Determining the adulteration of spices with Sudan I-II-II-IV dyes by UV–visible spectroscopy and multivariate classification techniques

Carolina V. Di Anibal [a], Marta Odena [b], Itziar Ruisánchez [a], M. Pilar Callao [a,*]

[a] Department of Analytical and Organic Chemistry, Rovira i Virgili University, Marcel·lí Domingo s/n Campus Sescelades, E-43007 Tarragona, Spain
[b] Public Health Laboratory, C/M, Cristina No. 54, 43002 Tarragona, Spain

ARTICLE INFO

ABSTRACT

We propose a very simple and fast method for detecting Sudan dyes (I, II, III and IV) in commercial spices, based on characterizing samples through their UV–visible spectra and using multivariate classification techniques to establish classification rules. We applied three classification techniques: *K*-Nearest Neighbour (KNN), Soft Independent Modelling of Class Analogy (SIMCA) and Partial Least Squares Discriminant Analysis (PLS-DA). A total of 27 commercial spice samples (turmeric, curry, hot paprika and mild paprika) were analysed by chromatography (HPLC-DAD) to check that they were free of Sudan dyes. These samples were then spiked with Sudan dyes (I, II, III and IV) up to a concentration of 5 mg L^{-1}. Our final data set consisted of 135 samples distributed in five classes: samples without Sudan dyes, samples spiked with Sudan I, samples spiked with Sudan II, samples spiked with Sudan III and samples spiked with Sudan IV.

Classification results were good and satisfactory using the classification techniques mentioned above: 99.3%, 96.3% and 90.4% of correct classification with PLS-DA, KNN and SIMCA, respectively. It should be pointed out that with SIMCA, there are no real classification errors as no samples were assigned to the wrong class: they were just not assigned to any of the pre-defined classes.

1. Introduction

Sudan I (1-[(2,4-dimethylphenyl)azo]-2-naphthalenol), Sudan II (1-(phenylazo)-2-naphthol), Sudan III (1-(4-phenylazophenylazo)-2-naphthol) and Sudan IV (*o*-tolyazo-*o*-tolyazo-betanaphthol) are an azo-family of synthetic dyes that are widely used for colouring agents such as waxes, floor and shoe polishes. They are categorized as class 3 carcinogens by the International Agency for Research on Cancer (IARC) [1]. As a result, Sudan dyes are illegal as additives in foodstuffs destined for human consumption according to both the FSA (Food Standards Agency) [2] and the European Union. Unfortunately, in some countries, these dyes are still being used as additives in some foodstuffs to improve the colour for commercial benefits. Reports have indicated that high amounts of Sudan dyes, at least 1 g L^{-1}, are required to have an impact on visual colour [3].

Several methods have been proposed to detect the presence of some of these synthetic dyes in foodstuffs: for example, high performance liquid chromatography-diode array detection HPLC-DAD [4,5], gel permeation chromatography–mass spectrometry [6], pressurized capillary electrochromatography (CEC) with amperometric detection [7], HPLC with electrochemical detection [8], liquid chromatography–mass spectrometry [9–11], tandem mass spectroscopy and isotope dilution [12], multi-wall carbon nanotube-based electrochemical sensing [13], and the ELISA method [14]. Also, second order multivariate techniques that use a data matrix for each sample have been applied to determine Sudan I in chilly powder [15]. In most of these methods, time-consuming pre-separations are often needed, analysis is not so fast and the instrumentation is no so affordable.

Multivariate techniques have been successfully applied to the analysis of foodstuffs. Some recent papers on this issue are: the multivariate prototype approach for authenticating red wines [16], the classification of milks according to their origin [17], the authentication of salmon and salmon-based products [18], characterisation and discrimination among butters according to their fat content [19], the authentication and classification of olive oil [20–22], the classification of vinegar [23] and the classification of apple fruits and ciders [24,25].

In this study, we propose a very simple and fast method for detecting Sudan dyes in commercial spices, based on characterizing samples through their UV–visible spectra and using multivariate classification techniques to establish classification rules. Five classes are considered: class 1: samples without Sudan dyes, class 2: samples spiked with Sudan I, class 3: samples spiked with Sudan II, class 4: samples spiked with Sudan III, and class 5: samples spiked with Sudan IV. Therefore, an unknown sample (the possible Sudan dye content of which is unknown) is assigned to one of the five classes considered in this study.

* Corresponding author. Tel.: +34 977558199; fax: +34 977558446.
E-mail address: mariapilar.callao@urv.cat (M.P. Callao).

Table 1
Type of errors made according to the decision taken satisfying the accomplishment of a null hypothesis.

