Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1002 (2003) 179–192
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Decoding of complex isothermal chromatograms recovered from space missions
Identification of molecular structure

Maria Chiara Pietrogrande[a,*], Ilaria Tellini[a], Luisa Pasti[a], Francesco Dondi[a], Cyril Szopa[b], Robert Sternberg[c], Claire Vidal-Madjar[d]

[a]*Department of Chemistry, University of Ferrara, Via L. Borsari 46, 44100 Ferrara, Italy*
[b]*SA, UMR CNRS 7620, 91371 Verrières-le-Buisson, France*
[c]*LISA, UMR CNRS 7583, 94010 Créteil Cedex, France*
[d]*LRP, CNRS, 94320 Thiais, France*

Received 28 February 2003; received in revised form 22 April 2003; accepted 23 April 2003

Abstract

A chemometric approach, based on the study of the autocovariance function, is described to study isothermal GC chromatograms of multicomponent mixtures: isothermal GC analysis is the method of choice in space missions since it is, to date, the only method compatible with flight constraints. Isothermal GC chromatograms look inhomogeneous and disordered with peak density decreasing at higher retention times: a time axis transformation is proposed to make retention an homogeneous process so that CH_2 addition in terms of an homologous series yields a constant retention increment. The time axis is transformed into a new scale based on the retention times of *n*-alkanes, as they are the basis of the universal Kovats indices procedure. The order introduced into the chromatogram by retention time linearization can be simply singled out by the experimental autocorrelation function (EACF) plot: if constant inter-distances are repeated in different regions of the chromatogram, well-shaped peaks are evident in the EACF plot. By comparison, with a standard mixture it is possible to identify peaks diagnostic of specific molecular structures: study of the EACF plot provides information on sample chemical composition. The procedure was applied to standard mixtures containing compounds representative of the planetary atmospheres that will be investigated in the near future: in particular, those related to Titan's atmosphere (Cassini–Huygens mission) and cometary's nucleus (Rosetta mission). The employed experimental conditions simulated those applied to GC instruments installed on space probes and landers in space missions. The method was applied to two specific investigations related to space research, i.e., a comparison of retention selectivity of different GC columns and identification of the chemical composition of an unknown mixture.

Keywords: Chemometrics; Autocovariance function; Volatile organic compounds

1. Introduction

In previous papers it has been demonstrated that a chemometric approach based on Fourier analysis

*Corresponding author. Tel.: +39-0532-291-152; fax: +39-0532-240-709.
E-mail address: mpc@dns.unife.it (M.C. Pietrogrande).

0021-9673/03/$ – see front matter
doi:10.1016/S0021-9673(03)00727-1

(FA) is a powerful tool for decoding a complex chromatogram, i.e., to extract all the information contained therein regarding the mixture—number of components, abundance distribution—and separation—separation performance, retention pattern [1–12]. In particular, its power lies in its ability to magnify ordered retention patterns, if present in the chromatogram, and single them out from the ''disordered forest'' of random peaks [6,7,11,12]. In fact, a chromatogram of a multicomponent sample displays a crowd of peaks randomly distributed within the chromatographic space and usually overlapping so that the ordered structure, if present, is hidden and extremely difficult to detect. The order—peaks appearing at repeated distances as a sequence (homologous series), or at constant distances located in different parts of the chromatogram—is related to specific chemical structure variations in the separated compounds and can be used to identify the presence of such structures in the mixture. The procedure has been successfully applied to chromatograms obtained under temperature-programming conditions [5–11]. Recently, the method has also been extended to study isothermal GC chromatograms [12]: the statistical handling of such chromatograms is complicated by the fact that they look inhomogeneous and disordered—they are more crowded at small or intermediate retention times, while fewer peaks appear at higher retention times—and peak width is not constant, since it increases with retention time [13–16].

To date, isothermal GC separations have not been popular, due to their low separation power and efficiency: the best separation performance can be achieved by a proper optimization of temperature-programmed conditions [14–17]. However, isothermal GC analysis is the method of choice for the analysis of extra-terrestrial atmospheres in space missions: GC instruments are installed on space probes and landers for in situ analyses in solar system explorations to identify and quantify gaseous or vaporizable constituents of planetary atmospheres and surfaces [18–26]. Specific requirements are imposed by flight conditions: GC equipment must have an extremely low mass and consume little power, but must provide high efficiency, resolution, sensitivity and reliability, working automatically or under remote control [18]. The strongest constraint concerns energy savings which imposes temperature and pressure limitations; for these reasons, isothermal separations are currently the only ones compatible with flight constraints [19–26]. The low separation performance obtained under isothermal conditions results in an increase in the signal complexity and peak overlapping usually present in separations of multicomponent mixtures such as planetary atmospheres [14,27–29]. In this case, the use of a mathematical approach, to deconvolve incompletely resolved peaks and to interpret the chromatogram, is mandatory if one wants to extract all the analytical information hidden therein, that is ''to decode'' the complex chromatogram [8,10].

In this paper an extension of the FA procedure consisting of a time axis transformation is proposed to evaluate complex isothermal chromatograms [3,12]: the retention time inhomogeneity is removed and the presence of an ordered structure, i.e., retention time repetitiveness related to specific molecular structures, can be singled out. The reliability of the procedure was tested on chromatograms of standard and unknown mixtures obtained under isothermal conditions simulating those usually applied in in situ analysis of extra-terrestrial atmospheres or soils (lander). Simulation conditions concern the Cassini–Huygens mission, developed in order to study the atmospheric composition of Titan, Saturn's largest moon, in 2005 [19,20], and the cometary sampling and composition (COSAC) experiment on board the cometary nucleus lander of the Rosetta mission to be launched in 2003 [21,22,25,26]. The proposed procedure can also be applied to chromatograms showing inhomogeneity in retention pattern, i.e., obtained in conditions far from optimal temperature-programmed conditions, i.e., pseudo-isothermal conditions [14–17].

2. Experimental

Standard mixtures were selected to represent the chemical composition of comet's nucleus [20–22,25,26]: hydrocarbons and oxygenated compounds with carbon atom numbers ranging from two to eight. A detailed list of the compounds is reported in Table 1.

Sample analysis was performed using a Mega

Table 1
Standard compounds representing the chemical composition of comet's nucleus

Hydrocarbons	*Alcohols*
n-Heptane	1-Butanol
n-Octane	1-Pentanol
n-Nonane	1-Hexanol
n-Decane	1-Heptanol
n-Undecane	1-Octanol
n-Dodecane	
Ketones	*Acetates*
2-Butanone	Propyl acetate
2-Pentanone	Butyl acetate
2-Hexanone	Amyl acetate
2-Heptanone	Hexyl acetate

Series 5160 gas chromatograph (Fisons Instruments, Milan, Italy) equipped with a flame ionization detection (FID) system. The columns used were polysiloxane-based wall-coated open tubular (WCOT) capillary columns (Restek, Bellafonte, PA, USA): a MTX 1 silanized stainless-steel (10 m×0.15 mm I.D., d_f 0.8 μm) and a RTX-20 column (10 m×0.25 mm I.D., d_f 1 μm) similar to those used in the COSAC experiment [25,26]. Previous studies concerning space applications showed that these columns make it possible to separate a wide range of components and exhibit high resistance properties (long lifetime, low bleeding, thermal stability, good surface coating): therefore, they are two of the GC columns selected for the Rosetta mission [19–26]. The analyses were performed under isothermal conditions in the 30–60 °C temperature range, which is the operating temperature planned for the COSAC experiments. The carrier gas was hydrogen with a flow-rate of 25.6 cm s^{-1}. FID was used. A split mode injection was used (with a splitting ratio of 500:1).

