Evaluation of near infrared hyperspectral imaging for detection of tuna powder contaminated with shrimp powder

K Nonthawong¹, P Phiriyangkul¹,², A Terdwongworakul¹ and K Phuangsonbut³

¹Division of Bioproducts Science, Department of Science, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, 73140, Thailand.
²Division of Biochemistry, Department of Science, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, 73140, Thailand.
³Agricultural Engineering Department, Faculty of Engineering, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, 73140, Thailand.

E-mail: faasprm@ku.ac.th.

Abstract. Near-infrared hyperspectral imaging (NIR-HSI) was investigated to detect the contamination of shrimp powder (SP) in tuna powder (TP) with partial least squares regression (PLSR) model. The principal component analysis was performed with NIR-HSI data for classification of tuna and shrimp powder. Samples for NIR-HSI data analysis were prepared using tuna powder contaminated with shrimp powder in concentration of 0%, 0.01%, 0.05%, 0.1%, 0.5%, 1%, 5%, 10%, 25%, 50%, 75% and 100% (w/w). The NIR-HIS in a wavelength range 864.5 to 1695.1 nm of the samples were used to create a prediction model using a partial least squares regression (PLSR) model. The result showed that the best model was based on spectra pretreated with second derivative combined with standard normal variate pretreatments, The performance of the prediction was expressed with the following values; factor = 3, $R^2_c = 0.989$, $RMSEC = 3.48\%$, $R^2_cv = 0.984$, $RMSECV = 4.218\%$, $R^2_p = 0.991$, $RMSEP = 3.110\%$. The regression coefficients of the PLSR model from 2D+SNV spectral pre-treatments were used to identify functional groups from the chemical composition of each sample. The study demonstrated that the NIR-HSI can be used for quantitative analysis of TP contaminated with SP which rapid non-destruction technique.

1. Introduction

Tuna powder products are popular seafood used as an ingredient in various food and have high nutritional value. But the production process may be contaminated with other products. Since other food powder may be used as additives, accidental contamination to tuna powder will pose a risk and must be controlled to maintain the quality of tuna powder. Therefore detection of contamination of food products such as shrimp powder in Tuna powder is very important because some protein of shrimps is allergen protein for consumers. The major protein of shrimp is tropomyosin [1]. Whereas major protein of Tuna is parvalbumin [2]. Shrimp powder and tuna powder have different protein which can be separated into chemical compounds for each sample. So the development of efficient detection techniques for accurate, rapid and non-destructive assessment is needed in order to reduce problems of hidden allergy which is the main cause of food safety.

There are several quantitative methods of contamination detection. In current year, there have been many reports on analysis method such as real-time polymer chain reaction (PCR) for detection of pathogens in seafood products [3], a sandwich enzyme-linked immunosorbent assay (ELISA) for the detection of fish and fish products
[4], mass spectrometry (MS) methods for allergen detection [5], quantitative method of nitrofuran residues using liquid chromatography combined with tandem in space mass spectrometry (LC-MS/MS) [6]. This technique has limitations in analysis and is destructive to the sample.

Near Infrared Spectroscopy (NIRS) has been applied to detection of cocoa powder with carob flour and the detection of milk powder with melamine [7, 8]. NIR spectroscopy is rapid and non-destructive which offers advantages over other technique. NIRS can be used for analysis both qualitatively and quantitatively. The wavelength of NIRS is in a range of 750-2500 nm of which spectra are characterized by reflectance or transmittance [7]. Absorption is dependent on the overtone and vibration of molecules. NIR spectra may contain background noise and scattering effects and can be removed by applying multivariate analysis [9]. Multivariate analyses popularly applied are principal component analysis and partial least squares regression. Hyperspectral imaging (HSI) gives both spatial and spectral data of samples. Also, HSI consists of a substantial amount of data that can be analyzed to improve model performance.

The objective of this study was to investigate the efficiency of the NIR-HIS model in the quantitative analysis of Tuna powder contaminated with shrimp powder.

2. Material and method

2.1. Sample preparation

Tuna powder and shrimp powder (Chaijinda Seafood, Kokkham, Samutsakhon, Thailand) were purchased from a local supermarket of Thailand. The average particle size of tuna powder and shrimp powder was 0.25 mm. For each sample, twelve tuna-shrimp powder mixtures were prepared with shrimp concentration (w/w) of 0, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 25, 50, 75 and 100% to develop concentration prediction models. Fifty grams of sample was mixed by shaking for 10 min and 10 grams of dry powder mixture were placed in a glass petri-dish (diameter 5515 mm). A total of five replicates (sets) for each mixture were prepared.

2.2. Near-infrared hyperspectral imaging

2.2.1. NIR-HSI system.

A line scanning NIR-HSI system (VLNIR-CL-100- N17E, SPECIM SisuCHEMA, Spectral Imaging Ltd., Oulu, Finland) was used to obtain the spectra of tuna and shrimp powder mixtures. While samples were scanned absorbance hyperspectral images in ENVI formal were recorded over a wavelength range of 900 to 1700 nm. Multivariate data analysis was carried out in the Evince version 2.7.9 - 2.7.10 software (Prediktera) to develop the models.

