

Rapid Detection of Paprika Adulteration by FT-NIR Spectroscopy

Introduction

There is a long history of spice adulteration, which dates back thousands of years. The main incentive behind the adulteration is economic: either reduction of cost or increase in perceived value. Most adulterants are harmless, such as adulteration with products from non-authentic geographic origins. However, some adulterants are deadly. Paprika, as an example, can be adulterated with tomato skins and brick dust, but it can also be adulterated with lead oxide and carcinogenic chemical dyes, such as Sudan I dye. In 1994, Hungarian ground paprika was found to be contaminated with lead oxide, which can easily dissolve in the hydrochloric acid present in our stomach, making it toxic upon ingestion. Several people died in that incident and dozens more were taken ill. In 2005, Sudan I dye was found in Worcestershire sauce contaminated by adulterated chili powder. Sudan I dye is known to be a rodent carcinogen and has been banned as a food additive.



Incidents such as these put the adulteration of food, including spice adulteration, at the top of the list when it comes to food safety concerns. Traditional analytical methods are comparatively expensive and time-consuming. On the other hand, although near-infrared (NIR) spectroscopy is not as sensitive as some other analytical methods, in the case of economic adulteration the concentrations of adulterants are often quite high, and it has the advantages of no sample preparation, high speed, and ease of use. To increase the sensitivity of the NIR method, a patent-pending algorithm has been developed specifically for screening at concentrations as low as 0.01%. In this study, the Advanced-ID™ algorithm with an FT-NIR spectrometer will be evaluated for paprika adulterated with tomato skin, red brick dust, and Sudan I dye.

Experimental

Materials

Four paprika samples were purchased from local supermarkets: McCormick Paprika, McCormick Gourmet Hot Hungarian Paprika, Morton & Bassett Paprika and Spice Chain Pride of Szeged Hungarian Style Paprika. Sudan I (dye content $\geq 95\%$) was purchased from Sigma Aldrich (St. Louis, MO). Red brick was obtained from Home Depot and ground into fine powder in the lab. Tomato was purchased from a local supermarket and its skin was peeled, dried and ground into fine powder. Sudan I dye, red brick dust and ground tomato were added into a paprika sample manufactured by Spice Chain in various concentrations as listed in Table 1.

Table 1: Adulterated Paprika Sample Information

Sample Adulterant ID	Adulterant %
Tomato skin Mix 1	10.86
Tomato skin Mix 2	5.54
Tomato skin Mix 3	1.23
Tomato skin Mix 4	0.52
Tomato skin Mix 5	0.11
Brick Dust Mix 1	14.12
Brick Dust Mix 2	4.77
Brick Dust Mix 3	1.04
Brick Dust Mix 4	0.68
Brick Dust Mix 5	0.1
Sudan I Mix 1	10.32
Sudan I Mix 2	4.58
Sudan I Mix 3	0.88
Sudan I Mix 4	0.62
Sudan I Mix 5	0.11

Sample Measurement

FT-NIR spectra were collected using two QuasIR™ 3000 spectrometers (Galaxy Scientific, Nashua, NH, USA). Samples were stored in 25 x 95 mm glass vials and then placed on top of the 23 mm sample window of the integrating sphere. Each sample was measured twice on each instrument, with 4 cm^{-1} resolution and 200 scans. Samples were shaken between measurements.

Data Processing

Spectral Sage™ software was used for data collection and

the CLS-based Advanced-ID™ algorithm and software were used for the analysis.

Result and Discussion

For a sample comprising n components, its spectrum S can be modeled as the sum of the spectra of n components $K_1 \dots K_n$, assuming the Beer-Lambert law is obeyed.

$$S = \sum_{i=1}^n (C_i * K_i) + R$$

where K is the matrix of reference spectra of the sample components, $c_1 \dots c_n$ are unknown coefficients and R is a residual, or error. The least squares solution to this equation for the coefficients can be found by standard matrix algebra, and is otherwise known as Classical Least Squares (CLS), or K -matrix regression.

