Optimization of electronic nose drift correction applied to tomato volatile profiling

Journal:	*Analytical and Bioanalytical Chemistry*
Manuscript ID	ABC-00278-2021.R1
Type of Paper:	Research Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Valcárcel, Mercedes; Universitat Politècnica de València, COMAV Ibáñez, Ginés; Universitat Jaume I, Agricultural Sciences and Natural Environment Martí, Raúl; Universitat Politecnica de Valencia, COMAV Beltran, Joaquim; University Jaume I, Research Institut for Pesticides and Water Cebolla-Cornejo, Jaime; Universitat Politècnica de València, COMAV Roselló, Salvador; Universitat Jaume I, Agricultural Sciences and Natural Environment
Keywords:	electronic nose, drift correction, chemometrics, sequence standardization, tomato

Optimization of electronic nose drift correction applied to tomato volatile profiling

Mercedes VALCÁRCEL[a], Ginés IBÁÑEZ[b], Raúl MARTÍ[a], Joaquín BELTRÁN[c], Jaime CEBOLLA-CORNEJO[a], and Salvador ROSELLÓ[b]*

[a]*Joint Research Unit UJI-UPV - Improvement of agri-food quality. COMAV. Universitat Politècnica de València, Cno. de Vera s/n, 46022 València, Spain*

[b]*Joint Research Unit UJI-UPV - Improvement of agri-food quality. Agricultural Sciences and Natural Environment Department, Universitat Jaume I, Avda. Sos Baynat s/n, 12071 Castelló de la Plana, Spain*

[c]*Reserach Institute for Pesticides and Water (IUPA) , Universitat Jaume I, Avda. Sos Baynat s/n, 12071 Castelló de la Plana, Spain*

**Corresponding author: E-mail: rosello@uji.es. Postal address: Universitat Jaume I, ESTCE, Avda. Sos Baynat s/n 12071 Castellón de la Plana.*

M. Valcárcel ORCID no.: 0000-0002-9347-1500
G. Ibañez ORCID no.: 0000-0002-1787-8587
R. Martí ORCID no.: 0000-0002-3517-1948
J. Beltran ORCID no: 0000-0003-4387-101X
J. Cebolla-Cornejo ORCID no.: 0000-0002-2607-9920
S. Roselló ORCID no.: 0000-0002-7733-4178

Funding

This research was partially funded by Jaume I University with projects P1-1B2011-41 and COGRUP/2016/04. G. Ibáñez thanks Universitat Jaume I for funding his pre-doctoral grant (PREDOC/2015/45).

Abstract

E-noses can be routinely used to evaluate the volatile profile of tomato samples once the sensor drift and standardization issues are adequately solved. Short-term drift can be corrected using a strategy based on a multiplicative drift correction procedure coupled with a PLS adaptation of the Component Correction. It must be performed specifically for each sequence, using all sequence signals data. With this procedure, a drastic reduction of sensor signal %RSD can be obtained, ranging between 91.5% and 99.7%for long sequences and 75.7% and 98.8% for short sequences. On the other hand, long-term drift can be fixed up using a synthetic reference standard mix (with a representation of main aroma volatiles of the species) to be included in each sequence that would enable sequence standardization. With this integral strategy, a high number of samples can be analyzed in different sequences, with a 94.4% success in the aggrupation of the same materials in PLS-DA two-dimensional graphical representations. Using this graphical interface e-noses can be used to developed expandable maps of volatile profile similitudes, which will be useful to select the materials that most resemble breeding objectives or to analyze which preharvest and postharvest procedures have a lower impact on the volatile profile, avoiding the costs and sample limitations of gas chromatography.

Keywords: electronic nose, drift correction, chemometrics, sequence standardization, tomato.

Introduction

The objective evaluation of flavor in crops such as tomato is expensive and time-consuming, consequently, this trait has been usually disregarded. Today it is known that one of the main factors under the loss of flavor relies on the loss of alleles related to the contents of aroma volatiles [1], and the use of delayed ripening genes that alter the aroma profile, an effect that depends on the genetic background [2]. Additionally, tomato flavor can also be altered by the preharvest and postharvest management of the crop that also alter the production of volatiles [3 – 6], .

In order to satisfy the demands of high quality markets, it would necessary to include flavor evaluation, and especially the volatile profile, during the development of breeding programs [7], cultivation, and postharvest procedures. In this context, the use of trained panelists or the precise volatile quantifications by gas chromatography-mass spectrometry is discarded considering that these evaluations are too expensive and time-consuming and, consequently, not adequate to evaluate a high number of samples.

As an alternative, electronic noses (e-noses) were designed to evaluate the volatile profiles of agricultural products [8]. For this purpose, they have been usually applied to classify materials considering their quality characteristics, their origin, the variety or the presence of diseases, additives, adulterations, and off-flavors in different fruits and vegetables (tomato, kiwifruits, peach, nectarine, apple, banana, persimmon, grape, watermelon, strawberry, blackberry, onion, potato, pumpkin, broccoli, etc.), grains (wheat, rice, maize, peanuts, etc.), aromatic and medicinal plants (tea, coffee, saffron, cocoa, oregano, ginseng, etc.), processed products (oils, juices), livestock and poultry meat, and fish [8–12]. Most of these applications were modeled and tested in a short-term scenario, using a limited number of samples. However, the application of this technology to the evaluation of materials in breeding programs and food industry makes it is necessary to assure the capability to process a high number of samples in the same day, as well as being able to compare them with data obtained in previous assays. By doing so, it would be possible to apply e-noses to selection and quality control programs, in which each new sample is compared with reference

values or fingerprints obtained in previous assays with elite materials grown and handled in ideal conditions. From this point of view, the objective would not be centered on classifying a new sample, but to have an idea of its distance to elite reference samples. Consequently, it would be possible to select the best individuals or those preharvest or postharvest procedures that minimize their impact on the volatile profile.

In order to take advantage of the capabilities of e-noses, it would be necessary to overcome the effects of sensor drift. This phenomenon is defined as temporary or gradual changes in one or some sensor properties which causes distorted response measures and reduces the validity of the electronic fingerprints. It is inevitable and caused by complex and dynamic processes, such as changes in room environmental conditions (temperature or humidity), changes in the composition of measured samples (component interactions), instrument operational disturbances (sensors thermal and memory effects, aging or poisoning) [13, 14]. These changes can be noticed both, in signals within a work sequence (short-term drift) and signals obtained in different work sequences (long-term drift). The improvement of sensor technology at the manufacturing stage to enhance its stability over time has contributed to reduce these problems. However, despite the advances obtained, a regular calibration is still required to limit the effects of sensor drift. It can be performed using external standards and statistical multivariate calibration models. Nonetheless, multivariate calibration requires a large number of samples and frequent re-calibrations of the sensor arrays and this would limit the number of new samples analyzed. Therefore, a new model calibration transfer or update and signal standardization using only a small number of reference samples would represent an interesting solution to keep the system operative for long periods [14].

In the last two decades, an enormous research effort has been made on different methodologies aimed to properly process signals and data from e-noses (reviewed by [13–15]). Nevertheless, it seems clear that, despite the high amount of research on drift correction and calibration update methods developed, these proposals were not routinely used, except for component correction and directed standardization methods. The best solutions proposed up to now rely on the analysis of a

high number of samples to develop robust models or use simple volatile mixes. These approaches are distant from the real context of tomato evaluation. This species has a complex volatile profile with more than 400 compounds, with nearly 30 of them playing an important role in tomato aroma perception [2]. On the other hand, the need to develop models with a high number of samples would not be realistic in high-throughput evaluations, as the models would have to be recalculated each time a sensor has to be changed.

In this context, although the use of commercial electronic noses for the evaluation of volatile profiles has a huge potential, it is necessary to develop an operating methodology enabling the routine evaluation of wide collections of real samples. This is, in fact, the aim of this paper, to propose a practical methodology to correct drift within and between sequences, using a minimum number of samples to calibrate the models and a tomato-like complex synthetic reference mix to standardize sequences. Finally, the development of long-term expandable partial least squares discriminant analysis (PLS-DA) graphical maps of e-nose volatile profiles is proposed as a valuable tool to enable the routine evaluation of the volatile profile of new samples, analyzing the relative distance to reference points.

Materials and Methods

Plant material and tomato-like synthetic standards

Tomato-like synthetic standards were developed to obtain a synthetic mixture of main volatile compounds of an average real tomato sample, but with higher stability and reproducibility. For this purpose, a high concentration standard mixture was prepared (TomSSt_4), containing 30 individual volatile compounds at concentrations (Table 1) corresponding to the mean values of representative tomato cultivars with different aromatic profiles [16]. Three alternative standards were obtained diluting TomSSt_4 to 70% (TomSSt_3), 50% (TomSSt_2) and 30% (TomSSt_1). The dilutions were obtained to cover a wide range of volatile sample concentrations. TomSSt_2 was employed

as a reference sample for inter-sequence standardization in long-term drift correction. These working solutions where prepared by volume dilution from more concentrated stock solutions which were stored in the freezer at -30ºC in sealed vials. They have an established stability of one year for the main stock solutions (around 500 ppm or higher) and of 1 month for the ppb to sub ppm solutions. As preparation of synthetic standards is carried out by dilution in volume of stock standards, this process can be reliably and reproducibly performed producing adequate standard solutions in the routine laboratory. For sequences run in different months, the specific standard mixtures were prepared de novo to provide restrictive conditions.

Table 1

Composition of the tomato-like synthetic standard TomSSt_4.

