GLOBAL ANALYTICAL SOLUTIONS

GERSTEL

AppNote 3/2003

Use of Calibration Transfer Algorithms on a Mass Spectrometry Based Chemical Sensor - Preliminary Results

Arnd C. Heiden, Carlos Gil
Gerstel GmbH & Co. KG, Eberhard-Gerstel-Platz 1, D-45473 Mülheim an der Ruhr, Germany

Vanessa R. Kinton, Edward A. Pfannkoch
Gerstel, Inc., Caton Research Center, 1510 Caton Center Drive, Suite H, Baltimore, MD 21227, USA

L. Scott Ramos, Marlana Blackburn
Infometrix, Inc., P.O. Box 1528, Woodinville, WA 98072, USA

Keywords

Chemometrics, ChemSensor, electronic nose, mass spectrometer (MS), calibration transfer, discrimination, fingerprint mass spectra.

Abstract

A benefit of using a mass spectral based chemical sensor is its fast analysis time. The successful use of this type of instrument requires training the sensor with standard samples. The mass spectra from the standards are used to create chemometric models and classification of unknown samples is obtained by projecting their mass spectra into these models. Predictions are very dependent on the quality of the model and reliable models are created with numerous replicas. Construction of these calibration models requires considerable amount of laboratory time and is therefore costly. If the sensor is disturbed (e. g., filament is changed

or maintenance is required) recalibration of the instrument is necessary to compensate for the new instrument conditions. An alternative to recalibration is the use of calibration transfer algorithms.

A series of food samples and individual compounds were used to monitor the reliability of the calibration transfer algorithms for a period of 10 weeks. The instrument was disrupted in 3 ways, (a) filaments were replaced; (b) tuning algorithms were changed and (c) preventive maintenance was applied. Instrument drift over the 10-week period was also investigated in a separate system and different types of calibration transfer algorithms were examined.

Preliminary results indicate that the successful use of calibration transfer is highly dependent on the type and number of transfer samples and model type. Overall KNN models appear to be more robust than SIMCA ones. The best accuracy was obtained using KNN models and SIMCA models with calibration transfer.

Introduction

Mass spectrometry is an analytical technique widely used for qualitative and some quantitative analysis. A mass selective detector (MSD) can be directly coupled to a headspace sampler with excellent time savings benefits. This type of hyphenated instrumentation is the basis of mass spectrometry based chemical sensors. These sensors introduce the entire headspace of samples into the MSD without chromatographic separation and classification of samples is performed with multivariate analysis. Because the data analysis is based on the mass spectral fingerprints, a robust fingerprint is desired. This study presents preliminary results of a time study in which mass spectral fingerprints of different compounds were monitored over time. We will investigate the effects of tuning, filament change and ion source maintenance on the mass spectral profiles.

The effects of tuning on mass spectral fingerprints will be investigated by changing tuning algorithms. The goal of tuning an MSD is to optimize sensitivity, mass assignment and relative ion abundance. It is therefore possible that mass spectral fingerprints could change slightly after each tuning. The parameters that can be adjusted during tuning are: lenses, quadrupole filter, electron multiplier, mass axis, and peak width calibration. Tuning requires several iterations since optimizing one parameter affects the optimal setting of another. For this study a GERSTEL Headspace ChemSensor (Figure 1) that includes a 5973 N (Agilent Technologies) detector was used. This MSD uses PFTBA (perfluorotributylamine) as the tuning compound for the MS tunings. PFTBA has ions at 31, 50, 69, 100, 131, 219, 264, 414, 464, 502, 576, and 614 amu. Depending on what m/z needs to be optimized, different tuning algorithms are provided with the instrument control software.

Figure 1. Gerstel Headspace ChemSensor.

A vital part of the MSD that is not optimized with tuning is the electron source or filament. An emission current heats the filament and causes them to emit electrons. The electrons emitted by the filament are responsible for the ionization of the sample molecules. If a filament is exchanged small differences in the mass spectral profiles could be obtained.

Another crucial part that determines the MSD performance is the ion source; a dirty ion source can cause poor sensitivity. The rate at which the ion source will require preventive maintenance will depend on three parameters [1]:

- The quantity of samples analyzed, usually the more samples analyzed the quicker the ion source will need cleaning.
- The nature of samples and sample matrix, usually highly concentrated samples in complex matrices dirty the ion source more quickly than trace samples in clean matrices.
- Operating conditions (mainly ion source temperature), higher ion source temperatures generally keep the ion source cleaner longer.

The minor differences in mass spectral fingerprints obtained from changing the above parameters can significantly reduce the reliability of chemometric models. To compensate for these minor differences in mass spectral profiles two alternatives exist. One option is to create a new model with new data, which can be time consuming, or a different alternative is to use a computational adjustment that compensates for instrument differences. The second approach is known as transfer of calibration (TOC).

The TOC algorithms available with the GERSTEL ChemSensor adjusts profiles obtained with the new set of instrument conditions to look like those collected before any parameters were changed. For this study we will apply the TOC known as direct standardization. Direct TOC relates all variables in the data collected with initial instrument settings to each corresponding variable measured after slightly different instrument conditions. Calibration transfer algorithms can be performed for either quantitative (e.g., PLS or PCR), or qualitative models (KNN or SIMCA). In this study we will investigate the efficiency of the calibration transfer using SIMCA and PLS models.

