



Underwater hyperspectral imaging technology has potential to differentiate and monitor scallop populations

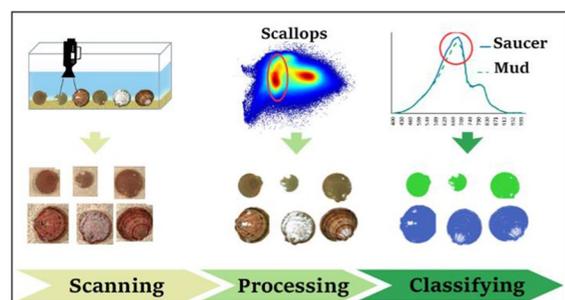
Iman Tahmasbian · Matthew N. McMillan · Jonathan Kok · Anthony J. Courtney

Received: 6 March 2023 / Accepted: 31 October 2023 / Published online: 4 January 2024
© Crown 2024

Abstract Accurate and low-impact monitoring of scallop abundance is critical for stock assessment, especially in sensitive habitats. The possibility of using low-impact hyperspectral imaging (HSI) for differentiating scallop species in the marine environment was investigated. Live saucer (*Ylistrum balloti*) and mud (*Ylistrum pleuronectes*) scallops (N = 31) were scanned inside a sea simulator using a visible to near infrared (400–1000 nm) line-scanner HSI camera. Partial least square discriminant analysis (PLS-DA) was trained to distinguish between the species using their spectral signatures. Important wavelengths were identified and new models were developed using these wavelengths to reduce the model complexity and potentially increase the imaging speed when applied under at-sea conditions. The PLS-DA model distinguished between saucer and mud scallops using

any area of the left valve that was exposed above the sediments, with 90.73% accuracy when all 462 available wavelengths were used. Using the subset of important wavelengths (N = 13) reduced the classification accuracy to 84%. Overall, our results showed that HSI has potential for detecting, distinguishing and counting commercially important saucer scallops for low-impact monitoring and resource management, and to complement RGB imaging that relies solely on morphological properties.

Graphical abstract



Scopus affiliation ID: 60028929

I. Tahmasbian (✉)
Department of Agriculture and Fisheries, Queensland
Government, Toowoomba, QLD 4350, Australia
e-mail: iman.tahmasbian@daf.qld.gov.au

M. N. McMillan · A. J. Courtney
Sustainable Fisheries, Animal Science Queensland,
Department of Agriculture and Fisheries, EcoSciences
Precinct, Dutton Park, QLD 4102, Australia

J. Kok
Australian Institute of Marine Science, Townsville,
QLD 4810, Australia

Keywords Hyperspectral imaging · Non-destructive · Non-invasive · Scallop · Sustainable management · VNIR

Introduction

Scallops are economically valuable marine bivalves with approximately 400 known species, inhabiting coastal marine environments worldwide (Brand 2016; Delargy et al. 2023; Stewart and Howarth 2016). The combined factors of rising demand for wild-captured scallops and the impacts of climate change have added pressure on scallop stocks (Cheung et al. 2013; FAO 2020; Zang et al. 2023). Due to their commercial significance and ecological importance, accurate assessment and monitoring of scallop populations are crucial for their sustainable management.

Trawling and dredging are the most used methods for surveying scallop populations. However, these methods can impose habitat impacts and incidental mortality on bycatch species (Delargy et al. 2023; Freese et al. 1999; Kaiser and Spencer 1996; Probert et al. 1997; Thrush and Dayton 2002). Visual survey alternatives, such as towed camera systems, remotely operated vehicles and autonomous underwater vehicles (Courtney et al. 2021; Miller et al. 2019; Taylor et al. 2008) impose fewer impacts, but the resulting census estimates are heavily reliant upon accurately interpreting the collected images. This includes the ability to differentiate between morphologically similar species. Additionally, imaging methods can be problematic for species that partially bury in sediments, such as scallops. Therefore, there is a need to explore alternative approaches that can effectively overcome the constraints associated with the existing methods.

Hyperspectral imaging (HSI) is a low-impact technology that combines conventional imaging and near infrared spectroscopy (NIRS), providing both spatial and spectral information (Grahn and Geladi 2007; Manley 2014). HSI utilises the reflection of light across hundreds of narrow wavelengths, encompassing both visible and non-visible ranges, to identify and measure materials based on their interaction with light (Park and Lu 2015). When light interacts with materials, it generates a characteristic pattern within the measured reflected spectrum, known as a spectral signature. Spectral signatures are specific to each material and can be leveraged, in combination with multivariate analysis and machine learning techniques, to identify and distinguish between various materials (Manley 2014). The spectral signature obtained through HSI offers an advantage by

minimising reliance on morphology-based classification methods. As a result, it provides enhanced capability to differentiate between species that share similar morphological characteristics. In addition, utilising the spectral signature enables identifying species that may be partially buried in sediment, such as scallops, through analysing the signature of parts that remain exposed.

Hyperspectral imaging has been applied in aquatic environments, including determining the distribution of microphytobenthos, mapping of benthic habitats and monitoring the health of sea corals (Chennu et al. 2013, 2017; Letnes et al. 2019; Montes-Herrera et al. 2021). Applications of HSI for analysing harvested shellfish in dry laboratory conditions include differentiating between the oyster valves grown in four different environments (Mehrübeoglu et al. 2013) and two relatively similar species of oysters grown in the same environment (Tahmasbian et al. 2022). While these studies provide valuable information on the possibility of using HSI in aquatic environments and for differentiating between harvested oysters, there is no evidence that HSI can differentiate between living scallops that are partly buried in sediments in aquatic environments.

In this paper we evaluated HSI as a potential alternative towed camera methodology for identifying saucer scallops (*Ylistrum balloti*) in their marine environments. Saucer scallops have been the main target species of a trawl fishery on the central Queensland coast, Australia, for several decades (Dredge et al. 2016). We specifically compared the spectral signatures of live saucer scallops with those of mud scallop (*Ylistrum pleuronectes*), a similar co-occurring but smaller and less valuable species. We hypothesised that these two scallop species have distinct spectral signatures that can be used to differentiate between them independent of their morphological similarities, using any area of their left valve that is exposed above the sediment.

