



The use of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ isotopic analyses combined with chemometrics as a traceability tool for the geographical origin of bell peppers



E. de Rijke^{a,*}, J.C. Schoorl^b, C. Cerli^b, H.B. Vonhof^c, S.J.A. Verdegaal^c, G. Vivó-Truyols^d, M. Lopatka^{e,f}, R. Dekter^g, D. Bakker^g, M.J. Sjerps^{e,f}, M. Ebskamp^g, C.G. de Koster^a

^a Mass Spectrometry of Biomacromolecules, Swammerdam Institute of Life Sciences, University of Amsterdam, Sciencepark 904, 1090 GE Amsterdam, The Netherlands

^b Earth Surface Science, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Sciencepark 904, 1090 GE Amsterdam, The Netherlands

^c Faculty of Earth and Life Sciences, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands

^d van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Sciencepark 904, 1090 GE Amsterdam, The Netherlands

^e Korteweg-de Vries Institute, University of Amsterdam, Sciencepark 904, 1090 GE Amsterdam, The Netherlands

^f Netherlands Forensic Institute, P.O. Box 24044, 2490 AA The Hague, The Netherlands

^g Naktuinbouw Laboratories, Sotaweg 25, P.O. Box 40, 2370 AA Roelofarendsveen, The Netherlands

ARTICLE INFO

Article history:

Received 24 February 2015

Received in revised form 25 January 2016

Accepted 29 January 2016

Available online 4 February 2016

Chemical compounds studied in this article:

Heptacosane (PubChem CID: 11636)

Nonacosane (PubChem CID: 12409)

Hentriacontane (PubChem CID: 12410)

Keywords:

Isotope ratio mass spectrometry

Gas chromatography

Geographic origin

Food products authenticity

Linear discriminant analysis

Likelihood ratio

Bell peppers

Capsicum annuum

ABSTRACT

Two approaches were investigated to discriminate between bell peppers of different geographic origins. Firstly, $\delta^{18}\text{O}$ fruit water and corresponding source water were analyzed and correlated to the regional GNIP (Global Network of Isotopes in Precipitation) values. The water and GNIP data showed good correlation with the pepper data, with constant isotope fractionation of about -4 . Secondly, compound-specific stable hydrogen isotope data was used for classification. Using n-alkane fingerprinting data, both linear discriminant analysis (LDA) and a likelihood-based classification, using the kernel-density smoothed data, were developed to discriminate between peppers from different origins. Both methods were evaluated using the $\delta^2\text{H}$ values and n-alkanes relative composition as variables. Misclassification rates were calculated using a Monte-Carlo 5-fold cross-validation procedure. Comparable overall classification performance was achieved, however, the two methods showed sensitivity to different samples. The combined values of $\delta^2\text{H}$ IRMS, and complimentary information regarding the relative abundance of four main alkanes in bell pepper fruit water, has proven effective for geographic origin discrimination. Evaluation of the rarity of observing particular ranges for these characteristics could be used to make quantitative assertions regarding geographic origin of bell peppers and, therefore, have a role in verifying compliance with labeling of geographical origin.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The Netherlands is the second largest exporter of agricultural products in the world (after the US) (LEI, 2013). Tomatoes, cucumbers and bell peppers (*Capsicum annuum*) are among the most important Dutch agricultural products. Within the European Union (EU), the Netherlands and Spain are the largest bell pepper producers, however, there is increasing competition from other countries

* Corresponding author.

E-mail addresses: e.derijke@uva.nl (E. de Rijke), j.c.schoorl@uva.nl (J.C. Schoorl), c.cerli@uva.nl (C. Cerli), h.b.vonhof@vu.nl (H.B. Vonhof), s.j.a.verdegaal-warmerdam@vu.nl (S.J.A. Verdegaal), g.vivotruyols@uva.nl (G. Vivó-Truyols), m.lopatka@uva.nl (M. Lopatka), r.dekter@naktuinbouw.nl (R. Dekter), d.bakker@naktuinbouw.nl (D. Bakker), m.j.sjerps@uva.nl (M.J. Sjerps), m.ebskamp@naktuinbouw.nl (M. Ebskamp), c.g.dekoster@uva.nl (C.G. de Koster).

<http://dx.doi.org/10.1016/j.foodchem.2016.01.134>

0308-8146/© 2016 Elsevier Ltd. All rights reserved.

such as Israel and Egypt. Increase in global trade and free markets has created an increased need for better regulation in the distribution chain of agricultural products. EU law on the provision of food information to consumers EEC No. 1169/2011 (2011) came into effect in December 2014. This law includes EU regulation 1337/2013 regarding the mandatory country of origin (COO) or place of provenance labeling for unprocessed meat from pigs, sheep, goats and poultry. This paves the way for new regulations regarding COO labeling of other food products. This regulatory shift has spurred the development of new techniques for the discrimination between foods of different origins, sources and farming systems. Stable isotope analysis of specific elements, including: hydrogen, carbon, nitrogen, oxygen, and sulfur, has been applied in food authentication for more than 20 years (Kelly, Heaton, & Hoogewerff, 2005; Rossmann, 2001). These methods are based on

stable isotope ratio measurements of a bulk product or a specific component such as an ingredient or target molecule. The measurements provide information on the botanical and geographical origin.

Fractionation of hydrogen and oxygen isotopes takes place during the evaporation and condensation processes of the water cycle: ocean water passing from liquid to vapor in the atmosphere undergoes strong isotope fractionation resulting in depletion of the heavier isotopes (^2H or ^{18}O) in vapor and clouds. The extent of the fractionation is dependent on the temperature, and thus is influenced by the local climate. While water is the only source of hydrogen for photosynthesis, oxygen is taken in by plants from several sources including atmospheric oxygen and carbon dioxide, and soil water. Consequently, the ^2H - and ^{18}O -contents of water from fruits should reflect the isotopic composition of ground water and, therefore, their place of origin. Plant physiology plays an important role in isotope fractionation processes within the plant, for example the opening or closing of stomata in response to water availability in the environment (atmospheric water and soil moisture). Evapotranspiration taking place during the maturation period of a plant also plays a factor in isotopic fractionation (Kahmen, Schefuß, & Sachse, 2013); this process causes the enrichment of heavy isotopes of oxygen and hydrogen in the water of fruits. In previous work, the stable isotopes of wine water ($\delta^{18}\text{O}$, $\delta^2\text{H}$) and ethanol ($\delta^{13}\text{C}$) have been used to authenticate the location of wine production (Christoph, Rossmann, Schlicht, & Voerkelius (2009, chap. 11); Calderone & Guillou, 2008; Raco, Dotsika, Poutoukis, Battaglini, & Chantzi 2015).