EXPERIMENTAL

A. Tuning study Data set 1 - Camphor. Data collected using three different tunings in the SCAN mode. Two camphor solutions were prepared (10 and 100 ppm in 1% Methanol). Ten replicas of each concentration were analyzed using 1 ml aliquots. The aliquots were placed in 10 mL vials which were crimped and equilibrated for 20 minutes at 80 °C before headspace sampling.

Three different tuning algorithms were tested: Autotune (atune), BFB Tune and Standard Spectra Tune (stune).

B. Filament replacement Data set 2 - Citrus oils. Data collected with two different filaments in the SCAN mode. 6 replicas of three different citrus oils were tested. 7.5 µL aliquots were placed in 20 mL vials which were equilibrated for 5 min at 80 °C before headspace sampling.

C. Ion source cleaning Data set 3 - Limonene. Data collected before and after cleaning the ion source in the SCAN mode. Three different solutions of Limonene (1000, 5000 and 10000 ppm in Methanol) were prepared. 6 replicas of each level were analyzed using 10 µL aliquots of the stock solutions. The samples were equilibrated for 15 min at 55 °C before headspace sampling.

Results and Discussion

A. Tuning study Data set 1 - Camphor. The mass spectral fingerprint for the 100 ppm solution of Camphor acquired with three different tunings is shown in Figure 2. Small differences in ion abundances are found throughout this profile, similar results were found for the 10 ppm solution, not shown. BFB tuning is designed to provide higher intensities to low-mass ions; this can be seen in Figure 2A, in which the BFB profile appears to have higher intensity than either atune or stune. Since the raw abundance of all ions appears to be higher we investigated the effect of normalization on this data set. After normalizing the dat set to 100% the samples are placed in the same scale. Figure 2B shows that after the vector length normalization, variation due to the tuning used was partially removed from the data set.

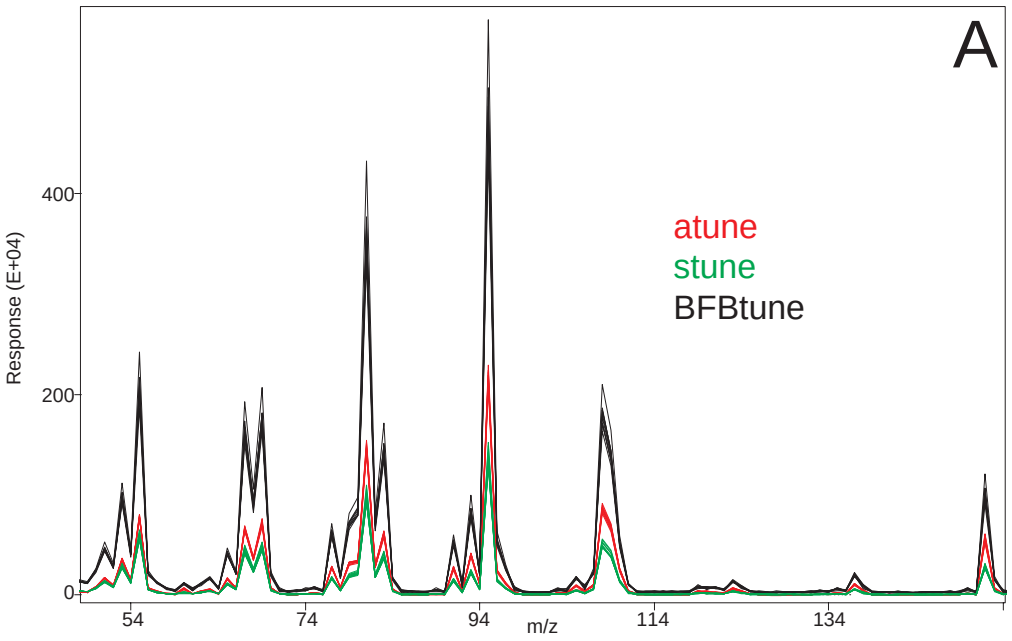

Figure 2A. Mass spectral fingerprints for 100 ppm Camphor solution acquired with three different tuning algorithms.

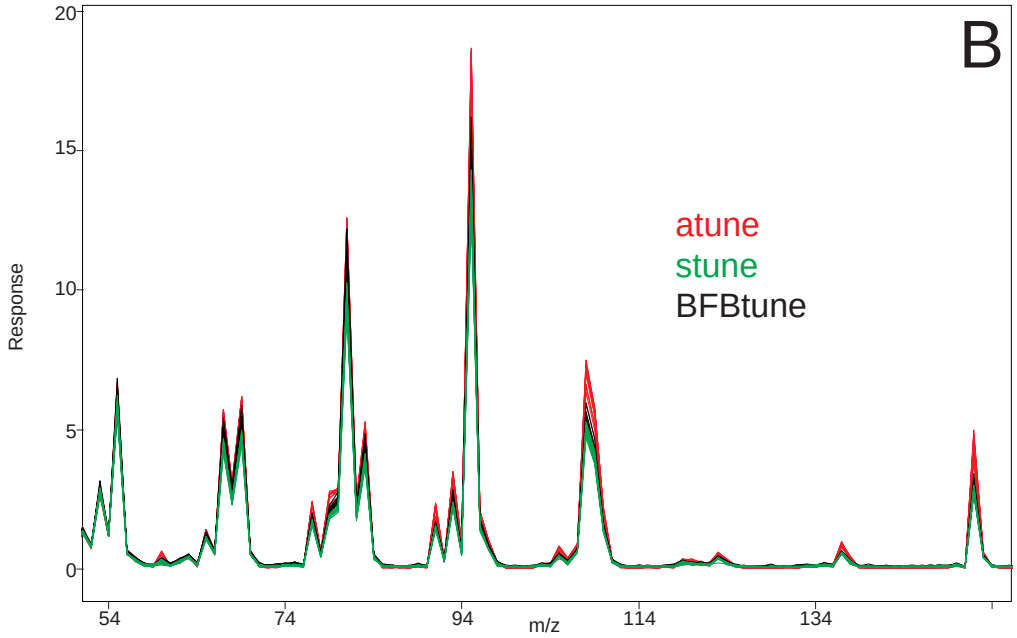

Figure 2B. Normalized mass spectral fingerprints for 100 ppm Camphor solution acquired with three different tuning algorithms.