Materials and methods

Sampling procedure

Samples of saucer and mud scallops were obtained from a commercial trawl vessel off the Townsville region of Queensland, Australia, in June 2020 on

two occasions to increase the variability among the scallops. Immediately after capture, the samples were transferred to 60 L holding tanks on board the vessel and kept oxygenated by maintaining pumped seawater to the tanks. The tanks were covered to reduce animal stress during transport to land which lasted approximately 12 h. Onshore, samples were transferred to 32 L containers filled with aerated seawater of ambient temperature for transport to the Australian Institute for Marine Science (AIMS) National Sea Simulator (SeaSim) facility at Townsville. At the AIMS SeaSim facility, the scallops were transferred to aerated holding tanks (Fig. 1a) filled with pumped flow-through seawater maintained at 20 °C and acclimated overnight prior to obtaining the in-situ hyperspectral images.

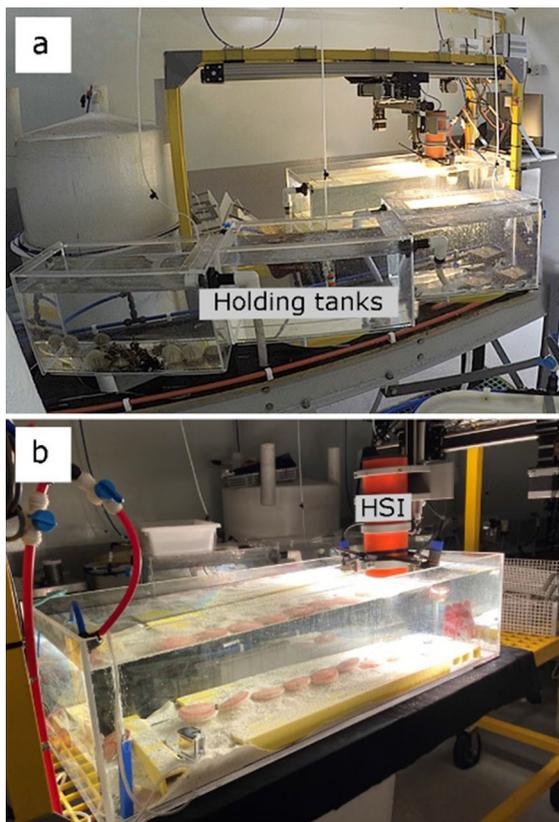


Fig. 1 Australian Institute of Marine Science (AIMS) Live Aquarium holding tanks (a) and hyperspectral imaging (HSI) data collection setup (b)

Hyperspectral imaging system and image acquisition procedure

The AIMS Live Aquarium HSI data collection setup was comprised of an imaging tank filled with 5 cm of sediment containing medium to coarse sand derived from the same habitat as the scallops. The tank was supplied with flow-through seawater at 20 °C and was equipped with a rail-mounted HSI system fitted above the tank (Fig. 1b).

The HSI system consisted of various components, including a 12-bit line scanner visible-near infrared (VNIR, 400–1000 nm) camera (Pika XC2, Resonon, USA) with a quartz lens. The camera had a sampling resolution of approximately 1.3 nm, generating 461 wavelengths. It was securely housed in a waterproof enclosure. Additionally, the system included two direct current (DC) DeepSea MultiLite underwater halogen lights (100 W), a computer numerical control router, and a Spectralon® reflectance target (Labsphere Inc., USA) capable of reflecting 99% of the incident light.

Four to eight live saucer and mud scallops were randomly transferred into the imaging tank (Fig. 1b) and spaced out in the middle of the tank within the field of view of the camera. The left brown pigmented valves of the scallops were faced up while the white or right-side valves were faced down, reflecting their natural orientation in the wild. The scallops were scanned immediately at 50 frames per second and exposure time of 17.82 ms. To ensure precise imaging and to maintain a stable lighting environment, the room's alternating current (AC) fluorescent lights were switched off during the imaging process. This prevented stray light interfering with the camera's observations and potential degradation of image quality due to the high-frequency flickering fluorescence. Instead, DC halogen lights, were utilised to maintain a steady and reliable lighting environment during the imaging process. After imaging, the samples were returned to the holding tank. This process was repeated with new individuals until all individuals were scanned.

The HSI system was controlled using a MATLAB program.

The spectral corrections were performed using Eq. 1 (Farrar et al. 2023).

$$R = \frac{R_0 - D}{W - D} \quad (1)$$

where R is the corrected/relative reflectance of the samples, R_0 is the raw reflectance, D is the dark current of the camera captured with the lights off and the lens cap on, and W is the reflectance of the white Spectralon® calibration target.

Image pre-processing and background removal

Low quality/noisy images were removed from the dataset. The images from a total of 31 scallops, including 8 saucer and 23 mud scallops, were used for further analysis. The original images that contained multiple scallops were cut into smaller images, each containing one scallop. The images were then imported into Evince software (Version 2.7.11, Prediktera, Sweden) for image pre-processing and data analysis. The images were subjected to size reduction, in Evince, by cropping the labels and sediments, leaving only the scallops and their immediate surrounding sediments. The last step of size reduction was conducted by using every tenth wavelength, retaining 47 wavelengths (spectral dimension), and every second column and row (spatial dimension). The above-mentioned size reduction procedures were performed to reduce data size for computation. This process reduced the data size from 80.1 to 1.4 GB.

A principal component analysis (PCA) model with three components was used to remove shading errors, saturated pixels, edge effects and the remaining background sediments (Tahmasbian et al. 2021a; Williams et al. 2009). PCA is a statistical technique used for simplifying and exploring complex datasets by extracting the important information and representing it as a set of new orthogonal variables called principal components (Wold et al. 1987). In addition to traditional applications in data dimension reduction, PCA is utilised for visualising and overviewing variations in datasets through clustering the related samples (Wold et al. 1987). In this study, we employed PCA to project samples (pixels in the images) onto vectors t . Plotting the first few t vectors displayed the pixel patterns in the images that were clustered based on their spectral similarities. The clusters were then used to identify and remove unwanted pixels from the images.

Training and evaluating classification algorithm

The scallops were divided into calibration and test datasets (Fig. 2). Five saucer and fifteen mud scallops were assigned to the calibration set and the remaining three and eight were assigned to the test set (63–65% calibration vs. 35–37% test). The test samples were mainly selected from the scallops that were partly buried for more stringent evaluation of the trained models (Fig. 2).

