Wiley
Aquaculture Research
Volume 2024, Article ID 1524148, 10 pages
https://doi.org/10.1155/2024/1524148

WILEY

Research Article

Non-Invasive Prediction of Barramundi (*Lates calcarifer*) Flesh Lightness Using Near-Infrared Spectroscopy

Weiyu Chen[1,2] **Dean R. Jerry**[1,2,3] **Ronald D. White**[1,2] **and Leo Nankervis**[1,2]

[1]*ARC ITRH for Supercharging Tropical Aquaculture Through Genetic Solutions, James Cook University, Queensland 4811, Townsville, Australia*
[2]*College of Science and Engineering, James Cook University, Queensland 4811, Townsville, Australia*
[3]*Tropical Futures Institute, James Cook University, Singapore, Singapore*

Correspondence should be addressed to Weiyu Chen; weiyu.chen@my.jcu.edu.au

Received 22 March 2024; Accepted 25 September 2024

Academic Editor: Christyn Bailey

Barramundi (*Lates calcarifer*) is an important aquaculture species extensively farmed throughout its natural distribution of Australia and Southeast Asia, as well as being increasingly farmed in the Middle East, USA, and Europe. Barramundi has a firm, pink-white flesh; however, fillets from farmed barramundi often exhibit grey colouration. This grey colouration detracts from its market appeal, leading to challenges in consumer acceptance and competitiveness of the product against other white fillet fish. Selective breeding, environmental manipulation, and dietary adjustments are being investigated to reduce grey flesh colouration. Yet, the absence of a rapid, noninvasive approach to predict greyness in flesh means that large numbers of samples cannot be quickly evaluated, and issues cannot be mitigated preharvest and noninvasively to preserve the fish. To address this issue, rapid analysis of flesh greyness was developed using noninvasive near-infrared (NIR) spectroscopy through the fish skin. Thirty fish were purchased from a barramundi farm, filleted, and divided into 3 cm sections, yielding a total of 335 samples from both dorsal and ventral fillet regions. NIR spectral data were obtained from the skin side of all samples, and colouration data were collected from the flesh side of the same samples. Data were randomised into a training set (256 spectra) and a validation set (79 spectra). Predictive models were developed using flesh colour as the training input for skin NIR spectra. The refined partial least squares regression model explained 78% of the variation in the medial flesh colour (R^2_{pe} of 0.776, an RMSEP of 2.820, and an RPD_{pe} of 2.122) demonstrating the ability to adequately predict the flesh quality through skin spectra. This highlights the potential of NIR spectroscopy as a dependable, noninvasive tool, enabling the rapid evaluation of large samples and offering the potential to address flesh colouration issues in barramundi preharvest.

Keywords: barramundi; fillet colour; near-infrared; noninvasive assessment; rapid analysis

1. Introduction

Barramundi (*Lates calcarifer*), also known as Asian seabass, is an important aquaculture species in southeast Asia and tropical Australia and is widely distributed in the Indo-Pacific region [1]. The global production of barramundi has undergone more than fivefold growth in the 20 years since 2000, with 105,800 tonnes produced in 2020 [2]. Wild barramundi generally has a firm, pink-white flesh; however, farmed barramundi fillets often exhibit greyness (Figure 1). Market analyses indicate that the appearance of grey colouration in barramundi flesh decreases the attractiveness to customers when the product is sold alongside white filleted produce and in some instances imported barramundi [3]. Thus, measuring the greyness of barramundi flesh as a commercial trait and identifying those fish with excessive greyness are important to the industry, as it will allow them to investigate the potential of selective breeding to reduce flesh greyness, as well as other management options such as diet and environmental manipulation. However, currently, the only way to quantify flesh greyness is through slaughtering and filleting the fish, which is invasive and labour-intensive. While noninvasive sampling not only

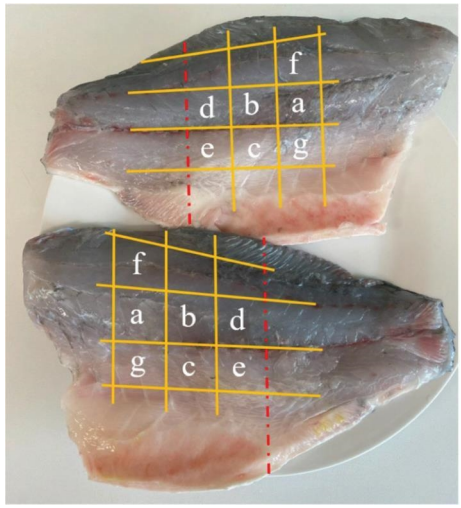


FIGURE 1: Barramundi fillet showing present of flesh greyness and example of how the fillet was cut into multiple samples for NIR and colour measurements.

allows for the rapid sampling of higher numbers of fish, but it also reduces impact and improves animal ethics outcomes in research [4]. From a commercial standpoint, identification of fish with grey flesh by a through skin approach will facilitate the assessment of potential downgrades. Industry can also supply barramundi fillet with differentiated colouration to different markets to obtain optimal returns.

In the past, the degree of grey colouration in fish flesh has been quantified through the L^* of Commission Internationale de l'Éclairage (CIE) $L^*a^*b^*$ colour space by taking images of fish fillet, where L^* represents the lightness. However, this analysis is time-consuming, laborious, and destructive. Alternatively, near-infrared (NIR) spectroscopy is a powerful analytical tool for measuring biochemical information that has been applied in a wide range of different fields, including food science, agriculture, and medicine [5–7]. The range of electromagnetic wavelengths within the NIR spectrum is from 780 to 2565 nm, which is beyond the range of visible wavelength [8]. Within the delineated spectral range, molecules exhibit absorption at characteristic wavelengths, where the absorption is highly related to molecular vibrational modes based on their structural configurations and specific functional groups, such as the carbon–hydrogen (C—H) and nitrogen–hydrogen (N—H) bonds [8–10]. Thus, analysis of NIR spectral signatures can provide information on the presence and concentration of organic compounds in the sample. Similar approaches have yielded effective through skin measurement of fat and moisture in Atlantic salmon and rainbow trout [11, 12]. NIR technology also has the ability to detect the bruises of avocado and predict colouration of pork muscle and peach flesh, where the colouration could be affected by more than one chemical compound [13–15]. This gives NIR technology the advantage to predict degree of grey colouration without knowing which exact chemicals are causing the grey colouration. Melanin is a pigment that has been linked to grey colouration in several species and was initially suspected to be the factor for grey colouration in barramundi flesh [16, 17]. While melanin has been identified in barramundi flesh, its concentration explains approximately 18% of the variation in grey colouration in dorsal flesh indicating that further unknown factors influence grey colouration in barramundi flesh [3].

