Food Control 161 (2024) 110397

Contents lists available at ScienceDirect

Food Control

journal homepage: www.elsevier.com/locate/foodcont

Authentication strategy for paprika analysis according to geographical origin and study of adulteration using near infrared spectroscopy and chemometric approaches

Ana Castell, Natalia Arroyo-Manzanares **, Ignacio López-García, Félix Zapata, Pilar Viñas *

Department of Analytical Chemistry, Faculty of Chemistry, University of Murcia, Regional Campus of International Excellence "Campus Mare Nostrum", E-30100, Murcia, Spain

ARTICLE INFO

Keywords:
Spices
Sudan dyes
Congo red
Food fraud
Food analysis
NIR spectroscopy

ABSTRACT

Paprika is a spice whose composition and characteristics vary with its geographical origin and additionally is illegally adulterated with dyes to improve its appearance. This work proposes a strategy based on Fourier-transform near infrared (FT-NIR) analysis and chemometric tools for its authentication and detection of fraud. A total of 115 paprika samples were analyzed, including paprika with protected designation of origin (PDO) labels from Spain, France and Hungary, and samples from China and Zambia. The proposed orthogonal partial least squares-discriminant analysis (OPLS-DA) models allow to distinguish paprika according to its PDO and variety, as well as to identify adulteration with Sudan dyes or Congo red. Partial least squares regressions allow to quantify the adulterant in paprika from 0.1 to 5 %. Chemometric models achieved high classification success rates and suitable linearities. The proposed strategy is presented as a comprehensive and effective tool to ensure paprika quality and authenticity, including the detection and quantification of adulteration with commercial dyes.

1. Introduction

Paprika is a spice obtained by grinding dried fruits of the genus *Capsicum annuum* L. widely used in the food industry as a natural coloring and flavouring agent, due to its distinctive taste, flavor and high coloring capacity. Paprika is a natural source of nutrients such as minerals and vitamins C and E, and bioactive compounds as carotenoids, capsaicinoids and phenolic compounds (Baenas et al., 2019). These compounds provide anti-inflammatory and antioxidant effects and play a key role in the prevention of several diseases, including cardiovascular diseases and several types of cancer (Hayman & Kam, 2008; Saini et al., 2020). The composition of this spice is determined by several factors such as the variety of pepper, the geographical origin, the climatic conditions or the production process. Therefore, the protection of the geographical area has become one of the main tools of certifying the authenticity of foods, recognizing their unique and distinctive characteristics. This is indicated by the Protected Designation of Origin (PDO) label. In Europe, three areas of paprika production with PDO are from Spain (*Pimentón de La Vera*, *Pimentón de Murcia* and *Pebre bord de Mallorca*), two from Hungary (*Kalocsai fűszerpaprika-őrlemény* and *Szegedi fűszerpaprika-őrlemény*), one from Slovakia (*Paprika Žitava*) and one from France (*Piment d'Espelette*) (Monago-Maraña et al., 2022).

Due to their high quality and prize, paprika with PDO is an attractive target for adulteration and mislabelling (Sun et al., 2022; Van Asselt et al., 2018). Thus, studies to ensure the authentication of paprika have grown over the years. The differentiation between Spanish paprikas from "La Vera" and "Murcia" has been studied by spectroscopic techniques as ultraviolet–visible (UV–Vis) based on color measurements (Palacios-Morillo et al., 2016) and visible near-infrared spectroscopy (Vis-NIR) (Monago-Maraña et al., 2021). High-performance liquid chromatography (HPLC) coupled to UV (Cetó et al., 2018, 2020), mass spectrometry (MS) (Barbosa et al., 2020) and fluorescence detector (FLD) (Campmajó et al., 2021) was applied for the differentiation of several paprikas with PDO based on phenolic compounds profile of the samples. Lighter sample treatment has been proposed, mixing the paprika with wax before its determination by energy dispersive X-ray fluorescence (ED-XRF) for paprika from "La Vera" (Fiamegos et al., 2021).

* Corresponding author.
** Corresponding author.
E-mail address: pilarvi@um.es (P. Viñas).

https://doi.org/10.1016/j.foodcont.2024.110397
Received 22 November 2023; Received in revised form 12 February 2024; Accepted 17 February 2024
Available online 17 February 2024

On the other hand, adulteration of paprika can be carried out by the addition of an external materials, own materials of inferior production and colorants (Oliveira et al., 2019). Analytical methodologies have been purpose for the identification of adulterated paprika with gum arabic (Horn et al., 2018; Oliveira et al., 2020), potato starch and annatto (Oliveira et al., 2020), corn flour (Zaukuu et al., 2019), tomato skin and red brick dust Galaxy Scientific Inc., 2016), lead oxide (II, IV), lead chromate, silicon dioxide and polyvinyl chloride (Horn et al., 2018). Galvin-King et al. proposed the detection of paprika adulterated with spent paprika (after extracting oleoresin), considered a waste product (Galvin-King et al., 2020). Moreover, paprika fraud in terms of authentication was studied by mixing different regions and cultivars of paprika samples (Sun et al., 2022, 2023). The addition of synthetic colorants as Congo red (Lohumi et al., 2018), azorubine and cochineal red A, as well as natural (sumac and beetroot) colorants (Horn et al., 2021) has also been explored to ensure the authenticity of paprika. The addition of colorants to paprika provide a bright color making it more attractive to the consumer as well as being durable and low cost (Monago-Maraña et al., 2022). Particular importance has been given to the detection of paprika adulterated with Sudan dyes (Sudan I, II, III, IV) (Di Anibal et al., 2009, 2011, 2012, 2014, 2015; Galaxy Scientific Inc., 2016; Gao et al., 2015; Horn et al., 2018; Hu et al., 2017; Jahn et al., 2015; Lohumi et al., 2017, 2018; Márquez et al., 2019; Mohamed et al., 2021; Monago-Maraña et al., 2019; Vera et al., 2018). These synthetic compounds are illegally food additives commonly used as colorants in textiles, plastics and other products, and its degradation in the organism may pose a health risk (Liu et al., 2015; Xu et al., 2007).

Spectroscopic techniques combined with chemometrics have demonstrated numerous advantages for the detection of adulterated paprika with dyes, being considered the most effective methodology for this purpose (Monago-Maraña et al., 2022). Among them, UV–Vis spectroscopy has been the most used one (Di Anibal et al., 2009, 2014; Márquez et al., 2019; Vera et al., 2018). However, sample pre-treatment is necessary prior to analysis by this technique, which require more time and involves the use of solvents. Other methods employ Raman (Lohumi et al., 2018; Monago-Maraña et al., 2019) and Surface Enhanced Raman Spectroscopy (SERS) (Di Anibal et al., 2012; Jahn et al., 2015). Although SERS has been used to improve the sensitivity of conventional Raman, quantification can be difficult because it depends on the interaction analyte-nanoparticles (Monago-Maraña et al., 2019). ^{1}H Nuclear Magnetic Resonance (^{1}H NMR) (Di Anibal et al., 2011; Hu et al., 2017) and infrared techniques such as Fourier-transform mid-infrared (FT-MIR) (Horn et al., 2018; Lohumi et al., 2017) are other spectroscopic techniques used for this purpose. Besides spectroscopic techniques, liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) was applied for quantification of dyes in spices including paprika (Mohamed et al., 2021). However, this methodology is destructive to the sample and requires more analysis time than spectroscopic methods.

This work proposes a strategy to ensure the authenticity and quality of paprika using the sample fingerprint provided by FT-NIR spectroscopy. From the chemometric models developed based on orthogonal partial least squares-discriminant analysis (OPLS-DA) and partial least squares (PLS) regressions, the aim is to detect and quantify paprika samples adulterated with Sudan II, III, IV and Congo Red, as well as to distinguish between different PDOs from Spain, France and Hungary in a single analysis. In addition, the variety of paprika samples (hot, sweet and smoked) can also be identified by the proposed chemometric strategy. To the best of our knowledge, this is the first time that paprika authentication has carried out in terms of adulteration and mislabelling simultaneously including samples of five different PDO paprikas. Thus, the technological challenge that has been achieved is to carry out the authentication of paprika samples with a non-invasive instrumental strategy which does not require a previous sample treatment and takes only a few seconds, thus surpassing the current state of the art.

