University of Modena and Reggio Emilia

Research Doctorate Course in Agri-food Science, Technology and Biotechnology

XXXVII cycle

Multivariate analysis of Big Data from non-destructive analytical techniques in the agri-food sector

Ph.D. Candidate: Camilla Menozzi

Tutor: Dr. Rosalba Calvini

Co-Tutor: Prof. Alessandro Ulrici

Co-Tutor: Dr. Giorgia Foca

PhD Course Coordinator: Prof. Fabio Licciardello

Riassunto

L’implementazione di metodi analitici veloci, economici e sostenibili svolge un ruolo cardine nel progresso tecnologico del settore agroalimentare. Siccome sia l'aspetto esteriore che la composizione chimica degli alimenti sono decisivi nel valutarne la qualità, le applicazioni di tecniche di imaging e spettroscopia nell’agricoltura di precisione e per il monitoraggio in tempo reale sono in aumento, consentendo una caratterizzazione oggettiva, rapida e automatica dei campioni alimentari.

L’utilizzo dell’imaging digitale RGB è oggi favorito grazie alla diffusione di sistemi tecnologicamente avanzati, facili da usare ed economici, ad esempio uno smartphone, in grado di catturare l’intensità della regione visibile dello spettro elettromagnetico nei canali del rosso, verde e blu. Colore e texture sono fra le principali proprietà che possono essere ricavate da un’immagine RGB e legate a aspetti qualitativi in diversi prodotti alimentari.

In base al tipo di applicazione, la qualità alimentare può dipendere da alcune proprietà esplorabili attraverso regioni dello spettro elettromagnetico diverse da quella visibile. A questo riguardo, la spettroscopia NIR è una tecnica veloce, non distruttiva e versatile che consente di determinare più costituenti chimici nella maggioranza dei prodotti alimentari. È inoltre possibile affrontare con successo alcune sfide in ambito alimentare tramite dispositivi economici e portatili già in commercio che operano in regioni spettrali ristrette.

Le tecniche di imaging e spettroscopiche tendono a generare in poco tempo grandi quantità di dati, e i dataset risultanti possono rientrare nella definizione di Big Data, che non possono essere elaborati con strumenti tradizionali di analisi dei dati, ma che richiedono un approccio statistico multivariato per estrarre l’informazione utile.

Il filo conduttore di questo progetto di dottorato è costituito dallo sviluppo e applicazione di metodi analitici avanzati basati su statistica multivariata e algoritmi di machine learning per analizzare spettri NIR e immagini RGB per la caratterizzazione di diverse matrici alimentari. In particolare, il focus di questa tesi verte sulla valutazione di strategie chemiometriche in grado di gestire efficacemente i Big Data ottenuti da tecniche di imaging e spettroscopiche, al fine di massimizzare la quantità di informazioni che possono essere acquisite e minimizzare lo sforzo necessario per l'acquisizione delle informazioni stesse.

Si valuta l'applicazione di algoritmi di riduzione della dimensionalità dei dati per estrarre e correlare le proprietà relative a colore e texture di immagini RGB di uve rosse con il corrispondente contenuto di antociani tramite modelli di calibrazione multivariata mediante PLS. Le immagini sono state acquisite utilizzando un semplice smartphone tramite un'interfaccia user-friendly e un dispositivo portatile. L’ampio dataset di immagini così ottenuto è stato convertito in una matrice di segnali chiamati colorigrammi, che è stato quindi analizzato con tecniche multivariate. È stato anche sviluppato un metodo alternativo per estrarre e codificare le proprietà di texture di immagini RGB in matrici di segnali denominati texturegrammi. La sua efficacia è stata testata su un subset di immagini di uve rosse, confrontando i risultati ottenuti da diversi metodi di estrazione di features e valutando possibili vantaggi della combinazione dell’informazione legata a colore e texture mediante diversi approcci di data fusion.

Un’altra parte del lavoro di ricerca è stato dedicato alla valutazione dell'efficacia della spettroscopia NIR accoppiata alla chemiometria per affrontare sfide caratteristiche nell’ambito agroalimentare, considerando anche diversi dispositivi di acquisizione, da banco e portatili.

Abstract

The implementation of fast, economic, and sustainable analytical methods plays a key role for technological advances in the agri-food sector. Since both visual aspect and chemical composition of food has a decisive impact on its quality evaluation, the applications of imaging and spectroscopic techniques in contexts like precision agriculture and real-time food monitoring are rapidly increasing, allowing an objective, fast and automatic characterization of food samples.

The use of RGB digital imaging is now favoured thanks to technically sophisticated, user-friendly and affordable systems, like common smartphones, whose sensors can capture the intensity of the visible region of the electromagnetic spectrum in the red, green and blue channels. Colour and texture are the main features that can be derived from RGB images and related to quality aspects in a wide variety of food products.

According to the application at hand, food quality could depend on some properties related to regions of the electromagnetic spectrum other than the visible. In this context, NIR spectroscopy is a fast, non-destructive and versatile technique allowing to determine many chemical constituents in most food products. Some food-related applications could be even successfully handled through cheap, portable and commercially available devices working in narrow spectral regions.

Imaging and spectroscopic techniques tend to generate large amounts of data in a very short time, and the resulting datasets can be covered by the definition of Big Data. Indeed, Big Data refers to extremely large and complex datasets that cannot be elaborated with traditional data processing tools but require a multivariate statistical approach to extract useful information.

The main topic of this PhD project consists in the development and application of advanced analytical methods based on multivariate statistics and machine learning algorithms to analyse NIR spectra and RGB image data for the characterization of different food matrices. More in detail, this PhD thesis is focused on the evaluation of chemometric strategies able to effectively manage Big Data resulting from imaging and spectroscopic techniques, in order to maximize the amount of information that can be acquired, while minimizing the effort required to acquire that information.

The thesis investigates the application of data dimensionality reduction algorithms to extract, codify and relate the colour- and texture- related information in RGB images of red grape samples with the corresponding anthocyanins content through PLS calibration models. Images were acquired using a common smartphone by means of a user-friendly interface and a portable device. A large dataset of images was acquired and converted into matrices of signals named colourgrams, subsequently analysed using multivariate approaches. Research activities focused also on the development of an alternative method to extract and codify the texture properties of RGB images into matrices of signals named texturegrams. The effectiveness of this approach has been tested on a subset of red grape sample images, by comparing the results obtained from different feature extraction methods and investigating the possible advantages of combining colour and texture information through different data fusion approaches.

Another part of the research work was devoted to evaluating the effectiveness of NIR spectroscopy coupled with chemometrics to address distinct challenges in the agri-food sector and considering also different acquisition devices, i.e., benchtop and portable instruments.

Table of contents

Chapter One

Introduction

1.1 Background of the Thesis

The development of rapid, cost-effective and sustainable analytical methods stands out as a major need for the technological progress in the agri-food sector, especially considering the recent Agriculture 4.0 paradigm.

Quality evaluation of agri-food products is a complex issue, requiring consideration of different aspects for proper assessment. In particular, both the visual appearance of food and its chemical composition represent two main basic pillars for the consumers' choices and for an effective evaluation of food quality.

In this framework, the application of computer vision techniques through Red-Green-Blue (RGB) imaging has gained an increasing relevance to evaluate colour-related food properties, due to the multiple advantages provided, like the possibility to get a fast, objective and automated analysis of the visual aspect of food samples (Brosnan and Sun, 2004; Ohta and Robertson, 2006; Ornberg et al., 1999; Wu and Sun, 2013). Furthermore, image acquisition can be performed with inexpensive and commonly used devices, such as webcams, smartphones, flatbed scanners or cameras, which in turn can be coupled with user-friendly software to integrate image acquisition and elaboration. Sample colour and texture, related to quality aspects in a wide variety of food products, are the main properties that can be obtained from RGB images.

While computer vision systems are very effective in assessing colour-related properties of the analysed samples, the analytical information of interest is sometimes associated with regions of the electromagnetic spectrum beyond the visible range. In this framework, near infrared spectroscopy (NIR) often stands out as a recommended choice. In fact, NIR spectroscopy allows to determine many chemical constituents in most food products in a fast, cheap, reagent-free and non-destructive manner, resulting more sustainable and advantageous compared to traditional analytical methods. These characteristics make NIR spectroscopy a profitable tool in many food-related applications.

The use of imaging and spectroscopic techniques often results in the production of large amounts of data, also referred to as "Big Data". This term stands for extremely large and complex datasets that cannot be processed by means of traditional univariate tools, but they require a multivariate statistical approach to extract meaningful information effectively. This aspect, in combination with the

multivariate nature of image and spectroscopic data, makes the application of chemometric strategies a mandatory analysis approach for achieving reliable, interpretable and adequate outcomes.

In the agri-food context, the wine sector clearly represents a field of key importance due to its economic relevance worldwide. At the same time, extreme climatic conditions and the spread of fungal diseases represent current challenges for winemakers. Indeed, these issues caused a decline of vineyard surface areas and, consequently, a decrease in wine production (International Organization of Vine and Wine (OIV), 2023a, 2023b).

In this context, the wine sector can greatly benefit from advanced analytical methods based on imaging and spectroscopic techniques to perform real-time and on-site quality assessment throughout the production chain. For these reasons, the applications presented in this PhD Thesis are mainly related to the wine sector.

1.2 Outline of the Thesis

This PhD Thesis is focused on the application and development of advanced analytical methods based on multivariate statistics for the analysis of image and spectral data, as well as the characterization of different food matrices, with special emphasis on the wine sector.

In particular, the following issues are investigated:

- development and application of a smartphone-based portable device for the determination of grape phenolic maturity. RGB images were processed at the image-level using the *colourgrams* method, a data reduction technique capable of encoding colour-related features of the corresponding images into signals;
- development of the *texturegrams* approach, an innovative method to extract and codify texture-related information in RGB images into matrices of signals, named *texturegrams*. Considering the quantification of anthocyanins from RGB images of red grapes as a case study, the effectiveness of the texturegrams approach was compared with state-of-the-art methods to extract texture features from images. In addition, the benefits of coupling colour and texture features through data fusion were also evaluated;
- application of NIR spectroscopy coupled with chemometrics to monitor alcoholic and acetic fermentation in red grape musts, considering also different spectral ranges potentially covered by different devices, i.e., benchtop and portable instruments.

The present Thesis is organized as follows:

- **Chapter 2** describes the key chemometric strategies used in this Thesis for the elaboration of image and spectral data;
- **Chapter 3** presents an overview of RGB imaging, with a focus on multivariate image analysis (MIA) and the image-level feature extraction techniques used in this work;
- **Chapter 4** presents an overview about NIR spectroscopy, considering both theoretical and instrumental aspects;
- **Chapter 5** shows the application of chemometrics coupled with RGB imaging and NIR spectroscopy to address different challenges in the wine and vinegar production sector. The applications are presented in form of preprints of the following articles:
    - Menozzi, C., Calvini, R., Nigro, G., Tessarin, P., Bossio, D., Calderisi, M., Ferrari V., Foca G., Ulrici, A. (2023). Design and application of a smartphone-based device for in vineyard determination of anthocyanins content in red grapes. *Microchemical Journal*, *191*, 108811. doi: 10.1016/j.microc.2023.108811;
    - Menozzi, C., Prats-Montalbán, J.M., Calvini, R., Ulrici, A. (2025). Comparison of colour and texture feature extraction methods to predict anthocyanins content in grapes. Submitted to *Chemometrics and Intelligent Laboratory Systems*;
    - Menozzi, C., Foca, G., Calvini, R., Catellani, L., Bezzecchi, A., & Ulrici, A. (2024). Comparison of Different Spectral Ranges to Monitor Alcoholic and Acetic Fermentation of Red Grape Must Using FT-NIR Spectroscopy and PLS Regression. *Food Analytical Methods*, 17, 1171-1182. doi: 10.1007/s12161-024-02636-3.
- **Chapter 6** reports the general conclusions derived from the outcomes obtained in the frame of this research work.

1.3 References

Brosnan, T., Sun, D.-W., 2004. Improving quality inspection of food products by computer vision—a review. Journal of Food Engineering 61, 3–16. https://doi.org/10.1016/S0260-8774(03)00183-3

International Organization of Vine and Wine (OIV), 2023a. Annual assessment of the world vine and wine sector in 2023.

International Organization of Vine and Wine (OIV), 2023b. State of the WOrld Vine and WIne Sector in 2023.

Ohta, N., Robertson, A., 2006. Colorimetry: fundamentals and applications. John Wiley & Sons.

Ornberg, R.L., Woerner, B.M., Edwards, D.A., 1999. Analysis of Stained Objects in Histological Sections by Spectral Imaging and Differential Absorption. J Histochem Cytochem. 47, 1307–1313. https://doi.org/10.1177/002215549904701010

Wu, D., Sun, D.-W., 2013. Colour measurements by computer vision for food quality control – A review. Trends in Food Science & Technology 29, 5–20. https://doi.org/10.1016/j.tifs.2012.08.004

Chapter Two

Chemometrics

2.1 Multivariate analysis methods

Food quality evaluation is certainly a complex concept and the characterization of the properties of food matrices usually requires performing analytical measurements on a high number of samples and obtain a high number of variables with a single measurement. In other words, multiple parameters are measured for each sample, and this is the case, for example, of spectroscopic and image datasets.

The possibility to characterize the sample properties by means of multiple variables, however, does not necessarily involve a correct extraction of the relevant information. Indeed, the choice of the proper data elaboration approach plays a key role in the effective assessment of food-related applications. By focusing on one-variable-at-a-time (OVAT), univariate data analysis often provides a limited knowledge of the system under investigation, with the consequent risk of losing valuable information. In fact, OVAT approaches preclude the possibility to evaluate real interactions occurring between samples and variables, counterintuitively to the world we live and experience.

Proper management of multivariate datasets, as described above, needs a simultaneous processing of all the data. Multivariate statistics can make the most of the relevant information contained in complex data matrices, by means of a simultaneous exploration of the effects owed to all the variables and their possible interactions to characterize the considered samples and according to the real data structure.

Chemometrics can be defined as "*the chemical discipline that uses mathematical and statistical methods for the obtention in the optimal way of relevant information on material systems*" (Frank and Kowalski, 1982). In the context of multivariate analysis of image and spectroscopic data, the use of chemometric methods allows to perform exploratory data analysis by identifying groups of objects sharing similar features, classify objects according to the categories of interest, calculate calibration models to quantify analytical parameters of interest and to visualise their spatial distribution, identify and select the relevant information through variable selection algorithms, apply data dimensionality reduction strategies to facilitate the simultaneous analysis of large datasets and apply data fusion strategies to combine the information coming from different sources to optimize model performances and properly address some given issues.

In the agri-food sector the applications of chemometrics are wide and reach multiple contexts, like quality evaluation, food control and process monitoring.

Many advantages can be ascribed to the use of chemometric strategies, including:

- Proper and relatively fast elaboration of Big Data and extraction of high-quality information from rapid, objective and automated techniques (e.g., imaging and spectroscopic methods);
- Deep knowledge of the system or process under investigation;
- Improvement of experimental procedures and data acquisition methods by Design of Experiments.

2.2 Data preprocessing

Prior to the application of any chemometric method, data are required to be properly pre-processed in order to remove or reduce the effect owed to undesired variability (such as noise), which could influence the interpretability and the quality of the results, and to facilitate the extraction of useful and quality information in view of the successive data analysis. Multiple preprocessing methods are available and the choice of the most appropriate is not always trivial, depending on the type of data at hand. At the same time, the preprocessing methods should not introduce troublesome artifacts and, in general, remove relevant information.

The wide variety of preprocessing methods can be essentially outlined into two main categories: column-wise and row-wise preprocessing.

Concerning column-wise preprocessing methods, they are performed on each variable over all the objects of the dataset, and they are used to correct systematic variations among variables. The most common column-wise preprocessing methods are:

- Mean centring (MC): for each variable j of the data matrix, the mean value of the corresponding column vector $\bar{\boldsymbol{x}}_{\boldsymbol{j}}$ is subtracted from each vector element x_{ij}:

  $$x_{ij,MC} = x_{ij} - \bar{x}_j \qquad (2.\ 1)$$

  As shown in Eq. 2.1, mean centring enhances the information related to the variation of each variable, by removing the contribution of the corresponding average value;
- Autoscaling (AUTO): for each variable j, the corresponding column vector $\boldsymbol{x}_{\boldsymbol{j}}$ is mean centred and scaled to unit variance by subtracting from each vector element x_{ij} the mean value $\bar{\boldsymbol{x}}_{\boldsymbol{j}}$ and then by dividing the result by the standard deviation s_j:

$$x_{ij,AUTO} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (2.\ 2)$$

Autoscaling allows to give an equal weight to all the dataset variables. This column preprocessing is useful when dealing with variables of different nature (such as variables with different measurements units but supposed to be of comparable meaning), since these scale differences can hide other relevant sources of information.