Volatile compound	ng mL^{-1}	Volatile compound	ng mL^{-1}
E-2-hexen-1-ol acetate	0.70	eugenol	13.92
3-methyl thiopropanal	1.12	nonanal	11.12
terpineol (alpha+beta+gamma)	0.56	2-isobutylthiazole	26.40
E-2-hexen-1-ol	1.10	E-2-heptenal	24.96
1-hexanol	2.02	methyl salicylate	892.00
3-carene	2.11	guaiacol	480.00
3-methylbutyl acetate	2.04	E-2-hexenal	702.00
alpha-pinene	1.98	6-methyl-5-hepten-2-one	590.00
gamma-terpinene	2.08	hexanal	800.00
2-carene	7.20	Z-3-hexenal	824.00
linalool	6.60	E-2-octenal	102.00
phenylacetaldehyde	9.20	citral (Z+E)	170.40
2-phenylethanol	12.04	R-limonene	98.00
6-methyl-5-hepten-2-ol	13.64	Z-3-hexen-1-ol	216.80
beta-ionone	13.16	geranyl acetone	114.80

Tomato varieties evaluated in this work represented a wide diversity of fruit shapes, colors, genotypic structures (commercial hybrids and landraces), and origins (Table 2). The plant material included four commercial hybrids, “Zayno RZ”, “Divyne RZ”, “Vinchy RZ” (Rijk Zwaan Iberica, Almería, Spain), and “Caramba” (De Ruiter Seeds, Almería, Spain). Four experimental tomato breeding lines (UJI008, UJI011, UJI014, and UJI028) with different fruit sizes. One cherry tomato type accession (BGV004587). Five accessions of local landraces, UJI023 of “de penjar” landrace, BGV005477 accession of a “Morado” landrace, BGV005651 an accession of “Muchamiel” landrace, BGV005718 an accession of “Amarillo” landrace, and BGV005655 an accession

belonging to the "Valenciano". The "de penjar" landrace carries with *alcobaça*, *alç*, long-life mutation allelic to the *nor* gene [17] and it results in a very specific aroma volatile evolution [18], "Morado" landrace has external pink color due to the transparent peel typical of the *yellow*, *y,* mutation which alters the synthesis of polyphenols and "Amarillo" has yellow flesh color typical of the impairment of carotenoid synthesis resulting from the presence of the *yellow-flesh*, *r*, mutation (reviewed by [19]) and it, therefore, affects the synthesis of apocarotenoid volatiles.

UJI accessions were obtained from Universitat Jaume I and BGV accessions from the genebank of the Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV).

Table 2

Description of the tomato accessions tested in the different assays performed.

Code	Type of material	Accession	Number of sequences			Fruit characteristics
			1st assay	2nd assay	3rd assay	
1	Commercial hybrid	"Zayno RZ"[a,z]	3	1	3	Large, rounded, green-red
2	"Amarillo" landrace	BGV005718[b,x]	3	1	3	Large, slightly flattened, yellow
3	Commercial hybrid	"Caramba"[a,y]	1	1	1	Large, flattened, green-red
4	Breeding line	UJI011[c,u]	1	1	1	Large, rounded, red
5	Commercial hybrid	"Divyne RZ"[a,z]	1		1	Medium-large, rounded, red
6	Commercial hybrid	"Vinchy RZ"[a,z]	1		1	Large, rounded, red, long life
7	"De penjar" landrace	UJI023[b,u]	1	1	1	Small, rounded, red, long life
8	"Morado" landrace	BGV005477[b,x]	1	1	1	Large, slightly flattened, pink
9	"Muchamiel" landrace	BGV005651[b,x]	1	1	1	Large, flattened, red-orange,
10	"Valenciano" landrace	BGV005655[b,x]	1		1	Medium-large, heart-shaped, red-orange
11	Cherry tomato	BGV004587[b,x]	1		1	Small, rounded, orange-brownish
12	Breeding line	UJI008[c,u]	1		1	Small, rounded, red
13	Breeding line	UJI014[c,u]	1		1	Medium-large, slightly flattened, red
14	Breeding line	UJI028[c,u]	1		1	Small, rounded, red
TomSSt1	Tomato like standard (30%)		3		3	
TomSSt2	Tomato like standard (50%)		3	1	3	
TomSSt3	Tomato like standard (70%)		3		3	
TomSSt4	Tomato like standard (100%)		3		3	

Tomato types: [a]commercial hybrid, [b]local landraces, [c]breeding lines
Origin: [z]Rijk Zwaan Iberica S.A., [y]De Ruiter Seeds S.A., [x]Instituto Universitario de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV) seed bank, [u]Univesitat Jaume I seed collection.

Experimental design

Three different assays were performed. In the first assay, 18 samples with different compositions were used. These samples included real tomato samples from 14 varieties obtained homogenizing whole fruits (Table 2) and the four tomato-like synthetic standards (TomSSt) with variable volatile composition. Three sequences were run on different days. Each sequence included four specific varieties (that were included only in one sequence) and two varieties that were included as controls in the three sequences. The 4 tomato-like synthetic standards were also included in all the sequences. Tomato samples were replicated 7 times and tomato-like synthetic standards 4 times in each sequence. All the samples were randomly distributed in each working sequence.

For a deeper study of the short-term drift, a second assay was designed to include a higher number of repetitions (12) per sample. Two consecutive long work sequences (22 hours each) were planned to test seven tomato and one tomato-like synthetic standard (TomSSt_2). All the samples were also randomly distributed within the first replicate of each sequence, and the order was maintained in the rest of the replicates. This design provided data to compare the performance in a whole sequence (12 repetitions/sample in 22 hours) or a short sequence (4 repetitions/sample in 8 hours approximately) to test the performance of the drift correction strategy proposed in different scenarios.

Finally, a third assay was performed to analyze the effect of long-term drift. To ensure the inclusion of long-term drift in the signal responses, the sequences of this trial were carried out in a 3 months period (one sequence per month) included in the normal routine usage of the equipment. During this period other samples from tomato and other vegetable crops were analyzed in the equipment. The short-term drift correction was applied before analyzing the results.

In a first step, the effect of long-term drift was analyzed using the four tomato-like standard solutions in three sequences. Then the effect of long-term drift was also checked adding two tomato varieties analyzed in three sequences. Long-term drift correction via sequence standardization was then applied and its validity checked.

The independent study of each one of these three sequences was used to test and correct short-term drift within a work-day sequence. The joint data of all these sequences were used to test the

performance of the long-term drift correction between sequences and standardization strategies proposed in this work.

Once the reliability of the long term-drift correction had been checked it was applied to analyze the data obtained from the analysis of the 14 tomato varieties distributed in three sequences, using the data from TomSST2 for sequence standardization.

Electronic nose and data acquisition

A FOX 4000 (Alpha MOS, Toulouse, France) e-nose system was used. The system included 18 metal oxide semiconductor sensors (MOS) installed in three chambers, an autosampler system (CombiPAL HS100, CTC Analytics, Zwingen, Switzerland), and a software package (AlphaSoft v11) to control and process initial data. The sensor response in MOS sensors is a resistance variation due to a reaction caused by the chemical species on the surface of the active layer of the sensor. As usual for MOS sensors, the signal was expressed as normalized resistance variation of the signal highest point $((R_i - R_{max})/R_i)$, where R_i is resistance at time zero and R_{max} is resistance in the signal highest point of the sensor [20].

The analysis parameters related to general aspects of equipment operation were fixed following manufacturer recommendations, while those that directly determine the response quality (influence headspace generation) were established from previous tests based on the methodology for the analysis of tomato aroma developed by the group [16]. For each sample, 2 g of homogenate (2 mL in the case of tomato synthetic standards) were introduced into a 10 mL vial and sealed. Each sample replicate corresponded to an independent vial. Samples were incubated in the autosampler at 45ºC for 10 minutes to generate the headspace and then 2 mL of it were injected into the sensors chambers for analysis. The sensors' response was recorded over two minutes with 18 minutes between each measurement to allow the baseline recovery. Between samples, dry clean synthetic air flowed over the sensor array for 2 minutes to remove residues of the previous sample, following manufacturer recommendations. The gas flow rate was 150 mL min^{-1}. Instrument maintenance (daily auto test

and two-week diagnosis) were routinely performed following supplier protocols to ensure proper operation.

Drift correction and inter-sequence standardization

A multivariate adaptation of the multiplicative drift correction procedure proposed by Salit and Turk [21], combined with a partial least squares (PLS) adaptation of the component correction strategy [22] to model time-dependent drift was used both to remove intra-sequence short-term drift and to perform inter-sequence standardization to counteract long-term drift. Salit and Turk method is based on an interpolative projection of sample signal onto a smooth function defined by fitting to signals from regularly interspersed standards. Component correction strategy is based on the assumption that there is a subspace direction that captures only the drift variance and can be modelled (they use Principal Component analysis) and substracted from the measurement matrix X to provide drift corrected signals. Two assumptions were considered: i) drift, regardless of its type, is a function of time, and ii) drift for our electronic nose instrument is multiplicative (i.e. the magnitude of the perturbations is dependent on the signal level). Additionally, it had to be considered that the nature of the samples being analyzed could not be contemplated by the model, as they were unpredictable.

A practical guide of our proposed intra-sequence drift correction methodology is included in Supp. Fig. 1. According to [21], when multiplicative drift appears, the signal measured in a sample i evaluated with j repetitions in each of the k sensors of the system ($S_{i_{(j)},\,k\,\text{measured}}$) could be decomposed as:

$$S_{i_{(j)},\,k\,\text{measured}} = S_{i,k\,\text{truth}}(1 + E_{\text{drift}}(t) + E_{\text{noise}}) \qquad (1)$$

Being $S_{i\,\text{truth}}$ the true signal for sample i, $E_{\text{drift}}(t)$ the drift estimation as a function of time and E_{noise} the estimation of the background noise (independent of time). $S_{i\,truth}$ can be estimated using the mean of all $\hat{S}_{i,k\,\text{measured}}$

Then the multiplicative deviation pretreatment for each measured signal ($\frac{S_{i_{(j)},k\ \text{measured}}}{\hat{S}_{i,k\ \text{measured}}}$) allowed to model the deviations from 1 as an estimate of $E_{i,k\ \text{drift}}\ (t) + E_{i,k\ \text{noise}}$

To estimate time-dependent drift, a multivariate PLS regression between the pretreated signal measurements for all system sensors as independent variables (X matrix) and the time of analysis as a dependent variable (Y vector) was performed. As PLS drift model finds latent variables that explain the variability in the deviation of electronic signals due only to time evolution, this model function provides the estimate of $E_{i,k\ \text{drift}}\ (t)$ and the residuals of these model provide the estimation of $E_{i,k\ \text{noise}}$.