A partial least square discriminant analysis (PLS-DA) model was trained using the calibration dataset to discriminate saucer scallops from mud scallops using their spectral signatures. PLS-DA combines dimensionality reduction and discriminant analysis into one algorithm and is especially applicable to modelling high dimensional intercorrelated datasets, such as HSI data (Lee et al. 2018). PLS-DA is based on the partial least square regression (PLSR) algorithm that searches for latent variables (LVs) with a maximum covariance with the Y-variables, which in this case were categorical (Ballabio and Consonni 2013; Wold 1966; Wold et al. 1984).

Data transformation algorithms such as, mean-centring, standard normal variation (SNV), multiple scatter correction (MSC) and Svitsky–Golay derivative (SG) were used and compared against un-transformed data (Chen et al. 2021; Farrar et al. 2021; Rinnan et al. 2009). The PLS-DA model was initially trained with all pixels of both saucer and mud scallops in the calibration sets and then compared with another PLS-DA model that was developed using a balanced class-size. Using the balanced class-size data gives data from both species an equal chance to influence the model and prevent a biased calibration towards the larger class (Brereton and Lloyd 2014; Chicco and Jurman 2020). In this study, there were more data points for the mud scallops and, therefore, the Equal-Size Class function in Evince was used to remove some pixels from the mud scallops until the number of pixels were comparable with those of saucer scallops. Since the performance of the balanced class-size model was significantly better, only the balanced class-size model is reported and discussed.

A random k-fold ($k=10$) cross-validation was used for selecting the optimal number of LVs for the PLS-DA model to avoid over/under fitting (Tahmasbian et al. 2021b). The 10-fold cross-validation divided the image pixels into 10 random groups,

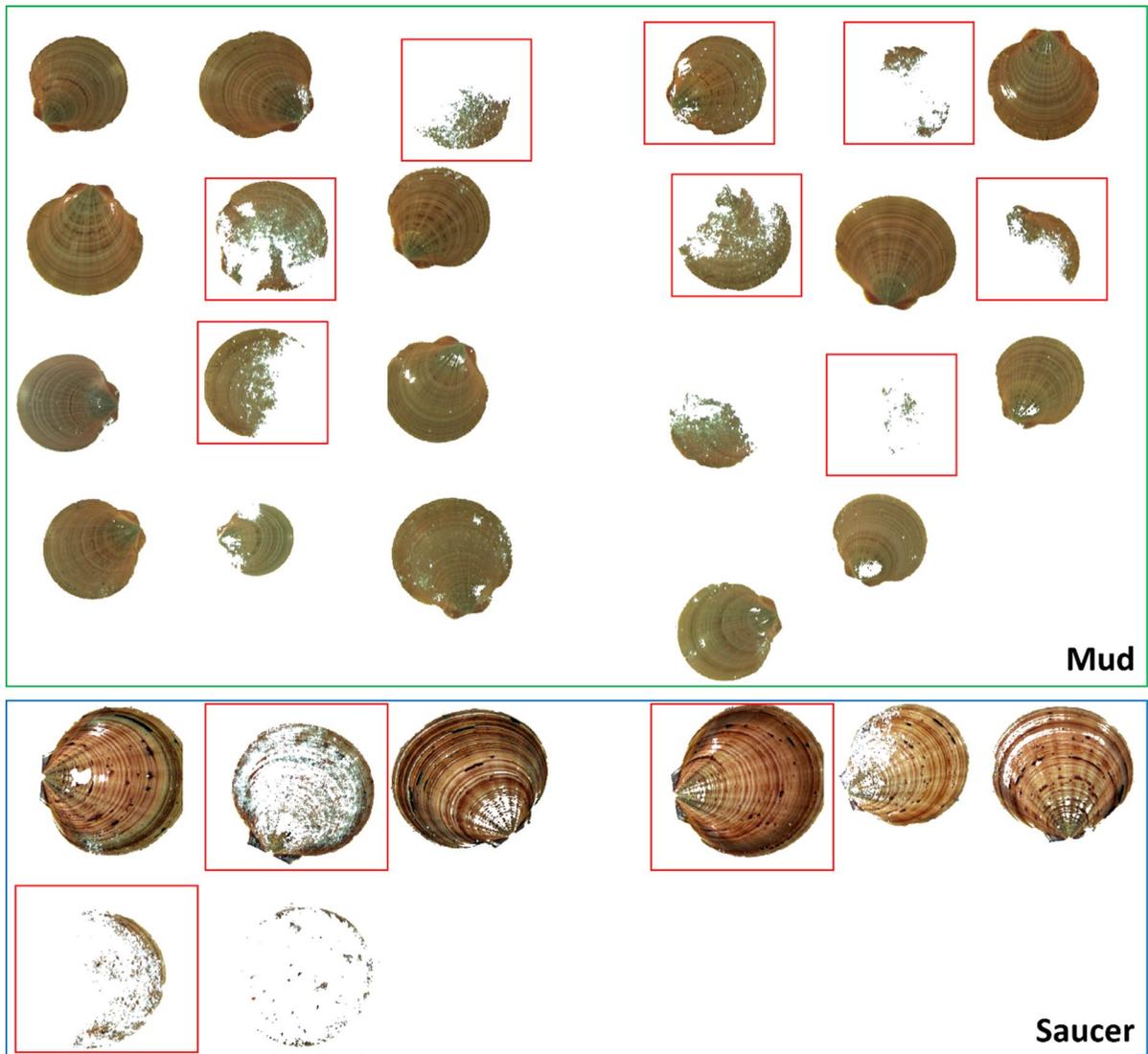


Fig. 2 The RGB image of saucer and mud scallops after removing background pixels. The samples in the boxes were the test samples and the remaining were calibration samples

trained the model with nine groups and tested it with the tenth group. This process was repeated until each of the 10 groups was left out of the models once. The number of LVs was selected where the coefficient of determination (R^2) of the cross validation reached its maximum.

