Italian Journal of Food Safety

https://www.pagepressjournals.org/index.php/ijfs/index

eISSN 2239-7132

Publisher's Disclaimer. E-publishing ahead of print is increasingly important for the rapid dissemination of science. The Early Access service lets users access peer-reviewed articles well before print/regular issue publication, significantly reducing the time it takes for critical findings to reach the research community.
These articles are searchable and citable by their DOI (Digital Object Identifier).

The Italian Journal of Food Safety is, therefore, E-publishing PDF files of an early version of manuscripts that have undergone a regular peer review and have been accepted for publication, but have not been through the copyediting, typesetting, pagination, and proofreading processes, which may lead to differences between this version and the final one.

The final version of the manuscript will then appear in a regular issue of the journal.

The E-publishing of this PDF file has been approved by the authors.

Ital J Food Saf 2026 [Online ahead of print]

Please cite this article as:

Marcolin Roncato M, Currò S, Breda T, et al. **Spectroscopic control of fish products: simultaneous recognition of species and thawed status in large-scale distribution**. *Ital J Food Saf* doi:10.4081/ijfs.2026.14840

Submitted: 14-01-2026
Accepted: 01-04-2026

Spectroscopic control of fish products: simultaneous recognition of species and thawed status in large-scale distribution

Martina Marcolin Roncato,[1*] Sarah Currò,[1*] Tatiana Breda,[1] Aurora Franceschi,[1] Mirko Bernardini,[2] Sara Strazzacappa,[2] Jacopo Ferlito,[3] Matteo Gazziero,[3] Enrico Novelli,[1] Federico Fontana,[1] Laura Da Dalt,[1] Luca Fasolato,[1] Stefania Balzan[1]

[1]Department of Comparative Biomedicine and Food Science, University of Padua, Legnaro; [2]Alì SPA, Padua; [3]ITPhotonics S.r.l., Fara Vicentino, Italy

[*]These authors contributed equally to this work.

Correspondence: Luca Fasolato, Department of Comparative Biomedicine and Food Science, University of Padua, Viale dell'Università 16, 35020 Legnaro (PD), Italy.
Tel.: +39 0498272936. E-mail: luca.fasolato@unipd.it

Key words: traceability, near infrared spectroscopy, seafood, film, classification.

Contributions: all the authors made a substantial intellectual contribution, read and approved the final version of the manuscript, and agreed to be accountable for all aspects of the work.

Conflict of interest: the authors declare no potential conflict of interest.

Ethics approval and consent to participate: not applicable.

Availability of data and materials: data and materials are available from the corresponding author upon request

Conference presentation: this paper was presented at the XXXIV National Conference of the Italian Association of Veterinary Food Hygienists (AIVI), September 10-12, 2025, Parma, Italy.

Abstract
In the context of globalized food production, traceability is a key requirement in the fisheries sector, particularly for prepared and packaged fish products, where processing may hinder visual inspection and increase the risk of mislabeling. Reliable analytical tools are essential, therefore, to discriminate product status (fresh or frozen-thawed) and to ensure accurate species identification, which is essential for food safety and consumer protection. This study investigates near-infrared spectroscopy as a rapid, non-destructive approach to support the traceability of packaged fish products in large-scale retail chains. An innovative in-field methodology was developed, enabling the acquisition of spectroscopic fingerprints directly at the point of sale using a handheld instrument operating in contact with the transparent plastic film. A total of 218 samples were analyzed across 40 different retail points. Spectral data were acquired both with and without packaging to generate two datasets. Chemometric and machine learning approaches were applied to process spectral data and develop classification models. The best models achieved satisfactory classification accuracies ranging from 0.92 to 0.69 and showed strong agreement between predictions obtained from spectra acquired with and without plastic film (Cohen's κ=0.83 for fresh *vs*. frozen-thawed classification and κ=0.77 for species identification). These results demonstrate the ability of portable near infrared spectroscopy to identify both fish species and physical status, supporting self-monitoring and official controls by reducing response times and product handling, with potential integration into blockchain-based traceability systems.

Introduction
The global fishery sector produced 185.4 million tons of product in 2022, and 164.6 million were destined for human consumption (FAO, 2024). In the same year, the Food and Agriculture Organization of the United Nations estimated a consumption of 20.6 kg per capita per year, an increment of 126.37% compared to 1961, with an expected growth to 21.3 kg per capita in 2032 (FAO, 2024). The worldwide production of fishery products is continuously increasing. In the globalization context, the European Regulation (EU) No. 1379/2013 plays an essential role in introducing a clear set of obligatory labelling information throughout the entire fishery supply chain. Accurate information on species identification and storage conditions prevents fraud and promotes traceability, thereby supporting inter-branch organizations activities and the coordinated regulatory framework for seafood distribution within the EU market (Hopkins *et al*., 2024; European Parliament and Council of the European Union, 2013). Avoiding mislabeling is crucial for maintaining consumer trust, valuing honest producers, and preventing health risks (Currò *et al*., 2022). Moreover, specific information improves traceability by enabling clearer identification of batch origin; end-to-end or full-chain batch tracking is increasingly requested not only by seafood producers but also by consumers and regulators. Technological innovations could further enhance traceability and verification across the seafood supply chain (Hopkins *et al*., 2024; Smit *et al*., 2025).
To define the status of fresh or frozen-thawed fish products and identify their species, several techniques are proposed, including the hydroxyacyl-coenzyme A dehydrogenase test, histology, DNA barcoding, and nuclear magnetic resonance (Arias *et al.*, 2022; European Commission, 2022). Such procedures are frequently expensive, destructive, time-consuming and require specialized personnel. Moreover, the rapid and cost-effective simultaneous classification of species and physical status remains challenging. Near infrared spectroscopy (NIRS) is used to obtain a product fingerprint spectrum, capturing absorbance or reflectance values in the 730-2500 nm wavelength range (Smit et al. 2025). The advantages compared to classical techniques include the possibility of performing rapid, eco-friendly, non-destructive analysis without the use of solvents and being directly applicable in the field using portable instruments (Ghidini *et al*., 2019; Arias *et al*., 2022). However, NIR libraries acquisition is an alternative, easy approach for food traceability; the main challenges are acquiring high-quality spectra and handling complex data processing, which requires chemometric tools (Qi *et al*., 2025). Chemometrics plays a key role in spectroscopic data interpretation by applying statistical and mathematical methods, including both supervised and unsupervised techniques, to

