Feasibility of Visible Near-Infrared Hyperspectral Imaging in Detection of Calcium Hypochlorite in Sago Flour

Ming Hao Lee[1], Agus Saptoro[1,2], King Hann Lim[3], Han Bing Chua[1], Tuong Thuy Vu[3], Nurleyna Yunus[4], Hasnain Hussain[5]

[1]Department of Chemical and Energy Engineering, Curtin University Malaysia, CDT 250, Miri 98009, Sarawak, Malaysia
[2]Curtin Malaysia Research Institute (CMRI), Curtin University Malaysia, CDT 250, Miri 98009, Sarawak, Malaysia
[3]Department of Electrical and Computer Engineering, Curtin University Malaysia, CDT 250, Miri 98009, Sarawak, Malaysia
[4]Downstream Technology Division, CRAUN Research Sdn. Bhd., Lot 3147, Block 14, Jalan Sultan Tengah, Petra Jaya, 93050 Kuching, Sarawak, Malaysia
[5]Centre for Sago Research, Universiti Malaysia Sarawak (UNIMAS), 94300 Kota Samarahan, Sarawak, Malaysia

Corresponding author: leeminghao@postgrad.curtin.edu.my

Abstract. The general public perspective on sago flour quality is based on the perceived colour appearances. This contributed to the potential of food fraud by excessive usage of bleaching agents such as calcium hypochlorite (CHC) to alter the product's colour. Conventional methods to detect and quantify CHC such as titration and chromatography are time-consuming, expensive and limited to laboratory setups only. In this research, visible near-infrared hyperspectral imaging (Vis-NIR HSI) was combined with partial least squares regression (PLSR) model to quantify CHC in pure sago flour accurately and rapidly. Hyperspectral images with the spectral region of 400 nm to 1000 nm were captured for CHC-pure sago mixture samples with CHC concentration ranging from 0.005 w/w% to 2 w/w%. Mean reflectance spectral data was extracted from the hyperspectral images, and was used as inputs to develop the PLSR model to predict the CHC concentration. The PLSR model achieved the commendable predictive results in this study, with Rp = 0.9509, RMSEP = 0.1655 and MAPEP of 3.801%, proving that Vis-NIR HSI can effectively predict the concentration of CHC in sago flour.

1. Introduction

Food integrity has been one of the concerning issues in the food industry in recent years. Due to the scandalous incidents of food fraud, such as melamine adulteration in baby milk powder, usage of gutter oil in the food and beverage industry and horse meat adulteration in beef patties, public awareness of the authenticity and safety of their food product has increased over the years [1]–[3]. Authorities also played their roles by monitoring and ensuring food safety was achieved. For the sago industry, Malaysia

is the world's largest sago starch exporter. Mukah in Sarawak is the hub of the sago industry, producing up to 96% of the sago production in the country [4]. With such a scale of production and exportation, Malaysia's sago industry and correlated authorities should be proactive in maturing the technology of sago quality monitoring to ensure the products are of good quality.

When most public consumers purchase sago flour, one major indication of the product quality is the colour, in which white-coloured starch is preferred over brown-coloured starch [5]. However, to gain economic advantages, unethical acts such as overusing bleaching additives (benzoyl peroxide and calcium hypochlorite) to sago starch are carried out to whiten the product. Apart from that, there is also potential for excessive usage of other varieties of chemical additives, for example, sodium metabisulfite, to prolong the shelf-life of sago flour and L-cysteine as a reducing agent to improve the textural properties [6]. Nevertheless, the excessive usage of these additives may pose public health hazards, and such applications should be limited and monitored.

In Malaysia, permitted food additives and their corresponding maximum limits are listed in Food Regulations 1985. The regulation allows minimal usage of nutrient supplements (e.g. L-cysteine and sulfites), preservatives (e.g. sodium metabisulfite, sorbic acids and proprionic acids) and a maximum of 50 mg/kg of benzoyl peroxide (BPO) for flour products. The regulation of BPO usage in flour production has differed for many different countries; for example, the maximum allowable BPO limit in Canada, the United States and the United Kingdom is 50 mg/kg, in Japan is 300 mg/kg, and the usage of BPO is banned in China [7]. On the other hand, the regional standard for edible sago starch (CXS 301R-2011) suggested by the Food and Agricultural Organization of the United Nations follows the General Standard for Food Additives (CXS 192-1995) which suggested the maximum allowable limited of BPO to be 75 mg/kg, sodium metabisulfite (SMBS) to be 200 mg/kg and chlorine to be 2500 mg/kg [8]. A summarisation of chemical additives regulations suggested by FAO is tabulated in **Table 1**.