2. Materials and methods

2.1. Paprika samples and adulterants

In this study, 115 paprika samples from five different countries were analyzed: Spain (65), France (14), Hungary (16), China (12) and Zambia (8). Spanish samples from two regions with PDO “Murcia” (38) and “La Vera” (27), French paprika from “Espelette”, Hungarian samples from the two main producing regions “Kalocsa” (9) and “Szeged” (7) with PDO, and China and Zambia paprikas, were obtained directly from the producers to ensure the authenticity of such samples. Paprika samples were of different varieties: hot (26), sweet (51), smoked (22), sweet smoked (11) and hot smoked (5). In addition, paprika samples with organic certification from Murcia (4) were included. Different types and varieties of paprikas were used to enrich the chemometric models by covering a wide range of paprika characteristics. Supplementary Table 1 contains the information of all paprika samples analyzed. Each sample was pounded and stored at room temperature until analysis. Sudan dyes (II, III and IV) and Congo red adulterant were purchased from Sigma Aldrich (St. Louis, MO, USA).

2.2. Instrumentation and software

The analysis of paprika samples was carried out using a Multi Purpose Analyzer (MPA) FT-NIR spectrometer from Bruker Optik GmbH (Ettlingen, Germany) using the solid sample compartment. The system was in combination with OPUS software version 8.5 for the acquisition of the spectra.

Spectral pre-treatment and PLS regressions were carried out by The Unscrambler X software version 10.4 from CAMO Software (Oslo, Norway). The construction of the OPLS-DA models was performed using the SIMCA software (Umetrics, Sartorius Stedim Biotech AS, Umea, Sweden) version 14.1. Python 3.8.8 was used to apply the Kennard-Stone algorithm to the samples to split them into training and validation sets.

2.3. NIR analysis

Each paprika was adulterated with each dye (Sudan II, III, IV and Congo red) at concentration levels of 0.1, 1 and 5 g/100 g (w/w). Therefore, 1495 samples of paprika were analyzed by the proposed methodology (115 pure and 1380 adulterated samples). Each mixture (0.4 g of sample) was homogenized, placed in a glass vial and immediately measured by the FT-NIR system using the sphere macrosample acquisition mode. A total of 10 aliquots of each paprika sample were collected and analyzed.

Data were collected in reflectance mode with spectral data output measured in absorbance units. Spectral data were acquired in a wavelength range from 12500 to 3600 cm-1 with 8 cm^{-1} resolution. For each spectrum, 32 scans were collected. A background scan was performed before the analysis of each different sample.

2.4. Multivariate processing

Firstly, the wavelength interval from 9000 to 3840 cm^{-1} was selected as the optimum range for chemometric analysis as it contained the characteristic NIR bands of the paprika samples. Higher wavelengths did not provide relevant information, thus were excluded from further processing. The spectra of the 10 aliquots measured for each paprika sample were averaged in order to use a representative spectrum.

The validation procedure was carried out in accordance with the recommendations of (McGrath et al., 2018; Riedl et al., 2015). Two types of chemometric models were performed for paprika authentication: OPLS-DA models for the discrimination of adulterated samples as well as their differentiation according to their PDO and variety, and PLS regressions for each adulterant to determine the dye content in adulterated paprika samples. Chemometric models were performed using the

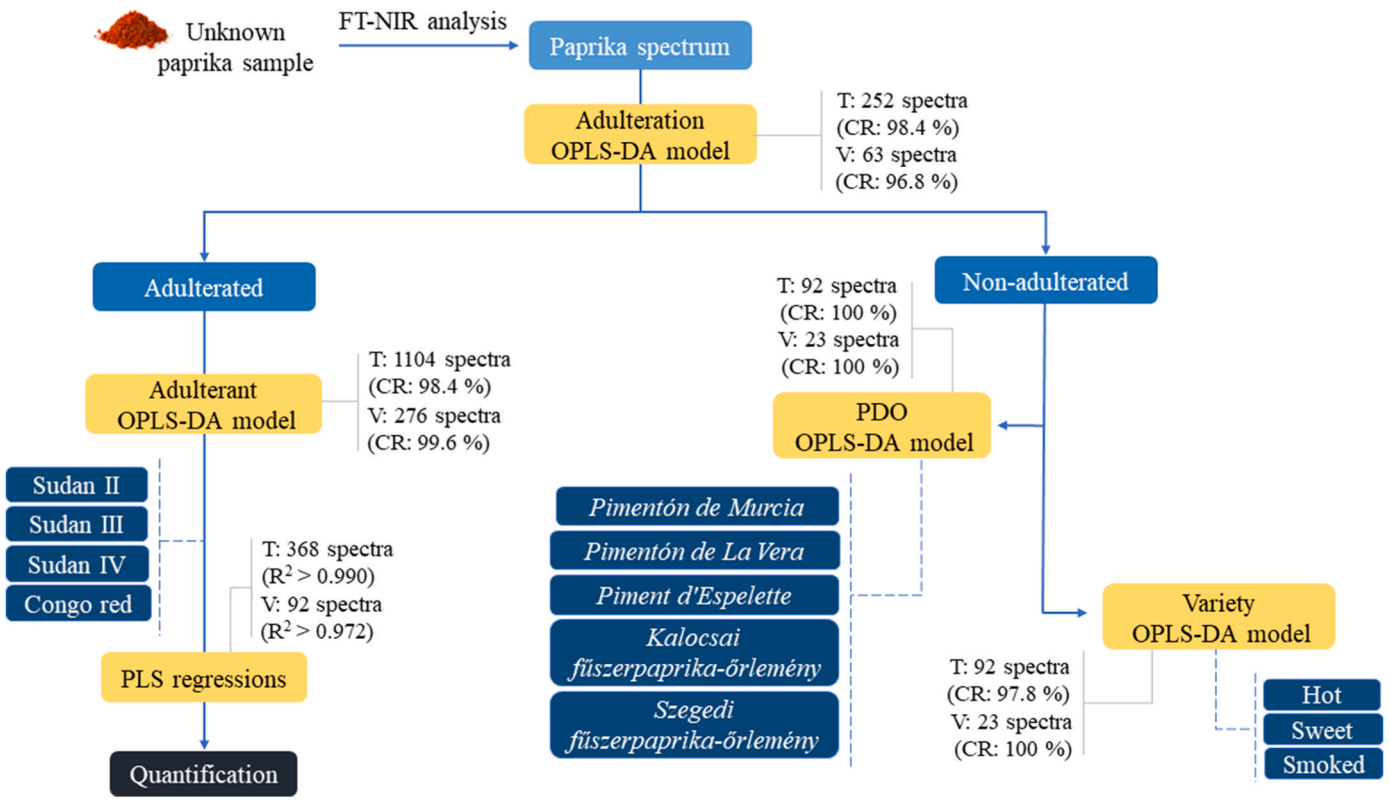

Fig. 1. Proposed strategy based on FT-NIR analysis and chemometric models for paprika authentication (T: training set; V: validation set; CR: classification rate).

80 % of paprika samples for the model training and the remaining 20 % were used for the validation. For the sample separation into both sets, the Kennard-Stone algorithm was applied. This algorithm aims to cover the design space as uniformly as possible by first selecting the two samples with the largest distance to be included in the training set. Next, the distances between the remaining samples and the two already selected samples are calculated. For each candidate, the minimum distance to the two selected samples is chosen. Then, the sample with the maximum value of the minimum distances (maximum-minimum distance criterion) is selected for the training set until the requested number of samples is obtained (Kennard & Stone, 1969). The processed spectra of validation and training sets used for the construction of each chemometric model can be seen in Figs. S1 and S2.

The pre-processing of the spectra was carried out consisting of a baseline correction using the "baseline offset" function to remove interference and background information followed by a normalization of the data using the "unit vector normalization". Savitzky-Golay smoothing using a window length of 3 points and a polynomial order 2 was subsequently applied to reduce the instrumental noise. PLS regressions were performed individually for each adulterant using the three percentages of adulteration (0.1, 1 and 5 %) and pure samples. The suitability of these models was evaluated in terms of statistical parameters as the correlation and coefficient of determination (R^2) to check the linearity, the root-mean-square error (RMSE) to measure the dispersion of residual values, and the standard error (SE) and bias of calibration and validation sets of samples to measure the accuracy of the model.

OPLS-DA models were built exploring six different scales: unit variance (UV), pareto (Par), centring (Ctr), unit variance none (UVN), pareto none (ParN) and freeze (Arroyo-Manzanares et al., 2019). The values of R2X (cum), R2Y (cum) and Q2 (cum) were used to evaluate the adequacy of the models. R2Y and Q2 are relevant parameters indicating the accuracy and the predictive ability, respectively. They all range between 0 and 1, with values closer to 1 indicating best model fitness. The model is acceptable when the value of Q2 is above 0.5 (Bajoub et al., 2016; Wang et al., 2014). Prior to model building, a normal probability plot of residuals was performed to verify the normal distribution of the data, and therefore, no logarithmic transformation was necessary. The sensitivity of the models was expressed as $\sum$ True positive/($\sum$ True positive + $\sum$ False negative) × 100.