transform complex raw signals into meaningful and interpretable information. Chemometric tools are widely used for significant wavelength identification, data visualization through graphical clustering, and the development of classification and prediction models. Beyond facilitating data interpretation, chemometric processing also enables the assessment of spectral library quality, which is crucial for ensuring the robustness of future analyses (Roggo *et al*., 2007; Beć *et al*., 2020). Several studies conducted on whole, prepared, and processed seafood products have successfully identified specific wavelengths capable of discriminating fresh from frozen-thawed fish (Currò *et al*., 2021; Atanassova *et al*., 2024), a relevant aspect for a regulatory perspective, as Implementing Regulation (EU) No. 2022/2503 formally recognizes NIRS as an analytical technique available to competent authorities for thawed fish inspection (European Commission, 2022). However, NIRS technologies have found application in different agri-food sectors (dos Santos *et al*., 2013; Aouadi *et al*., 2020; Fodor *et al*., 2024), such as meat (Arias *et al*., 2022; Fodor *et al*., 2024), vegetables (Chen *et al.*, 2023), forage (Rukundo *et al*., 2020), and dairy (Manuelian *et al*., 2022; Silva *et al.*, 2022).

The study investigates the direct in-field application of NIRS to acquire spectral data at the point of sale by placing the probe in contact with transparent film wrapped fishery products. This trial was conducted in a retail environment across multiple large-scale distribution stores, allowing for minimal product handling, which reduces the risk of contamination and increases the feasibility of routine controls. To the best of our knowledge, no previous studies have evaluated the effect of supermarket packaging on species authentication or on the discrimination between fresh and frozen-thawed seafood. Moreover, the construction of spectral libraries directly within the large-scale retail trade (LSRT) environment provides an opportunity to strengthen traceability systems, addressing industry requirements and increasing demands from consumers and regulatory authorities for improved transparency along the supply chain. In this context, the present study demonstrates the applicability of multi-species and multi-product spectral libraries, including items that have lost anatomical integrity and are marketed as prepared, sliced, or filleted products.

Materials and Methods

The present study considered a sampling period of 8 months from November 2023 to June 2024, conducted across 40 different Nord Est Italian points of sale in the LSRT (Alì S.p.a.; *Supplementary Table 1*). The sampling refers to fish stored at 0-2°C and sold as prepared and wrapped products by the point of sale, using 12 µm PVC-P food-grade film [compliant with UNI EN 13431:2005 (Eu 2005)]. The study considered a total of 218 products of eight different commercial fish species, based on availability in the LSRT, divided into 128 fresh and 90 frozen-thawed samples: *Gadus macrocephalus* (no. 1 fresh; no. 40 frozen-thawed), *Gadus morhua* (no. 6 fresh), *Pollachius virens* (no. 33 fresh), *Lophius piscatorius* (no. 16 fresh; no. 4 frozen-thawed), *Lophius vomerinus* (no. 11 frozen-thawed), *Sepia officinalis* (no. 25 fresh; no. 32 frozen-thawed), *Salmo salar* (no. 47 fresh) and *Oncorhynuchus nerka* (no. 3 frozen-thawed) (Ministero dell'Agricoltura della Sovranità Alimentare e delle Foreste, 2020).

After sampling from the refrigerated counter, each sample was analysed using a handheld NIRS device (PoliSPECNIR, ITPhotonics, Breganze, Italy) operating in reflectance mode between the 900-1680 nm spectral range with a resolution of 2 nm. Spectral data of each sample were acquired a) with film and b) without film. The data were acquired through a 3.2 cm² round window placed in direct contact with the sample surface in continuous mode with an acquisition time of 5 seconds and an integration time of 10 milliseconds. Using poliDATA 3.0.1 software (ITPhotonics, Breganze, Italy), the spectral dataset was converted to absorbance units [as log(1/R)] (Currò *et al*., 2023). This study evaluates the capability of the instrument acquisition range (900-1680 nm) to classify samples through chemometric analysis. Previous studies have shown that it is effective in discriminating between classes, such as fresh vs. frozen–thawed products and seafood origin (Currò *et al*., 2021; Currò *et al*., 2022; Currò *et al*., 2023). Although NIR extends up to ~2500 nm, the instrument range provides adequate information while ensuring stable signals and efficient acquisition (Rukundo *et al*., 2020). Chemometrics analysis of the spectral data, considering the labelling information, was

conducted using R software, version 4.3.1, 2023. This study evaluated both unsupervised and supervised methods, such as principal component analysis (PCA) and support vector machine (SVM), respectively. To investigate the spatial distribution of spectral raw data and the explained variance, the PCA was performed for each dataset. PCA allows the exploration of data patterns and possible segregation identification, as well as samples with or without film (Currò *et al*., 2022). SVM models were employed to assess the classification and prediction capability of raw NIRS data using a machine learning approach. Model performances were tested using the hold-out method, which divides the entire dataset into: training set composed of 70% of samples (no. 153) for model development, and the test set, which is composed of 30% of samples (no. 65). To maintain the class proportion was used the "createDataPartition" function in the "caret" R package. The training set was processed using a 10-fold cross-validation (repeated three times) and tested with five SVM models: Linear, Linear Grid, Radial, Radial cost, and Radial kernel (Lantz, 2020). Model performances were assessed through confusion matrices and several evaluation metrics, including accuracy, sensitivity, specificity, and precision (Warner *et al*., 2016; Meyer, 2025). To verify the agreement between the classifications with and without films, the Cohen k test was applied to the validation data results of the best SVM models for the acquisitions with film, using GraphPad software. The results were classified based on the agreement scale (Landis and Koch, 1977; Dotmatics, 2025).