Table 1. Allowable Chemical Additives in Edible Sago Starch (FAO Standards)

Chemicals	Maximum Allowable Concentration
Ascorbic acid	300 mg/kg
Benzoyl peroxide	75 mg/kg
Chlorine	2500 mg/kg
Phosphates	2500 mg/kg
Sulfites	200 mg/kg

In practice, the determination of the chemical additives applied in the flour industry is performed in the lab by following set standards by the International Organization of Standardization (ISO), Association of Agricultural Chemists (AOAC) or National Standards [9]. These processes involved traditional lab analytical methods such as titration techniques or complex instrumentations, for example, atomic absorption spectrometry, high-performance liquid chromatography and liquid chromatography-mass spectrometry [10]–[12]. Although these methods offer high detection accuracy, they are often cost-ineffective, time-consuming, require complex instrumentations to be operated by trained personnel and the analytical processes are limited in the laboratory. Thus, a rapid, cost-effective and accurate sago quality monitoring system is required to achieve a rigorous monitoring process of the sago industry.

In recent years, hyperspectral imaging (HSI) has been introduced to various food industries as a rapid quality monitoring solution, including the flour industry. HSI combines conventional imaging with spectroscopy, allowing each spatial pixel to carry the spectral information of the sample. With both the spatial and spectral information possessed by hyperspectral images, the application of HSI would allow the integration of computer vision to select an optimal region of interest (ROI) to efficiently analyse the desired physical or chemical parameters of a sample. Recent studies have shown the application of HSI

in various flour and starch industries, such as micronutrients quantification in wheat flour (375 nm to 1050 nm, Rp>0.6), BPO and talcum powder detection in wheat flour (900 nm to 1700 nm, discrimination successful at subpixel level), peanut powder detection in whole wheat flour (900 nm to 2500 nm, discrimination successful at subpixel level), detection of limestone adulteration in tapioca starch (935 nm to 1720 nm, Rp = 0.996) and discrimination of binary mixture of corn flour and icing sugar (880 nm to 1720 nm, Rp>0.997) [7], [13]–[16].

To date, no studies based on HSI have been performed on sago flour. The prior studies showed that HSI is effective in quality monitoring of various flour and starch products. However, most studies had chosen HSI with a spectral range within the NIR region of 900 to 1700 nm and above. Limited studies are performed on flour and starch products with an HSI system governing the spectral range of the Visible + Near Infrared (Vis-NIR) region. Therefore, this study aims to evaluate the effectiveness and feasibility of the Vis-NIR HSI system with a spectral range of 400 nm to 1000 nm in detecting low calcium hypochlorite (CHC) levels in sago flour to address the existing research gap.

2. Methodology

2.1. Samples Preparation

Pure and unprocessed sago starch was sourced from a local mill in Mukah, Sarawak. The unprocessed sago starch was washed and sieved over a cloth filter to remove excess fibre from the milling process. The washed sago starch was then dried in a tray dryer to achieve a moisture level below 15%. The chemical additives CHC (≤ 100%) was purchased from Sigma-Aldrich. The CHC concentration added to the sago samples were 0.005 w/w%, 0.01 w/w %, 0.02 w/w %, 0.04 w/w %, 0.06 w/w %, 0.08 w/w %, 0.1 w/w %, 0.25 w/w %, 0.5 w/w %, 0.75 w/w %, 1 w/w %, 1.25 w/w %, 1.5w/w%, 1.75 w/w % and 2 w/w % respectively. Three replicates for each concentration of CHC-pure sago mixture were prepared. To ensure the homogeneity of the mixtures, each mixture was placed in sterilised 50 mL centrifuge tube and shaken with a vortex mixer for 10 minutes. A total of 48 CHC-pure sago mixtures were prepared in this study.