In order to test if the normalization of the data set was sufficient to provide reliable predictions, chemometric models were created using only the data acquired with a specific tuning. Three classification models were tested, KNN, SIMCA and SIMCA with TOC. Table 1 lists the correct predictions obtained with each of the classification models.

Table 1. Correct prediction of unknowns for the camphor data set.

Tuning	atune			BFBtune			stune		
	KNN Norm	SIMCA	SIMCA* TOC	KNN Norm	SIMCA	SIMCA* TOC	KNN Norm	SIMCA	SIMCA* TOC
atune	27 of 27	19 of 27	25 of 27	27 of 27	0 of 27	26 of 27	27 of 27	0 of 27	27 of 27
BFB	28 of 28	0 of 28	26 of 28	28 of 28	28 of 28	27 of 28	28 of 28	6 of 28	28 of 28
stune	28 of 28	0 of 28	26 of 28	28 of 28	3 of 28	28 of 28	28 of 28	27 of 28	27 of 28
% total	100%	23%	93%	100%	37%	98%	100%	40%	99%

*Using direct transfer of calibration (2 transfer samples)

Using KNN with normalization all samples predicted accurately. Using SIMCA with normalization only few samples acquired with different tuning are classified correctly. For example, a SIMCA model created only with the atune data is shown in Figure 3A. Projections of the samples acquired with BFB and stune project outside the boundaries in this figure. A SIMCA model with direct TOC was created in which 2 calibration transfer samples were used for each level. When TOC is used over 90% correct classifications are obtained regardless of the tuning used for data acquisition. This can also be observed in Figure 3B, in which BFB and stune samples are projected within the boundaries of the atune model when using TOC.

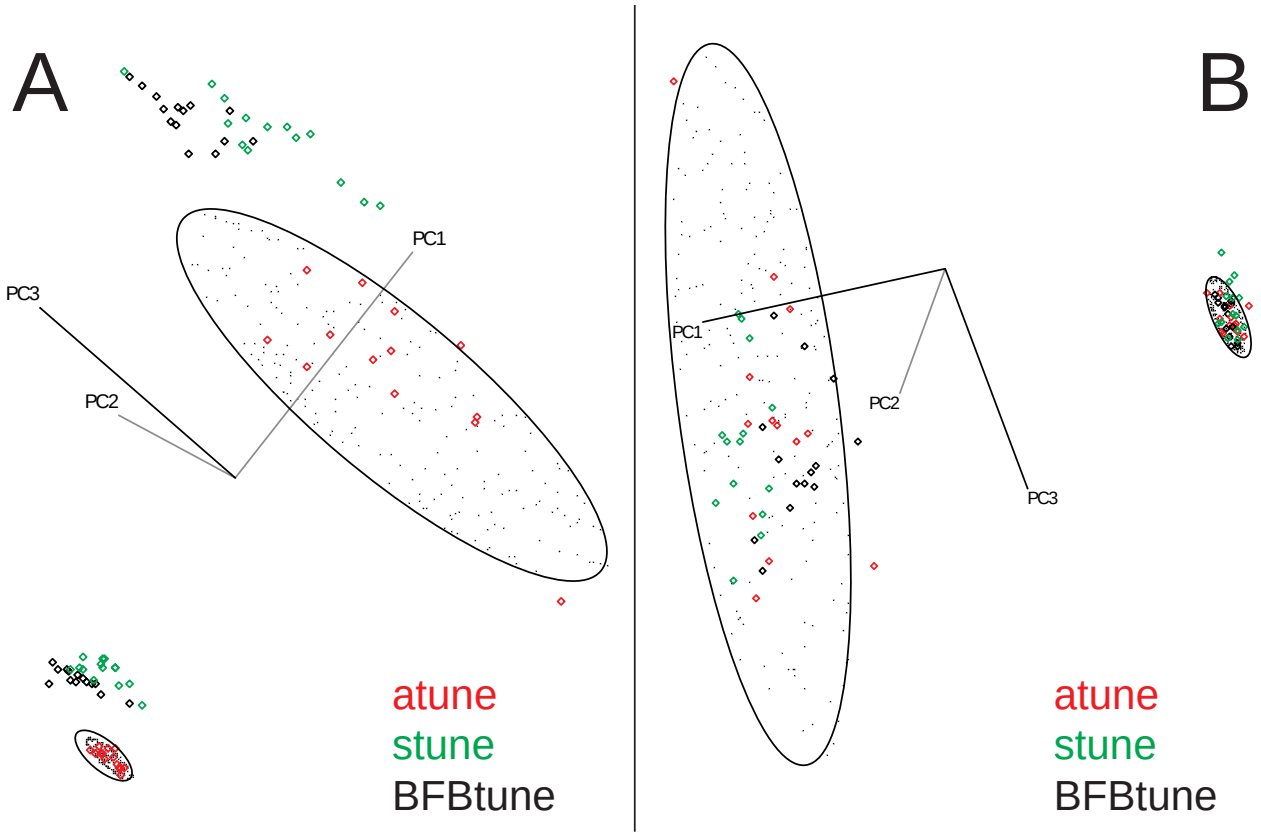

Figure 3. Mass spectral fingerprint projections into the space of the first three principal components. SIMCA-model created with data acquired with atune; A) original data, B) using two calibration transfer samples for each level and each tuning.