NIR measurements are inexpensive and much faster than wet chemistry for the quantification of organic compounds [18, 19], and the use of NIR is increasingly being applied into high-throughput phenotyping systems [20]. Further, NIR possesses a greater penetration depth compared to mid-infrared spectroscopy [21]. NIR light with wavelengths ranging from 900 to 1600 nm has been shown to penetrate up to 3.9 mm in chicken breast tissue [22] and approximately 3.3 mm in avocado at a wavelength of 1076 nm [23]. Reflectance spectral measurements in the 900–1600 nm range generally achieve penetration depths between 1 and 5 mm [24]. This indicates the potential for NIR spectroscopy to effectively analyse the flesh beneath the skin of barramundi. Even if the NIR does not penetrate deep into the flesh, there can be correlations between the spectral signature of the skin and flesh interface that relate to the chemical properties below the skin. All of these advantages make NIR spectroscopy a potential tool for predicting barramundi flesh greyness in a rapid, cheap, and noninvasive manner.

Therefore, the aim of this research was to develop a reliable, rapid, and nondestructive method of flesh greyness determination in barramundi by using NIR spectroscopy. Here, NIR measurements were taken from selected locations on the barramundi skin and correlated with the flesh colouration (L^*) at the same location in the fillet to build a predictive model.

2. Methods and Materials

2.1. Sample Preparation and Collection. A total of 30 commercially grown barramundi were obtained from a local farm in North Queensland, Australia. The weight of each barramundi varied from 1 to 3 kg, and the length was between 40 and 55 cm. In case culture environment influenced flesh greyness and in order to potentially capture a broader range of flesh colouration, 30 barramundi were sourced from two distinct freshwater ponds, with 15 fish from each pond. From each fish, two whole fillets were excised, and then from this fillet, vertical cuts were made starting at the site of the anus every ~ 3 cm anteriorly from this cut. From each vertical flesh strip, horizontal cuts were then excised that followed the red muscle line in the fillet. Depending on the size of each barramundi, between five (a–e) and seven (a–f) sections could be sampled from each fillet strip (Figure 1).

2.2. Reference Measurement of Colouration. The L^* value of CIE $L^*a^*b^*$ colour space has become a standard measure of the inverse of grey colouration since L^* represents the perceptual lightness by the human eye [25–27]. Hence, L^* was selected to describe the level of grey colouration in barramundi flesh in this study.

Images of deep flesh colouration (medial surface of fillets) were taken under standardised light conditions [28]. All flesh samples were placed in a white tray in the centre of field

of view surrounded by four colour checker passports (X-Rite Pantone, USA). A digital camera (Sony a5000 ILCE-5000, Japan) was used to take images with settings: ISO = 100, focal length = 16 mm, f-stop = F/10, exposure time = 1/60 s, and focus = 0.8 m.

All images were calibrated for colour using the colour checker passport in MATLAB [27]. This image calibration ensured that the colouration information of each image was not affected by any unexpected changes in ambient temperature and light irradiance [29]. The average lightness value (L^*) of the central area of each fillet was measured on the calibrated image by the software ImageJ [30]. The range of L^* value is from 0 (pure black) to 100 (pure white).

2.3. NIR Spectral Acquisition. NIR spectral data were collected 3 mm above the skin side of each fillet via a portable MicroNIR device (N1-00148, VIAVI Solutions). The spot diameter of the MicroNIR device was 1 cm, and the wavelength range was between 908 and 1676 nm, with a resolution of 6.2 nm. Three independent NIR spectral readings were obtained for each location. The spectral data from these three readings were then averaged and used in regression analyses. A total of 335 spectra and corresponding lightness (L^*) were obtained and were randomised into training set (256 spectra) and validation set (79 spectra). The impacts of external lighting were minimised, and all measurements were taken at a consistent standard room temperature (22°C).

2.4. Chemometric Methods in NIR Spectra Analysis. Partial least squares regression (PLSR), principal component regression (PCR), and multiple linear regression (MLR) methods were applied to the data using HYPER-Tools Version 3.0 within the MATLAB environment [31] to create a predictive model by using the preprocessed data and reference data (L^* value of barramundi flesh) [32]. Cross-validation was used to optimise the regression model for better prediction when using PLSR, PCR, and MLR methods. The number of latent variables or principal components was determined with two different methods. Initial models were developed using the lowest root mean square error of cross-validation (RMSECV) to ascertain the number of latent variables or principal components for regression analysis [10, 19]. Raw spectra were preprocessed by three different methods where necessary to enhance the spectral features, including Savitzky–Golay (SG) smoothing, standard normal variate (SNV, [33, 34]), multiplicative signal correction (MSC, [18]), and first and second derivative [21].

The performances of initial models developed by the training set were assessed based on four essential regression indicators (the coefficient of determination for calibration (R^2_{ca}), root mean square error of calibration (RMSEC), coefficient of determination for cross-validation (R^2_{cv}), and RMSECV) [10]. Models that showed a similar value between R^2_{ca} and R^2_{cv} with lower RMSEC and RMSECV and higher R^2_{cv} were selected [35, 36]. $|R^2_{ca}–R^2_{cv}|$ vs R^2_{cv} plot was used to further determine the optimal number of latent variables or principal components for higher R^2_{cv} and smaller difference between R^2_{ca} and R^2_{cv} which can avoid overfitting and capturing the noise in spectral data [37, 38]. From this analysis, a predictive model was created.

