Increasing the robustness of SIFT-MS volatilome fingerprinting by introducing notional analyte concentrations

Amina Benchennouf[1, ‡], Matthias Corion[2, ‡], Angelica Dizon[1, ‡], Yijie Zhao[3, ‡], Jeroen Lammertyn[2], Barbara De Coninck[3], Bart Nicolaï[1, 4], Joeri Vercammen[5], Maarten Hertog[1,*]

[1] KU Leuven, BIOSYST-MeBioS Postharvest group, Willem de Croylaan 42, B-3001 Leuven, Belgium

[2] KU Leuven, BIOSYST-MeBioS Biosensors group, Willem de Croylaan 42, B-3001 Leuven, Belgium

[3] KU Leuven, BIOSYST-Crop Biotechnics, Willem de Croylaan 42, B-3001 Leuven, Belgium

[4] Flanders Centre of Postharvest Technology, Willem de Croylaan 42, B-3001 Leuven, Belgium

[5] UGent, Department of Materials, Textiles and Chemical Engineering, Technologiepark Zwijnaarde 125 B-9052 Zwijnaarde, Belgium

ABSTRACT: Selected-ion-flow-tube-mass-spectrometry (SIFT-MS) is an analytical technique for volatile detection and quantification. SIFT-MS can be applied in a 'white box' approach, measuring concentrations of target compounds, or as a 'black box' fingerprinting technique, scanning all product ions during a full scan. Combining SIFT-MS full scan data acquired from multi-batches or large-scale experiments remains problematic due to signal fluctuation over time. The standard approach of normalizing full scan data to total signal intensity was insufficient. This study proposes a new approach to correct SIFT-MS fingerprinting data. In this concept, all the product ions from a full scan are considered individual compounds for which notional concentrations can be calculated. Converting ion count rates into notional analyte concentrations accounts for any changes in instrument parameters. The benefits of the proposed approach are demonstrated on three years of data from both multi-batches and long-term experiments showing a significant reduction of system-induced fluctuations providing a better focus on the changes of interest.

INTRODUCTION SECTION

Selected ion flow tube mass spectrometry (SIFT-MS) is a relatively underexploited analytical technique for real-time analysis of volatile organic compounds (VOCs) and inorganic gases and has been extensively reviewed elsewhere[1–6].

In short, SIFT-MS relies on soft chemical ionization of sample molecules by multiple reagent ions (such as: H_3O^+, NO^+ and O_2^+) formed by microwave discharge of moist air. Inside a flow tube, these reagent ions react consecutively with analyte molecules to form characteristic product ions. All ions, i.e. both analyte cations and unreacted reagent ions, enter the downstream mass spectrometer where they are filtered based on their m/z ratio before reaching the detector. Ion count rates (in s^{-1}) are determined using a channel electron multiplier, which results in a mass spectrum of all product ions of interest. SIFT-MS quantifies components based on the ratio of product ion count rates to reagent ion count rates taking into account the reaction chemistry between reagent ions and the analytes. SIFT-MS instruments rely on a built-in reaction kinetics library, containing information on target analytes with their reagent ion specific product ions and associated kinetic reaction rate constants. Furthermore, ion count rates are corrected for flow tube transmission efficiency as a function of product ion molecular weight, through the instrument calibration function (ICF), which is determined by frequent analysis of a certified gas standard (the calibrant gas). The concept of ICF is approximate and not rigorous, as it combines diffusion losses (dependent on the geometry and cross-section of the ions) with mass discrimination[7].

SIFT-MS is being used in a wide range of application areas such as biomedical applications[8], food science[9], environment (i.e., the emission of anthropogenic[10] or biogenic[11] VOCs), and microbiology[12]. Within the different fields, it is applied for both targeted analysis (e.g., looking for biomarkers[13,14]) and fingerprinting applications (e.g., checking for food adulteration[15,16]). The first mode, also referred to as multiple ion monitoring (MIM) mode, is most widely used because it allows for direct quantification of compounds in real-time using prior knowledge on sample composition, library tabulated reaction schemes and associated kinetic reaction rate constants, expressing the analyte concentrations as ppm values.

In case prior knowledge on target analytes is not available, the analysis has to be executed in full scan mode. The main aim of such full scan is not to identify individual compounds, but to profile the overall composition of the volatilome for classification purposes. Generally, for each of the reagent ions, a wide range of m/z values is measured, e.g. from 15 m/z to 300 m/z, and the product ion responses are expressed as count rates without being converted into analyte concentrations.