B. Filament replacement Data set 2 - Citrus oils. Figure 4A shows the mass spectral fingerprint for citrus oil #1 acquired with two different filaments. Slightly higher abundances were obtained with filament #1 (red trace) than with filament #2 (blue trace). Similar results were found with citrus oils #2 and #3, not shown. Since visual inspection of the oils fingerprints indicate that the overall intensity decrease throughout the spectrum, we decided to normalize the data set.

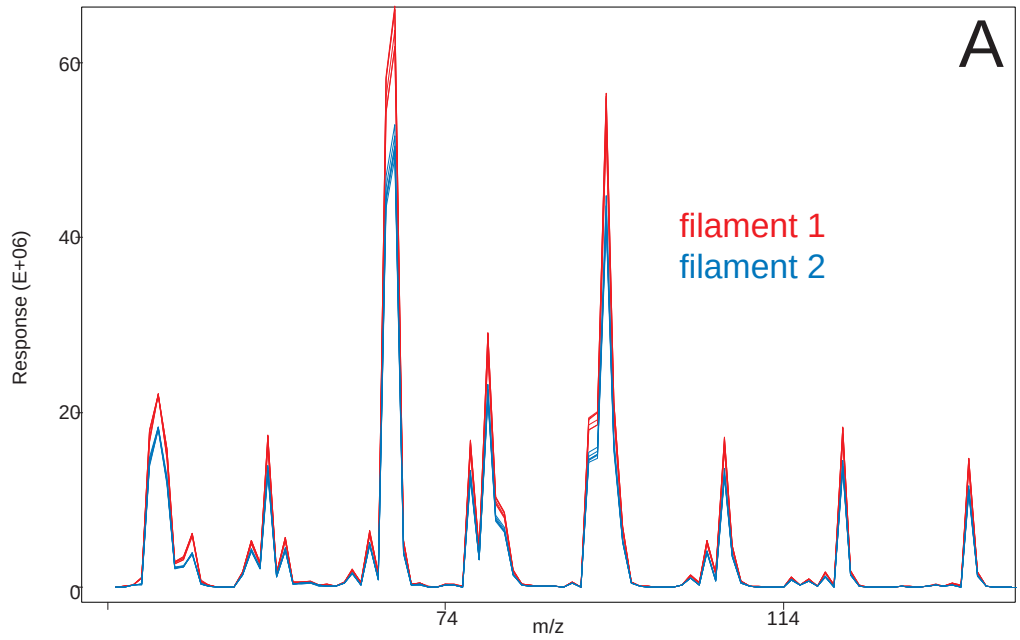

Figure 4A. Mass spectral fingerprints for citrus oil #1 acquired using two different filaments.

Figure 4B shows the normalized citrus oil #1. After normalization only smaller differences in the mass spectral fingerprints are observed.

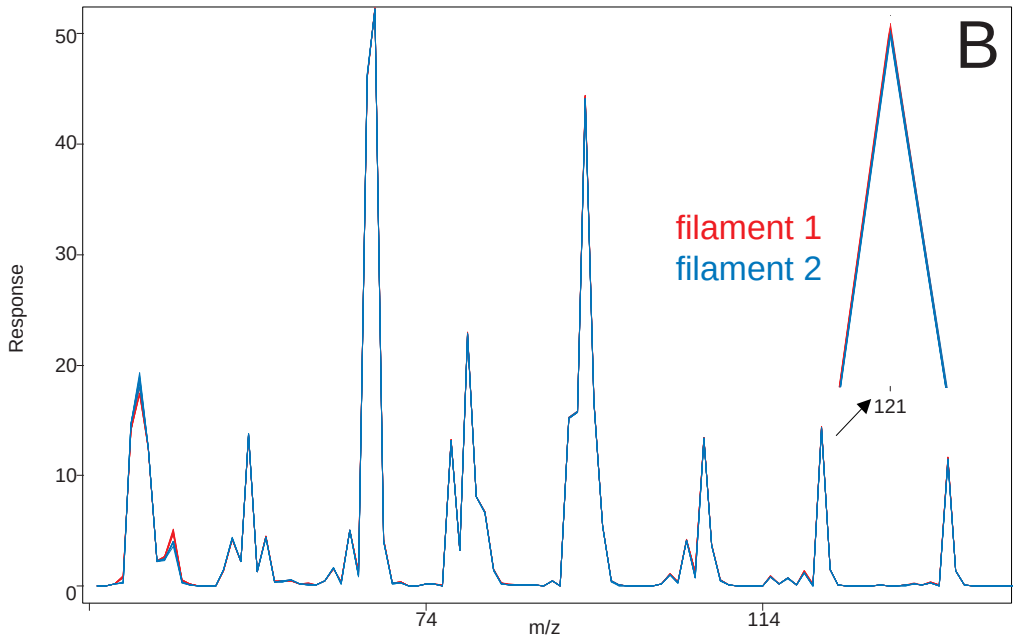

Figure 4B. Normalized mass spectral fingerprints for citrus oil #1 acquired using two different filaments.

