Talanta 83 (2011) 1302–1307

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Quantitative analysis by comprehensive two-dimensional gas chromatography using interval Multi-way Partial Least Squares calibration

Luiz Antonio Fonseca de Godoy, Marcio Pozzobon Pedroso, Leandro Wang Hantao, Ronei Jesus Poppi, Fabio Augusto*

Institute of Chemistry, State University of Campinas (Unicamp) and Instituto Nacional de Ciência e Tecnologia de Bioanalítica (INCTBio), CP 6154, 13084-971 Campinas, São Paulo, Brazil

ARTICLE INFO

Article history:
Available online 21 August 2010

Keywords:
Comprehensive two-dimensional gas chromatography
Multivariate calibration
Quantitative analysis
Allergens

ABSTRACT

A new approach for target quantitative analysis for comprehensive two-dimensional gas chromatography (GC × GC), interval Multi-way Partial Least Square (iNPLS) is presented and evaluated in this paper. In iNPLS, the two-dimensional chromatogram is split in small sections; each of these pieces is treated as an independent new chromatogram. Separated conventional NPLS calibration models for the concentration of the target analyte are built for each of the pieces of the whole chromatogram, and the best model is selected for quantitative analysis. An algorithm for iNPLS running on MatLab platform was written, preliminarily evaluated with using solutions of model compounds with different chemical properties and subsequently applied to quantify some allergens in perfume samples. The results were found to be adequate, and good precision and accuracy was obtained even for poorly resolved peaks.

1. Introduction

One of the most significant advances on the characterization of complex mixtures of volatile and semi-volatile compounds is comprehensive two-dimensional gas chromatography (GC × GC) [1–4]. A GC × GC instrument consists of two chromatographic columns interconnected by an interface called the modulator. The peak capacity of GC × GC is larger than the conventional GC, because co-eluted analytes in the first column can be separated in the second dimension. GC × GC has been applied for qualitative and/or fingerprint purposes [5–8], mainly coupled to mass spectrometry detectors; however, quantitative analysis by GC × GC is still less developed [9]. Due to the modulation process a single peak is split in several small peaks, and its quantification would demand the integration of each of these elements to obtain the total area for the target compound. Although this determination of peak areas is not as straightforward as in one-dimensional GC [10], GC × GC quantitative analysis has been using this simple strategy of adding the peak areas of modulated peaks [11]. Moreover, in spite of the high peak capacity of GC × GC, it is still possible to occasionally have co-elution between the target compound and interferents, which will affect the quantitation [11].

Other problem related to quantitative data processing in GC × GC is related to the size of data files. Since the detector signal must be digitized using high data acquisition rates (due to the small typical, peak widths, which could be as narrow as short as 80 ms at the baseline), the amount of data generated is also far greater than one-dimensional GC [4]. While the data from one-dimensional GC with a single-channel detector are basically a vector (e.g. a first order data), chromatograms generated by GC × GC equipped with the same detector can be seen as a two-dimensional matrix, which causes manual interpretation of a data set into a very difficult task (or, in many cases, virtually impossible). An alternative for analyzing this highly complex amount of data is the use of multivariate chemometric tools. Some of the suitable algorithms for analyzing high-order data already employed on GC × GC are Parallel Factor Analysis (PARAFAC [12] or PARAFAC2 [13]), Generalized Rank Annihilation Method (GRAM) [14] and Multi-way Partial Least Squares calibration (NPLS) [13,15]. Advantages such as handling missing values, achieving second order advantage, handling trilinear deviations or using incomplete calibration data can be achieved when some of these algorithms are selected for data treatment [16]. Also these chemometric methods have been applied to GC × GC data for sample classification [17–20], signal deconvolution [21–24] or target quantification [9,25–27].

In some instances, it is not always desirable to submit the entire data set to chemometric analysis. Since only small portions of the chromatogram may be relevant to a specific problem, while the remaining amount regards no significant information of the sample misleading the classification or quantification assets. Therefore, the aim of this work was to propose a multi-way algorithm called interval Multi-way Partial Least Square (iNPLS) which uses inter-

* Corresponding author. Tel.: +55 19 35213057; fax: +55 19 35213023.
E-mail address: augusto@iqm.unicamp.br (F. Augusto).

0039-9140/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2010.08.015

vals of multi-way data set to build calibration models. In the first part of this work the developed iNPLS algorithm was evaluated with standard solutions of compounds with different chemical properties. In the second step, some compounds in a real complex sample were quantified (allergens in perfume samples), in order to show the feasibility of the algorithm.