Row-wise preprocessing methods are applied to each object over all its variables. These techniques are typically applied on analytical signals, like NIR spectra and chromatograms, to minimize the effect of uninformative systematic variation sources between the signals (like e.g., baseline offset or drift) and better resolve overlapping bands or peaks.

When dealing with NIR spectroscopy, the physical properties of the sample can have a significant impact on the corresponding signal and can introduce useless information in the final output. For example, the particle size can affect light diffusion and cause scattering phenomena, and multiplicative deviations could be produced also when dealing with emulsions and dispersions. Furthermore, the instrumental sensitivity or changes in the optical pathlength could result in other sources of undesired contributions/perturbations. A proper row-wise pretreatment is required to minimize these irrelevant contributions and to work on comparable spectra. This is particularly important when evaluating properties of samples which are not dependent on physical appearance, like the sample chemical composition. Although row-wise pretreatments are often necessary and profitable, one must be aware that some of them could sometimes alter and confuse even part of the useful information present in the signals. Indeed, the choice of the optimal preprocessing method is fundamental. It depends on the nature of the data and it is always important to avoid the introduction of artifacts as well as the removal of relevant information (Engel et al., 2013).

Multiple row-wise pretreatments are available to reduce and correct different noisy and useless contributions (i.e., those sources of variation not bringing useful information) and some of the most common and applied in this Thesis are described below along with some recommendations:

- Standard Normal Variate (SNV): this method reduces the multiplicative effects owed to scattering, particle size and multicollinearity changes over the NIR spectra (Barnes et al., 1989; Oliveri et al., 2019). In fact, this pretreatment consists in centring and scaling to unit variance the row vector $\boldsymbol{x_i}$ that corresponds to the spectrum of each object i, following the equation:

$$x_{ij,SNV} = \frac{x_{ij} - \bar{x}_i}{s_i} \quad (2.\ 3)$$

where $x_{ij,SNV}$ is the element of the preprocessed spectrum, x_{ij} is the corresponding element of the original spectrum of object i at variable j, $\bar{x}_i$ and s_i are the mean and the standard deviation of $\boldsymbol{x_i}$, respectively. This preprocessing can be considered a "row-wise autoscaling", since it can be regarded as mean centring and scaling to unit variance in the object direction;

- Detrend: this pretreatment is intended to correct the baseline shift when a constant, linear or curved offset is present (Barnes et al., 1989). It consists in fitting a polynomial of a given order to the whole spectrum and then subtracting the polynomial from the corresponding spectrum;
- Derivatives: this pretreatment is generally applied to highlight the separation of overlapping peaks and to eliminate uninformative baseline offsets. First derivative removes the effects of an additive constant background, while second derivative eliminates both additive effects and baseline linear slope variations (Oliveri et al., 2019). The derivative procedure is often performed by means of the Savitzky-Golay algorithm (Savitzky and Golay, 1964), which works by fitting sub-sets of data points (defined by a moving window, whose size is user-defined) to a polynomial function calculated using least squares regression. The coefficients of the polynomial function are used to calculate the derivative value of each variable (i.e., the wavelength in the case of NIR spectra) at the centre of the moving window. Furthermore, the algorithm can also be used for smoothing, aimed at filtering out the high frequency noise of the original signal; in this case, the value at the centre of the moving window is given by the corresponding value calculated by the polynomial.

2.3 Principal Component Analysis

The implementation of fast and automated analytical techniques implicitly entails the generation of large amounts of data in a relatively short time. Indeed, effective and reliable sample characterization is often achieved thanks to the possibility to rely on multiple descriptors, as in the case of spectroscopic and imaging techniques.

Within this framework, proper exploration of the resulting datasets is clearly as crucial as nontrivial to reach an adequate knowledge of the real data structure and of the information contained in the system under investigation. A multivariate exploratory approach is strictly required to address data analysis in the most possible correct and exhaustive way.

Principal Component Analysis (PCA) can be regarded as a pillar of multivariate analysis, in many ways (Wold et al., 1987). Indeed, when dealing with new multivariate datasets PCA usually represents

the first and basic step of data analysis, being generally performed before applying calibration or classification algorithms. Due to the importance of multivariate measurements in chemistry and in agrifood sciences, PCA is a very efficient data exploration and interpretation tool that allows to simultaneously consider all the data variables, describing both objects structure, variables structure and possible correlations between objects and variables.

Given a multivariate data matrix $\boldsymbol{X}$, the latter is usually described by a huge number of original variables, which are often highly correlated. PCA basically computes a transformation of the original variables into a new and smaller set of uncorrelated variables, namely the *Principal Components* (PCs), which are linear combinations of the original variables and describe the major trends in the original data structure. PCs are intended to describe the real dimensionality of the data matrix, accounting for the relevant information of the system and enabling an easy and effective/profitable graphical description (by means of bi- or tri-dimensional plots) and interpretation of the data structure in a low-dimensional space, the latter something often impossible to be achieved when strictly focusing only on the original data matrix for the reasons explained above.

From a geometrical point of view, PCA describes, simplifies and projects the information contained in the data from an original m-dimensional space (with m = number of original variables) in a simpler A-dimensional one (with A = number of PCs). Indeed, PCA allows to represent the relevant information contained in the original data in a low-dimensional space, whose axes are not represented by the m original variables, but by a smaller set of A variables, i.e., the PCs. In other words, PCA is an unsupervised exploratory data analysis tool that allows to both explore the real data structure and profitably compress data dimensionality while improving the analysis of relevant information contained in the data and its visualization.

By definition, PCs are orthogonal, thus they describe different types of information, and are ordered according to the decreasing percentage of variance they account for. More in detail, each PC explains a given variability source in the original data. The first principal component (PC1) describes the main direction of data variability, accounting for the maximum percentage of explained variance of the analysed dataset. The second principal component (PC2) is defined to be orthogonal to PC1 and accounts for the maximum amount of the residual variance, i.e., of the variance not already described by PC1 and could be likewise relevant to describe the data structure. The same principles apply to define the subsequent principal components (Figure 2.1). Those PCs accounting for the smallest data variation, likely to be associated with noise, are then discarded achieving two main goals: data dimensionality reduction and noise reduction.

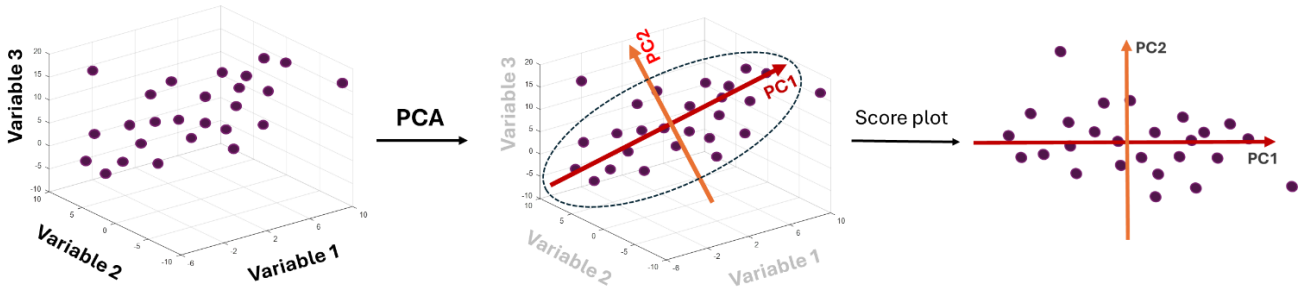


Figure 2.1 *Definition of the PC space with respect to the original domain.*

Before applying PCA it is mandatory to apply proper preprocessing techniques, so that the results of the analysis reflect the real data structure and the systematic variability of the data (*see* Section 2.2)

From a mathematical point of view, the original data matrix can be seen a matrix $\boldsymbol{X}$ with dimensions $\{n \times m\}$, corresponding to m variables measured on a set of n objects. PCA decomposes the variance of the original data matrix $\boldsymbol{X}$ (suitably pre-processed) in three matrices: the score matrix $\boldsymbol{T}$, accounting for the variance associated to the samples, the loading matrix $\boldsymbol{P}$, accounting for the variance associated to the variables, and the residuals matrix $\boldsymbol{E}$, accounting for the non-systematic variation (Figure 2.2). Such decomposition can be expressed according to the following equation:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{T} + \boldsymbol{E} = \sum_{a=1}^{A} \boldsymbol{t_a}\boldsymbol{p_a^T} + \boldsymbol{E} \qquad (2.\ 4)$$

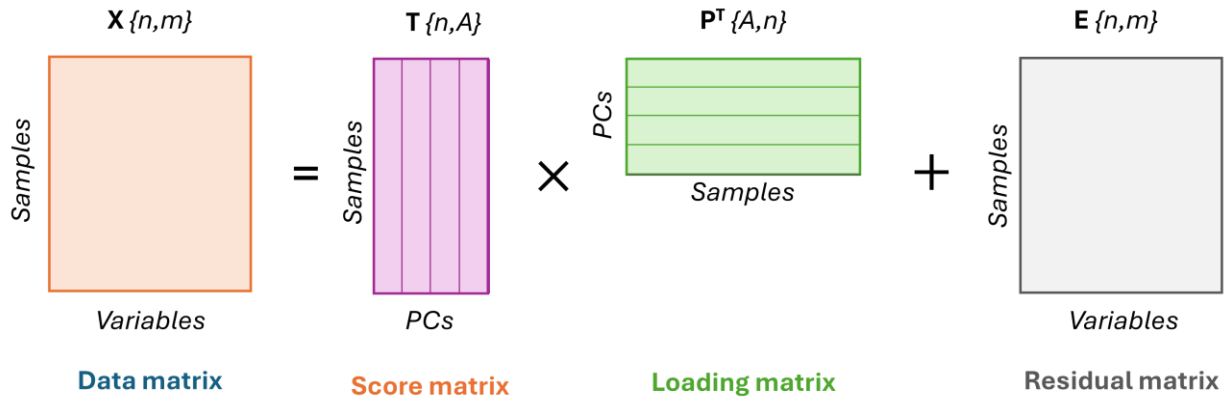


Figure 2.2 *Data matrix decomposition performed by PCA.*

In this manner, each PC can be described by two vectors, namely a score vector and a loading vector.

Original objects can effectively be observed with regard to the PCs directions in the new space, assuming a new set of values called scores. Indeed, concerning the score matrix $\boldsymbol{T}$, it has a number of rows equal to the number of objects and a number of columns equal to the number of PCs of the model. Each element of a score vector expresses the position of the orthogonal projection of the corresponding object onto the given PC. In general, the score values represent the coordinates of the objects when projected into the PCs space. In this regard, a score vector is computed for each PC, so it is possible to graphically represent the relative position of all the objects projected in the new PCs space, by means of a scatter plot (usually considering from one to three PCs at a time), named score plot. Objects score values along PCs can be positive or negative, since the origin of PCs correspond to the centroid of mean-centred data. Indeed, prior to perform PCA, data are required to be centred so that their variance is expressed with respect to zero mean.

PCA allows to visualize the real data structure and identify similarities and differences among samples, for example the presence of patterns, clusters, extreme samples (those having major absolute score values and playing a main influence on the PC direction) or outliers. An example of score plot is shown in Figure 2.1, where a 2 PCs space is considered to describe the informative part of the original three-dimensional data. The projection of each i-th object (with $1 \leq i \leq n$) into the 2-dimensional PCs space is described by two score values, $t_{i,1}$ and $t_{i,2}$ expressing its coordinates on PC1 and PC2, respectively.

Outliers are samples which locate far from the others and when present in the data could have a notable impact on the outcomes of the analysis, but not necessarily in a negative sense. Indeed, sometimes they could be relevant to explore some properties of the system under study, but they could adversely affect the direction of the PCs. Hence, it is strongly recommendable the detection of anomalous samples and their possible exclusion from the dataset. This is usually achieved by means of two statistical parameters:

- Q Residuals, which express the distance between each object and its projection on the hyperplane defined by the PCA model, i.e., the unusual variation of each object outside the model. In fact, high Q Residuals values correspond to objects not following the general pattern described by the PCA model (Figure 2.3). This parameter is calculated as the sum of squares of each row of the residuals matrix $\boldsymbol{E}$ according to the following equation for the i^{th} sample in $\boldsymbol{X}$ ($\boldsymbol{x_i}$):

$$Q_i = \boldsymbol{e_i e_i^T} = \boldsymbol{x_i}(\boldsymbol{I} - \boldsymbol{P P^T})\boldsymbol{x_i^T} \qquad (2.5)$$

where Q_i is the Q residual value of the i^{th} sample in $\boldsymbol{X}$ ($\boldsymbol{x_i}$), $\boldsymbol{e_i}$ is the i^{th} row of $\boldsymbol{E}$, $\boldsymbol{P}$ is the matrix of the A loading vectors of the PCA model (where each vector is a column of $\boldsymbol{P}$) and $\boldsymbol{I}$ is the identity matrix of appropriate size;

- Hotelling's T^2, which quantifies the distance of the projection of each object on the PCA model (hyperplane) from the centre of the model itself, i.e., it expresses the objects variability within the model and the contribution of each object to define the model itself. High Hotelling T^2 values indicate objects having a main influence on the model (Figure 2.3). This parameter is calculated as the sum of normalised squared score values according to this equation:

$$T_i^2 = \sum_{a=1}^{A} \frac{t_{ia}^2}{s_a^2} = \frac{t_{i1}^2}{s_1^2} + \frac{t_{i2}^2}{s_2^2} + \cdots + \frac{t_{iA}^2}{s_A^2} \qquad (2.6)$$

where t_{ia} is the score value of the i^{th} sample for the a^{th} PC and s_a^2 is the variance of the a^{th} score vector, t_a, in the model.

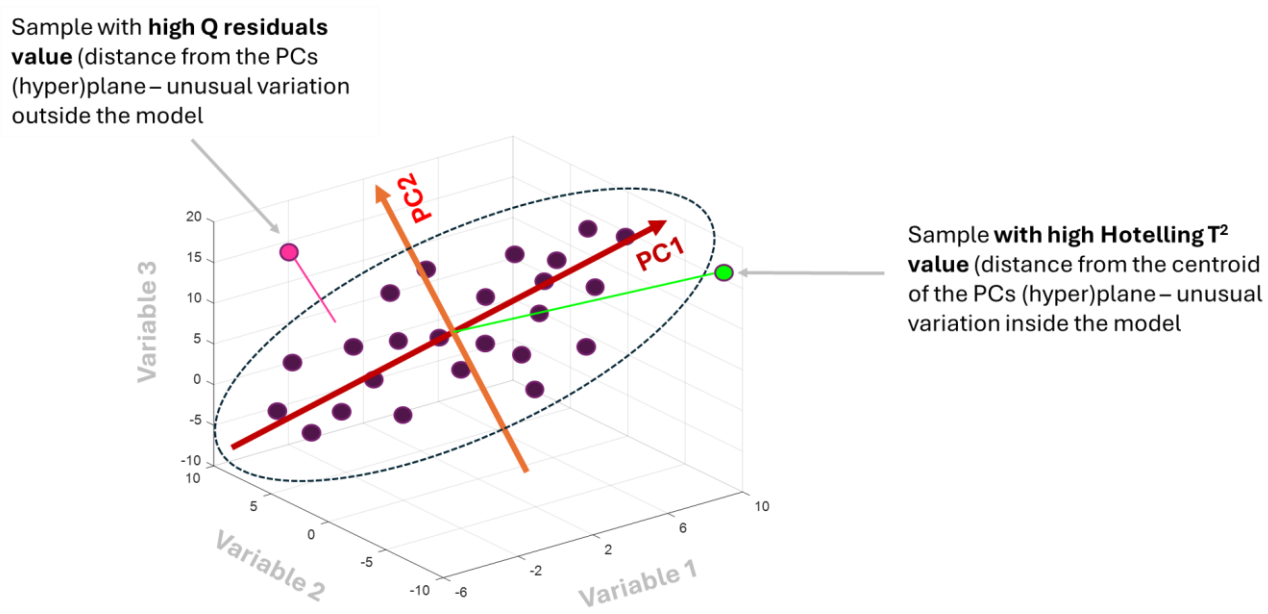


Figure 2.3 *Example of a sample with high Q value (in magenta colour) and of a sample with high T^2 value (in green colour).*

On the other side, the loading matrix $\boldsymbol{P}$ quantifies the contribution of each original variable to define the direction of each PC, i.e., accounts for the information related to the direction of each PC with respect to the direction of each original variable. The matrix $\boldsymbol{P}$ has as many rows as the number of original variables and as many columns as the number of PCs of the model. Each loading vector characterizes each PC, and each element of the vector quantifies the weight of each original variable

(once properly preprocessed) to define the direction expressed by the corresponding PC. As stated at the beginning of this section, each PC is a linear combination of the original variables, and the corresponding loading vector basically contains the coefficients of the original variables in defining the PCs. Therefore, high absolute values of a given loading coefficient correspond to high influence of the corresponding original variable for that PC direction.