In food authenticity studies, statistical analysis plays an important role in model generation and subsequent use of models for classification of unknown-origin samples (Oulhote, Le Bot, Deguen, & Glorennec 2011). These models may be either univariate (e.g. using only ^2H for classification) or multivariate (e.g. using ^2H and ^{18}O in combination with relative concentrations of several chemical compounds for classification). The variables are normally complementary (i.e. not fully correlated). Therefore, at least theoretically, the use of multivariate models should be advantageous over univariate ones. Multivariate methods normally require the extraction of latent information to reduce the dimensionality of the data and provide a rapid overview of data prior to multivariate analysis. This approach has been used in authenticity studies aiming to discriminate between organic and conventional plants (Laursen, Schjoerring, Kelly, & Husted 2014). In some techniques, the data reduction process and the classification models are intrinsically related, and completed in one step, e.g. linear discriminant analysis (LDA) and partial least squares discriminant analysis (PLS-DA). Unsupervised dimension reduction methods such as principal component analysis (PCA) are not specifically tuned to classification as an objective, but rather describe the underlying data as a reduced dimensional projection, defined as a linear combination of the original variables. On the other hand, supervised classification methods such as LDA and PLS-DA perform dimension reduction in such a way as to maximize separation between known groups of samples. In this way, supervised methods rely on a priori assertions regarding sample groupings (Berrueta, Alonso-Salces, & Héberger 2007). Common to both unsupervised and supervised classification methods is the importance of careful model validation to avoid over-fitting and thus over-optimistic classification results (Kjeldahl & Bro, 2010).

In this study, isotope ratio mass spectrometry (IRMS) was used to develop a traceability tool for the geographical origin of bell peppers in both compound-specific ($^2\text{H}/^1\text{H}$) and bulk ($^{18}\text{O}/^{16}\text{O}$) analysis mode. The $^{18}\text{O}/^{16}\text{O}$ isotope ratios of the bell pepper fruit water and corresponding source water was determined using a continuous flow IRMS technique. Because the bell pepper fruit water samples contained high concentrations of dissolved organic

substances, we use an isotope equilibration technique by which the $^{18}\text{O}/^{16}\text{O}$ isotope ratio of the fruit water is determined. This was achieved by analyzing the $^{18}\text{O}/^{16}\text{O}$ isotope ratio ($\delta^{18}\text{O}$ value) of CO_2 gas after isotope equilibration with the fruit water samples. This equilibration technique (Epstein & Mayeda, 1953) is a widely accepted method for high-precision $\delta^{18}\text{O}$ analysis of water samples. It works particularly well for waters that are difficult to analyze with other techniques due to high salt or dissolved organic content. Furthermore, the $^2\text{H}/^1\text{H}$ isotope ratio of n-alkanes in the surface wax layer of the peppers was determined using gas chromatography coupled to IRMS (GC-C-IRMS). The lipid composition of different varieties of peppers was recently investigated by Parsons et al. (2013). Isotope ratios vary per lipid and give a good reflection of the isotopic composition of the meteorological spring water in different plant species; moreover, the hydrogen atoms are not likely to exchange due to transpiration during transport and sample handling (McInerney, Helliiker, & Freeman 2011; Sachse, Gleixner, Wilkes, & Kahmen 2010).

On the basis of n-alkane fingerprinting data, two chemometric classification models were developed and compared based on their ability to discriminate between peppers from different countries of origin. Firstly, an LDA model was developed, which is a model that has been used in many food authenticity studies (e.g. Kelly et al. 2005; Laursen et al. 2014). Secondly, a naïve Bayesian classifier using kernel-density estimation of the variables was evaluated using the same data set. Since the problem of food authenticity is also of forensic interest, probabilistic assertions regarding classification outcomes are particularly useful. A common strategy for “soft” classification is the use of a likelihood ratio (LR) approach. This has been performed in both (LDA and kernel-density-based) classifiers. Similar methods using LR were recently developed for wine authentication (Martyna, Zadora, Stanimirova, & Ramos 2014). The classification performances of both models were externally validated using cross-validation.

2. Material and methods

2.1. Samples, standards and software

Dutch and Spanish peppers were collected from two greenhouses in the Netherlands (Naktuinbouw, Roelofarendsveen; Helderman, Middenmeer) and two greenhouses in Spain (Enza seeds, Almería; IMIDA, Murcia). Italian (Sicily) and Israeli peppers were collected from a local Italian Bio market, and a Dutch supermarket (Albert Heijn, Amsterdam). The Dutch supermarket-bought peppers specified Israel as the country of origin. The number of samples (n) indicated below is per sampling date. Peppers ($n=20$) and water samples ($n=8$) from the Roelofarendsveen greenhouse were collected each month from July 2013 to September 2014; peppers ($n=20$) and water samples ($n=4$) from the Middenmeer greenhouse were collected monthly from March to September 2014. The Spanish peppers ($n=20$) and water samples ($n=2$) were collected in May and August 2014 in both Murcia and Almería; Israeli and Italian peppers ($n=5$ and $n=6$, respectively) were bought in May; no corresponding water was available for these samples as they were obtained from supermarkets. All other water samples were collected from the main water supply in the greenhouses on the day the peppers were harvested. In the greenhouse in Middenmeer, water (and pepper) samples were collected from four different sections to test the intra-greenhouse variability. Water was sampled in this greenhouse both at the supply and at the drainage point of each section. As a control, the rainwater collection basin was also sampled as well as water from the underground storage and window condensate. The underground storage contained rainwater that is collected and stored in antici-

pation of dryer periods. In the 2014 season, less than 1% of the total water supply was from this source. All water samples were collected in 50 mL Greiner tubes (Greiner Bio-One B.V., Alphen a/d Rijn, The Netherlands) and stored at -20°C .