2. Theory

Partial Least Squares Calibration (PLS) is the most used chemometric method to build calibration models for first order data between independent (called **X**), instrumental responses, and dependent variables (called **Y**). The data set are organized in two matrices to perform the PLS model, one containing the independent and another one containing the dependent variables. Each row in both matrices corresponds to one sample. The algorithm decomposes the **X** ($I \times J$) and **Y** ($I \times N$) matrices into scores and loadings vectors, in which I is the number of samples, J is the number of variables and N is the number of columns in matrix **Y**, that corresponds to the numbers of compounds or physicochemical properties in analysis. The most important feature of PLS is that the decomposition is accomplished so that the successively computed score vectors have the maximum covariance between **X** and **Y** (Eq. (1)) [28].

$$\max_{w_a} \left[cov(t_a, y) \middle| \min \left(\sum_{i=1}^{I} \sum_{j=1}^{J} \left(x_{ij} - t_{a,i} w_{a,j} \right)^2 \right) \right] \quad (1)$$

where t_a ($I \times 1$) are the PLS scores of matrix **X** and w_a ($J \times 1$) are the PLS weight of matrix **X** for the ath latent variable.

In the last years a great development in analytical instrumentation has been disseminated worldwide due to the need of highly complex samples analysis. In this way, GC × GC was developed and when it is equipped with a flame ionization detector (GC × GC–FID) a matrix of second order data is obtained for each sample; the set of some samples can be organized and visualized as a box, where each level of the box consists of a sample. In consequence, Bro [29] proposed a chemometric method called NPLS to build calibrations models for second order data. The NPLS algorithm decomposes the data set **X** ($I \times J \times K$) into a set of triads. A triad consists of one score vector t, related to the first order of the data, and two weight vectors w^J and w^K, related to the second and third orders, respectively. These vectors are calculated to have the maximum covariance with the dependent variable y (Eq. (2)). NPLS does not present second order advantage, so any interferent present in an unknown sample must also be present in the calibration sample set.

$$\max_{w_a^J w_a^K} \left[cov(t_a, y) \middle| \min \left(\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(x_{ijk} - t_{a,i} w_{a,j}^J w_{a,k}^K \right)^2 \right) \right] \quad (2)$$

The regression models built by PLS and NPLS use the whole of the instrumental answer (independent variables) to find the maximum covariance with the dependent variables. But, in some cases, only one part of the instrumental data set is related with the dependent variables. So, Norgaard et al. [30] proposed a method called interval Partial Least Squares (iPLS) for first order data. It splits the data set into a number of intervals given by the analyst and a PLS model is calculated for each interval, being selected the interval with the lowest Root Mean Square of Cross-Validation (RMSECV) (Eq. (3)). Unfortunately, there is not an algorithm to perform interval selection for second order data, although iPLS algorithm could be used for this proposed through the unfolding of the second order data into a first order data. However, specifically for GC × GC data, some problems arise when the data is unfolded:

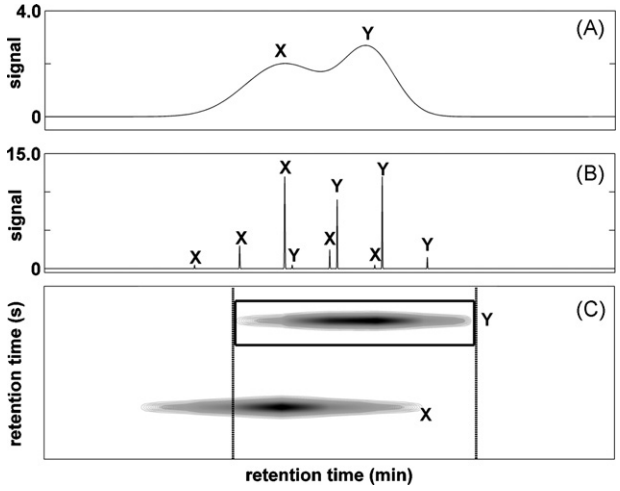

Fig. 1. (A) Co-elution between compounds **X** and **Y** at the end of the first column; (B) after modulation process, a single compound is transformed in several intense and narrow peaks; (C) GC × GC chromatogram visualized as contour plot regarding retention of first and second columns. The dashed lines delimits the part of the GC × GC chromatogram that would be used for an iPLS model while an iNPLS model would use the part of the chromatogram enclosed by the rectangle.

as a single compound is transformed in several small peaks during the modulation process, peaks from other compounds that had co-eluted in the first dimension with the target compound but were separated by the second column will also be included in the selected interval for iPLS calculation; although within the selected interval just the area of the analyte modulated peaks will vary proportionally to the analyte concentration, variation on the other peak areas may bias the calibration. Therefore, the separation achieved by the GC × GC second column is not fully exploited because the iPLS model included other compounds than the target analyte. An obvious alternative would be selecting as many intervals as the number of modulated peaks for the analyte in order to not include odd peaks in the calculation. It could be performed by siPLS [31], an extension of iPLS, which allow selecting several intervals simultaneously instead of just one and use the best combination among the all possible interval combinations. However, in a 60 min chromatographic run with detector operating at 100 Hz and typical modulated peak width at baseline of 250 ms, if siPLS were applied for a compound which is modulated five times there would be 12,000 intervals and it would result in more than 10^{17} combinations.