The loading values can be graphically represented, usually considering two or three PCs at a time, using a scatter plot (as for the score values), named loading plot. The loading plot allows to evaluate the structure of the original variables by identifying correlations among variables or highlighting those having the main influence in describing the data structure. From a graphical point of view, a set of variables is positively correlated when they lie in the same direction with respect to the PCs origin, while an opposite direction indicates a negative correlation. At the same time, variables registering the highest absolute loading values are those having a major influence on the corresponding PC direction.

When multivariate data matrices are characterized by a particularly high number of variables (e.g., when the dataset is composed of spectra or signals), the interpretation of the corresponding loading plot is very difficult. In this case, to improve interpretability, the loading vectors can be plotted as a function of the original variable domain (e.g., wavelengths in the case of spectra). In this manner, the loading profiles are directly compared with the original variables simplifying the detection of the most relevant descriptors based on their absolute values: the farthest from the origin (i.e., $y = 0$ line), the highest the significance.

The extensive exploration of the system under investigation using PCA can be achieved also evaluating the relationships occurring between objects and original variables, i.e., by comparing score and loading plots. For example, if a subset of samples is located at positive scores of a given PC the comparison with the corresponding loading plot allows to identify which original variables mainly characterize those samples, since they are likely to present high values of the variables showing high PC loading values but low values of the variables showing low PC loading values.

2.4 Partial Least Squares

Partial Least Squares (PLS) is a multivariate linear regression and projection method intended to identify and define the underlying mathematical linear relationship existing between a descriptor matrix $\boldsymbol{X}$ (a matrix of independent variables describing a set of objects) and a response matrix $\boldsymbol{Y}$ (a

matrix of dependent variables related to the same set of objects) (Geladi and Kowalski, 1986; Martens and Naes, 1989; Wold et al., 2001).

The mathematical model between $\boldsymbol{X}$ and $\boldsymbol{Y}$ describes how the descriptors in $\boldsymbol{X}$ can be used to model the response matrix $\boldsymbol{Y}$ in order to predict, after a proper model validation, the response variables starting from the independent variables describing unknown samples.

In a similar way to PCA, PLS assumes that the system under investigation can be described by a smaller set of variables, named *Latent Variables* (LVs) which are linear combinations of the original variables and are orthogonal to each other. On the other hand, unlike PCA, PLS simultaneously considers the information brought by both blocks. In other words, the whole set of original variables in $\boldsymbol{X}$ are projected onto the smaller set of LVs with the contextual aim to maximize the covariance with the response matrix $\boldsymbol{Y}$. Both $\boldsymbol{X}$ and $\boldsymbol{Y}$ matrices can be separately decomposed as described in the following equations:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\boldsymbol{T}} + \boldsymbol{E} \tag{2. 7}$$

$$\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{Q}^{\boldsymbol{T}} + \boldsymbol{G} \tag{2. 8}$$

As stated above, in PLS the directions of the LVs are defined based on the assumption that the $\boldsymbol{X}$ matrix has to be a predictor of the $\boldsymbol{Y}$ matrix. Hence, the object variation of the descriptor block $\boldsymbol{X}$ expressed by the score matrix $\boldsymbol{T}$, can be used to describe $\boldsymbol{Y}$**,** as follows:

$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{Q}^{\boldsymbol{T}} + \boldsymbol{F} \tag{2. 9}$$

Therefore, the decomposition of $\boldsymbol{X}$ is optimized in a such a way that matrix $\boldsymbol{T}$ expresses the variation of $\boldsymbol{X}$ enabling the best possible description of $\boldsymbol{Y}$, which in turn implies that the covariance between $\boldsymbol{T}$ and $\boldsymbol{U}$ must be maximised. To this aim, for each LV, a weight vector ($\boldsymbol{w}$) is calculated to weight the original variables according to their contribution to explain the $\boldsymbol{Y}$ matrix. Hence, Eq. 2.9 can be re-organised to calculate the estimate of $\boldsymbol{Y}$ ($\widehat{\boldsymbol{Y}}$) as follows:

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{W}\boldsymbol{Q}^{\boldsymbol{T}} = \boldsymbol{X}\boldsymbol{B} \tag{2. 10}$$

where $\boldsymbol{W}$ is the matrix of weights and $\boldsymbol{B}$ represents the set of PLS regression coefficients.

Among the different algorithms available to calculate PLS models, NIPALS (Geladi and Kowalski, 1986) and SIMPLS (de Jong, 1993) are the most used. NIPALS is an iterative method, and it is more time-consuming with respect to SIMPLS, the latter however being more complex from a theoretical point of view, since it calculates the PLS factors directly as linear combinations of the original variables. Both methods return the same results if $\boldsymbol{Y}$ is univariate, while slight (but negligible) differences are observed with multivariate $\boldsymbol{Y}$.

The importance of the contribution of each original variable, i.e., original descriptor, to the final PLS model can be estimated and expressed by means of the Variables Importance in Projection (VIP) scores, calculated for each $\boldsymbol{y}$ variable with respect to each k^{th} descriptor according to the following equation (Gosselin et al., 2010):

$$VIP_k = \sqrt{K \sum_{a=1}^{A} [(q_{ka}^2 \boldsymbol{t}_a^T \boldsymbol{t}_a)(w_{ak}/\|\boldsymbol{w}_k\|^2)] / \sum_{a=1}^{A} (q_{ka}^2 \boldsymbol{t}_a^T \boldsymbol{t}_a)} \quad (2.\ 11)$$

where K is the number of variables of the descriptor matrix $\boldsymbol{X}$, w_{ak} is the loading weight of the k^{th} variable at the a^{th} latent variable, $\|w_k\|^2$is the variance accounted from variable k, $\boldsymbol{t_a}$ and $\boldsymbol{w_a}$ are the a^{th} column vectors of $\boldsymbol{T}$ and $\boldsymbol{W}$ matrices, respectively, and $q_{\boldsymbol{ka}}$ is the element of matrix $\boldsymbol{Q}$ corresponding to the k^{th} variable and a^{th} latent variable.

The typical threshold to select the relevant variables according to the VIP values is usually set following the 'greater than one rule', which is a criterion based on the assumption that only variables with VIP score greater than 1 should be regarded as giving a significant contribution to the prediction of the y variable.

2.4.1 Model validation

A key step in the development of PLS models consists in the identification of the correct model dimensionality, i.e., in the determination of the adequate number of LVs. This step is necessary to find the best trade-off between including in the model as much as possible the useful information contained in the descriptor block while discarding noise.

As a matter of fact, a wrong choice in the number of LVs represents an error source in the model, and this can occur in two opposite ways, namely underfitting and overfitting. Underfitting occurs when the number of selected LVs is too small to properly model the y variable, while overfitting occurs when the number of LVs is excessive with a consequent inclusion of irrelevant experimental error in the model.

The correct model dimensionality is chosen using an internal validation procedure, referred to as cross-validation. Cross-validation is performed on the same set of samples used for model development, i.e., the training set, which are iteratively excluded for model calculation and used for internal prediction purposes. More in detail, cross-validation consists in splitting the training set samples in a user-defined number of subsets, named deletion groups. Then, PLS models are iteratively calculated by excluding one deletion group at a time and considering an increasing number of LVs,

from one until a maximum number of LVs (LV_{MAX}). For each iteration, the y values of the left-out objects are calculated and the difference between the predicted $\hat{y}$ and the corresponding measured value can be obtained.

For a given number a of LVs included in the model (with $1 \leq a \leq LV_{MAX}$), the sum of the squared values of the differences between the predicted ($\hat{y}_i$) and the corresponding experimentally measured value (y_i) can be computed on all the n samples (Eq. 2.12). This value is the Prediction Error Sum of Squares (PRESS) and it is an index of the predictive ability of the model.

$$PRESS_a = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \quad (2.12)$$

The optimal model dimensionality corresponds to the number of LVs returning the minimum PRESS value in cross-validation. However, from a practical point of view the choice of the proper model dimensionality should consider the best trade-off between the PRESS value and the parsimony of the model. In other words, if within a given interval of LVs the values of PRESS are very similar to each other, the model with the lowest number of LVs is preferable with respect to the model corresponding to the absolute minimum PRESS value but with higher LVs number.

Different cross-validation approaches are available, depending on the criterion based on which the deletion groups are determined. Leave one out (which consists in excluding one object at a time), Venetian Blinds (which consists in excluding in an alternating way) and Random groups (which consists in random attribution of the objects to the groups) are among the most common.

The internal validation of the model certainly represents an effective method to estimate the optimal model dimensionality but, truly, it is based on samples belonging to the training set. Indeed, the actual predictive capability of the final PLS model still has to be verified using samples with known y values and not considered in any step of the calculation of the model itself, i.e., not belonging to the training set.

Therefore, the final validation of a PLS model is performed through external validation, which consists in considering an external set of samples, namely test set. Usually, for a given dataset, objects are assigned to training and test set in a proportion of about two-thirds and one-third of the total samples number, respectively.

The error of PLS models can be quantified by means of the Root Mean Square Error (RMSE), which is strictly related to PRESS. RMSE can be referred to the values calculated by the model for the training set objects, to the prediction of the values estimated by cross-validation for training set objects, or to the prediction of the values for the external test set, and it is defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} = \sqrt{\frac{PRESS}{n}} \quad (2.\ 13)$$

where $\hat{y}$ are the predicted response values, y are the measured response values, and n is the number of samples. Three different RMSE values can be obtained depending on how $\hat{y}$ is obtained:

- RMSEC (Root Mean Square Error in Calibration) is the error in calibration of the training set, where $\hat{y}$ refers to the values calculated by the model on the training set samples, and n is equal to the number of training set samples;
- RMSECV (Root Mean Square Error in Cross-validation) is the error in cross-validation, where $\hat{y}$ refers to the values predicted in the cross-validation procedure, and n is equal to the number of training set samples;
- RMSEP (Root Mean Square Error in Prediction) is the error in prediction of the external test set, where $\hat{y}$ refers to the values predicted for the test set samples through the external validation procedure, and n is equal to the number of test set samples.

Besides, the performances of the regression model can be evaluated considering the R^2 statistics, which expresses the percentage of y variance that is accounted by the regression model. For a given response variable y, R^2 is defined as:

$$R^2 = 1 - \frac{PRESS}{SSY} \quad (2.\ 14)$$

where SSY is the sum of squares of the original $\boldsymbol{y}$ variable.

Similarly to RMSE, also the R^2 values can be obtained in calibration (R^2_{Cal}), cross validation (R^2_{CV}) and test set prediction (R^2_{PRED}).

RMSE is expressed with the same measurement unit of the dependent variable $\boldsymbol{y}$ and it allows a quantitative estimate of the error performed by the model, while R^2 statistic is independent from the scale unit of the $\boldsymbol{y}$ variable, enabling to evaluate in absolute terms the quality of the calibration model.

Finally, from a graphical point of view, an immediate and overall evaluation of the model performances can be achieved by plotting the y predicted values against the corresponding measured values. The closer are the samples to the bisector, the more satisfactory the model performances.

2.5 Data fusion

For a comprehensive characterisation of complex samples such as food matrices, it can be beneficial to combine different pieces of information obtained by acquiring the same samples with different analytical techniques (e.g., HPLC and UV-Vis spectroscopy) or different sensor technologies (e.g., electronic noses, tongues and eyes).

In this frame, the different data blocks can be jointly analysed through data fusion techniques to obtain an improved sample characterisation with respect to the use of single blocks (Calvini and Pigani, 2022).

The same principle can be extended to image analysis, where different feature extraction methods can be applied to extract different samples properties, such as colour or texture, from the same sample images. Then these features can be jointly analysed to simultaneously consider different visual properties of the analysed samples.

Data fusion can be carried out at three levels (Roussel et al., 2003):

- low-level (or concatenated data fusion);
- mid-level (or feature level data fusion);
- high level.

2.5.1 Low-level data fusion

In low-level data fusion, the different data blocks are simply concatenated row-wise, obtaining a unique data matrix with a number of rows equal to the number of samples, and a number of columns equal to the sum of the variables from each block (Figure 2.4).

The final data array can be subsequently analysed to build multivariate classification or calibration models.

Low-level data fusion is the simplest way to combine the information coming from different sources and it allows not only to have a direct interpretation of the contribution of the original variables but also to evaluate the existing correlations between variables of different blocks. However, it is more conservative than mid- and high-level, so no prior removal of the noise in the data blocks is provided.

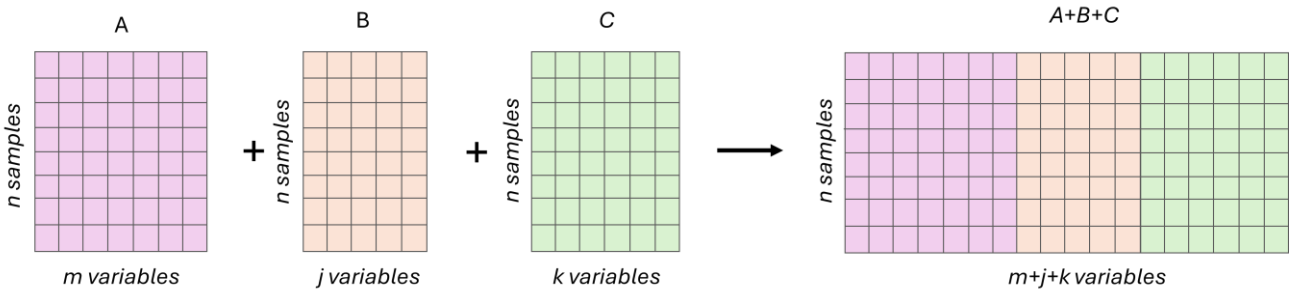


Figure 2.4 *Schematic representation of low-level data fusion.*

Hence, the use of low-level data fusion requires to pay particular attention to the preprocessing. Usually, it is performed following two consecutive steps. Firstly, each data block is separately pretreated based on its specific characteristics in order to minimize the contribution owed to noise or uninformative systematic variations (*see* Section 2.2). Then, scaling or weighting procedures have to be applied for proper combination of the different data blocks, i.e., to account for their different dimensionality. Indeed, when the data blocks are characterized by a different nature (e.g., with different measurements units) and/or by a notably different number of variables, the outcomes of the analysis would be clearly influenced by the block with the largest variance. To remove these effects, further preprocessing methods are required, and the most common one are listed below:

- autoscaling which gives the same relevance to all the variables (*see* Section 2.2);
- block-scaling, which consists in normalising each block to unit standard deviation, so it contributes with equal weight to the calculation of the model. In this manner, the same relevance is given to each source of information, i.e., each data block, while the relative weights of the variables within each block are preserved.

The choice of the proper preprocessing method depends on the nature of data and the specific issue.

2.5.2 Mid-level data fusion

When performing mid-level data fusion, a preliminary multivariate data analysis is applied separately to each data block to extract or select relevant features, which are then joined together to obtain the fused dataset. Then, the mid-level fused dataset can be analysed with multivariate statistical methods to provide the final classification or calibration output (Figure 2.5).

The features of interest can be identified from original data blocks by means of two main approaches:

- variable selection, which consists in applying proper variable selection algorithms (like e.g., sparse methods or interval PLS among others) separately to each one of the original data blocks in order to automatically select the most relevant variables for the problem under investigation. As a consequence, uninformative variables are discarded. Then, the variables selected from each block are concatenated to obtain the mid-level fused dataset;
- feature extraction, which is based on the concatenation of latent variables (i.e., score vectors) previously extracted through data decomposition of each single block through unsupervised or supervised methods like PCA or PLS, respectively. Latent Variables (LVs) account for underlying variable correlations and discard noise. Thus, it is generally possible to work with a low number of LVs without losing relevant information.

Therefore, the main goal of mid-level data fusion is to improve prediction models by considering only the relevant features extracted from each data block.