3. Theory

3.1. Autocovariance function under temperature-programming conditions

The chemometric approach, based on Fourier analysis, studies the autocovariance function (experimental ACVF, EACVF), which can be directly computed from the experimental chromatogram acquired in digitized form using the following expression [1,2]:

$$EACVF(t) = \frac{1}{M} \cdot \sum_{j=1}^{N-k} (Y_j - \hat{Y})(Y_{j+k} - \hat{Y}),$$
$$k = 0,1,2,\ldots M-1 \quad (1)$$

where Y_j is the digitized chromatogram signal, $\hat{Y}$ is its mean value, and M is the truncation point in the EACVF computation. The time t is the inter-distance between subsequent points in the chromatogram and assumes discrete k values ranging from 0 to $M-1$.

The autocorrelation function (ACF), representing the ACVF normalized to the value computed at time 0, is more frequently used than the ACVF itself. The ACF describes the short- and long-range correlation between subsequent peak positions. When EACF is computed on chromatograms obtained under temperature-programming conditions, i.e. nearly constant peak shape, its plot vs. the correlation inter-distance exhibits two informative regions [5–7,11].

(i) The first part of the EACF contains the shortest-range correlation and depends only on the shape of the single-component peaks. It resembles the descending half of a Gaussian peak describing the mean peak shape averaged on all the peaks in the chromatogram. Theoretical expressions have been derived [1–4] and a simplified procedure has been developed [7,8] to extract information from this part of the ACVF by a simple graphical inspection. Without any handling of particular regions of the chromatogram or specific peaks, whole chromatogram computation [Eq. (1)] makes it possible to obtain a correct estimate of chromatographic parameters such as the number of components, m, and average peak shape parameters (σ).

(ii) The second part (i.e., widest-range correlation inter-distance) is determined by the distribution of the retention increments: if an ordered retention pattern exists in the chromatogram—peaks appearing at constant inter-distances repeated in different parts of the chromatogram—some positive peaks are present at the corresponding distance values in the EACF plot [5–7,11]. In contrast to the original chromatogram where peak overcrowding hides repetitions in the retention pattern, the EACF plot retains only a small number of repeated peaks, corresponding to the most abundant/most repeated inter-

distances. The result is that the EACF plot is much simpler and a simple visual inspection allows one to identify regularities in the retention pattern which reflect specific structure variations in the separated compounds. Moreover, information on the quantitative composition of the samples can also be obtained: while the position of the EACF peaks is related to the specific molecular structure, their height is related to the abundance of repetitiveness in the chromatogram, i.e., the number of repeated peaks and their heights [7].

The basic assumption of this approach is that a complex chromatogram looks like a random series of peaks. In other words, single component (SC) peaks can be found with a constant probability per time unit at any point in the chromatogram [1,2]; this means that, along the chromatogram time axis, the single component density is constant. This assumption leads one to conclude that the SC peak position is described by a Poisson probability density function and the distribution of retention time increments is given by the exponential distribution. This is true for gradient LC or temperature-programmed GC separations [5–11,14]. Moreover, the most general expression of ACVF is obtained when the peak width is constant (gradient or temperature-programming conditions), while in isothermal GC separations peak width increases with retention time [3,12,27,28]. The FA model has been extended to describe these chromatograms by developing different theoretical models: a general expression has been derived which holds when peak width is randomly distributed along the time axis and is linearly dependent on retention time [3]. The FA procedure allows an accurate estimation of the number of components, m, and the standard deviation of the narrowest (σ_1) and the largest peak (σ_2): it has been successfully applied to isothermal chromatograms simulating space mission analyses [12].

3.2. Linearization procedure

In isocratic or isothermal separations the retention process is inhomogeneous, since under these conditions the partition free energy $\Delta\mu^0$ controlling the retention is linearly related to $\log t_r$ (logarithm of the retention time). Under these conditions the different components of a complex multicomponent mixture separation usually exhibit uniform density over the partition free energy $\Delta\mu^0$ axis, i.e. the ratio $\Delta m/\Delta(\Delta\mu^0)$, where m, the number of single components of the complex mixture, will prove to be constant. Consequently, the interval $\Delta m/\Delta\log t_r$ will also be constant [14–17]. Likewise for a given homologous series, $\log t_r$ increases linearly with the carbon number, n [15], since $\Delta(\Delta\mu^0)$ is constant for a CH_2 increment. As a consequence, peak density is higher at small or intermediate retention times and decreases at higher retention times. Likewise, when terms of n-alkane series are analyzed under isothermal conditions, an inhomogeneous chromatogram is obtained (Fig. 1a), where repetitivity diagnostic for the presence of a homologous series is totally absent: the autocovariance function computed from the isothermal–isocratic chromatogram will not exhibit the features described above for programming temperature/gradient elution conditions and will show a disordered pattern (Fig. 1a, inset).

Therefore, a time axis transformation into a new scale is required to make retention an homogeneous process yielding constant retention increments for CH_2 addition in terms of a homologous series. If $Y(x)$ represents the chromatographic signal, where x is the retention time, a transformation of the time axis into a new scale:

$$z = g(x) \tag{2}$$

is proposed. In order to preserve the total signal area, the following condition must be fulfilled:

$$Y_1(x)\mathrm{d}x = Y_2(z)\mathrm{d}z \tag{3}$$

which expresses the area conservation in the transformation of frequency functions (Hamilton), or:

$$Y_1(z) = \frac{Y_2[g(x)]}{g'(x)} \tag{4}$$

where $Y_1(x)$ is the original chromatogram in the original time axis x, $Y_2(z)$ is the transformed chromatogram in the new z-axis, $z = g(x)$ is the function which relates the original time axis to the new z-axis and $g'(x)$ is the first derivative vs. x of $g(x)$. One can see that, at a given z position of the transformed chromatogram, the value of the signal Y_2 is different from the corresponding value Y_1 in the original chromatogram, according to Eq. (4). As $z = g(x)$, one can assume the function $z = \log(x)$ as mentioned

