second dimension column 2t_M , the *n*-alkane mixture was analyzed by GC×GC-MS at 46 different column temperatures in isothermal conditions. Hold-up time 2t_M was then calculated as follows ⁽²³⁾:

$${}^2t_M = \frac{{}^2t_{R1} {}^2t_{R3} - ({}^2t_{R2})^2}{({}^2t_{R3} - {}^2t_{R2}) - ({}^2t_{R2} - {}^2t_{R1})} \quad (3)$$

where ${}^2t_{R1}$, ${}^2t_{R2}$, and ${}^2t_{R3}$ are the second dimension retention times of three continuous homologs such as C_{n-1} , C_n , and C_{n+1} alkanes.

Multiple *n*-alkanes were analyzed on the second dimension column at the same column temperature. Therefore, multiple hold-up times were obtained for each column temperature using equation (3). The mean hold-up time values at different column temperatures were then fitted as follows:

$${}^2t_M = -0.06000 \ln({}^2T_e) + 1.194 \quad (4)$$

where 2t_M is the fitted hold-up time in the second dimension column and 2T_e is the second dimension column temperature.

2.5. Construction of the second dimension retention index surface I_S

$C_7 - C_{30}$ *n*-alkanes were analyzed 46 times by GC×GC-MS in isothermal conditions. 2t_R and 2T_e were obtained for each reference compound in each analysis. The *I* value of a reference compound with *n* carbon atoms was defined as $100 \times n$. Therefore, each *n*-alkane compound was represented by $\{{}^2t_R, {}^2T_e, I\}$. A second dimension retention index surface, I_S ,

was then constructed using the data of all n -alkanes, where 2t_R and 2T_e formed the X - Y plane and the I was the Z -axis to form the I_S . A polynomial surface model in MATLAB (Mathworks, Natick, MA) named 'polyij' was used to construct the I_S , where i was the degree of polynomial in 2t_R , and j was the degree of polynomial in 2T_e . The maximum for both i and j was five in the MATLAB platform. The degree of polynomial was the maximum of i and j .

2.6. Evaluation criteria of surface fitting method

Three measures were used to evaluate the quality of surface fitting, including R-square (R^2), degrees of freedom adjusted R-square (R_{adj}^2), and root mean squared error ($RMSE$). R^2 measures how successful the fit is in explaining the variation of the data, and is defined as the ratio of the sum of squares of the regression (SSR) and the total sum of squares (SST).

$$SSR = \sum_{i=1}^n w_i (\hat{y}_i - \bar{y})^2 \quad (5)$$

$$SST = \sum_{i=1}^n w_i (y_i - \bar{y})^2 \quad (6)$$

$$R^2 = \frac{SSR}{SST} \quad (7)$$

where y_i is the original retention index value I_i , $\hat{y}$ is the predicted I value, $\bar{y}$ is the mean of the original retention index values, and the w_i is the weight on the value I_i . R^2 can take on any value between 0 and 1, with a value closer to 1 indicating that a greater proportion of variance is accounted for by the model.

R_{adj}^2 adjusts R^2 based on the residual degrees of freedom, v , that is defined as the number of response values, n , minus the number of fitted coefficients, m . m is estimated from the response values. The R_{adj}^2 can take on any value less than or equal to 1, with a value closer to 1 indicating a better fit. R_{adj}^2 is calculated as follows:

$$R_{adj}^2 = 1 - \frac{SSE(n-1)}{SST(v)} \quad (8)$$

$$SSE = \sum_{i=1}^n w_i (y_i - \hat{y})^2 \quad (9)$$

$RMSE$ is an estimate of the standard deviation of the random component in the data, and is defined as

$$RMSE = \sqrt{MSE} = \sqrt{\frac{SSE}{v}} \quad (10)$$

where MSE is the mean square error or the residual mean square. An $RMSE$ value closer to 0 indicates a fit that is more useful for prediction.

3. Results and discussion

$C_7 - C_{30}$ n -alkanes were chosen as the reference compounds to create a retention index surface I_5 in this study. All n -alkanes were identified from the GC×GC-MS data. A total of 102 compounds were identified from the mixture of MegaMix A and MegaMix B.