Therefore, it is proposed a multi-way algorithm to perform interval selection for second order data called iNPLS, which splits the data matrix in intervals in both dimensions, so a new reduced matrix is built from the initial one. A NPLS model is calculated for each new matrix and the model with the lowest RMSECV is selected.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \quad (3)$$

where n is the number of samples and $\hat{y}_i$ and y_i are the predicted and the reference values for ith sample, respectively.

Fig. 1(C) exemplifies which part of the data set would be used by NPLS, iPLS and iNPLS to build the calibration models. As the NPLS uses the entire matrix data, the whole chromatogram showed in Fig. 1(C) is used for the quantification. The iPLS selects one interval of the unfolded matrix that corresponds to the part of the GC × GC chromatogram enclosed by the dashed lines, which may contain other peaks. Finally, the iNPLS selects only the analyte peak that is enclosed by the rectangle inside Fig. 1.

Table 1
Composition of the sample set #1, % (v/v).

Sample	Toluene	Cyclohexanone	2-Octanone	1-Octanol	Undecane
1	2.0	0.0	3.0	1.0	1.0
2	2.0	3.0	0.5	0.0	0.5
3	0.0	0.0	0.0	0.0	0.0
4	0.0	1.0	0.5	3.0	1.0
5	3.0	1.0	2.0	0.0	2.0
6	3.0	0.0	1.0	0.5	0.5
7	3.0	3.0	3.0	3.0	0.0
8	1.0	3.0	2.0	0.5	1.0
9	0.5	1.0	3.0	2.0	0.5
10	2.0	1.0	0.0	0.5	3.0
11	1.0	2.0	0.0	3.0	0.5
12	3.0	0.5	0.0	2.0	1.0
13	0.0	0.5	2.0	1.0	0.5
14	0.0	3.0	1.0	2.0	3.0
15	3.0	2.0	0.5	1.0	3.0
16	0.0	2.0	3.0	0.5	2.0
17	0.5	0.0	2.0	3.0	3.0
18	2.0	2.0	2.0	2.0	0.0
19	1.0	0.5	3.0	0.0	3.0
20	0.5	2.0	1.0	0.0	1.0
21	1.0	1.0	1.0	1.0	0.0
22	0.5	3.0	0.0	1.0	2.0
23	1.0	0.0	0.5	2.0	2.0
24	0.5	0.5	0.5	0.5	0.0
25	2.0	0.5	1.0	3.0	2.0

Table 2
Composition of the perfume sample set #2 (concentration in ppm v/v).

Sample	Benzyl alcohol	Citronellol	Geraniol
1	50	0	50
2	100	75	25
3	0	50	75
4	50	50	0
5	25	25	0
6	100	25	75
7	75	75	0
8	50	25	25
9	75	50	25
10	25	75	75
11	75	100	100
12	0	25	100
13	25	100	50
14	50	75	100
15	100	0	100
16	50	100	75
17	75	25	50
18	0	100	25
19	100	50	50
20	0	0	0
21	25	0	25
22	0	75	50
23	25	50	100
24	100	100	0
25	75	0	75

3. Experimental

3.1. Samples

For a preliminary assessment of the applicability of iNPLS to GC × GC data, 25 synthetic test samples containing different concentrations of five model compounds (1-octanol, undecane, 2-octanone, cyclohexanone and toluene) were prepared in ethanol (sample set #1). The concentration of each test compound ranged between 0.0 and 3.0% (v/v) according to a Greco-Latin square planning (Table 1).

After this preliminary evaluation, iNPLS was applied to determine three allergens (geraniol, citronellol and benzyl alcohol) on a commercial perfume, which did not present the allergens analyzed. Twenty-five calibration samples (sample set #2) were prepared by spiking the perfume (50% v/v) with the allergens in concentration between zero and 100 ppm v/v according to a Greco-Latin squares planning and ethanol was added to complete 1.0 mL of total volume (Table 2).