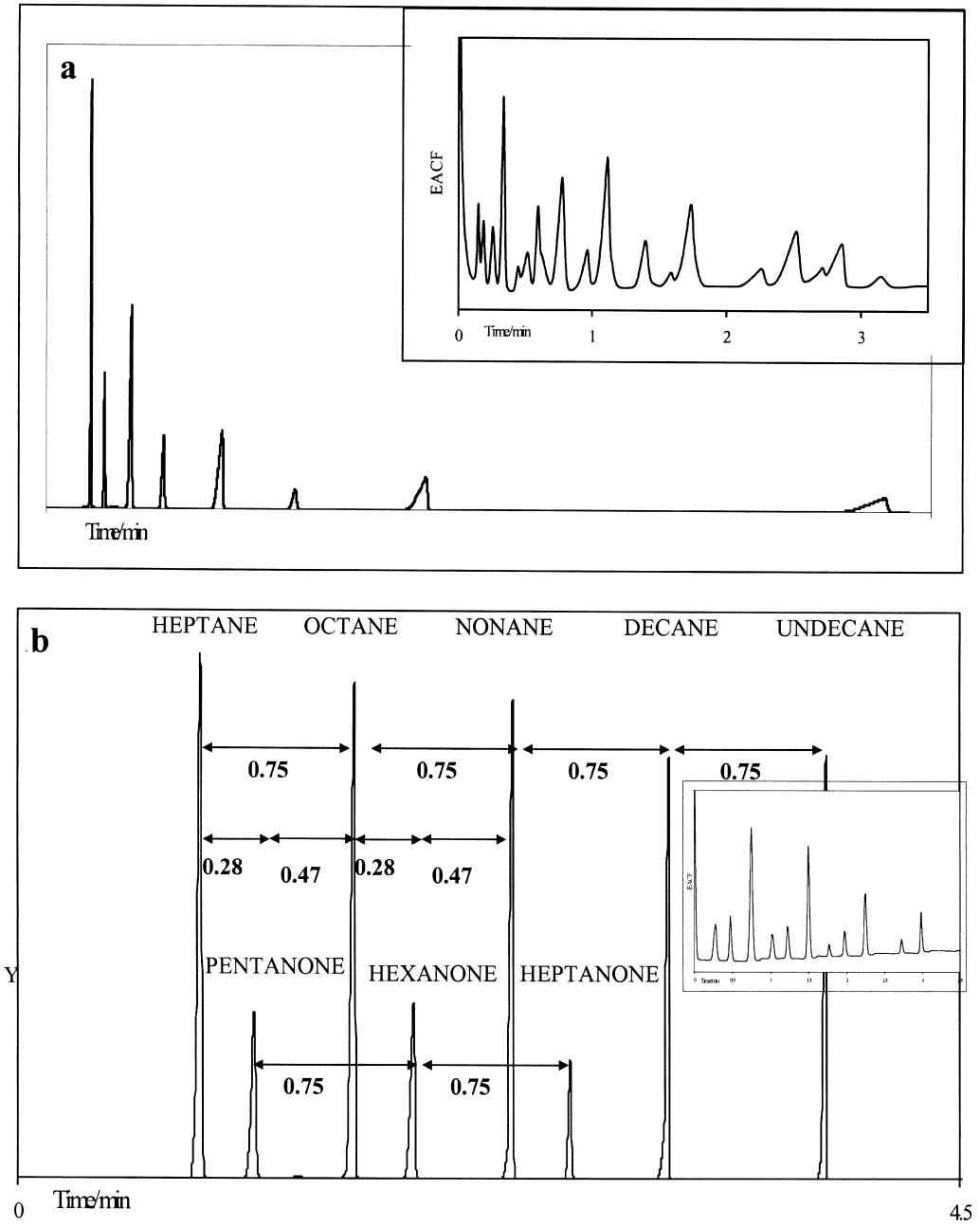

Fig. 1. Isothermal chromatogram of a standard mixture containing C_7–C_{11} *n*-alkanes and C_5–C_7 ketones. RTX-20 column, isothermal conditions at 60 °C, flow-rate 25.6 cm s^{-1}. (a) Original chromatogram and plot of EACF computed from it (inset); (b) linearized chromatogram and plot of EACF computed from it (inset).

above. Here, a different approach has been followed: an empirical transformation of the chromatogram which ensures a constant Δz increment between the subsequent terms of a homologous series, i.e., CH_2 addition. Therefore, the applied transformation has the property:

$$Y_1(x)\Delta x = Y_2(z)\Delta z \quad (5)$$

over subsequent regions of the chromatogram between subsequent members of the homologous series, in agreement with Eq. (3).

The retention time dependence of *n*-alkanes on *n* was selected as reference since it is the basis for the universal procedure of Kovats indices [16]. Components of the *n*-alkane series containing seven to eleven carbon atoms were submitted to isothermal GC analysis (chromatogram in Fig. 1a) and the inter-distance between subsequent terms was considered. The distance Δx_{7-8} in the original retention axis was assumed as a constant Δz retention increment in the new scale: subsequent retention regions, i.e., Δx_{8-9}, Δx_{9-10} and Δx_{10-11}, were linearly transformed into new Δz regions according to the plot reported in Fig. 2 (isothermal separation at 60 °C on a RTX-20 column). In practice, the chromatogram reported in Fig. 1b is the empirical transformation of the original chromatogram in Fig. 1a using a function corresponding to Eq. (2). In each Δx region the acquisition frequency was reduced to obtain constant Δz intervals. When some points of the original output signal $Y_1(x)$ are deleted, their values are added to the re-scaled signal $Y_2(z)$ in order to preserve the original chromatographic area, i.e., to fulfill Eq. (4) and conserve the chromatographic area under the transformation. If terms of an homologous series are present in the sample (Fig. 1a) a structured chromatogram characterized by retention repetition is obtained after the linearization procedure (Fig. 1b). If the sample contains incomplete homologous series or components that do not belong to any series, it yields a disordered retention pattern. It is well known that the superposition of different series of oligomers—i.e., the case of real mixtures consisting of assorted members of many classes of compounds—as well as the randomness of uncorrelated compounds leads to the domination of a random distribution along the $\Delta\mu^0$ axis [28,29]. The autocovariance function was then calculated on the re-scaled chromatogram: a simple visual inspection of the EACF plot (Fig. 1b, inset) makes it possible to single out constant repetitiveness diagnostic of the presence of homologous series.

The time axis transformation also yields a decrease in peak width, as a consequence of the reduction of the acquisition frequency to obtain constant Δz intervals. The most retained peaks, which appear strongly fronting in the original isothermal chromatogram (Fig. 1a), become narrower and more symmetrical after the linearization procedure, yielding a nearly constant peak width (Fig.