A confusion matrix (Table 1) of the test data set classified using the PLS-DA model was used to evaluate the performance of the model using Eqs. 2–7 (Da Conceição et al. 2021; Sokolova et al. 2006):

$$\text{Classification accuracy (\%)} = \frac{TP + TN}{Total} \times 100 \quad (2)$$

$$\text{False Positive Error (\%)} = \frac{FP}{Total} \times 100 \quad (3)$$

$$\text{False Negative Error (\%)} = \frac{FN}{Total} \times 100 \quad (4)$$

Table 1 Confusion matrix representing the number of pixels classified using the PLS-DA model (LV=7) versus their actual classes, and metrics used for evaluating the performance of

PLS-DA for classifying saucer from mud scallop in the test data set using 7 LVs and 14 LVs

Classes		Predicted (pixels)		
		Mud	Saucer	Unclassified
	Mud	95,751	4724	495
	Saucer	12,194	76,393	178
			LV = 7	LV = 14
	False positive error (%)		2.49	1.63
	False negative error (%)		6.43	5.19
	Sensitivity (%)		86.24	88.90
	Specificity (%)		95.30	96.94
	Precision (%)		94.18	96.23
	Mathews Correlation Coefficient (MCC)		0.82	0.86
	Classification accuracy (%)		90.73	92.90

LV Latent variable

$$\text{Sensitivity (\%)} = \frac{TP}{TP + FN} \times 100 \quad (5)$$

$$\text{Specificity (\%)} = \frac{TN}{TN + FP} \times 100 \quad (6)$$

$$\text{Precision (\%)} = \frac{TP}{TP + FP} \times 100 \quad (7)$$

where TP is true positive, TN is true negative, FP is false positive and FN is false negative.

We also used Mathews Correlation Coefficient (MCC, Eq. 8) to investigate the quality of the predictions. MCC is a more reliable statistical index compared to the classification accuracy (Eq. 2), because it involves multiple statistical measures represented in Eqs. 2–7 (Chicco and Jurman 2020; Park et al. 2021). MCC values vary between –1 and +1 with the completely accurate classification equal to +1 (Chicco and Jurman 2020).

$$\text{MCC} = \frac{(TP.TN) - (FP.FN)}{\sqrt{(TP + FP).(TP + FN).(TN + FP).(TN + FN)}} \quad (8)$$

Important wavelength selection

The β -coefficient generated by the PLS-DA model and the variable importance in projection (VIP) values (Eq. 9) were used individually and combined,

to identify the wavelengths that were important for the classification (Chong and Jun 2005; Hohrenk-Danzouma et al. 2022). The β -coefficient important wavelengths were selected where the wavelength's coefficient (absolute value) was larger than the standard deviation of the coefficients (β/β -standard deviation > 1). The VIP wavelengths larger than 1 were selected (Li et al. 2006; Tahmasbian et al. 2018; Wold 1995).

$$\text{VIP}_j = \sqrt{\frac{\sum_{f=1}^F w_{jf}^2 \cdot \text{SSY}_f \cdot J}{\text{SSY}_t \cdot F}} \quad (9)$$

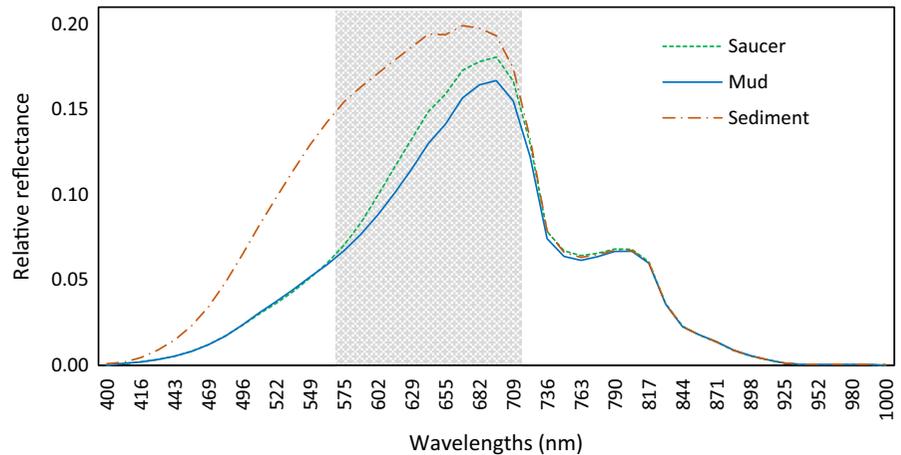
where, VIP_j was the value of j th wavelength in the model with F number of components, w_{jf} was the loading weight of the corresponding wavelength in the f th component, SSY_f was the explained sum of squares of the target variable in the f th component, SSY_t was the total sum of squares of the target variable and J was the total number of wavelengths.

Results and discussion

Spectral features

The average relative reflectance (spectral signatures) of saucer and mud scallops overlapped between 400 and 550 nm, and between 800 and 1000 nm (Fig. 3).

Fig. 3 Average relative reflectance (spectral signature) of the saucer scallops, mud scallops and sediments. The highlighted area represents the spectral region where significant differences between the saucer and mud scallops were observed



The spectral signatures of both saucer and mud scallops showed relatively similar pattern, increasing from 400 to 695 nm (major peak) followed by a sharp decline from 695 to 749 nm. There was a minor valley-plateau between 749 and 817 nm followed by another sharp decline to 857 nm. The relative reflectance of both species gradually declined from 857 to 1000 nm (Fig. 3). The magnitude of saucer scallops' reflectance was higher than that of mud scallops between 575 and 709 nm. This spectral region can be used to differentiate saucer and mud scallops (see “[Classification properties and important wavelengths](#)” section).

The spectral signature of sediments, however, exhibited notable dissimilarities, both in pattern and magnitude, when compared to those of saucer and mud scallops. These differences were prominently observed in the range of 400–709 nm (Fig. 3). Beyond this range, a relatively similar pattern was observed. The discernible spectral signature of sediments can be used to distinguish between the sediment and any part of the scallop valve exposed above the sediment.

Removing background using PCA

The PCA model used the mean-centred data for three components to distinguish and separate scallops from their background pixels (i.e., shade and sediments). The PCA explained 99.1% of the variation with the first component explaining 90.1%, the second component explaining 8.3% and the third component explaining 0.7%. The projection of principal component score vectors t_1 vs. t_3 was used to identify and