Predictions using a KNN model with 5 neighbors were accurate regardless of which filament was used to acquire the fingerprint. 100% accuracy was obtained with the KNN model. SIMCA predictions improved with the use of TOC, as seen in Figure 5.

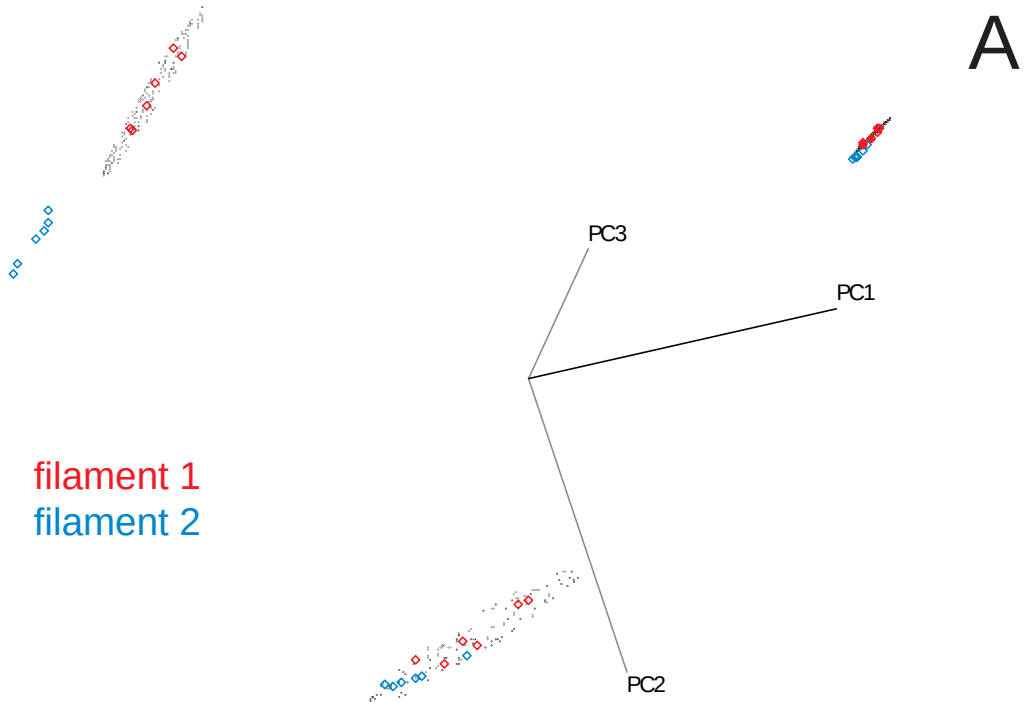

Figure 5A. SIMCA model created with original data.

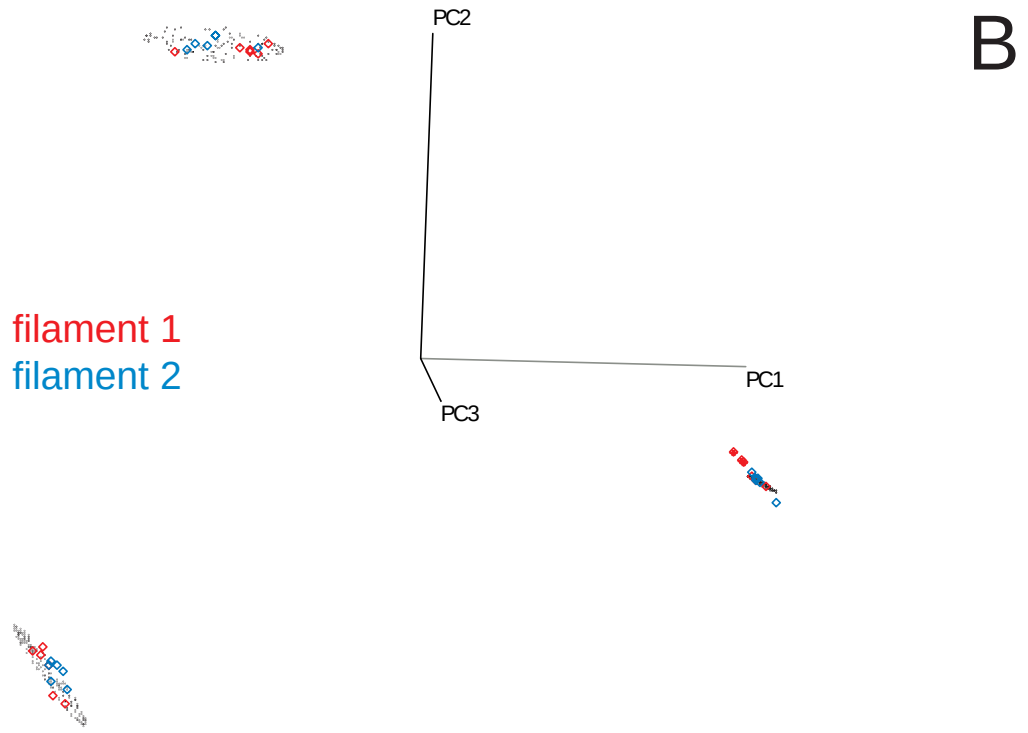

Figure 5B. SIMCA model created with original data (filament 1), using 2 samples per oil for calibration transfer (filament 2).

