

Study of quality controls for stability check of the ROIs of a ketones mixture in different GC-IMS measurement campaigns

Celia Mallafré-Muro
Signal and Information Processing for Sensing Systems
Institute for Bioengineering of Catalonia (IBEC)
Department of Electronics and Biomedical Engineering
University of Barcelona
Barcelona, Spain
cmallafr@ibecbarcelona.eu

Luis Fernández Romero
Signal and Information Processing for Sensing Systems
Institute for Bioengineering of Catalonia (IBEC)
Department of Electronics and Biomedical Engineering
University of Barcelona
Barcelona, Spain
lfernandez@ub.edu

Meryl Cruz
Signal and Information Processing for Sensing Systems
Institute for Bioengineering of Catalonia (IBEC)
Barcelona, Spain
mcruz@ibecbarcelona.eu

Antonio Pardo Martínez
Department of Electronics and Biomedical Engineering
University of Barcelona
Barcelona, Spain
a.pardo@ub.edu

Arnau Blanco I Borrego
Signal and Information Processing for Sensing Systems
Institute for Bioengineering of Catalonia (IBEC)
Barcelona, Spain
ablancob@ibecbarcelona.eu

Santiago Marco Colás
Signal and Information Processing for Sensing Systems
Institute for Bioengineering of Catalonia (IBEC)
Department of Electronics and Biomedical Engineering
University of Barcelona
Barcelona, Spain
smarco@ibecbarcelona.eu

Abstract—GC-IMS is a very good complementary technique to traditional GC-MS, that presents some advantages, but also, some disadvantages such as misalignments produced by many parameters affecting the equipment stability. The reproducibility of the measures has been studied in two different measurement campaigns with a set of automatized quality control parameters. Figures of merit from one region of interest present in the samples show that the saturation and asymmetry do not change between measurement campaigns, but the volume and area of the total ion spectra change. A correction of these changes between batches should be developed.

Keywords—GC-IMS, reproducibility, batches, ROIs, figures of merit

I. INTRODUCTION

Gas Chromatography coupled to Ion Mobility Spectrometer (GC-IMS) is one complementary technique to Gas Chromatography couple to Mass Spectrometry (GC-MS). GC-IMS initially separates the volatile compounds present in the sample using a chromatographic column. The molecules that elute from the column are ionized by interacting with the reactant ions, producing product ions. The humidity present in the air and the drift gas used in the instrument form the Reactant Ion Peak (RIP) and the reactant ions. Then, a second separation takes place according to the products ion mobility, which depends on the ion mass and cross-section [1]. The ion mobility is related to the time of flight of the ions inside the drift tube, wherein the lower the ion mobility, the lower the time they remain in the drift tube.

A detector is found at the end of the drift tube, where the

ions impact, and thus creating a mobility spectrum made of all the drift time reads. For each sample measure, a 2D data matrix is obtained with the intensities detected and the information of the retention times obtained from the chromatographic column in one axis, and the drift time information in the other. Within the matrix, there are regions of interest (ROIs) where peaks corresponding to molecules detected are present. The data produced is high dimensional and contains very complex information, which is hard to handle and needs analysis by an expert in the field.

The GC-IMS allows a fast, reliable, and more economical chromatographic method using either a standard capillary column or a multicapillary columns (MCC). Besides the previously mentioned advantages, it can be easily moved since GC-IMS avoids the use of vacuum systems and can work at ambient pressure [2] with the detection limit in the order of ppb. However, it also has some limitations, such as moderate selectivity, wide peaks, high degree of coelutions and equipment instability.

The objective of this study is to present a method that checks the reproducibility of the samples measured in different measurement campaigns using a set of automatized quality control parameters to determine if the samples between measurement campaigns need to be corrected.

II. METHODS

A. Data pre-processing

The raw data obtained needs previous steps before the data analysis since the dimensionality of the data is very

high, which makes the chemical information sparse. Much of the data matrix does not contain relevant information, so the data has been cut, keeping only the region of interest and reducing the size of the data matrix. With the objective of diminishing instrumental drifts, some pre-processing steps were applied (Fig. 1). These techniques were applied first in the drift time axis, and then in the retention time. The first step was a digital smoothing for improving the signal-to-noise ratio using a Savitzky-Golay filter. Then, the baseline was corrected using a modification of the traditional asymmetric least square algorithm, called psalsa [3], with the same parameters of a previous study [4]. Finally, the peaks of the samples were aligned using different methods according to the time axis. For the drift time axis alignment, Correlation Optimized Warping [5] with a segment length of 25 and a slack parameter equal to 10 was applied. In the retention time axis, Parametric Time Warping was performed, first with a global warping of first order, and a second individual warping of second order. Similar pre-processing workflows have been applied in food analysis [4], [6]