If each spectrum contains m data points, then we can write this in matrix notation as:

$$S = K * c + R$$

where S and R are $m \times 1$ matrices, K is an $m \times n$ matrix of reference spectra, and c is an $n \times 1$ matrix of coefficients. Often, all of the components represented in K are known to be present and the objective of the regression is to find the coefficients c that can then be used to calculate their relative concentrations. In certain cases, however, one of the components may be an unknown that needs to be identified, or a suspected component whose presence in the mixture needs to be confirmed. If we designate this component as a target component and the spectrum of this component as T (the target spectrum), then for convenience we can rewrite the equation as:

$$S = T * c_0 + K' * c' + R$$

where S , T , and R are $m \times 1$ matrices, K' is an $m \times (n-1)$ matrix of reference spectra of known components that does not contain the spectrum in T , c' is an $(n-1) \times 1$ matrix of coefficients, and c_0 is a scalar coefficient.

Various methods can be used to judge the quality of the model, which includes the common practice of examining the size of the residuals, R . However, if the contribution of the target component to the spectrum S is very small, then the residual is a very poor indicator of the presence of the target component. This is because the regression

of only the spectra of the known components K' will result in a very good fit to the sample spectrum S , resulting in a very small residual (close to zero).

The patent-pending Advanced-ID™ algorithm finds a new approach to resolve this issue. It first calculates an approximation to the target spectrum by performing a regression that includes the target and known spectra ($S = T * c_0 + K' * c' + R$), and then calculates a residual with the coefficient for the target spectrum, c_0 , set to 0, thus defining the extracted spectrum E :

$$E = S - K' * c'$$

This can be compared with the expression for the residual R :

$$R = S - T * c_0 - K' * c'$$

The residual R will be small if either the target component is not present and $K' * c'$ is a good approximation to S , or if the target component is present and $T * c_0 + K' * c'$ is a good approximation to S . As noted above, this is therefore not a good indicator of the presence of the target component. The extracted spectrum E will also be small and will resemble R if the target component is not present and $K' * c'$ is a good approximation to S . However, if the target component is present in the sample at any significant concentration and $T * c_0 + K' * c'$ is a good approximation to S , then the extracted spectrum will resemble the spectrum of the target component. Additionally, if the target component is not present and $K' * c'$ is not a good approximation to S because another component is present that was not included in the regression, then the extracted spectrum will not resemble either R or the target spectrum.

Comparison of the extracted and target spectrum, typically scaled by the regression coefficient c_0 , can therefore be a reliable indicator of the target's presence. The comparison could be mathematical or visual by overlaying the two spectra on the computer screen.

As long as the spectra of all components present are included in the regression, the method described above will also work if the sample contains more than one suspected component that needs to be confirmed. In this case, one of the target spectra is T while all other target spectra are included in K' and the extracted spectrum is calculated and compared with the target spectrum. This is then repeated for each of the other target spectra. The method described above may also work with more than

one unknown component, especially if the principal spectral features of the unknowns are in different spectral regions. In this fashion, individual components in a mixture may be identified.

In this study, the target spectra are for tomato skin, red brick dust, and Sudan I dye. The extracted spectrum is expressed as

$$E = S_{AMP} - K_{paprika} * c'$$

where S_{AMP} is the spectrum of adulterated paprika powder and $K_{paprika}$ are the spectra of various paprika powder components. The Advanced-ID™ package was used to solve the unknown coefficients and calculate the correlation coefficient between the extracted spectrum and the adulterants' reference spectra.

The average NIR spectra of commercial paprika powders and the spectra of tomato skin, brick dust, and Sudan I dye adulterants are shown in Figure 1. As shown, the NIR spectrum for brick dust, which is an inorganic material, is unremarkable, with its main absorbance an O-H band due to moisture, whereas Sudan I dye has distinctive absorbance around 4600 cm^{-1} and 6000 cm^{-1} . As a natural product, the spectral features of tomato skin is very similar to paprika. Using the Advanced-ID™ method, spectral regions were developed to analyze brick dust, Sudan I dye, and tomato skin at 4250-5000 cm^{-1} , 5800-6200 cm^{-1} , and 4200-6600 cm^{-1} , respectively.