3.1 GC×GC-MS experimental variability

The mixture of n -alkanes was analyzed in temperature-programmed mode 25 times in five continuous days. For the experiments performed within one day, the largest standard deviation of the second dimension retention time of n -alkanes is 0.007 s. Supplementary Figure S1 displays the cumulative probability curves of the day-to-day experiments, i.e. 25 replicate analyses in five days. The standard deviation of the second dimension retention time of the n -alkanes fluctuated from 0.001 s to 0.007 s. The standard deviation of 90% of the second dimension retention time values in all 25 experiments were less than 0.005 s. Such a small standard deviation of the second dimension retention time indicates that the instrument has a very good reproducibility in measuring the retention times of compounds separated on the second dimension column in temperature-programmed mode.

While the n -alkane mixture was analyzed in temperature-programmed mode, each chromatographic run was ramped to 285 °C in the first dimension column and held at 285 °C for 13 min. By design, the second dimension column was in fact operated in isothermal mode at 290 °C during the last 13 min in each of the 25 experiments. Five n -alkanes, i.e. $C_{26} - C_{30}$ n -alkanes, were eluted in isothermal mode in the second dimension column at 290 °C. The standard deviation of the second dimension retention times of these five n -alkanes were 0.004, 0.004, 0.004, 0.004 and 0.005 s, respectively. Such small standard deviations indicate that the reproducibility of the GC×GC-MS system is also quite good in isothermal mode.

3.2 Construction of the second dimension retention index surface

A total of 415 data points of $C_7 - C_{30}$ n -alkanes were used to construct the second dimension retention surface I_5 . Figure 1 depicts the calculation of hold-up time 2t_M , where means of hold-up times at each column temperature were fitted using a logarithm function. The fitting accuracy R^2 was 0.9843. It can be seen that 2t_M decreases with the increase of the second dimension column temperature 2T_c .

To obtain the optimal surface fitting for the construction of the second dimension retention index surface I_S , all combinations of i and j values in the polynomial surface model 'poly ij ' were tested. The performance of each polynomial surface model was then evaluated using the three evaluation parameters including R^2 , R^2_{adj} , and $RMSE$ (Supplementary Table S1). The optimal polynomial model in this study was 'poly25', i.e. $i = 2$ and $j = 5$. R^2 , R^2_{adj} , and $RMSE$ of the optimal surface fitting model 'poly25' were 0.9999, 0.9999, and 6.1 iu, respectively.

Ten-fold cross validation was performed to study how accurately the polynomial surface model 'poly25' would perform in fitting the n -alkane data points. The original 415 data points were randomly partitioned into 10 subsamples, where 9 subsamples had equal sample size (42 data points each) and the 10th subsample had 37 data points. Of the 10 subsamples, a single subsample was retained as the validation data for testing the model, and the remaining 9 subsamples were used as training data. Figure 2 is an example showing that the validation data set was randomly distributed. The cross validation process was repeated 10 times, with each of the 10 subsamples used exactly once as the validation data. The 10 validation results were then averaged to produce a single estimation. Table 1 lists the results of the 10-fold cross validation. The R^2 , R^2_{adj} , and $RMSE$ of the training data set were 0.9999, 0.9999 and 7.1 iu, respectively. The average $RMSE$ for the validation set of data was 16.0 iu. The 10-fold cross validation process was repeated three more times, and the Supplementary Table S2 shows the results of the second repeated trial. The third and fourth repeated trial's average $RMSE$ for the validation set of data was 15.8 iu and 15.7 iu, respectively (data not shown). The similar results of the four trials demonstrate that the polynomial surface model 'poly25' was very robust in fitting the n -alkane data points.

Figure 3A shows the second dimension retention index surface I_S constructed from the isothermal data of the n -alkanes using polynomial surface fitting model 'poly25'. Figure 3B is the residual plot of the surface fitting. A residual from the fitted surface model was defined as the difference between the original retention index I of an n -alkane and its fitted value. The residuals had a near normal distribution (Figure 3C), suggesting that no significant bias was introduced by the polynomial surface model 'poly25'. However, large residuals showed at the lower end of the Y -axis (Figure 3B), which were mainly caused by the inaccuracy of chromatographic peak picking as well as the small values of 2t_R of n -alkanes. Small variation in determining an n -alkane's 2t_R during peak picking resulted in large variation in positioning the data point of that n -alkane in the X - Y plane.