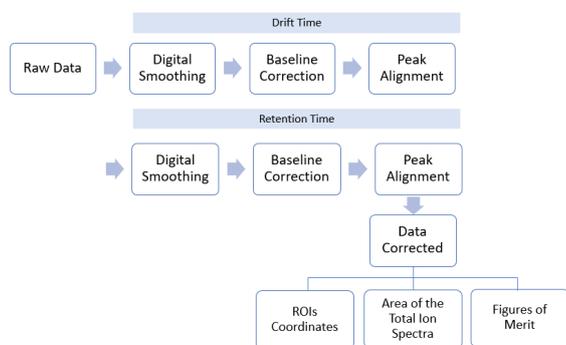


Fig. 1. Pre-processing steps applied, first in drift time and then in retention time, and parameters extracted from each sample.

B. Effect in GC-IMS reproducibility

It is very common to have experiments that need more than one day of measurements or even different measurement campaigns in different months. It is very important to check if the equipment is stable and the measures are comparable among them. For this purpose, two measurement batches of ketones mixtures have been done in different months, where the first batch was measured in June and the second batch was measured in November. These two batches were then compared.

As mentioned earlier, GC-IMS can work properly at ambient pressure. Nevertheless, there are many parameters that can affect the stability of the equipment and the reproducibility of the measures, including the temperature and the flow of the gases. The ambient pressure can also affect the RIP, decreasing the total amount of charge or shifting its drift time.

If the RIP is not at its optimal setup, the total amount of charge available will be lower than it is supposed to be, and consequently, the total reactant ions available will be different than they should be, affecting the intensity of the compounds detected.

In each measure, the different peaks or ROIs can be characterized by their asymmetry, the saturation with respect

to the RIP, and the volume. The asymmetry degree is calculated as the difference between the half the upper area and the lower area divided by the whole area. On the other hand, the volume is the sum of all the intensities within a ROI, where the logarithm of the volume is also used. Lastly, the saturation with respect to the RIP means that one ROI can saturate the RIP if the intensity of the RIP is lower than the 10% of its maximum. These characteristics are unique to each ROI and depend on the parameters used for the measurement. These ROIs can then be assessed through different figures of merit. Furthermore, the coordinates in retention and drift times of each ROI are unique and are also dependent on the measurement parameters. The combination of the figures of merit and the ROIs coordinates is a characteristic of each sample. With stable instrument conditions, these characteristics should be very similar for measurements of the same sample.

C. Ketones analysis with GC-IMS

The ketones chemical group presents a medium proton affinity, being able to be detected by the IMS as it shows a mid-level of selectivity.

The ketones group is widely used for testing the equipment and normalizing the retention indexes [7]. The typical image of a GC-IMS ketones mixture (Fig. 2) shows a good separation in the retention time axis of the different ketones, while in the drift time axis, the isomers of each ketone are easily identified.

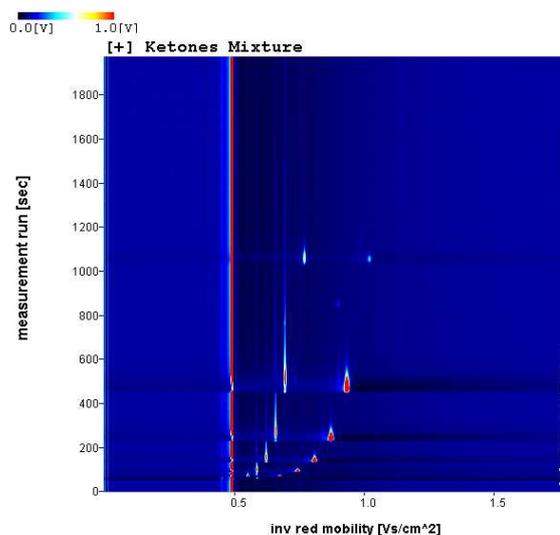


Fig. 2. Image obtained from a ketones mixture measurement.

As the molecular weight of the ketone increases, its retention time also increases. Also, changes in the mobility are noticed.