Calculated and experimental responses from the initial models were used to create a studentized residual plot. Any studentized residual bigger than 2 in absolute value was considered an outlier and removed. The total number of outliers might be different for each model due to different preprocessing methods and regression methods. Final models were established using the same method but with the training set after the removal of outliers. Then, the models were tested by the validation set and calculated corresponding R^2_{pe}, root mean square error of prediction (RMSEP), and the ratio of performance to deviation of the prediction (RPD_{pe}). The RPD_{pe} was employed to inform the predictive quality and robustness of the regression model [23, 39–41].

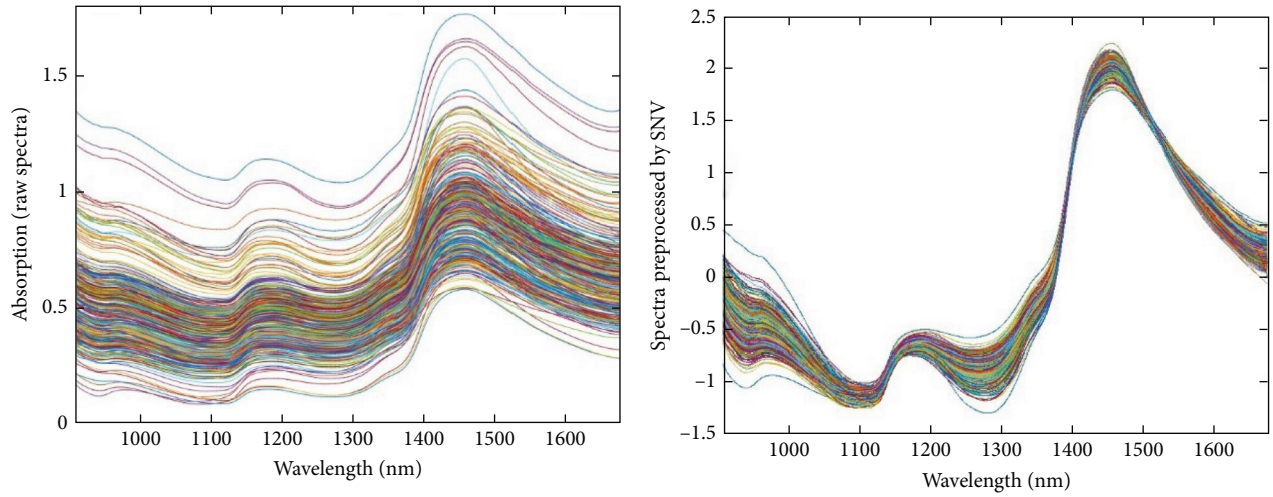
Last, the beta coefficient plot of the best-performed model was used to illustrate the crucial wavebands in the preprocessed spectrum that were significantly contributed to the prediction of grey colouration [42, 43]. Subsequently, based on the preprocessing method, the important wavebands identified in the preprocessed spectrum could facilitate an approximate identification of the corresponding wavebands in raw spectra.

3. Results

3.1. Barramundi Flesh Colour Reference. Lightness (L^*) values for barramundi flesh ranged from 59.6 (darkest) to 94.1 (lightest). Flesh lightness (L^*) across all ventral samples was from 64.7 to 94.1, and the range across all dorsal samples was from 59.6 to 90.3.

3.2. Spectral Features of Raw Spectra and Preprocessed Spectra. The trend in absorption spectra was similar between samples when evaluated grossly across the entire measured wavelength with prominent absorption peaks at approximately 1440–1460 nm, 1165–1180 nm, and 960–990 nm (Figure 2a). A baseline shift was apparent in this data, where the spectral absorption characteristics of each sample were offset across the spectral range. This was rectified through applying SNV clarifying absorption peaks at 900–1050 nm, 1160–1370 nm, and 1550–1660 nm (Figure 2b). The predictive model based on the PLSR method demonstrated the best predictive ability when preprocessed using the first derivative in conjunction with the SG smoothing technique (window size = 5, polynomial degree = 2). This approach outperformed other preprocessing techniques applied to PCR or MLR methods (Table 1). After applying the SG smoothing technique followed by the first derivative, both obvious and subtle features in the raw spectra were significantly enhanced, resulting in sharper and more distinct peaks and valleys. This improvement was particularly noticeable in the spectral regions from 930 to 1100 nm and from 1200 to 1400 nm (Figure 3).

3.3. Calibration, Cross-Validation, and Validation of the Best Models. The optimal number of latent variables was 13 for the PLSR model based on its low RMSECV and small difference between R^2_{ca} and R^2_{cv} (Supporting Information 2). After removing outliers, the final PLSR model trained with 13 latent



(a) (b)

FIGURE 2: (a) Raw skin NIR spectral data from the MicroNIR (pinpoint) device. Three spectral data were recorded in each sample, and each curve line in the plot represented an average NIR spectrum from one fillet sample. (b) Spectral data are preprocessed only by SNV.

variables had a R^2_{cv} of 0.783 and a RMSECV of 2.992. While R^2_{cv} of the final PCR model was 0.768, which was smaller than the R^2_{cv} of PLSR model by 1.91%, the RMSECV was bigger than the RESECV of the PLSR model by 3.41% (Table 1). The difference between R^2_{cv} and R^2_{cv} of best MLR model was the largest among the best PLSR, PCR, and MLR models, and it had a low R^2_{cv} of 0.640 and large RMSECV of 3.857. When using independent data to validate the model, the final PLSR model had the biggest R^2_{pe} of 0.776 and smallest RMSEP of 2.820 (Figure 4). Meanwhile, the quality of the regression model was described by RPD_{pe}, and the highest RPD_{pe} was 2.122 from the final PLSR model.