3.2. GC × GC–FID

The analysis was performed using a lab-made GC × GC–FID prototype. This prototype was based on an Agilent G6890 GC–FID (Hewlett-Packard, Wilmington, DE, USA) fitted with a split–splitless injector and using H_2 (0.6 mL min^{-1}) as the carrier gas at constant flow rate. The cryogenic modulator here used was designed based on devices previously described in the literature [32,33]. As cryogenic fluid, nitrogen cooled by LN_2 was employed, and the heating media for band demobilization was hot nitrogen at 250 °C. The flows of cold and hot nitrogen were toggled by two three-way Asco (Florham Park, NJ, USA) solenoid valves, commanded by a DAQPad-6015 16 bits AD/DA board controlled by self-made software developed using the LabView v.8.2 programming environment (National Instruments, Austin, TX, USA). This device also digitized the FID signal and was connected to an AMD Athlon 4600 GHz Dual Core microcomputer.

For the preliminary studies (sample set #1), the column set employed was a 30 m × 0.25 mm × 0.25 μm 5% phenyl-poly(dimethylsiloxane) column (HP-5, Agilent, Avondale, PA, USA) connected to a 1.0 m × 0.10 mm × 0.10 μm 50% phenyl-poly(dimethylsiloxane) column (BPX-50, SGE Incorporated, Austin, TX, USA). The injection volume was 1 μL at 1:250 split ratio and hydrogen head pressure was 13.0 psi. The oven temperature program was: 60 °C (3 min) to 120 °C at 3 °C min^{-1} to 210 °C at 20 °C min^{-1}. The injector and detector were operated at 250 °C. The modulation period was set to 4.0 s and the data acquisition rate was 100 Hz.

For the perfume analysis (sample set #2), the column set employed was a 30 m × 0.25 mm × 0.25 μm HP-5 column (Agilent, Avondale, PA, USA) connected to a 0.9 m × 0.10 mm × 0.10 μm poly(ethyleneglycol) column (SPwax, Supelco, Bellefonte, PA, USA). The injection volume was 1 μL and the injector was operated at 1:30 split ratio and hydrogen head pressure was 11.2 psi. The oven temperature program was 60 °C to 240 °C at 3 °C min^{-1}. The modulation period was set to 4.0 s and the data acquisition rate was 100 Hz.

3.3. Data processing

The raw chromatograms were generated and stored as ASCII vector files. All calculations and graph generation were performed using MatLab 6.5 platform (MathWorks, Natick, MA, USA) fitted with the *N*-way toolbox 2.11 [available in http://www.models.kvl.dk/source/nwaytoolbox/index.asp], iToolbox 1.1 [available in http://www.models.kvl.dk/source/itoolbox/index.asp] and running on a 2.41 GHz AMD Athlon 64 X2 4600 Dual Core Processor, 2 Gb RAM microcomputer.

4. Results and discussion

To create the quantitative models, the sample sets were split in two sub-sets: a calibration set consisting of 17 samples and a prediction set with 8 samples. For the sample set #1, used on the preliminary evaluation of this approach, for each analyte the prediction set included at least one sample where its concentration was 0.0%, two samples with 0.5%, two samples with 1.0%, two samples with 2.0% and one sample with 3.0%. For the determination of

the allergens in perfume, in the prediction set at least one sample had 0.0 ppm of one of the allergens tested, two samples 25 ppm, two samples 50 ppm, two samples 75 ppm and one sample 100 ppm. To build the iNPLS models for each compound, the chromatograms in the calibration and prediction sets were split in n equal sized (Signal $\times$ $^1t_R \times {}^2t_R$) rectangular discrete pieces, whose width and height were previously selected by the user considering the peak shape expected or observed for the target analyte. For each of these sub-sets of discrete chromatogram pieces, a NPLS model for the concentration of the target analyte is estimated. The best model according to its RMSECV is selected as representative for correlating the concentration of the analyte with the chromatographic data.

The size of these discrete pieces of the raw chromatogram should be selected taking in account several aspects, being a critical parameter for the success of the procedure. Their width and height should be consistent with the base width of the chromatographic peak on both dimensions, ^1w_b and ^2w_b, and ideally should enclose the whole peak. Increasing the size of the discrete pieces will also increase the possibility of signals corresponding to other chemical species than the target analyte (eluting near the target but otherwise well resolved from it) to be included inside the rectangle, decreasing the accuracy and precision of the results. However, excessive reduction of the dimensions of the discrete pieces to be processed is also inconvenient. Obviously, if the size of the rectangles is too small compared to the width of the two-dimensional peak of the target species, the quality of the results will deteriorate. Also, decreasing the size of the discrete pieces will enlarge the number of individual NPLS models to be estimated for each analyte, which would result on an excessive time to compute and process the chromatograms.