To select the most informative wavelengths for each raw dataset (with and without film), the random forest (RF) and the Boruta algorithm were applied as supervised feature selection methods for i) species and ii) physical status classification (European Network of Forensic Science Institutes, Drugs Working Group, 2020; Currò *et al*., 2021; Currò *et al*., 2022). The RF is based on building decision trees with the Bootstrap Sampling (Bagging) approach to ranking the most relevant spectral wavelengths (Breiman *et al*. 2022). The Boruta algorithm selects the most important variables for significant influence, based on the repetition of iterations to explain all the variables in a reaching run number (Currò *et al*., 2021; Currò *et al*., 2022; Kursa, 2025).

Summing up, this study applied the reported complement chemometrics approach to collect different types of information, based on spectral data. The PCA was used as an exploratory tool to visualise potential data clustering. Boruta was applied to identify the most informative variables, and the SVM classification algorithms were used to build predictive models based on a machine learning approach, suitable for the species attribution and the fresh or thawed status classification.

Results

To evaluate the effect of the plastic film on fresh and frozen-thawed products, the average spectra of all samples were reported. Figure 1 proposes two graphical reconstructions of the fish NIRS signal, while *Supplementary Figure 1* reports the same results for species classification. These results highlight differences between spectra acquired with and without film. PCA was applied to explore sample distribution in both datasets, explaining approximately 99% of the cumulative variance by the first three principal components. The 3D species distribution reported in Figures 2a and 2b, created with the R software, shows partial centroid separation with considerable overlap, whereas clearer clustering is observed for fresh versus frozen-thawed products reported in Figures 2c and 2d, both with and without film. SVM classification was applied using the five models, and the accuracy (Ac) and significance (p-values) are reported in Table 1, divided for species identification and physical status discrimination, each evaluated with and without film. The Linear model also provides satisfactory results. For species identification, the best performances were obtained with the Linear (Ac=0.79 with and 0.85 without film) and Linear Grid models (Ac=0.77 with and 0.85 without film), both showing high significance. The Radial Kernel model showed the best performance for fresh versus frozen–thawed discrimination, achieving accuracies of 0.85 with film and 0.75 without film, both with high statistical significance. Based on the confusion matrices (*Supplementary Tables 2a and 2b*), sensitivity (Se), specificity (Sp), and other performance metrics were calculated for each SVM model. For species identification (*Supplementary Table 3a*), samples analyzed with film showed the highest predictive performance using the Linear SVM model for *Gadus macrocephalus*

(Se=0.92; Sp=0.96), *Lophius piscatorius* (Se=0.83; Sp=1.00) and *Salmo salar* (Se=1.00; Sp=0.98), and the Linear Grid model for *Pollachius virens* (Se=0.88; Sp=0.98) and *Sepia officinalis* (Se=0.88; Sp=0.97). The Radial Kernel model showed the best performance for fresh versus frozen–thawed classification (*Supplementary Table 3b*), achieving Se=0.84 and Sp=0.85. Classification of samples as fresh or frozen–thawed showed high agreement between models developed with and without film, based on Cohen κ (κ=0.83; SE=0.05; 95% CI: 0.74-0.92), indicating almost perfect agreement according to the Landis and Koch scale. For species classification, substantial agreement was observed, with 84.5% observed agreement compared to 32.0% expected by chance (κ=0.77; SE=0.04; 95% CI: 0.69-0.85) (Landis and Koch, 1977).

In the meantime, RF and Boruta algorithms were applied to both, with and without film, to identify the most informative wavelengths for species and physical status discrimination. RF provided an initial variable ranking, while Boruta selected only statistically significant wavelengths without ranking them (Kursa and Rudnicki, 2010; Degenhardt *et al*., 2019; Kursa, 2025). The Boruta feature selection was reported in Table 2, to compare each class results.

Discussion

European legislation defines fresh seafoods as a product that has not undergone any treatment, but can be prepared and packaged under vacuum or in a modified atmosphere, and chilled (European Parliament and Council of the European Union, 2004). Freezing is a commercial practice that permits extending the shelf life of fishery products, but the thawed status information must be indicated on the label (European Parliament and Council of the European Union, 2004, 2013; European Commission, 2012) for consumer safety. Identifying previously frozen products is challenging; however, freezing and thawing induce physical and biochemical changes, such as lipid oxidation and protein denaturation, that can be detected (Strateva and Penchev, 2020; Atanassova *et al*., 2024). Furthermore, according to Council Regulation (EC) 1224/2009 Article 58 (Council of the European Union, 2009) and Regulation (EU) 1379/2013 (European Parliament and Council of the European Union, 2013), the traceability of each batch of non-prepacked fishery and aquaculture products must be ensured throughout the entire supply chain. Mandatory information includes commercial and scientific names of the species; production method; catch or farming area and the fishing gear category; whether the product has been defrosted; and the date of minimum durability, when applicable. The LSRT ensures the traceability data are transmitted to consumers with regular internal audits before and after product wrapping at the point of sale. Due to the complexity of the globalized seafood supply chain, additional methods are required to support food business operators in verifying suppliers and traceability information, and to enable the competent authority to perform checking activities. To test the NIRS classification performance and labelling information attribution, the use of spectral libraries with high product variability is essential. For frozen-thawed identification, NIRS and chemometrics analysis are defined as suitable applications by Implementing Regulation (EU) No. 2022/2503 (European Commission, 2022). Moreover, different studies have applied the NIRS technology for species identification in the fishing sector (Currò *et al*., 2021; Atanassova *et al*., 2024). This paper reports the feasibility of using spectral data acquired directly at the point of sale, with or without protective films.