3.4. Beta Coefficient Analysis. Beta coefficient plots provide physical insights into the importance of each waveband on the prediction of the flesh greyness. In the beta coefficient plot of the best predictive PLSR model, the lowest beta coefficient was 0.98, and it was located at 1372.7 nm, which meant the spectral reading at 1372.7 nm had a strong negative relationship between the lightness of the flesh (Figure 5). In other words, it demonstrated a strong positive relationship with the grey colouration in barramundi flesh. Other notable valleys were located at 1007, 1137, 1242, and 1533 nm. While the two dominant peaks were located at 1335 and 1397 nm with a beta coefficient of 0.85 and 1.03, respectively. So, the lightness of barramundi flesh had a strong positive relationship with spectral reading at the wavelength of 1335 and 1397 nm. The rest obvious peaks were located at 939, 1106, 1211, 1502, and 1626 nm.

4. Discussion

The best prediction model for assessing the greyness in barramundi fillets utilised the PLSR method, combined with preprocessing techniques of SG smoothing and first derivative. In this predictive model, around 78% of the measured colouration could be explained by the predicted colouration in the validation. The RMSEP of the PLSR model indicated that the average magnitude of the error between measured and predicted lightness was 2.820, which was acceptable compared with the given scale of fish colouration and mean lightness of 74.94. This represents a significant advancement in noninvasive methods, as the application of NIR spectra collected from the skin for predicting fillet greyness in fish is shown to have high prediction power for this trait. Previous studies with other food commodities have explored similar methodologies, such as using spectroscopy to determine peach flesh colour from the skin [14], but this is the first time such an approach has been used in fish.

All raw spectra displayed notable baseline shifts, which could be possibly caused by the scattering effect, changes in light conditions/source temperature, or differences in sample thickness [44, 45]. Three notable absorption peaks were observed at 1440–1460 nm, 1170−1190 nm, and 960–990 nm (Figure 2a). Significant absorption peaks around 980 nm and 1450 nm have been reported in previous research when collecting NIR spectra from the flesh samples of different species, such as barramundi, salmon, and chicken, which has been identified and attributed to moisture content [35, 46]. The absorption peaks around 980 and 1450 nm correspond to the second and first overtone of the O—H stretching vibration from the water molecule [47, 48]. So, the wavelength which ranges around 980–1450 nm may not be significant for grey colouration since they have been representing more to moisture, which makes the absorption variation between 1000 and 1400 nm and the wavelength range between 1500 and 1670 nm (Figure 2a) the potential indicators of the chemical components contributing to flesh greyness. This was confirmed later in the beta coefficient plot of the PLSR model that the three lowest beta coefficients were in these two ranges (Figure 5).

Since there is no specific rule about how to find the best preprocessing methods for a specific raw spectral data and regression method [34], many possible preprocessing techniques were tested on PCR, PLSR, and MLR methods,

Table 1: Performance of calibration, cross-validation, and validation of three best final regression models for predicting barramundi flesh colour.

Model	Initial model development								Prediction Validation		
	Method			Calibration		Cross-validation					
	n_{ca}	Regression	Preprocessing	R^2_{ca}	RMSEC	R^2_{cv}	RMSECV	n_{pe}	R^2_{pe}	RMSEP	RPD_{pe}
1	242	PLSR	Smoothing ($w = 5$, $p = 2$); 1st derivative ($w = 7$, $p = 2$)	0.825	2.691	0.783	2.992	79	0.776	2.820	2.122
2	243	PCR	Smoothing ($w = 5$, $p = 2$); 1st derivative ($w = 7$, $p = 2$)	0.808	2.813	0.768	3.094	79	0.771	2.851	2.099
3	247	MLR	SNV; 1st derivative ($w = 15$, $p = 1$)	0.906	1.965	0.640	3.857	79	0.632	3.999	1.496

Note: n_{ca}, the number of spectra to build model (after removing outliers for each model); n_{pe}, number of spectra for validation; R^2_{ca}, coefficient of determination of calibration; R^2_{cv}, coefficient of determination of cross-validation; R^2_{pe}, coefficient of determination of prediction; RPD_{pe}, ratio of performance to deviation of prediction (SD/RMSEP), where SD was the standard deviation of reference (flesh colour). Results of other models are in Supporting Information 1.

Abbreviations: *p*, polynomial order; RMSEC, root mean square error of calibration; RMSECV, root mean square error of cross-validation; RMSEP, root mean square error of prediction; *w*, window.

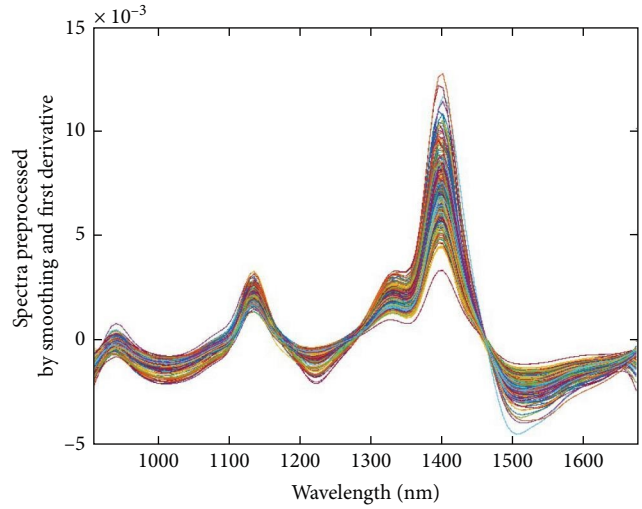


FIGURE 3: Skin spectral data preprocessed by smoothing (Savitzky–Golay, window = 5, polynomial degree = 2) and then the first derivative (Savitzky–Golay, window = 7, polynomial = 2), which was used in the PLSR and PCR model.