Like any mass spectrometric technique[17,18], SIFT-MS is not sufficiently stable when involved in large-scale untargeted fingerprinting studies. During such studies, samples are analyzed in different batches and then combined for data processing and statistical analysis. Systematic and random

variation between the measurement blocks over time hinders statistical analysis, leading to information loss. With LC-MS or GC-MS different strategies can help in correcting variations of signal intensity[17,19]. The simplest normalization approach is based on total intensity[20]. Some of the other techniques like the use of internal standards[21], or quality control samples[21,22] are difficult to implement for SIFT-MS. In the case of internal standards because of the lack of chromatographic separation and in the case of quality control samples because of working with intact perishable biological samples. Typically, data from large-scale SIFT-MS-based experiments is normalized into relative count rates by dividing each m/z abundance by the sum of the signal intensities per specific reagent ion[23,24]. We did notice, however, that this approach is not sufficient to address the prevailing variations, often resulting in inconsistent product to reagent ion ratios.

The most likely source of signal variations in SIFT-MS relates to inconsistent variation in product ion intensity levels induced by changes in instrument parameters due to maintenance and drift over time. Moreover, low-abundant product ions become undetectable when instrument sensitivity is compromised. Lehnert *et al*. illustrated the effect of certain instrument parameters, including carrier gas flow rate, flow tube voltage and sample flow rate, on product ion intensities[25]. These sources of variation are included when calculating absolute concentrations from the raw count data making the MIM data inherently stable and reproducible between instruments. This is in contrast to full scan data.

To the best of our knowledge, there are no available strategies to tackle the problem of signal drift over time or across multiple batches in large-scale SIFT-MS full scan studies used for fingerprinting. This manuscript introduces a new approach to correct for intensity drift in the data from large-scale SIFT-MS-based fingerprinting studies. The strategy we applied is to treat the untargeted full scan data similarly to the targeted MIM data. Each individual product ion is considered to be a single product of a reaction of the reagent ions with a notional compound and by affixing arbitrary kinetic reaction rate coefficients (*k*), notional analyte concentration (NAC) values were calculated (in ppm). This approach accounts for any unforeseen fluctuation in instrument parameters such as tube pressure, reaction time, and carrier gas flow, next to changes in the levels of generated ions.

A similar approach was applied by Granitto *et al*. in their study comparing full scan data of different commercial proton-transfer-reaction mass spectrometers (PTR-MS) also converting the count rates to concentration values[26]. However, in their study, the authors did not elaborate on the specific advantages of applying this correction, its generally positive impact on full-scan data approaches and its applicability to SIFT-MS.

Our approach was applied to a large data set consisting of repeated measurements of identical synthetic samples (calibrant gas, blank samples, filtered air) over a long time. Furthermore, the performance of the NAC approach was evaluated on original sets of biological data from two multi-batch experiments. The method was successfully used to remedy the negative effect of signal drift on data analysis compared to the normalization approach based on the total ion intensity.

EXPERIMENTAL SECTION

Methodological testing

Instrumentation

A Voice200 *ultra*® SIFT-MS (Syft™ Technologies, Christchurch, New Zealand) was used with helium as carrier gas at a flow rate of 316 mL/min (4.0 Torr L/s). Flow tube pressure was set to 80.0 Pa (0.6 Torr). All samples were introduced to the flow tube via a heated inlet (120 °C) at a set stable flow rate of 24 mL/min (at normal temperature, 293 K, and pressure, 101 kPa) which was checked routinely.

Before analysis, instrument validation tests were run as part of the instrument startup procedure. Besides verification of parameters like temperature, pressure, and quadrupole performance, the procedure also included a validation step in which a certified gas standard was measured. This calibrant gas consisted of a mixture of ethylene, isobutene, benzene, toluene, tetrafluorobenzene, hexafluorobenzene and octafluorotoluene in nitrogen all at 2.0 ppm ± 5 % by volume (Air Liquide America Specialty gases LLC). Labsyft 1.6.2 software (Syft™ Technologies) was used for data acquisition.

Signal intensity and stability

Signal intensities were determined daily in terms of the amount of product ions generated (in s^{-1}) for a given concentration of the corresponding analyte in the calibrant standard gas during instrument validation. To evaluate signal stability over time, signal intensities were collated over a three-year period (December 2018 to August 2021).

The validation was performed in MIM mode; after 20 s of settle time, 7 to 8 iteration cycles were run with 50 ms dwell time per product ion. Average ICF corrected ion count rates over all sampling cycles were calculated for each product ion of the target analytes from the calibrant gas (see supplementary section S1). Per reagent ion, the total ion count rate was calculated as the sum of all ICF corrected product ion count rates, including unreacted reagent ions.