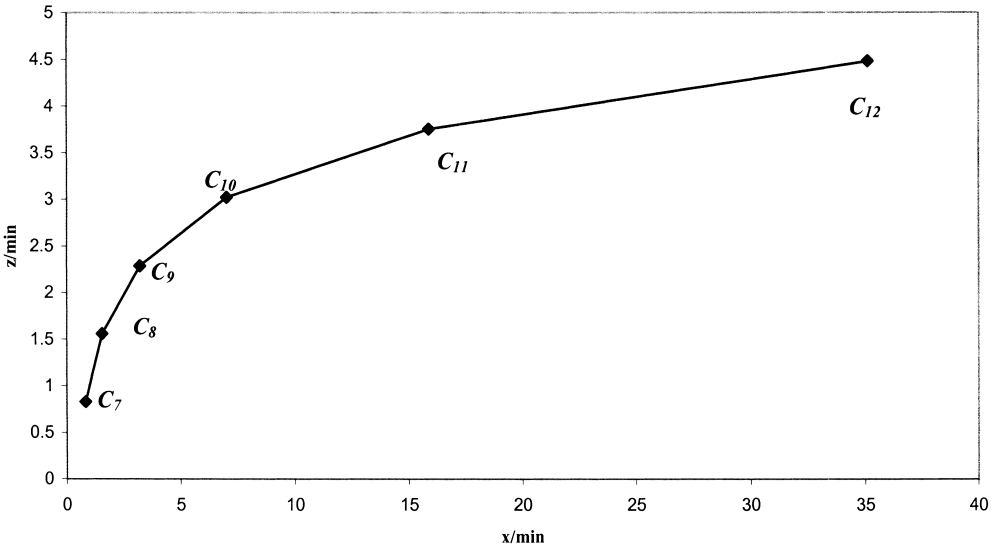

Fig. 2. Empirical linearization procedure for an isothermal chromatogram (isothermal separation at 60 °C on a RTX-20 column). *x*: original retention times for C_7–C_{12} alkanes; *z*: linearized retention times yielding constant intervals Δz.

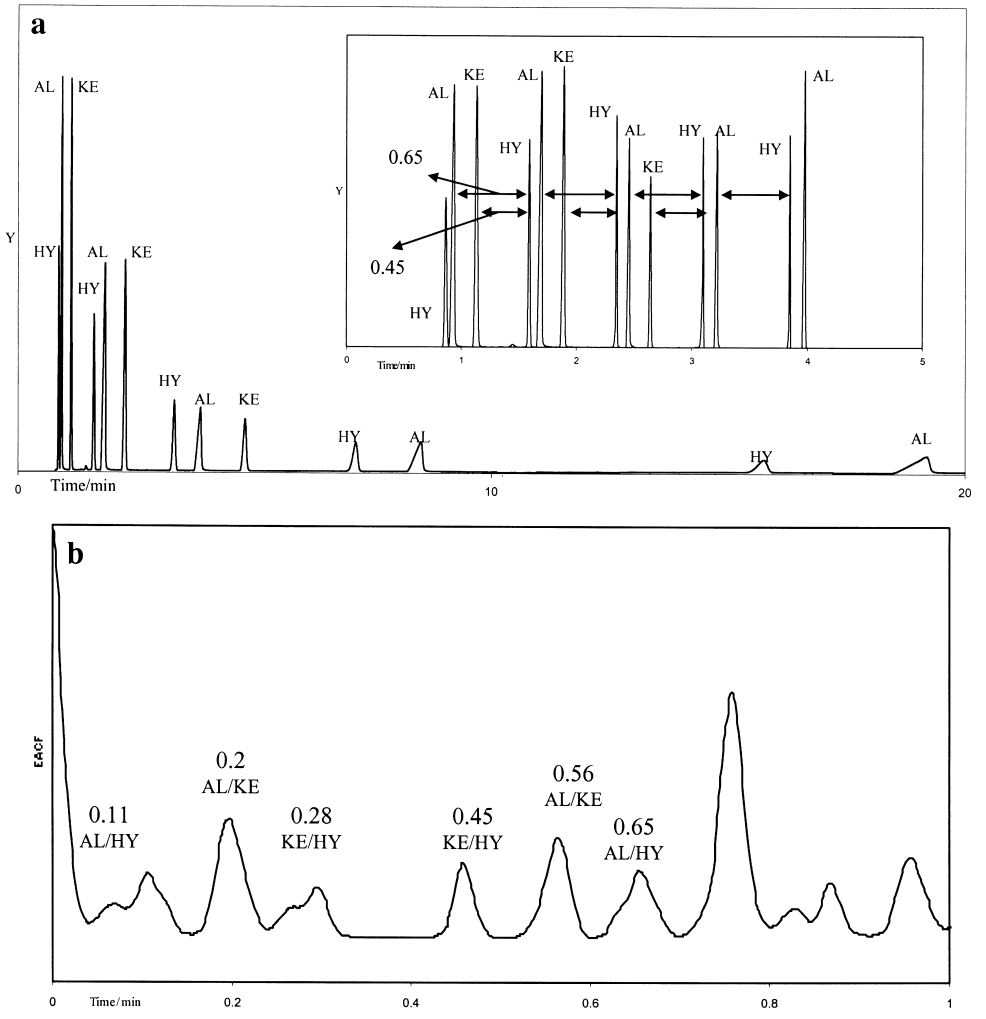

Fig. 3. Isothermal chromatogram of a standard mixture containing *n*-alkanes (HY), alcohols (AL) and ketones (KE). RTX-20 column, isothermal conditions at 60 °C, flow-rate 25.6 cm s^{-1}. (a) Original chromatogram, linearized chromatogram in the inset: inter-distance repetitiveness is indicated by the arrows; (b) plot of EACF computed from the linearized chromatogram.

1b). As a consequence, the EACF plot computed on the linearized chromatogram shows narrow, well-shaped peaks (Fig. 1b, inset).

4. Results

A standard mixture containing *n*-alkanes and ketones with a carbon number ranging from seven to eleven was analyzed under isothermal conditions at 60 °C on the RTX-20 column (flow-rate 25.6 cm s^{-1}): the obtained chromatogram clearly shows an inhomogeneous retention pattern characteristic of isothermal separation (Fig. 1a). After the linearization procedure, the chromatogram looks ordered (Fig. 1b): a constant inter-distance of 0.75 min was obtained for all the terms of the *n*-alkane series as well as for the terms of different homologous series, i.e., ketones, as shown in Fig. 1b. The order introduced into the chromatogram by retention time linearization can be simply singled out by the ACVF approach: the EACF plot computed on the linearized chromatogram clearly shows well-shaped peaks at 0.75 min and 1.5 and 2.25 min, i.e., multiple inter-

