Current Research in Food Science 9 (2024) 100813

Contents lists available at ScienceDirect

Current Research in Food Science

journal homepage: www.sciencedirect.com/journal/current-research-in-food-science

# Does hyperspectral always matter? A critical assessment of near infrared versus hyperspectral near infrared in the study of heterogeneous samples

Daniele Tanzilli [a,b], Marina Cocchi [a,*], José Manuel Amigo [c,d], Alessandro D'Alessandro [a,e], Lorenzo Strani [a]

[a] Dipartimento di Scienze Chimiche e Geologiche, Università di Modena e Reggio Emilia, Via Campi 103, 41125, Modena, Italy
[b] University of Lille, LASIRE, Cité Scientifique, Villeneuve-d'Ascq, 59650, France
[c] IKERBASQUE, Basque Society for the Promotion of Science, Plaza Euskadi, 5, Bilbao, 48009, Spain
[d] Department of Analytical Chemistry, University of the Basque Country, Barrio Sarriena S/N, Leioa, 48940, Spain
[e] Barilla G. and R. Fratelli, via Mantova 166, 43122, Parma, Italy

ARTICLE INFO

Handling editor: Dr. Maria Corradini

*Keywords:*
NIR
NIR-HSI
Heterogeneity
Classification
Pesto

ABSTRACT

Near Infrared spectroscopy (NIR), in combination with Chemometrics, has been used for many years in diverse scenarios, mostly focused on the classification and quantitation of properties in food, pharmaceutical preparations, artwork material, etc. This success has been possible due to their desirable properties: fast, reliable (under certain conditions), non-destructive, easy to implement from a hardware perspective, and able to create robust and transferable multivariate models.

For some years now, another modality has been gaining the attention of NIR users, especially in the Food sector. That is the plausibility of using NIR in the hyperspectral (HSI) domain. This adds to the previously mentioned abilities, the benefit of scanning the whole surface of samples, acquiring much richer spatial information and, therefore, assuring the quality of the final product more accurately by including parameters that depend on the surface distribution of certain components. This is especially relevant in heterogeneous samples. While this statement is generally true, there are certain situations where this oversampling feature is not strictly needed, and the problem can be easily solved with a classical NIR spectrophotometer. Besides, NIR-hyperspectral imaging (NIR-HSI), despite the abovementioned advantages, has several drawbacks that must be highlighted as well, like their measuring speed, instability, or price.

This manuscript will demonstrate that for certain situations, tuning the focal distance of a NIR spectrophotometer is a more feasible, reliable, and inexpensive strategy to collect all the needed information of samples with a certain degree of heterogeneity.

## 1. Introduction

It is not new to say that Near Infrared Spectroscopy (NIR) and Chemometrics have been part of the Food Industry for many years (dos Santos et al., 2013; Fernández Pierna et al., 2012; Grassi and Alamprese, 2018a; Nobari Moghaddam et al., 2022a; Porep et al., 2015). NIR and Chemometrics are a perfect combination for the determination of mostly macro-constituents in food and food products in practically every stage of the production line (Dixit et al., 2021; França et al., 2021; Grassi et al., 2023; Grassi and Alamprese, 2018a; Måge et al., 2023; Tanzilli et al., 2023; van den Berg et al., 2013). When it comes to solid products, the preferred measurement modality is using diffuse reflectance. This modality, even providing a less detailed spectrum than other modalities (like transmittance), has the property of collecting information on the measured surface (with a certain level of penetration) in questions of milliseconds. This has promoted its implementation in production lines to work in real-time, assessing the quality of the products at different stages of the production lines to determine the level of, for example, protein or carbohydrates in cereals (Kays et al., 2000; Ozaki et al., 2006), cheese (Bittante et al., 2022), vegetables and fruits (Sirisomboon, 2018), or to develop classification models to ascertain the origin or level of adulterations (Nobari Moghaddam et al., 2022b). NIR complies with the requirements of the well-known Quality By Design concept (QbD), demonstrating to be a perfect process analytical technology (PAT)

* Corresponding author.
*E-mail address:* marina.cocchi@unimore.it (M. Cocchi).

https://doi.org/10.1016/j.crfs.2024.100813
Received 18 March 2024; Received in revised form 4 June 2024; Accepted 18 July 2024
Available online 20 July 2024

(Gorla et al., 2023; Grassi et al., 2022; Grassi and Alamprese, 2018b; Pu et al., 2020).

In recent years, a new NIR modality has merged as an alternative to overcome one of the main problems of NIR, which is the need for more spatial information of the sample. NIR-Hyperspectral Imaging (NIR-HSI) cameras together with Chemometrics (Amigo, 2019) combine some of the properties of NIR (fast, non-destructive and reliable) with the major property of being able to obtain a spectral representation of the surface of a sample (Alamar et al., 2023). This property is especially welcome when the sample has a certain level of heterogeneity. Therefore, it is possible to collect that heterogeneity by measuring the whole sample and localising the adequate regions of interest (RoI)(Amigo et al., 2023; Ma et al., 2022; Squeo et al., 2022; Xu et al., 2023) that are more related to the property that wants to be measured.

Despite this increasing interest, several aspects must be considered when comparing a NIR-HSI with a more classical single-spot NIR.

