PAPER • OPEN ACCESS

Kovats Retention Index analysis of flavor and fragrance compound using Biplot Statistical method in gas chromatography systems

View the article online for updates and enhancements.

Kovats Retention Index analysis of flavor and fragrance compound using Biplot Statistical method in gas chromatography systems

R Idroes[1,2], A F Japnur[1], R Suhendra[3], A Rusyana[4,*]

[1]Department of Chemistry, Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Kopelma Darussalam, Banda Aceh 23111, Indonesia
[2]Department of Pharmacy, Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Kopelma Darussalam, Banda Aceh 23111, Indonesia
[3]Department of Informatics, Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Kopelma Darussalam, Banda Aceh 23111, Indonesia
[4]Department of Statistics, Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Kopelma Darussalam, Banda Aceh 23111, Indonesia

E-mail: asep.rusyana@unsyiah.ac.id

Abstract. A lot of retention parameters have been introduced and used in chromatographic instrumentation analysis, one of which is the retention index. Retention index is a secondary parameter which in its calculation requires data retention time and dead time. This study uses data from flavor and fragrance retention indices with a variation of columns, boiling points and log Kow which aims to identify the level of similarity between compounds based on retention index using biplot statistical method. Biplot analysis is used to view object relationships based on the variables. The results showed that biplot could describe the relationship of a compound as an object and its retention index as a variable of 99.45%. Based on the results, the compounds can be classified into three groups. The first group has a specific retention index that is lower than other groups. The second group has a high Log Kow whereas the third group is dominated by the influence of DB-Wax column compared to other columns. Columns DB-1, OV-101, and DB-5 have a positive correlation with the correlation level close to 1.00.

1. Introduction
Various parameters have been introduced and are used in the chromatographic analysis which can explain the analytical aspects of the data. The most commonly used parameter for explaining the data of the analysis results is retention factor. It can be used for both qualitative and quantitative analysis [1]. There are many retention parameters of chromatography, but only the retention time and dead time parameters as the primary parameters generated by the chromatography instrument, while the rest are secondary parameters [2].

The retention time is relatively sensitive to slight changes in experimental conditions such as column changes, column length, column diameter, carrier gas, flow rate, pre-column, and more, while

---

*Corresponding author : asep.rusyana@unsyiah.ac.id

the retention index Kovats is more reliable and reproducible. The Kovats retention index is a concept introduced in gas chromatography to convert retention time into a more reliable and reproducible system. This retention index was first introduced by E. Kovats in 1958 on a gas chromatographic system experiment. The retention index of a compound is calculated relatively to two standard compounds so that the effect of the instrument can be ignored and the precision can be increased [3]. Thus in this study, the data from retention index of flavor and fragrance compounds with a variation of the column is used to identify the level of similarity among compounds based on the pattern of Kovats retention index data using biplot statistic method.

2. Literature Review

Biplot, a two-dimensional mapping of the principal component analysis (PCA), was firstly introduced by Gabriel (1971), thus it is often called as Gabriel's biplot or PCA Biplot. Biplot is a descriptive statistical technique that is useful for simultaneously presenting observational objects and p variants in flat space, so that the characteristics of the variables, observation objects and relative positions between observed objects and variables are able to be analyzed visually. Biplot refers to two types of information contained in the data matrix. The information in the row corresponds to the sample and in the column to the variable [4].

The information obtained from Biplot is the degree of relative similarity among observation objects, the relationship among variables, the value of the variables in an object, and the variability of the variables. The relationship between the variables is that if the angle of two variables < 90º then the correlation is positive, if the angle of two variables > 90º the correlation is negative, and if the smaller the angle is, the stronger the correlation will become. An object has the property of a variable if the object is adjacent to that variable. The closer the position of the object to the variable, the more dominant character of the variable on the object [5].