To obtain relative product ion count rates for a specific reagent ion, the ICF corrected product ion count rates were normalized to the total ion count rates. In addition, ICF corrected count rates of individual m/z traces were converted to NACs as if they were single reaction products of notional compounds. The approach was implemented by introducing new 'compound entries' into the database, one for each m/z from 15 to 300 m/z, assigning an arbitrary reaction rate coefficient (*k*) of $1 \cdot 10^{-9}$ cm^3 s^{-1} (see supplementary data section S2). Subsequently, Labsyft was used to perform concentration calculations identical to a target analysis. The compound concentration was calculated as follows:

$$[A] = \frac{[P^+]\, ICF_p}{k\, t_r\, [R^+]\, ICF_R}$$

where [A] is the concentration of the analyte in gas phase, $[P^+]$ is the concentration of the protonated analyte at time t, *k* is the kinetic rate constant for the reaction of reagent ion with the analyte, t_r is the reaction time of the reagent ion with the analyte, $[R^+]$ is the concentration of the unreacted reagent ions

and *ICF* is the instrument calibration function for either the product or reagent ion.

$[P^+]$, $[R^+]$ and t_r are measured values, while *k* is used from the Syft™ compound library. t_r is the average time the ions spend in the flow tube, and is experimentally determined during the validation procedure of the machine using the certified standard's ethene component. The reaction of ethene (C_2H_4) with O_2^+ gives a single product ion P+ at m/z = 28. It is assumed that between m/z 28 to 32 there is no significant change in ion transmission, setting both the ICF_p for m/z 28 and the ICF_R for m/z 32 to 1, thus enabling the calculation of t_r from the equation above.

To investigate the influence of instrument parameters on the relative count rate data, a main factor least squares regression model was fit to the calibrant gas data. The main instrument parameters, being the sum of reagent ions, the carrier flow rate, the flow tube pressure and temperature, the downstream and upstream intensity, and the reaction time were extracted from the data and normalized between -1 and +1. The count rates obtained during multiple sampling cycles were normalized for the reagent ion abundance and were averaged over the sample cycles calculating a relative count rate sum. A standard least squares regression was performed including the main effects only, removing any non-significant effects. Finally, the instrument parameters were checked for multicollinearity using the variance inflation factor[27] applying a cut-off value of 5 to consider removal from the model. This analysis was performed in JMP®Pro 15.1.0 (SAS Institute Inc., Cary, NC, USA).

Case studies

Three case studies were performed on either (1) background and blank samples, (2) ripening strawberry fruit, or (3) microbially infected pear fruit. For the subsequent case studies, intact fruit were incubated in jars while juice samples were incubated in 20 mL vials, each equipped with two outlets. During headspace sampling, one outlet was connected to the heated inlet of the instrument while the other one was connected to a 5 L Tedlar bag with polypropylene fitting (MediSense, Netherlands) and filled with filtered air (see next section) to compensate for sampling-induced pressure changes. Blank samples were included using empty jars and vials flushed with filtered air for 60 s or 10 s, respectively. Prior to analysis, sample volatiles were allowed to accumulate in the headspace for different periods of time, depending on the sample type being analyzed. Further details can be found in the supplementary material section S6.

RESULTS & DISCUSSION

Methodological testing

Reagent ion intensity showed large changes over time without affecting absolute quantification

To evaluate the signal stability of H_3O^+, NO^+ and O_2^+ reagent ions, their respective total ion count rates were collated from the calibrant gas data between December 2018 and August 2021 (Figure 1A). Within this time frame, several distinct interferences were observed, due to routine instrument maintenance.

The total ion count rate for each reagent ion showed considerable variation over time, with the highest count rates observed directly after instrument maintenance, followed by a progressive decline. Apart from a possible gradual detector sensitivity loss due to aging, the fast drops in signal intensities after maintenance are not yet fully understood but are most probably due to a combination of factors (e.g., clean glassware, clean lenses, pump capacity, etc.).

Similarly, the signal intensity of selected calibrant gas product ions was plotted as a function of time (see supplementary data section S3 for all graphs). As shown, for example, for m/z 93 (H_3O^+), which corresponds to toluene (Figure 1B), a pattern similar to total ion count rate was observed. As ICF corrected product ion count rates are determined by the amount of reagent ions introduced into the flow tube, this is expected. Despite these variations, the calculated concentration values of 2 ppm remained stable for all the compounds as they should (Figure 1C and supplementary data section S4). As this absolute quantification of the calibrant gas is based on the ratio of the analyte product ion count rates to the reagent ion count rates involving the various instrument parameters (e.g., flow tube pressure, carrier gas flow rate, etc.), these calculated concentrations are not affected in their accuracy. Still, the drastic drop in overall signal intensity did lead to a loss of some product ions (data not shown).