3.3 Calculation of retention index of compounds in MegaMix A & MegaMix B

The mixture of MegaMix A and MegaMix B was analyzed by GC×GC-MS in temperature-programmed mode with the first dimension column started at 4 different initial temperatures, i.e. 60, 80, 100, and 120 °C, respectively. After data processing, 68 compounds were identified from MegaMix A and 41 compounds were identified from MegaMix B. Seven compounds were identified from both MegaMix A and MegaMix B. Pyridine is one of the compounds identified from MegaMix A. While pyridine was identified from the experimental data, it was excluded from the retention index calculation because it is one of

the carry-over compounds frequently detected in our GC×GC-MS system. Tables 2 and 3 list the I values of compounds identified from MegaMix A and MegaMix B, respectively. The standard deviation of I calculated from the four analyses ranged from 1 to 10 iu. The mean standard deviation of I was 1.6 iu for compounds in MegaMix A and 3.4 iu for compounds in MegaMix B. The average RSD for MegaMix A and MegaMix B was 0.10 % and 0.31%, respectively.

Compared with the compounds in MegaMix A, the relatively increased mean standard deviation of I of the compounds in MegaMix B may be related with the retention times of those compounds. The compounds in MegaMix A eluted in the middle to high temperature range whereas the compounds in MegaMix B eluted in the low to mid temperature range. To an individual compound, the large standard deviation of I was mainly caused by the inaccuracy of peak picking, during which the second dimension retention time 2t_R of the compound was determined. For instance, the standard deviation of I of compound 1, 4-dichlorobenzene was 7 iu. While 1, 4-dichlorobenzene was well separated from other compounds during the analysis with the initial temperature of the first dimension column set to 60 °C, its chromatographic peak overlapped with the chromatographic peaks of other compounds during the analysis when the initial temperature of the first dimension column was at 120 °C (Figure 4). Inaccuracy of deconvoluting overlapping chromatographic peaks increased the chance of introducing large variation in determining a compound's second dimension retention time 2t_R and therefore its I value.

Supplementary Tables S3 and S4 list the second dimension column temperature 2T_e and its corresponding adjusted second dimension retention time $^2t'_R$ of each identified compound in MegaMix A and MegaMix B, respectively. The average range of elution temperature of compounds in MegaMix A is 8 °C, and 25 °C for compounds in MegaMix B. Comparing with compounds in MegaMix A, compounds in MegaMix B had larger I variations (Table 3) and their elution temperature ranges were also markedly broader (Table S4) than those of the compounds in MegaMix A (Table 2 and Table S3).

Previous study demonstrated that retention index had temperature dependency, i.e. retention index of an analyte in general increases with the increase of elution temperature ⁽²⁴⁾. We observed that some of the identified compounds followed this trend in this study, while other compounds did not. The reason is that temperature dependency of retention index is also related to the structure of an analyte ^(24–27).

3.4 Robustness of surface fitting to variations of instrumental settings

To investigate the robustness of the developed surface fitting method in analyzing GC×GC-MS data, we analyzed the mixture of MegaMix A and MegaMix B in different instrumental conditions, by varying the temperature offset of the secondary oven, the ramp rate of temperature gradient, and the modulation period. Supplementary Tables S5 and S6 list the second dimension retention indices of compounds identified from MegaMix A and MegaMix B in different instrumental conditions, respectively. An average relative standard deviation (RSD) of the calculated second dimension retention index I for compounds in MegaMix A was 0.63% and 0.42 % for compounds in MegaMix B. The small RSD of

calculated I demonstrates that the developed surface fitting method is robust in calculating the second dimension retention index of compounds analyzed in different instrumental conditions.

As we analyzed the mixture of MegaMix A and MegaMix B in different temperature-programmed modes, the same compound eluted at different column temperatures. In general, the elution temperature of compounds in either MegaMix A or MegaMix B was increased with the increase of temperature ramp rate (data not shown). The low elution temperature corresponded to a ramp rate of 4 °C/min and the high elution temperature corresponded to a ramp rate of 6 °C/min. Therefore, the temperature range listed in Supplementary Tables S5 and S6 was calculated by subtracting the elution temperature of ramp rate 4 °C/min from the elution temperature of ramp rate 6 °C/min. The average range of the elution temperature for compounds in MegaMix A is 15.2 °C and 16.2 °C for compounds in MegaMix B.