2.2.2. Data assignment.

Near-infrared hyperspectral images of a sample were stored in ENVI formation files and imported into Evince software version 2.7.9 - 2.7.10. The hypercubes area were extracted from each sample. A calibration dataset was allocated by using mean spectra of 3 scanned area sample (n=36) per dataset and prediction set was assigned using the mean spectra of a 2 scanned sample (n=24) of the dataset.

2.2.3. Multivariate data analysis.

Multivariate data analysis including principal component analysis (PCA) and partial least squares regression (PLSR) was performed to study the structure of the data and to develop the models. The original spectra as well as pre-treated spectra were used to create the models. Pretreatments for model improvement included derivative (D), second derivative (2D), multiplicative scatter correction (MSC), standard normal variate (SNV) and Savitzky-Golay smoothing (SG). The performance of PLSR model was assessed by the root mean square error of calculated (RMSEC), the root mean square error of cross-validation (RMSECV), the root mean square error of prediction (RMSEP) on the prediction set as in equation (1) and the coefficient of determination of calculated (Rc²), coefficient of determination of cross-validation (Rcv²) and coefficient of determination of prediction (Rp²) as in equation (2).

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (1)
\]

\[
\text{R}^2 = 1 - \frac{\text{SSE}}{\sum_{i=1}^{n}y_i^2} \quad (2)
\]
SSE = $\sum (y - \bar{y})^2$  \hspace{1cm} (3)

Where $y$ are reference values of the samples, $\bar{y}$ are model-predicted values of the samples and $\bar{y}$ are mean reference values of the samples. While SSE is the sum of squared errors calculated.

3. Result and discussion

3.1. Principal component analysis (PCA)

Figure 1 shows 2D scatter plot of the PCA model which displays a structural distribution of samples of tuna and shrimp powder mixture in space of principal component (PC) 1 and 2 with a variance of 81.10% and 4.19% of total variance, respectively. Score plot PC1 of icing sugar and corn flour mixtures model represents possible sample concentration [9]. 2D Scatter plot shows differentiation between the samples, which can be visualized as separated groups of the sample while PC2 shows different samples set.

Figure 2 shows 2D images of tuna powder and shrimp powder by PCA. Each pixel of the images represented the value of PC1 which were displayed with a different color having corresponding value shown in the color chart [10]. Hyperspectral imaging by PCA clearly distinguished between tuna powder and shrimp powder.
3.2. Partial least squares regression (PLSR)

PLSR model developed from NIR-HSI accurately predict the concentration of shrimp powder of the samples. The PLSR model was based on the extracted spectra of all samples (n=60) with wavelength from 864.53 to 1695.08 nm in 256 bands. To reduce the signal-to-noise ratio of data, the start and end of spectra were removed and the remaining region of spectra was 952.18 to 1695.08 nm. The prediction performance of models with different pre-treatments in a prediction of shrimp powder concentration is shown in Table 1. It was found that the best model was developed using 2D combined with SNV pre-treatment yielding the following values of factor = 3, $R_c^2 = 0.989$, RMSEC = 3.48 %, $R_{cv}^2 = 0.984$, RMSECV = 4.218%, $R_p^2 = 0.991$, RMSEP = 3.110 %. The standard normal variate spectral pre-treatment and second derivative pre-treatment gave the best results for prediction of total phenolic content in wax jambu fruit and for discrimination between samples at different drying times of bread [14, 15].

| Pretreatments | Factor | $R_c^2$ | RMSEC (%) | $R_{cv}^2$ | RMSECV (%) | $R_p^2$ | RMSEP (%) |
|---------------|--------|---------|-----------|-----------|------------|---------|-----------|
| -             | 3      | 0.950   | 7.331     | 0.938     | 8.172      | 0.972   | 5.456     |
| D             | 3      | 0.977   | 5.481     | 0.983     | 4.286      | 0.983   | 5.210     |
| 2D            | 3      | 0.981   | 4.528     | 0.975     | 5.208      | 0.979   | 4.769     |
| MSC           | 3      | 0.991   | 3.206     | 0.986     | 3.871      | 0.965   | 6.163     |
| SNV           | 3      | 0.990   | 3.268     | 0.986     | 3.937      | 0.971   | 5.587     |
| SG            | 3      | 0.989   | 3.494     | 0.987     | 3.794      | 0.974   | 5.343     |
| D+MSC         | 3      | 0.987   | 3.757     | 0.977     | 4.963      | 0.989   | 3.371     |
| D+SNV         | 3      | 0.987   | 3.762     | 0.977     | 4.963      | 0.989   | 3.371     |
| D+SG          | 3      | 0.976   | 5.069     | 0.967     | 5.967      | 0.969   | 5.799     |
| 2D+MSC        | 3      | 0.989   | 3.477     | 0.984     | 4.217      | 0.991   | 3.112     |
| 2D+SNV        | 3      | **0.989** | **3.480** | **0.984** | **4.218** | **0.991** | **3.110** |
| 2D+SG         | 3      | 0.982   | 4.465     | 0.975     | 5.244      | 0.971   | 5.603     |
| MSC+SNV       | 3      | 0.990   | 3.268     | 0.986     | 3.937      | 0.971   | 5.587     |
| MSC+SG        | 3      | 0.993   | 2.830     | 0.991     | 3.088      | 0.989   | 3.406     |
| SNV+SG        | 3      | 0.993   | 2.830     | 0.991     | 3.088      | 0.989   | 3.406     |