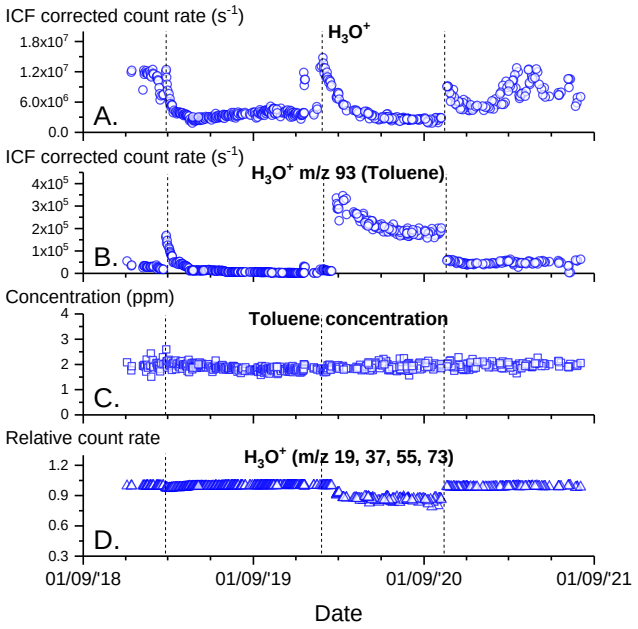

Figure 1. Results from methodological testing. **A**-**D** consider system validations performed from December 2018 to August 2021. **A**: Quantitative assessment of the sum of the unreacted reagent ions and generated product ions during the validation measurements taking H_3O^+ as an example. **B** & **C**: Intensity evolution of (**B**) the ion m/z 93 (in s^{-1}) and (**C**) the toluene concentration (in ppm). **D**: Relative count rates of the unreacted reagent ions and their water clusters during the system validation taking H_3O^+ as an example. The dashed lines indicate maintenance events.

Alongside the fluctuations observed for the total ion count rates of the reagent ions, variations were also observed in the ion count rates of the product ions. Looking into the blank control jars, it was noticed that the number of product ions for which the count rate exceeded a noise level of 200 s^{-1} changed in parallel to the total ion count rates of the reagent ions in response to maintenance.

Relative count rates are not sufficient to correct for fluctuations in signal intensities

Given the considerable changes in signal intensities after each maintenance, the effectiveness of the relative ion count signal approach was checked in more detail. Therefore, data were collated from the instrument validation files from December 2018 to August 2021. The results reveal an unexpected variation (Figure 1D). Typically, the summed relative count rates of the reagent ions remain around 98 to 99 %. However, the maintenance of February 2020 resulted in a significant deviation, indicating a reagent ion depletion with over 10 % bringing us potentially outside the linear range of the instrument[3].

In addition, the relative count rates of the compounds from the calibrant gas were calculated and plotted as a function of time, taking toluene as an example (Figure 2). Given that the concentrations of the compounds from the certified calibrant gas are by definition constant (2 ppm), assuming constant instrument parameters, the relative count rates were expected to be reproducible over time. As can be seen, the relative count rates still show four regions (Figure 2); the boundary of each region being marked by instrument maintenance. Similar to the reagent ion plots in Figure 1D, the variation was most pronounced after the maintenance of February 2020, showing a six-fold increase in the relative count rates of all product ions, presumably caused by changes in instrument parameters like flow tube pressure, flow tube temperature and carrier gas flow rate. Although these changes did not affect the accuracy of absolute quantification (Figure 1C, and supplementary data section S4), the relative count rates were significantly impacted (Figure 2).

From this, it can be concluded that relative count rates are not sufficient to correct for severe fluctuations in signal intensities. For long-term experiments, consistent count rates are, however, imperative.