The graphical representation of absorbance associated with wavelengths, for both species and fresh and frozen-thawed and reported in Figure 1 and *Supplementary Figure 1*, shows a differentiation between the spectral data acquired with and without film starting in the 1350-1390 nm range, suggesting possible interference effects from the film. This is in line with the absorption of food plastic polymers in the third overtone regions around 1100-1300 and 1380-1430 nm (Rukundo *et al.*, 2020; Arias *et al*., 2022). The significant wavelength determination is relevant to provide a suitable interpretation of NIR-light interaction with the chemical vibration bonds (Kursa and Rudnicki, 2010; Degenhardt *et al.*, 2019). As reported in Table 2, the Boruta algorithm identifies 89 significant wavelengths for species identification with film, and 130 without; these differences could be caused by a film matting effect. To classify fresh or frozen-thawed products, 31 robust wavelengths with

film and 64 wavelengths without film were selected. In all cases, the region 908-934 nm corresponds to the third overtones of C-H stretching modes was selected (Sannia *et al.*, 2019). Moreover, certain wavelength selections were associated with specific classifications, such as fresh and freeze-thawed status. The 902-1030 nm range, in general, indicates proteolysis and lipid modification caused by fish thawing, resulting in a shift of the water band to the 990-1060 nm range (Sannia *et al*., 2019; Currò *et al*., 2022; Giró-Candanedo *et al*., 2024;). The 1150-1200 nm range reported especially for interspecies differentiation can represent the second overtone of C-H bonds, but in meat, the range 1200-1450 nm can be associated with C-H, N-H and O-H bonds of proteins and water (Gonçalves *et al*., 2021; Arias *et al*., 2022; León *et al*., 2023). This information suggests that wavelengths related to proteins, water and interspecies differentiation may be covered by the PVC-P film absorbance. However, Boruta algorithms applied to both datasets, with and without film, select significant wavelengths within the film absorption range. This supports the hypothesis that transparent packaging interference has a negligible effect on wavelength selection and could mitigate the water interferences due to its high content in the fish tissue (75-80%) (Aouadi *et al*., 2020; Zhang *et al*., 2023), especially for freshness or thawed-state detection.

The 3D PCA explained a high level of variability, but as represented in Figure 2, it was not reflected in clear clustering, especially for species representation, which showed insufficient separation, except for *S. officinalis*, a mollusk that differs in muscle structure and composition. This pattern may reflect the presence of closely related commercial products, such as *L. piscatorius* and *L. vomerinus*, both ascribed to monkfish, *G. macrocefalus* and *morhua*, as cod, or *S. salar* and *O. nerka*, two salmon species. Similarly, data distribution was unclear for fresh or frozen-thawed discrimination, reflecting the certain classes availability, such as frozen-thawed salmon (no. 3), fresh cod (no. 6) and the absence of frozen-thawed pollock. Nevertheless, the study shows data separation into two areas. The datasets included a wide range of fishery products, demonstrating that a multi-species model can provide useful clustering. Overlaps in data distribution may be attributed to the high collinearity of certain variables shared across classes, such as the fishing area (Currò *et al*., 2021).

The SVM classification models reported high performances in both case studies, fresh versus frozen-thawed and species classification, with strong statistical significance for the products tested with the film. In particular, the most promising species classification analysis was performed by Linear and Linear grid methods. The models registered the best accuracies (Ac=0.87) with significant results ($p \leq 0.001$). Furthermore, the most fitting method for the state fresh or frozen and thawed analysis is the Radial kernel model (Ac=0.85 and $p<0.001$). Both results are close to those of models without the film, which achieved species identification accuracy of 0.89 and state determination accuracy of 0.75, respectively. Acquisition through film, as an innovative method, shows promising classification performances for other food matrices and applications (Arias *et al*., 2022; León *et al*., 2023). Accuracy data reflect model classification capability and class discrimination robustness (Arias *et al*., 2022). Results show limited differences between the data acquired with and without the film, suggesting a minimal packaging influence on NIRS data classification. Radial methods performed better on the status determination dataset acquired with film than the dataset without the film. Observations suggest a transparent packaging opacifying effect, which can reduce scattered light, reflectance noise and sample water interface (Aouadi *et al*., 2020; Zhang *et al*., 2023). However, as reported in the literature (Currò *et al*., 2021), the restricted variability in some fish categories, due to a small sample size, may affect performance, as shown in the confusion matrices in the *Supplementary Table 2a*. The limited variability of some species and the status confounding effects could constrain the predictive capability of the SVM models, particularly for *O. nerka* (n. 3 samples) and *G. morhua* (n. 6 samples). Notably, some *G. morhua* and *L. piscatorius* specimens were misclassified as *S. officinalis*. *Supplementary Table 2b* reports the prediction results for fresh versus thawed products. The fresh product determination is comparable, with a mislabeling of 5 and 6 products, respectively, analysed with and without film. In the meantime, the mislabeling of thawed products increases (+13%) without film. The results should be interpreted in light of the limitations due to low variability and small sample size of certain product classes. Nevertheless, the present study is based on SVM