For method evaluation and calibration of the $\delta^2\text{H}$ results, n-alkane standard solutions containing n-C₁₆ to n-C₃₀ were used (A5 and B3 mixtures, Schimmelmann, University of Indiana, USA), and n-C₂₀ (Sigma–Aldrich, 98%) was added to the samples as internal standard before extraction and served as quality control.

Statistical analysis was carried out using custom written scripts for modeling and evaluated in Matlab v. 2010b (The Mathworks, Natick, MA). For kernel-density estimation (KDE), the toolbox obtained from <http://www.ics.uci.edu/~ihler/code/kde.html> was used.

2.2. Extraction and sample clean-up

The peppers were freeze-dried for 3–5 days at a pressure below 5 kPa (50 mbar). Lipid extraction was performed on the freeze-dried peppers by rinsing the skin of the freeze-dried pepper with 10 ml chloroform. Each sample was spiked with an internal standard (icosane n-C₂₀) after extraction and each sample series was preceded by a blank sample containing chloroform plus the I.S. The extracts were evaporated to dryness under a gentle stream of nitrogen, dissolved in 3 mL hexane and filtered through a 0.45 μm syringe filter (Whatman, Dassel, Germany).

The extracts were purified on Teflon-coated Strata SI-1 SPE columns (Phenomenex, Torrance, CA, USA) placed on an SPE manifold. The columns were pre-conditioned twice with 3 ml of hexane before use, without letting them run dry. The extracts were added to the columns and eluted into a 10 ml sample vial. The original empty sample vial was rinsed three times with 0.5 ml of hexane, which was also transferred to the SPE column. Subsequently, the cartridge was eluted two times with 2 ml of hexane. The extract was evaporated to dryness under a gentle stream of nitrogen, redissolved in 1 ml hexane and stored at -20°C .

Constituents of the extract were identified and quantified using GC–MS by comparison to an external n-alkane standard mixture (n-C₂₁ to n-C₄₀) and corrected to the peak area of the internal standard. The recovery and the reproducibility of the sample clean-up method ($n = 10$) was tested using a standard mixture containing 4 n-alkanes plus internal standard in concentrations of 15 ng/ μL (n-C₁₈, n-C₂₄ and n-C₃₀) and 20 ng/ μL (5 α -androstane and n-C₂₀).

2.3. Gas chromatography-mass spectrometry

GC–MS was performed using a ThermoQuest Trace GC 2000 connected to a Finnigan Trace MS quadrupole mass spectrometer (ThermoFisher, San Jose, CA, USA). GC conditions were: 1.0 μL on-column injection at 5 $\mu\text{L}/\text{s}$, 2 m Siltek de-activated pre-column (i. d. 0.53 mm) connected to a 30 m Rtx-5Sil MS column (Restek, i. d. 0.25 mm; film thickness 0.1 μm), carrier gas Helium at a constant flow of 0.8 ml/min. The temperature program was as follows: 50 $^{\circ}\text{C}$ (2 min) to 80 $^{\circ}\text{C}$ at 40 $^{\circ}\text{C}/\text{min}$ (hold 2 min), to 130 $^{\circ}\text{C}$ at 20 $^{\circ}\text{C}/\text{min}$, to 350 $^{\circ}\text{C}$ at 4 $^{\circ}\text{C}/\text{min}$ (hold 10 min). MS conditions were: full scan (m/z 50–650), cycle time 0.65 s, electron ionization (70 eV). For quantification of the n-alkanes, an n-alkane standard mixture (n-C₂₀, n-C₂₇, n-C₂₉, n-C₃₁ and n-C₃₃) was used. Calibration plots of the n-alkanes were constructed by injecting a dilution series of the standard mixture at concentrations ranging from 0–25 ng/ μL (5 data points in triplicate).

2.4. Gas chromatography-hydrogen isotope ratio mass spectrometry

Gas chromatographic separation was performed using an Agilent 7890A at constant flow (0.8 ml/min) on a 30 m Agilent J&W,

DB-5MS column (J&W, 0.25 mm i. d.; film thickness 0.25 μm) with 0.2 m Siltek de-activated pre-column (0.53 mm i. d.). For each sample, 1.0 μL was injected at 6000 $\mu\text{L}/\text{min}$ in splitless mode at 275 $^{\circ}\text{C}$. The GC temperature program started at 60 $^{\circ}\text{C}$ for 2 min, then to 90 $^{\circ}\text{C}$ at 60 $^{\circ}\text{C}/\text{min}$ (hold 2 min), then to 130 $^{\circ}\text{C}$ at 20 $^{\circ}\text{C}/\text{min}$, and finally to 310 $^{\circ}\text{C}$ at 4 $^{\circ}\text{C}/\text{min}$ (hold 14 min). The eluting compounds were transferred to a high-temperature conversion furnace (Iso-prime, UK) operated at 1000 $^{\circ}\text{C}$ and quantitatively converted to H₂, which was then introduced into an isotope ratio mass spectrometer (IRMS, Isoprime 100, UK) for compound-specific analysis of $\delta^2\text{H}$ values. The H₃⁺ factor was determined at the start and the end of each 24hr sequence, and remained constant (SD 0.1275; $n = 32$) during the whole measurement period. Three replicate measurements were performed on each sample. An n-alkane standard (Mix A5, A. Schimmelmann, University of Indiana) of 15 externally calibrated n-alkanes was measured in triplicate every three samples and used daily to normalize all $\delta^2\text{H}$ values to the Vienna Standard Mean Ocean Water (VSMOW) scale.