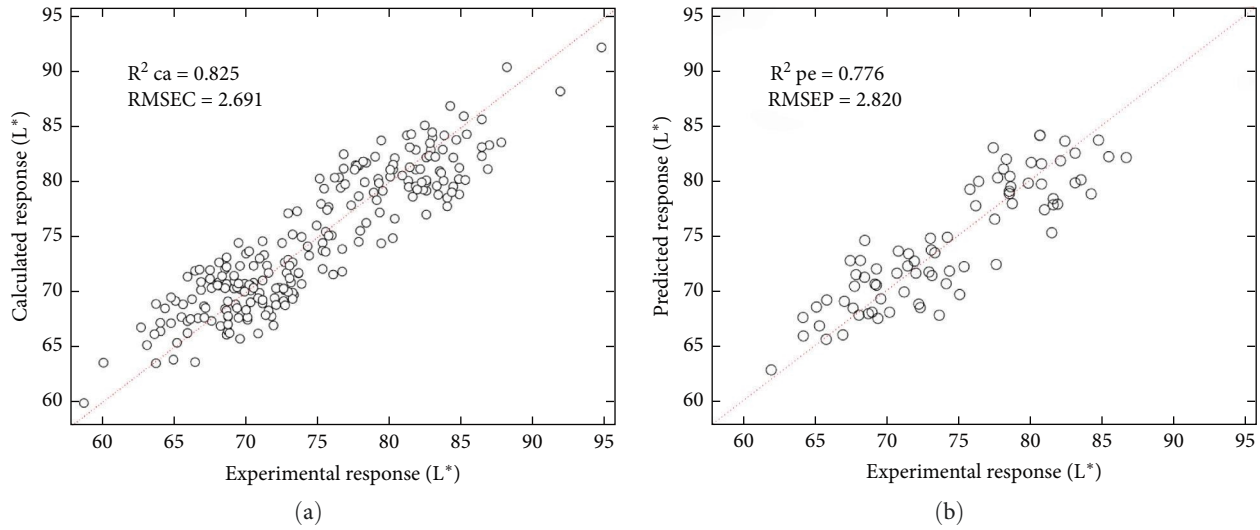


FIGURE 4: Performance of using skin spectral data to predict deep flesh greyness in the PLSR model. Image (a) is a scatter plot of the experimental response vs. the calculated response of the calibration model, and image (b) is the scatter plot of the experimental response vs. the predicted response of the validation.

respectively. Overall, first or second derivative preprocessing techniques could get relatively more robust models (smaller difference in R^2) than the models that use SNV or MSC (Supporting Information 1). In research studying NIR spectra of thawed pork, better regression models were also obtained by using first derivative preprocessing method instead of SNV or MSC [19]. This could be because that first or second derivative could solve the baseline shift issue like SNV and MSC and reduce the problem of overlapping peaks in raw spectral data [49]. For example, more specific peaks or valleys with narrow wavelength ranges could be observed at 957, 1025, 1330, and 1508 nm in Figure 3.

This study aimed to build a robust model for prediction of flesh greyness based on skin spectral data, thus avoiding overfitting noise and outliers in the NIR spectra was essential [50]. A higher R^2_{cv} is more critical than R^2_{ca} because it indicates how good a predictive model can explain the response variables (skin spectra) in the new data set. Meanwhile, a small difference between R^2 of calibration and cross-validation model was important as well when selecting the best predictive model [37]. This was because a bigger difference in R^2 between calibration model and cross-validation model revealed a potential problem of overfitting the calibration model, which might not be predictive in an unseen data set. The PLSR models trained with more than 14 latent variables were therefore not selected for training due to the higher deviation between R^2_{ca} and R^2_{cv} value despite slightly higher R^2_{cv} (Figure 5).

For a predictive model, it is essential to determine the model's reliability and the error when measuring. RPD is a

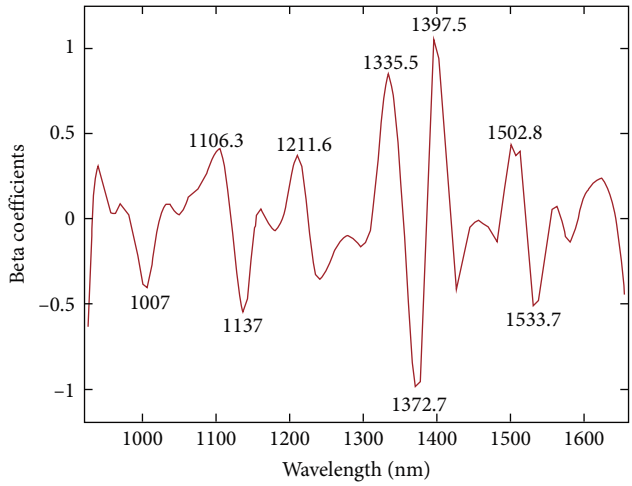


FIGURE 5: Beta coefficients plot of the PLSR model.

measure which looks at the estimated error to the standard deviation of the training samples. If the error of prediction is large compared to the standard deviation, then the model should not be used in prediction [23]. It was first used the quality of the calibration model for prediction with NIR spectra in the field of food science and suggested that RPD between 2 and 3 was adequate for rough screening [51]. As expected, the PLSR model with the highest R^2_{pe} and smallest RMSEP had the biggest RPD_{pe}, while the MLR model with obvious overfitting issues in calibration had a smaller RPD_{pe} that was smaller than 2 (Table 1). Judging from multiple indexes that were related to the performance, such as high R^2_{cv}, small RMSECV, high RPD_{pe}, and slight difference between R^2_{ca} and R^2_{cv}, the best predictive model in this study was the PLSR model preprocessed by smoothing and the first derivative (Table 1). Despite uncertainty about the specific factors or organic compounds causing the grey colouration in barramundi fillets, it was possible to predict the trait of flesh greyness using the skin NIR spectral data.

The NIR spectral wavelength range offered valuable insights into the wavebands corresponding to the first, second, and third overtones of organic functional groups, as well as their combination bands [21]. The wavelengths of the two highest positive and one lowest beta coefficient were located between 1335.5 and 1397.5 nm. The absorption wavebands of combination C—H stretching are around 1300 to 1420 nm [52], which indicated that the greyness of flesh was highly related to some organic compounds. Another group of wavelengths with notable beta coefficients was located at 1502 and 1533.7 nm. This could be caused by the N—H first overtone with an absorption range from 1450 to 1550 nm or overtone bands of secondary amides that have an absorption waveband from 1350 to 1550 nm. The rest of the wavelengths with notable beta coefficient at 1007 nm could be explained by the second overtone N—H or O—H stretching, while the wavelength at 1106 and 1137 nm might be because of the second overtone of C—H stretching based on the summary of Stuart [52]. It was difficult to identify specific organic compounds that are correlated to the grey flesh in barramundi based only on NIR spectra, so further analysis should be focused on MIR or IR spectra because it can detect organic functional groups with higher sensitivity and specificity and providing more distinct absorption peak instead of broad absorption wavebands in NIR spectra [53].