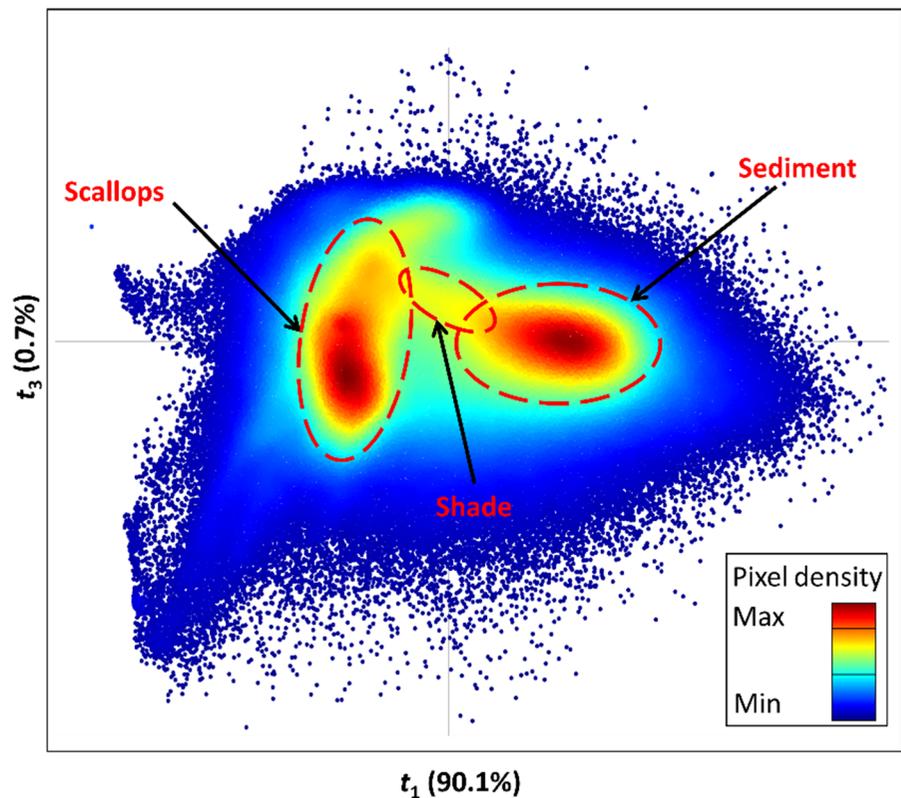
remove the background pixels, enabling the separation of shady areas, which was not possible using the projection of t_1 vs. t_2 (Fig. 4). This high (99.1%) value of the variance explained in three components was expected due to the distinctive spectral signatures of the scallops and their background.

Classification properties and important wavelengths

The data were only mean-centred for training the PLS-DA model. Using the balanced number of pixels improved the accuracy of the classification. The highest R^2 of the PLS-DA model was achieved at 14 LVs (Table 1). However, increasing the classification accuracy beyond 7 LVs had negligible benefit (Table 1). For example, the classification accuracy increased from 90.73 to 92.9% (2.17%) and Matthews correlation coefficient increased from 0.82 to 0.86 (0.04) when the number of LVs increased from 7 to 14 (Table 1). Therefore, 7 LVs was chosen for the final classification model, to maintain simplicity of the calculations, and used to classify the test scallop species using their HSI data (Fig. 5).

The PLS-DA model identified individual pixels and distinguished between them using their spectral signatures (Fig. 5). This enabled using HSI data from any part of the left valve (i.e., the upper facing valve), facilitating classifying scallop species with high accuracy even when most of the scallop was buried. This overcomes some of the limitations associated with RGB imaging and morphological sorting. Based on aquarium observations, both saucer and mud scallops spend most of their time on the substrate with their

Fig. 4 Scatter plot of the first and third principal components score vectors, t_1 and t_3 , used for identifying and removing the background pixels. The plot was generated by clustering pixels according to their spectral similarities. Clusters were formed in regions where the density of spectrally related pixels in the PCA scatter plot was high. These high-density areas are represented in red, while low-density regions appear in blue



feeding tentacles protruded and extended beyond the periphery of the valve. The tentacles have ability, although limited, to clear sediments from around the scallop as they filter the water for food. As a result, the peripheral edges of the valve tend to be less obscured by sediment and are therefore particularly useful for providing clear spectral images.

Reducing the dimensionality of the model and simplifying the calculations is critical to expedite the analysis for practical application under at-sea conditions. If the technology is applied to detect, distinguish, measure and count very large numbers of scallops distributed over thousands of hectares of the seafloor, the model needs to process massive HSI datasets. To improve the analysis speed, complex data pre-processing algorithms were avoided during the training of the PLS-DA model in this study. In addition, the number of LVs was reduced to the minimum functional value ($LV=7$) for a negligible decrease in the classification accuracy. Another possible limitation, should the technology be applied to assessing scallop populations, is the speed at which the HSI system measures light reflectance for hundreds

of wavelengths. The maximum achievable imaging speed of the camera used herein was 17 mm/sec (0.06 km/hour) which is not practical for a large-scale field operation at sea. Reducing the number of wavelengths used in the PLS-DA model, by recognising and excluding the uninformative wavelengths, allows a faster image acquisition process. This optimisation enables programming the camera to capture only the final selected wavelengths when operated under at-sea conditions, leading to faster camera movement. Furthermore, the resulting images are smaller in size (smaller dataset), reducing the duration of analyses and the required storage capacity.

In this study, we implemented a two-step process to minimise the number of wavelengths required for the PLS-DA model. Initially, we reduced the wavelengths to 47 by using every tenth wavelength during the image import into Evince software. Subsequently, we conducted a more refined analysis to pinpoint and select the important wavelengths from these 47, utilising the β coefficient and VIP (Fig. 6).

The classification accuracy of the PLS-DA model, in the test dataset, decreased to 84% where the model