2.5. Oxygen isotope ratio mass spectrometry

Source water and fruit water samples were analyzed on a Thermo Finnigan Delta + mass spectrometer equipped with a GASEBENCH-II preparation device (Thermo Scientific, Bremen, Germany). 0.5–1 milliliter of samples was injected through the septum cap of a 10 ml Exetainer vial (Labco Ltd, High Wycombe, UK) filled with a mixture of He and 0.2% CO₂. Within 24 h at 22 $^{\circ}\text{C}$, the oxygen in the headspace CO₂ was isotopically equilibrated with that of the water sample. Subsequently, the CO₂–He gas mixture was transported to the GASEBENCH-II using a He flow through a flushing needle system. In the GASEBENCH-II, water was extracted from the gas, using NAFION tubing, and CO₂ was analyzed in the mass spectrometer after separation of other gases on a GC column. Values are reported as $\delta^{18}\text{O}$ vs V-SMOW. The long-term reproducibility of a routinely analyzed lab water standard was better than 0.1% (1SD).

2.6. Analytical performance GC–MS and GC–IRMS methods

The linearity of the GC–C–IRMS method was tested by injecting a range of concentrations of the n-alkane A5 mix (0–200 ng on column, 5 levels in sixfold) and the stability of the $\delta^2\text{H}$ values determined. Calibration curves were constructed by plotting the detection response of the standard solutions versus the concentrations. Following regression analysis, the correlation coefficients (R^2) of the calibration curves were calculated. The criterion for good linearity was $R^2 > 0.990$. To determine the $\delta^2\text{H}$ stability, the same data were used and the normalized $\delta^2\text{H}$ values ($\delta_{\text{measured}}/\delta_{\text{average}}$ 0–200 ng) were plotted against concentration. The $\delta^2\text{H}$ values were considered stable when they landed within the 95th percentile range. The reproducibility of the procedure was tested by determination of the relative standard deviation of the sixfold measurements for each compound at each concentration.

To ensure good method performance of the $\delta^{18}\text{O}$ set-up, three water samples from a proficiency test organized by Agroisolab (Jülich, Germany) were analyzed on the EA–IRMS system and compared to the results of ten other labs.

3. Results and discussion

3.1. Analytical performance

The analytical method validation was based on three parameters: linearity, stability, and repeatability. Both the GC–MS and GC–IRMS methods were deemed satisfactory in terms of these criteria. The GC–MS system was linear between 0 and 25 ng/ μL for the

four main alkanes n-C₂₇, n-C₂₉, n-C₃₁ and n-C₃₃, exhibiting R² values between 0.994 and 0.997, and limits of detection (S/N = 3) between 0.05 and 0.15 ng/μL in all cases. The method was repeatable with a relative standard deviation for peak areas between 2 and 5%.

The GC–IRMS system was linear for values of 0–200 ng injected on-column for the four main alkanes, n-C₂₇, n-C₂₉, n-C₃₁ and n-C₃₃, with R² values between 0.992 and 0.995. The GC–IRMS method also demonstrated stability over the concentration range tested with all measurements falling within the 95th percentile range. The method was repeatable with a δ²H relative standard deviation ≤0.6% in all cases.

The mean δ¹⁸O values of the three water samples from the proficiency test were 0.3% (n = 9), –8.3% (n = 10) and –6.6% (n = 10), respectively, and the median values were 0.5%, –8.3% and –6.5%, respectively. The results obtained in our lab were 0.5%, –8.4% and –6.5%, respectively. These results are very much in line with the results of the other labs participating in the proficiency test.

3.2. δ¹⁸O source water and pepper fruit water

Monthly averages (May–Sep) for the source water δ¹⁸O values from the Middenmeer (NL) greenhouse varied between 5.2 and 6.6, and were found to be stable within a batch (n = 8, SD < 0.21%). The δ¹⁸O values of the other water sources all showed averages around –6.0%, with values between –5.9% and –6.3% for the underground storage, –2.8% and 10.5% for the condensate and –1.6% and –7.5% for the rainwater collection basin. The δ¹⁸O values of the storage water were quite constant over the measurement period, as were the values for the basin, with the only exception being the values for July (–1.6%), which was likely due to the low water level combined with high evaporation from the basin resulting from high temperatures experienced during that month. The δ¹⁸O values of the condensate varied to a larger extent, which is to be expected due to evaporation and condensation effects.

The δ¹⁸O values of the source water of the Roelofarendsveen (NL) greenhouse showed somewhat greater variation than those for the Middenmeer greenhouse; between 3.3% and 7.1%. This was most likely related to larger evaporation effects caused by the small size of the water basin and variable water source. During maintenance of the water basin in the months Dec/Jan water was taken from a nearby creek to irrigate the peppers.

δ¹⁸O values for the pepper fruit water varied between –4.3% and 3.3% (n = 20), with the highest values in Roelofarendsveen and Murcia. In Fig. 1, the δ¹⁸O values for the fruit water are plotted alongside the δ¹⁸O values of the corresponding source water. The graph also contains the δ¹⁸O data points for the water collected after freeze-drying of several pepper batches and the average monthly values per location of the GNIP database. For Murcia, GNIP monthly averages for May and Aug (2000–2006) were used and, for Middenmeer, monthly averages from Wieringerwerf, located 5 km from Middenmeer, were used (Mar–Sep, 1988–1992); for Roelofarendsveen no GNIP data were available. The results demonstrated a good correlation between the fruit- and source water data (Pearson's R = 0.774), with a constant isotope fractionation from source to fruit water of about 4%. The source water δ¹⁸O data also has a good correlation with the local long-term GNIP averages (Pearson's R = 0.833) and the δ¹⁸O values for water from the freeze-dryer correlated very well with the fruit water data (Pearson's R = 0.989). This indicates that combining δ¹⁸O values for locally sourced water (and/or monthly GNIP data) with water from different (freeze-dried) pepper fruits means pairs of samples can be compared and assertions deduced regarding origin.