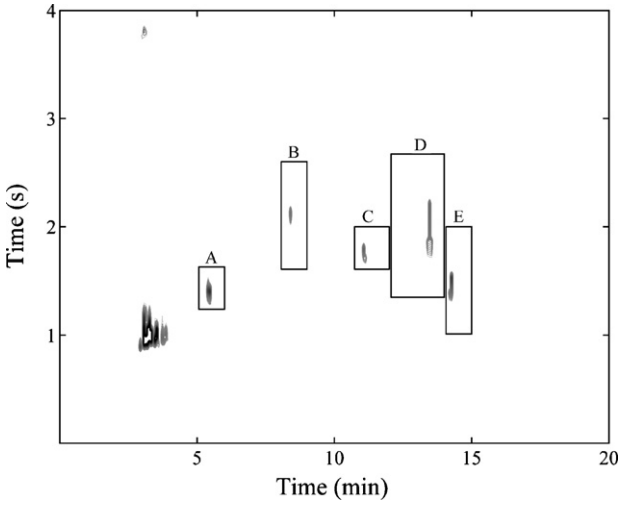

Fig. 2. GC × GC chromatogram obtained for sample 25 and the regions selected by iNPLS algorithm to build the calibrations models of toluene (A), cyclohexanone (B), 2-octanone (C), 1-octanol (D) and undecane (E).

4.1. Preliminary studies

Fig. 2 shows a typical GC × GC chromatogram for sample set #1 (sample 25). Inspection of this simple chromatogram reveals that the five model compounds are fully resolved in both dimensions. In this figure the regions of the chromatograms selected by iNPLS algorithm to build the quantitative models for each compound are also shown.

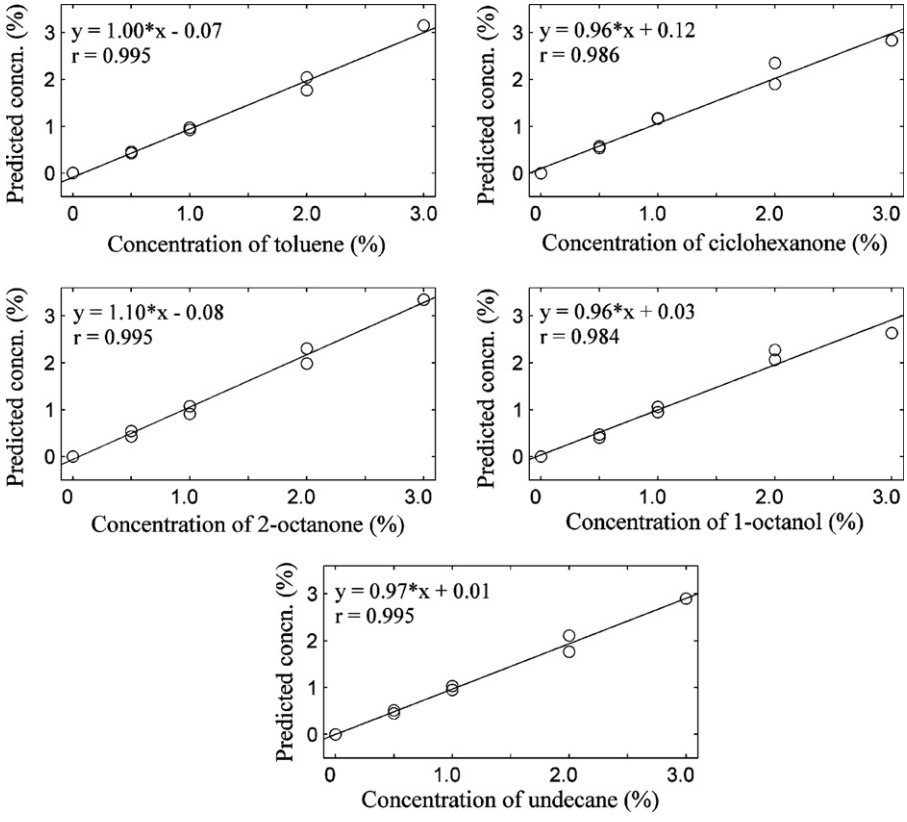

Fig. 3. Concentrations of the samples vs. predicted concentrations by iNPLS for each standard. r = correlation coefficient.

Table 3
Number of intervals in both dimensions and number of latent variables to build the iNPLS models and RMSEC, RMSECV and RMSEP values obtained.

	1D[a]	2D[b]	LV[c]	RMSEC (%)	RMSECV (%)	RMSEP (%)
Toluene	20	10	2	0.12	0.16	0.11
Cyclohexanone	20	5	2	0.19	0.25	0.17
2-Octanone	15	10	2	0.20	0.24	0.17
1-Octanol	10	3	3	0.20	0.29	0.17
Undecane	20	4	3	0.09	0.13	0.10

[a] Number of intervals in the first dimension of the GC × GC chromatograms.
[b] Number of intervals in the second dimension of the GC × GC chromatograms.
[c] Latent variables.