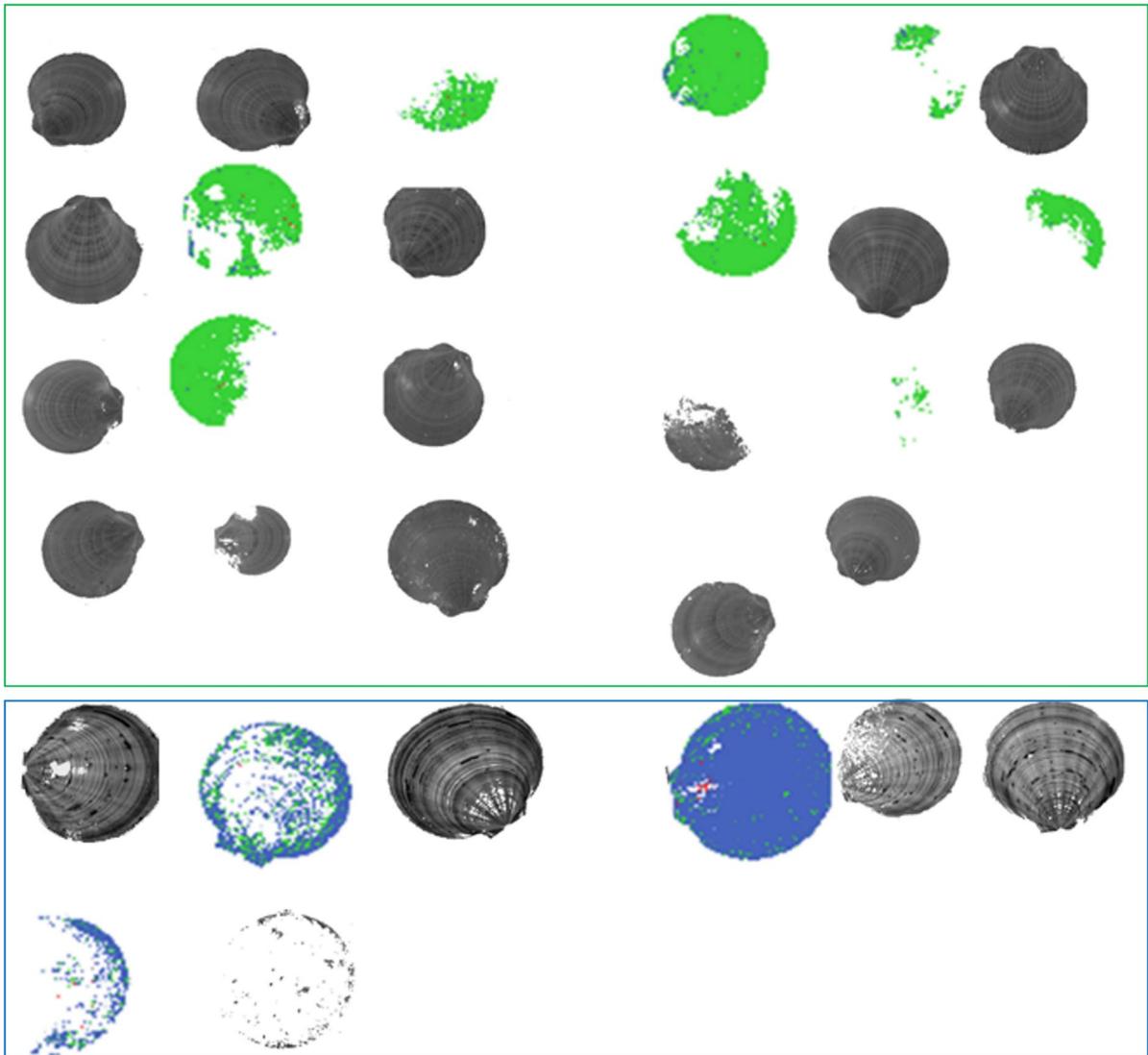


Fig. 5 PLS-DA-classified image of the scallop samples. Blue, green and red represent saucer, mud and unclassified pixels/samples, respectively. The grey scallops are the samples used for calibration and excluded from the test set

was only trained with wavelengths associated with large β -coefficients (β/β -coefficient standard deviation ratio > 1). The classification accuracy reduced to 74% where the PLS-DA model was only trained with wavelengths with a large VIP (VIP > 1). The accuracy of the model further reduced to 66% when the model was trained with wavelengths that had both large β -coefficient and VIP. The β -coefficient was the most efficient wavelength selection method, using only 13 wavelengths (versus 16 VIP wavelengths) of the original 462 wavelengths. We were unable to test the

imaging speed when measuring only these 13 wavelengths, due to equipment constraints. Given that only 13 of the 462 wavelengths were required to achieve 84% classification accuracy, a significant increase in the imaging speed (necessary for field-scale analysis) is expected.

The β -coefficient important wavelengths identified in this study included 403, 416, 482, 535, 602, 615, 629, 669, 682, 695, 709, 722 and 749 nm. Given that the main component of scallop valves is calcium carbonate, CaCO_3 (Martin et al. 2021;

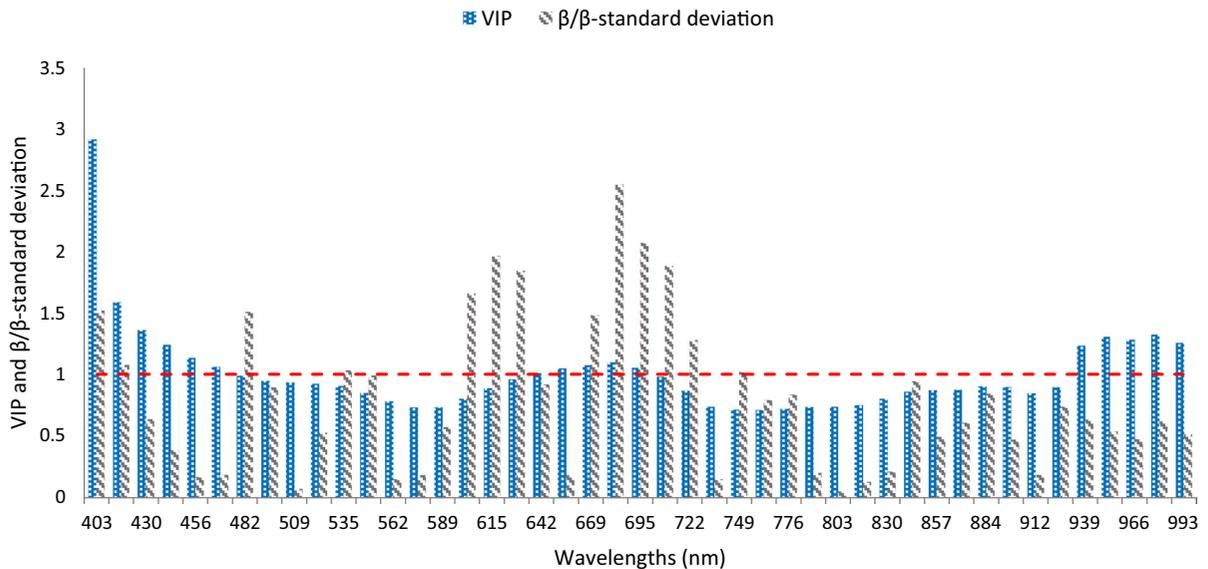


Fig. 6 The important wavelengths selected for the identification and distinguishing saucer scallops from mud scallops using PLS-DA models developed herein. The blue dotted bars are VIP values and the grey striped bars are β/β -standard

deviation calculated for each wavelength. Wavelengths that had VIP and β/β -standard deviation values larger than 1 (horizontal red dashed-line) were considered important for species identification

Sawai et al. 2001), it was expected that some of wavelengths were related to CaCO_3 . There is limited information about the wavelengths related to CaCO_3 in scallop valves. However, the 400, 500, 555, 575 and 600 nm wavelengths have been attributed to CaCO_3 in soil (Gomez and Coulouma 2018; Gomez et al. 2008). These are comparable to the 403 and 602 nm wavelengths that were found to be important for species identification in this study.