2.2. Hyperspectral Imaging Process

A benchtop line-scanning HSI system (Resonon Pika L, Resonon Inc., MT, USA) is used in this study. The HSI system consists of a spectrograph with a spectral range from 400 nm to 1000 nm and a spectral resolution of 3.3 nm, an objective lens with a focal length of 23 mm and aperture of f/2.4 was attached to the spectrograph, four halogen lamps with stabilised power supply and a linear translation stage driven by a stepper motor. The HSI system is connected to a computer installed with Spectronon software to control both the imager and the stepper motor.

Prior to the imaging process, a 'white' reference image is captured using a white Teflon board with 99.9% reflectance and a 'dark' reference image is captured by covering the camera lens with a lens cap. Each mixed sample was filled into a 66 mm diameter petri dish to full and the surface was levelled to the brim of the petri dish with a spatula. The Petri dishes were placed on the centre of the stage for imaging with the following HSI system settings: a) frame rate of 50 frames/sec b) shutter time of 16.49 ms, c) stage speed of 4.5 mm/s, and d) lines capture of 700. Radiometric calibration was performed concurrently with the Spectronon software based on the 'white' and 'dark' reference images to minimise signal noises and uneven dark current in the images.

2.3. Hyperspectral Images Pre-processing

Firstly, image segmentation was performed to identify the region of interest (ROI) and sub-sampling an image. The central region of the image was selected with a spatial dimension of 450×450 pixels, and the selected ROI was further broken down into 9 sub-samples with 150×150 pixels each. The mean spectrum of each sub-sample was then calculated to extract the spectral information from the ROI of the sago flour mixture.

2.4. Regression Analysis

The PLSR model was used in this study to predict the concentration of CHC in the sago flour mixtures based on the raw and pre-treated spectral data. In order to develop and optimise the hyperparameter (number of latent variables (LVs)) of the model, 432 sub-samples were split into training and testing datasets with a train:test ratio of 7:3. To ensure the replicability of the simulation, the splitting process was performed pseudo-randomly by using fixing a seed value to randomly split the data. To optimise the hyperparameter of the PLSR model, k-fold cross-validation strategy was applied with a recommended k value of 5 due to the small sample size ($n \leq 5000$) [17]. A selection of k value = 5 in the occasion of a smaller sample size would suffice in determining the optimal model hyperparemeter, while reducing the computational load. The number of LVs was selected when the mean value of root mean square error of cross-validation (RMSECV) is minimum OR the subsequent increment of latent variables does not greatly reduce the mean value of RMSECV.

After the PLSR model was optimised, the performance of the models was analysed based on the prediction results of the test data. The error score metrics – coefficient of determination of prediction (Rp), root mean square error of prediction (RMSEP) and mean average percentage error of prediction (MAPEP) were used to evaluate the predictive performances of the model mathematically. Generally, Rp values are desired when approaching +1, and RMSEP and MAPEP are desired when approaching 0 [18]. The desired value of Rp dictates good fitness of the model prediction, and desired values of RMSEP and MAPEP dictates good prediction accuracy. The model development and evaluation process were performed using Python 3.10.5 with Visual Studio Code (Microsoft, Redmond, WA, USA).

3. Results and Discussions

3.1. Characteristics of the Hyperspectral Reflectance Data

Fig. 1 shows the mean reflective spectra for CHC-pure sago mixture with CHC concentration varying from 0.01 w/w %, 0.1 w/w %, 0.5 w/w %, 1 w/w % and 2 w/w %. Across all concentrations of CHC, the spectra for all samples converged at the wavelength region of 650 nm to 750 nm. This wavelength interval is mainly within the region of red colour, hence dictating that the sago mixture with varying CHC concentration has minimal and undistinguishable characteristics in their 'redness'. Furthermore, notable peak can be noticed across all sago mixture spectra at the spectral region of: (a) 420 – 440 nm, (b) 540 – 560 nm, (c) 610 – 630 nm and (d) 930 – 970 nm. The spectral regions of (a), (b) and (c) are in the visible light region of the violet, green and orange colour, respectively. Although the trend is undeterminable based on the results, an inference could be made that those are the more prominent colour contributing to the visual characteristics of sago flour, and the addition of CHC may alter the visual characteristics within those regions. On the other hand, the peak observed in the spectral region of (d) is within the near-infrared region that corresponds mainly to the second N-H overtone of peptides and proteins in starches [19]. It is observed from **Fig. 1** that the reflectance of the said peak reduced in correlation to the concentration of CHC in the sago mixture. This occurrence could explain that CHC could destroy some of the protein and peptide chains within the sago flour, hence a reduction in the particular peak.