Biplot was developed based on Singular Value Decomposition (SVD). Suppose an X data matrix of size n x p containing n observations and p variables corrected to its mean and r-value can be written as:

$$X = ULA' \tag{1}$$

where: the matrix U dan A each sized (nxr) and (pxr) so that $U'U = A'A = I_r$. L is a diagonal matrix size (rxr) with its diagonal elements is the square root of eigen value of $X'X$ or XX' so:

$$\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \cdots \geq \sqrt{\lambda_r} \tag{2}$$

Column matrix A is an eigen vector corresponding to the eigen value λ from matrix $X'X$ or XX'. Matrix columns U can be calculated through:

$$U_i = \frac{1}{\sqrt{\lambda_i}} \times \alpha_i \tag{3}$$

with λ_i is eigen value i-th from matrix $X'X$ or XX', and α_i is column i-th matrix A.

$$X = UL^\lambda L^{1-\lambda} A' \tag{4}$$

if $G = UL^\lambda$ and $H' = L^{1-\lambda}A'$, so:

$$X = GH' \tag{5}$$

factor (i,j)-th matrix X can be written:

$$X_{ij} = g_i'h_j' \tag{6}$$

where: $i = 1,2,3, \dots, n$ and $j = 1,2,3, \dots, p$ with g_i' and h_j' each is a matrix row G and H'.

If $r(X) = 2$ so g_i' and h_j' depicted in a 2-dimensional space.

The percentage of diversity displayed by biplot can be defined as follows:

$$\rho_2 = (\lambda_1 + \lambda_2) \Big/ \sum_{i=1}^{r} \lambda_i \tag{7}$$

where:
λ_1 = the first largest eigen value
λ_2 = the second largest eigen value
λ_i = the largest eigen value i-th from $X'X$, $i = 1,2,3, ..., r$., [4]

3. Methods

3.1 Data Collection
In this study, the researchers took data on retention index of homologous n-alkane and flavor and fragrance compounds as chromatographed using OV-101, DB-1, DB-5, and DB-Wax columns, as well as Boiling Point (BP) and Log Kow data from the literature of Goodner K.L. 2008, *Practical retention index models of OV-101, DB-1, DB-5, and DB-Wax for flavor and fragrance compounds*, *LWT - Food Sci. Technol.*, vol. 41, no. 6, pp. 951–958 [6].

The columns used in this study were OV-101 column with dimethylpolysiloxane, DB-1 with 100% -dimethylpolysiloxane, DB-5 with (5%-Phenyl)-methylpolysiloxane and DB-Wax with polyethylene glycol as stationary phase respectively. The column information used is obtained from https://www.agilent.com/en/products/gas-chromatography/gc-columns.

3.2 Retention index determination
Retention index determination can be performed with a simple interpolation between the logarithm of the n-alkane compounds before and after being eluted solutes, at each temperature is defined as 100 times n_c (carbon number) [7,8].

3.3 Data Analysis
The data which have been collected, are initially converted or transferred into XLSTAT software, then the data outliers are removed. The next stage of preprocessing performed on the data to eliminate variations in data that are not related to analytical information. After preprocessing, the data analysis using biplot method is used to obtain data classification model.

4. Result and Discussion
Data on the retention index of n-alkane, flavor, and fragrance compounds with variations of OV-101, DB-1, DB-5, and DB-Wax columns, as well as Boiling Point and Log Kow data were used to identify similarity levels based on the pattern of Kovats retention index data using the Biplot statistics method.

Biplot statistical method was used to identify the similarity level up to 99.45%, PC1 87.79% and PC2 11.66% (figure 1). The result of Biplot which is based on the pattern of the compound retention index, can be grouped into three groups.