1) NIR-HSI cameras are much more affected by environmental conditions: temperature and moisture are two factors that must be controlled at every moment when dealing with NIR measurements. While these aspects are more controllable using a classical probe, even working with optical fibres, this is not the same with NIR-HSI cameras. They must be normally placed in controlled semi-sealed chambers that, placed on a conveyor belt, will make their operability difficult and more expensive.
2) NIR-HSI cameras are much more sensitive to vibrations: as in the previous case, the working environment largely affects the operability of the sensors. While NIR probes can also be affected by mechanical movement, this effect is much smaller compared to NIR-HSI cameras.
3) Models Calibration transfer: while there are hundreds of successful cases for calibration transfer implementations in NIR (Folch-Fortuny et al., 2017; Fonseca Diaz et al., 2022; Li et al., 2022; Qiao et al., 2023), it needs to be clarified how to perform this with NIR-HSI cameras.
4) Speed: while the NIR probes are extremely fast nowadays, the NIR-HSI still need to reach the same measuring speed with the same spectral resolution. One of the most common actions in NIR-HSI is to reduce the number of measured wavelengths and, thus, increasing the measuring speed. This, obviously, has a special impact in both the quality of the signal (Signal-to-noise-ratio) in the NIR-HSI systems and the complexity of the models to be developed.
5) Heterogeneity of the sample: one of the most common arguments for using a NIR-HSI camera instead of a much simpler NIR probe is the fact that NIR-HSI is able to measure the surface of the sample and, thus, collect a larger heterogeneity. While this is generally true, there are cases where NIR-HSI is not clearly justified as a breakthrough over NIR probes.

While all the points mentioned above must be carefully considered, this manuscript deals with the last one. Food samples are, by nature, heterogeneous at a certain spatial resolution level (da Silva et al., 2018; de Moura França et al., 2017), which is, in many cases, found at a microscopic level. Nevertheless, in order to achieve this level of spatial resolution (Cairós et al., 2009), the measuring speed must be sacrificed. In production lines, the process speed is at a level where the NIR-HSI cameras can only operate measuring pixels with a large spatial resolution, hampering the achievement of the detailed spatial resolution in mixtures that are heterogeneous at a small scale.

Besides, there is a factor with the classical NIR probes that is always forgotten (or merely considered). When measuring in diffuse reflectance and placing the NIR probe at a certain focal distance of the sample, the measured spectrum does not only capture a specific point of the sample, but it also contains information about a larger surface. The extent to which the sample surface is sampled with a single-spot NIR probe might be enough to give an analytical answer that can solve the analytical problem at hand. In other words, it is possible to measure heterogeneous samples as well, up to a certain level, with an NIR probe without underrepresenting the heterogeneity of the sample and avoiding all the issues mentioned above of NIR-HSI cameras.

This manuscript will demonstrate this point by comparing the results obtained in the classification of the origin of heterogeneous samples (at a macroscopic level) by using a NIR-HSI camera and an NIR probe working at the same focal distance. The samples are two different types of *Pesto alla Genovese*. Pesto is a well-known traditional Italian sauce consisting primarily of a heterogeneous mixture of fresh basil, pine nuts or cashew nuts, garlic, Parmesan cheese, and oil. These ingredients are normally mashed, giving a typical granule texture. The strategy followed in this manuscript is based on ascertaining the level of similarity of the signals collected by the NIR-HSI camera and the NIR probe, with the aim of classifying both types of Pesto. The amount of spatial information collected by the NIR probe will be quantified and compared to the one collected by the NIR-HSI camera.

## 2. Materials and methods

### 2.1. Sample Description

This work has considered jars of *Pesto alla Genovese* collected during the 2021 production campaign from May to October at the *Barilla G. e R. Fratelli S.p.A*. facility near Parma, Italy. Twenty-four samples of commercial jars were taken during the production, collecting two jars for each sample to ensure representativeness, resulting in a total of forty-eight samples. The collected Pesto differ in the variety of basil, the primary ingredient. To safeguard industrial confidentiality, the products are denoted as Product 1 and Product 2. The analysis examined thirty-four samples of Product 1 and fourteen samples of Product 2, focusing on the differences between the two pesto sauces. It is also important to highlight that this product has been manufactured and it is marketed as an out-of-fridge stable product.

### 2.2. NIR and NIR-HSI measurements and analysis

The sample inside each jar was gently stirred to homogenize the product as much as possible. Besides, the first layer of the sample in contact with the lid was removed since it has different characteristics from the rest of the product. After that, a portion of the sample was transferred to a Petri glass with a diameter of 32 mm. NIR-HSI images and NIR spectra were then acquired after letting the sample rests for a couple of minutes. To ensure a more comprehensive analysis, replicates were analysed for all samples for three days.

#### 2.2.1. HSI-Image

The NIR-HSI images were acquired using a line mapping system (Headwall Photonics, Inc. Massachusetts. USA, kindly donated by FOSS Analytics A/S, Denmark) working in the wavelength range of 938–1630 nm with a spectral resolution of 4.85 nm, recording a total of 142 wavelength channels (λ) for each spectrum. The line mapping scanning system consisted of 320 sensors measuring a 50 mm field-of-view-line moving the object in one direction, taking steps of 300 μm in each movement. The camera was placed at a right angle with respect to the samples, and the samples were illuminated with diffuse white light at an angle of 45° with respect to the samples. The measurements were made in diffuse reflectance mode. The camera was calibrated using the spectra obtained from a Spectralon plate (white reference) and the spectra obtained from the camera with the lenses covered by an opaque dark fabric (dark reference). The diffuse reflectance was transformed into log(1/R) units following standard procedures (Ozaki et al., 2006).

After the acquisition, different Regions of Interest (ROI), derived from the initial images of Petri dishes containing the pesto sauce, were taken, and the mean spectrum was calculated (Fig. 1).