3.3. Quantification and δ²H analysis of n-alkanes

Unlike oxygen, which plants acquire from atmospheric oxygen, carbon dioxide, and source water, the only source of hydrogen for photosynthesis is water. In addition to δ¹⁸O, the δ²H values of the bell pepper fruits were also investigated using GC–C-IRMS analysis of n-alkanes in extracts from the surface wax. Extraction and analytical method development were published in a separate paper (De Rijke et al. 2015). The main n-alkanes found in the pepper extracts were n-C₂₇, n-C₂₉, n-C₃₁ and n-C₃₃, corresponding to earlier findings (Bauer, Schulte, & Thier 2005; Parsons et al. 2013). Their concentrations varied between 0.01 and 105 μg/dm², and their total sum was between 0.19 and 177 μg/dm². In all cases, their concentrations were ranked n-C₃₁ > n-C₂₉ > n-C₂₇ > n-C₃₃, but their relative ratios varied. The GC–C-IRMS chromatogram of a typical pepper extract is presented in Fig. 2. As can be seen in the figure, the peaks were very well resolved, which is essential for a good IRMS measurement (and GC–MS quantification). In some cases the intensity of the IRMS signal dropped below 1 nA making it difficult to obtain a stable IRMS signal. This was the case for n-C₂₇ and n-C₃₃ in most samples despite their concentration by a factor of five. Therefore, for these alkanes only, the relative concentrations and no δ²H data were used in the data analysis. Furthermore, a slight increase in signal was observed for the n-C₂₉ IRMS data over the triplicate measurement series, which was most pronounced after injection of the B3 standard mixture. This was likely caused by discrepancies between triplicate measurements of the n-C₂₉ concentration in the sample (known to show large variations from sample to sample) and the standard mix, which might result in a memory effect across the system. This effect was not observed for n-C₃₁ as it was not present in the standard mixture. Given these observations, and in light of the strong correlation between n-C₂₉ and n-C₃₁ values for all samples, subsequent statistical data analysis was based only on the n-C₃₁ δ²H values, combined with the relative concentration ratios (in μg/dm²) of all major n-alkanes.

3.4. Statistical data analysis/classification

In order to check the importance of the inclusion of the n-alkanes in the classification model, a series of LDA models were build using four conditions: (a) considering only δ²H measurements, (b) considering δ²H and relative concentration of C₂₇, (c) δ²H and relative concentration of C₂₇ and C₂₉, (d) δ²H and relative concentrations of C₂₇, C₂₉ and C₃₁. The misclassification rates of these models were tested using Monte-Carlo 5-fold cross-validation. The cross-validation was stratified in order to compensate for the unbalanced representation of the four groups (Spain, Netherlands, Italy and Israel). Misclassification rates for the 4 models were: (a) 39%; (b) 18%; (c) 13%; (d) 12%. The classification performances were tested using the Friedman test, and differences were significant (p-value = 7.4e–9). A post hoc analysis (with a pair-wise multiple comparison) was performed to check which classification methods were considered significantly different from the others. Using the Tukey–Kamer correction method and average ranks for columns, we concluded that methods a and b were significantly different. Methods c and d were not significantly different. A value of α = 0.05 was used for all analyses. Based on Friedman's test a clear improvement in classification rates was observed between a and b. These results suggest that the use of the relative profile of n-alkanes provides information relevant for class discrimination. Also, the relatively low, but still significant decrease in error rates from model b to c, and the even lower (in this case not significant) decrease from c to d indicates some correlation between the values of the n-alkanes. In other words, the inclusion of the relative concentration of C₃₁ in the model does not improve

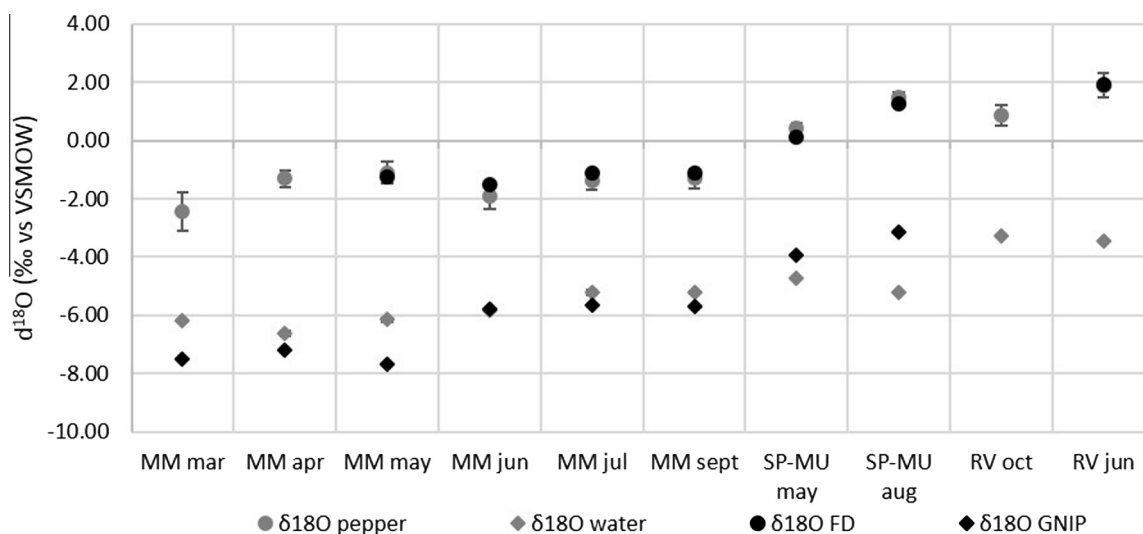


Fig. 1. Mean regional monthly $\delta^{18}\text{O}$ values of pepper fruit water (grey dots, $n = 20$), corresponding source water (grey diamonds, $n = 2$), water collected after freeze-drying peppers (black dots, $n = 1$: combined batch of 20 peppers) and long-term average monthly GNIP values (black diamonds, $n = 3$ –11). MM: Middenmeer, Netherlands; SP-MU: Murcia, Spain; RV: Roelofarendsveen, Netherlands.