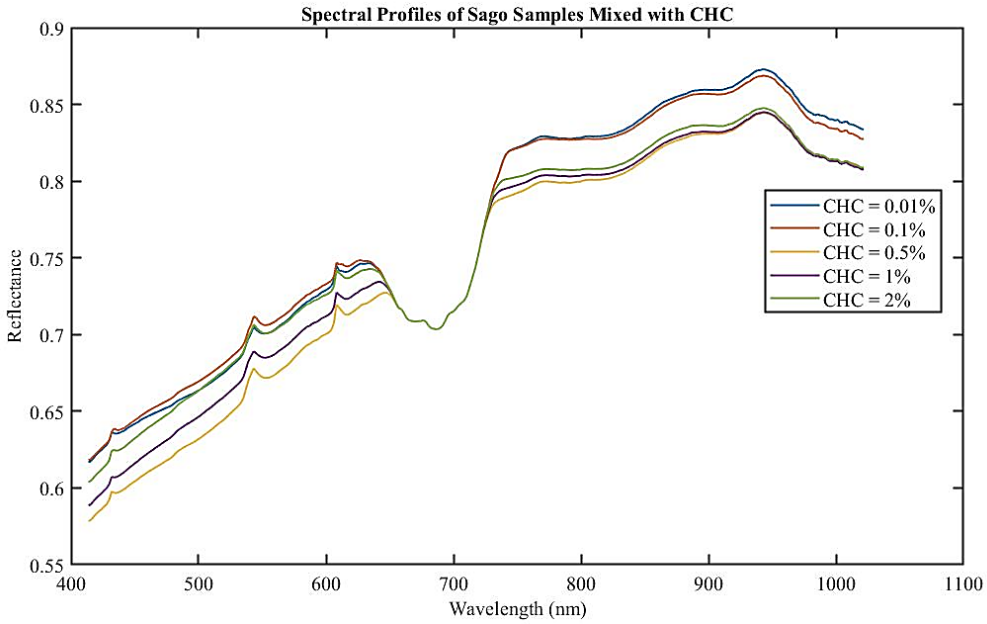

Fig. 1. Comparison of Spectral Profiles of CHC-Pure Sago Mixtures at Different Concentration of CHC

3.2. Prediction Models for CHC Concentration in Sago Flour

To establish an accurate and computationally efficient model, hyperparameters optimisation is performed using the k-fold cross-validation strategy to determine the optimal number of LVs to be included in each PLSR model. The hyperparameters optimisation process for PLSR is important as selecting a lesser number of LVs would introduce model underfitting; whereas selecting an excessive number of LVs would either introduce model overfitting or computational efficiency issues, or both. The hyperparameters optimisation process is shown in **Fig. 2**, where the optimal number of LVs was selected when the RMSECV is minimum or approaches a minimum. It was also observed from **Fig. 2** that the model suffered from slight performance degradation as the number of LVs increased from the optimal point due to overfitting.

The extracted mean reflectance spectral data was tested for predicting the CHC concentration in the CHC-sago flour mixture using PLSR models. The optimised parameter and performance metrics of the developed model are shown in **Table 2**. By examining the performances of the PLSR model developed with the spectral data, the score metrics showed good model performances with Rp = 0.9509, RMSEP = 0.1655 and MAPEP of 3.801%. The results reflected that the reflectance spectra extracted from the hyperspectral images with the spectral range of 400 nm to 1000 nm effectively predict the CHC concentration in the CHC-sago flour mixture.

