

# The Choice of Polar Stationary Phases for Gas/Liquid Chromatography by Statistical Analysis of Retention Data

Ervin sz. Kováts and Stephan Morgenthaler

*École Polytechnique Fédérale de Lausanne, FSB-IMA, Station 8, CH - 1015 Lausanne, Switzerland*

---

## Abstract

We disposed of retention indices of 127 volatile substances on a C<sub>78</sub>-paraffin and on its seven nearly isochor and isomorphous polar derivatives in a temperature range of 90-210°C. The retention index of a substance on the C<sub>78</sub> paraffin has been considered as the standard. The additional retention on the polar derivative was given by the difference of its retention index on the polar solvent and the C<sub>78</sub>-paraffin. Statistical analyses of the additional retention have shown that concerning retention, the seven polar solvents can be classified into three groups: type I: TTF (tetrakis(trifluoromethyl)), MTF (monotrifluoromethyl), type II: PCN (primary cyano), PSH (primary thiol) and type III: TMO (tetramethoxy), SOH (sec. alcohol), POH (primary alcohol). It is shown that the three types are best represented by the solvents TTF, PCN and TMO. PSH (primary thiol) is aligned with PCN at temperatures up to about 150°C, but is similar to TMO at 210°C.

*Keywords:* Gas/liquid partition data; Retention indices on polar stationary phases

---

## Introduction

Gas/liquid retention data have been determined on a family of nearly isochor and isomorphous liquid stationary phases shown in Table 1 [1 - 4]. Corrected retention indices and standard chemical potential differences related to molal Henry coefficients (Standard chemical POTential difference ≡ SPOT) of 127 solutes have been determined on all stationary phases. The data have been published on the branched hydrocarbon stationary phase of the carbon number 78, called "alkane C<sub>78</sub>", of the melting point of about 80°C and on seven thermally stable polar derivatives of this branched paraffin (see Table 1 and [5]). Because of a lack of thermal stability, the solvent with bromo derivative was left out of the analysis reported in this paper. Data in the polar solvents are given relative to those measured in the solvent C<sub>78</sub>. Latter has obviously no polar interacting group. As originally published, the determined indices contained small errors, especially at higher temperatures. The data have been corrected and are published in [5]. They can be downloaded from [doi:10.1016/j.chroma.2006.01.123].

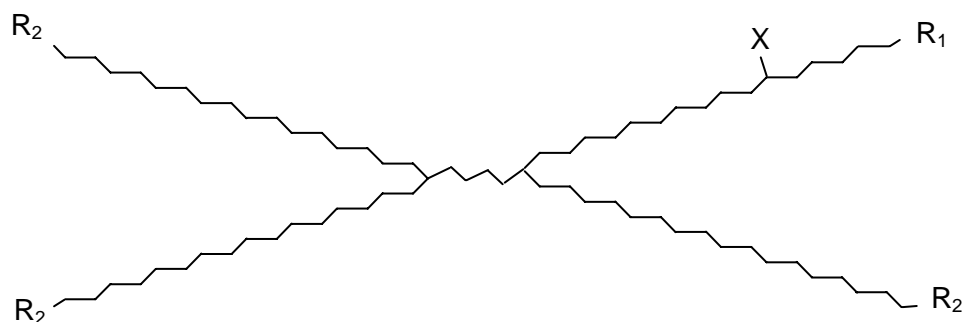
The retention index of the  $j^{\text{th}}$  substance,  $I_j$ , is calculated by interpolation of the logarithm of its net retention time,  $\lg(t_{N,j})$ , between those of the neighbouring  $n$ -alkanes measured under identical experimental conditions. It is now easy to show that in the case of ideal gas/liquid chromatography (no

adsorption) this is equivalent to the interpolation of the corresponding SPOT of a solute,  $\Delta\mu_j$ , between the SPOT-s of the two neighboring *n*-alkanes,  $\Delta\mu_z$  and  $\Delta\mu_{z+1}$ , of carbon number, *z* and *z*+1:

$$I_j = 100 \left[ \frac{\lg(t_{N,j}) - \lg(t_{N,z})}{\lg(t_{N,z+1}) - \lg(t_{N,z})} \right] + 100z = 100 \left[ \frac{\Delta\mu_j - \Delta\mu_z}{\Delta\mu_{z+1} - \Delta\mu_z} \right] + 100z \quad (1)$$

where the retention times  $t_{N,j}$ ,  $t_{N,z}$  and  $t_{N,z+1}$  are at the same temperature, *T*. At a given temperature the retention index of a solute designates on a logarithmic scale the place where a hypothetical *n*-alkane with the carbon number  $I_j/100$  would appear. Obviously, the retention index of a *n*-alkane of the carbon number, *z*, is defined at all temperatures as  $I_z=100z$ .

Table 1: Structure of the paraffin hydrocarbon  $C_{78}$  and its seven polar derivatives.



P	Polar interacting group(s)	R <sub>1</sub>	R <sub>2</sub>	X	Ref.	Corrected Indices
C <sub>78</sub>	-	CH <sub>2</sub> CH <sub>3</sub>	CH <sub>2</sub> CH <sub>3</sub>	H	[1]	[5]
TTF	Tetrakis(trifluoromethyl)	CH <sub>2</sub> CF <sub>3</sub>	CH <sub>2</sub> CF <sub>3</sub>	H	[2]	[5]
MTF	Monotrifluoromethyl	CH <sub>2</sub> CF <sub>3</sub>	CH <sub>2</sub> CH <sub>3</sub>	H	[2]	[5]
PCN	Primary cyano	CH <sub>2</sub> CN	CH <sub>2</sub> CH <sub>3</sub>	H	[3]	[5]
PSH	Primary thiol	CH <sub>2</sub> SH	CH <sub>2</sub> CH <sub>3</sub>	H	[3]	[5]
TMO	Tetramethoxy	OCH <sub>3</sub>	OCH <sub>3</sub>	H	[3]	[5]
SOH	Secondary alcohol	CH <sub>3</sub>	CH <sub>2</sub> CH <sub>3</sub>	OH	[4]	[5]
POH	Primary alcohol	CH <sub>2</sub> OH	CH <sub>2</sub> CH <sub>3</sub>	H	[1]	[5]

The temperature dependence of the retention index on a stationary phase P is given by the quadratic equation

$$I_j^P(T) = I_j^P(T^\dagger) + A_{T,j}^P \Delta T + A_{TT,j}^P \Delta T^2 \quad (2)$$

where the temperature  $T^\dagger = 150^\circ\text{C}$  is chosen in the middle of the temperature range of the determination of the experimental data and  $\Delta T = T - T^\dagger$ . The temperature dependence of the index is in general small and nearly linear. In our analysis, the indices determined in the polar derivative, are for each stationary phase P given as differences

$$\Delta I_j^P(T^\dagger) = I_j^P(T^\dagger) - I_j^{C_{78}}(T^\dagger) \quad (3)$$

where the symbol,  $\Delta$ , designates a difference with reference of the standard paraffin  $C_{78}$ .

A statistical analysis of the original data was published in [6]. It had the aim to predict the boiling point and certain distribution coefficients of the solutes. The corrections of the retention data at high temperatures are of an order that they do not invalidate the conclusions of [6]. The aim of our paper is to choose those polar phases that are different and represent all polar phase information.