Accordingly, as proposed by [22], after the drift model was fitted, the matrix product of resulting loadings and scores of the model was used to calculate the matrix of $E_{\text{drift}}\ (t)$ components. Then the initial signal measured values were corrected for multiplicative drift using the following equation from [21]:

$$S_{i_{(j)},k\ \text{corrected}} = [\hat{S}_{i,k\ \text{measured}}(1 - E_{i_{(j)},k\ \text{drift}}\ (t))] + S_{i_{(j)},k\ \text{measured}} \quad (2)$$

A similar strategy was used to perform inter-sequence standardization to correct long-term drift. A practical guide is included in Supp. Fig. 2. For different work sequences, a generalization of equation (1) was considered to decompose signal measured in a sample *i* evaluated with *j* repetitions in each of the *k* sensors of the system. This generalization assumes that in this case the truth signal can be estimated using two components, the mean of all $S_{i_{(j)}}$repetitions and an inter-sequence standardization coefficient. To calculate this inter-sequence standardization coefficient, the difference of the signals of the same reference sample measured in two different sequences was used. The tomato-like standard TomSSt_2 was used as a reference sample in all work sequences. Consequently, the multiplicative deviation pretreatment used for each measured signal was:

$$\frac{S_{i_{(j)},k\ \text{measured}}}{\hat{S}_{i,k\ \text{measured}} + (\hat{S}_{\text{TomSSt1},k} - \hat{S}_{\text{TomSStn},k})} \quad (3)$$

Where $\hat{S}_{TomSSt1,k}$ and $\hat{S}_{TomSStn,k}$ are the signal means of all repetitions for the tomato-like synthetic standard reference sample in sequences 1 and *n*, respectively, for each *k* sensor.

The generalization of equation (1) also assumes that, when considering several work sequences, the time-dependent drift can be decomposed in two components:

$$E_{\text{drift}}(t) = E_{\text{short}}(t) + E_{\text{long}}(t) \quad (4)$$

Where E_{short}(t) represents the short-term (between-sample within-run) signal drift and E_{long}(t) the long-term (between-run) drift. Inter-sequence standardization was applied to all sequences after short-drift correction. Doing that, time-dependent drift would be equivalent to the long-term drift that appears between sequences. Consequently, after applying pretreatment of equation (3) when drift was modeled by PLS regression as explained previously, it was possible to calculate the matrix of E_{long} components and to use it to standardize sequence signals applying equation (4).

The PLS regressions were performed using venetian blinds (with as many groups as samples evaluated) as resampling procedure, in order to calculate error models and to select the number of latent variables used in the model. Outliers were detected and removed, using Hotelling T^2 and Q Residuals [23].

Graphical maps and data analysis tools

Drift-corrected sensor signals were graphically plotted in a 2D PLS-DA scatterplot map as with this dimensional reduction representation technique the distance between projected points preserves sample similarities [24]. Confidence ellipsoids (p=0.05) were calculated and plotted for samples with more than four replicates. In some cases, after removing outliers there were not enough points to calculate these intervals, and data points were just linked with lines to provide rapid identification of groups. The closer the points, the higher the similarity between signals. This procedure enables the comparison of sample volatile profile similarities, for example, for selection purposes. The objective was not to classify samples in predefined groups. This would be a typical objective in a quality assurance control, but in breeding programs, the objective is to select those materials closer to specific volatile profile targets. Nevertheless, to assess the performance of the proposed drift correction strategy, classification results were compared with those obtained using other reputed

drift correction methods: the original method proposed by Salit and Turk [21], independent component analysis (ICA) and parallel factor analysis 2 (PARAFAC2) [25]. ICA is a signal processing method that separates a multivariate signal into additive subcomponents assuming that the subcomponents are non-Gaussian signals and that they are statistically independent from each other. PARAFAC methods are generalizations of Principal Component Analysis (PCA) to higher order arrays. PARAFAC2 is an improvement of the original PARAFAC method in which the strict trilinearity is no longer required. Compared with PCA methods, PARAFAC methods have the advantages of no rotation problem, as in PCA, easier to interpret and higher statistical robustness.

Once the correction was obtained, three frequent classification techniques were applied. K nearest neighbors (KNN) classification, soft independent modeling of class analogy (SIMCA), and discriminant analysis based on partial least square regression (PLS-DA) [24] . KNN is a non-parametric classification method in which a data point is assigned to the class most common among its k nearest neighbors. SIMCA classification is mainly based on principal component analysis and an object is assigned to a class if its residual distance is below the statistical limit for the class. In PLS-DA, the predictive modelling comprises two main procedures, a PLS component development (i.e. dimension reduction for selecting variables for classification) and a prediction model construction (i.e. discriminant analysis) to predict class assignment for the data.

KNN, SIMCA, PLS and PLS-DA, analysis and graphics were performed using PLS_Toolbox v 8.6 (Eigenvector Research Inc, Wenatchee, WA, USA) for Matlab v 9 (Mathworks Inc, Natick, MA, USA). ICA models [26] were calculated with the FastICA toolbox for Matlab developed at the Helsinki University of Technology. PARAFAC2 models were performed using a graphical user interface, SENSABLE [20].

To justify the need for standardization procedures, tests for significant differences between the same sample signals in different sequence work using MANOVA analysis and Roy test were used [24]. These analyses were performed using IBM SPSS v.24 (IBM Corp., Armonk, NY, USA).

Results and discussion

Short-term drift correction

In the first assay, high levels of short-term drift were observed leading to a high variation in the position of each sample replicate in the two-dimensional representation of the PLS analysis obtained with raw signals (Fig. 1 a, b, and c). This variation could be related to a possible lack of homogeneity of real tomato samples, but a considerable variability was also detected in tomato-like standards which are highly homogeneous. As a consequence, despite having a different aroma volatile profile, the confidence ellipsoids of each variety overlapped. Thus, it was impossible to discriminate the materials. This effect of short-term drift was detected in the three independent sequences tested, but it affected each sequence differentially. As an example, the confidence ellipsoid of the tomato-like standard TomSSt_2 was small and data points plotted close in the first sequence (Fig. 1a), while the ellipsoid was considerably wider in the second (Fig.1b) and third sequences (Fig. 1c). The contrary was observed in the case of TomSSt1, with higher variability in the first sequence and lower in the second and third. As the samples were randomly distributed for each replicate in the sequence, the differences observed in confidence ellipsoids suggest that the effect of drift changes between sequences. This spurious trend confirmed the difficulty of extrapolating short-term drift effects on different analysis sessions.

The effect of sequence duration on short-term drift was analyzed in-depth comparing the performance of long (22-hour) and short (8-hour) sequences using 8 samples, including 7 tomato varieties and one tomato-like standard (Table 2). This time, samples were randomized in the first replicate, but the order was maintained in the rest of the replicates to enable comparisons between varieties. The long sequence (22 hours), typical of situations where a high number of samples is to be analyzed, was obtained increasing the number of repetitions per sample up to 12. Raw sensor data from these analyses revealed, for all the samples and in all the sequences, the presence of an important drift effect that affected all the sensors. The drift affected differentially each variety, with the highest effects detected in the samples of the variety “Caramba” (Fig. 2).

These drift effects were more evident and important at the end of the sequence (Fig. 2a), showing a complex non-linear time-dependent variation, with positive and negative signals tending to converge to 0. In the case of "Caramba" samples, raw signals (12 repetitions distributed in a sequence of 60 analysis) showed a very high relative standard deviation (%RSD) for the complete sequence for all the sensors (Fig. 2b. first data in parenthesis).

In order to provide a reference, these values obtained with "Caramba", were compared to those obtained by Xu et al. [27] corresponding to 6 analyses with the same apparatus equipped with the same sensors (Fig. 2b in square brackets). %RSD values obtained in the present work were considerably higher. Thus, the use of long sequences such as these would be unacceptable. It should be considered though, that the material used by [27] was *Semen arecae*, a dried seed preparation from *Areca catechu* L. Therefore, differences in %RSD would be explained both by changes in the sample matrix and in the number of hours of work of the sensors per sequence.

When shorter sequences were considered (8 working hours, i.e. 4 "Caramba" samples distributed in a sequence of 18 injections) the drift levels were lower, but they continued to be excessive (Fig. 2d. first data in parenthesis).

The main factors contributing to e-nose drift effects in sensor performance are usually due to differences in temperature, humidity, changes in samples analyzed due to components interactions, or other uncontrolled effects. In the long term, the stability of MOS sensors could progressively be affected by sensor aging or poisoning affecting their performance. This includes changes in the morphology of the sensing layer and irreversibly bind of some sample compounds to metal oxides which diminish the catalytic oxidation of sample volatiles and affecting the sensors' resistance response [14]. In practice, the data distortion caused by sensor drift in short time scenarios (one or few work-sequences) has many times been avoided when the use of the data collected was strictly for classification purposes. In these cases, the use of advanced multivariate statistical classification methodologies makes it possible to obtain subjacent information from the raw signal characteristic of each sample group, discarding the rest of the signal information and thus diminishing the drift

distortions problems (see examples in [28–30]). Unconsciously, when using a multivariate classifying technique, the analysis identifies and, to preserve sample group characteristic information, it discards the "non-characteristic" part of raw sensor signal which is normally related to noise, drift, and other non-relevant information. Nevertheless, this "signal cleaning" is a collateral effect (unwanted effect) and, consequently, the success of this strategy is variable since the characteristic subjacent information of the group is highly dependent on the samples and the number of latent variables used to build up the classification model. When a reduced number of samples with important differences between them are evaluated or when the volatile composition of the samples is not complex the "signal cleaning" effect would work well, making it possible to classify the samples in a quite satisfactory way [31]. But, with this approach, it is not always possible to completely avoid drift distortion effects. It would be the case of complex samples (complex matrix and/or very complex mixtures of volatiles) or collections of samples with similar volatile profiles. Consequently, a drift correction strategy would be more convenient in those cases. In order to correct short-term drift effects, sensor drift of the second assay was modeled and subtracted from the raw signals. To do that, a multivariate adaptation of the multiplicative drift correction procedure proposed by Salit and Turk [21] combined with a PLS adaptation of the highly used component correction strategy [22] to specifically model each drift present in each sequence was performed. The following assumptions were considered: i) sensors of the array have similar drift behavior, ii) this drift has a specific direction in the data hyperspace which allows its modelization by regression, and iii) this drift is time-dependent. After modeling short-term drift for each sequence, drift components for each signal in the data matrix were calculated. Later, matrix subtraction was performed in a Matlab environment to remove drift from the raw sensor signal data, thus providing a corrected sensor data matrix, which was used to plot the data (Fig. 2e). Compared with the raw sensor signals (Fig. 2a), the corrected signals were much more stable during the whole sequence for all sensors, even those with higher %RSD. Accordingly, an impressive %RSD decrease was observed for all the sensors (Fig. 2b), ranging from between 91.5% and 99.7% for

long sequences and 75.7% and 98.8% for short sequences. Maximum %RSD values were 0.65% for long sequences and 0.72% for short sequences. Those values are were between one (T40/2 sensor) and 27 (LY2/LG sensor) times lower than those reported by Xu et al. [27] with a lower number of injections. As expected, the use of shorter work-sequences (18 injections in 8 hours sequence) resulted in better performance after drift correction (Fig. 2d and 2f), as it avoided the higher levels of drift detected at the end of long sequences.