Table 3 presents the numbers of intervals in both dimensions (selected by the analyst), the number of latent variables, RMSEC, RMSECV and RMSEP values obtained from iNPLS models for each standard. The relationships between reference concentration versus predicted concentration by iNPLS for the five standards are presented in Fig. 3. Correlation coefficients higher than 0.984 and low RMSE values (presented in Table 3) denote that the iNPLS algorithm selected the correct region for each model and the concentration prediction for the standards was performed accurately.

4.2. Perfume samples

The iNPLS algorithm was able to identify and quantify correctly the five compounds in a sample consisted of few standards; therefore it was evaluated with a more complex sample in which the co-elution may affect the interval selection as well as the calibration model. A commercial perfume was chosen as complex sample and it was spiked with three allergens (benzyl alcohol, citronellol and geraniol) in different concentrations, as described in Table 2. iNPLS was used to build calibration models for them, in which the selected region of the GC × GC chromatograms and the quantification results were examined. The results obtained for allergens were compared to those obtained with iPLS and NPLS in order to verify the advantage of not unfold the GC × GC matrix data and also to use only one part of the chromatograms instead of the entire ones, respectively.

GC × GC chromatogram of sample 19, which was spiked with the allergens, is shown in Fig. 4 (top); in this figure the regions selected by iNPLS to build the models are marked by rectangles. The benzyl alcohol peak was recognized and it was delimited in the GC × GC chromatogram by the rectangle A; the citronellol and geraniol peaks were delimited by the rectangles B and C, respectively. As had occurred to the set of five standards, the algorithm selected the correct region for each allergen. Fig. 4 (bottom) is a zoom from Fig. 4 (top); it shows that the benzyl alcohol peak (A) is partial overlapped with the constituents of the matrix (perfume). Because of the partial co-elution, the first dimension interval length was the shortest possible to avoid selection of other peaks. Unfortunately, it was not possible to select only parts of those peaks also included inside the rectangle area, but those matrix peaks do not affect the calibration model because their areas do not vary similarly to the benzyl alcohol area. Concerning the other two allergens, the selected intervals were correctly chosen and no other peak beyond the analyte peak was included in the calibration model.

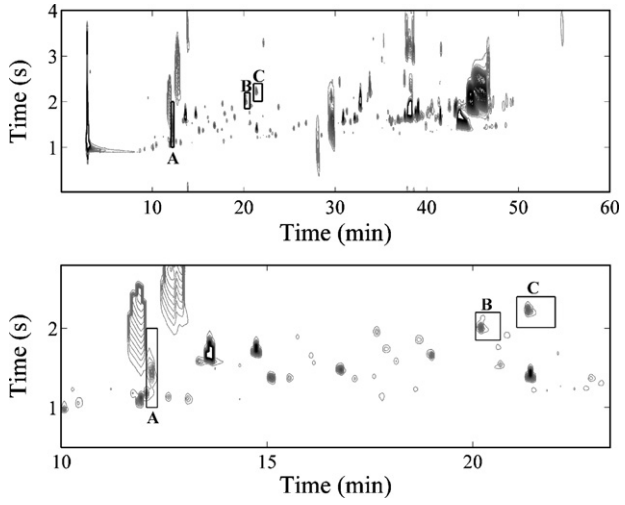

Fig. 4. GC × GC chromatogram obtained for sample 19 and the regions selected by iNPLS algorithm to build the calibrations models of benzyl alcohol (A), citronellol (B) and geraniol (C). Top all the chromatogram, bottom a zoom in the area of interest.

Table 4
Number of intervals in both dimensions and number of latent variables to build iNPLS models and RMSEC and RMSECV values obtained.

	1D[a]	2D[b]	LV[c]	RMSEC (ppm)	RMSECV (ppm)
Benzyl alcohol	180	4	2	11.4	15.2
Citronellol	90	11	5	6.9	9.3
Geraniol	60	10	2	8.9	12.3

[a] Number of intervals in the first dimension of the GC × GC chromatograms.
[b] Number of intervals in the second dimension of the GC × GC chromatograms.
[c] Latent variables.

After selection of the properly intervals, an iNPLS model was built for each allergen. The intervals in first and second dimension, the numbers of latent variables, RMSEC and RMSECV values are shown in Table 4. As the benzyl alcohol peak was partially co-eluted with other peaks, the numbers of intervals in first dimension was 180 with the purpose of select a small part of the chromatogram, as can be observed in Fig. 4 for rectangle A. When the number of intervals in first dimension were lower than 180, a part of the chromatogram was selected including the benzyl alcohol peak and a matrix peak that is co-eluted to it, as result the prediction errors increased. Regarding the rectangles B and C, the second one is the biggest because there are no peaks around it; as there are two small peaks close to the peak of citronellol, the rectangle B was a little small. In summary, the rectangles were adjusted to delimit just the analyte peak and even when co-elution in both dimensions occurred the RMSEP values were smaller than those obtained by iPLS and NPLS.