## Retention on the Polar Stationary Phases

### *Scatter Plots of the Additional Retention*

We put now forward the question if all polar stationary phases assure a different characterization of a mixture. Indeed, the  $\Delta I$  – values are different on all phases, but they can be correlated. A good correlation can be awaited between data on the two trifluoro substituted phases. The first panel in the first row of Fig. 2 shows that the additional retention on the tetrakis(trifluoromethyl) phase is systematically about 3.3 times higher than those on the monotrifluoromethyl derivative. This means, that the chromatogram on the MTF-phase can be reproduced by a chromatogram of a TTF-phase diluted by the  $C_{78}$ -hydrocarbon.

The individual panels in Fig. 1 show scatter plots of the additional retention in one polar phase as a function of the additional retention in a second polar phase. Since we have seven polar phases, constructing all these scatter plots gives a 7 by 7 matrix of panels. The first row of Fig. 2 shows the additional retention in TTF as a function of the additional retention in the other six stationary phases. The second row does the same for MTF, and so on. The slope for predicting TTF with the help of MTF is 3.45. Inversely, when explaining MTF with the help of TTF, or the ratio MTF/TTF, the slope is  $1/3.45 = 0.29$ . The other slopes are as follows: TTF/PCN = 1.14, TTF/PSH = 4.16, TTF/TMO = 0.99, TTF/SOH = 2.06, TTF/POH = 1.41, MTF/PCN = 0.31, MTF/PSH = 1.04, MTF/TMO = 0.25, MTF/SOH = 0.51, MTF/POH = 0.34, PCN/PSH = 3.28, PCN/TMO = 0.87, PCN/SOH = 1.7, PCN/POH = 1.19, PSH/TMO = 0.25, PSH/SOH = 0.5, PSH/POH = 0.34, TMO/SOH = 1.95, TMO/POH = 1.38, SOH/POH = 0.71.

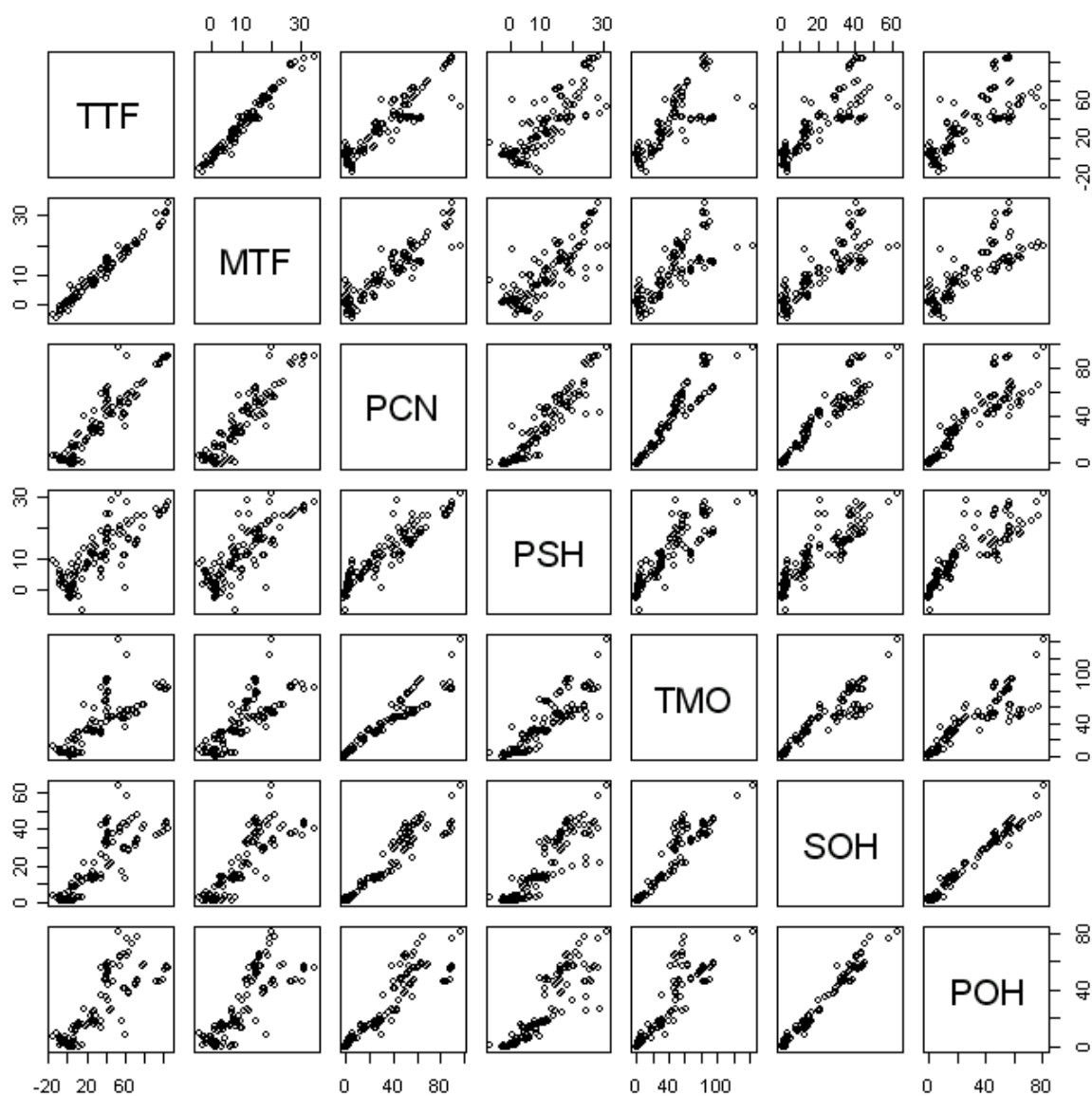


Fig. 1. The matrix of scatter plots of the additional retention of all pairs of stationary phases. In general, any pair of solvents produces additional retention index values that are tightly clustered around a straight line, which implies that any two solvents have positively correlated  $\Delta I_j^P$  values. The stationary phases are grouped by type I: TTF, MTF, type II: PCN, PSH and type III: TMO, SOH, POH.

*Box Plots of the Additional Retention*

Fig. 2 shows a summary of the 127 retention index differences,  $\Delta I_j^P$ , in the form of box plots for each of the seven polar solvents (see [7,8] for more details on box plots). With very few exceptions, the values of  $\Delta I_j^P$  are positive, that is, the retention index in a polar solvent is higher compared to  $C_{78}$ . This reflects the lack of interactive groups in  $C_{78}$ . Solvents with large  $\Delta I_j^P$ -values are TTF, TMO and PCN. The polar solvents MTF and PSH have the smallest  $\Delta I_j^P$ -values.

The box plots show the distribution of the variable  $\Delta I_j^P$  for each solvent separately, but they cannot show, how the variable  $\Delta I_j^P$  behaves for couples of solvents. This joint distribution of the additional retention index is exhibited in Fig. 2.