Identifying the mechanisms involved in differentiating between the scallop species using HSI was beyond the scope of this study. However, we think that differences in the valve chemical composition may not be the only mechanism through which HSI was able to differentiate between species. In another study, HSI distinguished between two similar species of oysters, namely black lip and Sydney rock, aged 19–24 weeks (Tahmasbian et al. 2022). The oyster valves exhibited similar average elemental and morphological compositions. Nevertheless, the number of layers composing the oyster valves varied between species, providing a distinctive characteristic for differentiating between them (Tahmasbian et al. 2022). Therefore, investigating the structure of the saucer and mud scallop valves may provide

additional information about the mechanisms that enabled HSI to differentiate between them.

This study used the spectral signature of the scallop valves, which may result in overestimation of the scallop populations due to the detection and inclusion of dead scallops. Future HSI analyses may be able to discern the condition of a scallop (i.e., live or dead) by examining the spectral signatures of scallop mantle and feeding tentacles. It is important to note that although the detection of feeding tentacle spectra can be indicative of a live state, the absence of such spectra does not necessarily indicate a deceased state. This is because there are instances where scallops withdraw their tentacles when they are disturbed or threatened.

This study offers valuable insights into the potential utilisation of HSI for real-time monitoring of scallop populations, as well as presenting solutions for enhancing imaging and data analysing speed through data size reduction and wavelength selection. However, it is important to acknowledge that operating HSI under at-sea conditions may present additional challenges that require thorough investigation in future research and development projects. These challenges encompass the dynamic nature of the benthos, surface conditions,

turbidity and currents, which introduce instability in the image acquisition process. Other factors, such as suspended sediment, detritus, dissolved organic matters and chlorophyll may interfere with the reflectance measurements, contributing to the complexity of imaging and data analysis. Furthermore, the optical characteristics of wild scallops may be influenced by factors such as life stages, health conditions, ocean acidification, and parasites on the valves. Conducting a pilot study that deploys the technology at sea may uncover additional challenges.

Conclusion

This study demonstrated that hyperspectral imaging has potential for detecting, distinguishing and counting commercially important saucer scallops for stock assessment and resource management. The ability of HSI to detect and identify a scallop when it is largely buried in sediments is advantageous and has potential to complement RGB imaging that relies on morphological properties. Application of this method would result in reduced impacts to the seabed compared to towed trawl nets or dredges that are commonly used to monitor scallop abundance. Moreover, this technology may hold promise for assessing other benthic fished species and monitoring marine ecosystems.

Acknowledgements This study was funded by Department of Agriculture and Fisheries, Queensland Government, Australia (AS11391).

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data availability The data supporting the finding of this study belong to the Department of Agriculture and Fisheries, Queensland Government, Australia and may be available from the corresponding author upon request and with permissions of the Queensland Government.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The

images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ballabio D, Consonni V (2013) Classification tools in chemistry. Part 1: linear models. *PLS-DA Anal Methods* 5:3790–3798
- Brand AR (2016) Chapter 11 - Scallop Ecology: Distributions and Behaviour. In: Shumway SE, Parsons GJ (eds) *Developments in Aquaculture and Fisheries Science*. Elsevier, pp 469–533
- Brereton RG, Lloyd GR (2014) Partial least squares discriminant analysis: taking the magic away. *J Chemom* 28:213–225
- Chen Y, Gao S, Jones EJ, Singh B (2021) Prediction of soil clay content and cation exchange capacity using visible near-infrared spectroscopy, portable X-ray fluorescence, and X-ray diffraction techniques. *Environ Sci Technol* 55:4629–4637
- Chennu A, Färber P, Death G, de Beer D, Fabricius KE (2017) A diver-operated hyperspectral imaging and topographic surveying system for automated mapping of benthic habitats. *Sci Rep* 7:7122
- Chennu A, Färber P, Volkenborn N, Al-Najjar MAA, Janssen F, de Beer D, Polerecky L (2013) Hyperspectral imaging of the microscale distribution and dynamics of microphytobenthos in intertidal sediments. *Limnol Oceanogr Methods* 11:511–528
- Cheung WWL, Watson R, Pauly D (2013) Signature of ocean warming in global fisheries catch. *Nature* 497:365–368
- Chicco D, Jurman G (2020) The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*. <https://doi.org/10.1186/s12864-019-6413-7>
- Chong I-G, Jun C-H (2005) Performance of some variable selection methods when multicollinearity is present. *Chemom Intell Lab Syst* 78:103–112
- Courtney AJ, Daniell J, French S, Leigh GM, Yang W-H, Campbell MJ, McLennan MF, Baker K, Sweetland T, Woof E (2021) Improving mortality rate estimates for management of the Queensland saucer scallop fishery. In: *State of Queensland, Australia*
- Da Conceição RRP, Simeone MLF, Queiroz VAV, De Medeiros EP, De Araújo JB, Coutinho WM, Da Silva DD, De Araújo Miguel R, De Paula Lana UG, De Resende Stoinoff MA (2021) Application of near-infrared hyperspectral (NIR) images combined with multivariate image analysis in the differentiation of two mycotoxicogenic *Fusarium* species associated with maize. *Food Chem* 344:128615
- Delargy AJ, Blackadder L, Bloor I, McMinn C, Rudders DB, Szostek CL, Dobby H, Kangas M, Stewart BD, Williams