It should be considered though, that the increase in stability entailed a small decrease in the absolute value of signals after correction. This side effect mainly affects long sequences (Fig. 2a vs. Fig. 2e), while this decrease is imperceptible in shorter sequences (Fig. 2c vs. Fig. 2f). Consequently, despite the powerful short-term drift correction capabilities obtained, it would be preferable to use short (8 hours) work sequences.

When this drift correction strategy was applied to the signals of the first assay, an impressive reduction of the sample signal variability was attained, enabling a clear comparison of similitudes between samples in the new PLS-DA similarity map obtained (corrected: Fig. 1d, 1e, and 1f vs. raw: Fig. 1a, 1b, and 1c). Indeed, after this correction, it was easy to ascertain similarities in the volatile signal profile between samples, and the confidence intervals did not overlap as it had happened with the raw data.

The use of similitude maps to compare volatile profiles is a novel alternative. Therefore, in order to compare this strategy with previous works it was necessary to assess its performance using classification methodologies, which are rather popular in e-nose preceding literature. Consequently, using the data of the second assay, the new drift correction strategy was compared with alternative drift correction methods including the original approach by Salit and Turk, [21], ICA [32], using KNN, SIMCA, and PLS-DA as classifying methods. In general, SIMCA outstood in the classifications. KNN and PLS-DA had a similar performance, which varied depending on the variety considered (Table 3). Considering different alternatives, the new correction proposed in this work offered the best results (classification effectiveness) compared to the alternatives evaluated

independently of the classification method. In fact, SIMCA and KNN classification with the proposed short-term drift correction allowed to classify correctly a 100% of the samples, assigning them to the variety to which they belonged.

Table 3

Percentage of samples correctly classified using KNN (K=8), SIMCA and PLS-DA classification methods for seven tomato cultivars and the tomato-like synthetic standard 2, before (raw data) and after intra-sequence drift correction using the proposed correction based on an adaptation of [21] and PLS component correction method, the Salit and Turk [21], ICA [32] and PARAFAC2 [20] methods. Average data of three work-sequences is provided (variation range in brackets).

Sample	Raw data			Proposed correction			Salit & Turk correction			ICA correction			PARAFAC2 correction		
	KNN	SIMCA	PLS-DA	KNN	SIMCA	PLS-DA	KNN	SIMCA	PLS-DA	KNN	SIMCA	PLS-DA	KNN	SIMCA	PLS-DA
TomSSt_2[a]	98.9 (97.9-100)	100	100	100	100	100	97.6 (93.8-100)	56.3 (51.1-60.6)	100	94.8 (86.5-100)	100	93.8 (81.3-100)	99.0 (97.9-100)	100	93.7 (88.3-99.0)
Zayno	85.4 (78.6-92.9)	82.0 (49-99)	90.4 (84.7-100)	100	100	99.3 (98.0-100)	92.2 (83.6-100)	60.1 (52-67.7)	93.5 (88.7-100)	83.6 (76.5-91.8)	96.6 (89.8-100)	80.9 (74.5-85.7)	80.3 (77.5-85.7)	81.6 (72.4-89.6)	72.7 (66.3-80.6)
"Amarillo" (BGV005718)	91.5 (87.8-99.0)	98.3 (94.8-100)	88.4 (83.5-91.8)	100	100	94.6 (90.8-100)	99.7 (99.0-100)	57.9 (52.1-59.4)	89.7 (85.6-91.8)	79.9 (64.3-98.0)	98.3 (94.8-100)	83.3 (76.5-87.8)	86.4 (74.4-93.9)	88.4 (85.7-90.6)	68.5 (55.1-80.6)
Caramba	60.5 (47.9-84.7)	76.5 (69-89.8)	76.8 (68.0-83.7)	100	100	94.5 (92.9-98.0)	89.0 (82.6-92.9)	52.4 (51-54.1)	80.8 (78.3-85.6)	54.8 (47-62.2)	76.5 (70.8-89.8)	81.3 (61.2-91.7)	66.8 (62.7-69.3)	64.9 (56.3-71.4)	60.5 (48.4-71.0)
Breeding line (UJI011)	78.2 (71.4-82.6)	79.8 (68-90.8)	77.7 (70.8-83.7)	100	100	92.4 (82.2-100)	94.2 (90.8-99.0)	53.8 (53.1-55.2)	78.3 (69.9-86.7)	63.6 (55.1-76.5)	83.2 (68-90.8)	65.2 (61.2-67.3)	61.2 (52-71.4)	67.5 (53.1-79.6)	67.9 (61.2-73.5)
De penjar (UJI023)	96.3 (91.8-99.0)	100	99.7 (99.0-100)	100	100	100	96.5 (90.6-100)	60.7 (56.1-65.6)	97.2 (91.7-100)	75.8 (56.1-86.6)	100	88.1 (71.4-100)	93.8 (90.6-98.0)	94.9 (89.8-98)	84.3 (81.6-89.8)
Morado (BGV005477)	88.8 (76.5-99.0)	100	92.8 (89.6-94.9)	100	100	98.0 (93.9-100)	97.3 (92.9-100)	72.4 (66.7-82.3)	94.5 (92.7-96.9)	79.2 (45.9-98.0)	100	90.1 (79.6-96.9)	95.2 (90.8-98.0)	91.5 (79.6-97.9)	85.3 (81.3-88.8)
Muchamiel (BGV005651)	82.8 (82.3-83.7)	96.9 (90.8-100)	77.9 (74.3-79.7)	100	100	95.9 (93.9-98.0)	91.4 (82.3-100)	57.8 (55.1-63.3)	84.3 (77.4-93.9)	79.7 (54.3-94.0)	96.9 (90.6-100)	83.7 (75.3-91.0)	85.8 (82.3-91.8)	93.2 (90.8-96)	77.9 (75.3-81.6)

Long-term drift and sequence standardization.

Once the problem of short-term drift was solved, the focus was set on the effects of the variability detected among sequences. This variability, as stated above, can be generated by different causes originating a long-term drift effect. A solution to this effect is critical when a high number of samples are to be analyzed, as samples have to be distributed in different sequences that would be run on several days.

Regardless of the cause of inter-sequence variability, the effects can be considerable and unpredictable, as was pointed out in the comparison of the three sequences of the first assay. Consequently, it seemed clear that some reference samples should be included in each sequence to assess how long-term drift affected the signal. At this point, it would not be advisable to use real tomato samples as references. The storage capability of these samples would be limited, and long-term evolution in a freezer would introduce an undesirable noise in the system, thus increasing long-term drift. Accordingly, it was decided to include tomato-like synthetic standard volatile solutions, which were designed and created for this purpose. As tomato volatile profile is highly complex, with more than 400 volatiles being involved, it was decided to focus on a group of compounds (Table 1) that had been suggested to hold a prominent role in the aroma perception [33, 34]. Standards were created from stock solutions for each session. Nonetheless, in the future and for practical reasons, standards can be created and stored in sealed vials at -30°C during one month with a high stability. In this case, over a 3-month span, the standards were created specifically for each session, thus providing more restrictive conditions.

In a first step, three different sequences with the tomato-like standards at different concentrations were run. After applying the proposed short-term drift correction, a PLS-2D map was obtained (Fig. 3a). Samples from the same tomato-like standard tended to group together, but still, a considerable level of variation was detected. In some cases, the confidence intervals of the same samples run in different sequences did not overlap and intervals of different standards did overlap in one case.

Considering the homogeneous nature of these standards, this variability would not be mainly related with the nature of the sample. To check this point, the analysis was repeated including samples from two tomato varieties "Rayno RZ" and "Amarillo". Again, wide variability was detected, which was not specifically higher in the real tomato samples than in the standard solutions, despite their more complex nature (Fig. 3b).

This time, even in the case of the control with lower variability (TomSSt_1), the fluctuations of signal values were rather high for some sensors, reaching RSD values above 20% (e.g. LY2/gCTl and LY2/GH sensors) or very close to this threshold (e.g. LY2/G sensor). In fact, a MANOVA analysis for TomSSt_1 using the data from the three sequences showed significant inter-sequence differences (Roy test $\alpha < 0.03$). Higher levels of variation were found in the rest of the controls. Consequently, despite the use of the routine instrument calibration recommended by the equipment manufacturer, the unacceptable inter-sequence variance for each sample caused important bias in the graphs constructed joining the data from several sequences. Therefore, it makes necessary the use of a data standardization step before merging data from different sessions.

In order to tackle this long-term drift effect, the data from the tomato-like synthetic standard TomSSt_2 was selected to standardize sequence signals. The use of a real sample as reference had to be discarded, as its volatile profile would evolve during their conservation and it would also have a finite nature. On the opposite, a homogeneous synthetic standard including main tomato volatiles, representing the complex nature of its aroma, can be generated expressly for each sequence.

Following this premise, in order to standardize sequences, sensor signals from each sample after short-term drift correction were transformed using the deviation observed between the corrected signals of the synthetic standard in the different sequences. Once the signals were transformed, they were related to a time vector using PLS regression. Time vector values were obtained adding the time of each analysis, including the different sequences consecutively.

New PLS-DA 2D maps were then obtained, and the efficiency of correction was evident (Fig. 3a

vs 3b). For five of the six controls no significant inter-sequence differences were found, and the confidence ellipsoids overlapped. Only in the case of the samples of the tomato landrace "Amarillo" (coded 2_1 in Fig. 3) significant differences (Roy test $\alpha < 0.001$) were found between the first sequence and the remaining two. Nonetheless, the three samples plotted at a short distance. The standardization procedure showed a grouping correction efficiency of 94.4%, as 17 of the 18 sample groups were correctly ascribed with their equals ran in different sequences and their confidence intervals overlapped. This result represents a similar efficiency compared to other strategies regarding long-term drift counteraction methods [15, 35–40] or better [41, 42]. It was confirmed, then, that data from different sequences could be pooled in order to work with a high number of samples.

Considering the good performance obtained with these controls, the sequence standardization procedure was applied to the data obtained with three sequences, with 14 tomato varieties and TomSSt_2 as a reference. When both short-term drift correction and sequence standardization was applied (Fig. 4b), the variation observed per sample was highly reduced compared to the use of raw data (Fig. 4a). Again, the replicates analyzed in different sequences tended to overlap their confidence intervals, and only one of the replicates of the "Amarillo" landrace could not be grouped with the rest of the corresponding replicates (coded 2.1 in Fig 4b). Therefore, this procedure enables a realistic comparison of similitudes and differences in the volatile signal profile between samples run in different sequences.