Table 5 shows the RMSEP values for iNPLS calibration models, which were compared to the results for iPLS and NPLS. Regarding the RMSEP values, it allowed concluding that iNPLS models were more reliable than iPLS or NPLS models. The RMSEP values showed the advantage of selecting only one part of the GC × GC chromatograms to build iNPLS calibration models instead of the entire chromatograms in the NPLS models. Moreover, it was evidenced that iPLS models had selected parts of the chromatogram that are not related to the analyte concentration (for iPLS were used 36 intervals for benzyl alcohol model and 72 intervals for citronellol and geraniol models). The graphs of predicted concentration versus reference concentration show good linearity and correlation coefficient (Fig. 5), which demonstrate the reliability of these models.

As the NPLS does not have second order advantage, the iNPLS algorithm has the same limitation. However, in the iNPLS is pos-

Table 5
RMSEP obtained for iNPLS, iPLS and NPLS models.

	RMSEP (ppm)		
	iNPLS	iPLS	NPLS
Benzyl alcohol	10.2	32.4	28.6
Citronellol	4.5	6.5	39.3
Geraniol	6.5	11.2	44.9

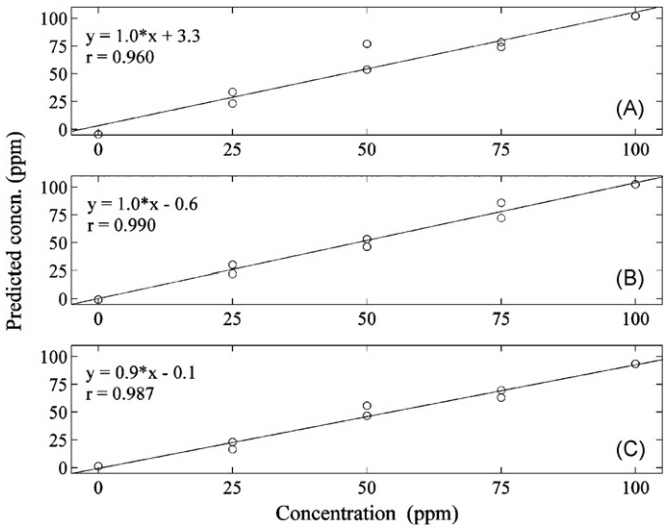

Fig. 5. Concentrations of the samples vs. predicted concentrations by iNPLS for each allergen. *r* = correlation coefficient.

sible to analyze an unknown sample that contains interferents not present in the calibration set since those interferents are not included in the sub-matrix selected by the iNPLS to build the model. Otherwise, if the interferents are included in the sub-matrix selected by the iNPLS, the algorithm is not able to quantify the analyte. The proposed approach is suitable for routine quality control analysis in which the sample is well known and not calibrated interferents will not be present. Otherwise, a procedure for outlier detection can be implemented in the quality control to exclude anomalous samples.

5. Conclusions

The iNPLS approach presented here was shown to allow precise and accurate quantitation of discrete analytes from GC × GC chromatograms. The main feature of this approach is that it is not necessary to perform any operation to detect the pertinent chromatographic peaks prior the quantitation: on its preliminary steps, the algorithm checks models for all sections of the chromatograms, discarding the part which contains the relevant signal based on inspection and comparison of RMSEC values. This is an automated operation, which does not need intervention of the analyst. Also, for GC × GC and from the final user's point of view this can be significantly less complex than performing conventional peak detection and integration: the resulting chromatograms can have literally thousands of separated peaks, and the analysis of the resulting integration tables can be cumbersome, at least for less known samples. The only significant user-selected parameter is the size of the discrete pieces, which ideally should match the first and second dimension widths of the chromatographic peak as much as possible (although this is not mandatory and easily adjusted); a mere visual inspection of the chromatogram can provide good initial estimates for the dimensions of the integration piece. Several extensions of this algorithm can be made to further improve its applicability and to make it more user-friendly, such as the inclusion of steps to pre-inspection the data and suggest (or even automatically select) the size of discrete pieces, to its extension to four-dimension GC × GC–MS chromatograms.

Acknowledgments

The authors thank FAPESP (Foundation for Research Support of the State of São Paulo), CNPq (Brazilian National Council for Research and Technological Development) and CAPES (Brazilian Ministry of Education Agency for Improvement of Graduate Personnel) for financial support and fellowships.