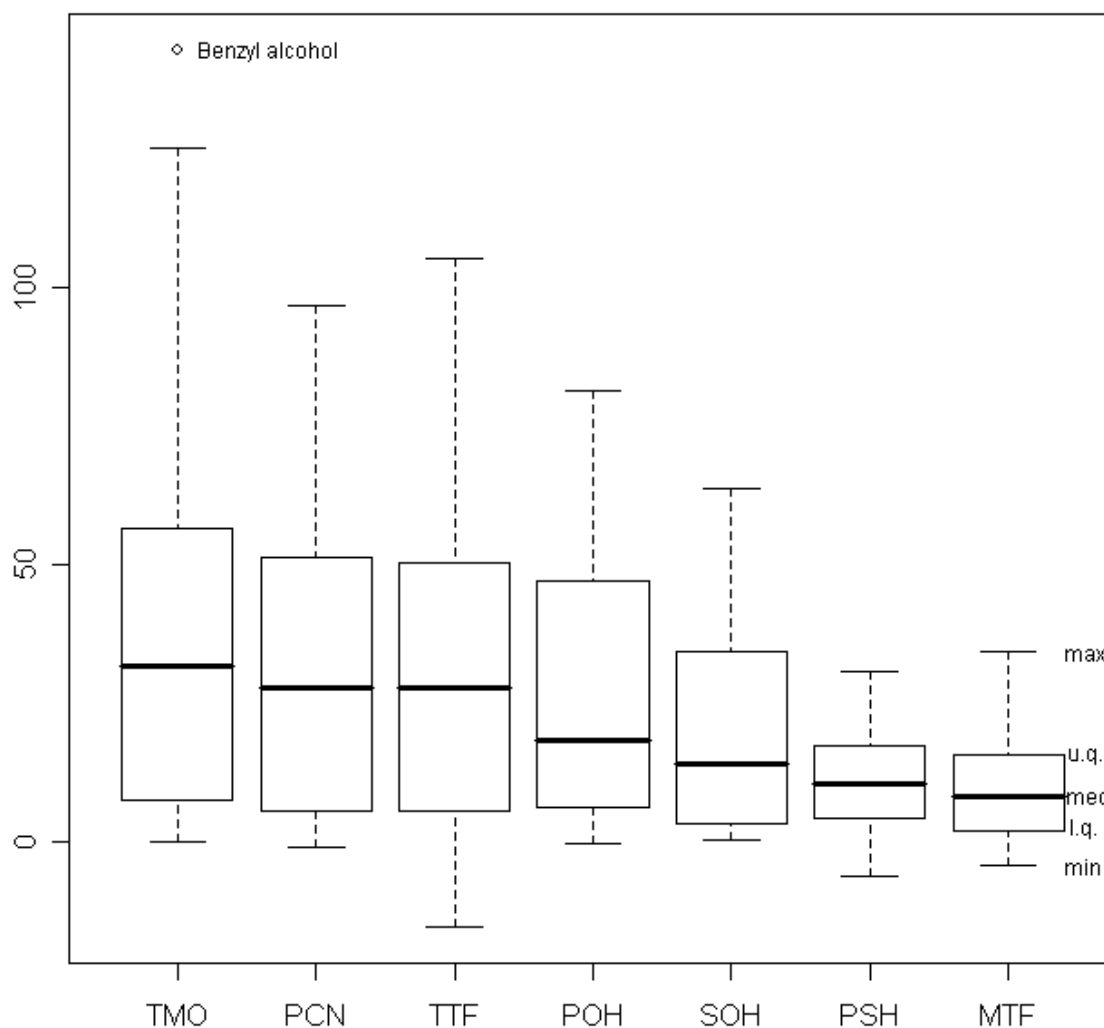


Fig. 2. Box plots of the variable  $\Delta I_j^P$  for the seven polar solvents, where  $\Delta I_j^P$  is the difference of the retention index between the solvent and the standard  $C_{78}$ . Each column corresponds to a stationary phase and shows the distribution of the  $\Delta I_j^P$  values across the substances  $j$ . The box extends from the lower quartile (l.q.) to the higher quartile (h.q.), with the median value (med) indicated by a horizontal separator. The whiskers attached to the box extend to the largest (max) and smallest (min) values. Recall that  $\Delta I_j^P$  is the difference of the retention index between the solvent and the standard  $C_{78}$ . A large majority of the  $\Delta I_j^P$ -values are positive and the median  $\Delta I_j^P$ -values are large for TTF, TMO and PCN, small for MTF and PSH and between the extremes for POH and SOH.

*Principal Components*

The analysis by principal components is a statistical procedure, which can be applied to data consisting of several variables measured on the same substances. If variables are intercorrelated, it is sometimes possible to reduce their number without much loss of information. As seen in Fig. 1 our seven extra retentions in different polar stationary phases are highly correlated and are such inter-related variables. How many of the seven are redundant and how many are really needed? We denoted the extra retentions by  $\Delta I_j^P$ , where the stationary phase P is either TTF, MTF, PCN, PSH, TMO, SOH, or POH. There are seven variables and also seven principal components. The  $k^{\text{th}}$  principal component is a linear combination of the variables, that is, of the form

$$C_{k,j} = v_k^{TTF} \Delta I_j^{TTF} + v_k^{MTF} \Delta I_j^{MTF} + v_k^{PCN} \Delta I_j^{PCN} + v_k^{PSH} \Delta I_j^{PSH} + v_k^{TMO} \Delta I_j^{TMO} + v_k^{SOH} \Delta I_j^{SOH} + v_k^{POH} \Delta I_j^{POH}. \quad (4)$$

The weights or loadings,  $v_k^P$ , are chosen in such a way that the first principal component maximises the variance, the second principal component is uncorrelated with the first and under this condition maximises the variance, the third is uncorrelated with the other two and maximises the variance, and so on. Of course, the constraints imposed on successive principal components become more and more limiting and the maximal variance achievable under the constraint becomes smaller and smaller. In this way, principal components analysis (PCA) can replace many variables by one or two principal components, which contain a very large part of the total variance. The relative importance of each component is determined by the fraction of the total variance it explains. All seven components together explain 100% of the variation in the original data, the first six a little less than 100%, and so on. A basic text on this statistical procedure, written for chemists, is found in [9].

In the data we analyse in this paper, the PCA is revealing. The first and most important components explains a very large 96.2% of the total variation in the variables  $\Delta I_j^P$ . This reflects the fact that the extra retentions in the seven polar phases are so strongly correlated that they can be successfully explained by an average extra retention. The first and the second components together increase this percentage to 98.8%. Hence, only 1.2% of the information is not captured by these two principal components.

The loadings of the first and second principal components are as follows:

	$v_k^{TTF}$	$v_k^{MTF}$	$v_k^{PCN}$	$v_k^{PSH}$	$v_k^{SOH}$	$v_k^{POH}$	$v_k^{TMO}$
first PC (k=1)	0.47	0.14	0.47	0.15	0.28	0.39	0.47
second PC (k=2)	0.78	0.19	0.04	-0.04	-0.16	-0.25	-0.51

It is not difficult to interpret the loading of the first principal component. The most important loadings are given to TTF, PCN and TMO (value of 0.47). The smallest loadings are given to MTF and PSH (value of

around 0.15). The weight 0.47 given to TTF is 3.3 times as large as the one given to MTF. Hence, these coefficients give large weight to the stationary phases with big extra retention (see Fig. 2) and small weight to those with small additional retention and the relative importance reflects the slopes we observed in Fig. 1. The second principal component has positive and negative loadings and describes a contrast between the extra retention in the stationary phases TTF and TMO. All other phases have small loadings.