According to the equipment manufacturer, a mixture of 6 ketones can be used to check the equipment: 2-Butanone, 2-Pentanone, 2-Hexanone, 2-Heptanone, 2-Octanone and 2-Nonanone. The analytical standard grade (purity $\geq 99.9\%$) ketones used in the analysis were purchased from Sigma-Aldrich. Stock solution was prepared using the same concentration of each ketone ($0.116 \text{ g}\cdot\text{ml}^{-1}$), and then working solutions, S1 ($23.2 \text{ mg}\cdot\text{l}^{-1}$) and M1 ($464 \text{ }\mu\text{g}\cdot\text{l}^{-1}$), were prepared by diluting it in milli-Q water. Nitrogen was used as carrier and drift gas. The experimental gas flow rates conditions were 200.0 [ml/min] for the drift gas and 11.0

[ml/min] for the carrier gas. The column and IMS temperature were settled at 60°C while the injector and transfer line were at 80°C.

The first batch of measures was done in June while the second measurement campaign was done in November the same year.

D. Data Analysis for checking reproducibility

With the objective of studying the reproducibility of the samples in different measurement campaigns, two statistical analyses have been performed. The distribution of the different figures of merit between samples has been studied with the Kolmogorov-Smirnov test. It is a non-parametric test that checks and compares the distributions among two samples [8]. The null hypothesis is that the two samples are drawn under the same distribution. Complementary, one ROI has been selected in all the samples and the figures of merit have been compared in a t-test between measurement campaigns.

III. RESULTS

The first analysis performed was the comparison of the number of ROIs detected in each sample and the comparison of the retention and drift time coordinates of all the ROIs in the same batch. In all the samples from the first batch, 21 ROIs were detected. When comparing the coordinates ROI by ROI, the maximum time shift in retention time found was 1.14 seconds and 0.09 milliseconds in the drift time axis. The mean time drift among all the ROIs of all the samples in the first batch of measurements was 0.41 second in retention time and 0.04 milliseconds in drift time.

For the second batch of measures, the number of ROIs detected in each sample was also 21. When comparing the coordinates ROI by ROI, the maximum shift in retention time found was 0.24 seconds and 0.12 milliseconds in the drift time axis. The mean shift among all the ROIs of all the samples in the first batch of measurements was 0.08 second in retention time and 0.03 milliseconds in drift time.

When comparing the inter-batch time misalignment, which means comparing the samples from the first batch with the samples from the second batch, it was observed that the mean shift in the retention time was 6.8 seconds, while the mean shift in drift time was 2.05 milliseconds (Table I).

TABLE I. SUMMARY OF TIME DRIFTS IN DRIFT AND RETENTION TIME

	Intrabatch		Interbatch
	Batch 1	Batch 2	
Drift Time	Max: 0.09 ms	Max: 0.12 ms	2.05 ms
	Mean: 0.04 ms	Mean: 0.03 ms	
Retention Time	Max: 1.14 s	Max: 0.24 s	6.8 s
	Mean: 0.42 s	Mean: 0.08 s	

The Kolmogorov-Smirnov test was applied to check if the ROIs of the samples from different measurement campaigns was following the same distribution. When studying the asymmetry, the p-value obtained was lower than 0.05 (p-value < 2.2e-16), which means that the distribution of the asymmetry is different between the two batches. The

same result was obtained when analysing the volume of the ROIs with a p-value lower than 0.05 (p-value < 2.2e-16). In contrast, the saturation distribution was the same for both batches, with a p-value of 0.067.

The comparison of the figures of merit of one ROI between measurement campaigns was done with a t-test. Using the logarithm of the volume, it showed a p-value lower than 0.05, which means that the volumes are different between batches. In the first batch, the logarithm of the volume tends to have higher values, what means that the volume of the ROI in the second measurement campaign is smaller (Fig. 3). In contrast the saturation and the asymmetry t-test do not show a significant p-value when comparing them between the two measurement campaigns (Table II).

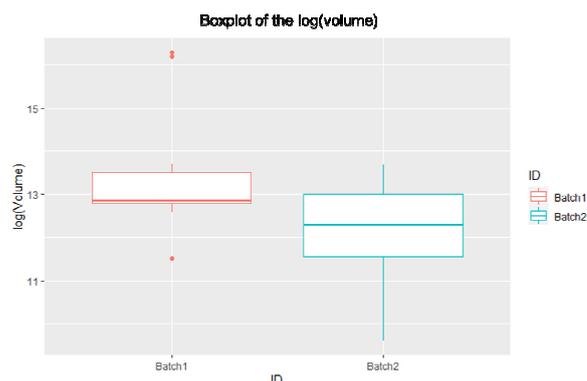


Fig. 3. The logarithm of the volume from one ROI was compared between the 2 batches of measurement campaigns. In the first batch, the logarithm of the volume tends to have higher values in blue.