3.5. Comparison with literature reported I values

We previously identified 47 compounds from MegaMix A by GC×GC-MS, where a DB-17ms column was used as the secondary column ⁽²³⁾. The second dimension retention index of each identified compound was then calculated from the isovolatility curves of reference compounds, i.e. *n*-alkanes. From these 47 compounds, 44 compounds were identified in this study. Figure 5 shows the overall relation of the I values of the 44 compounds in MegaMix A calculated by surface fitting and by isovolatility curves, respectively. The R^2 value of the fitted curve was 0.9995, indicating a very good agreement between these two methods.

While the I values calculated by the conventional isovolatility curves had a good correlation with those calculated by the surface fitting method, there are some difference between the I values of these two methods. Twenty five compounds had the differences of I values within 5 iu; 8 compounds within 6–10 iu; 5 compounds within 11–15; and 5 compounds within 16–25 iu; one compound had the variation of 36 iu. Compound benzyl butyl phthalate had the largest I difference, i.e. 36 iu. The mean I value of this compound calculated by the isovolatility curve was 2965 iu with standard deviation 2 iu ⁽²³⁾, while its mean I value calculated by the surface fitting approach was 2929 iu with standard deviation 1 iu.

The standard deviations of the retention index of the 44 compounds calculated from the conventional isovolatility curves ranged from 1 – 12 iu with average standard deviation 5.2 iu ⁽²³⁾. However, the standard deviations of the retention index of those 44 compounds calculated by the surface fitting method ranged from 0 – 5 iu with average standard deviation 1.6 iu. While the true values of the retention indices of these compounds remain unknown, the small value of average standard deviation in the calculated retention index using surface fitting method shows that the surface fitting method outperforms the conventional isovolatility curve method by introducing much decreased variation during data analysis.

5. Conclusion

With the increased use of GC×GC-MS for analysis of complex samples, it is crucial to have a robust method to calculate the second dimension retention index I for every compound, to aid the accuracy of compound identification. A surface fitting method was developed to calculate the I values from the second dimension retention time 2t_R and the secondary column temperature 2T_c . In the surface fitting approach, 2t_R and 2T_c of all reference compounds formed an X - Y plane and the I was the Z -axis to form the I surface. The predictability of the surface fitting method was validated using 10-fold cross validation. The average $RMSE$ for the validation set of data was 16.0 iu. Analysis of the GC×GC-MS data acquired from the mixture of MegaMix A and MegaMix B showed that the mean standard deviation of I values calculated by the surface fitting approach was 1.6 iu for compounds in MegaMix A and 3.4 iu for compounds in MegaMix B. Compared with the literature reported results, there was a very good agreement between the I values of the compounds in MegaMix A as calculated by the surface fitting and by the conventional isovolatility curves with R^2 value of 0.9995. However, the small value of standard deviation in the calculated retention index using surface fitting method shows that the surface fitting method has less measurement variability than the isovolatility curve approach.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Mrs. Marion McClain for review of this manuscript. This work was supported by NIH grant nos. 1P20GM113226 (CJM), 1P50AA024337 (CJM), 1U01AA021893-01 (CJM), 1U01AA021901-01 (CJM), 1U01AA022489-01A1 (CJM) and 1R01AA023681-01 (CJM). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was also supported by the Department of Veterans Affairs 1101BX002996-01A2 (CJM).

References

1. Winnike JH, Wei XL, Knagge KJ, Colman SD, Gregory SG, Zhang X. Comparison of GC-MS and GC×GC-MS in the Analysis of Human Serum Samples for Biomarker Discovery. *Journal of Proteome Research*. 2015; 14:1810–1817. [PubMed: 25735966]
2. Shi X, Wei X, Yin X, Wang Y, Zhang M, Zhao C, Zhao H, McClain CJ, Feng W, Zhang X. Hepatic and fecal metabolomic analysis of the effects of *Lactobacillus rhamnosus* GG on alcoholic fatty liver disease in mice. *J Proteome Res*. 2015; 14:1174–1182. [PubMed: 25592873]
3. Shi X, Wei X, Koo I, Schmidt RH, Yin X, Vaughn A, Kim SH, McClain CJ, Arteel GE, Zhang X, Watson WH. Metabolomic analysis of the effects of chronic arsenic exposure in a mouse model of diet-induced fatty liver disease. *J Proteome Res*. 2014; 13:547–554. [PubMed: 24328084]
4. Shi X, Wahlang B, Wei XL, Yin XM, Falkner KC, Prough RA, Kim SH, Mueller EG, McClain CJ, Cave M, Zhang X. Metabolomic Analysis of the Effects of Polychlorinated Biphenyls in Nonalcoholic Fatty Liver Disease. *Journal of Proteome Research*. 2012; 11:3805–3815. [PubMed: 22686559]
5. Han J, Datla R, Chan S, Borchers CH. Mass spectrometry-based technologies for high-throughput metabolomics. *Bioanalysis*. 2009; 1:1665–1684. [PubMed: 21083110]
6. Marriott P, Shellie R. Principles and applications of comprehensive two-dimensional gas chromatography. *TrAC Trends in Analytical Chemistry*. 2002; 21:573–583.