A visual analysis of the loadings is obtained by plotting the vectors  $[v_1^P, v_2^P]$  for the seven polar stationary phases  $P = \text{TTF, MTF, PCN, PSH, TMO, SOH, POH}$ . Fig. 3 shows the vectors of loadings at different temperatures.



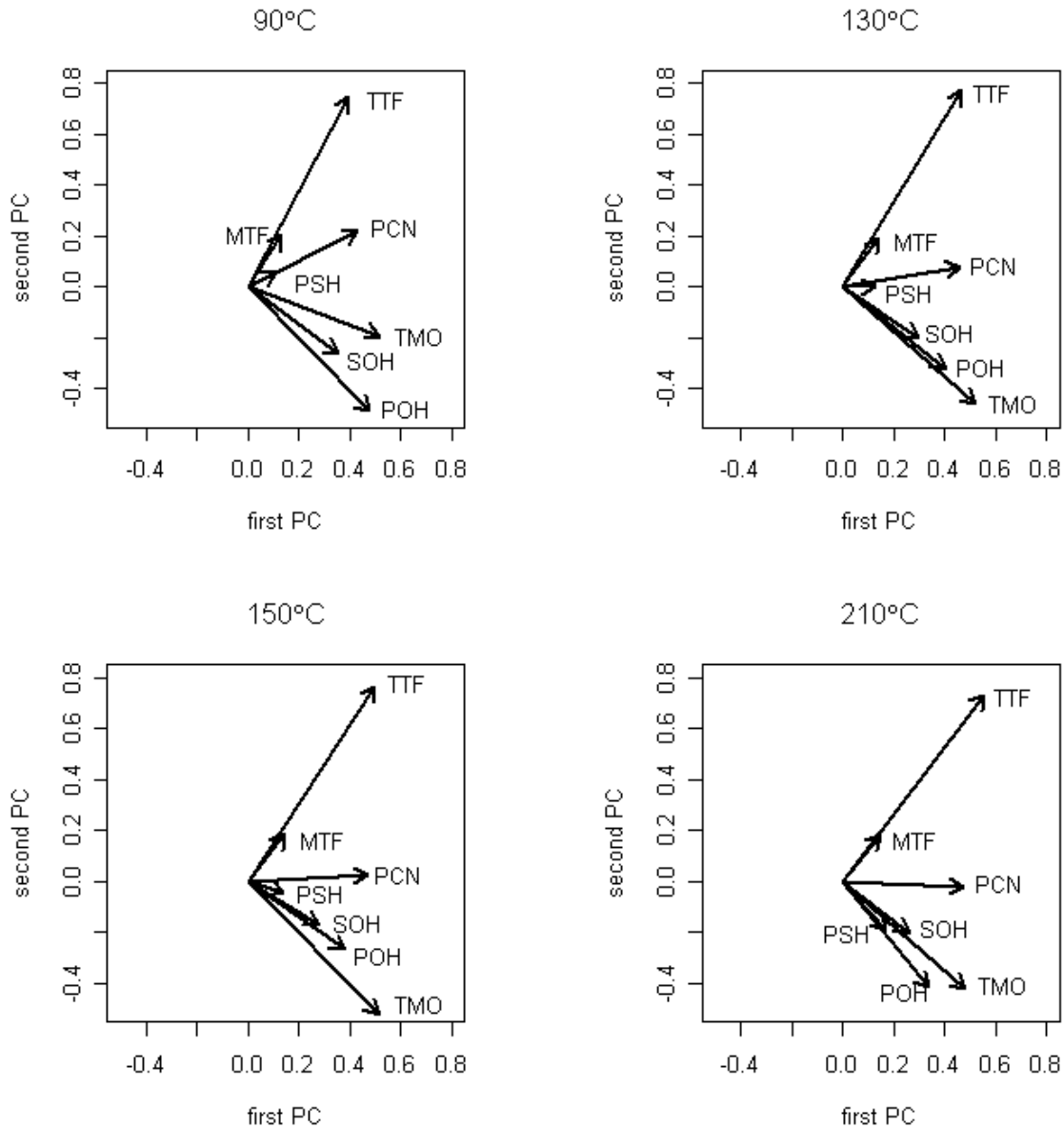


Fig. 3. Plot of the loadings for the two first principal components for different temperatures in °C. The grouping of the loading vectors changes across the temperature range. One attribute that remains constant is the separation between TTF and TMO as well as the fact that PCN remains in between these two. Of the other stationary phases, MTF stays with TTF, while PSH, SOH and POH change their behavior across the temperature range. At low temperatures, PSH is similar to PCN, at high temperatures it is similar to TMO. SOH and POH are close to each other and fluctuate around TMO. It is of interest to note that the fraction of the total variation explained by the first two principal components increases with temperature.

### Conclusions

This analysis shows that the seven polar stationary phases are well represented by TTF, PCN and TMO. This remains true across a large temperature range. Of the remaining solvents, MTF is grouped with TTF and POH, SOH are grouped with TMO. PSH has a small polarity effect and behaves similar to PCN at low temperatures and to TMO at higher temperatures. Because MTF and TTF are closely linked, one could also choose MTF as a representative. However, larger and more clearly visible effects will be obtained with TTF.

Table 2

The groups of substances shown in Figs. 4 - 9 are as follows:

Isoalkanes	1-Alkenes	1-Alkynes	Alkynes	Monocyclic hydrocarbons (HC)
2,2-Dimethylbutane 2,3-Dimethylbutane 2,2-Dimethylpentane 2,3-Dimethylpentane 2,4-Dimethylpentane 2,3-Dimethylhexane 2,4-Dimethylhexane 3,4-Dimethylhexane 2,2,3-Trimethylbutane 2,2,4-Trimethylpentane 2,3,4-Trimethylpentane	1-Hexene 1-Heptene 1-Octene 1-Nonene 1-Decene	1-Pentyne 1-Hexyne 1-Heptyne 1-Octyne 1-Nonyne 1-Decyne	2-Hexyne 3-Hexyne 4-Octyne	Cyclopentane Cyclohexane Cycloheptane Cyclooctane
Methylcyclohexanes	Cyclohexenes	Alkylbenzenes	1-Fluoroalkanes	1-Chloroalkanes
Methylcyclohexane <i>cis</i> -1,2-Dimethylcyclohexane <i>trans</i> -1,2-Dimethylcyclohexane <i>cis</i> -1,4-Dimethylcyclohexane <i>trans</i> -1,4-Dimethylcyclohexane	Cyclohexene 1,3-Cyclohexadiene 1,4-Cyclohexadiene	Benzene Toluene Ethylbenzene	1-Fluoropentane 1-Fluorohexane 1-Fluorooctane	1-Chlorobutane 1-Chloropentane 1-Chlorohexane
1-Bromoalkanes	1-Cyanoalkanes	1-Nitroalkanes	1-Acetoxyalkanes	1-Alkanols
1-Bromopropane 1-Bromobutane 1-Bromopentane	Cyanoethane 1-Cyanopropane 1-Cyanobutane 1-Cyanopentane	Nitroethane 1-Nitropropane 1-Nitrobutane 1-Nitropentane	1-Acetoxypropane 1-Acetoxybutane 1-Acetoxy-pentane	1-Butanol 1-Pentanol 1-Hexanol 1-Heptanol
2-Alkanols	2-Methyl-2-alkanols	1-Alkanethiols	2-Alkanons	Aldehydes
2-Butanol 2-Pentanol 2-Hexanol 2-Heptanol	2-Methyl-2-propanol 2-Methyl-2-butanol 2-Methyl-2-pentanol 2-Methyl-2-hexanol	1-Butanthiol 1-Pentanthiol 1-Hexanthiol	2-Butanone 2-Pentanone 2-Hexanone 2-Heptanone	Pentanal Hexanal