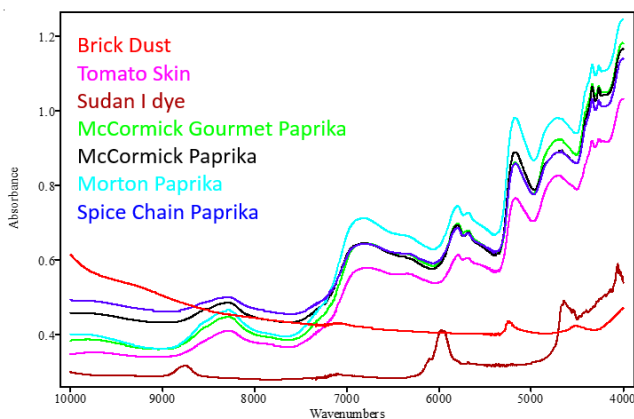


Figure 1: Spectra of pure components

Sudan I Dye Mixture

Sudan I dye was added to Spice Chain paprika powder in five concentrations: 10.32%, 4.58%, 0.88%, 0.62%, and 0.11%.

Apparent spectral changes can be observed around Sudan I dye's distinctive absorbance regions, especially between 5800 cm^{-1} and 6200 cm^{-1} as shown in Figure 2.

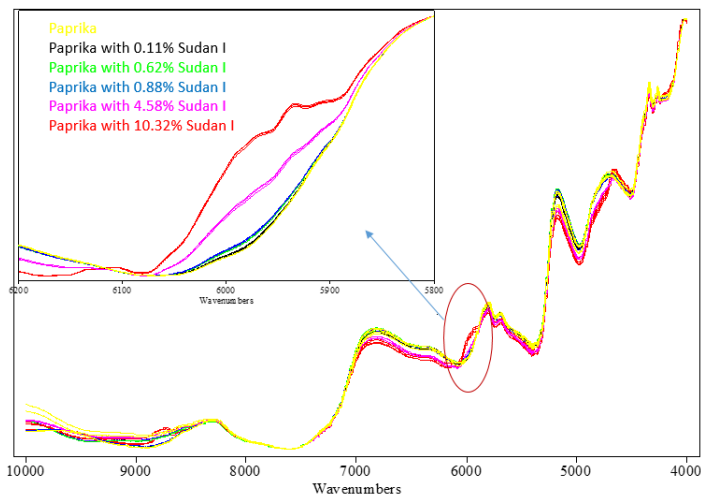


Figure 2. Original spectra of Sudan I mixtures

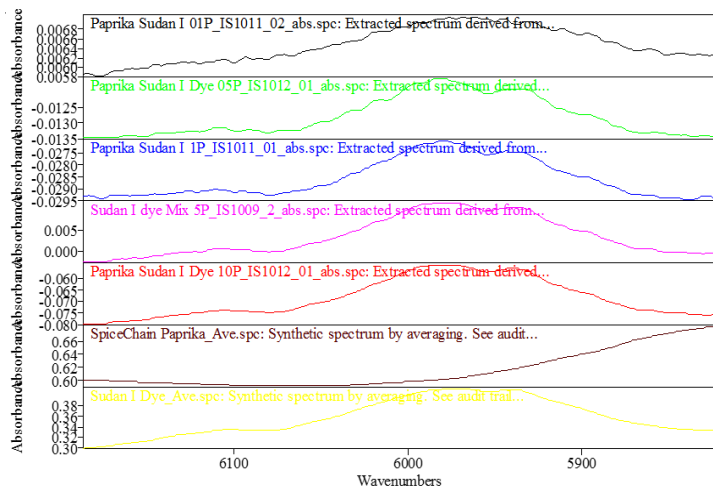


Figure 3. Extract spectra of Sudan I mixtures

Figure 3 presents the extracted spectra for the mixtures using the Advanced-ID™ method, along with the original spectra of Sudan I and paprika in the region of 5800–6000 cm^{-1} . As can be seen, the extracted spectra of the five mixtures have apparent absorbance peaks around 6000 cm^{-1} similar to Sudan I, whereas the Spice Chain paprika sample does not.