The first group (I) consists of relatively shorter carbon n-alkane compounds (pentane, hexane, heptane, octane, nonana, and decane). The n-alkane compounds or homologous series of n-alkanes are often used in the determination of kovats retention index because they have robust retention index results. This is based on the retention index calculation process performed by simple interpolation logarithms between the two n-alkana, before and after eluted solutes, at any temperature defined as 100 times nc (carbon number), such as Retention Index for Ethane = 200, Butane = 400, Hexane = 600, and so on [9,10].
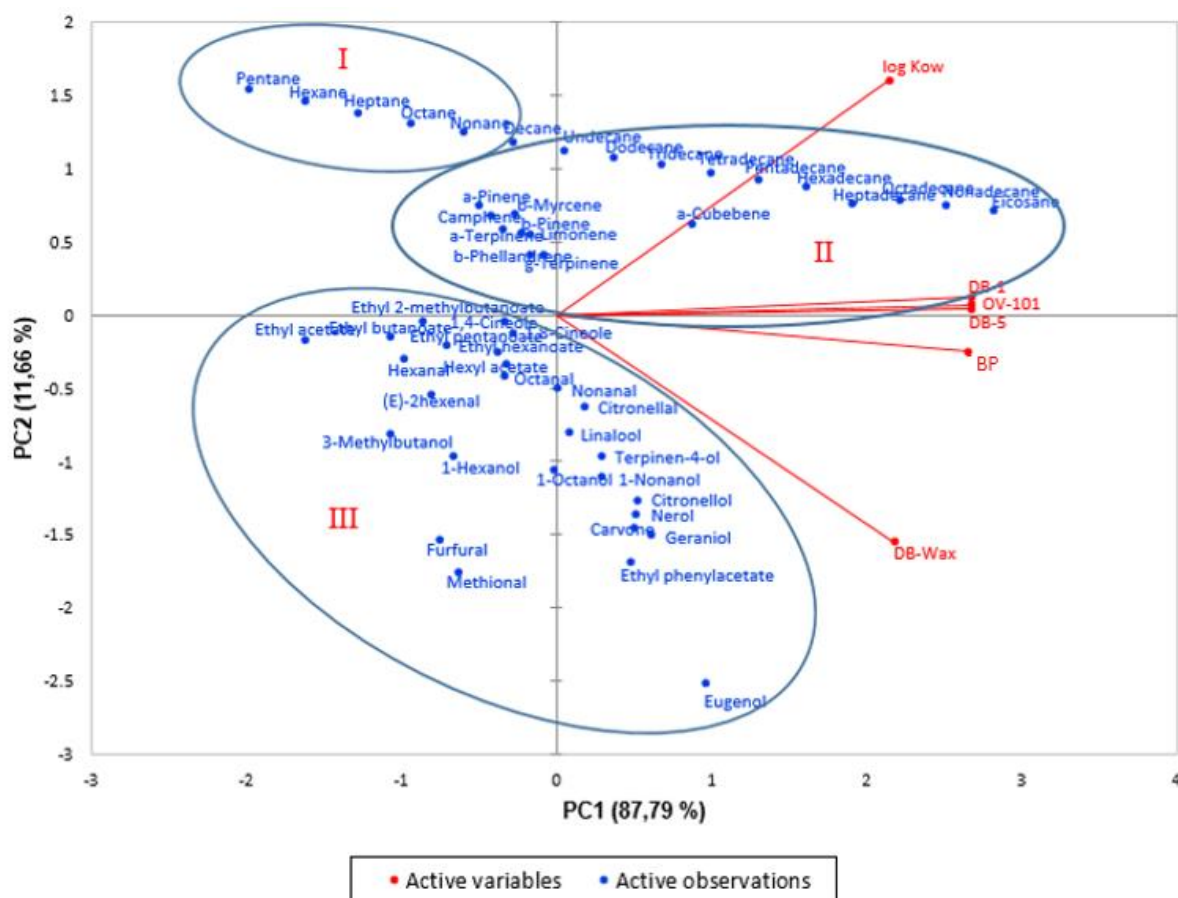
Figure 1. Biplot analysis results of retention index data of flavor and fragrance compounds

The second group (II) has a higher Log Kow value than the others. Based on the retention index pattern, the compounds in this group consist of an n-alkane compound with a higher carbon chain (undecane, dodecane, tridecane, tetradecane, pentadecane, hexadecane, heptadecane, octececane, nonadecane, and eicosane) as the result, these compounds have higher retention index values as opposed to the first group. The terpenoid compounds comprising (a-pinen, b-mycenen, b-pinen, camphen, a-terpenen, limonene, b-phellandene, g-terpenen, and a-cubebene) are considered to have a degree of polarity that is close to the n-alkane homologous series with a higher carbon chain.

The third group (III) comprises alcohol group compounds (linalool, 3-methylbutanol, 1-hexanol, 1-octanol, 4-ol, 1-nonanol, citronellol, nerol, geraniol, and eugenol), carboxylic acids (ethyl butanoate, ethyl 2- methylbutanoate, ethyl pentanoate, ethyl hexanoate), esters (ethyl acetate, hexyl acetate, ethyl phenylacetate), and aldehydes (hexanal, octanal, furfural, methional, nonanal). This group has the dominant similarity to the pattern of retention index data generated from the DB-Wax column. This is because groupings of these compound groups have a higher retention index than other compounds in the DB-Wax column.