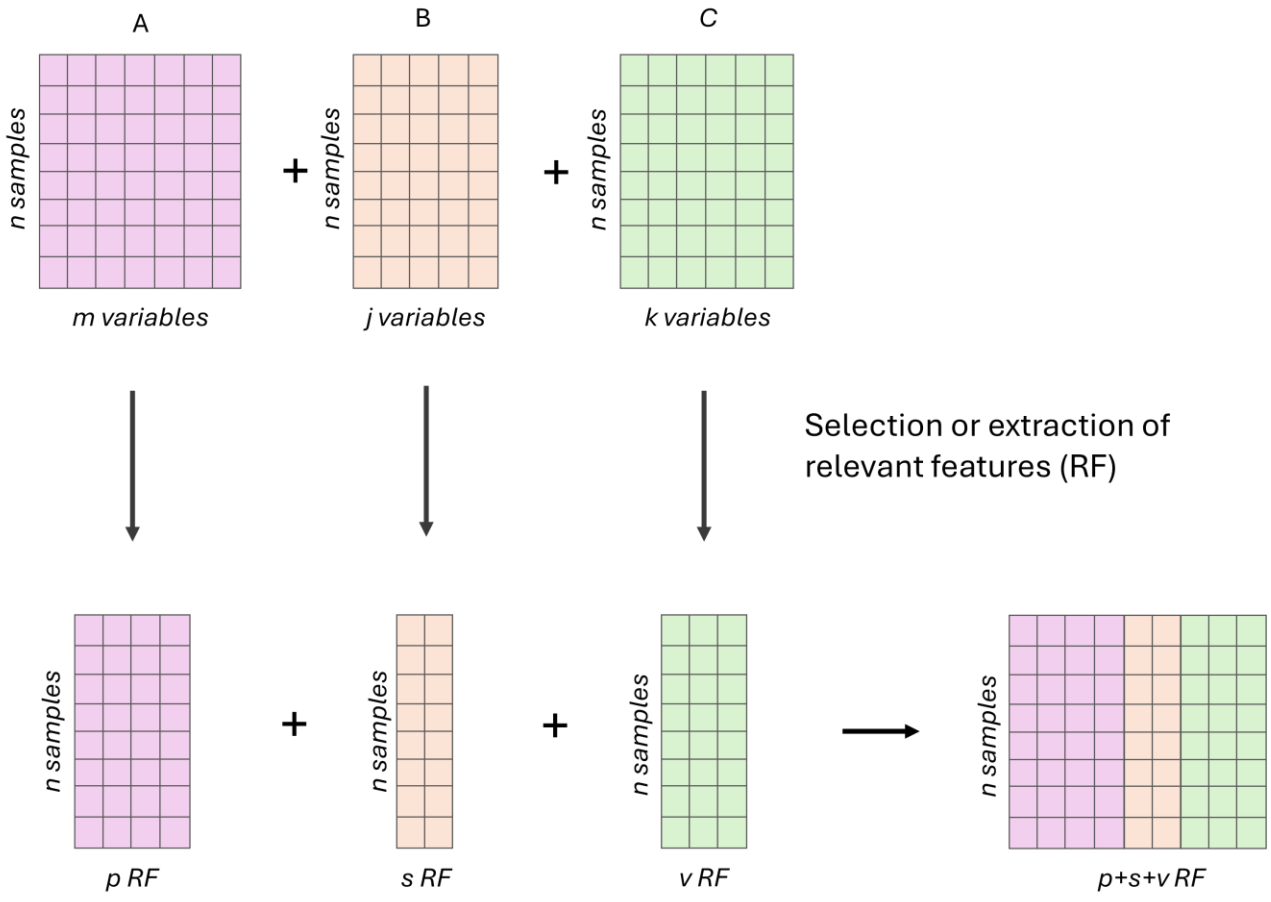


Figure 2.5 *Schematic representation of mid-level data fusion.*

As previously described for low-level data fusion, also for mid-level data fusion it is necessary to preprocess the data blocks using autoscaling or block scaling prior to the calculation of the models on fused data.

2.5.3 High-level

In high-level data fusion the information coming from the different data blocks is combined at the decision level, i.e., each block of data is independently analysed, and the outputs of the different models are then joined and elaborated to produce a final response (Figure 2.6). Multiple strategies are available to combine the predictions resulting from the single models, like majority voting, or methods based on Bayesian statistics or evidence theory (Calvini and Pigani, 2022; Kuncheva, 2004; Roussel et al., 2003; Shafer, 1976).

This level of data fusion has been found quite interesting in many scientific sectors, like sensory analysis and for classification purposes. Notwithstanding the improvement of the results that can be achieved through this level of data fusion, it is not frequently applied in chemometrics, since it implies the loss of the initial information contained in the original datasets and the consequent difficult interpretation of the variables having a major impact to solve the problem at hand.

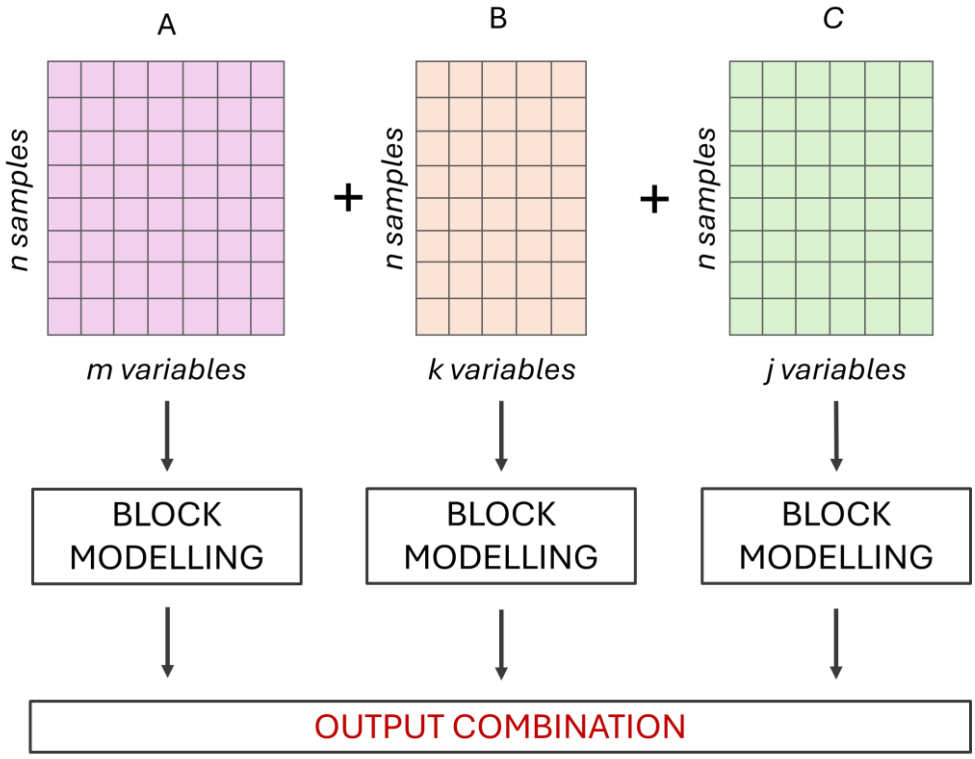


Figure 2.6 *Schematic representation of high-level data fusion.*

2.6 References

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. Appl Spectrosc 43, 772–777. https://doi.org/10.1366/0003702894202201

Calvini, R., Pigani, L., 2022. Toward the Development of Combined Artificial Sensing Systems for Food Quality Evaluation: A Review on the Application of Data Fusion of Electronic Noses, Electronic Tongues and Electronic Eyes. Sensors 22, 577. https://doi.org/10.3390/s22020577

de Jong, S., 1993. SIMPLS: An alternative approach to partial least squares regression. Chemometrics and Intelligent Laboratory Systems 18, 251–263. https://doi.org/10.1016/0169-7439(93)85002-X

Engel, J., Gerretzen, J., Szymańska, E., Jansen, J.J., Downey, G., Blanchet, L., Buydens, L.M.C., 2013. Breaking with trends in pre-processing? TrAC Trends in Analytical Chemistry 50, 96–106. https://doi.org/10.1016/j.trac.2013.04.015

Frank, I.E., Kowalski, B.R., 1982. Chemometrics. Analytical Chemistry 54, 232–243.

Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Analytica Chimica Acta 185, 1–17. https://doi.org/10.1016/0003-2670(86)80028-9

Gosselin, R., Rodrigue, D., Duchesne, C., 2010. A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications. Chemometrics and Intelligent Laboratory Systems 100, 12–21. https://doi.org/10.1016/j.chemolab.2009.09.005

Kuncheva, L.I., 2004. Combining pattern classifiers: methods and algorithms. J. Wiley, Hoboken, NJ.

Martens, H., Naes, T., 1989. Multivariate calibration, Wiley, New York, 1989., Ed. Wiley&Sons Ltd. ed. Chichester.

Oliveri, P., Malegori, C., Simonetti, R., Casale, M., 2019. The impact of signal pre-processing on the final interpretation of analytical outcomes – A tutorial. Analytica Chimica Acta 1058, 9–17. https://doi.org/10.1016/j.aca.2018.10.055

Roussel, S., Bellon-Maurel, V., Roger, J.-M., Grenier, P., 2003. Fusion of aroma, FT-IR and UV sensor data based on the Bayesian inference. Application to the discrimination of white grape varieties. Chemometrics and Intelligent Laboratory Systems 65, 209–219. https://doi.org/10.1016/S0169-7439(02)00111-9

Savitzky, Abraham., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. 36, 1627–1639. https://doi.org/10.1021/ac60214a047

Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press.

Wold, S., Esbensen, K., Geladi, P., 1987. Principal Component Analysis. Chemom. Intell. Lab. Syst. 2, 37–52.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58, 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Chapter Three

RGB imaging for food analysis

3.1 Relevance of colour in food analysis

Food quality is certainly a complex and multidimensional concept, embracing both subjective (e.g., sensory properties) and objective (e.g., chemical, physical and nutritional composition, presence and level of microbiological and toxicological contamination) attributes (Lawless, 1995; Molnár, 1995).

Either way, food quality evaluation relies on multiple sample properties, among which colour plays a major role. The visual aspect of food has a decisive influence on the consumer choices (Pereira et al., 2009). Indeed, it allows to evaluate aspects like colour, texture, shape and absence of defects.

Truly, the visual aspect of food is often the first and unique evidence in the consumers' possession leading their own purchasing decisions about food quality. For example, the evaluation of other sensory properties like taste or flavour is often unfeasible/hindered due to the packaging used to protect and preserve a vast majority of products marketed in food stores or supermarkets.

Hence, food colour and its variability has a leading role to evaluate most of the food products and are closely associated with quality aspects like degree of ripening, desirability, food safety and freshness (McCaig, 2002). More in detail, colour properties are often and strongly associated to the presence and amount of chemical compounds, the latter in their turn possibly relatable to pH effects, oxidation effects, quantification of pigments, defect detection, food preservation and many other quality aspects. Regarding pigments, anthocyanins are generally responsible of the red and blue-purple colour of different fruits and vegetables, but their appearance is influenced by the specific pH. Chlorophylls, responsible of the green colour of most vegetables, are often regarded as an index of vegetable freshness and of their ripening stage, since they are quite unstable (for example during cooking or blanching of vegetables) and a modification of their structure can easily result in different green tones. Concerning meat products, myoglobin is a protein found in muscle tissue and is the main responsible of meat colour. Indeed, cooking processes or variation in oxygen concentration can modify the myoglobin structure and determine a different colour of most meat products.

3.1.1 Colour perception

Colour can be defined as the property possessed by an object that produces different sensations in the eye because of the way it reflects or emits light.

Human eye can perceive a region of the electromagnetic spectrum, intuitively referred to as “visible”, ranging from about 380 nm to about 780 nm. As shown in Figure 3.1, shorter wavelengths correspond to blue-violet colour, while longer wavelengths correspond to red colour.

It follows that the perception of colour basically relies on the interaction between the visible spectrum of light emitted or reflected by an object and the human eye.

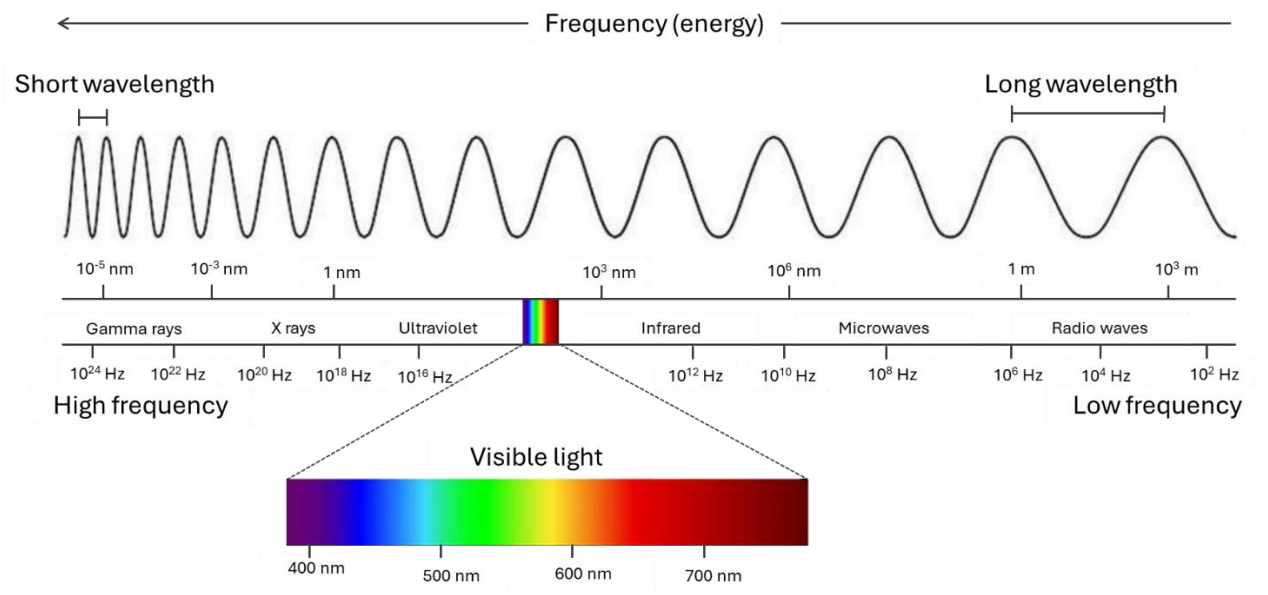


Figure 3.1 *Electromagnetic spectrum.*

More in detail, each monochromatic radiation in the visible spectrum is perceived as a single colour by the human eye. For example, a radiation at λ=500 nm is perceived as green and a radiation at λ=420 nm is perceived as violet. Interestingly, the combination of radiations at different wavelengths results in the perception of other colours that are not present in the visible spectrum. This is the case, for example, of white, grey or magenta.

The retina is a fundamental component for vision. It is the light-sensitive tissue located at the back of the eye, and it contains the photoreceptor cells capable of capturing light (emitted or reflected by an object) and to transform it into a bioelectric signal which passes via the optic nerve to the brain, where the visual image is formed. Photoreceptors in the retina can be distinguished in two different categories:

- the rod cells (or rods), which are responsible for night vision being sensitive to low lighting conditions (scotopic vision). They contain only one type of light-sensitive pigment with a

maximum absorption peak at about 510 nm (green part of the visible spectrum) providing a monochromatic vision;

- the cone cells (or cones), which are responsible for colour vision under optimal lighting conditions (photopic vision). Cones in the human eye can be divided into three types, being sensitive to different wavelength ranges in the visible spectrum:
    - S-cones (short-wavelength sensitive cones), containing a light-sensitive pigment with a maximum absorption peak at about 430 nm;
    - M-cones (middle-wavelength sensitive cones), containing a light-sensitive pigment with a maximum absorption peak at about 530 nm;
    - L-cones (long-wavelength sensitive cones), containing a light-sensitive pigment with a maximum absorption peak at about 560 nm.

The bioelectric signal generated from the photoreceptors is then transmitted to the brain, undergoing a very complex elaboration processes which lead to colour perception.