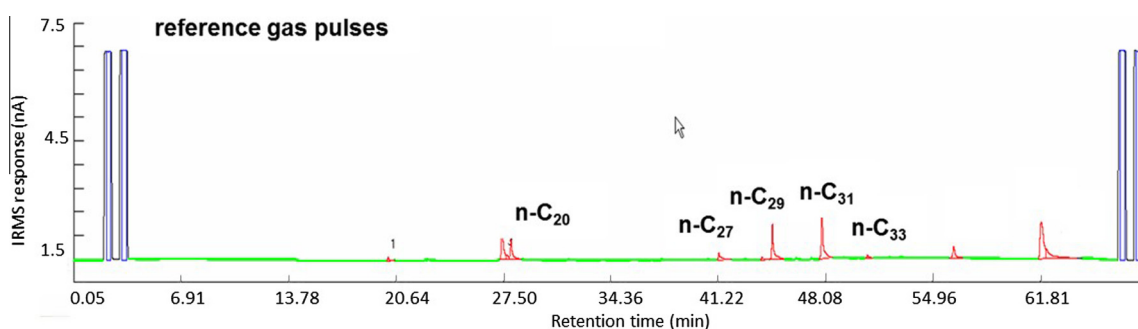


Fig. 2. GC-C-IRMS $\delta^2\text{H}$ chromatogram of a pepper extract.

the model significantly once $\delta^2\text{H}$ and the relative concentrations of C₂₇ and C₂₉ have been included. Table 1 shows the confusion matrix for the LDA model using the four variables.

Peppers from Italy and Israel can be clearly distinguished from the rest. The distinction between Spain and the Netherlands proved to be more difficult, especially in the case of Dutch peppers being misclassified as Spanish ones. This can be explained by inspecting the distribution of $\delta^2\text{H}$ values for the four groups (see Fig. 3). In this figure, it can be clearly seen that the isotope ratios are informative enough to distinguish between the origin of peppers from Italy and Israel, but some overlap occurred between the Netherlands and Spain. Inclusion of the n-alkanes measurements improves this distinction (the misclassification rates drop from 0.39 to 0.12).

In order to focus on the domestic authenticity problem, a classification task was defined with two groups: Dutch versus non-Dutch. The confusion matrix resulting from performing the LDA

classification with the newly defined (two-class) labels is shown in Table 2A.

Improved performance was observed for the LDA model, as it no longer had to classify between Italy, Spain and Israel. Three samples from the Dutch group were systematically misclassified as non-Dutch. On the other hand, one (Spanish) sample from the non-Dutch group was systematically misclassified as Dutch. Inspection of data from these samples showed that a sample from Murcia-Spain (harvested in August) had a $\delta^2\text{H}$ value more similar to average Dutch values and the three misclassified Dutch samples had $\delta^2\text{H}$ values that were more similar to the average Murcian values. This may indicate a change in local irrigation source, such that Murcian peppers harvested in August generate $\delta^2\text{H}$ values that are similar to the Dutch peppers harvested in spring (all misclassified Dutch samples were from April).

One of the assumptions implicit to LDA is the use of a multivariate Gaussian to model the probability density of each class. An extra premise (which distinguishes it from Quadratic Discriminant Analysis) is the assumption that the variance-covariance matrix is the same for all classes. Although this might be a strong assumption, LDA still exhibits satisfactory classification performance (Hastie, Tibshirani, & Friedman 2009).

LDA is a soft classifier by nature (Hastie et al., 2009). However, it can easily be converted to a hard classifier by assuming prior probabilities for each class, and defining a threshold for the posterior probability. In this way, a sample is assigned to a specific class when its posterior probability for that class is greater than that

Table 1
Confusion matrix of the LDA model using the four groups. This is the result of 100 Monte-Carlo repetitions in a 5-fold cross-validation strategy.

		Result of the test			
		Spain	The Netherlands	Italy	Israel
Ground truth	Spain	19	1	0	0
	Netherlands	6	24	0	0
	Italy	0	0	6	0
	Israel	0	0	0	5

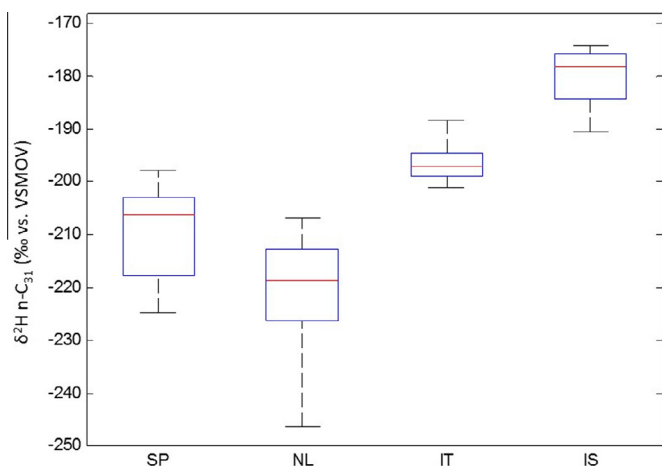


Fig. 3. $\delta^2\text{H}$ boxplots of $n\text{-C}_{31}$ for the different locations. SP: Spain, NL: Netherlands, IT: Italy, IS: Israel. The box represents the 1st to 3rd IQR (interquartile range), the whiskers represent $1.5 \times \text{IQR}$ and the red bar is the median value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2

Confusion matrix of the A) linear discriminant analysis (LDA) model using two groups and B) kernel-density estimation (KDE) model using two groups. The results of 100 Monte-Carlo repetitions in a 5-fold cross-validation strategy are presented.

		Result of the test	
		Somewhere else	Netherlands
A: LDA			
Ground truth	Somewhere else	30	1
	Netherlands	3	27
B: KDE			
Ground truth	Somewhere else	28	3
	Netherlands	1	29

of all other classes. In most of applications of LDA, a sample is assigned to the class showing the highest likelihood, which implicitly assumes an equal prior probability of belonging to each class. In the particular case of a two-class problem, as the one addressed here, the LR is inspected, and the sample is assigned to one of the classes depending whether this LR is above or below 1.