Prediction accuracy increased slightly when 3 transfer samples were used instead of two. Table 2 summarizes these findings.

Table 2. Correct prediction of unknowns for the citrus oils data set.

	KNN	SIMCA	SIMCA*	SIMCA**
oil 1-Film1	6 of 6	6 of 6	3 of 6	3 of 6
oil 1-Film2	6 of 6	6 of 6	4 of 6	4 of 6
oil 2-Film1	6 of 6	6 of 6	6 of 6	6 of 6
oil 2-Film2	6 of 6	0 of 6	4 of 6	3 of 6
oil 3-Film1	6 of 6	0 of 6	4 of 6	6 of 6
oil 3-Film2	6 of 6	0 of 6	6 of 6	6 of 6
Total	100%	50%	75%	78%

*Using direct transfer of calibration (2 transfer samples)
**Using direct transfer of calibration (3 transfer samples)

C. Ion source cleaning Data set 3 - Limonene. Figure 6 shows the mass spectral profiles for the 100 µg sample of Limonene in Methanol before source clean (BSC) and after source clean (ASC).

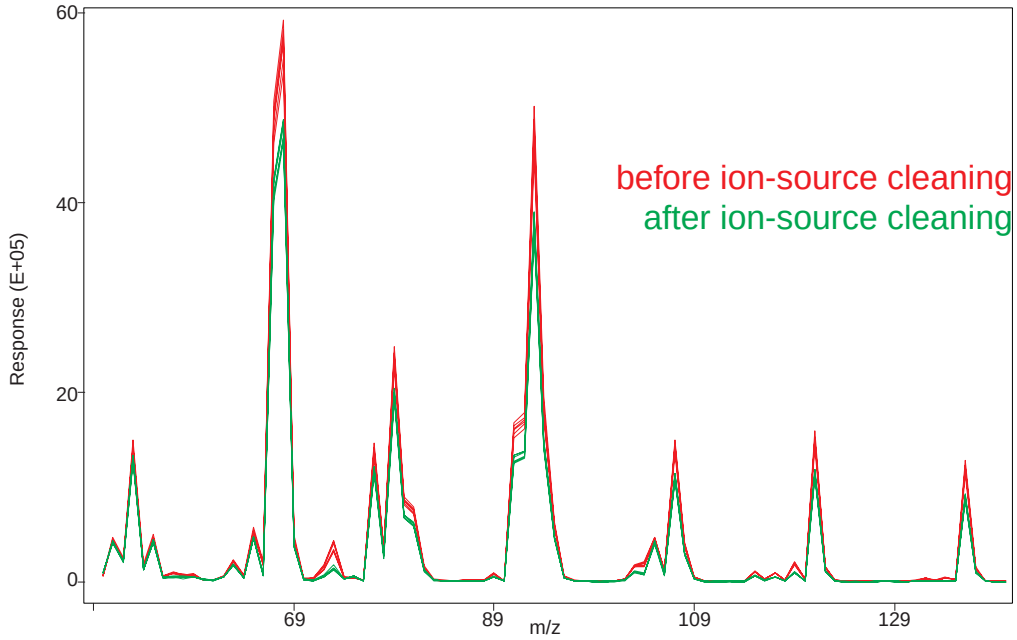

Figure 6. Mass spectral fingerprint for 100 µg of Limonene in Methanol.

A slight difference in the raw abundance of the majority of ions is also present in this data set. Similar results were found with the 10 and 50 µg samples, not shown. Exploratory analysis using hierarchical cluster analysis (HCA) indicates that even though the samples were measured with different ion source conditions they still cluster according to their level (Figure 7).