Ethers	Halomethanes	Halobenzenes	Pyridine derivatives	Organosilicon compounds
Dipropyl-ether Dibutyl-ether	Dichloromethane Trichloromethane Tetrachloromethane CF <sub>2</sub> Br <sub>2</sub>	Fluorobenzene Hexafluorobenzene Trifluoromethylbenzene Chlorobenzene Bromobenzene	Pyridine 2-Methylpyridine 3-Methylpyridine 4-Methylpyridine 3-Chloropyridine	Tetramethylsilane Hexamethyldisilane Hexamethyldisiloxane
Miscellaneous A	Cyclohydrocarbon derivatives	Cyclohydrocarbons	Dimethylpyridines	Miscellaneous B
Carbondisulfide Tetramethyltin Tetrahydrofurane 1.4-Dioxane Thiophene	Cyclopentanone Cyclohexanone Cyclohexanol	Cyclodecane <i>cis</i> -Hydrindane <i>trans</i> -Hydrindane <i>cis</i> -Decaline <i>trans</i> -Decaline Adamantane Naphthalene Azulene	2,3-Dimethylpyridine 2,4-Dimethylpyridine 2,5-Dimethylpyridine 2,6-Dimethylpyridine 3,4-Dimethylpyridine 3,5-Dimethylpyridine	Nitrobenzene Benzyl alcohol 2-Phenylethanol Anisole Ethoxybenzene

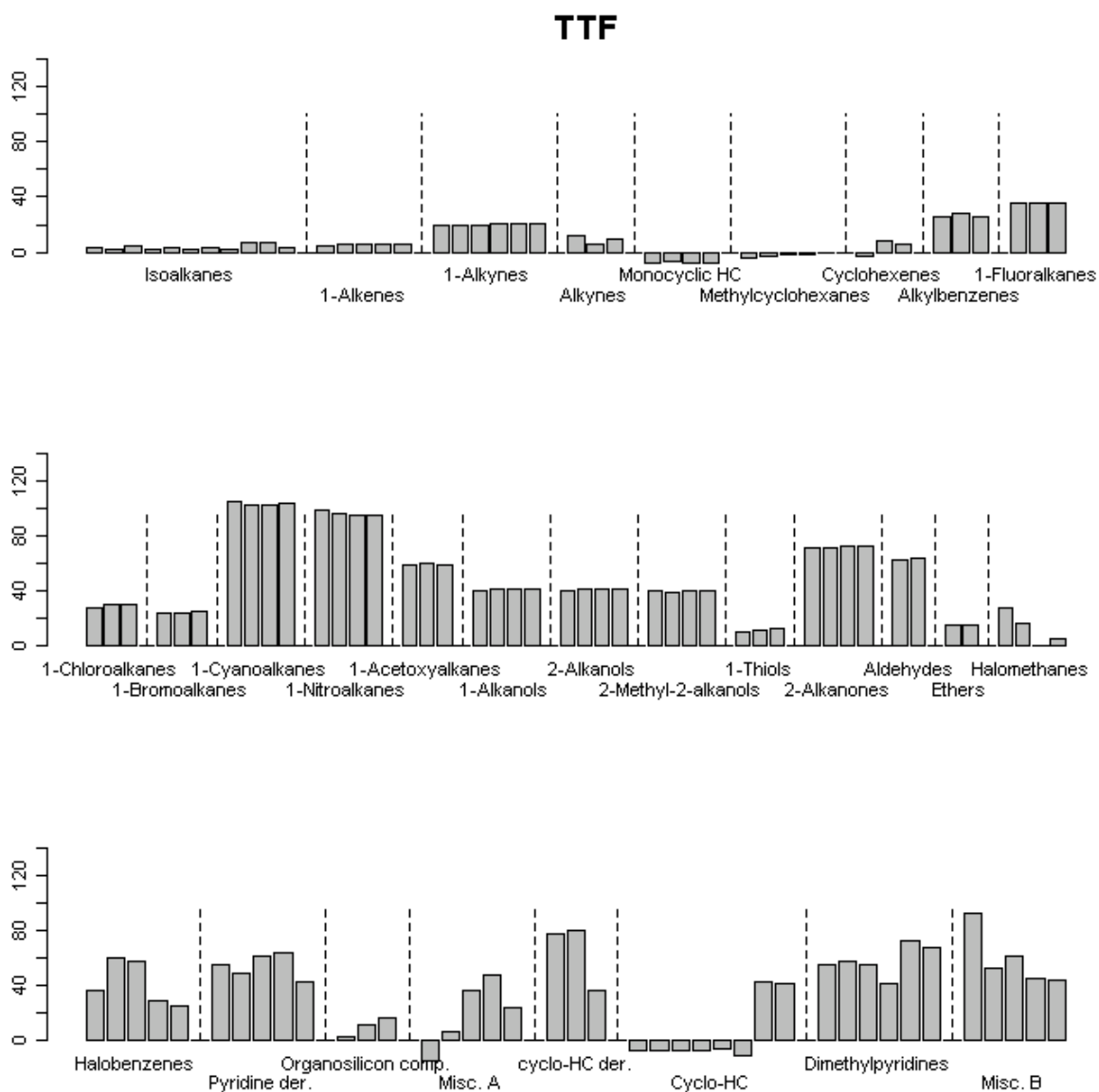


Fig. 4. Values of  $\Delta I_j^{TTF}$  at 150 °C as a function of the substances on the stationary phase TTF.