Twenty-five distinct ROIs were considered for each sample, varying

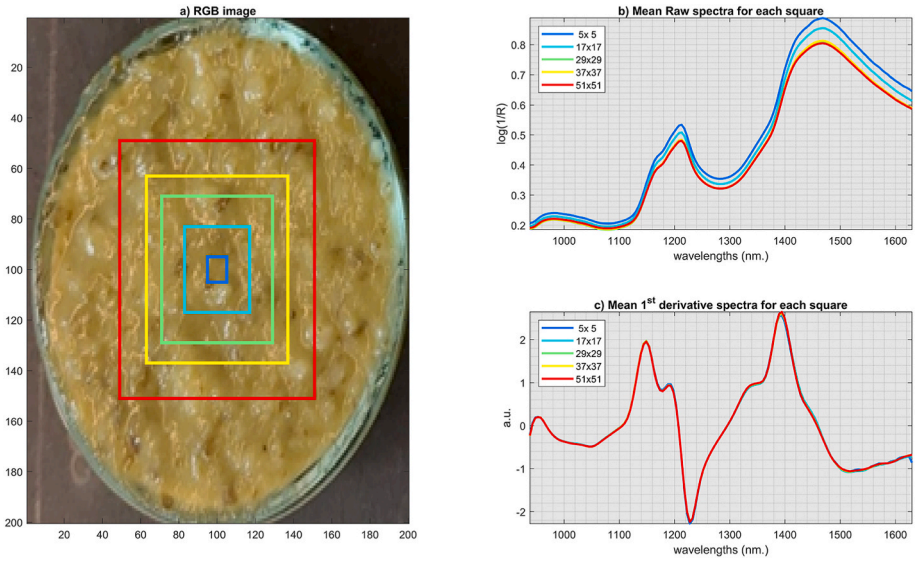



**Fig. 1. a)** Pesto in Petri disk with some of the considered ROI highlighted. **b)** mean spectra of the shown ROI **c)** mean of pre-processed spectra with first derivative for each ROI shown.

the area as shown in Fig. 1a. The ROIs were generated by starting from the central pixel and expanding in squares with dimensions ranging from 3x3 to 51x51 pixels, incrementing by 2 pixels. For each ROI, we calculated the mean of the spectra within the ROI. To explore the differences between ROIs, the augmented dataset $\mathbf{X_{aug}}$ was created by row-wise concatenation of the individual $\mathbf{X_{ROIn}}$ datasets.

$$\mathbf{X_{aug}} = [\, X_{3x3}; X_{5x5} \ldots; X_{51x51}] \quad (1)$$

### 2.2.2. NIR probe

The Vis-NIR spectrometer used to acquire pesto spectra was the ASD FieldSpec4 (Analytical Spectral Devices, Inc.), covering a spectral range from 350 to 2500 nm with a spectral resolution of 1 nm. Prior to acquisition, the spectrometer was calibrated using a Spectralon reflectance standard. The light source was positioned at an incident angle of 45° relative to the sample. Spectra were collected using a fibre optic positioned at 90° relative to the sample. The distance between the fibre optic and the sample was maintained at a constant 20 cm for each acquisition. The choice of the distance has been done based on the following consideration: the longer the focal distance is, the wider the area scanned. Nevertheless, there is a loss of energy that arrives to the detector. Therefore, 20 cm was chosen as an ideal distance for a) capturing the heterogeneity of the current samples, and b) obtaining a good spectrum in the measurement.

## 2.3. Data analysis

To study the similarities between both measuring approaches and the spectral variability found, Principal component analysis (PCA) and k-Nearest Neighbours (kNN) were used for explorative and classification purposes, respectively. All the data analysis has been performed using in-house routines and the PLS_Toolbox (Eigenvector research, WA, USA) working under Matlab v.2023b (The Mathworks, MA, USA).

### 2.3.1. Principal Component Analysis

Principal Component Analysis (PCA) is a method to reduce the dimensionality of the data while retaining the most critical information, achieving this by a projection of the data from the original variables into a new set of orthogonal variables, i.e. Principal Components (PC). PCA allow a better understanding of data structure, simplifying the exploration phase, enhancing visualisation, and aiding pattern recognition.

$$\mathbf{X} = \mathbf{TP}^{\mathrm{T}} + \mathbf{E} \quad (2)$$

Mathematically, as expressed in *equation (2)*, we have a decomposition of the original matrix **X** with *n* row, as the samples, and *m* columns, as the variables, into scores matrix **T**, with *n* rows and *p* columns as the number of PC, and into matrix **P** called loading matrix with *p* rows, as PC number, and *m* columns, as variables number. The scores are the coordinates of the sample in the new system, and they are useful to understand the structure of the data, for instance, recognising a pattern. Meanwhile, the loadings matrix represents the contribution of each variable to each PC. The analysis of the loadings matrix allowed us to understand the correlation structure of the variables. Last, the residual matrix **E**, obtained by the subtraction of recalculated data from the PCA model ($\mathrm{TP}^{\mathrm{T}}$) from X, represents the unmodeled information.

### 2.3.2. K-nearest neighbours

k-Nearest Neighbours (kNN) is a non-parametric discriminant classification algorithm based on the distances between unknown samples and a set of calibration samples. The first step in assigning a label to a new sample is to identify its k-nearest neighbours from the calibration set, determined by their proximity in feature space using, for example, the Euclidean distance. Once the nearest neighbours have been identified, the new sample is assigned to the class that prevails among the majority of its k neighbours in the feature space. The choice of the parameter k is crucial and requires optimisation. Typically, the optimal value for k is determined by cross-validation, where the number of neighbours that results in the lowest classification error is chosen. This ensures robust and accurate classification performance across different data sets. To evaluate the classification performance Sensitivity, Specificity and Efficiency were used as metrics. Sensitivity represents the proportion of samples from the modelled class correctly identified by the model. At the same time, Specificity signifies the proportion of samples not belonging to the modelled class that the model correctly rejects. Lastly, Efficiency combines Specificity and Sensibility through a geometric mean.