7. Tranchida PQ, Purcaro G, Dugo P, Mondello L, Purcaro G. Modulators for comprehensive two-dimensional gas chromatography. *TrAC Trends in Analytical Chemistry*. 2011; 30:1437–1461.
8. Murray JA. Qualitative and quantitative approaches in comprehensive two-dimensional gas chromatography. *J Chromatogr A*. 2012; 1261:58–68. [PubMed: 22647189]
9. Phillips JB, Beens J. Comprehensive two-dimensional gas chromatography: a hyphenated method with strong coupling between the two dimensions. *J Chromatogr A*. 1999; 856:331–347. [PubMed: 10526795]
10. Ong RCY, Marriott PJ. A Review of Basic Concepts in Comprehensive Two-Dimensional Gas Chromatography. *Journal of Chromatographic Science*. 2002; 40:276–291. [PubMed: 12049157]
11. Adahchour M, Beens J, Vreuls RJJ, Brinkman UAT. Recent developments in comprehensive two-dimensional gas chromatography (GC×GC). *TrAC Trends in Analytical Chemistry*. 2006; 25:821–840.
12. Adahchour M, Beens J, Brinkman UAT. Recent developments in the application of comprehensive two-dimensional gas chromatography. *J Chromatogr A*. 2008; 1186:67–108. [PubMed: 18207152]
13. von Muhlen C, Marriott PJ. Retention indices in comprehensive two-dimensional gas chromatography. *Analytical and bioanalytical chemistry*. 2011; 401:2351–2360. [PubMed: 21826459]
14. Jiang M, Kulsing C, Nolvachai Y, Marriott PJ. Two-Dimensional Retention Indices Improve Component Identification in Comprehensive Two-Dimensional Gas Chromatography of Saffron. *Analytical Chemistry*. 2015; 87:5753–5761. [PubMed: 25936794]
15. Vandendool H, Kratz PD. A Generalization of the Retention Index System Including Linear Temperature Programmed Gas-Liquid Partition Chromatography. *Journal of chromatography*. 1963; 11:463–471. [PubMed: 14062605]
16. Kováts E. Gas-chromatographische Charakterisierung organischer Verbindungen. Teil 1: Retentionsindices aliphatischer Halogenide, Alkohole, Aldehyde und Ketone. *Helvetica Chimica Acta*. 1958; 41:1915–1932.
17. Beens J, Tijssen R, Blomberg J. Comprehensive Two-Dimensional Gas Chromatography (GC×GC) as a Diagnostic Tool. *Journal of High Resolution Chromatography*. 1998; 21:63–64.
18. Beens J, Tijssen R, Blomberg J. Prediction of comprehensive two-dimensional gas chromatographic separations. *J Chromatogr A*. 1998; 822:233–251.
19. Western RJ, Marriott PJ. Retention correlation maps in comprehensive two-dimensional gas chromatography. *Journal of Separation Science*. 2002; 25:831–838.
20. Western RJ, Marriott PJ. Methods for generating second dimension retention index data in comprehensive two-dimensional gas chromatography. *J Chromatogr A*. 2003; 1019:3–14. [PubMed: 14650600]
21. Bieri S, Marriott PJ. Dual-Injection System with Multiple Injections for Determining Bidimensional Retention Indexes in Comprehensive Two-Dimensional Gas Chromatography. *Analytical Chemistry*. 2008; 80:760–768. [PubMed: 18177017]
22. Bieri S, Marriott PJ. Generating Multiple Independent Retention Index Data in Dual-secondary Column Comprehensive Two-dimensional Gas Chromatography. *Analytical Chemistry*. 2006; 78:8089–8097. [PubMed: 17134144]
23. Zhao Y, Zhang J, Wang B, Kim SH, Fang A, Bogdanov B, Zhou Z, McClain C, Zhang X. A method of calculating the second dimension retention index in comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry. *J Chromatogr A*. 2011; 1218:2577–2583. [PubMed: 21429495]
24. Pacáková, V., Felzl, L. Chromatographic retention indices: an aid to identification of organic compounds. Ellis Horwood Ltd; 1992.
25. Rang S, Kuningas K, Orav A, Eisen O. Capillary gas chromatography of n-alkynes: II. Variation of retention indices with temperature. *J Chromatogr A*. 1976; 128:53–58.
26. Hala S, Eyem J, Burkhard J, Landa S. Retention Indices of Adamantanes. *Journal of Chromatographic Science*. 1970; 8:203–209.
27. Döring CE, Estel D, Weber J, Zimmermann G, Zschummel D. Kapillar-gaschromatographische Untersuchungen an homo- und codimeren C5-Dienen. –Einige Zusammenhänge zwischen