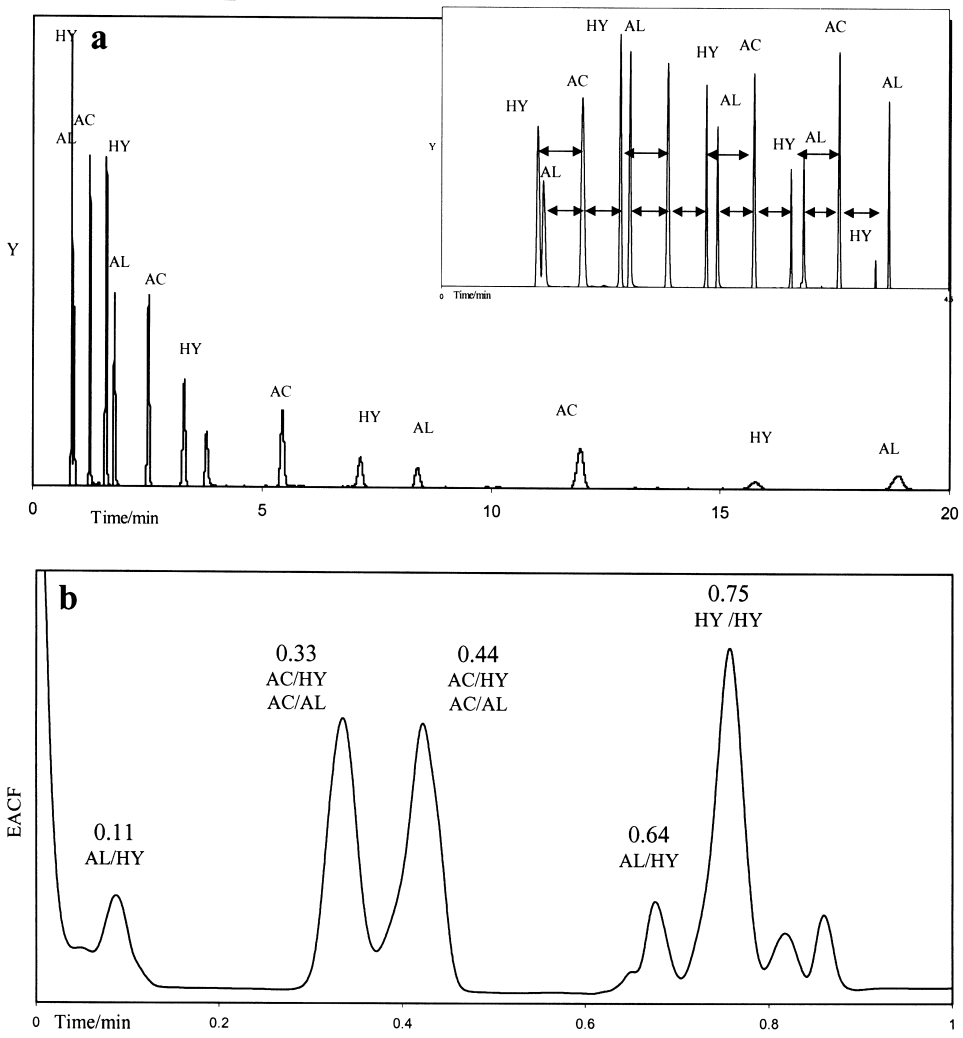

Fig. 4. Isothermal chromatogram of a standard mixture containing *n*-hydrocarbons (HY), alcohols (AL) and acetates (AC). RTX-20 column, isothermal conditions at 60 °C, flow-rate 25.6 cm s^{-1}. (a) Original chromatogram, linearized chromatogram in the inset: inter-distance repetitiveness is indicated by arrows; (b) plot of EACF computed from the linearized chromatogram.

distances (Fig. 1b, inset). Moreover, Fig. 1b shows that, after the linearization procedure, the inter-distance between compounds belonging to components of different homologue classes is also constant: ketones fall into the chromatogram between two subsequent hydrocarbons with constant inter-distance values of 0.28 and 0.47 min repeated in different parts of the linearized chromatogram. As a consequence, the EACF plot computed on the linearized chromatogram shows peaks at 0.28 and 0.47 min (Fig. 1b, inset) and these can be used to identify the presence of *n*-alkanes and ketones in the analyzed mixture.

This approach is also useful for more complex mixtures: a mixture containing three different classes of compounds—*n*-alkanes, alcohols and ketones—was analyzed at 60 °C (the obtained chromatogram is shown in Fig. 3a). After the linearization procedure an ordered chromatogram is obtained (Fig. 3a, inset): its properties can be identified in the EACF plot (Fig.

3b) where specific peaks corresponding to the repeated inter-distances in the chromatogram are evident (indicated by arrows). In addition to the previously identified peaks at 0.28 and 0.45 min corresponding to *n*-alkanes and ketones (Ke/Hy), two peaks at 0.11 and 0.65 min (Al/Hy) and two peaks at 0.2 and 0.56 min (Al/Ke) can be singled out, representing repeated inter-distances between alcohols and *n*-alkanes and between alcohols and ketones, respectively. Fig. 4a shows a chromatogram of a standard mixture containing *n*-alkanes, alcohols and acetates. The chromatogram was linearized (Fig. 4a, inset) and EACF computed: the plot (Fig. 4b) shows a large peak at 0.75 min revealing the presence of a homologous series. Moreover, the two peaks at 0.11 and 0.65 min (Al/Hy) represent repeated inter-distances between alcohols and *n*-alkanes and the two well-shaped peaks at 0.33 and 0.44 min (Ac/Hy, Ac/Al) are diagnostic for the presence of acetates, since they are due to the repeated acetate–alcohol and acetate–*n*-alkane peak inter-distances (Fig. 4b). It should be emphasized that the relative heights of the EACF peaks are related to the abundance of the corresponding repetitiveness in the chromatogram, i.e., the number of terms of the same homologous series and/or the relative abundance of each compound in the mixture.

Study of the EACF plot makes it possible to identify the chemical composition of the mixture: the peak at 0.75 is diagnostic for the presence of an homologous series in the analyzed mixture, while different peaks at smaller inter-distance values can be related to the specific chemical structure of the compounds.

The same standard mixtures simulating the composition of the extra-terrestrial atmosphere—contain-

Tables 2
EACF peaks (min) diagnostic for specific molecular structures identified in standard mixtures analyzed under GC conditions compatible with in situ analyses in space: isothermal at 60, 40 and 30 °C. Mixture composition is obtained by combining compounds in rows with those in columns

	Hydrocarbons	Alcohols	Ketones
$T = 60$ °C			
Alcohols	0.11/0.64		
Ketones	0.28/0.47	0.2/0.56	
Acetates	0.33/0.41	0.33/0.44	0.13/0.63
Alcohols/ketones	0.11/0.2/0.28/0.45/0.56/0.65		
Alcohols/acetates	0.11/0.33/0.41/0.64		
Acetates/ketones	0.13/0.33/0.41/0.63	0.13/0.21/0.33/0.44/ 0.56/0.65	
$T = 40$ °C			
Alcohols	0.28/1.49		
Ketones	0.72/1.11	0.44/1.34	
Acetates	0.8/1.0	0.8/1.01	0.36/1.42
Alcohols/ketones	0.28/0.42/0.7/1.11/1.42		
Alcohols/acetates	0.28//0.8/1.01/1.50		
Acetates/ketones	0.36/0.75/1.1/1.49	0.35/0.44/0.8/1.0/ 1.34/1.43	
$T = 30$ °C			
Alcohols	~(0.3/0.35)/2.49		
Ketones	1/1.7	0.68/2.17	
Acetates	1.2/1.60	1.25/1.60	0.59/2.25
Alcohols/ketones	0.35/0.68/1.02/1.85/ 2.17/2.49		
Alcohol/acetates	~0.35/~1.25/~1.60/~2.49		
Acetates/ketones	0.59/1/1.25/~1.7/2.25	0.59/~1.25/~1.58/ ~2.17/2.25	

ing *n*-alkanes, alcohols, acetates and ketones—were also analyzed under different isothermal conditions (30 and 40 °C) compatible with in situ analyses in space [25,26]. On the linearized chromatograms the EACF was computed and specific EACF peaks can be identified: the obtained results (summarized in Table 2) show that, under each experimental condition, it is possible to identify EACF peaks diagnostic of the presence of two or more classes of compounds present in the analyzed mixture.