3.1.2 RGB and HSI colour spaces

As stated above, human colour vision relies on the varying sensitivity peaks of cone cells in the retina to visible light corresponding to short (blue 420–440 nm), middle (green, 530–540 nm), and long (red, 560–580 nm) wavelengths, respectively (Hunt, 2005). Hence, the three colour components perceived by the human eye (generally referred to as tristimulus values, derived from the extent of the stimulation caused by radiation in each of the three types of cones) can be used to simply describe colour perception. A colour space is a mathematical representation associating tristimulus values to each colour. Based on the application, three different types of colour spaces have been proposed (Wu and Sun, 2013):

- *Hardware-oriented colour spaces* are used for data processing in machines (e.g., image acquisition, storage and visualisation). Their widespread use in the measurement of food colour is encouraged by their ability to detect even a tiny colour variation, which often occurs during food processing, ripening or storage. RGB (red, green and blue), YIQ (luminance, in-phase, quadrature) and CMYK (cyan, magenta, yellow, black) are among the most popular. However, the latter two are mainly used for television transmission and in printing and copying output, respectively;
- *Human-oriented colour spaces* as the name suggests, have been developed to better reflect human colour perception, describing the colour properties of a given object in terms of concepts such as tint, shade and tone, rather than the amounts of red, green, and blue. In

general, they are hue-saturation (HS)-based spaces, like HSI (hue, saturation, intensity), HSV (hue, saturation, value), HSL (hue, saturation, lightness), and HSB (hue, saturation, brightness). Even if HS-based spaces boast a notable relationship to the visual significance of food surfaces, they lack sensitivity to small colour variations, hindering their suitability to evaluate food colour changes over time;

- *Instrumental oriented colour spaces* are generally used by classical instruments for colour measurement (e.g., colorimeters) and have been largely standardized by the Commission Internationale d'Eclairage (CIE) under specific conditions (e.g., illuminants, observers, and methodology spectra) (Viscarra Rossel et al., 2006). In food-related applications, the L*a*b* colour space (also known as CIELAB) is the most widely used due to its uniform distribution of colours and perceptual uniformity (i.e., the Euclidean distance occurring between two different colours approximately matches the colour difference perceived by the human eye) (León et al., 2006).

In this PhD Thesis, image elaboration has been performed considering RGB and HSI colour spaces, as detailed below.

RGB is the most widespread hardware-oriented colour space, used in computer graphics and monitors. It is an additive colour system based on the trichromatic theory. Each colour is reproduced and defined by the combination of different intensity levels of the three additive primary colours: Red (R), Green (G) and Blue (B). As shown in Figure 3.2, the RGB colour space can be graphically represented as a cube, whose coordinates are defined on the three R, G and B axes. The vertices of the cube represent the three additive primary colours (red, green and blue), the three subtractive primary colours (cyan, magenta and yellow), the sum of the additive primary colours (white) and the absence of colour (black). The diagonal of the cube connecting the black and white vertices represents the different grey levels (greyscale). Each point of the cube corresponds to a colour achieved by combining red, green and blue colour intensities in specific proportions.

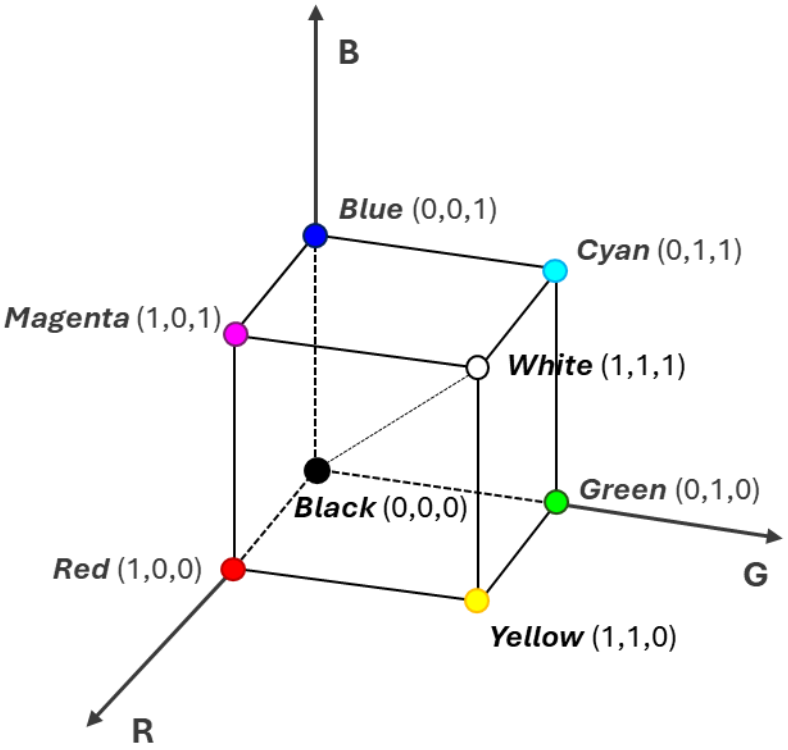


Figure 3.2 *RGB colour space.*

Typically, R, G, and B intensity values are expressed as integers ranging from 0 to 255 rather than a normalized scale ranging from 0 to 1. Since this colour space is commonly used in informatics, a digital RGB image is usually represented in binary form, with 1 byte (equal to 8 bit) allocated for each R, G and B component. In other words, each component is coded by 8 bits, requiring 256 ($= 2^8$) integer values ranging from 0 to 255. This system allows to represent 16.7 million colours ($= 256 \times 256 \times 256$), a variety comparable or even exceeding the sensitivity of the human eye, able to distinguish up to 10 million colours.

While the RGB colour space is particularly suitable for handling digital images and mimics the physiological functioning of the human retina, the HSI colour space offers a better description of human perception of colours. This colour space defines the colours in terms of:

- *Hue (H)*: The attribute of the human visual sensation that defines the colour similarity to one of the perceived colours in the visible spectrum (e.g., red, yellow, green, blue, etc…) or a combination of two. It depends only on the dominant wavelength and is independent from intensity or lightness;

- *Saturation (S)*: The attribute of human visual sensation that defines the perceived "purity" of a colour relative to its brightness (Calvini and Pigani, 2022; Wu and Sun, 2013). For example, a low saturation colour has a high amount of grey mixed with the hue;
- *Intensity (I)*: The attribute of human visual sensation describing the amount of light present in a colour.

The graphical representation of the HSI space is shown in Figure 3.3, by means of a double cone:

- Hue is described by the angle measured from a reference point (usually, red is associated to an angle of 0° and hue increases counterclockwise starting from the red point);
- Saturation corresponds to the distance from the vertical axis;
- Intensity corresponds to the vertical axis, connecting the black and white vertices.

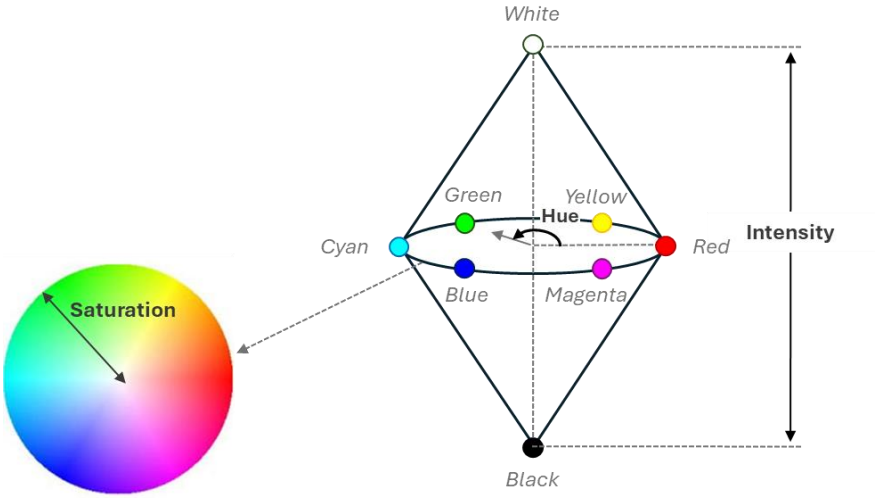


***Figure 3.3** HSI colour space.*

There is a biunivocal correspondence between RGB and HSI colour spaces. Indeed, the mathematical equations that allow to convert the RGB coordinates to the HSI ones are detailed in Section 3.4.1.

3.2 RGB imaging

Trained operators have been employed for a long time for food quality inspection, grading and control. Even if the conveyance of radiation by light can be regarded as on objective phenomenon, the same cannot be said concerning the human vision process. The intrinsic subjectivity of the

conversion and elaboration processes involved must be taken into account when evaluating food quality, since extensive training and experience are required to minimize the variability of human assessment in control procedures, resulting in laborious, time-consuming and expensive methods (Gunasekaran, 1996). For the same reasons, the transfer of knowledge about monitoring and assessment procedures between human operators (and similarly between laboratories, industries and different product chains) should not be regarded as an easy task either.

Within this framework, this issue could be overcome thanks to the availability of instrumentation able to provide an objective, fast, reproducible and cost-effective evaluation of food properties and intended to mimic the response of the human eye. The use of computer vision techniques has been rapidly spreading, offering the ability to achieve objective, comprehensive, quick, non-invasive and reproducible measurements of the visual aspect of food, with minimal human intervention. Furthermore, image-based systems are able to provide both spatial-related and colour-related information, enabling a detailed analysis of samples with heterogeneous appearances (like the vast majority of food products/matrices), contrary to traditional systems employed for colour measurement, like tristimulus spot colorimeters (Brosnan and Sun, 2004, 2002; V. G. Narendra and Hareesh, 2010; Pedreschi et al., 2006). Modern cameras can acquire images with tens of millions of pixels, allowing the description of both local information accounted for by few pixels (e.g., very small details) and the global information contained in the whole image scene.

RGB imaging systems have been extensively applied to evaluate attributes of several food matrices other than colour, like texture, size, shape and presence of surface defects. Some examples of applications of RGB imaging are listed in Table 3.1, in order to illustrate the wide range of food related issues and food products that can be investigated and addressed by means of this technique. As shown in Table 3.1, a selection of applications involving the use of common smartphones is also reported.

Table 3.1 *Selection of RGB imaging applications for food quality evaluation.*

Matrix	**Main goal**	**Device**	**Reference**
Chicken meat	Defect detection	CCD camera	(Barni et al., 1997)
Muffins	Separate dark from light samples	CCD camera	(Abdullah et al., 2000)
Biscuits	Estimate physical features of chocolate chip cookies	CCD camera	(Davidson et al., 2001)
Potato chips	Colour grading	Digital camera	(Pedreschi et al., 2006)
Pomegranate arils	Sorting of good pomegranate arils	Scan cameras	(Blasco et al., 2009)
Raw hams	Detection of red skin defects	Digital camera	(Ulrici et al., 2012)
Potatoes	Detection of irregular potatoes	CCD camera	(ElMasry et al., 2012)
Citrus	Estimate the standard colour index	Smart camera	(Vidal et al., 2013)
Pigmented carrots	Predict antioxidant activity and total phenols content	3CCD digital camera	(Pace et al., 2013)
Meat	Measurement of colour	Digital camera with CMOS sensor	(Girolami et al., 2013)
Potato chips	Identification of acrylamide	Digital camera	(Dutta et al., 2015)
Grape	Assess bunch compactness	Digital camera	(Cubero et al., 2015)
Pork meat	Defect detection	Digital camera	(Chmiel and Słowiński, 2016)
Hazelnut	Identification of defective hazelnuts	Digital camera	(Giraudo et al., 2018)
Olive	Detection of fly stings in infected olive fruit samples	CCD camera	(Beyaz et al., 2019)
Strawberry yoghurt purée	Evaluation of strawberry yoghurt puree colour changes	Flatbed scanner	(Rolando et al., 2023)
Grapevine	Count of the flower number in grapevine inflorescence	Smartphone + App (for image acquisition, analysis and results display)	(Aquino et al., 2015)
Cured meat	Estimation of fat content	Smartphone camera	(Cruz-Fernández et al., 2017)
Beef	Prediction of fresh beef tenderness	Smartphone + App (for image acquisition, analysis and results display)	(Hosseinpour et al., 2019)
Bitter orange peel	Determine chlorophyll and carotenoid contents	Smartphone camera	(Taghadomi-Saberi et al., 2019)
Freshness sensors	Spoilage detection of packaged meat	Smartphone + App (for image analysis)	(Perez De Vargas-Sansalvador et al., 2020)
Starfruit	Classification of ripening stage	Smartphone camera	(Saha et al., 2023)
Fish	Fish quality grading	Smartphone cameras	(Jayasundara et al., 2023)

3.2.1 RGB image acquisition

Computer vision systems for image acquisition contain a sensor array (usually the Charge Coupled Device, CCD, or the Complementary Metal–Oxide Semiconductor, CMOS), which consists of an array of photosensitive elements able to capture and record for each picture element (i.e., each pixel) the intensity values of the visible region of the electromagnetic spectrum in the red (R, at $\lambda \approx 630$ nm) green (G, at $\lambda \approx 545$ nm) and blue (B, at $\lambda \approx 435$ nm) channels. These sensors are designed to mimic the human colour perception (Brosnan and Sun, 2004; Ornberg et al., 1999; Wu and Sun, 2013).

As a result, a digital RGB image consists of a three-dimensional data array, recording the intensity values in the R, G and B channels for each element of the two-dimensional pixel grid.

Usually, a computer vision system contains the following components:

- an illumination system, which is the light source illuminating the sample under investigation and whose performance and characteristics significantly impact image quality and can affect the subsequent image processing steps. Indeed, lighting type, light source location and spectral power distribution require particular attention to ensure an accurate image of the object (i.e., of the actual sample colour) and avoid the presence of shadows or specular reflections (Bachelor, 1985; Gunasekaran, 1996);
- an RGB camera for image acquisition incorporating a CCD or a CMOS sensor. The minute photosensitive elements of both arrays are able to convert the intensity of incident light into an electric signal;
- sample holder or conveyor belt (in industrial setups);
- hardware and software for image processing and disk storage (Brosnan and Sun, 2004; V. Narendra and Hareesh, 2010).

The main components of an RGB image acquisition system are schematically depicted in Figure 3.4.

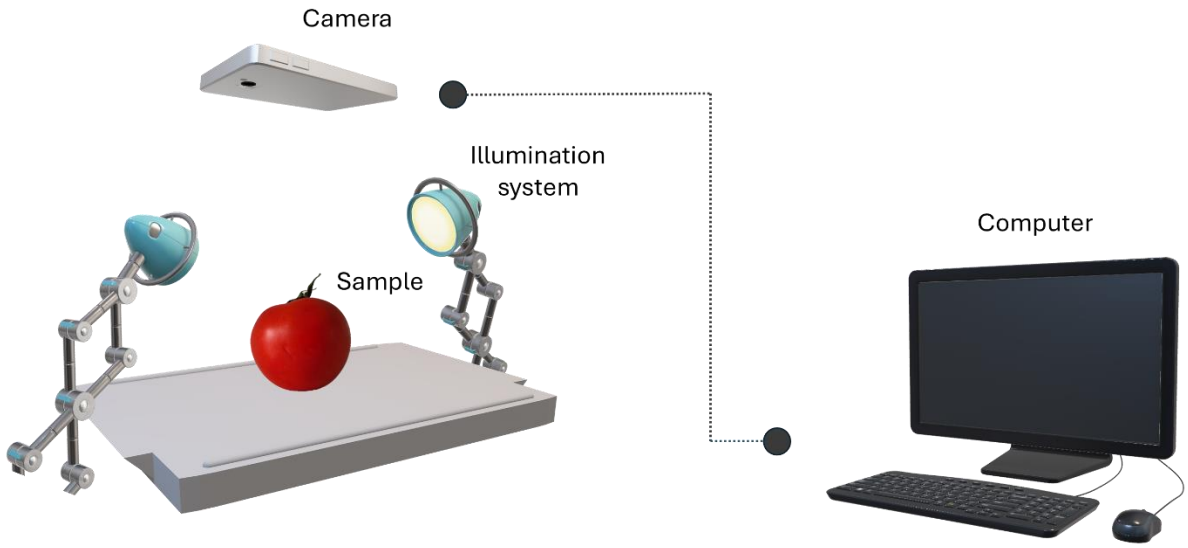


Figure 3.4 *Schematic representation of an RGB imaging system.*

With regard to the type of sensors that can be implemented in RGB cameras, CMOS sensors boast a high integration (i.e., one chip with many functions), low power dissipation and the possibility to be included in small acquisition systems. Indeed, they are commonly used in smaller consumer cameras or mobile phones cameras. CCD sensors, on the other side, are typically employed in photography as well as in scientific and industrial applications requiring the highest possible image quality, at the expense of bulkier devices.

No matter of the specific sensor type, two main configurations of colour cameras are commonly used:

- Three-chips cameras, which are based on combinations of prisms or filters coupled with three sensors, one for each channel (R, G, B). The resulting image is obtained by merging, for each pixel, the R, G and B values measured separately by the three sensors in correspondence to the same pixel;
- Bayer-based cameras which contain one sensor covered by a Bayer filter, which consists in a mosaic of red, green and blue filters (Marszalec and Pietikäinen, 1996). More in detail, Bayer filters are intended to mimic the physiology of the human eye, which is more sensitive to green than to red or blue. They consist of 50% green elements, 25% red elements and 25% blue elements. Thus, for each pixel only one channel is directly measured, while the values of the other two channels are interpolated according to the values measured in the surrounding pixels. Despite the wide variety of colour filter arrays, the Bayer filter is by far the most used on consumer digital cameras.