Other works [15, 35, 41] deal with adaptations of the component correction strategy applied to a long-term drift counteraction. These works use a group of training samples to model the drift using different regression methodologies (PLS, OSC, or CPCA) and, then subtract the drift modeled from the signals of new samples. These strategies assume that with a good training set, the calibration model can be useful for a long time for practical purposes. However, it is obvious that to extend the period of use, large training sets are needed. Gutierrez-Osuna [43] used a training set of 5 to 10

samples for a drift correction period of 3 months in samples of 4 very different spices. Padilla et al., [35] used training sets higher than 100 samples for a drift correction period of 10 months in samples of individual chemical compounds at different concentrations. A similar application was tested by Ziyatdinov [41] with training sets higher than 1000 samples for a drift correction period of 7 months. Nevertheless, it seems also obvious that when sensor degradation increases, the usefulness of these calibration models will decrease and, at any moment, they would need an update. Additionally, training sets have been used with mixes of a few volatiles, and real tomato samples consist of more than 400 volatiles [34].

In the present study, specific training set samples were not used. Instead, the information of the samples evaluated in each sequence was used to calculate the specific drift correction model. Four injections per sample would be enough to model short-term drift and at the same time providing a reliable confidence interval. By doing so, each sequence would have its proper model and, consequently, it would always be up to date. The unpredictable nature of short-term drift in different sequences using tomato matrices would limit the efficiency of other alternatives.

On the other hand, the use of one reference synthetic standard has proven to be highly efficient to standardize sequences in order to reduce inter-sequence variability, enabling the comparison of samples analyzed in different sequences. This strategy would also be useful when a replacement of sensors is performed or when different instruments are used to enlarge the processing capabilities of the lab. Tomic et al., [37] tried a similar component correction strategy based on PCA and complemented with multiplicative drift correction to accomplish a successful calibration transfer between instruments. Other calibration transfer strategies which use sophisticated correction methods and algorithms have been also applied to the expansion of calibration update models [38, 39, 44] but they need a higher number of training samples (10 to more than 400 depending on the methodology) and were tested only for the detection of individual chemical compounds, so the efficiency in more complex samples still needs to be tested [14].

Combining short-term and sequence standardization and PLS-DA 2D similitude maps it is possible to easily identify differences in the volatile signal profiles of the samples. It is then possible to make rapid identification of those samples with a volatile profile more similar to high quality reference materials. This procedure would enable the use of e-noses for example in breeding programs. It would be possible to select which genetic backgrounds have a lower negative impact on the aroma profile. From an agronomic point of view, it would also enable a rapid identification of which preharvest and postharvest procedures have the lowest impact on the volatile profile. These maps would be expandable, offering the possibility of including new reference points. In fact, when Fig. 3c and 4b are compared, the relative position of the real tomato samples of "Zayno RZ" (coded 1 in the figures) and the "Amarillo" landrace (coded 2 in the figures) were not altered.

In the present work, this strategy has been successfully applied to a combination of different tomato materials, selected to represent a wide variability of volatile profiles, especially in the case of tomato landraces. The landraces included in the study had already shown a clearly different volatile profile [45], and especially important as they are frequently commercialized in quality markets in which consumers are willing to pay a price premium for excellent flavor [46]. Interestingly, "Muchamiel", which had previously shown a less intense volatile profile in gas chromatography analysis compared to "Valenciano" and "Morada", plotted in the PLS-DA 2D map in an area corresponding to materials with lower volatile concentration (Fig.4b). The next step in future works will be centered on the comparison of the volatile profile obtained with the e-nose and GC-MS data in order to confirm this trend.

Conclusions

Short- and long-term drift compromises the application of e-noses to the evaluation of volatile profiles. These effects are variable and unpredictable. Consequently, general models are not useful, and the performance registered in each sequence has to be used in order to model drift effects. The

distribution of 4 replicates per sample and sequence enables the development of an effective and sequence-specific short-term drift correction. On the other hand, the unpredictable nature of the variation between sessions makes it necessary to use reference materials to standardize sequences. By doing so it would be possible to analyze a high number of samples distributed in different sequences. The use of a tomato-like synthetic has proven to be for this purpose. The two-step correction methodology proposed here, combined with PLS-DA two-dimensional similitude maps, will enable rapid and reliable identification of samples with a volatile signal profile similar to references selected as ideal targets.

Declarations

Funding

This research was partially funded by Jaume I University with projects P1-1B2011-41 and COGRUP/2016/04. G. Ibáñez also thanks Universitat Jaume I for funding his pre-doctoral grant (PREDOC/2015/45).

Conflicts of interest

The authors declare that there is no conflict of interest.

Availability of data and material

Data available on request to the authors

Code availability

Not applicable

Authors' contributions

Mercedes Valcárcel: Conceptualization, Methodology, Investigation, Formal analysis, Writing - Original Draft, and Supervision. Ginés Ibáñez: Formal analysis and Visualization. Raúl Martí: Formal analysis and Visualization. Joaquín Beltrán: Resources and Validation. Jaime Cebolla-Cornejo: Writing-Reviewing and Editing. Salvador Roselló: Conceptualization, Formal analysis, Writing - Original Draft and Supervision

Ethics approval

Not applicable

Consent to participate

All the authors have consented to participate on the research. The research does not include humans nor animals as subjects of research.

Consent for publication

All the authors consent the publication of the research

References

1. Tieman D, Zhu G, Resende MFR, Lin T, Nguyen C, Bies D, Rambla JL, Beltran KSO, Taylor M, Zhang B, Ikeda H, Liu Z, Fisher J, Zemach I, Monforte A, Zamir D, Granell A, Kirst M, Huang S, Klee H (2017) A chemical genetic roadmap to improved tomato flavor. Science (80-) 355:391–394 . https://doi.org/10.1126/science.aal1556
2. Baldwin EA, Scott J, Shewmaker CK, Schuh W (2000) Flavor trivia and tomato aroma: biochemistry and possible mechanism for control of important aroma components. HortScience 35:1013–1022
3. Davies JN, Hobson GE (1981) The constituents of tomato fruit — the influence of environment, nutrition, and genotype. C R C Crit Rev Food Sci Nutr 15:205–280 . https://doi.org/10.1080/10408398109527317
4. Lahoz I, Pérez-de-Castro A, Valcárcel M, Macua JI, Beltrán J, Roselló S, Cebolla-Cornejo J (2016) Effect of water deficit on the agronomical performance and quality of processing tomato. Sci Hortic (Amsterdam) 200: . https://doi.org/10.1016/j.scienta.2015.12.051
5. Schouten RE, Woltering EJ, Tijskens LMM (2016) Sugar and acid interconversion in tomato fruits based on biopsy sampling of locule gel and pericarp tissue. Postharvest Biol Technol 111:83–92 . https://doi.org/10.1016/j.postharvbio.2015.07.032
6. Boukobza F, Taylor AJ (2002) Effect of postharvest treatment on flavour volatiles of tomatoes.

Postharvest Biol Technol 25:321–331 . https://doi.org/10.1016/S0925-5214(02)00037-6

7. Zhao J, Sauvage C, Zhao J, Bitton F, Bauchet G, Liu D, Huang S, Tieman DM, Klee HJ, Causse M (2019) Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. Nat Commun 1–43 . https://doi.org/10.1038/s41467-019-09462-w

8. Loutfi A, Coradeschi S, Mani GK, Shankar P, Rayappan JBB (2015) Electronic noses for food quality: A review. J Food Eng 144:103–111 . https://doi.org/10.1016/j.jfoodeng.2014.07.019

9. Kiani S, Minaei S, Ghasemi-Varnamkhasti M (2016) Application of electronic nose systems for assessing quality of medicinal and aromatic plant products: A review. J Appl Res Med Aromat Plants 3:1–9 . https://doi.org/10.1016/j.jarmap.2015.12.002

10. Sun Y, Wang J, Cheng S (2017) Discrimination among tea plants either with different invasive severities or different invasive times using MOS electronic nose combined with a new feature extraction method. Comput Electron Agric 143:293–301 . https://doi.org/10.1016/j.compag.2017.11.007

11. Majchrzak T, Wojnowski W, Dymerski T, Gębicki J, Namieśnik J (2018) Electronic noses in classification and quality control of edible oils: A review. Food Chem 246:192–201 . https://doi.org/10.1016/j.foodchem.2017.11.013

12. Jia W, Liang G, Jiang Z, Wang J (2019) Advances in Electronic Nose Development for Application to Agricultural Products

13. Marco S, Gutierrez-Galvez A (2012) Signal and data processing for machine olfaction and chemical sensing: A review. IEEE Sens J 12:3189–3214 . https://doi.org/10.1109/JSEN.2012.2192920

14. Rudnitskaya A (2018) Calibration Update and Drift Correction for Electronic Noses and Tongues. Front Chem 6: . https://doi.org/10.3389/fchem.2018.00433

15. Gutierrez-Osuna R (2002) Pattern analysis for machine olfaction: a review. Sensors Journal, IEEE 2:189–202

16. Beltran J, Serrano E, López FJ, Peruga a, Valcarcel M, Rosello S (2006) Comparison of two quantitative GC-MS methods for analysis of tomato aroma based on purge-and-trap and on solid-phase microextraction. Anal Bioanal Chem 385:1255–64 . https://doi.org/10.1007/s00216-006-0410-9

17. Casals J, Pascual L, Cañizares J, Cebolla-Cornejo J, Casañas F, Nuez F (2012) Genetic basis of long shelf life and variability into Penjar tomato. Genet Resour Crop Evol 59:219–229 . https://doi.org/10.1007/s10722-011-9677-6

18. Roselló S, Nuez F, Casals J, Beltrán J, Casañas F, Cebolla-Cornejo J (2011) Long-term postharvest aroma evolution of tomatoes with the alcobaça (alc) mutation. Eur. Food Res. Technol. 233:331–342

19. Cebolla-Cornejo J, Roselló S, Nuez F (2013) Selection of tomato rich in nutritional terpenes

20. Skov T, Bro R (2005) A new approach for modelling sensor based data. Sensors Actuators, B Chem 106:719–729 . https://doi.org/10.1016/j.snb.2004.09.023