## 3. Results and discussion

The raw spectra (Fig. 2a) were pre-processed to remove variability not linked to useful information. In particular, Savitzky–Golay First derivative (polynomial order = 2 and width = 6) was applied. Moreover,

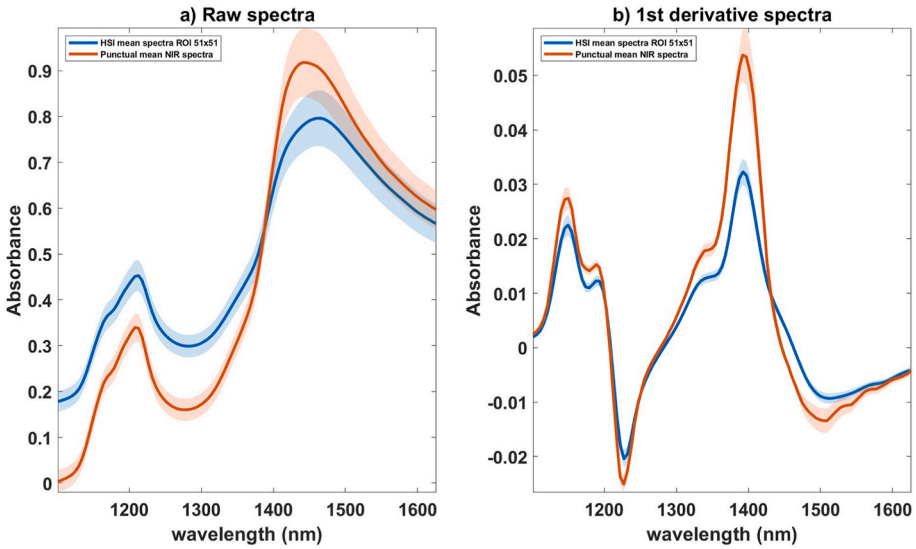



**Fig. 2.** Comparison of mean spectra of ROI 51x51 and Punctual NIR **a)** raw spectra **b)** pre-processed. The shaded areas correspond to the standard deviation of the spectra.

to make a fair comparison between the two techniques, only the common wavelength region, namely 1100–1625 nm, was used.

### 3.1. Measuring the surface area covered by the NIR spot probe

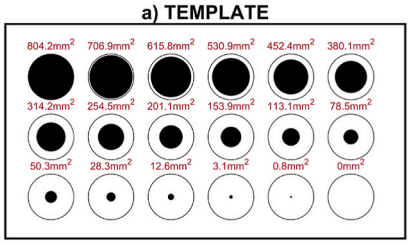
To calculate the exact region measured from the single-spot NIR, a template, shown in Fig. 3a, was created and printed on white cellulose paper. This template consisted of twenty-four paper discs 32 mm in diameter (the same size as the Petri glass). Each disc was characterised by a different size of a concentric black circle, from 32 mm in diameter down to 0.2 mm, with a step of 0.2 mm.

To evaluate the measured area, the reflectance behaviour as a function of the area of the black area in each spot was inspected at a fixed spectral wavelength, corresponding to one of the bands in the spectrum profile of the white paper and to a maximum in the loadings plots (1420 nm). This procedure was repeated 5 times. Fig. 3b shows that there is a sudden change in the reflectance value when the black circle area is lower than 452 $mm^2$ (as highlighted by the red dot in Fig. 3b), with an increasing slope of the reflectance as a function of the disc area. This means that at this point, the sensor starts to measure part of the white area. Thus, the reflectance value increases. Therefore, it can be concluded that for the considered setup the spatial range measured by the single-spot NIR is 452 $mm^2$ at 20 cm focal distance.

### 3.2. PCA on the NIR-HSI data

PCA was conducted on the ROI augmented dataset to explore potential variations across different ROI areas. Fig. 4a reveals a trend in the dispersion of the score values with the ROI area. Specifically, as the ROI area expands, reaching approximately 17.3 $mm^2$ (highlighted by the red circle), there is a discernible decrease in the variability of the scores.

The boxplot shows a consistently stable interquartile range (IQR) beyond the ROI area of approximately 17.3 $mm^2$, highlighted by the red circle. These findings demonstrate a strong association between ROI size and the observed variability, providing valuable insights into the underlying dataset dynamics and implications for subsequent analyses.



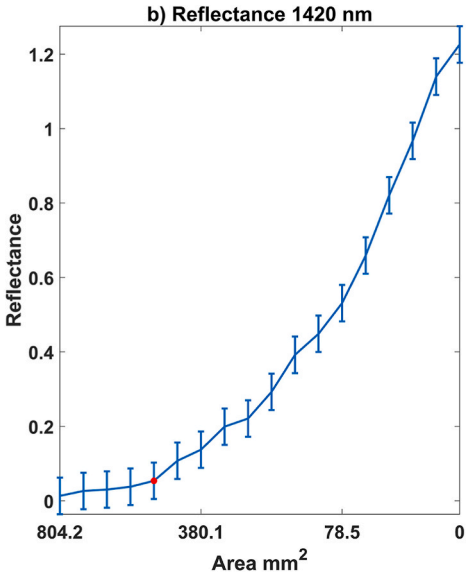




**Fig. 3. a)** Template of 18 disks used to evaluate Punctual NIR measurement area; **b)** Reflectance trend (mean and standard deviation) at 1420 nm as a function of the black area for measurements conducted on the template disks.