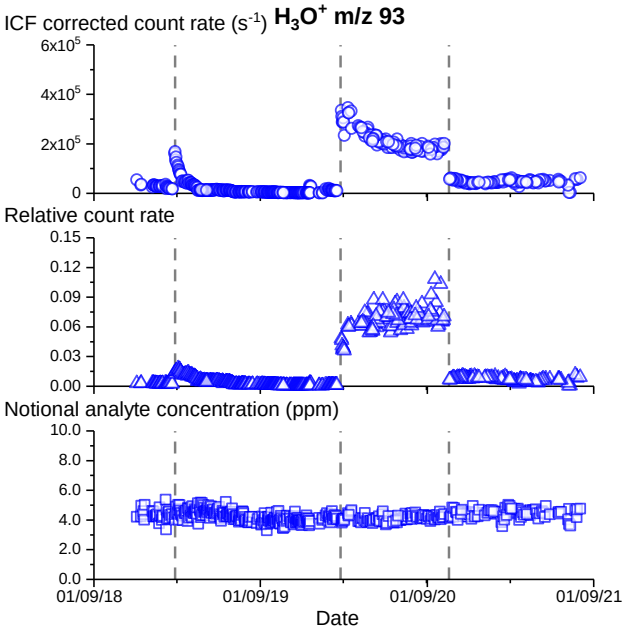

Figure 2. Comparison of ICF corrected count rate (○), relative count rate (Δ) and notional analyte concentration (□) of the exemplary product ion m/z 93 reacted with H_3O^+ generated during the reaction with toluene. Measurements were taken from SIFT standard validation data files from December 2018 to August 2021. The dashed lines indicate maintenance events.

Fluctuations in count rates can be linked to specific instrument parameters

To identify which instrument parameters impact data output, a linear main effects regression model was fitted to the relative count rates of the calibrant gas data (see supplementary data section S5). Main factors considered were sum of reagent ions, carrier flow rate, flow tube pressure and temperature, downstream and upstream intensity, and reaction time. After reducing the full model to the significant factors only, an R^2 adjusted was obtained of 0.975, with the main effects being reaction time ($p < 0.0001$) and reagent ion sum ($p = 0.03$). In this model, multicollinearity was reduced to a minimum, with the variance inflation factors of both instrument parameters remaining below five. The significance of the reagent ion sum is unexpected since relative count rates are already normalized for the reagent ion sum. Clearly, mere correction for changes in the sum of reagent ions does not provide a robust enough data correction approach.

Table 1. Statistical performance of relative count rate data and notional analyte concentration of calibrant gas' product ions.

Reagent ion	Product ion	Relative count rate	NAC	CV Relative data/ CV NAC
		CV (%)	CV (%)	
H_3O^+	79	146	7.7	18.90
H_3O^+	93	142	7.7	18.19

NO^+	57	145	6.1	23.91
NO^+	92	146	5.8	24.71
O_2^+	28	135	7.5	17.56
O_2^+	78	134	5.9	22.73
O_2^+	92	133	11.6	11.47
O_2^+	150	134	6.5	20.56
O_2^+	186	135	6.9	19.48
O_2^+	236	135	8.2	16.45

Statistical data is presented in terms of their coefficient of variation (CV), with the ratio of the two coefficients of variation indicating the reduction in variation thanks to the introduction of the concept of notional analyte concentration (NAC).

Converting count rates into NACs provides stable results

Using data from the validation measurements from December 2018 to August 2021, NACs were calculated for the product ions of the calibrant gas. The rate coefficient used, $1 \cdot 10^{-9}$ cm^3 s^{-1}, was in the range between $1 \cdot 10^{-13}$ and $6 \cdot 10^{-9}$ cm^3 s^{-1}, which is the range observed in the Labsyft library for real compounds. The value assigned to K can be any value within the range of rate coefficients, as the aim is not to calculate exact compound concentrations but only to account for all possible factors affecting the ICF corrected count rates.

Taking toluene as an example, relative counts normalization was not able to fully correct for the fluctuations in the data set (Figure 2). Using the NAC approach, however, the data set became stable over time. Similar results were observed for product ions of all seven compounds from the certified standard gas (see supplementary material section S3). The results are summarized in Table 1 showing that the coefficient of variation using relative count rates was 16 to 25 times higher compared to NACs, confirming suitability of the latter for long-term monitoring applications.

NACs fully account for the impacts of instrument parameters

Similar to the regression analysis of the relative count rates, a linear main effects model was fitted to the sum of the NACs of the calibrant product ions (see supplementary data in section S5). It was hypothesized that the instrument parameters would have a lower or even non-significant effect on the NACs. This was confirmed and resulted in a model with an R^2 adjusted of only 0.06, confirming that the full scan data correction approach was able to account for the instrumental fluctuation in the ICF corrected full scan data.

Case studies

The previous section showed that the NAC approach can correct for fluctuations in the full scan data over long periods and provide stable values for selected ions. Next, practical case studies are used to verify the feasibility of the NAC approach (see supplementary material section S6). The first case study focuses on background samples and blank samples to check the stability of the instrument and to further assess the benefits of using NACs. The additional two case studies cover multi-batch experiments on fruits (strawberry and pear) to verify the effectiveness of this NAC method in practice.

In each of the case studies it was shown systematic time-dependent influences exist resulting from maintenance. In all cases, relative count correction was not able to structurally address these impacts. Full scan data expressed as NACs were shown to be stable and able to successfully eliminate the impact of artifacts due to unwanted instrument fluctuations on SIFT-MS full scan data.