The developed procedure has been applied to two specific investigations related to space research: a

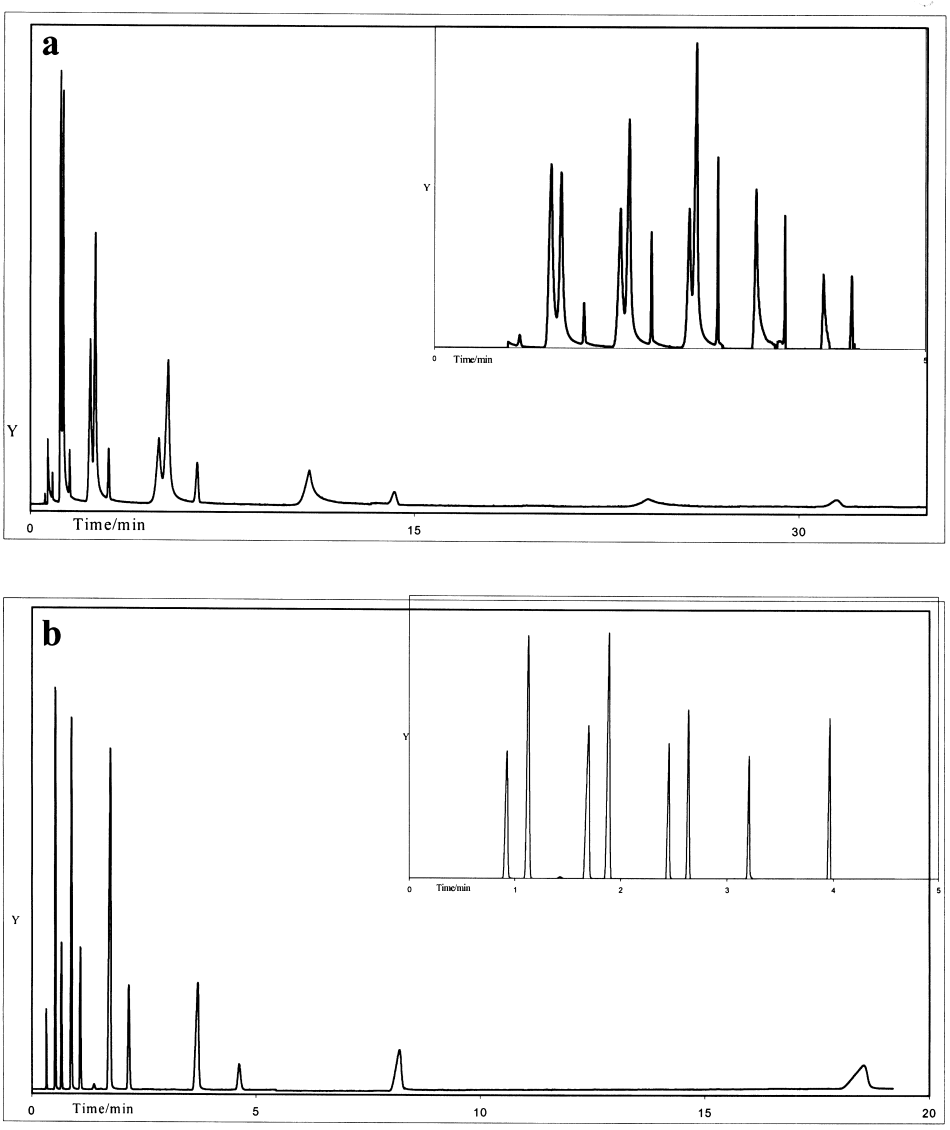

Fig. 5. Selectivity of MTX 1 and RTX-20 columns towards *n*-hydrocarbons, alcohols and ketones. Isothermal conditions at 60 °C, flow-rate 25.6 cm s^{-1}. (a) Experimental isothermal chromatogram obtained with the MTX 1 column (linearized chromatogram in the inset); (b) experimental isothermal chromatogram obtained with the RTX-20 column (linearized chromatogram in the inset); (c) plots of EACVF computed from the linearized chromatograms obtained with the MTX 1 (———) and RTX-20 (· · ·) columns.

comparison of two different GC columns in order to select the best retention selectivity and a study of the chemical composition of an unknown mixture.

4.1. Comparison of the retention selectivity of two different GC columns

A standard mixture was analyzed under the same separation conditions (60 °C with a flow-rate of 25.6 cm s^{-1}) on MTX-1 and RTX-20 columns (Fig. 5a and b). The EACF was computed for the linearized chromatograms (Fig. 5a and b, inset): a comparison of the obtained EACF plots (Fig. 5c) makes it possible to investigate the different selectivity of the two stationary phases in a very immediate way. The EACF peaks related to CH_2 retention increment—i.e., selectivity vs. carbon number—for both columns are very close (0.68 and 0.75 min for MXT 1 and RTX-20, respectively), while peaks related to the inter-distance between different classes—i.e., selectivity vs. different chemical classes—are located in different regions. For alcohol–ketone pair selectivity, for the RTX-20 column the inter-distance between alcohol and ketone peaks is well defined (0.20 and 0.56 min) as shown by the well-shaped EACF peaks. On the other hand, no baseline separation is obtained with the MXT 1 column, as shown by the shoulders of the EACF peak at 0.68 min (Fig. 5c). As a consequence, a significantly better separation can be obtained on the RTX-20 column than on MXT 1. This result is obvious for the separation of simple mixtures, such as those yielding the chromatograms shown in the inset of Fig. 5a and b, but it is not so clear with the forest of peaks present in a complex chromatogram having very crowded peaks [12]. In these chromatograms, extraction of information on specific compounds is very difficult and study of the EACF plot is very powerful in such cases.

4.2. Study of the chemical composition of an unknown mixture

An unknown sample was analyzed under isothermal conditions: a Light Petroleum (Fluka) (b.p. 100–160 °C) was submitted to GC analysis at 40 °C on the RTX-20 column (flow-rate 25.6 cm s^{-1}). The extended FA procedure, developed to study isothermal chromatograms [3,12], was applied to the obtained chromatogram (lower chromatogram in Fig. 6a): the number of components, m, can be estimated as well as the standard deviation of the narrowest, σ_1, and largest, σ_2, peaks. The obtained m value was 54: this value permits an estimate of the extent of overlapping $\gamma = p/m$, i.e., the fraction of components