The regression coefficients of the best model are shown in Figure 3 which highlights some strong peaks at 8 wavelengths (i.e. 1137 nm, 1177 nm, 1352 nm, 1369 nm, 1486 nm, 1571 nm, 1617 nm, 1672 nm). Some of the wavelengths were associated with the C-H methyl (1370 nm), the N-H for secondary amine (1486 nm), the N-H amide (1570 nm), the C-H vinyl (1617 nm) and the C-H aromatic (1672 nm) [11]. The peak absorption was caused by different bond groups to present chemical compositions. Hyperspectral imaging combined with near infrared spectra is sensitive to the presence of molecules from the functional group [12]. The large $R^2$ are a high coefficient of determination displayed the capacity of the detection of the product in the medium by the NIR [13]. The lower RMSE implied the best model and the estimation accuracy due to the less error between the sample value and the predicted value.
4. Conclusion

This research demonstrated the potential of NIR-HSI analysis to quantitatively assess contamination of shrimp powder in tuna powder. Principal component analysis (PCA) can be used to differentiate between tuna and shrimp powder. The PLSR best model to predict the concentration of shrimp powder concentration in tuna was based on 2D+SNV pre-treated spectra. Regression coefficients of the best PLSR model can be used to identify molecules of samples which contributed to the difference in shrimp and tuna powder. In the next study, the robustness of the model for the detection of tuna powder contaminated with shrimp powder will be further studied including the depth from a surface of the sample that can detect tuna powder contaminated with shrimp powder.

References

[1] Hoffman D R, Day E D, and Miller J S 1981 *Ann. Allergy* 47 pp 17-22
[2] Ruethers T, Taki A C, Johnston E B, Nugraha R, Le T T K, Kalic T, McLean T R, Kamath S D and Lopata A L 2018 *Mol. Immunol.* 100 pp 28-57
[3] Taminiau B, Korsak N, Lemaire C, Delcenserie V and Daube G 2014 *Food Control* 44 pp 130-7
[4] Chen Y T and Hsieh Y H P 2014 *Food Control* 40 pp 265-73
[5] Monaci L, Angelis E D, Montemurro N and Pilolli R 2018 *Trac-Trend Anal. Chem.* 106 pp 21-36
[6] Douny C, Widart J, Pauw E D, Silvestre F, Kestemont P, Tu H T, Phuong N T, Maghuin-Register G and Scippo M L 2013 *Aquaculture*, 376-9 pp 54-8
[7] Quelal-Vasconez M A, Perez-Eateve E, Arnaud-Bonachera A, Barat J M and Talens P 2018 *Food Control* 92 pp 183-9
[8] Huang M, Kim M S, Delwiche S R, Chao K, Qin J, Mo C, Esquerre C and Zhu Q 2016 *J. Food Eng.* 181 pp 10-9
[9] Achata E M, Esquerre C, Gowen A A and O'Donnell C P 2018 *Powder Technol.* 336 pp 555-66
[10] Sandasi M, Chen W, Vermaak L and Vijoen A 2018 *Phytochem. Lett.* 24 pp 94-101
[11] Burns D A and Ciurczak E W 2001 Handbook of Near Infrared Analysis (New York)
[12] Elmasry G, Sun D W and Allen P 2012 *J. Food Eng.* 110 pp 127-40
[13] Lopes L C Brandao I V Sanchez O C Franceschi E, Borges G, Dariva C and Fricks A T 2018 *Process. Biochem.* 71 pp 127-33
[14] chakravartula S S N, Cevoli C, Belestra F, Febbi A and Rosa M D 2019 *J. Food Eng.* 240 pp 29-37
[15] Viegas T R, Mata A L M L, Duarte M M L and Lima K M G 2016 *Food Chem.* 190 pp 1-4
Acknowledgement
This research was supported by Kasetsart University Kamphaeng Saen Campus (KPS-RDI 2018-007), Graduate School of Kasetsart University. We thank Laboratory of Non-destructive Engineering for Agricultural Products, Department of Agricultural Engineering, Faculty of Engineering, Kasetsart University Kamphaeng Saen Campus for NIRS facilities.