analyses, a robust and powerful machine learning approach that performs well even with unbalanced sample sets, as the algorithm is able to account for majority classes while still capturing meaningful patterns in minority classes (Farquad and Bose, 2012; Currò *et al*., 2022).
Compared to other similar studies, based on the identification of frozen-thawed fish (Sannia *et al*., 2019; Currò *et al*., 2022; Atanassova *et al.*, 2024) and species determination (Cavallini *et al*., 2022; Currò *et al*., 2023), these results report a lower accuracy due to species variability. However, the small differences observed between the spectral profiles with and without the film, along with the better results obtained with some models, are highly promising and suggest a possible positive film effect on the analysis. PVC film is a common packaging material used in the food industry; in this case, the fishery products were wrapped in 12 μm PVC film. In the same LSRT, a film with a different thickness was tested for dairy products with promising results, and other studies testify the NIRS capability on different packaging and conditions, such as polypropylene (Rukundo *et al*., 2020; Zhang *et al*., 2023), polyethylene (Chen *et al*., 2023; Zhang *et al*., 2023) and modified atmosphere packaging (Arias *et al*., 2022; León *et al*., 2023). For some of these analyses, the NIR spectra were affected by film interference, but the literature highlights the film matting effect and the potential application in different food supply chains. Based on the promising data acquired, the evidence suggests the possibility of generalizing the effect to other film types.
In conclusion, prospects could include handheld NIR applications in key control points to introduce a transverse digitalization of traceability and quality/authentication checking throughout the entire product line, such as the implementation of a blockchain in real production and distribution environments (Chen *et al.*, 2023; Freund, 2023; Lukacs *et al*., 2025). As literature reports, laboratory instruments have superior performance, but handheld field-application could be a more flexible method (dos Santos *et al*., 2013). These instruments have been applied in different aim investigation, such as authenticating food products, shelf life monitoring, and differentiating geographical origin (Freund, 2023; Lukacs *et al.*, 2025; dos Santos *et al*., 2013). Handheld NIR spectroscopy, applied as a low-cost, routine analysis and accessible quality control instrument, has acquired industrial value, especially for the possibility to apply this technology in different food supply chain integrated with real-time monitoring systems, such as the Internet of Things (dos Santos *et al*., 2013; Freund, 2023; Lukacs *et al*., 2025). The integration of blockchain with NIR spectral fingerprinting allows the storage of immutable spectral data for each batch, enhancing traceability, authenticity verification, and transparency across the food supply chain, particularly in chain-of-custody systems and certification schemes based on origin, production practices, and sustainability criteria. This approach could significantly reduce vulnerability to fraud in highly fraud-prone products such as fish (Fox *et al.*, 2018), by ensuring traceability and authentication through immutable spectral data.

Conclusions

These findings confirm the capability of NIRS to classify fish products according to species and fresh or frozen–thawed status, addressing the growing need for traceability and labelling control in an increasingly globalized fisheries sector. The proposed approach, based on spectral acquisition directly through the plastic film in an LSRT, reflects a realistic scenario for food inspection and self-monitoring activities. This rapid and eco-friendly methodology allows the simultaneous evaluation of multiple parameters, including species identification and conservation state. The limited differences observed between acquisitions performed with and without film, together with the improved performance of some models, suggest a potentially positive effect of packaging on the analysis. Although increased dataset variability, balanced class representation, and the inclusion of different commercial films would further strengthen discrimination, the high agreement observed between classifications obtained with and without film indicates comparable and reliable performance. Despite possible wavelength absorption effects induced by packaging, the use of film-specific spectral libraries combined with appropriate models provides consistent predictions. Overall, this technique contributes to reducing analysis time and product handling during LSRT inspections

and represents a step forward in simplifying traceability operations and enhancing the valorization of fish products, with future potential integration into blockchain-based traceability systems.

Online supplementary material:
Supplementary Figure 1. Graphical reconstruction of wavelength for the absorbance into the NIRS range of the eight species tested, both with and without film.
Supplementary Table 1. Samples collected description: samples were collected based on availability from large-scale retail outlets. For each specimen, the scientific and commercial name, and the number of fresh and frozen and thawed product analyses were reported (Ministero dell'agricoltura, 2020).
Supplementary Table 2a. Support vector machine confusion matrix results with and without film investigating species using the hold-out (70-30%) and an internal cross-validation.
Supplementary Table 2b. Support vector machine confusion matrix results with and without film investigating status using the hold-out (70-30%) and an internal cross-validation.
Supplementary Table 3a. Support vector machine results for both datasets, with and without film, representing species identification tested using the hold-out (70-30%) and an internal cross-validation.
Supplementary Table 3b. Support vector machine results for both datasets, with and without film, representing status identification tested using the hold-out (70-30%) and an internal cross-validation.

References

Aouadi B, Zinia Zaukuu JL, Vitális F, Bodor Z, Fehér O, Gillay Z, Bazar G, Kovacs Z, 2020. Historical evolution and food control achievements of near infrared spectroscopy, electronic nose, and electronic tongue—critical overview. Sensors 20:5479.

Arias E, Sierra V, Prado N, González P, Fiorentini G, Díaz J, Oliván M, 2022. Development of a portable near-infrared spectroscopy tool for detecting freshness of commercial packaged pork. Foods 11:3808.

Atanassova S, Yorgov D, Stratev D, Veleva P, Stoyanchev T, 2024. Near-infrared spectroscopy for rapid differentiation of fresh and frozen–thawed common carp (Cyprinus Carpio). Sensors 24:3620.

Beć KB, Grabska J, Huck CW, 2020. Near-infrared spectroscopy in bio-applications. Molecules 2020;25:2948.