Applying NIR spectral technology to measure the trait of flesh colour in barramundi can be cost-effective and a rapid analytical method because there is no need for filleting the fish and taking images. The successful results from the portable NIR device suggested the potential of using a hyperspectral camera to predict and map out the flesh colour of the whole barramundi fillet and correlate to the spatial variation of the NIR spectra. High-throughput phenotyping of barramundi flesh colour can also be achieved by combining hyperspectral camera technology. In the past, a selective breeding programme that studied the composition of Atlantic salmon required over 8000 fish, and with the help of the NIR device, it could sample up to 400 fish per day [54]. Hence, the practical application of barramundi skin NIR spectra to measure the colouration trait can be useful when the experiments require sampling a lot of fish. This technology is also essential when studying how barramundi flesh colour changes over a period of time under different environmental conditions or nutrients in feed. Without the tool for noninvasive quantification of barramundi flesh colour, the experiment would have to measure the change of flesh colour from different fish through a period of time or only measure the colour at the beginning and ending of the experiment, which could cause either less data or not independent observations for analysis.

5. Conclusions

This study has shown the utility of using a portable NIR device with a scan range from 900 to 1670 nm to predict the flesh colour of barramundi through scanning the skin NIR spectra. The best-performed PLSR model could give a predictive accuracy of 78% with an average error of 2.820 (L^*) between predicted colour and experimental colour in the validation, indicating a strong correlation between the predicted flesh colour and skin NIR spectra. RPD_{pe} was 2.122, which demonstrated the robustness and reliability of applying this model in barramundi for rough colour prediction. Nevertheless, in contexts demanding higher precision, such as aquaculture research, enhancing the predictive model's accuracy is crucial. This improvement can be achieved by incorporating a larger and more diverse set of barramundi samples, particularly with a wider range of flesh colouration variation into training data sets.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Disclosure

The presentation of manuscript has been made available as a preprint, and the link to the preprint is on SSRN [55].

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

This project was funded by the Australian Research Council Industrial Transformation Research Program IH210100014.

Acknowledgments

We thank Dr. Bronson Philippa (James Cook University), Dr. Brett Wedding (Department of Agriculture and Fisheries, Queensland), Steve Grauf, and Russell Withers for their assistance with the NIR device and Dr. Jose Manuel Amigo Rubio for statistical inquiries regarding regression methods in NIR analysis and the use of HYPER-Tools software. This project was funded by the Australian Research Council Industrial Transformation Research Program IH210100014.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information 1. Table A1: models from number 4 to 64 that have been tried with different preprocessing techniques and regression method were displayed in this table. No. = number; cv = cross-validation; PC = principal component; LV = latent variable; w = window; p = polynomial (first, second, or third degree), d = derivative (first or second). Table A2: performance index to show how well each model (from number 4 to 64) could predict barramundi flesh colouration. Difference in $R^2 = R^2_{ca} - R^2_{cv}$; Difference (%) = (Difference in R^2)/R^2_{ca}.

Supporting Information 2. Figure S1: bar plot of RMSECV vs number of latent variables for PLSR model. The *Y*-axis represents the root mean square error of the cross-validation model, and the *X*-axis represents the number of latent variables used in training the PLSR model. Figure S2: $|R^2_{ca} - R^2_{cv}|$ versus R^2_{cv} plot, where the label next to each point indicated the number of latent variables used in PLSR analysis. R^2_{ca} represented the coefficient of determination for calibration, and R^2_{cv} was the coefficient of determination for cross-validation.

References

[1] D. R. Jerry, *Biology and Culture of Asian Seabass Lates calcarifer* (Taylor & Francis, 2014).

[2] FAO, *The State of World Fisheries and Aquaculture 2022* (Rome: Food and Agriculture Organization of the United Nations, 2022).

[3] J. Howieson, B. Glencross, S. Little, A. Ais, et al., "Understanding & Minimising "Greying" of Farmed Barramundi Fillets," Project No. 2011/721 Australian Seafood Cooperative Research Centre, 2013.

[4] A. H. Pálsdóttir, A. Bläuer, E. Rannamäe, S. Boessenkool, and J. H. Hallsson, "Not a Limitless Resource: Ethics and Guidelines for Destructive Sampling of Archaeofaunal Remains," *Royal Society Open Science* 6, no. 10 (2019): 191059–191059.

[5] R. Li, G. Rui, W. Chen, S. Li, P. E. Schulz, and Y. Zhang, "Early Detection of Alzheimer's Disease Using Non-Invasive Near-Infrared Spectroscopy," *Frontiers in Aging Neuroscience* 10 (2018): 366–366.

[6] P. Pallav, G. G. Diamond, D. A. Hutchins, R. J. Green, and T. H. Gan, "Near-Infrared (NIR) Technique for Imaging Food Materials," *Journal of Food Science* 74, no. 1 (2009): E23–E33.

[7] J. U. Porep, D. R. Kammerer, and R. Carle, "On-Line Application of Near Infrared (NIR) Spectroscopy in Food Production," *Trends in Food Science & Technology* 46, no. 2 (2015): 211–230.

[8] G. Reich, "Near-Infrared Spectroscopy and Imaging: Basic Principles and Pharmaceutical Applications," *Advanced Drug Delivery Reviews* 57, no. 8 (2005): 1109–1143.

[9] M. Blanco, J. Cruz, and M. Bautista, "Development of a Univariate Calibration Model for Pharmaceutical Analysis Based on NIR Spectra," *Analytical and Bioanalytical Chemistry* 392, no. 7-8 (2008): 1367–1372.

[10] Y. Zhang, J. F. Nock, Y. Al Shoffe, and C. B. Watkins, "Non-Destructive Prediction of Soluble Solids and Dry Matter Contents in Eight Apple Cultivars Using Near-Infrared Spectroscopy," *Postharvest Biology and Technology* 151 (2019): 111–118.