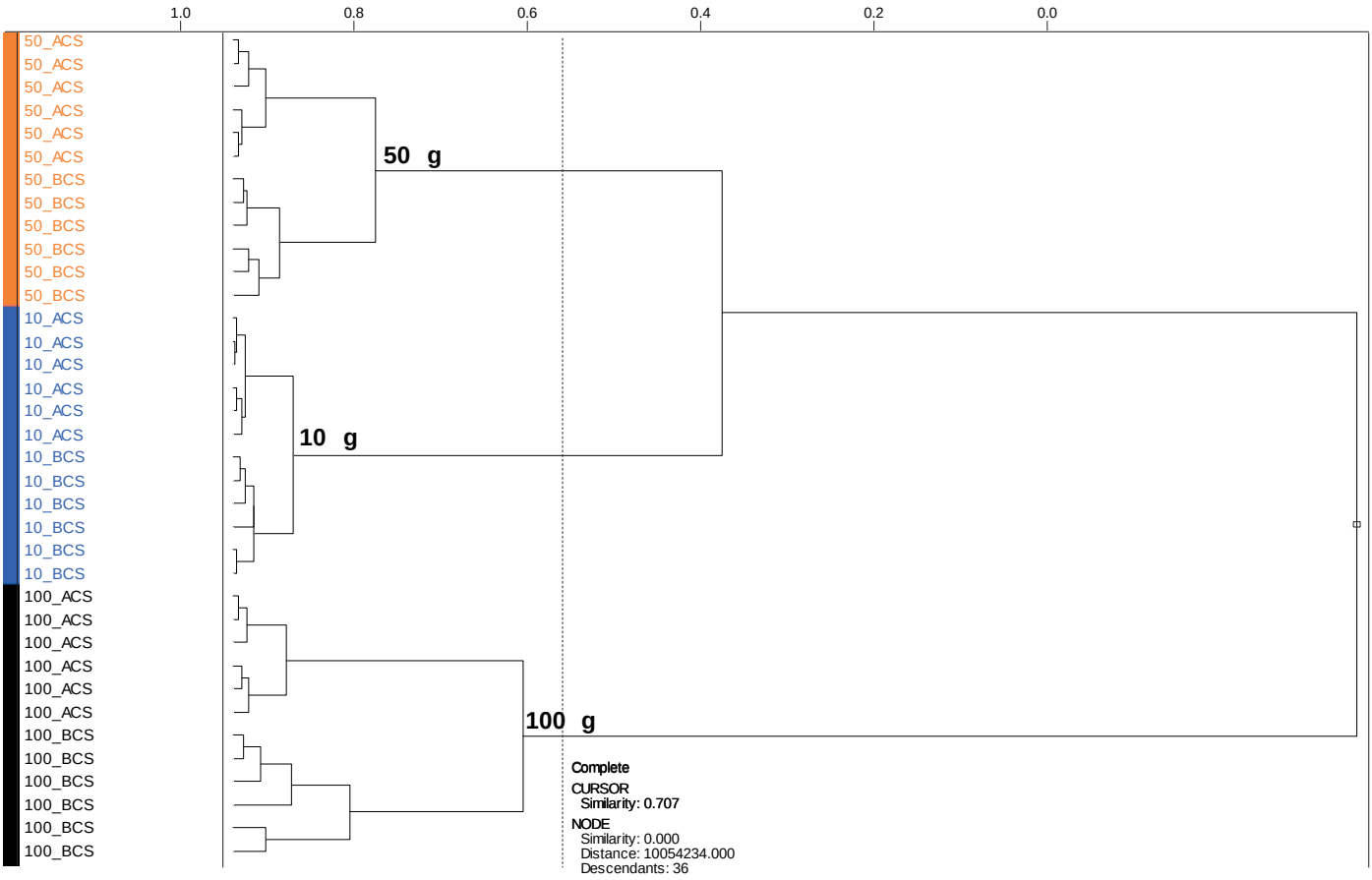

Figure 7. HCA model for three Limonene solutions. HCA obtained using Euclidean distance and complete linkage.

Inspection of this figure also indicates that the experimental variation appears to be lower after the ion source has been cleaned. This is evident by the inspection of the two clusters within each level; samples acquired before ion source clean are more spread out in their clusters.

Four chemometric models were created with the samples acquired before the ion source clean; KNN and SIMCA-TOC and PLS with and without TOC. For the TOC predictions two calibration transfer samples were used per level. Using the KNN model all samples were predicted in the correct class, 10, 50 or 100 µg. SIMCA-TOC predicted 89% of the samples correctly.

Since we wanted to test the quantitative predictions using the PLS models, we only mean centered the data sets. As an overall measure of the accuracy of the models we decided to use the standard error of prediction (SEP) [2]. SEP is related to the PRESS (Prediction Residual Error Sums of Squares) but SEP takes into account the number of samples and has the same units as the dependent variable. Table 3 contains the predictions results using PLS and PLS-TOC.

Table 3. Prediction of unknowns for the limonene data set using a 5-factor PLS model.

BCS Amount [µg]	PLS Amount [µg] @5	PLS_TOC Amount [µg] @5	ACS Amount [µg]	PLS Amount [µg] @5	PLS_TOC Amount [µg] @5
10	10	9	100	104	100
50	50	47	50	50	50
100	100	108	100	105	101
10	10	10	50	49	49
50	50	50	10	8	12
100	100	100	100	103	101
10	10	11	50	49	49
50	49	50	10	8	11
100	101	108	10	8	11
10	11	11	100	99	100
50	50	50	10	8	11
100	101	101	100	101	99
10	11	11	100	102	98
50	50	52	50	49	48
100	100	106	10	8	11
10	9	12	10	8	11
50	49	51	50	48	47
100	100	106	50	47	49
			SEP	38	37
			PRESS	51877	49954

Overall, slight different predictions and lower SEP were obtained using the PLS model with TOC. Predictions using PLS-TOC are also shown in Figure 8.