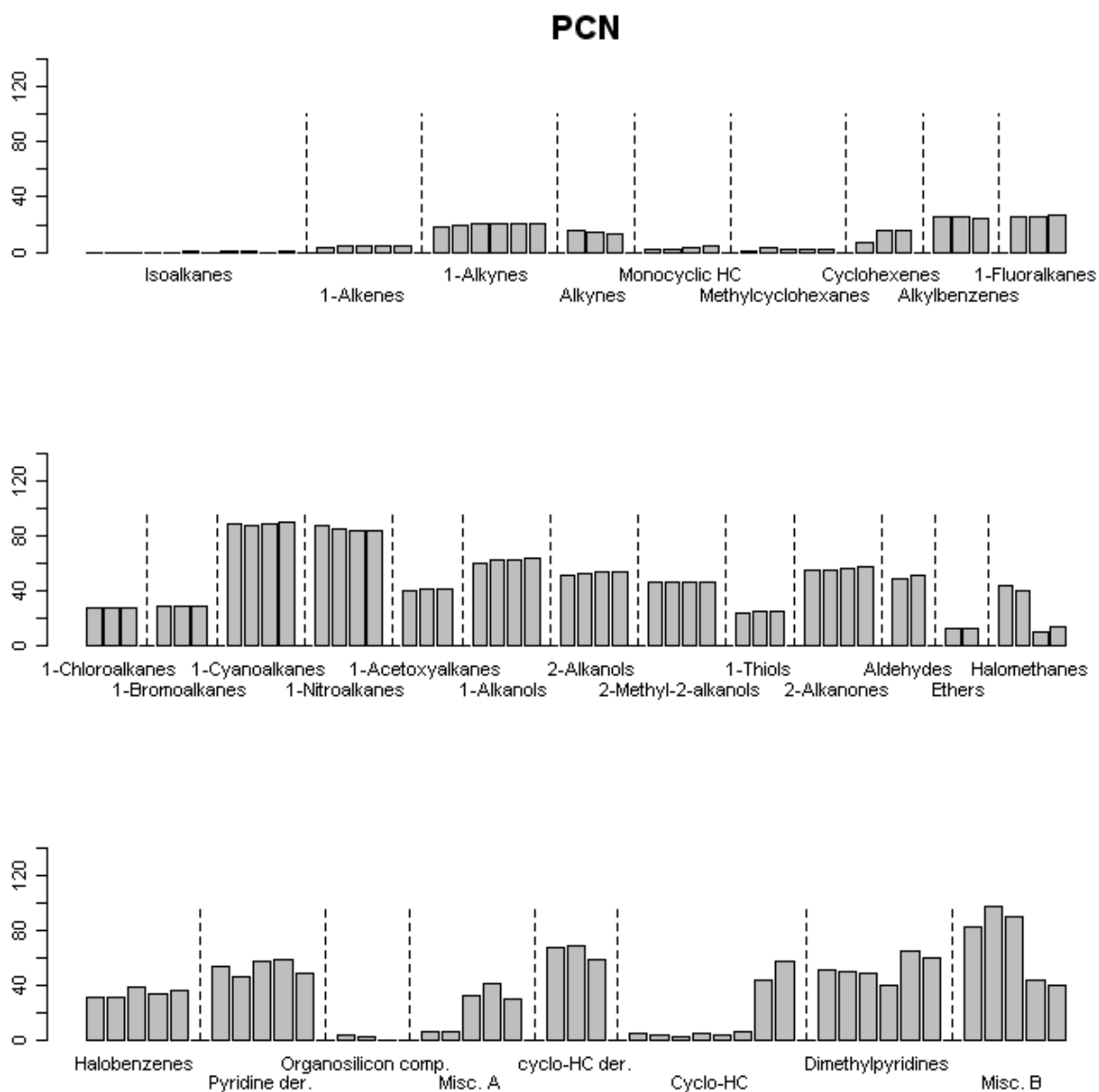


Fig. 5. Values of  $\Delta I_j^{PCN}$  at 150 °C as a function of the substances on the stationary phase PCN.

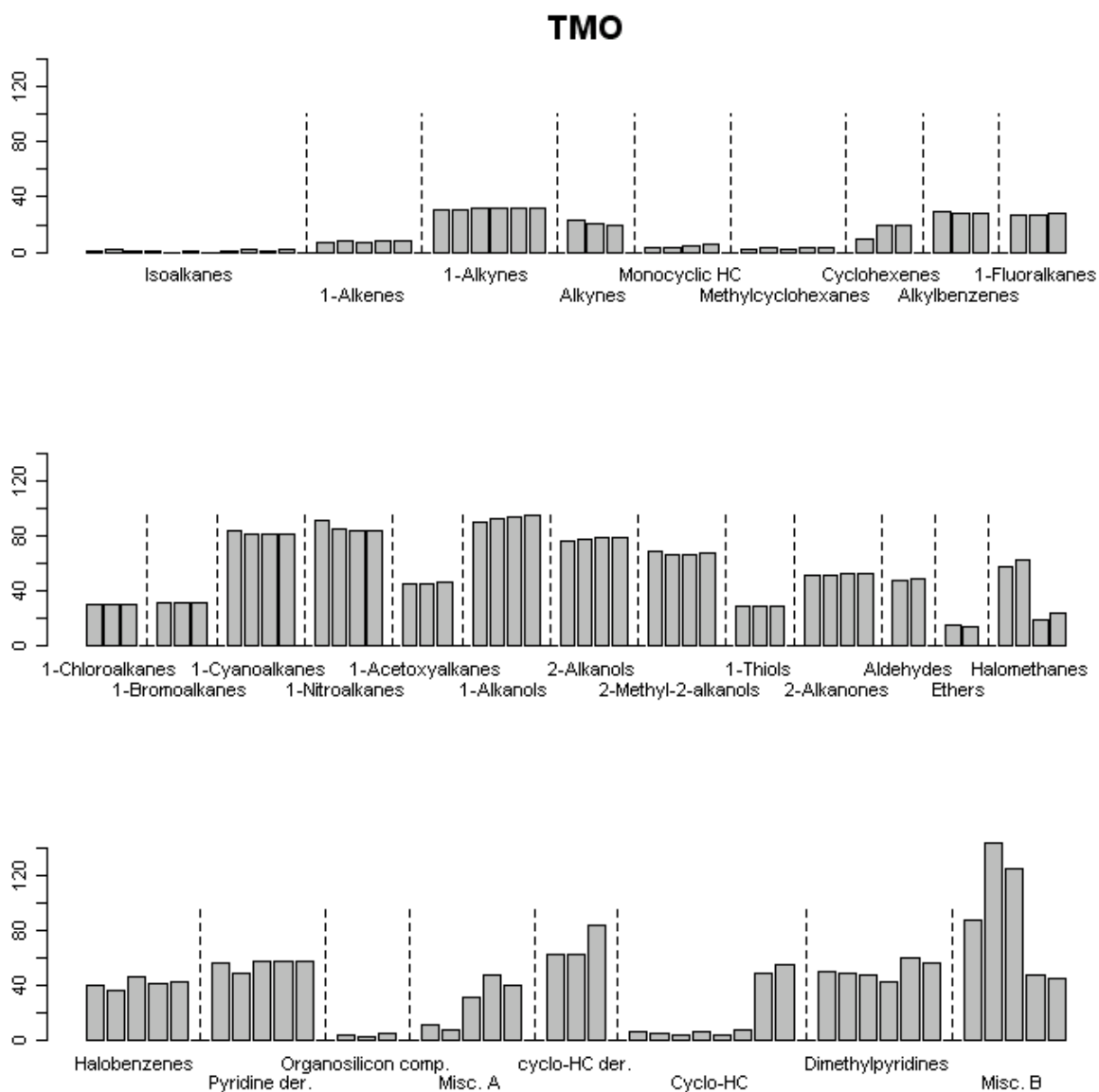


Fig. 6. Values of  $\Delta I_j^{TMO}$  at 150 °C as a function of the substances on the stationary phase TMO.