[11] M. H. Lee, A. G. Cavinato, D. M. Mayes, and B. A. Rasco, "Noninvasive Short-Wavelength Near-Infrared Spectroscopic Method to Estimate the Crude Lipid Content in the Muscle of Intact Rainbow Trout," *Journal of Agricultural and Food Chemistry* 40, no. 11 (1992): 2176–2181.

[12] J. P. Wold and T. Isaksson, "Non-Destructive Determination of Fat and Moisture in Whole Atlantic Salmon by Near-Infrared Diffuse Spectroscopy," *Journal of Food Science* 62, no. 4 (1997): 734–736.

[13] D. Cozzolino, N. Barlocco, A. Vadell, F. Ballesteros, and G. Gallieta, "The Use of Visible and Near-Infrared Reflectance Spectroscopy to Predict Colour on Both Intact and Homogenised Pork Muscle," *LWT—Food Science and Technology* 36, no. 2 (2003): 195–202.

[14] D. C. Slaughter, C. H. Crisosto, and G. Tiwari, "Nondestructive Determination of Flesh Color in Clingstone Peaches," *Journal of Food Engineering* 116, no. 4 (2013): 920–925.

[15] B. B. Wedding, C. Wright, S. Grauf, P. Gadek, and R. D. White, "The Application of FT-NIRS for the Detection of Bruises and the Prediction of Rot Susceptibility of 'Hass' Avocado Fruit," *Journal of the Science of Food and Agriculture* 99, no. 4 (2019): 1880–1887.

[16] R. Fujii, "The Regulation of Motile Activity in Fish Chromatophores," *Pigment Cell Research* 13, no. 5 (2000): 300–319.

[17] M. Luo, G. Lu, H. Yin, L. Wang, M. Atuganile, and Z. Dong, "Fish Pigmentation and Coloration: Molecular Mechanisms and Aquaculture Perspectives," *Reviews in Aquaculture* 13, no. 4 (2021): 2395–2412.

[18] C. Chariskou, E. Vrochidou, A. J. Daniels, and V. G. Kaburlasos, "Variable Selection on Reflectance NIR Spectra for the Prediction of TSS in Intact Berries of Thompson Seedless Grapes," *Agronomy* 12, no. 9 (2022): 2113.

[19] Q. Ouyang, L. Liu, M. Zareef, L. Wang, and Q. Chen, "Application of Portable Visible and Near-Infrared Spectroscopy for Rapid Detection of Cooking Loss Rate in Pork: Comparing Spectra from Frozen and Thawed Pork," *Food Science & Technology* 160 (2022): 113304.

[20] G. Spielbauer, P. Armstrong, J. W. Baier, et al., "High-Throughput Near-Infrared Reflectance Spectroscopy for Predicting Quantitative and Qualitative Composition Phenotypes of