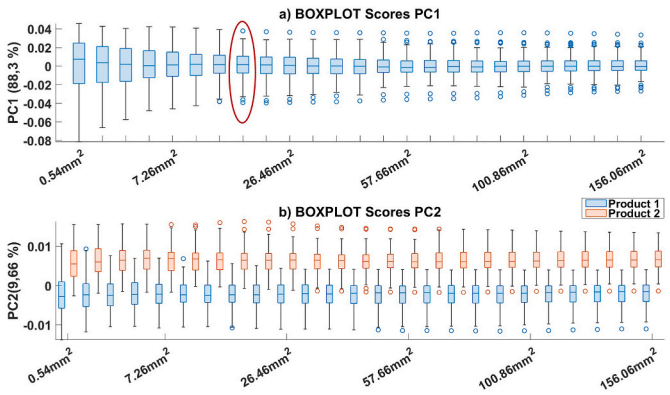



**Fig. 4.** *Boxplot of Scores of Augmented Dataset.* ***a)*** *PC1 scores. It highlights that increasing the ROI area reduces the score values spread.* ***b)*** *PC2 scores of augmented dataset coloured by product type.*

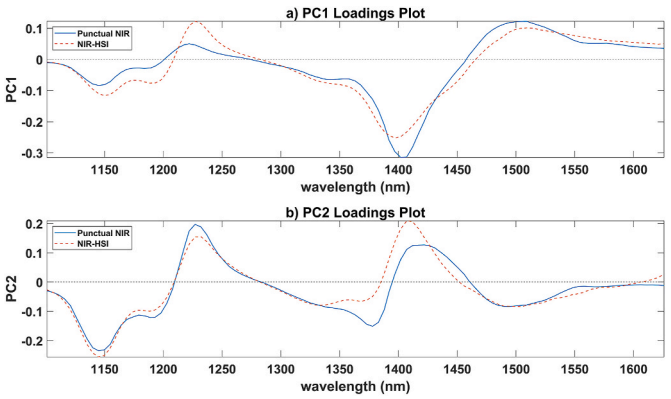



**Fig. 6.** Loadings plot of PCA analysis on Punctual NIR and HSI *Augmented Dataset.* ***a)*** *PC1 and* ***b)*** *PC2.*

Focusing on the boxplot of PC2 (Fig. 4b), which explains 9.66% of the variance, it is possible to observe that with increasing the ROI area the separation between the two different classes of pesto slightly increases, especially promoted by the change in the IQR range of Product 1.

### 3.3. PCA on the NIR probe data

PCA analysis was performed on the single-spot NIR using the setup described in section 2.2.2. The scores plot is shown in Fig. 5a, with scores coloured by the type of Pesto. PC1 did not provide any useful information for the separation between the classes, while PC2 as for the HSI case (Fig. 5b), allowed for the differentiation of the two classes of Pesto. A comparison of the loadings acquired with the two different instrumental setups, as shown in Fig. 6, reveals that they are very similar, exhibiting a comparable loadings profile, especially for PC2 loadings. Thus, both techniques are capable of capturing the same information and highlighting differences between the two types of products.

### 3.4. Classification of pesto types

To assess the classification performance of the k-nearest neighbours (kNN) algorithm, we divided the dataset into calibration and test sets with a ratio of 70/30. The number of neighbours, *k,* was established by Cross-Validation (Venetian blind, 5 splits), maximising Efficiency. In Fig. 7a are reported, for class 1, the values of Efficiency in Cross-Validation vs. the number of neighbours, *k* (varied from 11 to 21 with step of 2). Whereas in Fig. 7b the median value of the Efficiency calculated along the ROI dimensions for each *k* number is shown. From an overall point of view, the Efficiency in classification is maximised with 11 neighbours. Consequently, it was selected as *a k* number to perform the classification.

kNN classification with 11 neighbours was conducted on both types of data, NIR-HSI data, considering each Region of Interest (ROI) and the single-spot NIR data separately.

The analysis of the NIR-HSI dataset, as illustrated in Fig. 8, shows a consistent and notable trend. Sensitivity and Specificity levels, holding steadfast at around 100% and 97%, emphasise the critical role of adequately sized ROIs. This reaffirms earlier observations from the exploratory survey in Section 3.1, suggesting that, for the considered case, an optimal ROI should exceed 17 $mm^2$.

Furthermore, the classification performance obtained by using single NIR data, which for class 1 gives a Sensitivity of 97.5% and a Specificity of 100 %, is perfectly comparable with the results obtained with an ROI area over 17.3 $mm^2$. This means that the setup selected to execute the analysis (described in 2.2.2) is not merely measuring the reflectance from a coherent narrow area, but instead, the diffuse reflectance contains information of a larger area of the sample, as demonstrated in Fig. 3. Hence, the distance between the measuring device (fibre optic in this work) and the sample determines how large this area is.

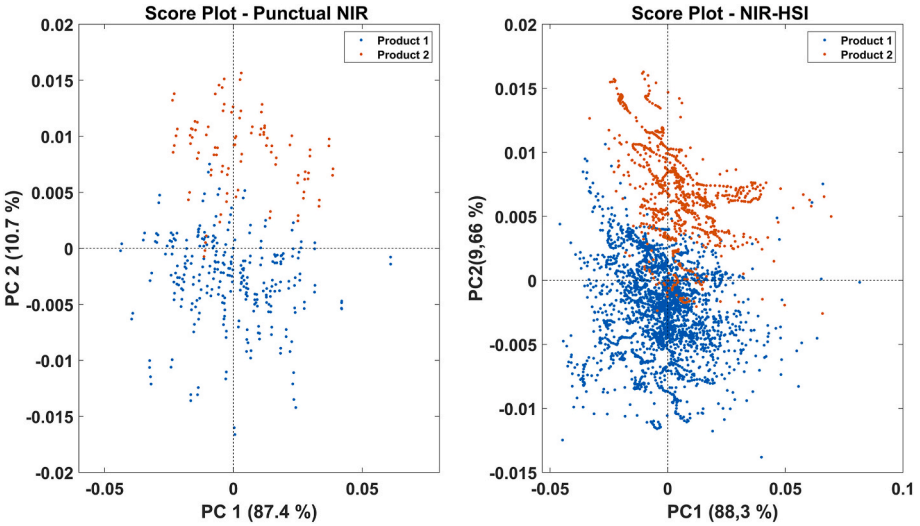



**Fig. 5.** Scores plot of PCA analysis with the scores coloured by the Pesto type. **a)** Punctual NIR **b)** HSI *Augmented Dataset.*