Retentionsverhalten und Struktur bzw. Konfiguration. Journal für Praktische Chemie. 1971; 313:1081–1091.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- Surface fitting for calculation the secondary retention index in GC×GC-MS
- Isovolatility curve construction is not needed
- Retention index variation was significantly reduced
- The proposed method was validated using a mixture of compound standards

Figure 1.

The second dimension hold-up time calculation. The means of hold-up times at each column temperature were fitted using a logarithm function.

Figure 2. Distribution of one validation data set used in 10-fold cross validation. The filled circles are the training data set and the open circles are the validation data set.

A. **B.**

C.

Figure 3.

The second dimension retention index surface constructed using the GC×GC-MS data of *n*-alkanes. The retention index surface is a function of the elution temperature of each *n*-alkane in the second dimension column 2T_e and the corresponding adjusted retention time 2t_R . (A) is the second dimension retention index surface. (B) is the residual plot, and (C) is the distribution of the residuals.

Figure 4.

GC×GC-MS separation of 1, 4-dichlorobenzene (black arrow) and other compounds in MegaMix A analyzed in temperature-programmed mode at different column temperatures. With the increase of the second dimension column temperature from 65 °C to 125 °C, the separation between 1, 4-dichlorobenzene and other compounds became worse and resulted in chromatographic peak overlapping. The overlapping chromatographic peaks introduce a high degree of variation in the retention time and therefore the calculated retention index. (A) are the three-dimensional images of the two dimensional chromatographic peaks, and (B) are the corresponding contour images.

Figure 5.
Correlation of second dimension retention index of 44 compounds calculated by conventional isovolatility curves and by surface fitting approach.

Table 1

Results of a ten-fold cross validation of surface fitting using the second dimension retention indices of *n*-alkanes

Fold count	Goodness of fit			Goodness of validation
	R^2	R_{adj}^2	$RMSE$	$RMSE$
1	0.9999	0.9999	6.3	17.2
2	0.9999	0.9999	6.2	15.7
3	0.9999	0.9999	6.3	17.3
4	0.9999	0.9999	6.3	17.1
5	0.9999	0.9999	10.1	16.0
6	0.9999	0.9999	6.3	16.4
7	0.9999	0.9999	10.6	14.5
8	0.9999	0.9999	6.4	17.3
9	0.9999	0.9999	6.1	17.0
10	0.9999	0.9999	6.4	11.9
Average	0.9999	0.9999	7.1	16.0

Table 2

The second dimension retention indices of compounds identified from MegaMix A by surface fitting