CONCLUSION

In this study, a novel approach for SIFT-MS full scan data correction was presented, as the classical approach of using relative data was demonstrated not to suffice. Using an arbitrarily fixed kinetic reaction rate coefficient, the count rates of all product ions were transformed to NACs. In this way, the impact of system variations was effectively removed from the full scan data. The applicability of this approach was demonstrated using three years of data, containing long term fluctuations in reagent ion intensities and other instrument parameters affecting product ion abundances. Our NAC approach resulted in far lower coefficients of variation removing drift and the impact of discrete maintenance events from the full scan data. As was shown for actual case studies, interpretation of the experimental data was dramatically improved, providing focus on differences that mattered.

The proposed NAC approach is crucial to any long-term fingerprinting screening experiments where historical data needs to be compared, especially in the case of a decision support system. As long as SIFT-MS is applied in the context of a targeted white box approach, the results are already intrinsically stable over time. Finally, the proposed approach is also transferable to PTR-MS instruments. Here, the option exists in the data preprocessing software to create a custom library with nominal masses that are considered compounds with a fixed reaction rate. When exporting the spectra, NACs can be calculated with the same benefits as reported in this work.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the JASMS Publications website.

A single document containing supplementary section S1-S6 (PDF)
S1. Composition calibrant gas
S2. Addition of m/z based compounds in LabSyft for calculating notional analyte concentrations (NACs)
S3. Comparison of ICF corrected count rate, relative count rate and ‘dummy’ concentration of the seven remaining product ions generated from reaction of calibrant gas with the reagent ions.
S4. Absolute concentration of the six other compounds comprising the calibrant gas
S5. Linear main effects models on calibrant gas data
S6. Case studies

AUTHOR INFORMATION

Corresponding Author

* Maarten Hertog, maarten.hertog@kuleuven.be

Author Contributions

‡These authors contributed equally.
AB, **MC**, **AD** and **YZ**: Methodology, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. **JL** and **BDC**: Funding acquisition, Writing - Review & Editing. **BN**: Resources, Funding acquisition, Writing - Review & Editing. **JV**: Conceptualization, Writing - Review & Editing. **MH**: Conceptualization, Supervision, Writing - Review & Editing.

Notes

The authors declare no competing financial interest.
The data that support the findings of this study are available from the corresponding author upon request.

ACKNOWLEDGMENT

The authors acknowledge the valuable discussions with support staff from Interscience BENELUX. **AB** acknowledges research funding through the VLAIO (HBC.2017.0831); **MC** has received a scholarship from the Research Foundation – Flanders [FWO-SB project 1SC7219N]; **YZ** has been supported by the China Scholarship Council at KU Leuven (No. 201706990031). **BDC** and **BN** acknowledges financial support from the BelOrta chair (ITP-LSBEL1-O2010).

REFERENCES

(1) Smith, D.; McEwan, M. J.; Španěl, P. Understanding Gas Phase Ion Chemistry Is the Key to Reliable Selected Ion Flow Tube-Mass Spectrometry Analyses. *Anal. Chem.* **2020**, *92* (19), 12750–12762. https://doi.org/10.1021/acs.analchem.0c03050.

(2) Smith, D.; Spanel, P. The Novel Selected-Ion Flow Tube Approach to Trace Gas Analysis of Air and Breath. *Rapid Commun. Mass Spectrom.* **1996**, *10* (10), 1183–1198. https://doi.org/10.1002/(SICI)1097-0231(19960731)10:10<1183::AID-RCM641>3.0.CO;2-3.

(3) Smith, D.; Španěl, P. Ambient Analysis of Trace Compounds in Gaseous Media by SIFT-MS. *Analyst*. 2011, pp 2009–2032. https://doi.org/10.1039/c1an15082k.

(4) Belluomo, I.; Boshier, P. R.; Myridakis, A.; Vadhwana, B.; Markar, S. R.; Spanel, P.; Hanna, G. B. Selected Ion Flow Tube Mass Spectrometry for Targeted Analysis of Volatile Organic Compounds in Human Breath. *Nat. Protoc.* **2021**, *16* (7), 3419–3438. https://doi.org/10.1038/s41596-021-00542-0.

(5) Slingers, G.; Vanden Eede, M.; Lindekens, J.; Spruyt, M.; Goelen, E.; Raes, M.; Koppen, G. Real-Time versus Thermal Desorption Selected Ion Flow Tube Mass Spectrometry for Quantification of Breath Volatiles. *Rapid Commun. Mass Spectrom.* **2021**, *35* (4), 1–9. https://doi.org/10.1002/RCM.8994.