2.5. Proposed strategy for paprika authentication

An overview of the proposed strategy is shown in Fig. 1. The first step is the analysis of the investigated paprika sample using the FT-NIR system under the conditions indicated in section 2.3. The analysis does not require a previous sample treatment and takes only a few seconds. As a result, the NIR spectrum characteristic of the sample is obtained. The spectra of paprika in the wavelength range of 9000 to 3840 cm^{-1} were used to build the chemometric models since it contained the characteristic bands of the samples. The first proposed OPLS-DA model allows to detect whether the investigated sample is adulterated with dyes; otherwise, the paprika sample is classified as pure. If the sample is considered as pure, two OPLS-DA models are proposed to guarantee their authenticity. Paprika is classified according to their variety (hot, sweet or smoked) and their PDO in order to ensure its quality certification and avoid fraud by false designation of origin. On the other hand, if paprika is identified as adulterated, the OPLS-DA model of classification according to the type of adulterant allows to assign the dye present in the sample. The last proposed step consists in determining the amount of dye in the investigated sample, using PLS regression models constructed for this purpose. In summary, the proposed strategy consists of four OPLS-DA models, and four PLS regression models (one for each adulterant) that would guarantee the authenticity of paprika in terms of its registered PDO and adulteration by adding dyes with the identification and quantification of adulterated samples.

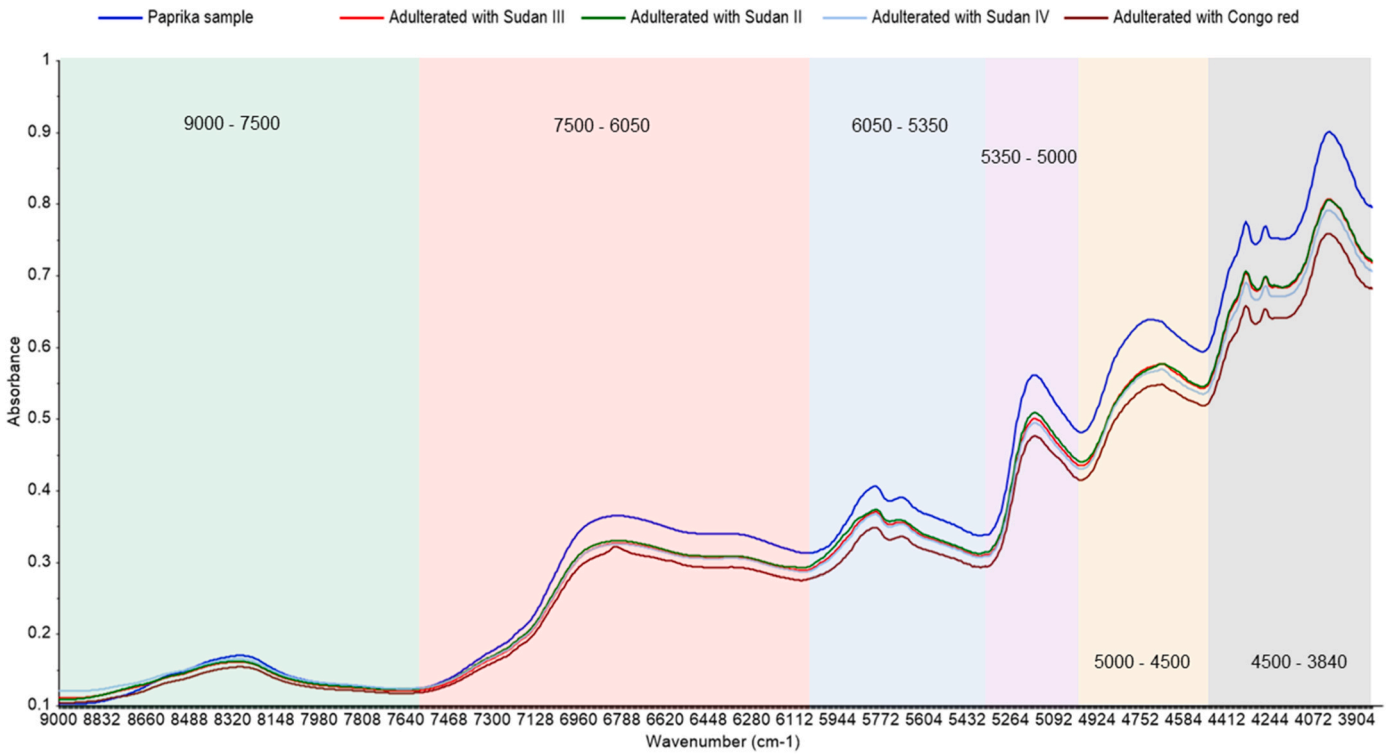

Fig. 2. Raw FT-NIR spectra of a pure sample of paprika and adulterated with Sudan II, III, IV and Congo red at 5 % (w/w).

2.6. Classification analysis

For the construction of the first proposed OPLS-DA model, which allows the differentiation between pure and adulterated paprika samples, a total of 315 samples (115 pure and 200 adulterated) were used. Regarding adulterated samples, 200 were selected to ensure equal groups: 50 of each adulterant (Sudan II, III, IV and Congo red) including the three percentages of adulteration (0.1, 1 and 5 %). Samples were partitioned by the Kennard-Stone method into the two sets. The training set was composed of 252 samples and the remaining 63 samples were used as the validation set.

The paprika samples classified as pure can belong to paprika with or without PDO. The OPLS-DA model for the PDO certification was composed of all pure paprika samples with PDO from "Murcia", "La Vera", "Espelette", "Szeged" and "Kalocsa", as well as paprikas from China and Zambia. Most samples were from "Murcia" (38 samples), 27 were from "La Vera", 14 from "Espelette", 7 from "Szeged", 9 from "Kalocsa", 12 from China and 8 from Zambia. Samples were classified according to their PDO and China and Zambia paprikas were labelled as "Others". Thus, 115 samples divided into training (92) and validation (23) sets were finally used. In addition, an OPLS-DA model was built to classify pure samples according to their variety into sweet, hot and smoked paprika. Paprika samples analyzed were 26 of the hot variety, 38 of the smoked variety and 51 of the sweet variety. The training set consisted of 92 samples, and the remaining 23 samples were used for the validation.

On the other hand, the OPLS-DA model proposed for the identification of the adulterant present in paprika samples was composed of 1380

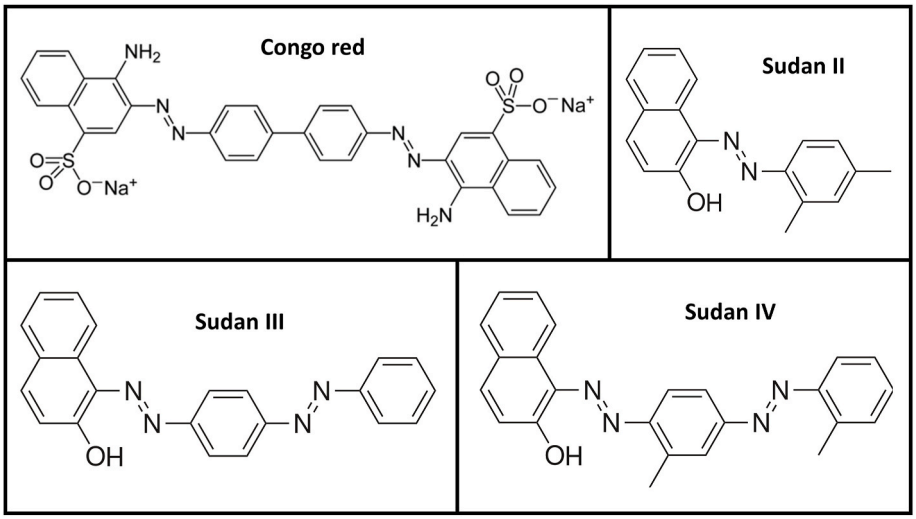

Fig. 3. Chemical structures of the Congo red, Sudan II, Sudan III, and Sudan IV molecules.