The definition of a common variance–covariance matrix for all classes is quite a strong assumption. In order to avoid this

assumption a purely data-driven approach was evaluated and the classification results were compared with those produced using LDA. This classification model was defined by modeling the probability densities of both groups using kernel-density estimation (KDE). The probability densities of each group (Dutch and non-Dutch) are first modeled using a Gaussian kernel. A large number of samples is required to model these distributions correctly, especially when the number of variables (4 in our case) is relatively large in relation to the number of samples (61 in our case). To ensure stable densities, PCA was used to reduce the number of variables. This approach was particularly effective as we assumed that relative concentrations of C_{27} , C_{29} and C_{31} were correlated. Following dimension reduction by PCA, KDE was performed using only two variables: the values of $\delta^2\text{H}$ and the scores of the first principal component of a PCA model of the values of the relative concentrations of C_{27} , C_{29} and C_{31} .

Similar to LDA, KDE yields a soft classifier by nature, but it can be transformed into a hard classifier considering prior distributions and assigning each sample to the class that shows the highest posterior probability. By assuming flat priors (as in the case of LDA), the classification step is performed as follows. An unknown sample is classified as belonging to one group or another by measuring the likelihood at the experimental point using both (KDE-modeled) densities. The LR is then calculated, and the sample assigned to one group or another depending whether the LR is greater or less than 1. In our example, predicted labels were based on the LR of the Dutch group versus the non-Dutch group. LR greater than 1 indicated the Dutch group whilst LR less than 1 indicated the non-Dutch group.

Compared to LDA, this method makes no assumptions about the variance–covariance structure of either group, as such it is more flexible in terms of modeling the probability density function. In order to validate the model correctly, both the PCA dimension reduction and the kernel-density estimation was performed exclusively using the training subset excluding the hold-out portion of each cross-validation fold. The same 5-fold cross-validation strategy was followed (as with the LDA model) using 100 Monte-Carlo repetitions. Table 2B presents the confusion matrix achieved using this classification method.

In this case, the model was more prone to “non-Dutch” peppers misclassified as “Dutch” compared to the LDA model. In contrast, it was less prone to “Dutch” peppers misclassified as “non-Dutch”. There was quite some overlap in samples misclassified by both methods; two out of three misclassified non-Dutch samples were from Murcia (harvested in August) and the Dutch sample (misclas-

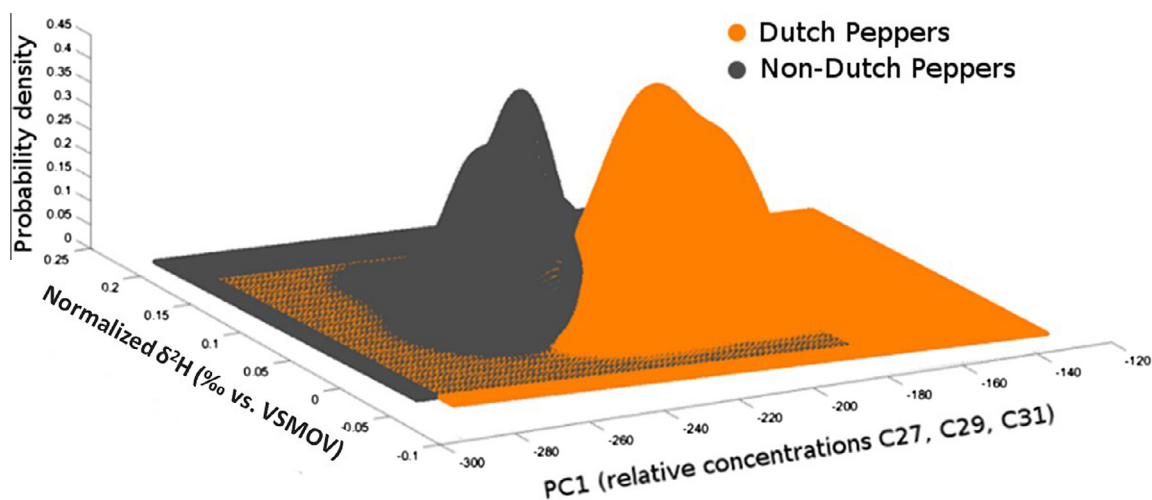


Fig. 4. Density of both groups (Dutch vs. non-Dutch peppers), modeled using KDE (Gaussian kernel).

sified as “non-Dutch”, harvested in the Netherlands during April) had similar $\delta^2\text{H}$ values.

Because the kernel-density model was less prone to misclassification of Dutch peppers, it seemed preferable in an application where the main goal was to prevent Dutch products being non-Dutch in origin. Conversely, the LDA method was less prone to errors where non-Dutch peppers were classified as Dutch in origin. From the perspective of using the model as a soft classifier (i.e., getting the probability of belonging to each of the groups), the kernel-density estimator provided much better sensitivity to the heterogeneous distribution of different paprika populations. This would also mean the probability density function is better modeled, given sufficient data. The non-linear classification threshold and two-class probability densities are shown in Fig. 4. In this figure, the density of both groups is shown, estimated using a Gaussian kernel. Both densities exhibit non-Gaussian shapes and, moreover, it is quite unlikely that the variance–covariance of both densities could be the same. Despite these observations, comparable total error was observed using the LDA model as a hard classifier (Hastie et al., 2009).

It should be noted that, implicit in our use of soft classifiers to perform “hard” label assignments classifiers, we have assumed equal prior probabilities over the classes of interest. In the absence of compelling insight into the prevalence of erroneous geographic labeling of peppers, we have opted for an uninformative (or flat) prior. In reality, we may want to consider the likely motivation for misreporting the geographic origin of peppers by factoring in the relative cost of peppers from different regions as well as the availability of Dutch versus non-Dutch peppers. This can easily be done by adapting the prior probabilities of the classes. With increasing sample size, the LDA model might be expected to exhibit worse performance than the kernel density-based estimate (because of the assumption of having a unique variance–covariance matrix). In our case, more samples were needed to effectively make such conclusions. More samples would also allow sensible use of performance metrics that are targeted at comparing LR and soft classification methods, such as empirical cross entropy plots (see Martyna et al. 2014 for an application in wine authenticity).