Breiman L, Cutler A, Liaw A, Wiener M, 2025. Package Random Forest, CRAN. Available from: https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Cavallini N, Pennisi F, Giraudo A, Pezzolato M, Esposito G, Gavoci G, Magnani L, Pianezzola A, Geobaldo F, Savorani F, Bozzetta E, 2022. Chemometric differentiation of sole and plaice fish fillets using three near-infrared instruments. Foods 11:1643.

Chen Y, Li Y, Williams RA, Zhang Z, Peng R, Liu X, Xing T, 2023. Modeling of soluble solid content of PE-packaged blueberries based on near-infrared spectroscopy with back propagation neural network and partial least squares (BP–PLS) algorithm. J Food Sci 88:4602-19.

Council of the European Union, 2009. Council Regulation (EC) No 1224/2009 of 20 November 2009 establishing a Community control system for ensuring compliance with the rules of the common fisheries policy, amending Regulations (EC) No 847/96, (EC) No 2371/2002, (EC) No 811/2004, (EC) No 768/2005, (EC) No 2115/2005, (EC) No 2166/2005, (EC) No 388/2006, (EC) No 509/2007, (EC) No 676/2007, (EC) No 1098/2007, (EC) No 1300/2008, (EC) No 1342/2008 and repealing Regulations (EEC) No 2847/93, (EC) No 1627/94 and (EC) No 1966/2006. In: Official Journal, L 343/1, 22/12/2009.

Currò S, Balzan S, Novelli E, Fasolato L, 2023. Cuttlefish species authentication: advancing label control through near-infrared spectroscopy as rapid, eco-friendly, and robust approach. Foods 12:2973.

Currò S, Balzan S, Serva L, Boffo L, Ferlito JC, Novelli E, Fasolato L, 2021. Fast and green method to control frauds of geographical origin in traded cuttlefish using a portable infrared reflective instrument. Foods 10:1678.

Currò S, Fasolato L, Serva L, Boffo L, Ferlito JC, Novelli E, Balzan S, 2022. Use of a portable near-infrared tool for rapid on-site inspection of freezing and hydrogen peroxide treatment of cuttlefish (Sepia Officinalis). Food Control 132:108524.

Degenhardt F, Seifert S, Szymczak S, 2019. Evaluation of variable selection methods for random forests and omics data sets. Brief Bioinform 20:492-503.

dos Santos CAT, Lopo M, Páscoa RNMJ, Lopes JA, 2013. A review on the applications of portable near-infrared spectrometers in the agro-food industry. Appl Spectroscopy 67:1215-33.

Dotmatics, 2025. GraphPad. Available from: https://www.graphpad.com/quickcalcs/kappa1/.

EU, 2005. UNI EN 13431:2005 - imballaggi - requisiti per imballaggi recuperabili sotto forma di recupero energetico compresa la specifica del potere calorico inferiore minimo.

European Commission, 2012. Commission Regulation (EU) No 16/2012 amending Annex II to Regulation (EC) No 853/2004 of the European Parliament and of the Council as Regards the requirements concerning frozen food of animal origin intended for human consumption. In: Official Journal, L 8/29, 12/01/2012.

European Commission, 2022. Commission Implementing Regulation (EU) 2022/2503 Amending and Correcting Implementing Regulation (EU) 2019/627 as regards practical arrangements for the performance of official controls in live bivalve molluscs, fishery products, or related to UV-radiat. Official Journal of the European Union L 325/58.

European Network of Forensic Science Institutes Drugs Working Group, 2020. Guideline for the use of chemometrics in forensic chemistry. Available from: https://enfsi.eu/wp-content/uploads/2021/09/Guideline-for-the-use-of-Chemometrics-in-Forensic-Chemistry.pdf.

European Parliament, Council of the European Union, 2013. Regulation (EU) No 1379/2013 on the common organisation of the markets in fishery and aquaculture products, amending Council Regulations (EC) No 1184/2006 and (EC) No 1224/2009 and repealing Council Regulation (EC) No 104/2000. In: Official Journal, L 354/1, 28/12/2013.

European Parliament, Council of the European Union, 2004. Regulation (EC) No 853/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific hygiene rules for food of animal origin. In: Official Journal, L 139/55, 30/04/2004.

FAO, 2024. The state of world fisheries and aquaculture. Available from: https://openknowledge.fao.org/server/api/core/bitstreams/a10e81b3-3fbd-4393-b7b6-6a926915a19a/content.

Farquad MAH, Bose I, 2012. Preprocessing unbalanced data using support vector machine. Decis Support Syst 53:226-33.

Fodor M, Matkovits A, Benes EL, Jókai Z, 2024. The role of near-infrared spectroscopy in food quality assurance: a review of the past two decades. Foods 13:3501.

Fox M, Mitchell M, Dean M, Elliott C, Campbell K, 2018. The seafood supply chain from a fraudulent perspective. Food Security 10:939-63.

Freund A, 2023. Effects of digitalisation on food safety. Stud Agric Econ 125:154-68.

Ghidini S, Varrà MO, Zanardi E, 2019. Approaching authenticity issues in fish and seafood products by qualitative spectroscopy and chemometrics. Molecules 2019;24:1812.

Giró-Candanedo M, Cruz J, Comaposada J, Barnés-Calle C, Gou P, Fulladosa E, 2024. Differentiation between fresh and frozen-thawed mackerel fish using low-cost portable near infrared spectrometry devices. J Food Engin 379:112131.

Gonçalves DB, Santos CSP, Pinho T, Queirós R, Vaz PD, Bloore M, Satta P, Kovács Z, Casal S, Hoffmann I, 2021. Near infrared reflectance spectroscopy coupled to chemometrics as a cost-effective, rapid, and non-destructive tool for fish fraud control: monitoring source, condition, and nutritional value of five common whitefish species. J AOAC 104:53-60.