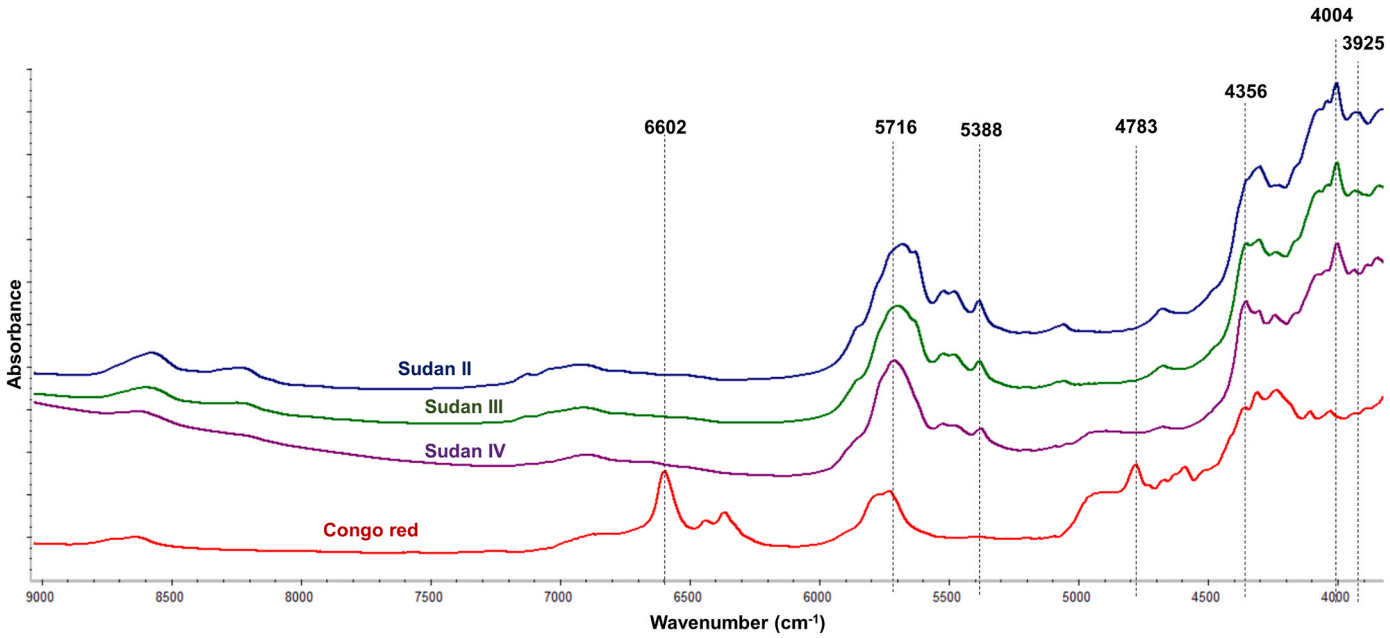

Fig. 4. FT-NIR spectra of Sudan II, III, IV and Congo red.

spectra, being 345 of each dye (Sudan II, III, IV and Congo red). Model training was carried out using 1104 samples and the remaining (276) were used for the validation. Finally, once adulterated paprika samples are classified according to their adulterant, PLS regressions allow to quantify their content in the sample. Pure and adulterated paprika samples with the three different concentrations tested were used to perform the regressions. Each model was composed of a total of 460 samples (115 pure and 345 adulterated), which were split into the two calibration (368) and validation (92) sets.

Finally, the proposed chemometric strategy was evaluated by analyzing a set of 50 paprika samples from different classes (20 pure samples and 30 adulterated samples).

3. Results and discussion

3.1. NIR spectra interpretation

Fig. 2 shows the NIR spectra in the selected region 9000-3840 cm^{-1} of a pure sample of paprika and adulterated with Sudan II, Sudan III, Sudan IV and Congo red at 5 %. All of them show a similar pattern, differing in the absorbance values. Six regions can be distinguished in the spectrum of paprika according to the NIR bands: 9000-7500 cm^{-1}, 7500-6050 cm^{-1}, 6050-5350 cm^{-1}, 5350-5000 cm^{-1}, 5000-4500 cm^{-1} and 4500-3840 cm^{-1}. The first region (9000-7500 cm^{-1}) shows a wide band at 8250 cm^{-1} and a slight shoulder at 8600 cm^{-1} approximately, which are both due to the second overtone of the stretching mode of the *C*–H bond. Bands observed in the next region (7500-6050 cm^{-1}) are associated to the first overtone of the stretching modes of water molecule and OH groups, and to the *C*–H second combination region (which combines stretching and vibrational modes). The two bands located within the 6050-5350 cm^{-1} region are mainly attributed to the first overtone of *C*–H stretching modes either from CH_3, CH_2, aliphatic CH, or aromatic CH groups. A band corresponding to the first combination region of water appears in the 5350-5000 cm^{-1} region. Bands appearing in the 5000-4500 cm^{-1} region are attributed to the first combination region of *C*–H aliphatic bonds (stretching + bending modes), while the bands located within the 4500-3840 cm^{-1} region are due to the first combination region of CH_3, CH_2, and aromatic *C*–H stretching and bending modes (Westad et al., 2008; Workman & Weyer, 2012).

Regarding the spectra of adulterated paprika samples, they were visually compared in order to examine whether the slight spectral differences among them could be correlated to the different chemical structure of the azo dye adulterant used in each case. As displayed in Fig. 3, the chemical structures of Congo red, Sudan II, Sudan III, and Sudan IV molecules share an azo group bonded (on one side) to naphthyl and (on the other side) to phenyl ring. Sudan II, Sudan III and Sudan IV additionally share a OH group in position 2 of the naphthyl ring, which does not appear in Congo red. Congo red has an amino and a sulphonate groups attached to the naphthyl ring instead. Sudan III and Sudan IV have an additional phenyl-azo group linked to the phenyl ring, while Sudan II does not. Finally, Sudan II and Sudan IV have methyl groups in their structure, whereas Sudan III does not. The theoretical NIR spectral differences expected for previous structural divergences (particularly those based on the NIR active vibrations) are summarized in Supplementary Table S2.

In brief, as evidence in Fig. 4, in which the NIR spectra of the four pure adulterants are visually compared, the NIR bands due to amino group, located at 6700-6400 cm^{-1} and 5100-4700 cm^{-1}, are exclusive to Congo red. The NIR bands of methyl group, specific to Sudan II and Sudan IV, are not evident to the naked eye. On the contrary, the NIR bands due to OH group, which are common to Sudan II, Sudan III, and Sudan IV, are clearly observed within the 5750-5300 cm^{-1} and 4050-3900 cm^{-1} regions. These spectral differences are easily recognizable in the NIR spectra of the pure azo dye standards. However, they are less evident in the NIR spectra of adulterated paprika samples, in which the adulterant mass percentage ranges from 0.1 to 5 % of the composition. As visually observed in Supplementary Fig. S3, the NIR spectrum is mostly due to the major component (*i.e.*, paprika at 95–99.9 %), whereas the characteristic NIR bands of adulterants appear as little intense bands and shoulders. Regarding the NIR spectra of adulterated paprika samples with Sudan II, Sudan III and Sudan IV, slight unspecific baseline spectral differences are observed between the pure paprika and the adulterated paprika samples, especially within the 5750-5550 cm^{-1} region due to the OH group of Sudan azo dyes. Regarding the Congo red adulterated paprika, the NIR bands of amino group (NH_2) from Congo red are visually identifiable as little intense shoulders in the NIR spectra

Table 1
Data treatment tested and chemometric model information for paprika authentication.