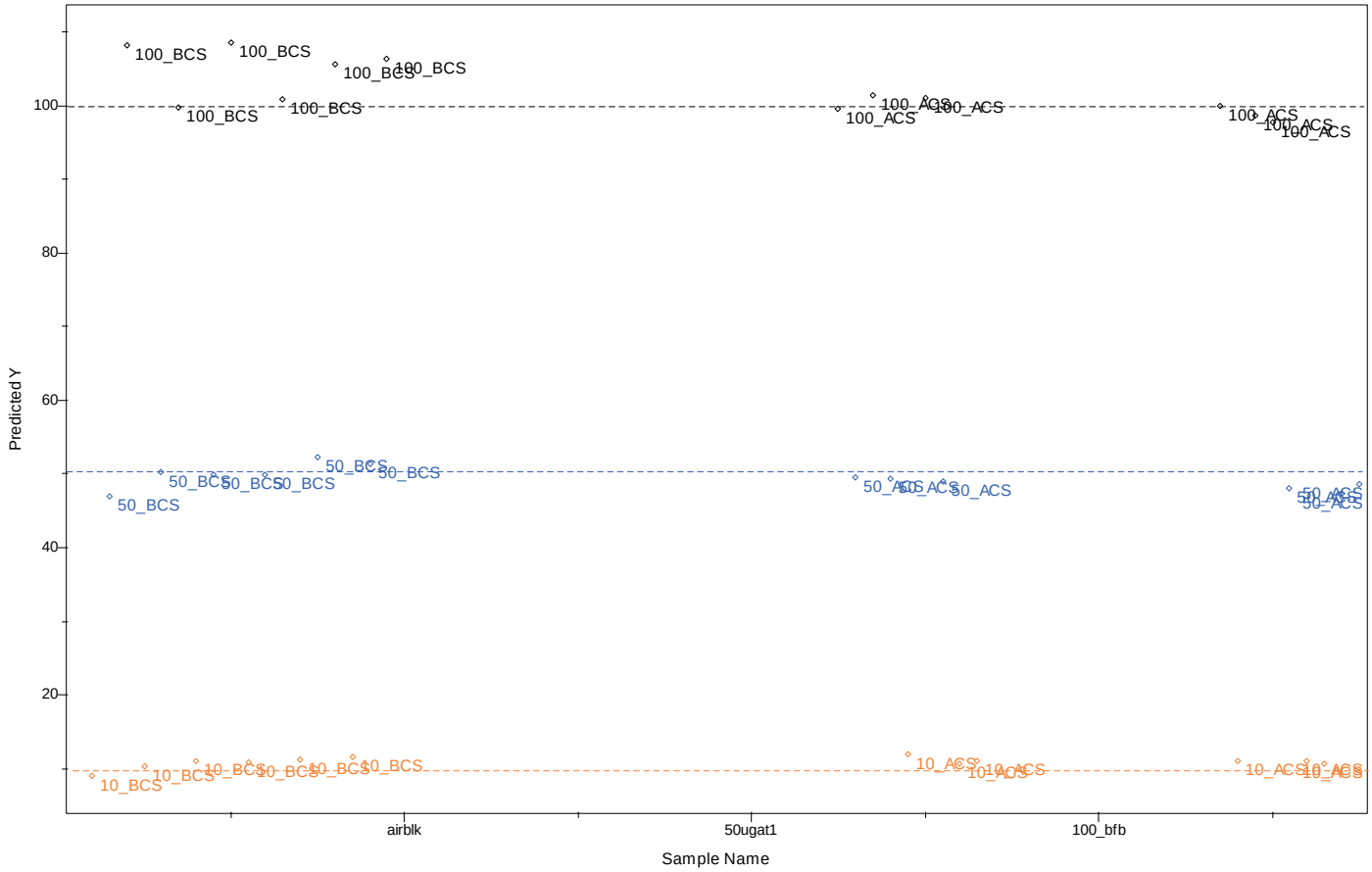

Figure 8. PLS predictions for Limonene solutions before source cleaning-samples on the right side of the figure and after source cleaning-left side samples. TOC was used with 2 calibration transfer samples per level.

Conclusions

Slight differences in mass spectral profiles were found depending on the tuning algorithm, filament used and ion source condition. In all the three data sets studied, KNN models provided good accurate predictions regardless of slight changes in instrument conditions.

The use of direct TOC improved all the classifications using SIMCA models. In this study, two calibration transfer samples per class or level appear to greatly improve the correct classifications. Our goal is to continue monitoring the above samples over longer periods of time and test the ability of TOC algorithms to compensate for slight changes in the mass spectral fingerprints.

References

[1] ChemStation enhanced data analysis software Help menu, Tuning the MSD. G1701CA, vC.00.00 (Agilent Technologies).

[2] Pirouette chemometrics software Help menu. V3.11(Infometrix, Woodinville, WA).

GERSTEL GmbH & Co. KG
Eberhard-Gerstel-Platz 1
45473 Mülheim an der Ruhr
Germany
+49 (0) 208 - 7 65 03-0
+49 (0) 208 - 7 65 03 33
gerstel@gerstel.com
www.gerstel.com

GERSTEL Worldwide

GERSTEL, Inc.
701 Digital Drive, Suite J
Linthicum, MD 21090
USA
+1 (410) 247 5885
+1 (410) 247 5887
sales@gerstelus.com
www.gerstelus.com

GERSTEL AG
Wassergrabe 27
CH-6210 Sursee
Switzerland
+41 (41) 9 21 97 23
+41 (41) 9 21 97 25
swiss@ch.gerstel.com
www.gerstel.ch

GERSTEL K.K.
1-3-1 Nakane, Meguro-ku
Tokyo 152-0031
SMBC Toritsudai Ekimae Bldg 4F
Japan
+81 3 5731 5321
+81 3 5731 5322
info@gerstel.co.jp
www.gerstel.co.jp

GERSTEL LLP
10 Science Park Road
#02-18 The Alpha
Singapore 117684
+65 6779 0933
+65 6779 0938
SEA@gerstel.com
www.gerstel.com

GERSTEL (Shanghai) Co. Ltd
Room 206, 2F, Bldg.56
No.1000, Jinhai Road,
Pudong District
Shanghai 201206
+86 21 50 93 30 57
china@gerstel.com
www.gerstel.cn

GERSTEL Brasil
Av. Pascoal da Rocha Falcão, 367
04785-000 São Paulo - SP Brasil
+55 (11)5665-8931
+55 (11)5666-9084
gerstel-brasil@gerstel.com
www.gerstel.com.br

Awarded for the active pursuit of environmental sustainability