In order to obtain a digital RGB image, a frame-grabber is required to elaborate and convert the electric signal generated from the sensor into Picture Elements (pixels for short), the latter digitized to achieve the final RGB image. Hence, RGB images are three-dimensional data arrays with size $\{r, c, 3\}$, where r is the number of pixel rows, c the number of pixel columns, and 3 corresponds to the R, G and B channels. Each pixel is defined by integer values ranging from 0 to 255, representing the intensity values in the R, G and B channels.

The quality of digital images depends on three fundamental properties:

- image resolution, which refers to the number of pixels contained in the image, is generally measured in pixels per inch (ppi), as the product of the number of column pixels by the number of row pixels (e.g., 8000 × 6000 pixel), or as the resulting product in million pixels units, i.e., in megapixel, MP (e.g., 48 MP). Hence, higher resolutions enable the detection of smaller objects with distinct boundaries;
- image size, which refers to the width and height of the image and it is measured in centimetres or inches;
- image file size, which measures how much data are contained in the image file, expressed in kilobytes (KB) or megabytes (MB), depending on the overall number of image pixels, the image file format and the image data compression level.

Different types of computer vision systems are available and can be mentioned, like webcams, flatbed scanners or cameras. However, given the recent developments in image acquisition technology the most interesting and compelling tools in image acquisition are nowadays clearly represented by smartphones, which are very sophisticated, portable and cost-effective everyday tools. Even if the cost of smartphones is relatively low, they have high-resolution cameras, and their processing capabilities are now enhanced thanks to advanced processors and memory storage devices. In addition, the use of a smartphone camera enables to integrate the response returned by the computer vision system with proper and accurate elaboration steps, thanks to a profitable implementation in user-friendly smartphone apps, designed *ad hoc* for the intended data processing. Besides, the handy size of a smartphone allows an advantageous combination with portable instruments.

In this PhD Thesis, image acquisition has been performed by means of an innovative portable device conceived to acquire RGB images under controlled lighting conditions and using a common smartphone. Furthermore, the whole workflow, from image acquisition to data elaboration and visualization of the results, has been implemented in a user-friendly software interface. This includes a smartphone app allowing to acquire geolocated RGB images, visualize multiple outcomes of the analysis, store data and share relevant information (*see* Section 5.1)

3.2.2 Image standardization

The use of RGB digital imaging as an analytical measurement tool in food-related applications requires an objective, accurate and reliable quantification of the colour properties of food samples. Indeed, the accuracy and robustness of any calibration or classification model closely rely on the quality of data provided as input.

In this scenario, computer vision systems should be designed to provide controlled and constant lighting conditions during image acquisition. Nevertheless, it is difficult to fully satisfy this condition, not only in the field but even in the laboratory (Menesatti et al., 2012).

Several factors can undesirably affect the RGB colour values, resulting in noisy or even misleading variations. For example, it is truly difficult to completely avoid or have an absolute control of phenomena like instrumental instability which could occur over time (possibly due to a drift of the detector or camera settings accidentally modified) or variations of lighting conditions.

For these reasons, image standardization is highly recommendable as a first pre-processing step of image-based data to reduce variations whose occurrence and extent cannot be completely prevented and controlled *a priori*.

In this regard, different procedures for colour calibration have been proposed over time based on different polynomial algorithms, multivariate statistics, and neural networks approaches (Chang and Reid, 1996; Van Poucke et al., 2010). The most appropriate standardization procedure can be chosen according to the issue at hand.

Section 5.1 of this PhD Thesis reports the development of a portable smartphone-based device to estimate anthocyanins content of red grapes. Even if the acquisition box was designed to grant the most possible standardized lighting conditions during image acquisition, image standardisation using calibration correction method was also implemented.

Similarly to other image correction methods, the calibration method relies on standard colour references (SCR) included in the acquired image scene. SCR typically consist in a white paper sheet including a given number of patches of different colours. In this PhD Thesis image correction has been performed by means of SCR composed of a series of rectangles of different colours: the primary additive colours (red, green and blue), as well as white, black and grey tones, as shown in Figure 3.5. An example of a SCR cropped from a master and a slave image, respectively, is reported in Figure 3.5.

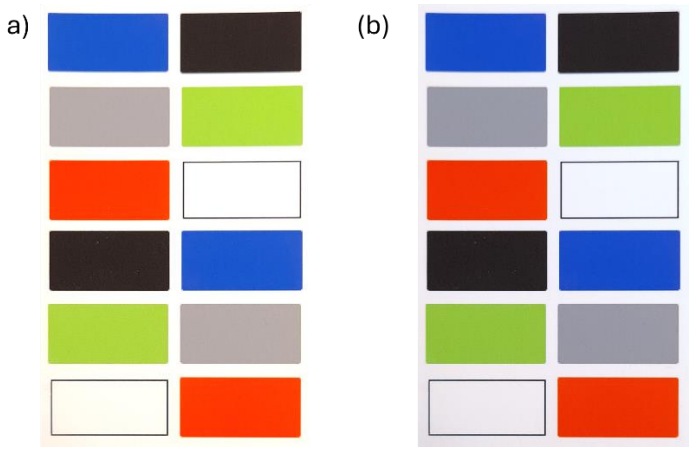


Figure 3.5 *Example of the Standard Colour Reference of a (a) Master image and (b) slave image.*

At first, the SCR of an image referred to as master (usually the first one acquired), is chosen as a reference. The correction of the remainder images, referred to as slave images, is performed by comparing the RGB values of the SCR of the master image with the RGB values of the SCR of each slave image. The correction procedure aims at making the SCR of each slave image as much similar as possible to the SCR of the master image using a proper mathematical function, which in turn is applied to correct the whole slave images.

The calibration procedure used in this Thesis was initially developed for processing hyperspectral images (Ulrici et al., 2013) and lately extended to RGB images. Essentially it has been conceived to adjust the possible variations of the colour dynamic ranges by computing appropriate and case specific regression models. Indeed, this approach can be regarded as a flexible correction tool, allowing to handle linear and non-linear variations in the colour dynamic range.

More in detail, the m colour patches in the SCR are selected and each of them is divided in g pixel groups according to a user-defined grid, in order to consider possible spatial variability within the same colour patch. Then, for each pixel group, the corresponding median values of each R, G and B channel are computed and recorded in a median matrix with size $\{(g$ median values $\times m$ colour patches), 3 channels$\}$.

The calibration step consists in calculating, separately for each single channel c, a regression model between the median vector of the master image, $\boldsymbol{y}(\boldsymbol{c})$, and the median vector of each slave image i, $\boldsymbol{x}_i(c)$, as described in the following equation:

$$\boldsymbol{y}(c) = b_{0,i}(c) + b_{1,i}(c)\boldsymbol{x}_i(c) + b_{2,i}(c)\boldsymbol{x}_i^2(c) + \cdots + b_{n,i}(c)\boldsymbol{x}_i^n \qquad (3.1)$$

where n is the degree of the polynomial and b_0, b_1, b_2 and b_n are the regression coefficients calculated for the channel c of the i^{th} slave image.

The calculation of the regression model also includes the evaluation of the statistical significance of each regression coefficient according to the 95% confidence limit. Non-significant regression coefficients are discarded by means of backward elimination procedure. Clearly, the correction procedure is applied to a slave image only if the algorithm detects significant differences with respect to the SCR of the master image.

Once computed the proper regression model, the sample area of the image scene containing the slave SCR is corrected, channel-by-channel, according to the following equation:

$$\boldsymbol{CS}_i(c) = b_0(i,c) + b_{1,i}(c)\boldsymbol{S}_i(c) + b_{2,i}(c)[\boldsymbol{S}_i(c)]^2 + \cdots + b_{n,i}(c)[\boldsymbol{S}_i(c)]^n \qquad (3.2)$$

where $\boldsymbol{CS}_i(c)$ is the channel c of the i^{th} standardized sample image, $\boldsymbol{S}_i(c)$ is the corresponding channel of the original sample image, and the b terms are the statistically significant regression coefficients calculated starting from Equation 3.1.

Usually, the optimal value of the degree of the polynomial, *n*, is chosen based on some preliminary tests performed on a representative subset of sample images.

A schematic overview of the calibration correction procedure is shown in Figure 3.6.

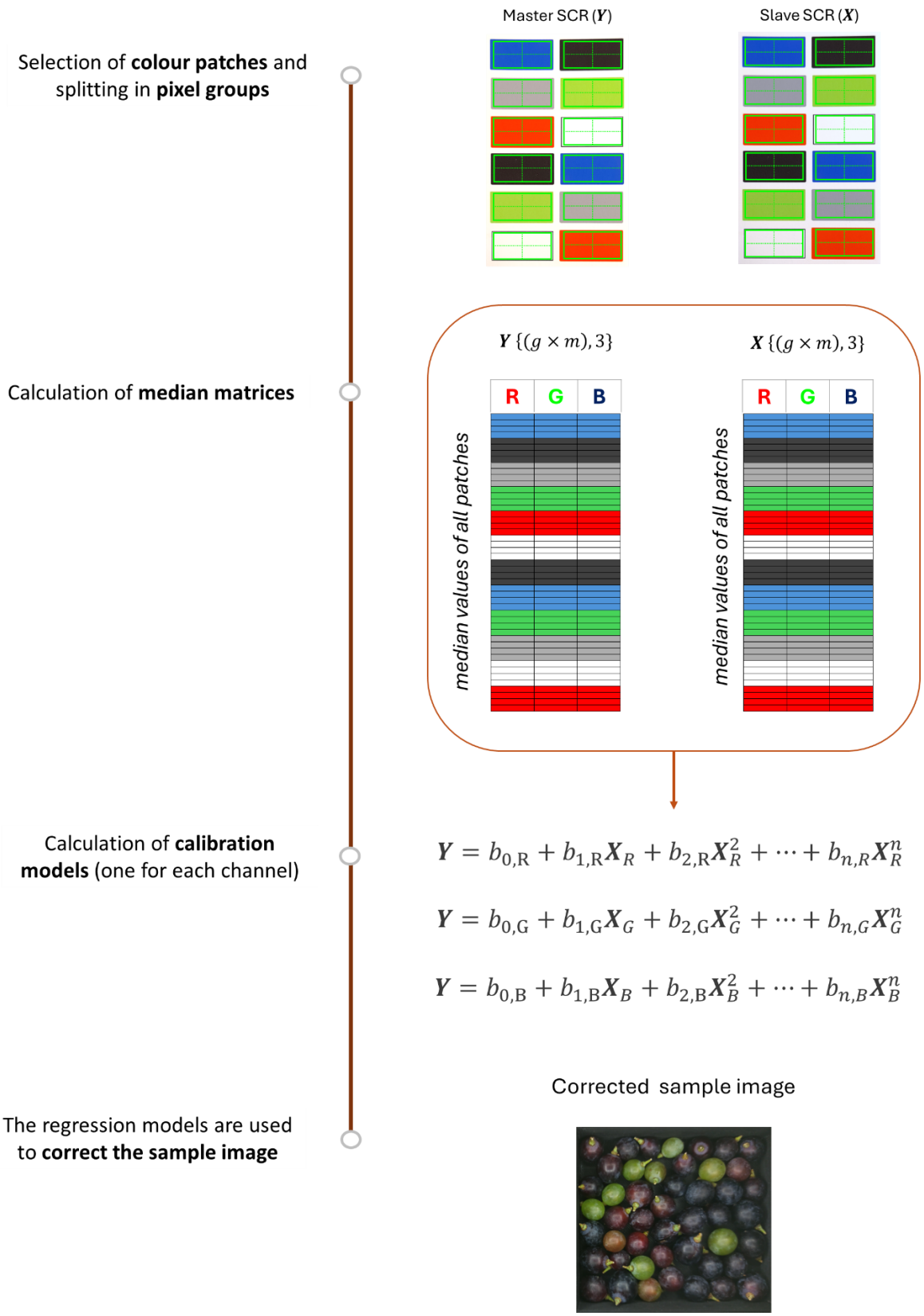


Figure 3.6 *Scheme of the main steps of image correction by means of calibration method.*

3.3 Multivariate Image Analysis

The application of chemometric strategies is crucial to properly extract and process the useful information contained in digital image data. This becomes even more important when dealing with large datasets of images, as often occurs in real-world case scenarios, due to the huge amount of data that needs to be analysed.

Multivariate Image Analysis (MIA) can be regarded as the development and application of chemometric methods (e.g., PCA or PLS) to analyse the data contained in images whose pixels are described by more than one variable (i.e., multivariate images), like RGB images (Esbensen and Geladi, 1989; Prats-Montalbán et al., 2011a).

MIA can be performed based on different types of image representation: pixel-level, object-level or image-level. By the way, independently of the level at which MIA is performed, the use of chemometric tools allows to conveniently extract relevant information and achieve the different goals to be faced in food-related applications, like image segmentation, exploration, classification or calibration, as well as the identification and quantification of defects and the application of data dimensionality reduction strategies to facilitate the simultaneous analysis of large datasets of images.

The first research studies related to MIA considered the information contained in single pixels as individual objects, leading to a pixel-level approach. According to this approach, each image pixel is considered as a sample/object on its own, described by a defined number of variables, which, for a typical RGB image, correspond to the R, G and B channels. Clearly, it is possible to consider different colour spaces other than RGB (like e.g., L*a*b*, HSI or CMYK) or even add further spectral channels (for example, those related to NIR images resulting in multispectral and hyperspectral images) (Prats-Montalbán et al., 2011b). Indeed, this type of image representation could remain a good option when the information of interest is found at the pixel-level.

Object-level MIA constitutes a somehow intermediate step between pixel-level and image-level, where the single objects contained within the image scene (e.g., single coffee beans/maize kernels/etc.) are separated from the background through a proper segmentation procedure, and then each object is characterized as a whole, based on the properties of the corresponding pixels.

However, many practical applications of imaging techniques require the acquisition of a considerable number of images that have to be simultaneously analysed. To face this issue, different MIA methods have been proposed to provide a global characterisation of image properties by extracting relevant metrics/features that represent the information of interest (Antonelli et al., 2004; Calvini et al., 2020; Prats-Montalbán et al., 2011b). This kind of approach is referred to as image-level analysis and it considers the whole image as a unique object to provide an appropriate overall image characterisation.

Starting from each image of the dataset, the descriptors of interest (e.g., colour or texture descriptors) are extracted and collected into a feature vector that will be used to characterise the corresponding image. The selection of the most suitable features for the issue at hand can be certainly regarded as a critical and crucial aspect, playing a key role on the final outcomes of image-level analysis (*see* Section 3.4). This kind of image representation stands out as a good solution when dealing with large datasets of images that are required to be simultaneously compared, i.e., when the final goal of the specific application implies a decision making based on the information provided by the whole image and a deep understanding of the whole dataset structure. Truly, in real case scenarios and applications, each image is often expected to represent the properties of a given sample and the possibility to get an immediate and overall estimate of a sample is preferable, focusing on the single pixels only when necessary (Calvini et al., 2016). For all these reasons, the present PhD Thesis has been focused on the development and application of MIA feature extraction methods at the image-level, as will be explained in the following section.

In general, prior to multivariate analysis, the three-dimensional RGB image cube is subjected to unfolding, which consists in converting the image data array with size $\{r, c, 3\}$ (where r corresponds to the number of row pixels, c corresponds to the number of column pixels and 3 corresponds to the R, G and B channels) into a data matrix with size $\{r \times c, 3\}$, i.e., with as many rows as the overall number of pixels and three columns.

Even if the spatial organisation of the image is lost during unfolding, it is always possible to visualise back into the original image domain the outcomes of the analysis. For example, in the case of PCA, the PCs score vectors can be refolded back in the corresponding score images. In this way, the information present in the image can be profitably explored and interpreted.

An example of a PCA carried out at the pixel-level can be illustrated using the RGB image depicted in Figure 3.7a, in which each pixel is characterized by the values of the three colour channels.

The whole set of individual image pixels can be graphically represented as shown in Figure 3.7b, by means of a space defined by three Cartesian axes corresponding to the original R, G and B channels, respectively.