Table 1. Correlation between variables

Variables	OV-101	DB-1	DB-5	DB-Wax	BP	log Kow
OV-101	1	0,998	0,999	0,794	0,983	0,811
DB-1	0,998	1	0,999	0,782	0,981	0,821
DB-5	0,999	0,999	1	0,801	0,986	0,803
DB-Wax	0,794	0,782	0,801	1	0,855	0,312
BP	0,983	0,981	0,986	0,855	1	0,734
log Kow	0,811	0,821	0,803	0,312	0,734	1

Table 2. The standard deviation of retention index compound using OV-101, DB-1, DB-5, and DB-Wax columns

			SDV Retention Index		
Compound	SDV^1	SDV^2	Compound	SDV^1	SDV^2
(E)-2hexenal	149,67	21,20	1,8-Cineole	102,26	11,37
Ethyl acetate	147,72	9,85	Limonene	93,43	8,50
Ethyl butanoate	142,34	54,00	g-Terpinene	101,38	25,79
3-Methylbutanol	238,47	9,87	1-Octanol	248,63	10,00
Hexanal	148,89	10,07	Citronellal	174,79	12,29
Ethyl 2-methylbutanoate	108,44	5,86	Terpinen-4-ol	216,23	12,17
Ethyl pentanoate	126,39	12,12	1-Nonanol	247,35	3,79
a-Pinene	52,13	6,81	Citronellol	269,20	5,03
1-Hexanol	252,96	9,61	Ethyl phenylacetate	276,88	17,69
Camphene	62,52	2,00	Carvone	267,69	20,53
Furfural	308,68	30,62	Geraniol	299,23	14,42
b-Pinene	67,70	2,52	a-Cubebene	64,69	16,17
Methional	294,45	26,31	Eugenol	403,11	23,03
b-Myrcene	83,83	0,58	1,4-Cineole	92,50	11,79
Ethyl hexanoate	121,35	11,27	Linalool	224,93	8,14
Hexyl acetate	131,43	10,12	Nerol	283,49	11,72
Octanal	150,10	10,97	Nonanal	151,39	10,02
a-Terpinene	91,92	12,66	b-Phellandrene	115,00	31,09

[SDV1] Standard deviation from the column retention index of OV-101, DB-1, DB-5, and DB-Wax
[SDV2] Standard deviation from the column retention index of OV-101, DB-1, and DB-5

    DB-1, OV-101, and DB-5 columns have a positive correlation among them because they are close to each other (see Figure 1). The data of Table 1 shows the column correlation is close to 1. The DB-1, OV-101, and DB-5 columns are relatively non-polar so they have a higher similarity level, whereas DB-Wax column has different polarity. Based on the data Table 1, DB-Wax column has a lower correlation level against other columns.

    From the Table 2, we could identify that the standard deviation of retention index compound using DB-Wax column is greater than without using it, because the data of retention index compound using DB-Wax column is strongly influenced by the column broadening effect.

5. Conclusion

Based on kovats retention index data, flavor and fragrance compounds used and OV-101, DB-1, DB-5, and DB-Wax columns, as well as boiling point and Kow log can be grouped into three groups using the Biplot statistics method. The first group consists of the n-alkanes group with lower carbon numbers and a lower retention index. The second group consists of the n-alkanes group with higher carbon numbers and terpenoid group compounds that have a higher Kow log values. The third group consists of alcohols, aldehydes, esters and carboxylic acids which are dominantly the retention index pattern in the DB-Wax column. Columns DB-1, OV-101, and DB-5 have a positive correlation close to 1.00.

6. References

[1]   Domínguez J G and Diez-Masa J 2001 *Pure Appl. Chem.* **73** 969–92
[2]   Furr H C 1989 *J. Chromatogr. Sci.* **27** 216–20
[3]   Mahmudi, Saiful, khairan and Idroes R. 2015 *J. Nat.* **15** 1–9
[4]   Johnson R A and Wichern D W 2007 *Applied Multivariate Statistical Analysis*
[5]   Gabriel K R 1971 *Biometrika* **58** 453–67
[6]   Goodner K L 2008 *Food Sci. Technol.* **41** 951–8
[7]   Idroes R 2009 *Maj. Farm. Indonesia.* **20** 133–40
[8]   Pratiwi S U T, Lagendijk E L, de Weert S, Idroes R, Hertiani T and Van den Hondel C 2015
         Int. J. Appl. Res. Nat. Prod. **8** 1–13
[9]   Idroes R 2005 *Malaysian Journal of Analytical Science.* **9** 224–2
[10]   Idroes R 2010 *Indonesian Journal of Chemistry.* **1** 58–10