Hopkins, CR, Roberts SI, Caveen AJ, Graham C, Burns NM, 2024. Improved traceability in seafood

supply chains is achievable by minimising vulnerable nodes in processing and distribution networks. Mar Policy 159;104910.
Kursa MB, 2025. Package Boruta : wrapper algorithm for all relevant feature selection. CRAN.
Kursa MB, Rudnicki WR, 2010. Feature Selection with the Boruta Package. J Stat Softw 36:1-13.
Landis JR, Koch GG, 1977. The Measurement of observer agreement for categorical data published. Biometrics 33:159-74.
Lantz B, 2020. Machine learning con R - conoscere le tecniche per costruire modelli predittivi. Apogeo, Milano, Italy. [Book in Italian].
León L, Ortiz A, Ezquerro S, Tejerina D, 2023. NIRS (near infrared spectroscopy) classification of sliced duroc dry-cured ham under various packaging systems and storage temperature and time. Meat Sci 206:109348.
Lukacs M, Toth F, Horvath R, Solymos G, Alpár B, Varga P, Kertesz I, Gillay Z, Baranyai L, Felfoldi J, Nguyen QD, Zoltan Kovacs Z, Laszlo Friedrich L, 2025. Advanced digital solutions for food traceability : enhancing origin , quality , and safety through NIRS , RFID , blockchain , and IoT. J Sens Actuator Netw 2025;14:21.
Manuelian CL, Ghetti M, De Lorenzi C, Pozza M, Franzoi M, 2022. Feasibility of Pocket-sized near-infrared spectrometer for the prediction of cheese quality traits. J of Food Compos Anal 105:104245.
Meyer D, 2025. Support vector machines (SVM). CRAN.
Ministero dell'Agricoltura della Sovranità Alimentare e delle Foreste, 2020. D.M 175216 - Modalità di attribuzione delle denominazioni in lingua italiana delle specie ittiche di interesse commerciale.
Qi W, Jiang Q, Erihemu, 2025. Research progress on the application of near - infrared spectroscopy in liquid food quality testing. Beni-Suef University J Basic Appl Sci 14:75.
Roggo Y, Chalus P, Maurer L, Lema-Martinez C, Edmond A, Jent N, 2007. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. J Pharm Biomed Anal 44:683-700.
Rukundo IR, Danao MGC, Mitchell RB, Masterson SD, Wehling RL, Weller CL, 2020. Effect of scanning samples through polypropylene film on predicting nitrogen content of forage using handheld NIR. AIMS Agric Food 5:835-49.
Sannia M, Serva L, Balzan S, Segato S, Novelli E, Fasolato L, 2019. Application of near-infrared spectroscopy for frozen-thawed characterization of cuttlefish (Sepia Officinalis). J Food Sci Technol 56:4437-47.
Silva LKR, Santos LS, Ferrão SPB, 2022. Application of infrared spectroscopic techniques to cheese authentication: a review. Int J Dairy Technol 75:490-512.
Smit I, Vosmann K, Weber L, Truberg B, Muders K, Bülbül MK, Demirel U, Çalışkan ME, Haase NU, 2025. Potential of near-infrared spectroscopy (NIRS) for prediction of acrylamide formation in French fries in the potato breeding process. Food Chem 463:141214.
Strateva M, Penchev G, 2020. Histological, physicochemical and microbiological changes in fresh and frozen/thawed fish. Trakia J Sci 1:69-80.
Warner K, Mustain P, Lowell B, Geren S, Talmage S, 2016. Deceptive dishes : seafood swaps found worldwide. Oceana. Available from: https://usa.oceana.org/wp-content/uploads/sites/4/global_fraud_report_final_low-res.pdf.
Zhang W, Pan L, Lu L, 2023. Prediction of TVB-N content in beef with packaging films using visible-near infrared hyperspectral imaging. Food Control 147:109562.

Table 1. SVM results for both datasets, with and without film, are represented for both analyses: species identification and state fresh or frozen- thawed discrimination. Each method tested used the hold-out (70-30%) and an internal cross-validation.

Method		Specie		Fresh or frozen-thawed	
		With film	*Without film*	*With film*	*Without film*
Linear	***Accuracy***	0.87	0.89	0.79	0.85
	p	****	****	***	****
Linear grid	***Accuracy***	0.87	0.90	0.77	0.85
	p	****	****	**	****
Radial	***Accuracy***	0.68	0.84	0.82	0.77
	p	****	****	****	**
Radial cost	***Accuracy***	0.76	0.82	0.80	0.77
	p	****	****	***	**
Radial kernel	***Accuracy***	0.73	0.87	0.85	0.75
	p	****	****	****	**

****p<0.0001; ***p<0.001; **p<0.01; *p<0.05.

Table 2. Significant wavelength (nm) selected by the Boruta approach to investigate the species, with and without film, and to evaluate the fish state, with and without film. The feature selection was here reported divided for each class investigated.