References

[1] Z. Lui, J.B. Phillips, J. Chromatogr. Sci. 29 (1991) 227.
[2] J.B. Phillips, J. Beens, J. Chromatogr. A 856 (1999) 331.
[3] T. Górecki, J. Harynuk, O. Panic, J. Sep. Sci. 27 (2004) 359.
[4] J. Dallüge, J. Beens, U.A.T. Brinkman, J. Chromatogr. A 1000 (2003) 69.
[5] C. Cordero, C. Bicchi, P. Rubiolo, J. Agric. Food Chem. 56 (2008) 7655.
[6] T. Cajka, J. Hajslova, F. Pudil, K. Riddellova, J. Chromatogr. A 1216 (2009) 1458.
[7] L. Mondello, A. Casilli, P.Q. Tranchida, P. Dugo, R. Costa, S. Festa, G. Dugo, J. Sep. Sci. 27 (2004) 442.
[8] L.R. Bordajandi, L. Ramos, M.J. González, J. Chromatogr. A 1078 (2005) 128.
[9] V.G. van Mispelaar, A.C. Tas, A.K. Smilde, P.J. Schoenmakers, A.C. van Asten, J. Chromatogr. A 1019 (2003) 15.
[10] M. Kallio, T. Hyötyläinen, J. Chromatogr. A 1148 (2007) 228.
[11] O. Amador-Muñoz, P.J. Marriott, J. Chromatogr. A 1184 (2008) 323.
[12] L.R. Snyder, J.C. Hoggard, T.J. Montine, R.E. Synovec, J. Chromatogr. A 1217 (2010) 4639.
[13] L.A.F. de Godoy, E.C. Ferreira, M.P. Pedroso, C.H.V. Fidelis, F. Augusto, R.J. Poppi, Anal. Lett. 41 (2008) 1603.
[14] C.G. Fraga, B.J. Prazen, R.E. Synovec, Anal. Chem. 73 (2001) 5833.
[15] M.P. Pedroso, L.A.F. de Godoy, E.C. Ferreira, R.J. Poppi, F. Augusto, J. Chromatogr. A 1201 (2008) 176.
[16] G.M. Escandar, N.M. Faber, H.C. Goicoechea, A.M. de la Peña, A.C. Olivieri, R.J. Poppi, Trends Anal. Chem. 26 (2007) 752.
[17] K.J. Johnson, B.J. Prazen, R.K. Olund, R.E. Synovec, J. Sep. Sci. 25 (2002) 297.
[18] L. Xie, P.J. Marriott, M. Adams, Anal. Chim. Acta 500 (2003) 211.
[19] C.G. Fraga, B.J. Prazen, R.E. Synovec, J. High Res. Chromatogr. 23 (2000) 215.
[20] C.G. Fraga, B.J. Prazen, R.E. Synovec, Anal. Chem. 73 (2003) 5833.
[21] A.E. Sinha, J.L. Hope, B.J. Prazen, E.J. Nilsson, R.M. Jack, R.E. Synovec, J. Chromatogr. A 1058 (2004) 209.
[22] R.E. Mohler, K.M. Dombek, J.C. Hoggard, K.M. Pierce, E.T. Young, R.E. Synovec, Analyst 132 (2007) 756.
[23] J.C. Hoggard, W.C. Siegler, R.E. Synovec, J. Chemometr. 23 (2009) 421.
[24] J.C. Hoggard, J.H. Wahl, R.E. Synovec, G.M. Mong, C.G. Fraga, Anal. Chem. 82 (2010) 689.
[25] K.J. Johnson, B.J. Prazen, D.C. Young, R.E. Synovec, J. Sep. Sci. 27 (2004) 410.
[26] R.E. Mohler, K.M. Dombek, J.C. Hoggard, E.T. Young, R.E. Synovec, Anal. Chem. 78 (2006) 2700.
[27] S. Yanga, M. Sadilek, R.E. Synovec, M.E. Lidstrom, J. Chromatogr. A 1216 (2009) 3280.
[28] R.G. Beretron, Analyst 125 (2000) 2125.
[29] R. Bro, J. Chemometr. 10 (1996) 47.
[30] L. Norgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Appl. Spectrosc. 54 (2000) 413.
[31] L. Munck, J.P. Nielsen, B. Moller, S. Jacobsen, I. Sondergaard, S.B. Engelsen, L. Norgaard, R. Bro, Anal. Chim. Acta 446 (2001) 169.
[32] J. Harynuk, T. Górecki, J. Chromatogr. A 1019 (2003) 53.
[33] M. Pursch, P. Eckerle, J. Biel, R. Streck, H. Cortes, K. Sunc, B. Winniford, J. Chromatogr. A 1019 (2003) 43.