4. Conclusions

Two analytical strategies were investigated, based on bulk $\delta^{18}\text{O}$ elemental analysis of source and pepper fruit water, and on compound-specific, n-alkane, $\delta^2\text{H}$ GC–C–IRMS analysis. Both strategies were able to discriminate between Dutch bell peppers and bell peppers from other countries. LDA delivered a satisfactory overall classification performance (93% accuracy), despite simplified class distribution assumptions. A density-based method was evaluated for the same classification task. Comparable results were achieved in terms of overall classification accuracy. However, some peppers were misclassified.

The difference in classification thresholds observed originated from specific assumptions made regarding the distribution of bell peppers in terms of their distinctive compound-specific GC–C–IRMS and relative concentration ratio data. The accuracy of density models (and therefore classification performance) could be improved with more complete and representative sampling of the population of bell peppers in geographic regions pertinent to the classification task.

These results show it is possible to establish classification models for bell pepper from either $\delta^{18}\text{O}$ elemental analysis IRMS or compound-specific GC–C–IRMS based data. The $\delta^{18}\text{O}$ data from different (freeze-dried) pepper fruits correlated well with the long-term GNIP averages and the source water (and/or monthly GNIP

data) and, therefore, pairs of samples could be compared and assertions deducted regarding origin. The proposed compound-specific GC–C–IRMS based method provides a tool for pepper classification according to geographical origin and could serve as a technique to verify compliance regarding geographic origin labeling of peppers.

Acknowledgements

We thank the Dutch Horticultural Inspection Service Naktuinbouw, Barry Helderma from Kwekerij Helderma VOF, Antonio Sances Lopez from Enza Zaden Almeria and Francisco del Amor from Instituto Murciano de Investigación y Desarrollo Agrario y Alimentario for donation of the peppers. The European Fund for Regional Development Kansen voor West programme, Province North-Holland and the Amsterdam Economy Board are acknowledged for financial support (137323/156855 EFRO Herkomstbepaling).

References

- Bauer, S., Schulte, E., & Thier, H.-P. (2005). Composition of the surface waxes from bell pepper and eggplant. *European Food Research and Technology*, 220, 5–10.
- Berrueta, L. A., Alonso-Salces, R. M., & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatogr. A*, 1158, 196–214.
- Calderone, G., & Guillou, C. (2008). Analysis of isotopic ratios for the detection of illegal watering of beverages. *Food Chemistry*, 106, 1399–1405.
- Christoph, N., Rossmann, A., Schlicht, C., & Voerkelius, S. (2009). Wine Authentication Using Stable Isotope Ratio Analysis: Significance of Geographic Origin, Climate, and Viticultural Parameters. In *ACS symposium series: Vol. 952. Authentication of food and wine* (pp. 166–179).
- De Rijke, E., Fellner, C., Westerveld, J., Cerli, C., Kalbitz, K., & de Koster, C. G. (2015). Determination of n-alkanes in *C. annuum* (bell pepper) fruit and seed using GC–MS: comparison of extraction methods and application to samples of different geographical origin. *Analytical and Bioanalytical Chemistry*, 407, 5729–5738.
- EEC No. 1169/2011. EEC No. 1169/2011 on the provision of food information to consumers. *Official Journal of the European Communities L* 304, 18–63.
- Epstein, S., & Mayeda, T. (1953). Variations of ^{18}O content of waters from natural sources. *Geochimica et Cosmochimica Acta*, 4, 213–224.
- Hastie, T., Tibshirani, R., & Friedman, J. (Eds.). (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York: Springer.
- Kahmen, A., Schefuß, E., & Sachse, D. (2013). Leaf water deuterium enrichment shapes leaf wax n-alkane δD values of angiosperm plants I: Experimental evidence and mechanistic insights. *Geochimica et Cosmochimica Acta*, 111, 39–49.
- Kelly, S., Heaton, S., & Hoogewerff, J. (2005). Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. *Trends in Food Science & Technology*, 16, 555–567.
- Kjeldahl, K., & Bro, R. (2010). Some common misunderstandings in chemometrics. *Journal of Chemometrics*, 24, 558–564.
- Laursen, K. H., Schjoerring, J. K., Kelly, S. D., & Husted, S. (2014). Authentication of organically grown plants – advantages and limitations of atomic spectroscopy for multi-element and stable isotope analysis. *Trends in Analytical Chemistry*, 59, 73–82.
- LEI (2013). The agricultural trade in the Netherlands in 2013. De agrarische handel van Nederland in 2013. Cited on 11-2-2015. <<http://www.rijksverheid.nl/documenten-en-publicaties/rapporten/2014/01/17/de-agrarische-handel-van-nederland-in-2013.html>>.
- Martyna, A., Zadora, G., Stanimirova, I., & Ramos, D. (2014). Wine authenticity verification as a forensic problem: An application of likelihood ratio test to label verification. *Food Chemistry*, 150, 287–295.
- McInerney, F. A., Helliker, B. R., & Freeman, K. H. (2011). Hydrogen isotope ratios of leaf wax n-alkanes in grasses are insensitive to transpiration. *Geochimica et Cosmochimica Acta*, 75, 541–554.
- Oulhote, Y., Le Bot, B., Deguen, S., & Glorennec, P. (2011). Using and interpreting isotope data for source identification. *Trends in Analytical Chemistry*, 30, 302–312.
- Parsons, E. P., Popovskiy, S., Lohrey, G. T., Alkalai-Tuvia, S., Perzelan, Y., Bosland, P., ... Jenks, M. A. (2013). Fruit cuticle lipid composition and water loss in a diverse collection of pepper (*Capsicum*). *Physiologia Plantarum*, 149, 160–174.
- Raco, B., Dotsika, E., Poutoukis, D., Battagliani, R., & Chantzi, P. (2015). O–H–C isotope ratio determination in wine in order to be used as a fingerprint of its regional origin. *Food Chemistry*, 168, 588–594.
- Rossmann, A. (2001). Determination of stable isotope ratios in food analysis. *Food Reviews International*, 17, 347–381.
- Sachse, D., Gleixner, G., Wilkes, H., & Kahmen, A. (2010). Leaf wax n-alkane dD values of field-grown barley reflect leaf water dD values at the time of leaf formation. *Geochimica et Cosmochimica Acta*, 74, 6741–6750.