21. Salit ML, Turk GC (1998) A drift correction procedure. Anal Chem 70:3184–90 . https://doi.org/10.1021/ac980095b

22. Artursson T, Eklo T, Lundstro I, Sjo M (2000) Drift correction for gas sensors using multivariate methods ˚. J Chemom 711–723

23. Ballabio D, Consonni V (2013) Classification tools in chemistry. Part 1: linear models. PLS-DA. Anal Methods 5:3790 . https://doi.org/10.1039/c3ay40582f

24. Krzanowski W (2000) Principles of multivariate analysis: A User's Perspective. Oxford University Press

25. Di Natale C, Martinelli E, D'Amico A (2002) Counteraction of environmental disturbances of electronic nose data by independent component analysis. Sensors Actuators B Chem 82:158–165

26. Hyvärinen a, Oja E (2000) Independent component analysis: algorithms and applications. Neural Netw 13:411–30

27. Xu M, Yang SL, Peng W, Liu YJ, Xie DS, Li XY, Wu CJ (2015) A novel method for the discrimination of semen arecae and its processed products by using computer vision, electronic nose, and electronic tongue. Evidence-based Complement Altern Med 2015: . https://doi.org/10.1155/2015/753942

28. Gromski PS, Correa E, Vaughan AA, Wedge DC, Turner ML, Goodacre R (2014) A comparison of different chemometrics approaches for the robust classification of electronic nose data. Anal Bioanal Chem 406:7581–7590 . https://doi.org/10.1007/s00216-014-8216-7

29. Shi Y, Gong F, Wang M, Liu J, Wu Y, Men H (2019) A deep feature mining method of electronic nose sensor data for identification identifying beer olfactory information. J Food Eng 263:437–445 . https://doi.org/10.1016/j.jfoodeng.2019.07.023

30. Song J, Bi J, Chen Q, Wu X, Lyu Y, Meng X (2019) Assessment of sugar content, fatty acids, free amino acids, and volatile profiles in jujube fruits at different ripening stages. Food Chem 270:344–352 . https://doi.org/10.1016/j.foodchem.2018.07.102

31. Hong X, Wang J, Qi G (2015) E-nose combined with chemometrics to trace tomato-juice quality. J Food Eng 149:38–43 . https://doi.org/10.1016/j.jfoodeng.2014.10.003

32. Di Natale C, Martinelli E, D'Amico A (2002) Counteraction of environmental disturbances of electronic nose data by independent component analysis. Sensors Actuators B Chem 82:158–165 . https://doi.org/10.1016/S0925-4005(01)01001-2

33. Buttery R, Teranishi R, Ling LC (1987) Fresh tomato aroma volatiles: A quantitative study. J Agric Food Chem 35:540–544

34. Tieman D, Bliss P, McIntyre LM, Blandon-Ubeda A, Bies D, Odabasi AZ, Rodríguez GR, Van Der Knaap E, Taylor MG, Goulet C, Mageroy MH, Snyder DJ, Colquhoun T, Moskowitz H, Clark DG, Sims C, Bartoshuk L, Klee HJ (2012) The chemical interactions underlying tomato flavor preferences. Curr Biol 22:1035–1039 . https://doi.org/10.1016/j.cub.2012.04.016

35. Padilla M, Perera a., Montoliu I, Chaudry a., Persaud K, Marco S (2010) Drift compensation

of gas sensor array data by Orthogonal Signal Correction. Chemom Intell Lab Syst 100:28–35 . https://doi.org/10.1016/j.chemolab.2009.10.002

36. Tomic O, Ulmer H, Haugen JE (2002) Standardization methods for handling instrument related signal shift in gas-sensor array measurement data. Anal Chim Acta 472:99–111 . https://doi.org/10.1016/S0003-2670(02)00936-4

37. Tomic O, Eklöv T, Kvaal K, Haugen JE (2004) Recalibration of a gas-sensor array system related to sensor replacement. Anal Chim Acta 512:199–206 . https://doi.org/10.1016/j.aca.2004.03.001

38. Zhang L, Zhang D (2015) Domain Adaptation Extreme Learning Machines for Drift Compensation in E-Nose Systems. IEEE Trans Instrum Meas 64:1790–1801 . https://doi.org/10.1109/TIM.2014.2367775

39. Yan K, Zhang D (2016) Calibration transfer and drift compensation of e-noses via coupled task learning. Sensors Actuators, B Chem 225:288–297 . https://doi.org/10.1016/j.snb.2015.11.058

40. Solórzano A, Rodríguez-Pérez R, Padilla M, Graunke T, Fernandez L, Marco S, Fonollosa J (2018) Multi-unit calibration rejects inherent device variability of chemical sensor arrays. Sensors Actuators, B Chem 265:142–154 . https://doi.org/10.1016/j.snb.2018.02.188

41. Ziyatdinov A, Marco S, Chaudry A, Persaud K, Caminal P, Perera A (2010) Drift compensation of gas sensor array data by common principal component analysis. Sensors Actuators, B Chem 146:460–465 . https://doi.org/10.1016/j.snb.2009.11.034

42. Fernandez L, Guney S, Gutierrez-Galvez A, Marco S (2016) Calibration transfer in temperature modulated gas sensor arrays. Sensors Actuators, B Chem 231:276–284 . https://doi.org/10.1016/j.snb.2016.02.131

43. Gutierrez-Osuna R (2000) Drift reduction for metal-oxide sensor arrays using canonical correlation regression and partial least squares. Proc 7th Int Symp Olfaction Electron Nose 1–7

44. Vergara A, Vembu S, Ayhan T, Ryan MA, Homer ML, Huerta R (2012) Chemical gas sensor

drift compensation using classifier ensembles. Sensors Actuators, B Chem 166–167:320–329 . https://doi.org/10.1016/j.snb.2012.01.074

45. Cebolla-Cornejo J, Roselló S, Valcárcel M, Serrano E, Beltrán J, Nuez F (2011) Evaluation of genotype and environment effects on taste and aroma flavour components of Spanish fresh tomato. J Agric Food Chem 59:2440–2450

46. Cebolla-Cornejo J, Soler S, Nuez F (2007) Genetic erosion of traditional varieties of vegetable crops in Europe: tomato cultivation in Valencia (Spain) as a case Study. Int J Plant Prod 1:113–128

Fig. 1 Similarity PLS-DA maps of volatile electronic profiles from raw signals (on the left) and short-term drift-corrected signals (on the right) data from samples tested in three different work sequences assayed in the first assay. Sample codes as indicated in Table 2. TomSSt = Tomato-like synthetic standard. Ellipsoids represent confidence intervals (p=0.05) for samples with more than three replicates. For samples with a lower number of replicates confidence intervals cannot be calculated, and sample points are only connected with lines to easily identify them.

Fig. 2 Sensors response in the evaluation of tomato "Caramba" samples in (top) long work sequences (22h) and (bottom) short work sequences (8h). On the left: Raw sensor signals (a, c). On the right: intra-sequence drift corrected sensor signal (e, f). Legends (b, d) show the evolution of sensor signals %RSD before and after applying intra-sequence drift correction (first and second value in parenthesis). External reference %RSD values using the same equipment and 6 injections is provided in square brackets [27].

Fig. 3 PLS-DA similitude maps from electronic nose fingerprints obtained in three different sequences and applying short term-drift correction using (a) the 4 synthetic tomato-like standards and two tomato samples, and (b) the standards and two real tomato samples applying sequence standardization. Confidence ellipsoids (P=0.05) are represented for samples with more than 4 repetitions. TomSSt: tomato-like synthetic standards. 1: "Zayno RZ"; 2: BGV005718 (real tomato samples used as controls). _1, _2, _3: sequence number.

Fig. 4 PLS-DA similitude map obtained using 14 tomato varieties evaluated in three different work sequences. (a) using raw data, (b) using short-term intra-sequence drift correction + sequence standardization. Confidence ellipsoids (P=0.05). 1: "Zayno RZ"; 2: BGV005718; 3: "Caramba"; 4: UJI011; 5: "Divyne RZ"; 6: "Vinchy RZ"; 7: UJI023; 8: BGV005477; 9: BGV005651; 10: BGV005655; 11: BGV004587; 12: UJI008; 13: UJI014; 14: UJI028. _1, _2, _3 samples analyzed in different sequences.

Supp. Fig. 1. Schematic representation of short-term drift within a sequence.

Supp. Fig. 2. Schematic representation of long-term drift with several sequences.