	Model information				Classification rate (%)		Sensitivity (%)
Scaling[a]	Components	R2X(cum)	R2Y(cum)	Q2(cum)	Calibration	Validation	Validation[b]
OPLS-DA model for the detection of adulterated paprika							
UV	1 + 13+0	0.998	0.784	0.740	98.4	96.8	P (92.9), A (100)
UVN	1 + 14+0	1	0.716	0.647	97.6	96.8	P (92.9), A (100)
Par	1 + 11+0	0.996	0.746	0.703	96.4	92.1	P (85.7), A (97.1)
ParN	1 + 11+0	1	0.738	0.696	97.2	95.2	P (92.8), A (97.1)
Ctr	1 + 12+0	0.997	0.765	0.715	98.0	93.6	P (89.3), A (97.1)
Freeze	1 + 11+0	0.997	0.747	0.703	96.4	90.5	P (82.1), A (97.1)
OPLS-DA model for the classification of paprika according to their PDO							
UV	5 + 11+0	1	0.903	0.837	100	100	M (100), LV (100), S (100), K (100), E (100), O (100)
UVN	6 + 9+0	1	0.853	0.724	100	100	M (100), LV (100), S (100), K (100), E (100), O (100)
Par	6 + 7+0	0.999	0.845	0.733	98.9	100	M (100), LV (100), S (100), K (100), E (100), O (100)
ParN	6 + 7+0	1	0.840	0.734	98.9	100	M (100), LV (100), S (100), K (100), E (100), O (100)
Ctr	5 + 10+0	1	0.882	0.804	100	100	M (100), LV (100), S (100), K (100), E (100), O (100)
Freeze	5 + 11+0	1	0.903	0.837	100	100	M (100), LV (100), S (100), K (100), E (100), O (100)
OPLS-DA model for the classification of paprika according to their variety							
UV	3 + 12+0	0.999	0.832	0.639	97.8	100	H (100), SW (100), SM (100)
UVN	3 + 13+0	1	0.820	0.547	95.7	95.7	H (83.3), SW (100), SM (100)
Par	2 + 12+0	1	0.764	0.616	95.7	100	H (100), SW (100), SM (100)
ParN	3 + 10+0	1	0.728	0.611	94.6	95.7	H (100), SW (90), SM (83.3)
Ctr	2 + 10+0	0.999	0.713	0.611	89.1	91.3	H (83.3), SW (100), SM (83.3)
Freeze	2 + 10+0	0.999	0.711	0.625	93.5	95.7	H (100), SW (100), SM (83.3)
OPLS-DA model for the classification of adulterated paprika according to their adulterant							
UV	3 + 9+0	0.996	0.859	0.854	98.4	99.6	SII (100), SIII (98.3), SIV (100) CR (100)
UVN	3 + 10+0	1	0.834	0.828	98.2	99.3	SII (98.4), SIII (98.3), SIV (100) CR (100)
Par	3 + 9+0	0.996	0.863	0.858	98.5	99.6	SII (100), SIII (98.3), SIV (100) CR (100)
ParN	3 + 9+0	1	0.834	0.830	98.0	99.6	SII (100), SIII (98.3), SIV (100) CR (100)
Ctr	3 + 9+0	0.997	0.855	0.851	98.2	99.6	SII (100), SIII (98.3), SIV (100) CR (100)
Freeze	3 + 9+0	0.996	0.859	0.854	98.4	99.6	SII (100), SIII (98.3), SIV (100) CR (100)

[a] Unit variance (UV), unit variance none (UVN), pareto (Par), pareto none (ParN), centering (Ctr).
[b] Classes: pure (P), adulterated (A), Murcia (M), La Vera (LV), Szeged (S), Kalocsa (K), Espelette (E), Others (O), hot (H), sweet (SW), smoked (SM), Sudan II (SII), Sudan III (SIII), Sudan IV (SIV), Congo red (CR).

of adulterated paprika, including the first overtone of NH_2 stretching located at 6600 cm^{-1}, and the first combination region of NH_2 located around 5000-4700 cm^{-1}. These visual evidence are insufficient to confidently identify adulterated paprika samples, especially for low adulterated samples. Statistical and mathematical analyses are required for an effective and confident discrimination of paprika samples.

3.2. Authentication strategy based on chemometric models

In order to ensure the authenticity of paprika, this work proposes a strategy based on the use of different chemometric models, which allow the detection and quantification of adulterated paprika with dyes or the classification of the samples according to their PDO and variety (Fig. 1).

OPLS-DA models were investigated using different scales (UV, Ctr, UVN, Par, Freeze, ParN) as pre-treatment methods. Table 1 summarizes the results obtained for each OPLS-DA model, including the R2X (cum), R2Y (cum) and Q2 (cum) values and the classification rate of training and validation sets, including the percentage of sensitivity achieved in each category.

UV and Pareto scaling are commonly used as spectral data processing methods. The UV scale analyzes the data based on correlations, using the standard deviation as the scaling factor, while Pareto uses the square root, getting closer to the original data (Lee et al., 2018; Van den Berg et al., 2006). As can be seen in Table 1, all the scales tested provided suitable values of classification rates and Q2 values higher than 0.5. Generally, the highest values were obtained using the UV scale, therefore, it was selected as the optimum scale to carried out the proposed strategy for paprika authentication.

3.2.1. Detection of adulterated paprika samples

The first proposed OPLS-DA model allows the differentiation between pure and adulterated paprika samples. Model training was carried out using 252 pure and adulterated paprika samples at the different concentrations tested, while 63 were used for model validation. All the scales tested as pre-treatment methods provided good classification results, with classification success rates greater than 90 % in all cases. The validation of the model was 96.8 % successful, with all adulterated samples correctly classified (100 %). Fig. 5a shows the class distinction using the UV scale, which provided a Q2 value of 0.740, demonstrating the high accuracy of the model. Only two pure samples were misclassified as adulterated in the validation set (Supplementary Table S3). Therefore, it can be assumed that the proposed model would be a reliable tool for the identification of adulterated paprika.

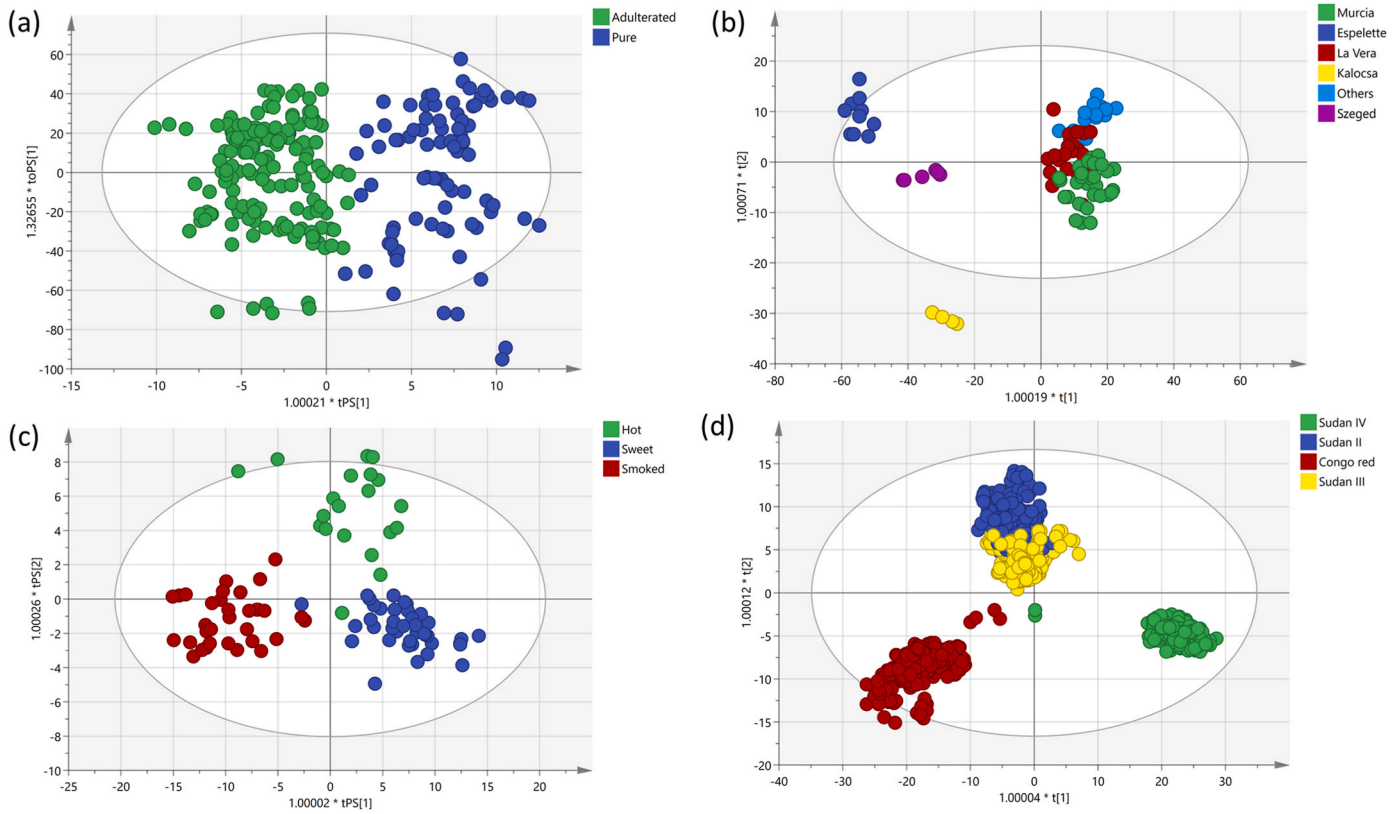

Fig. 5. OPLS-DA models for the discrimination between pure and adulterated paprika samples (a), the differentiation according to PDO (b), the classification according to their variety (c), and the classification of adulterated samples according to their adulterant (d) using a UV scale.