	H_0 TRUE	H_0 FALSE
Statistical decision: reject H_0	*Type I error*	*Correct decision*
Statistical decision: do not reject H_0	*Correct decision*	*Type II error*

First of all, we applied an exploratory analysis based on the well-known principal component analysis (PCA) technique [26,27] in order to detect natural sample grouping with no previous information. Then, we applied classification techniques, both hard (discriminating) and soft (modelling) techniques. In the first, the hyperspace is divided in as many regions as the number of existing classes so, if a sample falls in the region of space corresponding to a particular class; it is classified as belonging to only this class. In the last, frontiers are built between each class and the rest of the space [28], modelling each class separately. The decision rule for a given class is a class box that envelopes the position of the class and an object can be assigned in more than one class or not assigned to a class [29]. We applied *K*-Nearest Neighbours (KNN) [30,31], Soft Independent Modelling of Class Analogy (SIMCA) [32,33] that is a soft classification technique and Partial Least Squares Discriminant Analysis (PLS-DA) that are hard classification techniques. PLS-DA is a variant of PLS [34] in which instead of predicting a quantitative parameter a qualitative assignation (class) is done [35–37].

We are dealing with a contamination problem, so it is important to consider the type of classification error because the associated cost and practical implications are different. From a practical point of view, then, stating that a sample does not contain Sudan dye when it does is not the same as stating that a sample contains Sudan dye when it does not. In this scenario, we will present the results on the basis of the following null hypothesis, H_0: "Sample does not contain Sudan dye". Therefore, as is shown in Table 1, type I and type II errors will be studied when each classification technique is applied. Type I errors might have economic implications as samples will be removed from the market with no real need. The consequences of type II errors are worse as they involve a health risk: consumers will buy samples contaminated with Sudan dye. Type I and II errors are related with sensitivity and specificity [38].

To the best of our knowledge, the one-step determination of the four Sudan dyes on the basis of their UV–visible spectra and multivariate classification has yet to be reported.

2. Experimental

2.1. Reagents and samples

Table 2 shows the chemical structure of the four Sudan dyes. It can be seen that all four Sudan dyes have some common structure but that Sudan III and IV have an additional azo and benzene group. Also, Sudan II and IV have two methyl groups more than Sudan I and III. The Sudan I standard was purchased from ACROS (Geel, Belgium) and the other Sudan dyes were purchased from SIGMA (St. Louis, MO, USA). A total of 27 spices from different commercial trade were purchased from markets. Acetonitrile and chloroform (for HPLC and UV–visible analysis) and acetic acid (for HPLC) were all of HPLC grade.

Samples contaminated with Sudan dye were obtained from the non-contaminated samples. To prepare the working solutions that were to be used to spike the commercial spices, appropriate amounts of Sudan I, II, III and IV were diluted in acetonitrile. Sudan III and Sudan IV were initially dissolved in a small fraction of chloroform and then diluted with acetonitrile.

The samples to be analysed, either by HPLC or UV–visible, followed an extraction process: one gram of sample was weighed and

Table 2
Chemical structure of Sudan dyes.

50 mL of acetonitrile was added. The sample was extracted with magnetic stirring. Each extract was obtained by filtering twice, first with glass microfibre filters and then with syringe filters. In order to obtain the UV–visible spectra, a volume of this extract (350 μL for mild and hot paprika, 300 μL for curry and 70 μL for turmeric) was spiked with the appropriate amount of one Sudan dye in such a way that the final concentration was 5 mg L^{-1}. Table 3 contains the list of the commercial samples analysed, the corresponding spiked samples and the numbers assigned to them all.

2.2. Apparatus and software

The HPLC is a Varian Star (Varian Inc., USA) model equipped with the following modules: a Prostar 240 pump; a Prostar 410 automatic injector and a Prostar 335 diode array detector (DAD). The relevant parameters were: column, Agilent Zorbax ODS (Agilent Technologies, USA) 250 mm × 4.6 mm (5 μm particle size); mobile phase, acetonitrile/acid acetic 16% (70:30, v/v) in gradient mode; flow rate 1.0 mL min^{-1}; temperature 40 °C. Photometric detection was performed at 478 nm for Sudan I and 510 nm for Sudan II, III and IV. The software used was the STAR 6.41 Chromatography Workstation.

The UV–visible spectrophotometer (Agilent 8453, United States) was equipped with a diode array detector (DAD). UV–visible spectrum scanning was carried out in the wavelength range of 260–600 nm (each nm) which represents 340 variables.

The data measured was processed with Matlab 6.5 software (Version 6.5, The Math Works Inc., Natick, USA) and PLS Toolbox 3.5 (Eigenvector Research Incorporated). Data was pre-processed with mean centering before each chemometric treatment.