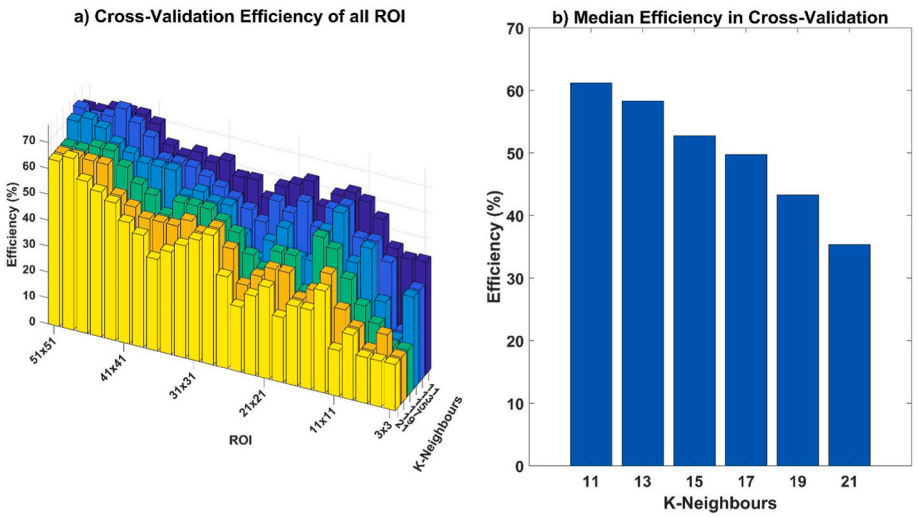



**Fig. 7.** **a)** Efficiency in Cross-Validation as a function of ROI area and the number *k* of neighbours tested; **b)** median efficiency in Cross-Validation calculated along the ROI dimension as a function of *k*.

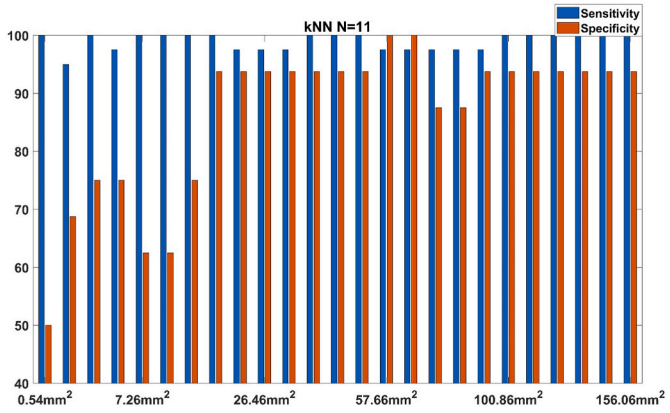



**Fig. 8.** Sensibility and Specificity of kNN-Classification for Product 1.

## 4. Conclusion

This paper aims at presenting a rationale for comparing classical NIR and NIR-HSI measurements in recording spectra from heterogeneous samples, establishing a procedure to assess when the analysis of heterogeneous samples would benefit from HSI imaging. In the cases where the sample is heterogeneous, but the focus of the analysis is not on studying the distribution of the constituents, as it is the case considered, where the focus was on distinguishing different types of Pesto. The obtained results show that the heterogeneity of the sample could be handled by the diffusive reflectance of the spot collected by a single-spot spectrophotometer by carefully planning the acquisition set up.

We suggested an easy setup to numerically assess the single-spot spectrophotometer collected area, by using a black/white paper template of the size of the sampling device, with decreasing black areas. This procedure can be generalized to different experimental setup.

Thus, there is evidence that a set up with a certain focal distance may cover a larger area, and actually the acquired area can be large enough as the ROI area for which homogeneity is reached by using a NIR-HSI camera.

In the present case, we corroborate this by also comparing the kNN classification performance obtained with the single-spot NIR with that of the NIR-HIS, which resulted to be the same.

From a PAT perspective, it could have a huge impact because with a simple explorative analysis is possible to define which instrument is better to handle the sample heterogeneity, considering also the different economic impact of the two instruments. The NIR-HSI Camera is more expensive than single-spot NIR. In addition, the focal distance at which to execute the measure with the single-spot NIR may be tuned depending on the heterogeneity of the sample at hand.

Furthermore, this work also suggests how to rationally select a proper ROI size, also in the case where an HSI camera is used. In fact, a simple PCA analysis conducted on the augmented data set (obtained by increasing ROI areas) can reveal the minimum area to be taken into account to handle the sample heterogeneity.

Although this study provides a detailed comparison between classical near-infrared (NIR) and hyperspectral NIR (HSI-NIR) measurements, it has some limitations that deserve to be discussed for a complete evaluation of the methodologies used. Firstly, for both techniques, there is a margin of error in the definition of the measured area, which cannot be exactly derived just from the focal distance in the case of punctual NIR, as well as the definition of ROI in HSI is limited by the square shape of the pixels. Furthermore, for the latter, a more accurate optimisation of the set-up is required to ensure optimal target focus. The results of this study clearly indicate that, for heterogeneous samples where the analysis is not aimed at the distribution of constituents, both NIR and HSI techniques can be used, but with significant differences in terms of time and cost. Single point NIR offers a faster and less expensive solution, making it ideal for rapid and less detailed analysis. Nowadays, it is easy to find NIR spectrophotometers covering the wavelength range used in this work for less than 10000 euros. In contrast, even though a great job is being done in the hardware of HSI-NIR systems in terms of acquisition speed, they are still expensive (more than 50000 euros). Nevertheless, the great advantage of HSI-NIR over NIR is its usefulness in cases where a more in-depth analysis of ingredient distribution is required.