Table 2. Performances Metrics for CHC-Pure Sago Estimator based on PLSR Model

Model	PLSR
Optimised Hyperparameters (Number of LVs)	22
Rp	0.9509
RMSEP	0.1655
MAPEP	3.801

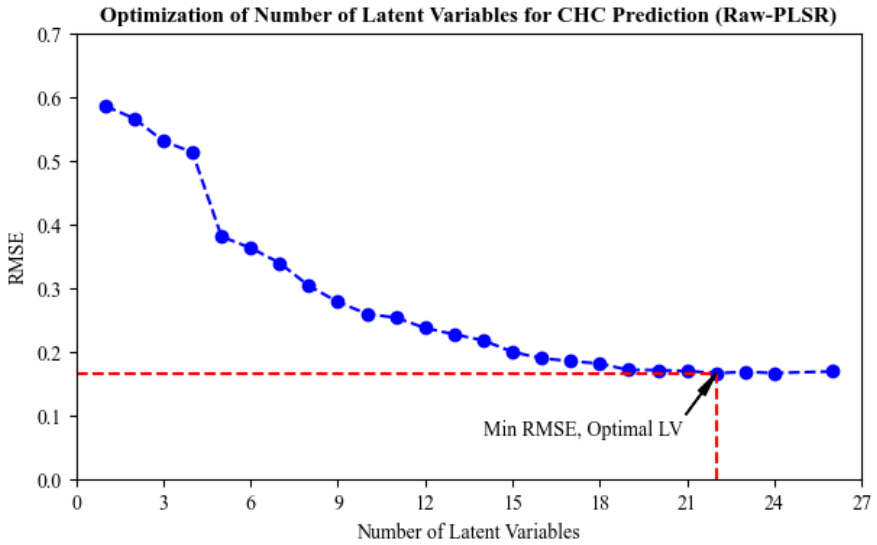

Fig. 2. Hyperparameter Optimization of PLSR Model

4. Conclusions

An HSI system with a spectral range of 400 nm to 1000 nm was used to capture the hyperspectral images of CHC-sago flour mixture samples. The mean reflectance spectrum was extracted from the ROI of the hyperspectral images and was then tested for predicting the CHC concentration in the CHC-sago flour mixture samples with the PLSR model. PLSR model developed based on the extracted spectral data yielded good predictive performances in this study with Rp = 0.9509, RMSEP = 0.1655 and MAPEP of 3.801%, in which both good fitness and predictive accuracy can be inferred based on the Rp value approaching 1 and low MAPEP value. This research results indicated that hyperspectral images with reflectance data over the spectral range of 400 nm to 1000 nm could effectively predict the concentration of CHC in sago flour with good model performances score metrics. However, the developed model showed further potential to be improved, and the implications of pre-processing algorithms such as spectral pre-treatment and spectral data selection and usage of alternative modelling algorithms could be ventured in the future.

Acknowledgement

The authors would like to acknowledge the financial support given by Sarawak Research and Development Council (SRDC) through SRDC Catalyst Fund (RDCRG/CAT/2019/03).

References

[1] H. Lee, M. S. Kim, S. Lohumi, and B.-K. Cho, “Detection of melamine in milk powder using MCT-based short-wave infrared hyperspectral imaging system,” *Food Addit. Contam. Part A*, vol. 35, no. 6, pp. 1027–1037, Jun. 2018, doi: 10.1080/19440049.2018.1469050.

[2] M. Strong, “Gutter oil scandal executive sentenced to 22 years in prison,” *Taiwan News*, Taipei, Sep. 13, 2017. [Online]. Available: https://www.taiwannews.com.tw/en/news/3252462

[3] R. Smith and G. McElwee, “The ‘horse-meat’ scandal: illegal activity in the food supply chain,” *Supply Chain Manag. An Int. J.*, vol. 26, no. 5, pp. 565–578, Jan. 2021, doi: 10.1108/SCM-08-2019-0292.

[4] N. M.Amin, N. Sabli, S. Izhar, and H. Yoshida, “A review: Sago Wastes and Its Applications,” *Pertanika J. Sci. Technol.*, vol. 27, pp. 1841–1862, Oct. 2019.

[5] D. Wahab, Muhidin, Ansharullah, Hermanto, and Surni, "The Effect of Whitening Chemical Properties on the Sago Starch Quality in Southeast Sulawesi Indonesia," *Int. J. ChemTech Res.*, vol. 10, no. 1, pp. 7–11, 2017.