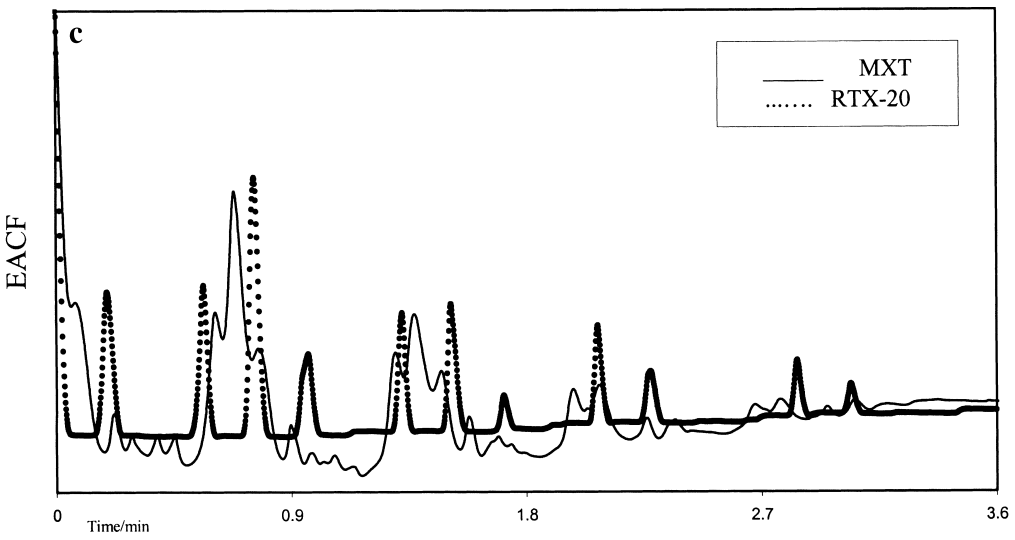

Fig. 5. (*continued*)

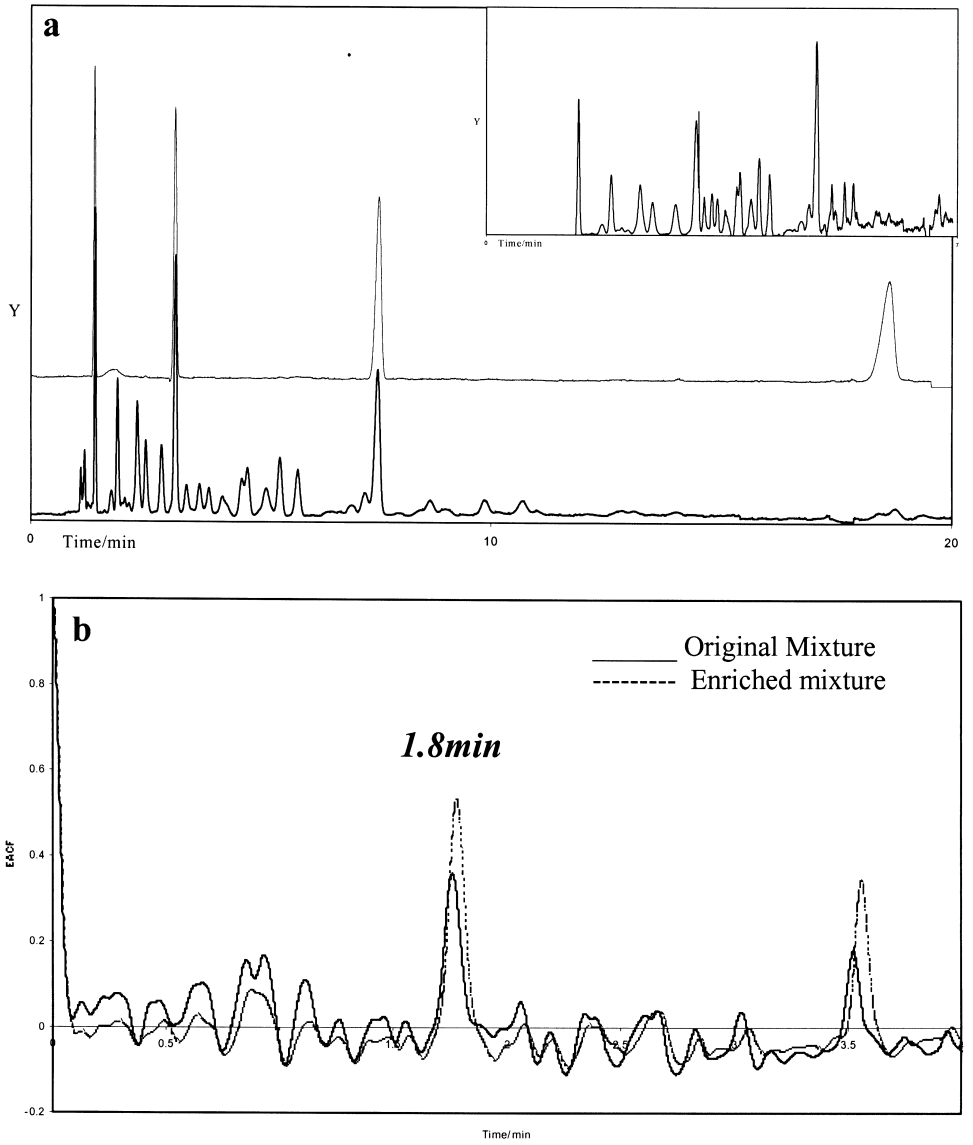

Fig. 6. GC separation of a Light Petroleum (b.p. 100–160 °C) sample. Column, RTX-20; isothermal temperature conditions, 40 °C; flow-rate, 25.6 cm s^{-1}. (a) Chromatograms of the Light Petroleum sample (lower curve) and a standard *n*-alkane mixture (upper curve). Inset: linearized chromatogram of the Light Petroleum sample. (b) Comparison between the plot of EACF computed on the linearized chromatograms of the original (———) and *n*-alkane-enriched mixture (· · ·). (c) Comparison between the plot of EACF computed on the linearized chromatograms of the original (———) and acetate-enriched mixture (· · ·).

present in the mixture, *m*, which appear in the chromatogram as peaks, *p*. Since 33 peaks can be detected in the chromatogram ($p = 33$) the obtained γ value is 0.61: this means that only 61% of the analytical information (number of components) present in the chromatogram can be observed. Also, the separation performance can be estimated by means of the standard deviation of the narrowest, $\sigma_1 = 0.01$

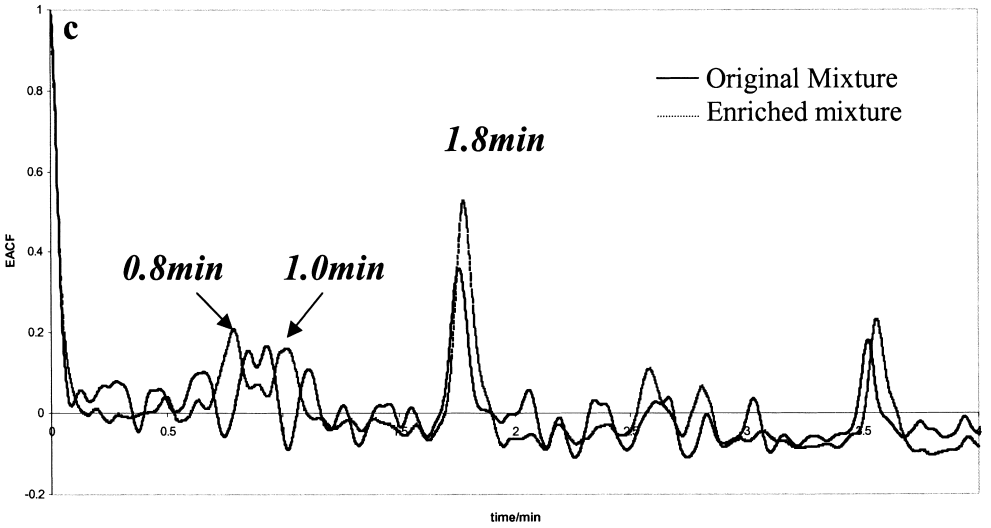

Fig. 6. (*continued*)