## CRediT authorship contribution statement

**Daniele Tanzilli:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing, Software, Formal analysis, Data curation, Visualization. **Marina Cocchi:** Conceptualization, Project administration, Formal analysis, Supervision, Validation, Writing – review & editing. **José Manuel Amigo:** Conceptualization, Project administration, Methodology, Formal analysis, Validation, Supervision, Writing – original draft, Writing – review & editing. **Alessandro D'Alessandro:** Investigation, Data curation, Validation, Visualization. **Lorenzo Strani:** Conceptualization, Methodology, Formal analysis, Data curation, Validation, Supervision, Writing –

review & editing.

## Declaration of competing 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.

## Data availability

Data will be made available on request.

## Acknowledgements

D. Tanzilli acknowledges support from the COST Action CA19145 “European Network for Assuring Food Integrity using Non-Destructive Spectral Sensors (SENSORFINT), in the form of a mobility Grant which allowed him to work in Prof. Amigo research group in Bilbao to this project. José Manuel Amigo thanks the Innovation Fund Denmark for supporting this research under “HyperSort - Hyper-spectral imaging for advanced textile sorting and food control” project. Besides, this research has been supported by the Elkartek Programme, Basque Government (Spain) (SMART-EYE, KK-2023/00021).

## References

Alamar, M.C., Aleixos, N., Amigo, J.M., Barbin, D., Blasco, J., 2023. Hyperspectral imaging techniques for quality assessment in fresh Horticultural Produce and Prospects for measurement of mechanical Damage. In: Pathare, P.B., Opara, U.L. (Eds.), Mechanical Damage in Fresh Horticultural Produce: Measurement, Analysis and Control. Springer Nature, pp. 69–90. https://doi.org/10.1007/978-981-99-7096-4_4.

Amigo, J., 2019. Hyperspectral Imaging, first ed. Elsevier Data Handling in Science and Technology. 32.

Amigo, J.M., Jespersen, B.M., van den Berg, F., Jensen, J.J., Engelsen, S.B., 2023. Batch-wise versus continuous dough mixing of Danish butter cookies. A near infrared hyperspectral imaging study. Food Chem. 414, 135731 https://doi.org/10.1016/j.foodchem.2023.135731.

Bittante, G., Patel, N., Cecchinato, A., Berzaghi, P., 2022. Invited review: a comprehensive review of visible and near-infrared spectroscopy for predicting the chemical composition of cheese. J. Dairy Sci. 105 (3), 1817–1836. https://doi.org/10.3168/jds.2021-20640.

Cairós, C., Amigo, J.M., Watt, R., Coello, J., Maspoch, S., 2009. Implementation of enhanced correlation maps in near infrared chemical images: application in pharmaceutical research. Talanta 79 (3), 657–664. https://doi.org/10.1016/j.talanta.2009.04.042.

da Silva, N.C., de Moura França, L., Amigo, J.M., Bautista, M., Pimentel, M.F., 2018. Evaluation and assessment of homogeneity in images. Part 2: homogeneity assessment on single channel non-binary images. Blending end-point detection as example. Chemometr. Intell. Lab. Syst. 180, 15–25. https://doi.org/10.1016/j.chemolab.2018.06.011.

de Moura França, L., Amigo, J.M., Cairós, C., Bautista, M., Pimentel, M.F., 2017. Evaluation and assessment of homogeneity in images. Part 1: Unique homogeneity percentage for binary images. Chemometr. Intell. Lab. Syst. 171, 26–39. https://doi.org/10.1016/j.chemolab.2017.10.002.

Dixit, Y., Hitchman, S., Hicks, T.M., Lim, P., Wong, C.K., Holibar, L., Gordon, K.C., Loeffen, M., Farouk, M.M., Craigie, C.R., Reis, M.M., 2021. Non-invasive spectroscopic and imaging systems for prediction of beef quality in a meat processing pilot plant. Meat Sci. 181, 108410 https://doi.org/10.1016/j.meatsci.2020.108410.

dos Santos, C.A.T., Lopo, M., Páscoa, R.N.M.J., Lopes, J.A., 2013. A review on the Applications of portable near-infrared spectrometers in the Agro-food industry. Appl. Spectrosc. 67 (11), 1215–1233. https://doi.org/10.1366/13-07228.

Fernández Pierna, J.A., Vermeulen, P., Amand, O., Tossens, A., Dardenne, P., Baeten, V., 2012. NIR hyperspectral imaging spectroscopy and chemometrics for the detection of undesirable substances in food and feed. Chemometr. Intell. Lab. Syst. 117, 233–239. https://doi.org/10.1016/j.chemolab.2012.02.004.

Folch-Fortuny, A., Vitale, R., de Noord, O.E., Ferrer, A., 2017. Calibration transfer between NIR spectrometers: new proposals and a comparative study. J. Chemometr. 31 (3), e2874 https://doi.org/10.1002/cem.2874.

Fonseca Diaz, V., Roger, J.-M., Saeys, W., 2022. Unsupervised dynamic orthogonal projection. An efficient approach to calibration transfer without standard samples. Anal. Chim. Acta 1225, Scopus. https://doi.org/10.1016/j.aca.2022.340154.

França, L., Grassi, S., Pimentel, M.F., Amigo, J.M., 2021. A single model to monitor multistep craft beer manufacturing using near infrared spectroscopy and chemometrics. Food Bioprod. Process. 126, 95–103. https://doi.org/10.1016/j.fbp.2020.12.011.

Gorla, G., Ferrer, A., Giussani, B., 2023. Process understanding and monitoring: a glimpse into data strategies for miniaturised NIR spectrometers. Anal. Chim. Acta 1281, 341902. https://doi.org/10.1016/j.aca.2023.341902.