The first principal component (PC1) accounts for the direction of maximum variance of the mean-centred data. The second principal component (PC2) accounts for the direction of the maximum residual (and smaller with respect to PC1) amount of variability of the cloud of data points and is orthogonal to PC1. Although it would theoretically be possible to consider a further third principal

component (PC3) orthogonal to the previous ones, it can be neglected when the third PC direction truly accounts for a small residual variance, which is likely due to noise.

In the specific case of an RGB image, PCA decomposed the mean-centred ***X*** matrix into:

- a score matrix (***T***), accounting for the variance associated with the pixels. It describes the position of each image pixel in the new PCs space;
- a loading matrix (***P***), accounting for the contribution/relevance of the original R, G and B channels in defining the PCs;
- a residual matrix (***E***) describing the residual variance (noise).

In line with these considerations, the score plot (Figure 3.7c) enables both a graphical representation of the position of the projections of all individual pixels in the PCs space (the colour coding is referred to the pixels density) and the possibility to investigate the objects structure, so as to identify clusters of similar pixels and/or outliers, which can significantly affect the results of the subsequent data analysis.

Interestingly, when dealing with images, the score vector of each PC can be refolded back into the original image domain, returning the corresponding score image. By this way, it is possible to simultaneously visualize the position of all the image pixels both in the score plot and in the original image domain and investigate the occurrence of possible interactions between the two representations with notable advantages in the image analysis. Indeed, it is possible to select a cluster of pixels in the score plot and visualize the corresponding location in the original image scene or, vice versa, select a specific image area and visualize the corresponding pixels position in the score plot. This technique is called *brushing* (Grahn and Geladi, 2007; Williams et al., 2010) and can be used in order to define specific image areas (often referred to as Regions of Interest, ROIs) corresponding to particular features.

Figure 3.8 shows an example of the brushing technique, applied to the RGB image of rolled wafers with chocolate topping. In fact, based on the pixel similarity in the PC1-PC2 score plot, it is possible to identify the pixels clusters belonging to the chocolate topping, and similarly those related to the background and to the wafer rolls.

Since the number of original variables is low, i.e., equal to three, the graphical representation of the loading coefficients is very easy by means of the loading plot in order to evaluate the contribution of each colour channel in the new PCs space.

In this way, PCA enables to investigate the relationships existing between pixels and original variables, i.e., to identify which original variables play a major influence in the groupings observed among the samples by means of a combined evaluation of both score and loading plots.

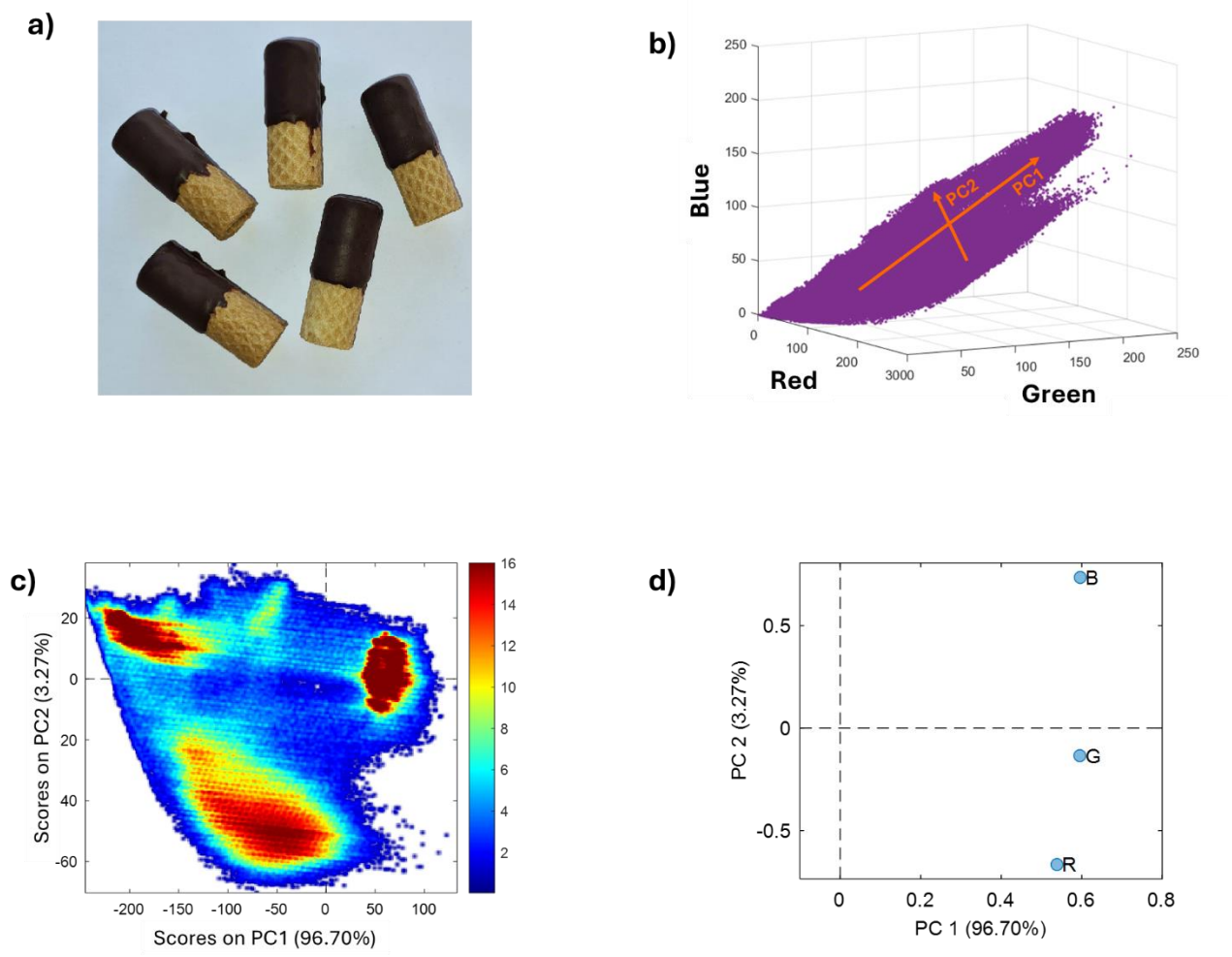


Figure 3.7 *Original RGB image (a), pixels of the image represented in the R, G and B space (b) and in the corresponding PC1-PC2 space (c), loading plot (d).*

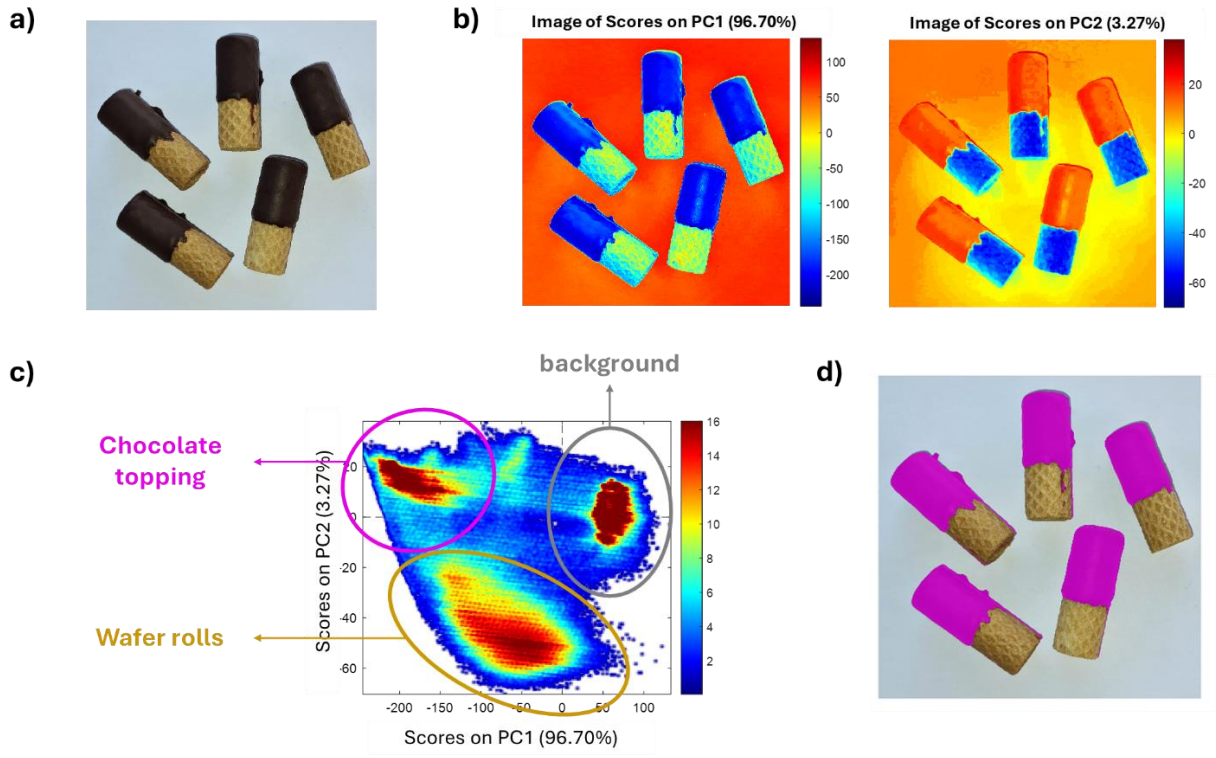


Figure 3.8 *An RGB image (a) is decomposed by PCA and the corresponding PC1 and PC2 score images are reported in (b). The selection on the score plot (c) of the cluster of pixels with low values of PC1 and high values of PC2 allows to identify and locate the pixels of the chocolate topping (d).*

3.4 Image-level extraction of colour and texture features

As discussed in Section 3.3, even if MIA was initially conceived to be performed at a pixel-level, the subsequent development of MIA techniques at the image-level can be regarded as addressing the practical needs of most food-related applications. In fact, in the vast majority of real case scenarios (e.g., food monitoring) it is required to simultaneously analyse a notable amount of images, each one often expected to enclose the overall properties of a specific sample. In this way, the possibility to encode the information of interest contained in an image into a feature vector stands out as an effective and suitable tool to address the challenge. Indeed, each feature vector can act like a "fingerprint" of the relevant properties of the corresponding image and the different feature vectors obtained from all the images of the dataset can be collected into a data matrix which in turn can be analysed with unsupervised or supervised chemometric techniques.

When dealing with RGB images, colour is the most immediate information type that can be derived from such data. In this PhD Thesis, the colourgrams approach (Antonelli et al., 2004; Calvini et al., 2020) was considered as a colour feature extraction method for the analysis of RGB images, and in particular this kind of approach was used in the development of a smartphone-based device to predict phenolic ripening in red grapes (*see* Section 5.1). Basically, the colourgrams method consists in converting each RGB image into a signal, the *colourgram*, obtained by merging in sequence the frequency distribution curves of RGB values and additional colour descriptors derived from them. A detailed description of the colourgrams method is provided in Section 3.4.1.

However, RGB images are data rich array allowing to account for several sample properties other than colour. Within this framework, one of the most relevant is certainly texture, being often relevant to detect, describe and distinguish objects in an image.

Even if there is not a universally accepted definition of texture, this property can be described for a given single-channel image as a function of the spatial variation of pixel intensities and involves the spatial distribution of intensity levels, i.e., the spatial information contained in the neighbourhood of each nominal pixel (Prats-Montalbán et al., 2011a).

Multiple and different approaches are available to explore image texture properties and, according to literature, they can be essentially distinguished based on four categories (Bharati et al., 2004):

- *statistical methods*, which describe texture by computing a set of statistical descriptors from the distribution of grey-scale levels of pixel values. Various statistical methods are available, allowing to describe different texture properties (Tomita and Tsuji, 1990);

- *structural methods*, which define/describe image texture based on the properties and placement rule of a composition of well-defined texture elements (e.g., regularly spaced parallel lines). From a practical point of view, the application of these methods is quite limited, since they can be used considering only regular texture patterns and are not really suitable for random or natural textures (Gonzalez and Woods, 1992). Moreover, the application of these methods strictly requires a prior definition of the textural elements and of the location rules (Rosenfeld and Lipkin, 1970);
- *model-based methods*, which extract texture feature descriptors as estimated parameters of empirical models generated assuming that each image pixel intensity corresponds to a weighted average of its neighbouring ones. Autoregressive models (Sarkar et al., 1997), Markov random fields (Cross and Jain, 1983; Winkler, 2012) and fractal models (Keller et al., 1989; Pentland, 1984) are among the most common;
- *transform-based methods*, which describe the image texture by a transformation of the original image based on the spatial frequency properties of the pixel intensities variations. Different types of transforms are available and the choice of the most suitable plays a crucial role in the success of the analysis. Examples of types of transform are Fourier transforms (Alsberg et al., 1997; Walczak, 2000), Wigner distribution (Boukouvalas et al., 1998; Fioravanti et al., 1995; Song et al., 1995) and Wavelets, within of which Gabor transforms can be highlighted (Bovik et al., 1990; Freeman and Adelson, 1991; Materka and Strzelecki, 1998; Simoncelli et al., 1992; Tzagkarakis et al., 2006; Walczak, 2000).

Statistical methods are certainly among the most common and widespread texture analysis approaches, due to their relative simplicity and immediacy. Furthermore, based on the assumption that local features are defined by the combination of intensities at specified positions relative to each point in the image, the derived statistics can be further classified into first-, second- and higher-order, depending on the number of points which define the local feature (Bevk and Kononenko, 2002; Kononenko and Kukar, 2007; Srinivasan and Shobha, 2008; Tomita and Tsuji, 1990):

- *First order statistics* are calculated based on the probability that, when observing an image pixel at a randomly chosen location in the image, it has certain properties. The most common statistical methods can be associated to the computation of mean, median and statistical central moments (variance, skewness, kurtosis, 5^{th} order moment, etc.) and can be regarded as a good option to describe well-spread random textures (Prats-Montalbán et al., 2011a). However, these statistics depend only on individual pixels values and are independent of relationships between neighbouring pixels. In other words, they provide information related to the intensity level distribution in the image, but not about the relative positions of the

various intensity levels within the image. Soft Colour Texture Descriptors (SCTD) described in Section 3.4.2 are an example of first order statistical texture descriptors;

- *Second order statistics* are calculated based on the probability that, when randomly selecting a pair of image pixels, they have certain predefined properties. The most well-known set of this type of statistics in image analysis is undoubtedly represented by Gray Level Co-occurrence Matrices (GLCM) (Haralick et al., 1973) described in Section 3.4.3. Contrary to first-order statistics, second-order statistical methods provide information about the spatial relationship between pairs of pixels with defined intensity values;
- *Higher order statistics* which are computed considering the properties of three or more image pixels values, occurring at specific locations relative to each other, at a time.

The identification of the most suitable colour and/or texture features to be extracted from RGB images is neither a trivial nor absolute choice. Such a selection strictly depends on the sample characteristics, issues and final goal to be addressed for the specific application and at the same time can make a crucial difference in the final outcomes. In other words, it is really difficult and critical to define the optimal method to extract the appropriate image features for the given specific case. In this context, the proposal and application of different image-level descriptors extraction techniques provide the possibility to focus on multiple and different image features. In addition, one or several methods might be employed together to acquire image features for some applications in food quality evaluation (Zheng et al., 2006). Besides, the search and development of other approaches and features to be extracted is ongoing.

Within this framework, this PhD Thesis is focused on the application of different image-level techniques intended to extract colour and texture properties from RGB images, as will be detailed in the following sections. In addition, an alternative method to extract and codify the image texture features into a one-dimensional signal, hereinafter named *texturegram,* is proposed.

3.4.1 Colourgrams

The colourgrams method is an example of an image-level feature extraction method accounting for colour properties of RGB images. Each RGB image is converted into a colourgram, which is a one-dimensional signal of 4900 points encompassing and codifying the colour-related properties of the corresponding image. This chemometric strategy has been developed to both reduce the data dimensionality and allow to analyse in an automated way a large number of images altogether (Antonelli et al., 2004).

Basically, the colourgrams approach is based on the computation for each RGB image of the frequency distribution curves of several colour-related descriptors, including R, G and B values, lightness, relative colours, hue, saturation and intensity values, and the score values of different PCA models calculated on raw, mean centred and autoscaled RGB data.

More in detail, at first each three-dimensional RGB image with size $\{r, c, 3\}$ is unfolded into a two-dimensional array with size $\{(r \times c), 3\}$, where r corresponds to the number of pixel rows, c to the number of pixel columns and 3 to the three colour channels.