[6] S. Lohumi, H. Lee, M. S. Kim, J. Qin, and B. K. Cho, "Raman hyperspectral imaging and spectral similarity analysis for quantitative detection of multiple adulterants in wheat flour," *Biosyst. Eng.*, vol. 181, pp. 103–113, May 2019, doi: 10.1016/j.biosystemseng.2019.03.006.

[7] X. Fu, J. Chen, F. Fu, and C. Wu, "Discrimination of talcum powder and benzoyl peroxide in wheat flour by near-infrared hyperspectral imaging," *Biosyst. Eng.*, vol. 190, pp. 120–130, Feb. 2020, doi: 10.1016/j.biosystemseng.2019.12.006.

[8] Food and Agriculture Organization of the United Nations, "General Standard for Food Additives," 2019. Accessed: Apr. 14, 2021. [Online]. Available: http://www.fao.org/food/food-safety-quality/scientific-advice/jecfa/jecfa-additives/en/.

[9] Food and Agriculture Organization of the United Nations, "Regional Standard for Edible Sago Flour (Asia)," 2011. Accessed: Apr. 14, 2021. [Online]. Available: http://www.codexalimentarius.net/web/members_area.jsp?lang=EN.

[10] Q. Wang, W. Shi, and C. Hou, "Determination of benzoyl peroxide content in wheat products by high-performance liquid chromatography," *J. Food Process. Preserv.*, vol. 34, no. 3, pp. 414–424, Jun. 2010, doi: 10.1111/j.1745-4549.2008.00325.x.

[11] Y. Abe-Onishi, C. Yomota, N. Sugimoto, H. Kubota, and K. Tanamoto, "Determination of benzoyl peroxide and benzoic acid in wheat flour by high-performance liquid chromatography and its identification by high-performance liquid chromatography–mass spectrometry," *J. Chromatogr. A*, vol. 1040, no. 2, pp. 209–214, 2004, doi: https://doi.org/10.1016/j.chroma.2004.03.059.

[12] T. Lin, M. Zhang, F. Xu, X. Wang, Z. Xu, and L. Guo, "Colorimetric detection of benzoyl peroxide based on the etching of silver nanoshells of Au@Ag nanorods," *Sensors Actuators B Chem.*, vol. 261, pp. 379–384, 2018, doi: https://doi.org/10.1016/j.snb.2018.01.172.

[13] N. Hu *et al.*, "Predicting micronutrients of wheat using hyperspectral imaging," *Food Chem.*, vol. 343, p. 128473, May 2021, doi: 10.1016/j.foodchem.2020.128473.

[14] A. Laborde *et al.*, "Subpixel detection of peanut in wheat flour using a matched subspace detector algorithm and near-infrared hyperspectral imaging," *Talanta*, vol. 216, p. 120993, Aug. 2020, doi: 10.1016/j.talanta.2020.120993.

[15] D. Khamsopha, S. Woranitta, and S. Teerachaichayut, "Utilizing near infrared hyperspectral imaging for quantitatively predicting adulteration in tapioca starch," *Food Control*, vol. 123, p. 107781, May 2021, doi: 10.1016/j.foodcont.2020.107781.

[16] E. M. Achata, C. Esquerre, A. A. Gowen, and C. P. O'Donnell, "Feasibility of near infrared and Raman hyperspectral imaging combined with multivariate analysis to assess binary mixtures of food powders," *Powder Technol.*, vol. 336, pp. 555–566, Aug. 2018, doi: 10.1016/j.powtec.2018.06.025.

[17] B. G. Marcot and A. M. Hanea, "What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis?," *Comput. Stat.*, vol. 36, no. 3, pp. 2009–2031, 2021, doi: 10.1007/s00180-020-00999-9.

[18] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation.," *PeerJ. Comput. Sci.*, vol. 7, p. e623, 2021, doi: 10.7717/peerj-cs.623.

[19] H. Zhao, B. Guo, Y. Wei, and B. Zhang, "Near infrared reflectance spectroscopy for determination of the geographical origin of wheat," *Food Chem.*, vol. 138, no. 2, pp. 1902–1907, 2013, doi: https://doi.org/10.1016/j.foodchem.2012.11.037.