Individual Maize Kernels," *Cereal Chemistry* 86, no. 5 (2009): 556–564.
[21] E. W. Ciurczak, *Handbook of Near-Infrared Analysis* (CRC Press, Fourth edition, 2021).
[22] L. Wu, J. Hu, Q. Zou, et al., "Synthesis and Optical Properties of a Y_3 $(Al/Ga)_5O_{12}$:Ce^{3+},Cr^{3+},Nd^{3+} Persistent Luminescence Nanophosphor: A Promising Near-Infrared-II Nanoprobe for Biological Applications," *Nanoscale* 12, no. 26 (2020): 14180–14187.
[23] B. Wedding, "The Non-Invasive Assessment of Avocado Maturity and Quality James Cook University]," (2018).
[24] J. Lammertyn, A. Peirs, J. De Baerdemaeker, and B. Nicolaï, "Light Penetration Properties of NIR Radiation in Fruit With Respect to Non-Destructive Quality Assessment," *Postharvest Biology and Technology* 18, no. 2 (2000): 121–132.
[25] S. E. Border, T. J. Piefke, R. J. Fialkowski, et al., "Color Change and Pigmentation in a Color Polymorphic Cichlid Fish," *Hydrobiologia* 832, no. 1 (2019): 175–191.
[26] H. N. Sköld, T. Amundsen, P. A. Svensson, I. Mayer, J. Bjelvenmark, and E. Forsgren, "Hormonal Regulation of Female Nuptial Coloration in a Fish," *Hormones and Behavior* 54, no. 4 (2008): 549–556.
[27] S. Sunoj, C. Igathinathane, N. Saliendra, J. Hendrickson, and D. Archer, "Color Calibration of Digital Images for Agriculture and Other Applications," *ISPRS Journal of Photogrammetry and Remote Sensing* 146 (2018): 221–234.
[28] F. Fernando, C. L. Candebat, J. M. Strugnell, N. Andreakis, and L. Nankervis, "Dietary Supplementation of Astaxanthin Modulates Skin Color and Liver Antioxidant Status of Giant Grouper (*Epinephelus lanceolatus*)," *Aquaculture Reports* 26 (2022): 101266.
[29] C. G. Andresen, C. E. Tweedie, and V. L. Lougheed, "Climate and Nutrient Effects on Arctic Wetland Plant Phenology Observed From Phenocams," *Remote Sensing of Environment* 205 (2018): 46–55.
[30] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 Years of Image Analysis," *Nature Methods* 9, no. 7 (2012): 671–675.
[31] The MathWorks Inc, "MATLAB Version: 9.13.0 (R2022b), Natick, Massachusetts: The MathWorks Inc." 2022, https://www.mathworks.com.
[32] N. Mobaraki and J. M. Amigo, "HYPER-Tools. A Graphical User-Friendly Interface for Hyperspectral Image Analysis," *Chemometrics and Intelligent Laboratory Systems* 172 (2018): 174–187.
[33] X. Bian, K. Wang, E. Tan, P. Diwu, F. Zhang, and Y. Guo, "A Selective Ensemble Preprocessing Strategy for Near-Infrared Spectral Quantitative Analysis of Complex Samples," *Chemometrics and Intelligent Laboratory Systems* 197 (2020): 103916.
[34] Å. Rinnan, F. den Berg, and S. B. Engelsen, "Review of the Most Common Pre-Processing Techniques for Near-Infrared Spectra," *TrAC Trends in Analytical Chemistry* 28, no. 10 (2009): 1201–1222.
[35] H.-J. He, D. Wu, and D.-W. Sun, "Non-Destructive and Rapid Analysis of Moisture Distribution in Farmed Atlantic Salmon (*Salmo salar*) Fillets Using Visible and Near-Infrared Hyperspectral Imaging," *Innovative Food Science & Emerging Technologies* 18 (2013): 237–245.
[36] M. Urbano-Cuadrado, M. D. Luque de Castro, P. M. Pérez-Juan, J. García-Olmo, and M. A. Gómez-Nieto, "Near Infrared Reflectance Spectroscopy and Multivariate Analysis in Enology: Determination or Screening of Fifteen Parameters in Different Types of Wines," *Analytica Chimica Acta* 527, no. 1 (2004): 81–88.
[37] K. M. Mendez, S. N. Reinke, and D. I. Broadhurst, "A Comparative Evaluation of the Generalised Predictive Ability of Eight Machine Learning Algorithms Across Ten Clinical Metabolomics Data Sets for Binary Classification," *Metabolomics* 15, no. 12 (2019): 150–115.
[38] R. E. Shaffer, "Multi- and Megavariate Data Analysis," in *Principles and Applications*, eds. I. Eriksson, E. Johansson, N. Kettaneh-Wold, and S. Wold, 16, no. 5, (Umetrics Academy, Umeå, 2002): 261–262.
[39] M. R. Brown, "Rapid Compositional Analysis of Oysters Using Visible-Near Infrared Reflectance Spectroscopy," *Aquaculture* 317, no. 1–4 (2011): 233–239.
[40] H. L. Galasso, M. D. Callier, D. Bastianelli, J.-P. Blancheton, and C. Aliaume, "The Potential of Near Infrared Spectroscopy (NIRS) to Measure the Chemical Composition of Aquaculture Solid Waste," *Aquaculture* 476 (2017): 134–140.
[41] Z. Li, H. Qi, Y. Yu, et al., "Near-Infrared Spectroscopy Method for Rapid Proximate Quantitative Analysis of Nutrient Composition in Pacific Oyster *Crassostrea gigas*," *Journal of Oceanology and Limnology* 41, no. 1 (2023): 342–351.
[42] R. Joshi, R. Sathasivam, P. K. Jayapal, et al., "Comparative Determination of Phenolic Compounds in *Arabidopsis thaliana* Leaf Powder Under Distinct Stress Conditions Using Fourier-Transform Infrared (FT-IR) and Near-Infrared (FT-NIR) Spectroscopy," *Plants* 11, no. 7 (2022): 836.
[43] J. Lim, G. Kim, C. Mo, et al., "Detection of Melamine in Milk Powders Using Near-Infrared Hyperspectral Imaging Combined With Regression Coefficient of Partial Least Square Regression Model," *Talanta* 151 (2016): 183–191.
[44] X. Luo, X. Yu, X. Wu, Y. Cheng, and H. Qu, "Rapid Determination of Paeoniae Radix Using Near Infrared Spectroscopy," *Microchemical Journal* 90, no. 1 (2008): 8–12.
[45] F. Zhang, X. Tang, and L. Li, "Origins of Baseline Drift and Distortion in Fourier Transform Spectra," *Molecules* 27, no. 13 (2022): 4287.
[46] D. Samuel, B. Park, M. Sohn, and L. Wicker, "Visible-Near-Infrared Spectroscopy to Predict Water-Holding Capacity in Normal and Pale Broiler Breast Meat," *Poultry Science* 90, no. 4 (2011): 914–921.
[47] B. Bowker, S. Hawkins, and H. Zhuang, "Measurement of Water-Holding Capacity in Raw and Freeze-Dried Broiler Breast Meat With Visible and Near-Infrared Spectroscopy," *Poultry Science* 93, no. 7 (2014): 1834–1841.
[48] D. Wu, D.-W. Sun, and Y. He, "Novel Non-Invasive Distribution Measurement of Texture Profile Analysis (TPA) in Salmon Fillet by Using Visible and near Infrared Hyperspectral Imaging," *Food Chemistry* 145 (2014): 417–426.
[49] A. Ambrose, S. Lohumi, W.-H. Lee, and B. K. Cho, "Comparative Nondestructive Measurement of Corn Seed Viability Using Fourier Transform Near-Infrared (FT-NIR) and Raman Spectroscopy," *Sensors and Actuators B: Chemical* 224 (2016): 500–506.
[50] A. D. Rocha, T. A. Groen, A. K. Skidmore, R. Darvishzadeh, and L. Willemen, "The Naïve Overfitting Index Selection (NOIS): A New Method to Optimize Model Complexity for Hyperspectral Data," *ISPRS Journal of Photogrammetry and Remote Sensing* 133 (2017): 61–74.
[51] P. Williams, "Grains and Seeds," in *Near-Infrared Spectroscopy in Food Science and Technology*, eds. Y. Ozaki, W. F. McClure, and A. A. Christy, (John Wiley & Sons, Inc., New Jersey, United States of America, 2007): 165–218.
[52] B. H. Stuart, *Infrared Spectroscopy Fundamentals and Applications* (Wiley, 2004).

[53] J. M. Soriano-Disla, L. J. Janik, R. A. Viscarra Rossel, L. M. Macdonald, and M. J. McLaughlin, "The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties," *Applied Spectroscopy Reviews* 49, no. 2 (2014): 139–186.

[54] N. G. Elliott and P. D. Kube, "Development and Early Results of the Tasmanian Atlantic Salmon Breeding Program," *Association for the Advancement of Animal Breeding and Genetics* 18 (2009): 362–365.

[55] W. Chen, D. R. Jerry, R. D. White, and L. Nankervis, "Non-Invasive Prediction of Barramundi (*Lates Calcarifer*) Flesh Lightness Using Near-Infrared Spectroscopy," 2024, https://ssrn.com/abstract=4758743.