Table 2 lists the correlation coefficients of the extracted spectra to the Sudan I reference spectrum at different concentrations. The extracted spectra for adulterated samples, with Sudan I levels as low as about 0.1%, have a high correlation with the Sudan I spectrum, with a correlation coefficient higher than 0.95.

Table 2: Advanced-ID™ Result for Sudan I Dye Mixture

Sample ID	Sudan I (%)	Avg. Corr. Coefficient
Sudan I Mix 1	10.32	0.98
Sudan I Mix 2	4.58	0.98
Sudan I Mix 3	0.88	0.95
Sudan I Mix 4	0.62	0.95
Sudan I Mix 5	0.11	0.96

Brick Dust Mixture

Ground brick dust was added to Spice Chain paprika powder in five concentrations: 14.12%, 4.77%, 1.04%, 0.68%, and 0.10%. By visually checking the NIR spectra of samples, no apparent spectral changes were observed, except a baseline tilt at higher wavenumbers (Figure 4). With brick dust at lower concentrations, no apparent correlation was found between extracted spectra and the brick dust reference spectrum. Figure 5 presents the extracted spectra for brick dust mixtures 1, 2, and 3 using the Advanced-ID™ method along with the original brick dust spectrum in the region of 5050-5400 cm^{-1} . The extracted spectra of mixtures 1 and 2 show spectral features similar to brick dust, while the extracted spectrum of mixture 3 does not. The extracted spectra for adulterated samples, with brick dust levels at 1% and below, have low correlation with the brick dust spectrum, whereas brick dust levels at around 5% or higher have correlation coefficients higher than 0.95. This low sensitivity is directly linked to the inorganic nature of brick dust, which has no distinctive absorbance in the near-infrared region.