## The Choice of Polar Stationary Phases

The effect of the temperature on the  $\Delta I_j^P$ -value over a limited temperature range is well-described by a quadratic equation, which is closely approximated by a linear function. The following Figs. show  $(\Delta I_j^P [175 \text{ }^\circ\text{C}] - \Delta I_j^P [125 \text{ }^\circ\text{C}])/5$ , that is, the change of the  $\Delta I_j^P$ -value of a substance for  $10 \text{ }^\circ\text{C}$  around the median temperature of  $150 \text{ }^\circ\text{C}$  for the three selected stationary phases.

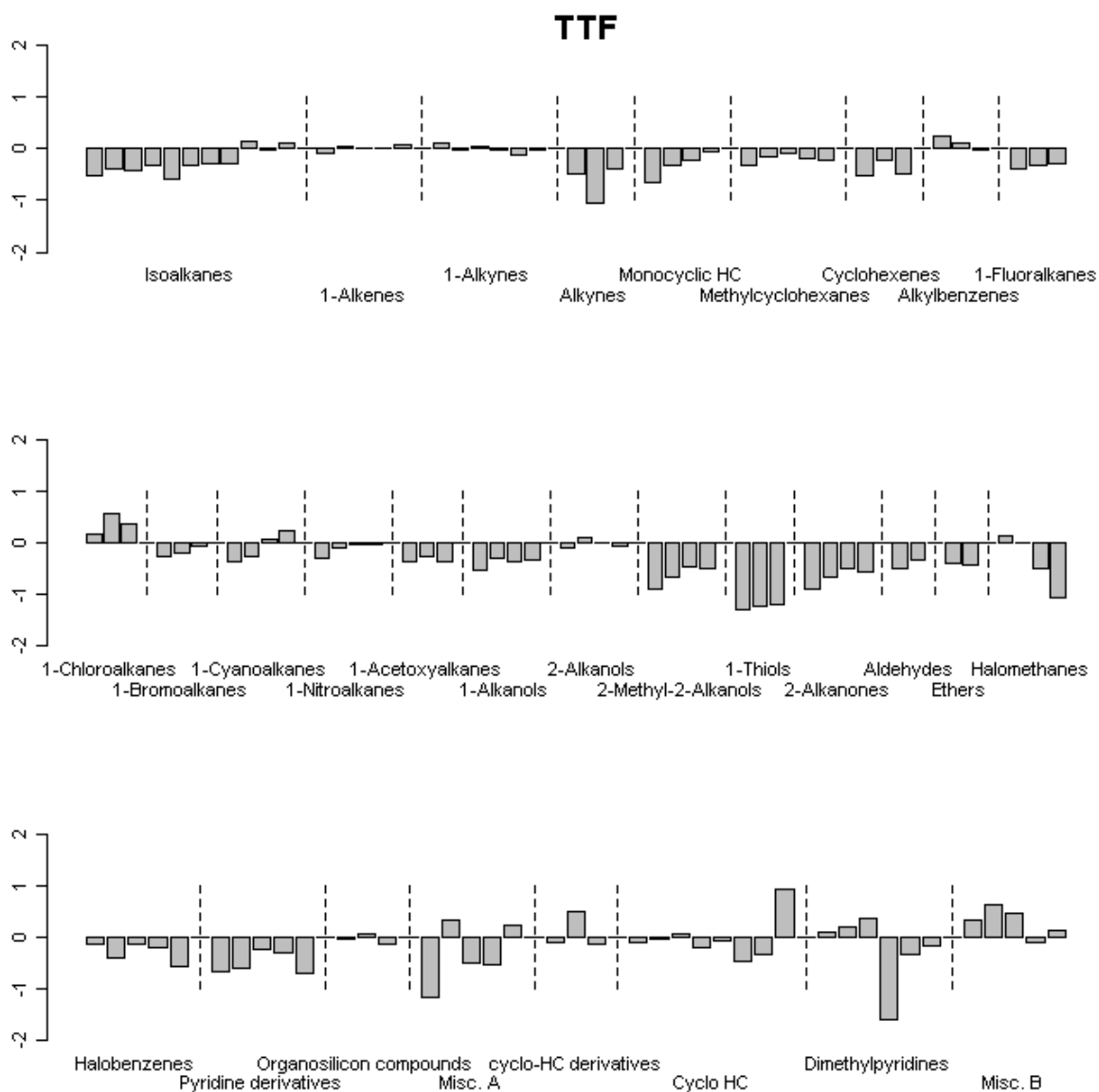


Fig. 7. Temperature dependence of the retention index per 10 °C between 125 and 175 °C on the stationary phase TTF: values of  $(\Delta I_j^{TTF}[175\text{ °C}] - \Delta I_j^{TTF}[125\text{ °C}])/5$ .