min, and largest peaks, $\sigma_2 = 0.06$ min. The reliability of the obtained values was checked by comparison with σ values estimated independently using a software package [12]. Moreover, in order to obtain information on sample chemical composition, the original chromatogram was linearized (Fig. 6a, inset) and the EACF function was computed (plot in Fig. 6b and c): the EACF peaks were compared with those obtained from standard mixtures (Table 2) as a diagnostic for the identification of specific molecular structures. The presence of *n*-alkanes was identified in the mixture, while alcohols, acetates and ketones were ruled out (they are beyond the abundance threshold level of the method). The presence of specific molecular structures can be checked by enriching the original sample with the corresponding standard compounds: if the original EACF peaks are significantly increased, the presence of the corresponding compounds is confirmed. This result was obtained by enriching the analyzed mixture with *n*-alkanes (upper chromatogram in Fig. 6a): their presence is confirmed by the correspondence between the EACF plots of the original and enriched chromatograms (Fig. 6b). On the contrary, the presence of acetates can be ruled out: in fact, if these standards are added to the sample, the EACF peaks obtained are significantly different (0.8 and 1 min, Table 2) from those shown in the original EACF plot (shoulders at 0.9 min, Fig. 6c).

5. Conclusions

The described extension of the chemometric FA approach to the study of isothermal chromatograms proves useful in characterizing complex chromatograms showing an inhomogeneous retention pattern, i.e., chromatograms obtained under isothermal or far from optimal temperature-programmed conditions. The reported linearization procedure makes it possible to obtain a chemical characterization of the analyzed sample since the presence of specific molecular structures can be easily identified through simple graphical inspection of the EACF plot as compared to analysis of standard mixtures. In addition, information on separation properties—i.e, mixture complexity, *m*, and system performance, σ—can be extracted from the EACF plot, as previously reported [12].

The method represents an exceptional tool for interpreting data recovered from space missions: simple handling of the original data makes it possible to extract information on the number and chemical structure of components of extra-terrestrial matter.

Moreover, the information obtained on separation performance can be used to develop proper GC instrumentation and columns that provide the best possible performance under the extreme flight conditions of future space missions.

The proposed method is the initial step in a general linearization approach: it is a simplified procedure based on an experimental reference chromatogram. Further development concerning the analytical derivatization of the mathematical function is currently being studied.

6. Nomenclature

ACVF	autocovariance function
EACVF	experimental autocovariance function
ACF	autocorrelation function
EACF	experimental autocorrelation function
FA	Fourier analysis

References

[1] A. Felinger, L. Pasti, F. Dondi, Anal. Chem. 62 (1990) 1846.
[2] A. Felinger, L. Pasti, P. Reschiglian, F. Dondi, Anal. Chem. 62 (1990) 1854.
[3] A. Felinger, L. Pasti, F. Dondi, Anal. Chem. 63 (1991) 2627.
[4] A. Felinger, L. Pasti, F. Dondi, Anal. Chem. 64 (1992) 2164.
[5] F. Dondi, A. Betti, L. Pasti, M.C. Pietrogrande, A. Felinger, Anal. Chem. 65 (1993) 2209.
[6] M.C. Pietrogrande, L. Pasti, F. Dondi, M.H. Bollain Rodriguez, M.A. Carro Diaz, J. High Resolut. Chromatogr. 17 (1994) 839.
[7] M.C. Pietrogrande, F. Dondi, A. Felinger, J. High Resolut. Chromatogr. 19 (1996) 327.
[8] F. Dondi, M.C. Pietrogrande, A. Felinger, Chromatographia 45 (1997) 435.
[9] M.C. Pietrogrande, D. Ghedini, G. Velada, F. Dondi, Analyst 123 (1998) 1199.
[10] A. Felinger, M.C. Pietrogrande, Anal. Chem. 73 (2001) 618A.
[11] M.C. Pietrogrande, P. Coll, R. Sternberg, C. Szopa, R. Navarro-Gonzalez, C. Vidal-Majar, F. Dondi, J. Chromatogr. A 939 (2001) 69.
[12] M.C. Pietrogrande, P. Coll, R. Sternberg, C. Szopa, R. Navarro-Gonzalez, C. Vidal-Majar, F. Dondi, J. Sep. Sci. 26 (2003) 569.
[13] J.M. Davis, J.C. Giddings, Anal. Chem. 55 (1983) 418.
[14] C. Giddings, Unified Separation Science, Wiley, New York, 1991.
[15] B.L. Karger, L.R. Snyder, Cs. Horváth, An Introduction to Separation Science, Wiley, New York, 1973.
[16] A.B. Littlewood, Gas Chromatography: Principles, Techniques and Applications, Academic Press, New York, 1970.
[17] L.M. Blumberg, M.S. Klee, J. Chromatogr. A 933 (2001) 13.
[18] S.O. Akapo, J.-M.D. Dimandja, D.R. Valentin, G.C. Carle, J. Chromatogr. A 842 (1999) 147.
[19] A. Aflalaye, S. Anguel, R. Sternberg, F. Raulin, C. Vidal-Madjar, J. Chromatogr. A 746 (1996) 63.
[20] R. Sternberg, C. Szopa, D. Coscia, S. Zubrzycki, F. Raulin, C. Vidal-Madjar, H. Niemann, G. Israel, J. Chromatogr. A 846 (1999) 307.
[21] C. Szopa, R. Sternberg, D. Coscia, H. Cottin, F. Raulin, F. Goesmann, H. Rosenbauer, J. Chromatogr. A 863 (1999) 157.
[22] R. Sternberg, F. Raulin, C. Vidal-Madjar, in: I.D. Wilson, E.R. Adlard, M. Cooke, C.F. Poole (Eds.), Encyclopedia of Separation Science, Vol. 9, Academic Press, London, 2000, p. 4212.
[23] P. Coll, D. Coscia, N. Smith, M.-C. Gazeau, S.I. Ramirez, G. Cernogora, G. Israel, F. Raulin, Planet. Space Sci. 47 (1999) 1331.
[24] P. Coll, J.-C. Guillemin, M.-C. Gazeau, F. Raulin, Planet. Space Sci. 47 (1999) 1433.
[25] C. Szopa, R. Sternberg, D. Coscia, F. Raulin, C. Vidal-Madjar, J. Chromatogr. A 904 (2000) 73.
[26] C. Szopa, R. Sternberg, D. Coscia, F. Raulin, C. Vidal-Madjar, H. Rosenbauer, J. Chromatogr. A 953 (2002) 165.
[27] M. Joll, D. Bowlin, J.M. Davis, Anal. Chim. Acta 302 (1995) 133.
[28] J.M. Davis, J. Microcol. Sep. 9 (1997) 193.
[29] A. Felinger, Data Analysis and Signal Processing in Chromatography, Elsevier, Amsterdam, 1998.