Compound name	Second dimension retention index (iu) ^a				Mean (iu)	STD (iu)	RSD	
	65 °C	85 °C	105 °C	125 °C			STD (%)	(%)
N-Nitrosodimethylamine	983	977	977	979	3	0.30		
Aniline	1250	1252	1251	1250	1251	1	0.08	
Phenol	1205	1202	1199	1199	1201	3	0.22	
Bis(2-chloroethyl) ether	1219	1216	1212	1208	1214	4	0.33	
2-Chlorophenol	1217	1213	1216	1219	1216	2	0.17	
Benzyl alcohol	1297	1299	1300		1298	1	0.09	
2,2'-Oxybis(2-chloropropane)	1247	1245	1242	1241	1244	3	0.21	
2-Methylphenol	1290	1290	1289	1286	1289	2	0.13	
N-nitrosodipropylamine	1305	1303	1306	1295	1302	4	0.32	
3-Methylphenol	1313	1314	1315	1309	1313	2	0.17	
Hexachloroethane	1276	1275	1272	1277	1275	2	0.13	
Isophorone	1374	1375	1373	1373	1374	1	0.05	
2-Nitrophenol	1418	1421	1424	1426	1422	3	0.22	
2,4-Dimethylphenol	1397	1397	1397	1394	1396	1	0.08	
Bis(2-chloroethoxy)-Methane	1422	1424	1422	1418	1422	2	0.14	
2,4-Dichlorophenol	1435	1434	1435	1436	1435	1	0.05	
4-Chloroaniline	1537	1538	1542	1549	1542	5	0.29	
4-Chloro-3-methylphenol	1588	1587	1587	1591	1588	2	0.10	
2-Methylnaphthalene	1609	1608	1611	1615	1611	3	0.16	
1-Methylnaphthalene	1641	1643	1645	1649	1644	3	0.19	
1,3-Cyclopentadiene, 1,2,3,4,5,5-hexachloro-	1574	1573	1572	1573	1573	1	0.05	
2,4,6-Trichlorophenol	1666	1665	1667	1668	1666	1	0.05	
2,4,5-Trichlorophenol	1674	1674	1673	1678	1675	2	0.13	
2-Chloronaphthalene	1726	1727	1729	1733	1729	3	0.15	
2-Nitroaniline	1835	1836	1838	1843	1838	3	0.16	

Compound name	Second dimension retention index (iu) ^a				Mean (iu)	STD (iu)	RSD (%)
	65 °C	85 °C	105 °C	125 °C			
1,4-Dinitrobenzene	1876	1876	1877	1880	1877	2	0.09
1,3-Dinitrobenzene	1888	1887	1889	1891	1889	2	0.08
Dimethyl phthalate	1843	1843	1843	1847	1844	2	0.10
2,6-Dinitrotoluene	1884	1884	1884	1888	1885	2	0.10
1,2-Dinitrobenzene	1932	1930	1932	1934	1932	1	0.07
Acenaphthylene	1852	1852	1853	1858	1854	2	0.13
3-Nitroaniline	1952	1952	1952	1958	1954	3	0.14
Acenaphthene	1884	1884	1885	1888	1885	2	0.08
2,4-dinitro-phenol	1945	1944	1943	1945	1944	1	0.05
Dibenzofuran	1915	1915	1915	1920	1916	2	0.10
2,4-Dinitrotoluene	1969	1968	1968	1968	1968	1	0.03
2,3,5,6-Tetrachlorophenol	1923	1923	1922	1924	1923	1	0.04
2,3,4,6-Tetrachlorophenol	1937	1936	1937	1938	1937	1	0.05
Diethyl Phthalate	1985	1983	1982	1983	1983	1	0.05
Fluorene	2008	2009	2008	2010	2009	1	0.05
4-Chlorophenyl phenyl ether	1980	1980	1979	1981	1980	1	0.04
4-Nitroaniline	2142	2143	214	2143	2142	1	0.04
4,6-Dinitro-2-methylphenol	2051	2051	2050	2051	2051	1	0.03
Diphenylamine	2063	2062	2064	2064	2063	1	0.04
Azobenzene	2035	2035	2035	2036	2035	0	0.02
4-Bromophenyl phenyl ether	2124	2121	2122	2123	2122	1	0.05
Hexachlorobenzene	2113	2110	2113	2113	2112	1	0.05
Pentachloro-phenol	2203	2200	2202	2201	2202	1	0.04
Anthracene	2299	2299	2299	2300	2300	0	0.01
Phenanthrene	2309	2307	2310	2310	2309	1	0.06
Carbazole	2414	2413	2411	2414	2413	1	0.04
di-n-Butyl phthalate	2371	2371	2370	2369	2370	1	0.04
Fluoranthene	2678	2679	2678	2677	2678	1	0.02

Compound name	Second dimension retention index (iu) ^a				Mean (iu)	STD (iu)	RSD (%)
	65 °C	85 °C	105 °C	125 °C			
Pyrene	2773	2773	2773	2775	2774	1	0.03
Benzyl butyl phthalate	2932	2933	2931	2929	2931	1	0.05
Bis(2-ethylhexyl) adipate	2643	2645	2641	2645	2643	2	0.07
Benz[a]anthracene	3200	3201	3202	3200	3201	1	0.03
Chrysene	3234	3234	3234	3232	3233	1	0.02
Bis(2-ethylhexyl) phthalate	2930	2928	2932	2932	2931	1	0.05
di-n-Octyl phthalate	3076	3081	3074	3079	3077	3	0.10

^aThe blank means the compound was not detected during the experiment.