3.2.2. *Classification of paprika according to their PDO*

The origin of pure paprika samples can be differentiated using the following proposed OPLS-DA model. Samples were classified according to their PDO ("Murcia", "La Vera", "Espelette", "Szeged" and "Kalocsa"), and China and Zambia paprikas were labelled as "Others". Model training was composed of 92 samples and validation of 23 samples. Classification rates for training and validation were higher than 98.9 and 100 %, respectively. The selected model based on the UV scale provided a Q2 value of 0.837. The classification graph of the different classes is shown in Fig. 5b. Regarding the validation of the model, all samples were correctly classified according to their PDO (Supplementary Table S4).

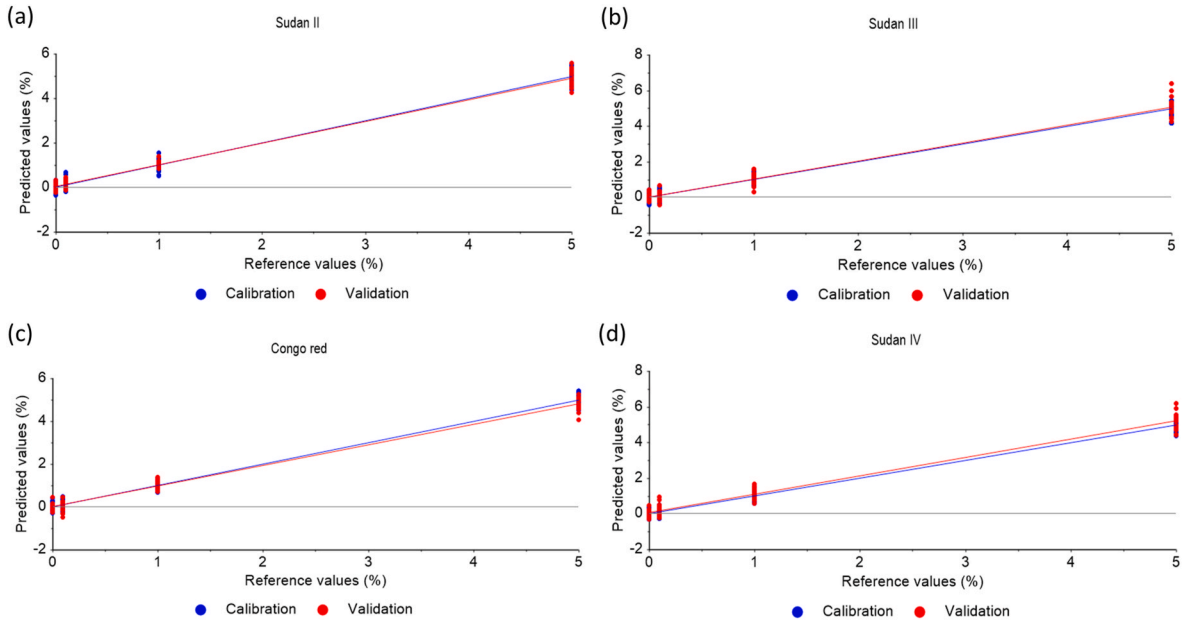

Fig. 6. PLS models selected for Sudan II (a), Sudan III (b), Congo red (c) and Sudan IV (d), including calibration and validation regressions.

Table 2
Summary of calibration and validation parameters in PLS models for each adulterant.

	Slope		R^2		Correlation		RMSEC	RMSEP	SEC	SEP	Bias	
	Cal	Val	Cal	Val	Cal	Val					Cal	Val
Sudan II	0.993	0.971	0.992	0.988	0.996	0.994	0.171	0.224	0.171	0.225	−1.8E-06	0.003
Sudan III	0.992	1.01	0.990	0.974	0.995	0.987	0.203	0.316	0.204	0.317	4.8E-06	0.025
Sudan IV	0.993	1.029	0.991	0.972	0.995	0.988	0.184	0.332	0.184	0.316	−8.7E-07	0.109
Congo red	0.996	0.959	0.994	0.986	0.997	0.993	0.145	0.231	0.145	0.229	−7.4E-06	−0.039

Cal: calibration set; Val: validation set.
R^2: coefficient of determination; RMSEC: root-mean-square error of calibration; RMSEP: root-mean-square error of prediction; SEC: standard error of calibration; SEP: standard error of prediction.

Table 3
Classification of test samples in the OPLS-DA models proposed for paprika authentication.

Sample	Classification										
	Pure										
	Pure	Adulterated	Murcia	La Vera	Espelette	Szeged	Kalocsa	Others	Hot	Sweet	Smoked
1	+						+				+
2	+			+						+	
3	+		+							+	
4	+							+			+
5	+				+				+		
6	+		+							+	
7	+				+					+	
8	+		+						+		
9		–			+				+		
10	+			+						+	
11	+						+			+	
12	+						+			+	
13	+			+						+	
14	+							+			+
15	+							+	+		
16	+		+								–
17	+		+								+
18	+						+			+	
19	+					+				+	
20	+			+						+	

Adulterated					
	Pure	Adulterated	Sudan II	Sudan III	Sudan IV	Congo Red
21		+				+
22		+		+		
23		+		+		
24		+	+			
25		+		+		
26		+			+	
27		+	+			
28		+				+
29		+				+
30		+		+		
31		+	+			
32		+			+	
33		+		+		
34		+				+
35		+	+			
36		+				+
37		+				+
38		+		+		
39		+	+			
40		+	+			
41		+				+
42		+		+		
43		+			+	
44		+	+			
45		+		+		
46		+				+
47		+				+
48		+			+	
49		+		+		
50		+	+			

The "+" sign means correctly classified.
The "-" sign means incorrectly classified.

3.2.3. Classification of paprika according to their variety

Pure paprika samples were classified using this OPLS-DA model into sweet, hot and smoked paprika. Suitable calibration and validation rates were achieved, higher than 89.1 and 91.3 %, respectively. A Q2 value of 0.639 was achieved using the selected UV scale. Fig. 5c shows the class separation as hot, sweet or smoked paprika. The classification rate achieved was 97.8 %, and all the hot, smoked and sweet samples were classified successfully during validation (100%) (Supplementary Table S5).

3.2.4. Classification of adulterated samples according to the type of adulterant

An OPLS-DA model was proposed for the identification of the adulterant present in paprika samples classified as non-pure. Model training was carried out using 1104 samples and the remaining (276) were used for the validation. In this case, the classification and validation rates achieved were in all cases higher than 98.2 and 99.3 %, respectively. The UV scale pre-treatment achieved a Q2 value of 0.854, demonstrating the suitability of the model. The classification graph is shown in Fig. 5d. Visually, samples adulterated with Sudan IV are clearly separated, while paprika samples adulterated with Sudan II and Sudan III were slightly overlapping. Moreover, Congo red samples showed a slight distinction between them, due to the different percentages of adulteration. As could also be observed spectrally, the paprika samples adulterated with 5 % of Congo red clearly showed the NIR band of the amino group (Fig. S3), while in the samples with lower adulteration, these bands may be overlapped by the paprika bands themselves. The low adulteration of the samples and, therefore, the slight spectral differences influenced this chemometric classification. Nevertheless, the chemometric model created allowed the successfully classification of the samples according to their adulterant, achieving a 98.4 % of success rate of model calibration and 99.6 % of the validation. The paprika samples adulterated with Sudan II, Sudan IV and Congo red were all correctly classified. Only one adulterated paprika with Sudan III was misclassified as Sudan II (Supplementary Table S6).

3.2.5. Quantification of adulteration in paprika samples

Once adulterated paprika samples are classified according to their adulterant, PLS regressions allow to quantify their content in the sample. For this purpose, after pre-treatment data based on baseline correction, normalization and smoothing of the spectra, PLS regressions were built for each dye. Pure and adulterated paprika samples with the three different concentrations tested were used to perform the regressions. Fig. 6 shows the PLS regressions of Sudan II, Sudan III, Sudan IV and Congo red, including calibration and validation sets for each dye. Models were evaluated in terms of the suitability of R^2, RMSE, SE and bias parameters. The number of factors used in these models was also investigated as it influences the values of the RMSE and R^2. The optimal number was considered the factor that provides the minimum error value avoiding overfitting and noise in the model (Sadergaski et al., 2022). Factors were selected by observing the plots of RMSE and variance versus the number of factors.

In all cases, good correlations were achieved between the reference and predicted adulteration values, being greater than 0.98. Suitable linearities with values of R^2 in calibration and validation sets higher than 0.97 were obtained for all the models. Calibration (RMSEC and SEC) and validation (RMSEP and SEP) errors, as well as bias, were closed to zero, demonstrating the accuracy of the PLS regressions. The results achieved for each model are summarized in Table 2.