Grassi, S., Alamprese, C., 2018a. Advances in NIR spectroscopy applied to process analytical technology in food industries. Curr. Opin. Food Sci. 22, 17–21. https://doi.org/10.1016/j.cofs.2017.12.008.

Grassi, S., Alamprese, C., 2018b. Advances in NIR spectroscopy applied to process analytical technology in food industries. Curr. Opin. Food Sci. 22, 17–21. https://doi.org/10.1016/j.cofs.2017.12.008.

Grassi, S., Strani, L., Alamprese, C., Pricca, N., Casiraghi, E., Cabassi, G., 2022. A FT-NIR process analytical technology approach for Milk Renneting control. Foods 11 (1). https://doi.org/10.3390/foods11010033. Article 1.

Grassi, S., Giraudo, A., Novara, C., Cavallini, N., Geobaldo, F., Casiraghi, E., Savorani, F., 2023. Monitoring chemical changes of Coffee Beans during Roasting using real-time NIR spectroscopy and chemometrics. Food Anal. Methods 16 (5), 947–960. https://doi.org/10.1007/s12161-023-02473-w.

Kays, S.E., Barton, F.E., Windham, W.R., 2000. Predicting protein content by near infrared reflectance spectroscopy in diverse cereal food products. J. Near Infrared Spectrosc. 8 (1), 35–43. https://doi.org/10.1255/jnirs.262.

Li, L., Huang, W., Wang, Z., Liu, S., He, X., Fan, S., 2022. Calibration transfer between developed portable Vis/NIR devices for detection of soluble solids contents in apple. Postharvest Biol. Technol. 183, 111720 https://doi.org/10.1016/j.postharvbio.2021.111720.

Ma, X., Luo, H., Zhang, F., Gao, F., 2022. Study on the influence of region of interest on the detection of total sugar content in apple using hyperspectral imaging technology. Food Science and Technology 42, e87922. https://doi.org/10.1590/fst.87922.

Måge, I., Wubshet, S.G., Wold, J.P., Solberg, L.E., Böcker, U., Dankel, K., Lintvedt, T.A., Kafle, B., Cattaldo, M., Matić, J., Sorokina, L., Afseth, N.K., 2023. The role of biospectroscopy and chemometrics as enabling technologies for upcycling of raw materials from the food industry. Anal. Chim. Acta 1284, 342005. https://doi.org/10.1016/j.aca.2023.342005.

Nobari Moghaddam, H., Tamiji, Z., Akbari Lakeh, M., Khoshayand, M.R., Haji Mahmoodi, M., 2022a. Multivariate analysis of food fraud: a review of NIR based instruments in tandem with chemometrics. J. Food Compos. Anal. 107, 104343 https://doi.org/10.1016/j.jfca.2021.104343.

Nobari Moghaddam, H., Tamiji, Z., Akbari Lakeh, M., Khoshayand, M.R., Haji Mahmoodi, M., 2022b. Multivariate analysis of food fraud: a review of NIR based instruments in tandem with chemometrics. J. Food Compos. Anal. 107, 104343 https://doi.org/10.1016/j.jfca.2021.104343.

Ozaki, Y., McClure, W.F., Christy, A.A., 2006. Near-Infrared Spectroscopy in Food Science and Technology. John Wiley & Sons.

Porep, J.U., Kammerer, D.R., Carle, R., 2015. On-line application of near infrared (NIR) spectroscopy in food production. Trends Food Sci. Technol. 46 (2), 211–230. https://doi.org/10.1016/j.tifs.2015.10.002. Part A.

Pu, Y.-Y., O'Donnell, C., Tobin, J.T., O'Shea, N., 2020. Review of near-infrared spectroscopy as a process analytical technology for real-time product monitoring in dairy processing. Int. Dairy J. 103, 104623 https://doi.org/10.1016/j.idairyj.2019.104623.

Qiao, L., Mu, Y., Lu, B., Tang, X., 2023. Calibration Maintenance application of near-infrared Spectrometric model in food analysis. Food Rev. Int. 39 (3), 1628–1644. https://doi.org/10.1080/87559129.2021.1935999.

Sirisomboon, P., 2018. NIR spectroscopy for quality evaluation of fruits and vegetables. Mater. Today: Proc. 5 (10), 22481–22486. https://doi.org/10.1016/j.matpr.2018.06.619. Part 3.

Squeo, G., De Angelis, D., Summo, C., Pasqualone, A., Caponio, F., Amigo, J.M., 2022. Assessment of macronutrients and alpha-galactosides of texturised vegetable proteins by near infrared hyperspectral imaging. J. Food Compos. Anal. 108, 104459 https://doi.org/10.1016/j.jfca.2022.104459.

Tanzilli, D., D'Alessandro, A., Tamelli, S., Durante, C., Cocchi, M., Strani, L., 2023. A Feasibility study towards the on-line quality assessment of pesto sauce production by NIR and chemometrics. Foods 12 (8). https://doi.org/10.3390/foods12081679. Article 8.

van den Berg, F., Lyndgaard, C.B., Sørensen, K.M., Engelsen, S.B., 2013. Process analytical technology in the food industry. Trends Food Sci. Technol. 31 (1), 27–35. https://doi.org/10.1016/j.tifs.2012.04.007.

Xu, H., Ren, J., Lin, J., Mao, S., Xu, Z., Chen, Z., Zhao, J., Wu, Y., Xu, N., Wang, P., 2023. The impact of high-quality data on the assessment results of visible/near-infrared hyperspectral imaging and development direction in the food fields: a review. J. Food Meas. Char. 17 (3), 2988–3004. https://doi.org/10.1007/s11694-023-01822-x.