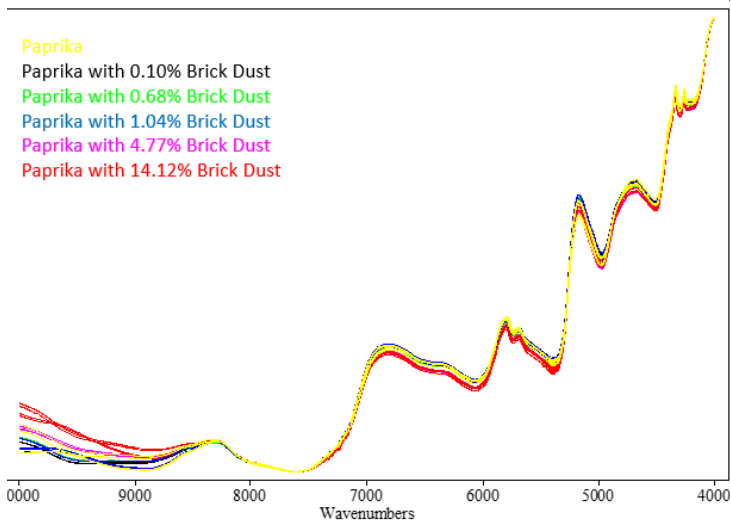


Figure 4. Original spectra of brick dust mixtures

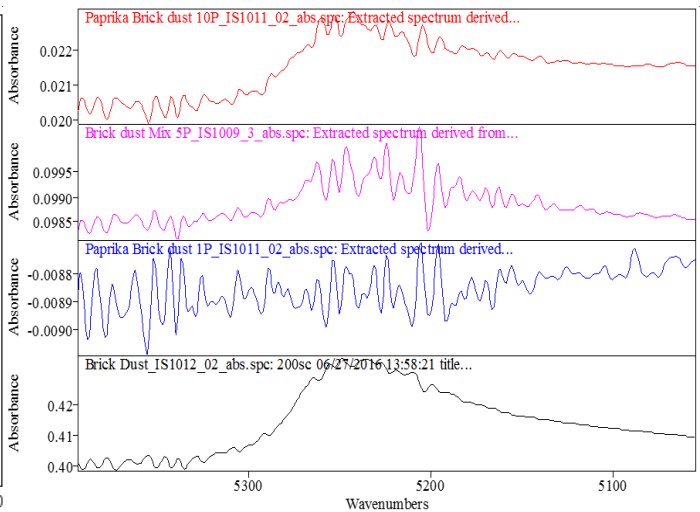


Figure 5. Extract spectra of brick dust mixtures

Tomato Skin Mixture

Ground tomato skin was added to Spice Chain paprika at concentrations of 0.11%, 0.52%, 1.23%, 5.54%, and 10.86%. Since the spectra of tomato skin and paprika look very similar, first derivative spectra were used to enhance small differences. Figure 6 displays the first derivative spectra of ground tomato skin, paprika, and the paprika-tomato skin mixtures in the region of 4200-6600 cm^{-1} . Spectral differences can be noticed between the ground tomato skin, paprika, and various mixtures.

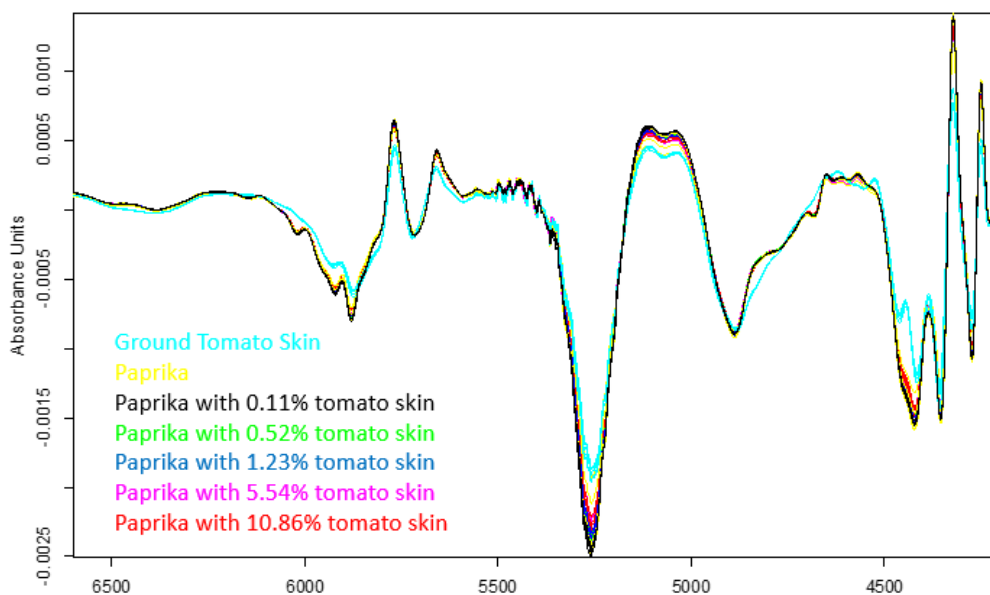


Figure 6. 1st derivative spectra of ground tomato skin, paprika and paprika mixtures

The extracted spectrum of the paprika-tomato skin mixture at a concentration of 0.11% does not show significant correlation with the tomato skin spectrum, whereas the extracted spectra of the mixture at concentrations of 0.52%, 1.23%, 5.54%, and 10.86% have average correlation coefficients of 0.93, 0.95, 0.96, and 0.98, respectively, to the spectrum of tomato skin. This demonstrates that at very low concentration (below 0.5%), it is very hard to detect tomato skin adulteration.

Conclusion

Using Galaxy Scientific's CLS-based Advanced-ID™ algorithm, we can detect paprika adulterated with Sudan I dye, tomato skin, and brick dust at high concentrations. However, with lower concentration adulteration, the detection limit varies with the nature of the adulterant. Sudan I dye has distinctive absorbance in the NIR region; thus, an adulteration level as low as 0.1% can be detected, whereas tomato skin adulteration can only be detected down to the 0.5% level. The worst is red brick dust, which has almost no absorbance in the NIR region and a detection limit of about 5%. However, to make economic sense, such adulteration is normally at high concentration, therefore, NIR can be used as a rapid screening tool for paprika adulteration at economically motivated concentrations.