3.3. Application of the proposed strategy

The proposed chemometric strategy was evaluated by analyzing a randomly set of 50 paprika samples from different classes. The classification of the samples in each category is shown in Table 3. A total of 44 paprikas were successfully classified and categorized according to their features. Only one pure sample (number 9) was misclassified as adulterated but correctly classified in the rest of the OPLS-DA models, and one hot paprika sample from Murcia was wrongly classified in the smoked category (number 16). Regarding adulterated spectra, all of them were correctly classified in their category and the adulterant was properly identified. Furthermore, adulterated samples were introduced into the PLS models to predict their adulterant concentration. Trueness results ranged from 89 to 107 %. The samples showing values furthest from the optimum corresponded to an adulteration of 0.1 %, while the 5 and 1 % adulterations showed concentration prediction values closer to the real values.

4. Conclusion

The chemometric strategy proposed in this work to ensure the authenticity of paprika has proven to be an effective tool. The set of chemometric models built for this purpose (4 OPLS-DA and 4 PL S regressions) have been constructed and validated, obtaining satisfactory classification and quantification results. Regarding OPLS-DA models, validation rates achieved ensure the accurate identification of adulterated paprika (96.8 %), the assignment of the type of adulterant (99.6 %) and the classification of pure samples according to their PDO label (100 %) and variety (100 %). As for the PLS regressions built for Sudan II, Sudan III, Sudan IV and Congo red, the error values close to zero and the good linearity obtained in all cases, with values of R^2 higher than 0.97, allow to quantify the adulterant content in paprika efficiently. Furthermore, FT-NIR technique allows a non-destructive, fast and simple analysis of the samples, avoiding the use of any solvent. Therefore, this work proposes a comprehensive, fast and efficient method to prevent paprika fraud and to guarantee its quality and safety for consumption.

CRediT authorship contribution statement

Ana Castell: Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Natalia Arroyo-Manzanares:** Writing – review & editing, Validation, Supervision, Software, Methodology, Investigation, Conceptualization. **Ignacio López-García:** Writing – review & editing, Supervision, Software, Project administration, Investigation, Funding acquisition. **Félix Zapata:** Writing – original draft, Software, Methodology, Investigation. **Pilar Viñas:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors acknowledge the financial support of the Spanish MCIN (Project PID 2021-123201NB-I00 financed by MCIN/AEI/10.13039/501100011033/FEDER, EU). A. Castell acknowledges a fellowship from University of Murcia (AEIP 2020-0005).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodcont.2024.110397.

References

Arroyo-Manzanares, N., Markiv, B., Hernández, J. D., López-García, I., Guillén, I., Vizcaíno, P., Hernández-Córdoba, M., & Viñas, P. (2019). Head-space gas chromatography coupled to mass spectrometry for the assessment of the contamination of mayonnaise by yeasts. *Food Chemistry, 289*, 461–467. https://doi.org/10.1016/j.foodchem.2019.03.083

Baenas, N., Belović, M., Ilic, N., Moreno, D. A., & García-Viguera, C. (2019). Industrial use of pepper (Capsicum annum L.) derived products: Technological benefits and biological advantages. *Food Chemistry, 274*, 872–885. https://doi.org/10.1016/j.foodchem.2018.09.047

Bajoub, A., Pacchiarotta, T., Hurtado-Fernández, E., Olmo-García, L., García-Villalba, R., Fernández-Gutiérrez, A., Mayboroda, O. A., & Carrasco-Pancorbo, A. (2016). Comparing two metabolic profiling approaches (liquid chromatography and gas chromatography coupled to mass spectrometry) for extra-virgin olive oil phenolic compounds analysis: A botanical classification perspective. *Journal of Chromatography A, 1428*, 267–279. https://doi.org/10.1016/j.chroma.2015.10.059

Barbosa, S., Campmajó, G., Saurina, J., Puignou, L., & Núñez, O. (2020). Determination of phenolic compounds in paprika by ultrahigh performance liquid chromatography-tandem mass spectrometry: Application to product designation of origin authentication by chemometrics. *Journal of Agricultural and Food Chemistry, 68*(2), 591–602. https://doi.org/10.1021/acs.jafc.9b06054

Campmajó, G., Rodríguez-Javier, L. R., Saurina, J., & Núñez, O. (2021). Assessment of paprika geographical origin fraud by high-performance liquid chromatography with fluorescence detection (HPLC-FLD) fingerprinting. *Food Chemistry, 352*. https://doi.org/10.1016/j.foodchem.2021.129397

Cetó, X., Sánchez, C., Serrano, N., Díaz-Cruz, J. M., & Núñez, O. (2020). Authentication of paprika using HPLC-UV fingerprints. *Lwt, 124*. https://doi.org/10.1016/j.lwt.2020.109153

Cetó, X., Serrano, N., Aragó, M., Gámez, A., Esteban, M., Díaz-Cruz, J. M., & Núñez, O. (2018). Determination of HPLC-UV fingerprints of Spanish paprika (Capsicum annuum L.) for its classification by linear discriminant analysis. *Sensors, 18*(12). https://doi.org/10.3390/s18124479

Di Anibal, C. V., Marsal, L. F., Callao, M. P., & Ruisánchez, I. (2012). Surface Enhanced Raman Spectroscopy (SERS) and multivariate analysis as a screening tool for detecting Sudan i dye in culinary spices. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 87*, 135–141. https://doi.org/10.1016/j.saa.2011.11.027

Di Anibal, C. V., Odena, M., Ruisánchez, I., & Callao, M. P. (2009). Determining the adulteration of spices with Sudan I-II-II-IV dyes by UV-visible spectroscopy and multivariate classification techniques. *Talanta, 79*(3), 887–892. https://doi.org/10.1016/j.talanta.2009.05.023

Di Anibal, C., Rodriguez, M. S., & Albertengo, L. (2014). UV-Visible spectroscopy and multivariate classification as a screening tool to identify adulteration of culinary spices with Sudan I and blends of Sudan I + IV dyes. *Food Analytical Methods, 7*(5), 1090–1096. https://doi.org/10.1007/s12161-013-9717-2

Di Anibal, C. V., Rodríguez, M. S., & Albertengo, L. (2015). Synchronous fluorescence and multivariate classification analysis as a screening tool for determining Sudan I dye in culinary spices. *Food Control, 56*, 18–23. https://doi.org/10.1016/j.foodcont.2015.03.010

Di Anibal, C. V., Ruisánchez, I., & Callao, M. P. (2011). High-resolution 1H nuclear magnetic resonance spectrometry combined with chemometric treatment to identify adulteration of culinary spices with Sudan dyes. *Food Chemistry, 124*(3), 1139–1145. https://doi.org/10.1016/j.foodchem.2010.07.025

Fiamegos, Y., Dumitrascu, C., Papoci, S., & de la Calle, M. B. (2021). Authentication of PDO paprika powder (Pimentón de la Vera) by multivariate analysis of the elemental fingerprint determined by ED-XRF. A feasibility study. *Food Control, 120*, Article 107496. https://doi.org/10.1016/j.foodcont.2020.107496

Galaxy Scientific Inc. (2016). *Fast detection of paprika adulteration using FT-NIR spectroscopy*. https://www.azom.com/article.aspx?ArticleID=13251.

Galvin-King, P., Haughey, S. A., & Elliott, C. T. (2020). The detection of substitution adulteration of paprika with spent paprika by the application of molecular spectroscopy tools. *Foods, 9*(7). https://doi.org/10.3390/foods9070944

Gao, F., Hu, Y., Chen, D., Li-Chan, E. C. Y., Grant, E., & Lu, X. (2015). Determination of Sudan I in paprika powder by molecularly imprinted polymers-thin layer chromatography-surface enhanced Raman spectroscopic biosensor. *Talanta, 143*, 344–352. https://doi.org/10.1016/j.talanta.2015.05.003

Hayman, M., & Kam, P. C. A. (2008). Capsaicin: A review of its pharmacology and clinical applications. *Current Anaesthesia and Critical Care, 19*(5–6), 338–343. https://doi.org/10.1016/j.cacc.2008.07.003

Horn, B., Esslinger, S., Fauhl-Hassek, C., & Riedl, J. (2021). 1H NMR spectroscopy, one-class classification and outlier diagnosis: A powerful combination for adulteration detection in paprika powder. *Food Control, 128*, Article 108205. https://doi.org/10.1016/j.foodcont.2021.108205

Horn, B., Esslinger, S., Pfister, M., Fauhl-Hassek, C., & Riedl, J. (2018). Non-targeted detection of paprika adulteration using mid-infrared spectroscopy and one-class classification – is it data preprocessing that makes the performance? *Food Chemistry, 257*, 112–119. https://doi.org/10.1016/j.foodchem.2018.03.007