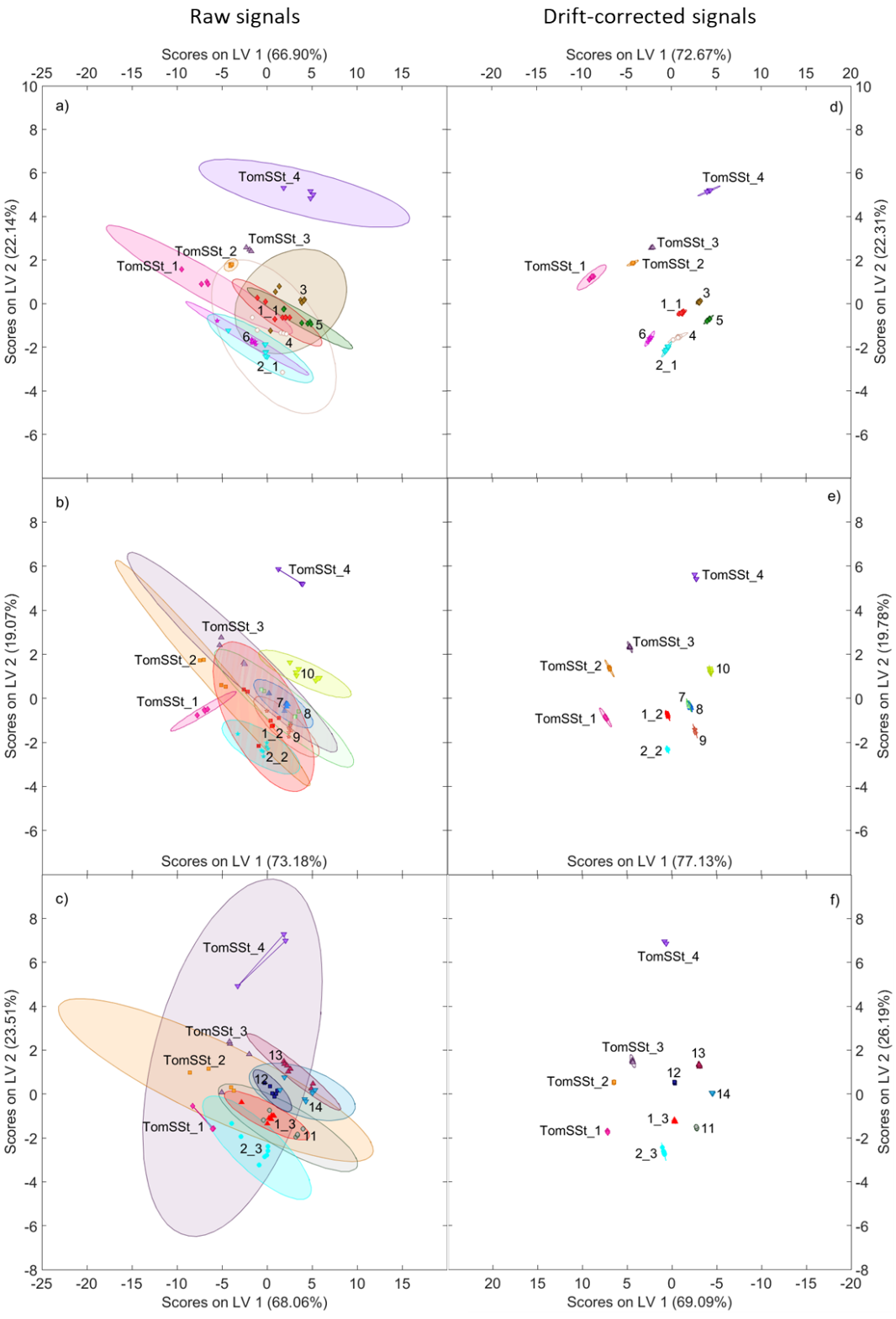

Raw signals
Drift-corrected signals
a)
b)
c)
d)
e)
f)
Scores on LV 1 (66.90%)
Scores on LV 2 (22.14%)
Scores on LV 1 (72.67%)
Scores on LV 2 (22.31%)
Scores on LV 2 (19.07%)
Scores on LV 1 (73.18%)
Scores on LV 2 (19.78%)
Scores on LV 1 (77.13%)
Scores on LV 2 (23.51%)
Scores on LV 1 (68.06%)
Scores on LV 2 (26.19%)
Scores on LV 1 (69.09%)
TomSSt_1
TomSSt_2
TomSSt_3
TomSSt_4

Fig. 1

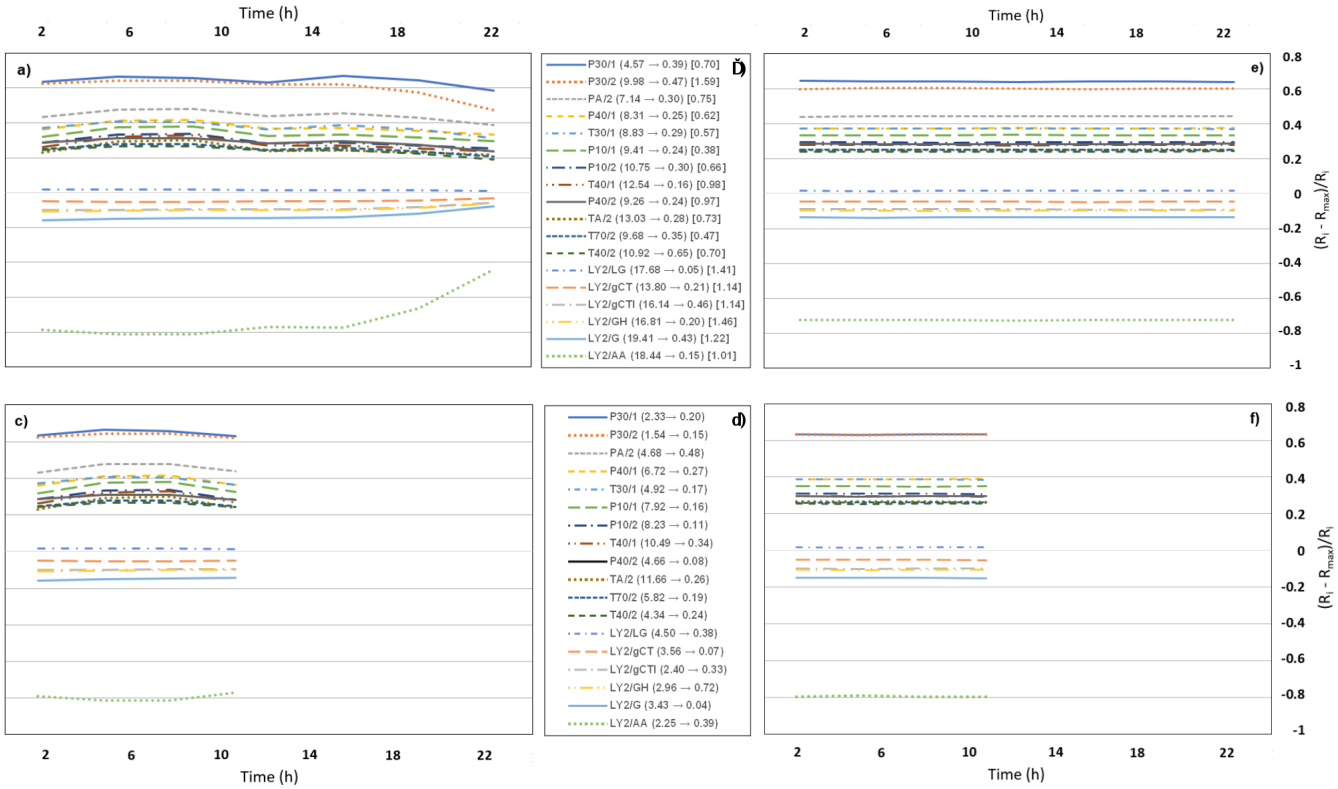

a)
b)
c)
d)
e)
f)
Time (h)
$(R_i - R_{max})/R_i$
P30/1 (4.57 → 0.39) [0.70]
P30/2 (9.98 → 0.47) [1.59]
PA/2 (7.14 → 0.30) [0.75]
P40/1 (8.31 → 0.25) [0.62]
T30/1 (8.83 → 0.29) [0.57]
P10/1 (9.41 → 0.24) [0.38]
P10/2 (10.75 → 0.30) [0.66]
T40/1 (12.54 → 0.16) [0.98]
P40/2 (9.26 → 0.24) [0.97]
TA/2 (13.03 → 0.28) [0.73]
T70/2 (9.68 → 0.35) [0.47]
T40/2 (10.92 → 0.65) [0.70]
LY2/LG (17.68 → 0.05) [1.41]
LY2/gCT (13.80 → 0.21) [1.14]
LY2/gCTI (16.14 → 0.46) [1.14]
LY2/GH (16.81 → 0.20) [1.46]
LY2/G (19.41 → 0.43) [1.22]
LY2/AA (18.44 → 0.15) [1.01]
P30/1 (2.33→ 0.20)
P30/2 (1.54 → 0.15)
PA/2 (4.68 → 0.48)
P40/1 (6.72 → 0.27)
T30/1 (4.92 → 0.17)
P10/1 (7.92 → 0.16)
P10/2 (8.23 → 0.11)
T40/1 (10.49 → 0.34)
P40/2 (4.66 → 0.08)
TA/2 (11.66 → 0.26)
T70/2 (5.82 → 0.19)
T40/2 (4.34 → 0.24)
LY2/LG (4.50 → 0.38)
LY2/gCT (3.56 → 0.07)
LY2/gCTI (2.40 → 0.33)
LY2/GH (2.96 → 0.72)
LY2/G (3.43 → 0.04)
LY2/AA (2.25 → 0.39)