(6) Smith, D.; Španěl, P.; Demarais, N.; Langford, V. S.; McEwan, M. J. Recent Developments and Applications of Selected Ion Flow Tube Mass Spectrometry (SIFT-MS). *Mass Spectrom. Rev.* **2023**, No. August 2022, 1–34. https://doi.org/10.1002/mas.21835.

(7) Španěl, P.; Dryahina, K.; Smith, D. A General Method for the Calculation of Absolute Trace Gas Concentrations in Air and Breath from Selected Ion Flow Tube Mass Spectrometry Data. *Int. J. Mass Spectrom.* **2006**, *249–250*, 230–239. https://doi.org/10.1016/j.ijms.2005.12.024.

(8) Casas-Ferreira, A. M.; Nogal-Sánchez, M. del; Pérez-Pavón, J. L.; Moreno-Cordero, B. Non-Separative Mass Spectrometry Methods for Non-Invasive Medical Diagnostics Based on Volatile Organic Compounds: A Review. *Anal. Chim. Acta* **2019**, *1045*, 10–22. https://doi.org/10.1016/j.aca.2018.07.005.

(9) Langford, V. S.; Padayachee, D.; McEwan, M. J.; Barringer, S. A. Comprehensive Odorant Analysis for On-line Applications Using Selected Ion Flow Tube Mass Spectrometry (SIFT-MS). *Flavour Fragr. J.* **2019**, *34*, 393–410. https://doi.org/10.1002/ffj.3516.

(10) Smith, D.; Spanel, P. The Novel Selected-Ion Flow Tube Approach to Trace Gas Analysis of Air and Breath. *Rapid Commun. Mass Spectrom.* **1996**, *10* (10), 1183–1198. https://doi.org/10.1002/(SICI)1097-0231(19960731)10:10<1183::AID-RCM641>3.0.CO;2-3.

(11) Biasioli, F.; Yeretzian, C.; Märk, T. D.; Dewulf, J.; Van Langenhove, H. Direct-Injection Mass Spectrometry Adds the Time Dimension to (B)VOC Analysis. *TrAC - Trends Anal. Chem.* **2011**, *30* (7), 1003–1017. https://doi.org/10.1016/j.trac.2011.04.005.

(12) Wang, Y.; Li, Y.; Yang, J.; Ruan, J.; Sun, C. Microbial Volatile Organic Compounds and Their Application in Microorganism Identification in Foodstuff. *TrAC - Trends Anal. Chem.* **2016**, *78*, 1–16. https://doi.org/10.1016/j.trac.2015.08.010.

(13) Ioannidis, A. G.; Walgraeve, C.; Vanderroost, M.; Van Langenhove, H.; Devlieghere, F.; De Meulenaer, B. Non-Destructive Measurement of Volatile Organic Compounds in Modified Atmosphere Packaged Poultry Using SPME-SIFT-MS in Tandem with Headspace TD-GC-MS. *Food Anal. Methods* **2018**, *11* (3), 848–861. https://doi.org/10.1007/s12161-017-1061-5.

(14) Španěl, P.; Smith, D. Quantification of Volatile Metabolites in Exhaled Breath by Selected Ion Flow Tube Mass Spectrometry, SIFT-MS. *Clin. Mass Spectrom.* **2020**, *16*, 18–24. https://doi.org/10.1016/j.clinms.2020.02.001.

(15) Bajoub, A.; Medina-Rodríguez, S.; Ajal, E. A.; Cuadros-Rodríguez, L.; Monasterio, R. P.; Vercammen, J.; Fernández-Gutiérrez, A.; Carrasco-Pancorbo, A. A Metabolic Fingerprinting Approach Based on Selected Ion Flow Tube Mass Spectrometry (SIFT-MS) and Chemometrics: A Reliable Tool for Mediterranean Origin-Labeled Olive Oils Authentication. *Food Res. Int.* **2018**, *106* (December 2017), 233–242. https://doi.org/10.1016/j.foodres.2017.12.027.

(16) Kharbach, M.; Kamal, R.; Mansouri, M. A.; Marmouzi, I.; Viaene, J.; Cherrah, Y.; Alaoui, K.; Vercammen, J.; Bouklouze, A.; Vander Heyden, Y. Selected-Ion Flow-Tube Mass-Spectrometry (SIFT-MS) Fingerprinting versus Chemical Profiling for Geographic Traceability of Moroccan Argan Oils. *Food Chem.* **2018**, *263* (December 2017), 8–17. https://doi.org/10.1016/j.foodchem.2018.04.059.