Method	Species		Fresh or Frozen-thawed	
	With film	*Without film*	*With film*	*Without film*
Boruta	902 - 934; 1090 - 1098; 1164 - 1220; 1372 - 1426; 1570; 1572; 1590; 1606; 1626; 1646; 1674 - 1680	902 - 940; 946; 948; 952; 1162 - 1220; 1302; 1312; 1328 - 1360; 1400 - 1512; 1520	902 - 918; 922; 1132; 1248; 1264; 1266; 1270 - 1274; 1280; 1290; 1296; 1396; 1414; 1420; 1422; 1426; 1428; 1432; 1440; 1442; 1446; 1454	904 -914; 920; 932; 934; 1108;1252 -1286; 1290 -1310; 1314; 1316; 1400 - 1406; 1410 - 1422; 1432 - 1438; 1448; 1452; 1464; 1468; 1470 - 1476

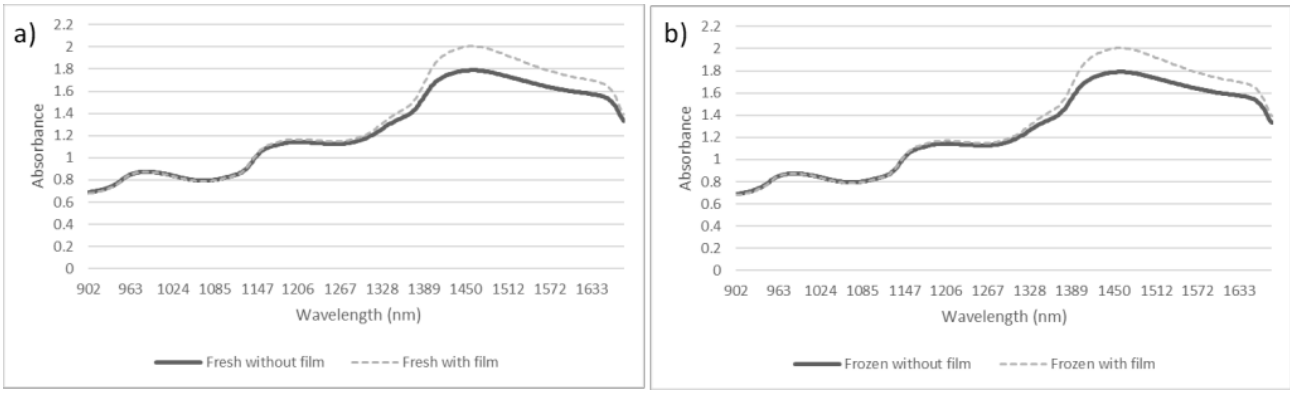

Figure 1. Graphical reconstruction of wavelength for the absorbance into the NIRS range, of fresh or frozen-thawed products tested both with and without film: a) Fresh samples spectra, in dark without film and in grey with film; and b) frozen-thawed samples main absorption, in dark without film and in grey with film.

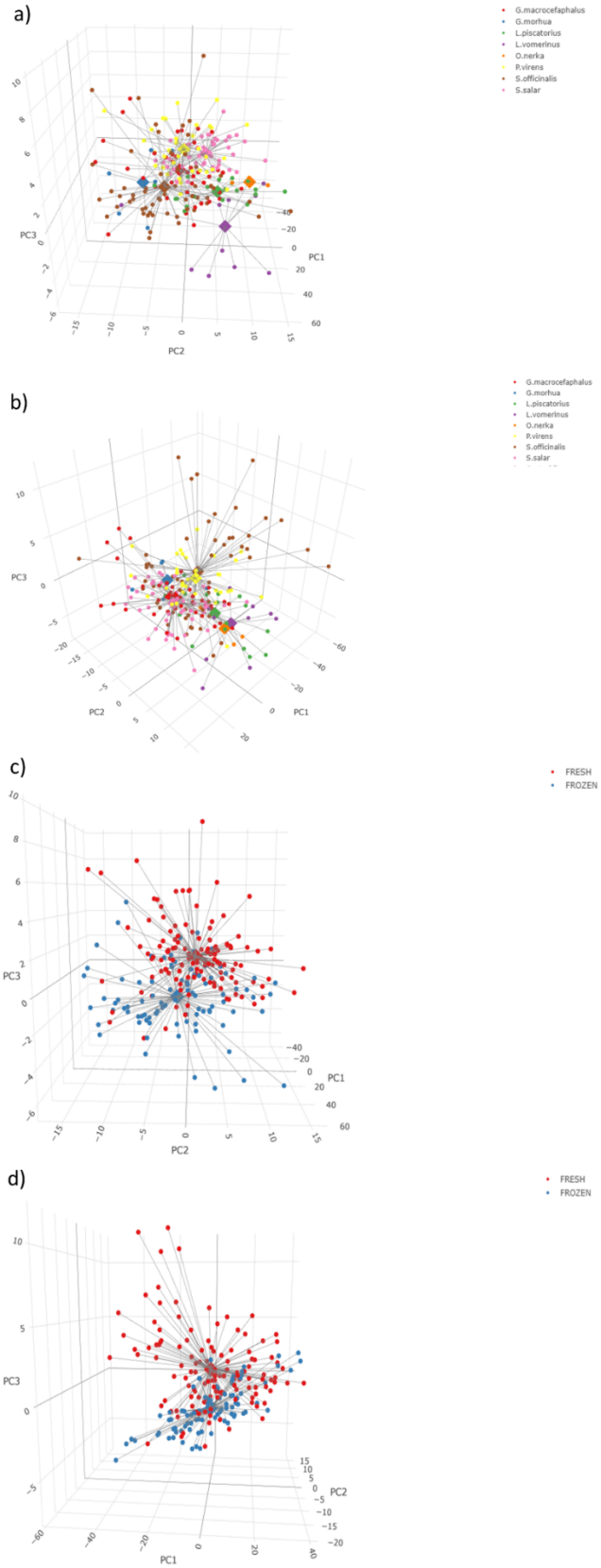

Figure 2. The image reports the graphical representation of the three-dimensional PCA reported as a) species distribution from the dataset acquired without film (86.95% PC1, 10.37% PC2, 1.74% PC3); b) species distribution from the dataset acquired with film (82.78% PC1, 14.10% PC2, 2.31% PC3); c) fresh and frozen-thawed distribution from the dataset acquired without film (86.95% PC1, 10.37% PC2, 1.74% PC3); d) fresh and frozen-thawed distribution from the dataset acquired with film (82.78% PC1, 14.10% PC2, 2.31% PC3).