Table 3

The second dimension retention indices of compounds identified from MegaMix B

Compound name	Second dimension retention index (iu) ^d				Mean (iu)	STD (iu)	RSD (%)	
	65 °C	85 °C	105 °C	125 °C				
cis-1,3-Dichloro-1-Propene, (Z)	888	891			890	2		0.19
trans-1,3-Dichloro-Propene(E)	927	924			925	1		0.15
Toluene	899				899	0		0.00
1,1,2-Trichloroethane	937	941			939	2		0.22
Ethyl methacrylate	906	915			910	5		0.50
1,3-Dichloropropane	954	950	966		956	7		0.70
Dibromochloro-Methane	987	984	994		988	4		0.42
Tetrachloroethene	927	931	950		936	10		1.09
1,2-Dibromoethane	1001	994	1007		1000	5		0.51
Chlorobenzene	1016	1013	1022		1017	4		0.37
1,1,1,2-Tetrachloroethane	1012	1002			1007	5		0.47
Ethylbenzene	1008	1005			1006	1		0.10
p-Xylene	1014	1015	1019		1016	2		0.19
Styrene	1072	1070	1075	1075	1073	2		0.19
o-Xylene	1054	1052	1058	1046	1053	4		0.41
Bromoforn	1129	1126	1125	1140	1130	6		0.54
cis-1,4-Dichloro-2-butene,, (z)-	1114	1111	1106	1118	1112	4		0.38
1,1,2,2-Tetrachloroethane	1119	1115	1110	1120	1116	4		0.33
1,2,3-Trichloropropane	1140	1136	1131	1136	1136	3		0.27
Isopropylbenzene(cumene)	1072	1070	1070	1088	1075	8		0.72
trans-1,4-Dichloro-2-butene, (E)-	1153	1149	1139	1147	1147	5		0.42
Bromobenzene	1140	1138	1135	1145	1140	4		0.32
2-Chlorotoluene	1141	1140			1141	0		0.01
n- Propylbenzene	1110	1108	1105		1108	2		0.15
4-Chlorotoluene	1142	1142	1141	1149	1143	3		0.29

Compound name	Second dimension retention index (iu) ^d				Mean (iu)	STD (iu)	RSD (%)
	65 °C	85 °C	105 °C	125 °C			
1,3,5-Trimethylbenzene	1123	1121	1120	1131	1124	4	0.39
Pentachloroethane	1172	1168	1165	1182	1172	6	0.53
tert-Butylbenzene	1148	1147	1146	1151	1148	2	0.18
1,2,4-Trimethylbenzene	1164	1160	1157	1167	1162	4	0.34
sec-Butylbenzene	1167	1167	1165	1169	1167	1	0.12
p-Cymene(4-isopropyl toluene)	1182	1179	1177	1186	1181	3	0.30
n-Butylbenzene	1221	1220	1214	1218	1218	3	0.21
1,2-Dibromo-3-Chloropropane	1364	1367	1366	1369	1366	2	0.13
Benzene, 1,2,3-trichloro-	1494	1495	1497	1500	1497	2	0.15
1,3-Dichlorobenzene *	1217	1215	1214		1215	1	0.09
1,4-Dichlorobenzene *	1232	1231	1227	1213	1226	7	0.59
1,2-Dichlorobenzene *	1267	1267	1268	1269	1267	1	0.07
Nitrobenzene *	1381	1385	1389	1391	1386	4	0.27
Benzene, 1,2,4-trichloro- *	1434	1436	1437	1437	1436	1	0.07
Naphthalene *	1485	1488	1490	1496	1490	4	0.27
1,3-Butadiene, 1,1,2,3,4,4-hexachloro- *	1417	1417	1416	1415	1416	1	0.06

^dThe blank means the compound was not detected during the experiment.

* Compounds detected in both MegaMix A and MegaMix B.