Finally, the area of the total ion spectra of the samples from the 2 different batches were compared, observing a decreasing in the total area of the samples from the second measurement campaign.

TABLE II. SUMMARY OF THE RESULTS OBTAINED FROM THE STATISTICAL ANALYSIS

Statistical Analysis	Figures of Merit	p-Value
T-Test from one ROI	log(volume)	< 0.05
	Asymmetry	> 0.05
	Saturation	> 0.05
Kolmogorov-Smirnov	log(volume)	< 2.2e-16
	Asymmetry	< 2.2e-16
	Saturation	> 0.05

IV. CONCLUSIONS

Comparison between samples from different measurement campaigns is necessary due to some misalignments that can be produced throughout the measurement. Visual inspection can be performed if the number of samples is reduced. However, when a large number of samples is present, some automatize quality control parameters are more suitable to evaluate the data.

In this study, it has been proven that the data obtained from different measurement campaigns may vary, making the results not repeatable. Reproducibility can be affected by several experimental parameters including time. The misalignment in the drift time found in the samples was very

small compared to the shifts in the retention time, making it more difficult for future peak matching between samples. After performing all the pre-processing steps, the analysis of the figures of merits of the ROIs of each sample showed that in the measures from different batches, the distribution of these figures of merit change. For the volume figure of merit, the volume showed changes between batches, which can be caused by any change in various experimental parameters. In contrast, the asymmetry and saturation figures of merit remained constant among the batches, which is consistent with the fix concentration of the ketones in the original mixture. The figures of merit can work as automatized quality controls that determine the reproducibility of the samples. It is a necessity to make the samples from different measurement campaigns comparable due to the different projects that can be running at the same time in different countries, or even, in different times. These shifts and changes of distributions of the figures of merits should be corrected for this purpose.

ACKNOWLEDGMENT

This research was supported by the Spanish MINECO coordinated project TensorChrom (Total2DChrom, ref. RTI2018-098577-B-C21 and TENSOMICS RTI2018-098577-B-C22). CM is thankful for her pre-doctoral fellowship from the UB APIF program (ref. APIF_18_19).

REFERENCES

- [1] Gary A. Eiceman, Z. Karpas, and J. Herbert H. Hill, *Ion Mobility Spectrometry, 3rd Edition*. 2014.
- [2] A. B. Kanu and H. H. Hill, "Ion mobility spectrometry detection for gas chromatography.," *J. Chromatogr. A*, vol. 1177, no. 1, pp. 12–27, Jan. 2008, doi: 10.1016/j.chroma.2007.10.110.
- [3] S. Oller-Moreno, A. Pardo, J. M. Jimenez-Soto, J. Samitier, and S. Marco, "Adaptive Asymmetric Least Squares baseline estimation for analytical instruments," *2014 IEEE 11th Int. Multi-Conference Syst. Signals Devices, SSD 2014*, pp. 1–5, 2014, doi: 10.1109/SSD.2014.6808837.
- [4] R. Freire *et al.*, "Full workflows for the analysis of gas chromatography—ion mobility spectrometry in foodomics: Application to the analysis of iberian ham aroma," *Sensors*, vol. 21, no. 18, 2021, doi: 10.3390/s21186156.
- [5] G. Tomasi, F. Van Den Berg, and C. Andersson, "Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data," *J. Chemom.*, vol. 18, no. 5, pp. 231–241, 2004, doi: 10.1002/cem.859.
- [6] L. Fernandez, A. Martin-Gomez, M. Mar Contreras, M. Padilla, S. Marco, and L. Arce, "Ham quality evaluation assisted by gas chromatography ion mobility spectrometry," 2017, doi: 10.1109/ISOEN.2017.7968852.
- [7] E. Krylov, E. G. Nazarov, R. A. Miller, B. Tadjikov, and G. A. Eiceman, "Field dependence of mobilities for gas-phase-protonated monomers and proton-bound dimers of ketones by planar field asymmetric waveform ion mobility spectrometer (PFAIMS)," *J. Phys. Chem. A*, vol. 106, no. 22, pp. 5437–5444, 2002, doi: 10.1021/jp020009i.
- [8] M. Aslam, "Introducing Kolmogorov-Smirnov Tests under Uncertainty: An Application to Radioactive Data," *ACS Omega*, vol. 5, no. 1, pp. 914–917, 2020, doi: 10.1021/acsomega.9b03940.