(17) Ejigu, B. A.; Valkenborg, D.; Baggerman, G.; Vanaerschot, M.; Witters, E.; Dujardin, J.-C.; Burzykowski, T.; Berg, M. Evaluation of Normalization Methods to Pave the Way Towards Large-Scale LC-MS-Based Metabolomics Profiling Experiments. *Omi. A J. Integr. Biol.* **2013**, *17* (9), 473–485. https://doi.org/10.1089/omi.2013.0010.

(18) Fernández-Albert, F.; Llorach, R.; Garcia-Aloy, M.; Ziyatdinov, A.; Andres-Lacueva, C.; Perera, A. Intensity Drift Removal in LC/MS Metabolomics by Common Variance Compensation. *Bioinformatics* **2014**, *30* (20), 2899–2905. https://doi.org/10.1093/bioinformatics/btu423.

(19) Chen, M.; Rao, R. S. P.; Zhang, Y.; Zhong, C. X.; Thelen, J. J. A Modified Data Normalization Method for GC-MS-Based Metabolomics to Minimize Batch Variation. *Springerplus* **2014**, *3* (1), 439. https://doi.org/10.1186/2193-1801-3-439.

(20) Brunius, C.; Shi, L.; Landberg, R. Large-Scale Untargeted LC-MS Metabolomics Data Correction Using between-Batch Feature Alignment and Cluster-Based within-Batch Signal Intensity Drift Correction. *Metabolomics* **2016**, *12* (11), 173. https://doi.org/10.1007/s11306-016-1124-4.

(21) Broadhurst, D.; Goodacre, R.; Reinke, S. N.; Kuligowski, J.; Wilson, I. D.; Lewis, M. R.; Dunn, W. B. Guidelines and Considerations for the Use of System Suitability and Quality Control Samples in Mass Spectrometry Assays Applied in Untargeted Clinical Metabolomic Studies. *Metabolomics* **2018**, *14* (6), 72. https://doi.org/10.1007/s11306-018-1367-3.

(22) Dudzik, D.; Barbas-Bernardos, C.; García, A.; Barbas, C. Quality Assurance Procedures for Mass Spectrometry Untargeted Metabolomics. a Review. *J. Pharm. Biomed. Anal.* **2018**, *147*, 149–173. https://doi.org/10.1016/j.jpba.2017.07.044.

(23) Henderson, B.; Slingers, G.; Pedrotti, M.; Pugliese, G.; Malásková, M.; Bryant, L.; Lomonaco, T.; Ghimenti, S.; Moreno, S.; Cordell, R.; Harren, F. J. M.; Schubert, J.; Mayhew, C. A.; Wilde, M.; Di Francesco, F.; Koppen, G.; Beauchamp, J. D.; Cristescu, S. M. The Peppermint Breath Test Benchmark for

PTR-MS and SIFT-MS. *J. Breath Res.* **2021**, *15* (4), 046005. https://doi.org/10.1088/1752-7163/ac1fcf.

(24) Wang, M. H.; Yuk-Fai Lau, S.; Chong, K. C.; Kwok, C.; Lai, M.; Chung, A. H. Y.; Ho, C. S.; Szeto, C. C.; Chung-Ying Zee, B. Estimation of Clinical Parameters of Chronic Kidney Disease by Exhaled Breath Full-Scan Mass Spectrometry Data and Iterative PCA with Intensity Screening Algorithm. *J. Breath Res.* **2017**, *11* (3), 1–8. https://doi.org/10.1088/1752-7163/aa7635.

(25) Lehnert, A. S.; Behrendt, T.; Ruecker, A.; Pohnert, G.; Trumbore, S. E. SIFT-MS Optimization for Atmospheric Trace Gas Measurements at Varying Humidity. *Atmos. Meas. Tech.* **2020**, *13* (7), 3507–3520. https://doi.org/10.5194/AMT-13-3507-2020.

(26) Granitto, P.; Biasioli, F.; Aprea, E.; Mott, D.; Furlanello, C.; Mark, T.; Gasperi, F. Rapid and Non-Destructive Identification of Strawberry Cultivars by Direct PTR-MS Headspace Analysis and Data Mining Techniques. *Sensors Actuators B Chem.* **2007**, *121* (2), 379–385. https://doi.org/10.1016/j.snb.2006.03.047.

(27) Salmerón, R.; García, C. B.; García, J. Variance Inflation Factor and Condition Number in Multiple Linear Regression. *J. Stat. Comput. Simul.* **2018**, *88* (12), 2365–2384. https://doi.org/10.1080/00949655.2018.1463376.

Table of Contents artwork