Fig. 2

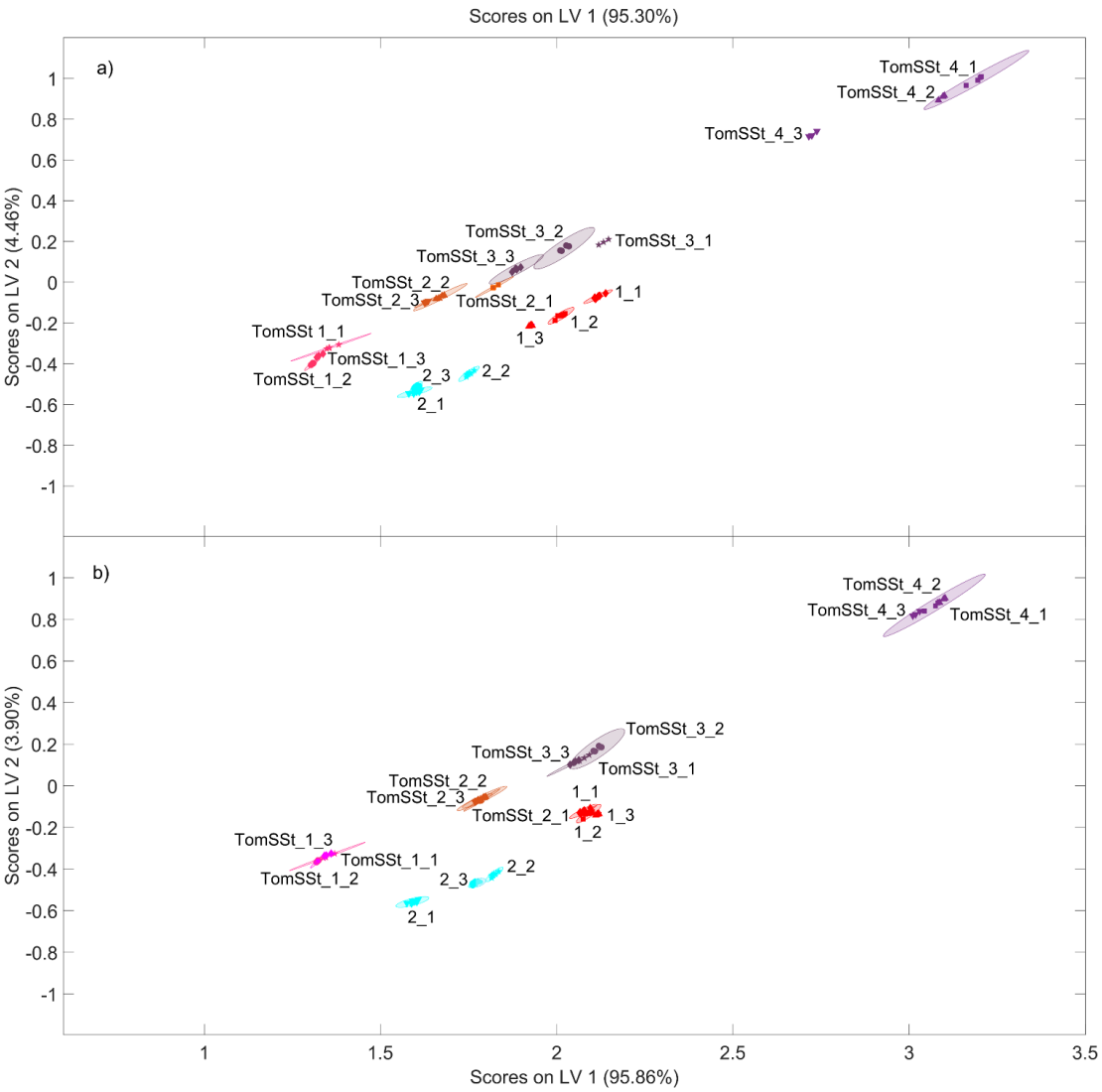

Scores on LV 1 (95.30%)
a)
Scores on LV 2 (4.46%)
TomSSt_4_1
TomSSt_4_2
TomSSt_4_3
TomSSt_3_2
TomSSt_3_1
TomSSt_3_3
TomSSt_2_2
TomSSt_2_3
TomSSt_2_1
1_1
1_2
1_3
TomSSt 1_1
TomSSt_1_3
TomSSt_1_2
2_3
2_2
2_1
b)
Scores on LV 2 (3.90%)
TomSSt_4_2
TomSSt_4_3
TomSSt_4_1
TomSSt_3_2
TomSSt_3_3
TomSSt_3_1
TomSSt_2_2
TomSSt_2_3
1_1
TomSSt_2_1
1_3
1_2
TomSSt_1_3
TomSSt_1_1
TomSSt_1_2
2_3
2_2
2_1
Scores on LV 1 (95.86%)

Fig. 3

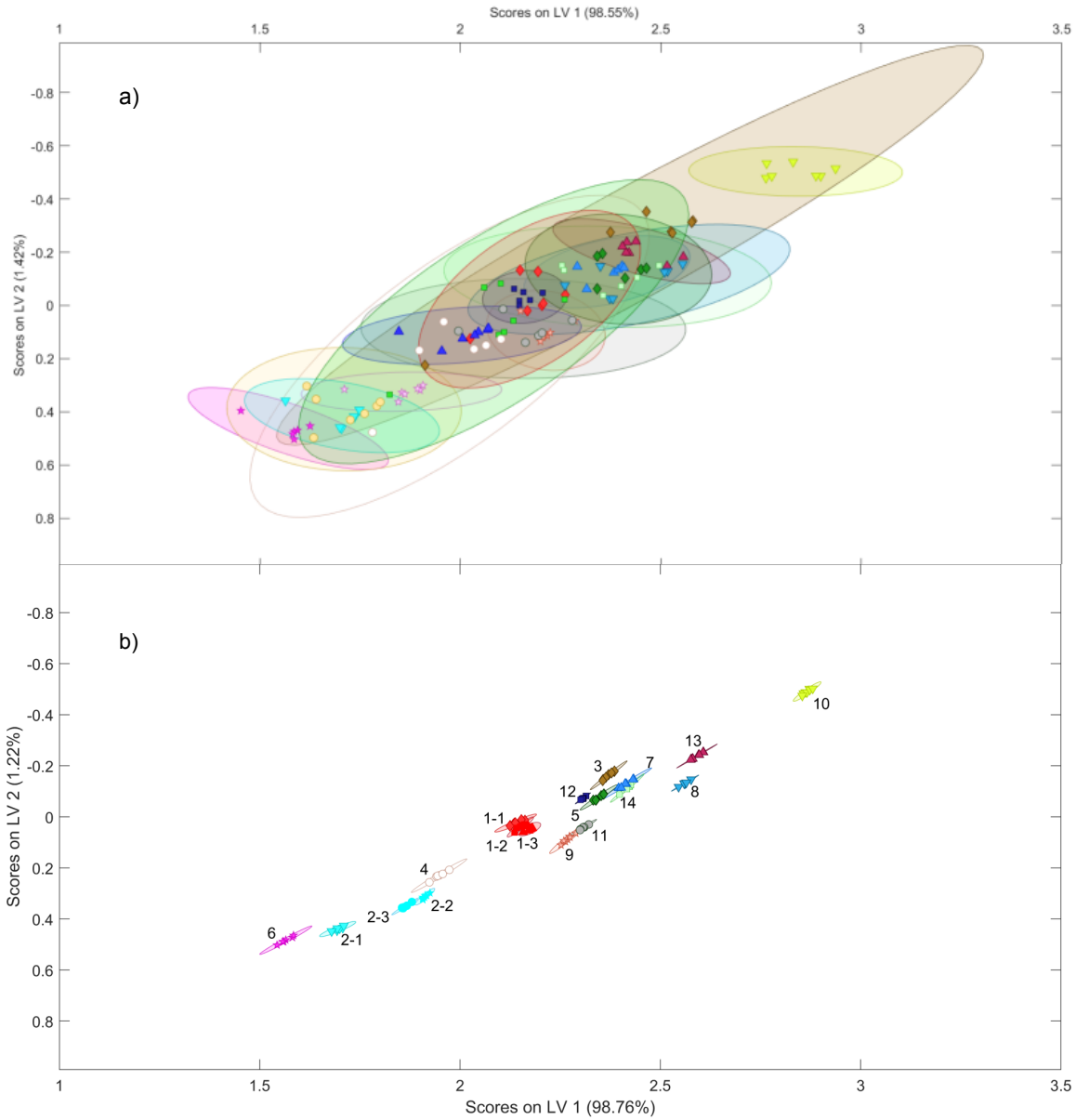

a)
Scores on LV 1 (98.55%)
Scores on LV 2 (1.42%)
b)
Scores on LV 2 (1.22%)
Scores on LV 1 (98.76%)
1
1-1
1-2
1-3
2-1
2-2
2-3
3
4
5
6
7
8
9
10
11
12
13
14

Fig. 4

Supp. Fig. 1. Schematic representation of short-term drift within a sequence.

1.- Getting raw data from e-nose

			Sensors response							
Injection nº	Sample	Injection time (h)	LY2/LG	LY2/G	LY2/AA	LY2/Gh	LY2/gCTI	LY2/gCT	T30/1	...
1	TomSST 2	0,00	0,667	-2,17	-2,06	-2,03	-2,03	-0,905	0,817	...
2	3	0,35	0,146	-0,477	-0,554	-0,384	-0,295	-0,113	0,649	...
3	6	0,70	0,106	-0348	-0,428	-0,273	-0,203	-0,079	0,622	...
4	2_1	1,03	0,392	-0,528	-0,64	-0,403	-0,301	-0,104	0,645	...
...	...	...	...	...	...	...	...	...	...	...

2.- Appliying multiplicative pretreatment for each sample class

$$S = \left(\frac{S_{i_{(j)},k\ \text{measured}}}{\hat{S}_{i,k\ \text{measured}}}\right)$$

Eg.: $$LY2/LG_{\ \text{TomSST2 injection 1}} = \left(\frac{0{,}667}{\hat{S}_{TomSST2\ class}}\right) = \left(\frac{0{,}667}{0{,}643}\right)$$

3.- Performing PLS regression to model intra-sequence drift (in PLSToolbox for Matlab)

Y = Injection time (h) (n x 1 vector)

X = Pretreated sensor response matrix (n samples x 18 sensor matrix)

(Outliers removed from the model based on Q residuals and T2 Hotelling statistics if necessary)

4.- Obtaining intra-sequence drift (in Matlab)

Obtaining scores and loadings matrices from de PLS model (extracting data from de model as raw data)

Scores (matrix of n samples x number of latent variables selected)

Loadings (matrix of 18 sensors x number of latent variables selected)

Intra-sequence **drift** matrix (E_{drift})= scores * loadings'

(matrix product in Matlab command window giving a n samples x 18 sensors matrix)

5.- Substracting intra-sequence drift from raw data signal (in Matlab or Excel)

$$S_{i_{(j)},k\ \text{corrected}} = [\hat{S}_{i,k\ \text{measured}}(1 - E_{i_{(j)},k\ \text{drift}}(t))] + S_{i_{(j)},k\ \text{measured}}$$

Eg.: $$LY2/LG_{\ \text{TomSST2 injection 1 corrected}} = 0{,}643\ (1 - E_{\text{drift } 1,1}) + 0{,}667 = 0{,}6419$$

Supp. Fig. 2. Schematic representation of long-term drift with several sequences.

1.- Getting raw data from e-nose for each sequence

Seq matrix 1 | Seq matrix 2 | ... | Seq matrix m

2.- Performing intra-sequence drift correction for each sequence data matrix

Seq matrix 1 intra-seq corrected | Seq matrix 2 intra-seq corrected | ... | Seq matrix m intra-seq corrected

3.- Appliying multiplicative pretreatment for each sample class

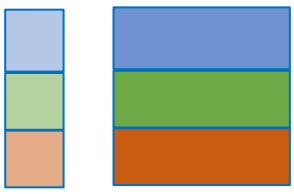
$$\frac{S_{i_{(j)},k\,measured}}{\hat{S}_{i,k\,measured}+(\hat{S}_{TomSSt1,k}-\hat{S}_{TomSStn,k})}$$

Eg.: $$LY2/LG_{\text{ TomSST2 injection 1 seq 2}} = \left(\frac{S_{1,1\,TomSST2}}{\hat{S}_{TomSST2\,class\,seq\,1}-S_{TomSST2\,class\,seq\,2}}\right) = \left(\frac{0{,}652}{0{,}643-0{,}627}\right)$$

Seq matrix 1 intra-seq corrected | Seq matrix 2 intra-seq corrected + multiplicative pretreated | ... | Seq matrix m intra-seq corrected + multiplicative pretreated

4.- Performing PLS regression to model inter-sequence drift with all corrected matrices joined (in PLSToolbox for Matlab)

Y vector
(injection time in hous; continuous for the whole trial)



X matrix
(sensor signals)

5.- Obtaining inter-sequence drift (in Matlab)

6.- Substracting intra-sequence drift from raw data signal (in Matlab or Excel)

Similar as described for intra-sequence drift correction

Graphical abstract