- JR, Stokesbury KDE (2023) A global review of catch efficiencies of towed fishing gears targeting scallops. *Rev Fisheries Sci Aquacult* 31:296–319
- Dredge MC, Marsden ID, Williams JR (2016) Scallop fisheries, mariculture, and enhancement in Australasia. *Developments in Aquaculture and Fisheries Science*, Elsevier, pp 1127–1170
- FAO (2020) The state of world fisheries and aquaculture 2020. Sustainability in action. In, Rome
- Farrar MB, Wallace HM, Brooks P, Yule CM, Tahmasbian I, Dunn PK, Hosseini Bai S (2021) A performance evaluation of Vis/NIR hyperspectral imaging to predict curcumin concentration in fresh turmeric rhizomes. *Remote Sensing* 13:1807
- Farrar MB, Wallace HM, Tahmasbian I, Yule CM, Dunn PK, Bai SH (2023) Rapid assessment of soil carbon and nutrients following application of organic amendments. *Catena* 223:106928
- Freese L, Auster PJ, Heifetz J, Wing BL (1999) Effects of trawling on seafloor habitat and associated invertebrate taxa in the Gulf of Alaska. *Mar Ecol Prog Ser* 182:119–126
- Gomez C, Coulouma G (2018) Importance of the spatial extent for using soil properties estimated by laboratory VNIR/SWIR spectroscopy: examples of the clay and calcium carbonate content. *Geoderma* 330:244–253
- Gomez C, Lagacherie P, Coulouma G (2008) Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. *Geoderma* 148:141–148
- Grahn H, Geladi P (2007) Techniques and applications of hyperspectral image analysis. John Wiley & Sons
- Hohrenk-Danzouma LL, Vosough M, Merkus VI, Drees F, Schmidt TC (2022) Non-target analysis and chemometric evaluation of a passive sampler monitoring of small streams. *Environ Sci Technol* 56:5466–5477
- Kaiser MJ, Spencer BE (1996) The effects of beam-trawl disturbance on infaunal communities in different habitats. *J Anim Ecol* 65:348–358
- Lee LC, Liong C-Y, Jemain AA (2018) Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst* 143:3526–3539
- Letnes PA, Hansen IM, Aas LMS, Eide I, Pettersen R, Tassara L, Receveur J, le Floch S, Guyomarch J, Camus L (2019) Underwater hyperspectral classification of deep sea corals exposed to 2-methylnaphthalene. *PLoS ONE* 14:e0209960
- Li B, Liew OW, Asundi AK (2006) Pre-visual detection of iron and phosphorus deficiency by transformed reflectance spectra. *J Photochem Photobiol, B* 85:131–139
- Manley M (2014) Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chem Soc Rev* 43:8200–8214
- Martin KR, Waits LP, Parent CE (2021) Teaching an old shell new tricks: extracting DNA from current, historical, and ancient Mollusk shells. *Bioscience* 71:235–248
- Mehrübeoglu M, Smith DK, Smith SW, Smee DL, Simionescu P-A (2013) Investigating oyster shell thickness and strength using three imaging modalities: hyperspectral imaging, thermal imaging and digital photography. In, *Imaging Spectrometry XVIII, SPIE*, pp 174–193
- Miller TJ, Hart DR, Hopkins K, Vine NH, Taylor R, York AD, Gallager SM (2019) Estimation of the capture efficiency and abundance of Atlantic sea scallops (*Placecten magellanicus*) from paired photographic–dredge tows using hierarchical models. *Can J Fish Aquat Sci* 76:847–855
- Montes-Herrera JC, Cimoli E, Cummings V, Hill N, Lucieer A, Lucieer V (2021) Underwater hyperspectral imaging (UHI): a review of systems and applications for proximal seafloor ecosystem studies. *Remote Sens* 13:3451
- Park B, Lu R (2015) *Hyperspectral imaging technology in food and agriculture*. Springer
- Park B, Shin T-S, Cho J-S, Lim J-H, Park K-J (2021) Characterizing hyperspectral microscope imagery for classification of blueberry firmness with deep learning methods. *Agronomy* 12:85
- Probert PK, Mcknight DG, Grove SL (1997) Benthic invertebrate bycatch from a deep-water trawl fishery, Chatham Rise, New Zealand. *Aquat Conserv Mar Freshwat Ecosyst* 7:27–40
- Rinnan Å, Berg FVD, Engelsen SB (2009) Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal Chem* 28:1201–1222
- Sawai J, Satoh M, Horikawa M, Shiga H, Kojima H (2001) Heated scallop-shell powder slurry treatment of shredded cabbage. *J Food Prot* 64:1579–1583
- Sokolova M, Japkowicz N, Szpakowicz S (2006) Beyond Accuracy, F-Score and ROC: a family of discriminant measures for performance evaluation, Springer Berlin Heidelberg, pp 1015–1021
- Stewart BD, Howarth LM (2016) Quantifying and managing the ecosystem effects of scallop dredge fisheries. *Developments in Aquaculture and Fisheries Science*, Elsevier, pp 585–609
- Tahmasbian I, Morgan NK, Hosseini Bai S, Dunlop MW, Moss AF (2021a) Comparison of hyperspectral imaging and near-infrared spectroscopy to determine nitrogen and carbon concentrations in wheat. *Remote Sens* 13:1128
- Tahmasbian I, Stewart P, Mellor A, Wingfield M (2022) Non-destructive identification of oyster species for Queensland hatcheries. In: *In press: State of Queensland, Australia*
- Tahmasbian I, Wallace HM, Gama T, Hosseini Bai S (2021b) An automated non-destructive prediction of peroxide value and free fatty acid level in mixed nut samples. *LWT* 143:110893
- Tahmasbian I, Xu Z, Boyd S, Zhou J, Esmailani R, Che R, Bai SH (2018) Laboratory-based hyperspectral image analysis for predicting soil carbon, nitrogen and their isotopic compositions. *Geoderma* 330:254–263
- Taylor R, Vine N, York A, Lerner S, Hart D, Howland J, Prasad L, Mayer L, Gallager S (2008) Evolution of a benthic imaging system from a towed camera to an automated habitat characterization system. In, *OCEANS 2008, IEEE*, pp 1–7

- Thrush SF, Dayton PK (2002) Disturbance to marine benthic habitats by trawling and dredging: implications for marine biodiversity. *Annu Rev Ecol Syst* 33:449–473
- Williams P, Geladi P, Fox G, Manley M (2009) Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. *Anal Chim Acta* 653:121–130
- Wold H (1966) Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, pp 391–420
- Wold S (1995). PLS for multivariate linear modeling. In: HVD Waterbeemd (Ed), *Chemometric Methods in Molecular Design*, VCH, pp 195–218
- Wold S, Esbensen K, Geladi P (1987) Principal component analysis. *Chemom Intell Lab Syst* 2:37–52
- Wold S, Ruhe A, Wold H, Dunn WJ (1984) The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J Sci Stat Comput* 5:735–743
- Zang Z, Ji R, Hart DR, Jin D, Chen C, Liu Y, Davis CS (2023) Effects of warming and fishing on Atlantic sea scallop (*Placopecten magellanicus*) size structure in the Mid-Atlantic rotationally closed areas. *ICES J Mar Sci* 80:1351–1366

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.