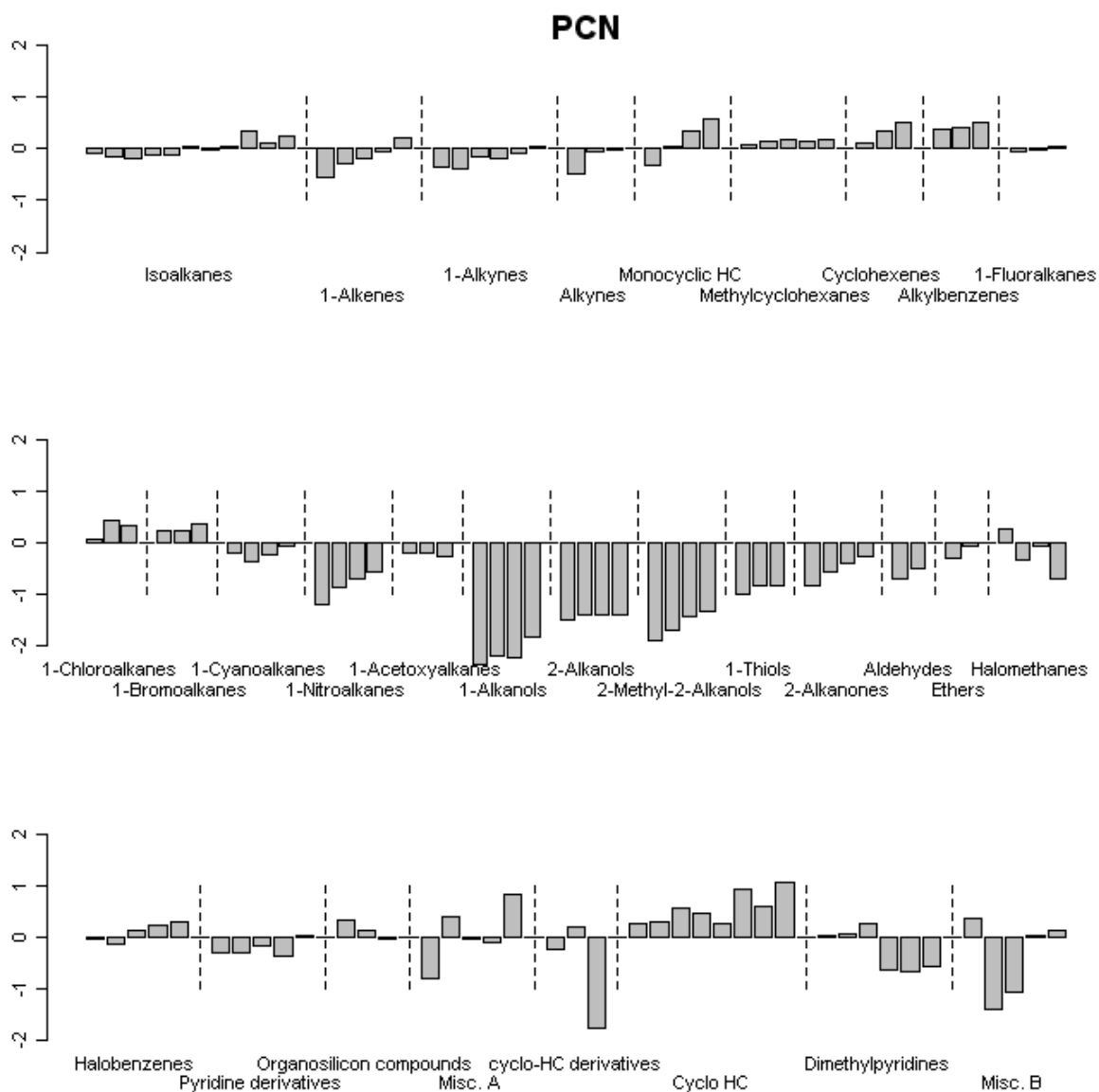


Fig. 8. Temperature dependence of the retention index per 10 °C between 125 and 175 °C on the stationary phase PCN: values of  $(\Delta I_j^{PCN}[175\text{ °C}] - \Delta I_j^{PCN}[125\text{ °C}])/5$ .

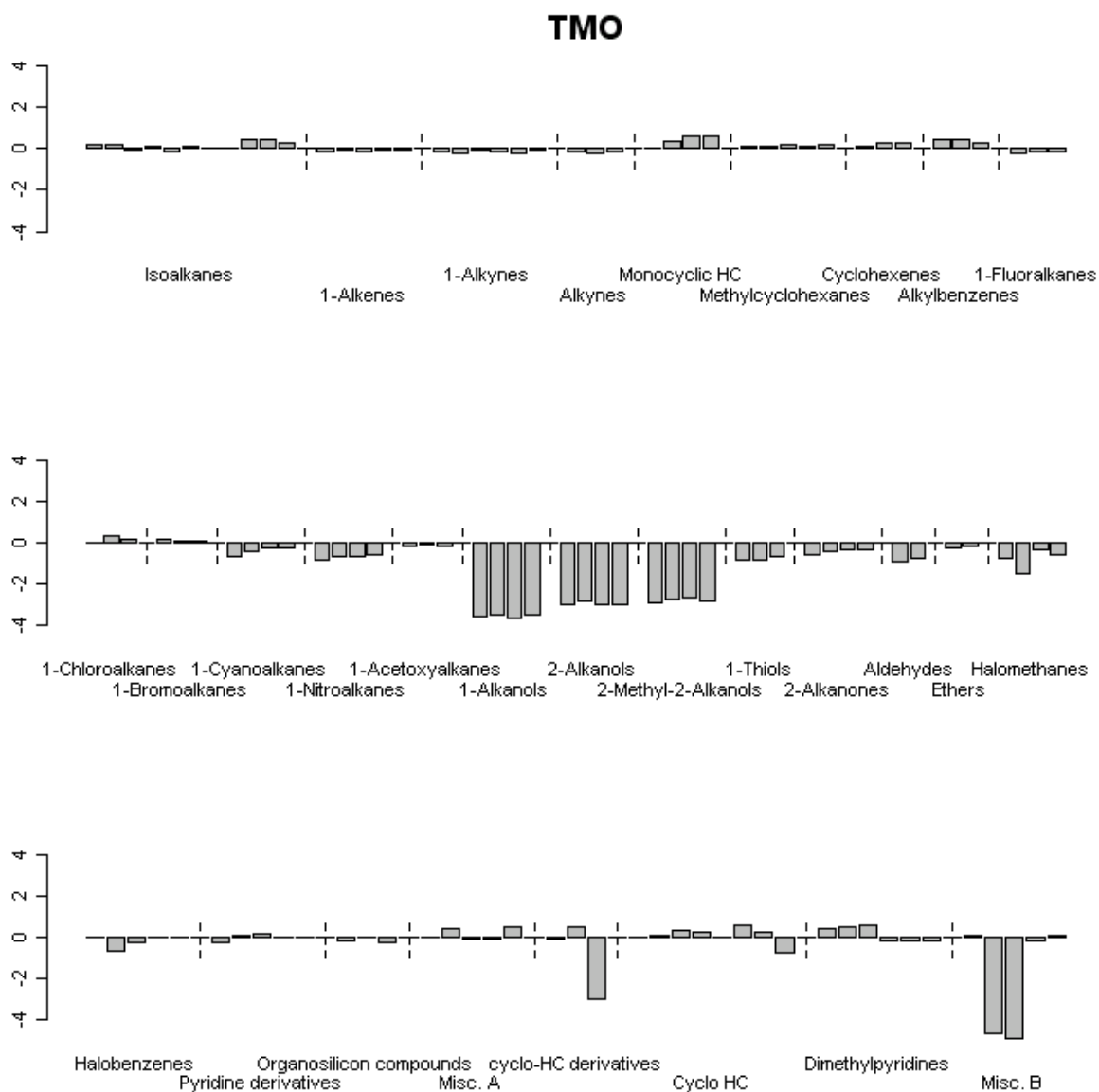


Fig. 9. Temperature dependence of the retention index per 10 °C between 125 and 175 °C on the stationary phase TMO: values of  $(\Delta I_j^{TMO}[175\text{ °C}] - \Delta I_j^{TMO}[125\text{ °C}])/5$ .

## Bibliography

1. K.S. Reddy, J.-Cl. Dutoit, E.sz. Kováts (1992) *J. Chromatogr. A* 609: 229-259. DOI 10.1016/0021-9673(92)80167-S
2. K.S. Reddy, R. Cloux, E.sz. Kováts (1994) *J. Chromatogr. A* 673: 181-209 DOI 10.1016/0021-9673(94)85037-2
3. K.S. Reddy, R. Cloux, E.sz. Kováts (1995) *J. Chromatogr. A* 704: 387-436. DOI 10.1016/0021-9673(95)93206-B
4. A. Dallos, A. Sisak, Z. Kulcsár, E.sz. Kováts (2000) *J. Chromatogr. A* 904: 211-242. DOI 10.1016/S0021-9673(00)00908-0
5. E.sz. Kováts, R. Kresz J. (2006) *J. Chromatogr. A* 1113: 206-219. DOI 10.1016/j.chroma.2006.01.123
6. A. Dallos, H.S. Ngo, R. Kresz, K. Héberger (2008) *J. Chromatogr. A* 1177: 175-182. DOI 10.1016/j.chroma.2007.11.034
7. J.W. Tukey (1977) *Exploratory Data Analysis*. Academic Press, USA.
8. R. McGill, J. W. Tukey, W. A. Larsen (1978) *The American Statistician* 32: 12-16
9. E.R. Malinowski (2002) *Factor Analysis in Chemistry* (3<sup>rd</sup> ed.). Wiley, USA

### Appendix A. Data

Corrected data used in this article can be found, in the online version, in [5] at doi:10.1016/j.chroma.2006.01.123.