Hu, Y., Wang, S., Wang, S., & Lu, X. (2017). Application of nuclear magnetic resonance spectroscopy in food adulteration determination: The example of Sudan dye I in paprika powder. *Scientific Reports, 7*(1), 1–9. https://doi.org/10.1038/s41598-017-02921-8

Jahn, M., Patze, S., Bocklitz, T., Weber, K., Cialla-May, D., & Popp, J. (2015). Towards SERS based applications in food analytics: Lipophilic sensor layers for the detection of Sudan III in food matrices. *Analytica Chimica Acta, 860*, 43–50. https://doi.org/10.1016/j.aca.2015.01.005

Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics, 11*(1), 137–148. https://doi.org/10.1080/00401706.1969.10490666

Lee, L. C., Liong, C. Y., & Jemain, A. A. (2018). Effects of data pre-processing methods on classification of ATR-FTIR spectra of pen inks using partial least squares-discriminant analysis (PLS-DA). *Chemometrics and Intelligent Laboratory Systems, 182*, 90–100. https://doi.org/10.1016/j.chemolab.2018.09.001

Liu, F., Zhang, S., Wang, G., Zhao, J., & Guo, Z. (2015). A novel bifunctional molecularly imprinted polymer for determination of Congo red in food. *RSC Advances, 5*(29), 22811–22817. https://doi.org/10.1039/c4ra14719g

Lohumi, S., Joshi, R., Kandpal, L. M., Lee, H., Kim, M. S., Cho, H., Mo, C., Seo, Y. W., Rahman, A., & Cho, B. K. (2017). Quantitative analysis of Sudan dye adulteration in paprika powder using FTIR spectroscopy. *Food Additives & Contaminants Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 34*(5), 678–686. https://doi.org/10.1080/19440049.2017.1290828

Lohumi, S., Lee, H., Kim, M. S., Qin, J., Kandpal, L. M., Bae, H., Rahman, A., & Cho, B. K. (2018). Calibration and testing of a Raman hyperspectral imaging system to reveal powdered food adulteration. *PLoS One, 13*(4), 1–16. https://doi.org/10.1371/journal.pone.0195253

Márquez, C., Ruisánchez, I., & Callao, M. P. (2019). Qualitative and quantitative multivariate strategies for determining paprika adulteration with Sudan I and II dyes. *Microchemical Journal, 145*, 686–692. https://doi.org/10.1016/j.microc.2018.11.034

McGrath, T. F., Haughey, S. A., Patterson, J., Fauhl-Hassek, C., Donarski, J., Alewijn, M., van Ruth, S., & Elliott, C. T. (2018). What are the scientific challenges in moving from targeted to non-targeted methods for food fraud testing and how can they be addressed? – Spectroscopy case study. *Trends in Food Science and Technology, 76*, 38–55. https://doi.org/10.1016/j.tifs.2018.04.001

Mohamed, S. H., Salim, A. I., Issa, Y. M., Atwa, M. A., & Nassar, R. H. (2021). Evaluation of different Sudan dyes in Egyptian food samples utilizing liquid chromatography/tandem mass spectrometry. *Food Analytical Methods, 14*(10), 2038–2050. https://doi.org/10.1007/s12161-021-02036-x

Monago-Maraña, O., Durán-Merás, I., Muñoz de la Peña, A., & Galeano-Díaz, T. (2022). Analytical techniques and chemometrics approaches in authenticating and identifying adulteration of paprika powder using fingerprints: A review. *Microchemical Journal, 178*. https://doi.org/10.1016/j.microc.2022.107382

Monago-Maraña, O., Eskildsen, C. E., Afseth, N. K., Galeano-Díaz, T., Muñoz de la Peña, A., & Wold, J. P. (2019). Non-destructive Raman spectroscopy as a tool for measuring ASTA color values and Sudan I content in paprika powder. *Food Chemistry, 274*, 187–193. https://doi.org/10.1016/j.foodchem.2018.08.129

Monago-Maraña, O., Eskildsen, C. E., Galeano-Díaz, T., Muñoz de la Peña, A., & Wold, J. P. (2021). Untargeted classification for paprika powder authentication using visible – near infrared spectroscopy (VIS-NIRS). *Food Control, 121*. https://doi.org/10.1016/j.foodcont.2020.107564

Oliveira, M. M., Cruz-Tirado, J. P., & Barbin, D. F. (2019). Nontargeted analytical methods as a powerful tool for the authentication of spices and herbs: A review. *Comprehensive Reviews in Food Science and Food Safety, 18*(3), 670–689. https://doi.org/10.1111/1541-4337.12436

Oliveira, M. M., Cruz-Tirado, J. P., Roque, J. V., Teófilo, R. F., & Barbin, D. F. (2020). Portable near-infrared spectroscopy for rapid authentication of adulterated paprika powder. *Journal of Food Composition and Analysis, 87*. https://doi.org/10.1016/j.jfca.2019.103403

Palacios-Morillo, A., Jurado, J. M., Alcázar, A., & Pablos, F. (2016). Differentiation of Spanish paprika from protected designation of origin based on color measurements and pattern recognition. *Food Control, 62*, 243–249. https://doi.org/10.1016/j.foodcont.2015.10.045

Riedl, J., Esslinger, S., & Fauhl-Hassek, C. (2015). Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Analytica Chimica Acta, 885*, 17–32. https://doi.org/10.1016/j.aca.2015.06.003

Sadergaski, L. R., Hager, T. J., & Andrews, H. B. (2022). Design of experiments, chemometrics, and Raman spectroscopy for the quantification of hydroxylammonium, nitrate, and nitric acid. *ACS Omega, 7*(8), 7287–7296. https://doi.org/10.1021/acsomega.1c07111

Saini, R. K., Keum, Y. S., Daglia, M., & Rengasamy, K. R. (2020). Dietary carotenoids in cancer chemoprevention and chemotherapy: A review of emerging evidence. *Pharmacological Research, 157*, Article 104830. https://doi.org/10.1016/j.phrs.2020.104830

Sun, X., Zhang, M., Wang, P., Chen, J., Yang, S., Luo, P., & Gao, X. (2022). Detection and quantitation of adulterated paprika samples using second-order HPLC-FLD fingerprints and chemometrics. *Foods, 11*(15). https://doi.org/10.3390/foods11152376

Sun, X. D., Zhang, M., Zhang, S., Wang, P. J., Chen, J. H., & Gao, X. L. (2023). Non-targeted HPLC–FLD fingerprinting for the classification, authentication, and fraud quantitation of Guizhou paprika by chemometrics. *Journal of Food Composition and Analysis, 120*, Article 105346. https://doi.org/10.1016/j.jfca.2023.105346

Van Asselt, E. D., Banach, J. L., & Van der Fels-Klerx, H. J. (2018). Prioritization of chemical hazards in spices and herbs for European monitoring programs. *Food Control, 83*, 7–17. https://doi.org/10.1016/j.foodcont.2016.12.023

Van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., & Van der Werf, M. J. (2006). Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics, 7*, 1–15. https://doi.org/10.1186/1471-2164-7-142

Vera, D. N., Ruisánchez, I., & Callao, M. P. (2018). Establishing time stability for multivariate qualitative methods. Case study: Sudan I and IV adulteration in food spices. *Food Control, 92*, 341–347. https://doi.org/10.1016/j.foodcont.2018.04.057

Wang, Q., Gao, P., Wang, X., & Duan, Y. (2014). The early diagnosis and monitoring of squamous cell carcinoma via saliva metabolomics. *Scientific Reports, 4*, 1–9. https://doi.org/10.1038/srep06802

Westad, F., Schmidt, A., & Kermit, M. (2008). Incorporating chemical band-assignment in near infrared spectroscopy regression models. *Journal of Near Infrared Spectroscopy, 16*(3), 265–273. https://doi.org/10.1255/jnirs.786

Workman, J. J., & Weyer, L. (2012). *Practical guide and spectral atlas for interpretative Near-Infrared spectroscopy* (2nd ed.). Boca Raton: CRC Press.

Xu, H., Heinze, T. M., Chen, S., Cerniglia, C. E., & Chen, H. (2007). Anaerobic metabolism of 1-amino-2-naphthol-based azo dyes (Sudan dyes) by human intestinal microflora. *Applied and Environmental Microbiology, 73*(23), 7759–7762. https://doi.org/10.1128/AEM.01410-07

Zaukuu, J. L. Z., Bodor, Z., Vitalis, F., Zsom-Muha, V., & Kovacs, Z. (2019). Near infrared spectroscopy as a rapid method for detecting paprika powder adulteration with corn flour. *Acta Periodica Technologica, 50*, 346–352. https://doi.org/10.2298/APT1950346Z