Then, four colour-related parameters are calculated for each pixel starting from the corresponding R, G and B values as listed below using the corresponding equations:

- Lightness (L):

$$L = R + G + B \tag{3.3}$$

- Relative red (rR):

$$rR = \frac{R}{L} \tag{3.4}$$

- Relative green (rG):

$$rG = \frac{G}{L} \tag{3.5}$$

- Relative blue (rB):

$$rB = \frac{B}{L} \tag{3.6}$$

Subsequently, three additional series of colour-related parameters are computed for each pixel and correspond to hue (H), saturation (S) and intensity (I) in order to convert the RGB space in the corresponding HSI space.

Hue values are calculated using a quite complex algorithm which is difficult to represent in an easily understandable analytical form. The hue ranges from 0 to 1, corresponding to a colour transition from red thorough yellow, green, cyan, blue and magenta, then back to red.

Saturation values are calculated according to the following equation:

$$S = \frac{\max(RGB) - \min(RGB)}{\max(RGB)} \tag{3.7}$$

while intensity values are calculated according to the following equation:

$$I = \max(RGB) \tag{3.8}$$

Afterwards, the unfolded RGB matrix is analysed by means of PCA, whose principal components directions are due to the distribution of the pixels in the original RGB space. A total of three PCA models are calculated, starting from raw, mean centred and autoscaled RGB data, respectively.

Different data pretreatments result in different PCA outcomes. For example, in the case of raw data, PC1 basically accounts for the distance between the centroid of the objects in the original RGB space and the axes origin, while in case of mean centred data, the direction of the principal components principally depends on the variance of the original variables (R, G and B channels) while neglecting their corresponding average level. On the other side, autoscaling could be convenient to account for the correlation between the original variables regardless of their variance. Typically, there is no *a priori* knowledge of which pretreatment performs best for a given issue. Hence the three different pretreatments are all considered in order to preserve as much as possible the general applicability of the colourgrams approach.

At this point, the initial two-dimensional RGB data matrix is expanded by adding a total of 16 descriptors, hence achieving a data matrix with 19 variables on the whole: R, G, B, L, rR, rG, rB, H, S, I and three score vectors for each PCA model calculated on the raw (SC1R, SC2R, SC3R), mean centred (SC1M, SC2M, SC3M) and autoscaled data matrix (SC1A, SC2A, SC3A), respectively.

As a following step, for each one of the 19 colour-related parameters the corresponding frequency distribution curve is calculated considering 256 bins and its entire range of variability. For example, for the G parameter, the frequency distribution curve is always calculated considering the whole range between 0 and 255 even when extreme values are not effectively registered.

Next, the frequency distribution curves are merged in sequence obtaining the first part of a one-dimensional signal, which is finally completed by adding at the end the loading vectors and the eigenvalues of the three PCs for each one of the three PCA models. The complete colourgram is now obtained, encoding the colour properties of the corresponding image and reaching a final length equal to 4900 ($= 256 \times 19 + 36$) points.

The whole procedure to convert an RGB image into the corresponding colourgram is schematically represented in Figure 3.9, where in the lower part all the regions of an image colourgram have been labelled according to the detailed description given in Table 3.2.

The conversion of an RGB image into a one-dimensional signal boasts multiple advantages in terms of data analysis. The colourgrams approach allows a notable data compression: 4900 points is a considerably low number of data compared to the millions characterizing a typical RGB image while preserving the corresponding relevant colour-related information. Truly, the several colour-related

descriptors included in the signal allow a simultaneous and comprehensive investigation of the colour-related properties of an image with no need to know *a priori* the specific parameters required to address the issue at hand. However, each colourgram accounts for several colour-related image properties which may be partially redundant, as some have similar meanings, and not all are useful for a given task. Nevertheless, it is possible to simply reduce the number of considered variables and select the most suitable descriptors *a posteriori* by means of variable selection methods.

In this way, the colour-related properties of large datasets of RGB images can be easily handled and explored with a relatively low computational effort. In fact, the colourgrams matrix can be analysed by means of multiple multivariate statistical methods, enabling the visualization of the overall structure of the dataset, the identification of possible outlier images, the computation of multivariate calibration and classification models based on specific colour-related properties of the imaged samples, and the selection of the colour features relevant for a specific problem. Regarding the latter, the selected colourgrams features can be represented back into the original image domain allowing to visually evaluate the results in the original image scene and to localize the image areas that are characterized by the features selected on the colourgram.

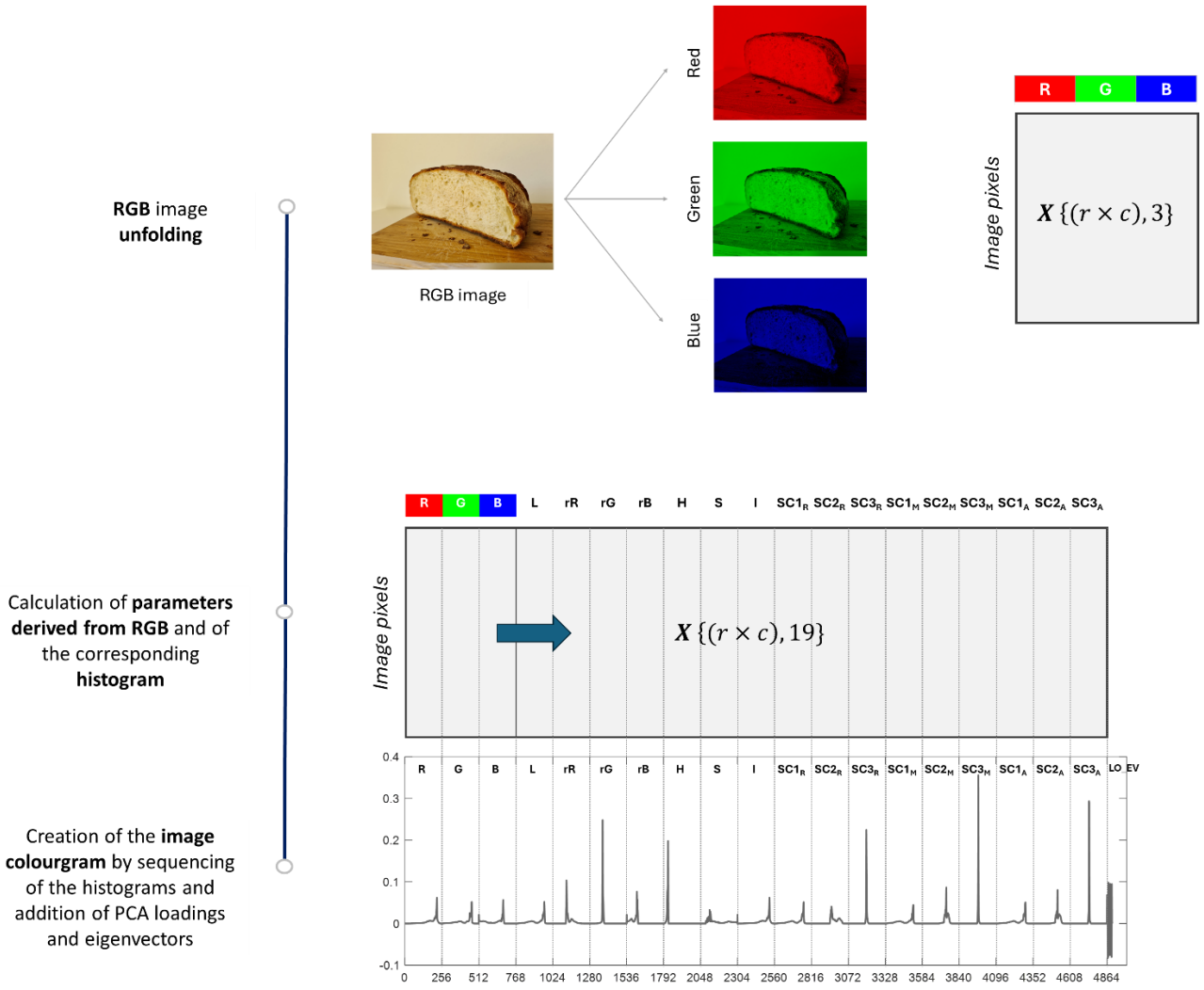


Figure 3.9 *Schematic representation of the key steps involved in the conversion of a RGB image into the corresponding colourgram. For the description of the colourgram regions, please refer to the corresponding abbreviations in Table 3.2.*

Table 3.2 *List of the colourgram region and the corresponding indexing and description.*

Abbreviation	Colourgram region	Description
R	1:256	Frequency distribution curve of the red (R) values
G	257:512	Frequency distribution curve of the green (G) values
B	513:768	Frequency distribution curve of the blue (B) values
L	769:1024	Frequency distribution curve of the lightness (L) values and $L = R + G + B$
rR	1025:1280	Frequency distribution curve of the relative red (rR) values and $rR = R/L$
rG	1281:1536	Frequency distribution curve of the relative green (rG) values and $rG = G/L$
rB	1537:1792	Frequency distribution curve of the relative blue (rB) values and $rB = B/L$
H	1793:2048	Frequency distribution curve of the hue (H) values calculated according to the *rgb2hsv* MATLAB® function
S	2049:2304	Frequency distribution curve of the saturation (S) values and $S = [\max(R,G,B) - \min(R,G,B)]/\max(R,G,B)$
I	2305:2560	Frequency distribution curve of the intensity (I) values and $I = \max(R,G,B)$
SC1R	2561:2816	Frequency distribution curve of the first score vector of the PCA model calculated on the raw unfolded RGB matrix (SC1R)
SC2R	2817:307	Frequency distribution curve of the first second vector of the PCA model calculated on the raw unfolded RGB matrix (SC2R)
SC3R	3073:3328	Frequency distribution curve of the third score vector of the PCA model calculated on the raw unfolded RGB matrix (SC3R)
SC1M	3329:3584	Frequency distribution curve of the first score vector of the PCA model calculated on the mean centred unfolded RGB matrix (SC1M)
SC2M	3585:3840	Frequency distribution curve of the second score vector of the PCA model calculated on the mean centred unfolded RGB matrix (SC2M)
SC3M	3841:4096	Frequency distribution curve of the third score vector of the PCA model calculated on the mean centred unfolded RGB matrix (SC3M)
SC1A	4097:4352	Frequency distribution curve of the first score vector of the PCA model calculated on the autoscaled unfolded RGB matrix (SC1A)
SC2A	4353:4608	Frequency distribution curve of the second score vector of the PCA model calculated on the autoscaled unfolded RGB matrix (SC2A)
SC3A	4609:4864	Frequency distribution curve of the third score vector of the PCA model calculated on the autoscaled unfolded RGB matrix (SC3A)
LO_EV	4865:4900	Normalized loading vectors and eigenvalues of the three PCA models

3.4.2 Soft Colour Texture Descriptors

Soft colour texture descriptors (SCTD) (López et al., 2005) are a set of first order statistical features that can be extracted from digital images in a defined colour space. The name soft colour texture descriptors highlights the simplicity and low computational effort required for their calculation. This approach resulted particularly successful to describe well-spread random textures (López, 2008; López et al., 2005). In the application of SCTD reported in this PhD Thesis, these statistical texture features have been extracted considering the RGB colour space, but in principle also different colour spaces can be considered based on the specific problem.

Basically, for a given RGB image the feature extraction process starts with unfolding and calculating the histograms of the R, G and B channels. Then, for each channel the mean value and different mathematical statistical central moments (from 2^{nd} to 5^{th} order) are computed. Therefore, 5 statistical descriptors accounting for texture are obtained from each channel and combined into a single feature vector with 15 elements, corresponding to 5 descriptors × 3 RGB channels.

The same procedure can be applied to all the images of the dataset, obtaining one feature vector with 15 SCTD for each image.

3.4.3 Grey Level Co-occurrence Matrix

Grey Level Co-occurrence Matrix (GLCM) is a second-order statistical method used to analyse image texture by considering the spatial relationships between pixels (Albregtsen, 2008; Haralick et al., 1973). For a grey-scale image, a GLCM is a square matrix with as many rows and columns as the number of discrete grey level intensities of the image being examined. The GLCM accounts for the frequencies of all possible combinations of grey levels that can occur between two image pixels located at a fixed spatial position relative to each other (Prats-Montalbán et al., 2011b). Therefore, the GLCM depends on the spatial relationship between pairs of pixels, which can be expressed as a function of two parameters: relative distance measured in pixel numbers (d) and their relative orientation (θ). The orientation θ is represented with four possible angle values, i.e., 0°, 45°, 90° and 135°, corresponding to horizontal, diagonal, vertical and anti-diagonal directions, respectively (Figure 3.10).

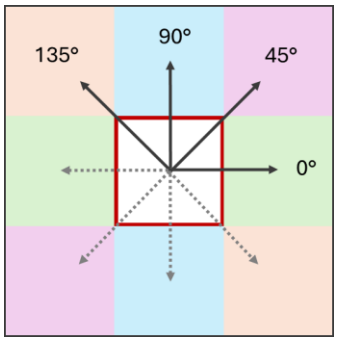


Figure 3.10 *Example of neighbouring pixels for a given pixel of interest (edged in red) according to a distance of 1 pixel. Green pixels are regarded as 0° neighbours, lilac pixels as 45° neighbours, light blue pixels as 90° neighbours, orange pixels as 135° neighbours.*

A general example of the calculation of GLCM from a grey-scale image considering a 1 pixel distance and the four possible directions is shown in Figure 3.11.

0	0	1	2
1	3	2	2
0	3	3	3
2	2	3	1

(a)

0°

	0	1	2	3
0	2	1	0	1
1	1	0	1	2
2	0	1	4	2
3	1	2	2	4

(b)

45°

	0	1	2	3
0	0	1	0	1
1	1	0	0	1
2	0	0	2	4
3	1	1	4	2

(c)

90°

	0	1	2	3
0	0	2	1	1
1	2	0	1	1
2	1	1	2	3
3	1	1	3	4

(d)

135°

	0	1	2	3
0	0	0	2	1
1	0	0	1	2
2	2	1	0	1
3	1	2	1	4

(e)

Figure 3.11 *Example of GLCM calculation from a 4 × 4 grey-scale image (a) with four intensity grey-tone values from 0 to 3. In (b), (c), (d) and (e) the corresponding GLCM matrices computed for 1 pixel distance and 0°, 45°, 90° and 135°, respectively.*

For a quantitative evaluation of the information contained in the GLCM, Haralick (Haralick et al., 1973) proposed to extract several textural descriptors from a single GLCM, among which contrast, correlation, energy, homogeneity and entropy are the most common ones and were considered in this PhD Thesis. Complete equations for the computation of GLCM descriptors are reported in Table 3.3, where N_g is the number of grey levels in the image and $P_d^\theta(i,j)$ is the $(i, j)^{th}$ entry of GLCM, corresponding to co-occurrence probability of having pairs of pixels with i^{th} and j^{th} grey levels at distance d and angle θ (Fongaro and Kvaal, 2013; Malegori et al., 2016; Singh et al., 2017).

Table 3.3 *Equations and definitions of the texture descriptors extracted from GLCM.*

Name	**Equation**	**Description**
Contrast	$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} P_d^\theta(i,j)\lvert i-j\rvert^2$	Also known as *variance* or *inertia*, contrast measures the intensity difference between a pixel and its neighbours. Contrast is equal to zero for a constant image, while high values indicate significant intensity variations.
Correlation	$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} \frac{(i-\mu_i)(j-\mu_j)P_d^\theta(i,j)}{\sigma_i\sigma_j}$	Correlation provides a measure of linear dependency between pixels intensities. It can assume values ranging from -1 to 1.
Energy	$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} P_d^\theta(i,j)^2$	Also known as *uniformity* or *angular second moment*, energy returns the sum of squared elements in the GLCM and it measures the uniformity of the texture. Energy is equal to 1 for a constant image.
Homogeneity	$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} \frac{1}{1+(i-j)^2} P_d^\theta(i,j)$	Homogeneity provides a measure of how close the GLCM distribution is to the diagonal. Homogeneity is equal to 1 for a diagonal GLCM
Entropy	$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} P_d^\theta(i,j)\log P_d^\theta(i,j)$	Entropy is related to statistical randomness in the image texture and it measures the image disorder

To fully describe image texture properties, it is recommended to compute multiple GLCM based on different distances and directions. For a fixed distance value, GLCMs are commonly calculated according to the four possible directions, obtaining four distinct GLCMs for a single grey-scale image. However, when the textural features under investigation are expected to be invariant to the direction, the mean values of the descriptors extracted over all the four GLCMs are computed.

All these principles can be easily extended to RGB images, by calculating GLCM features for each single channel. Since each R, G and B component is coded by 8 bits, 256 discrete intensity levels (ranging from 0 to 255) are considered resulting in a GLCM with dimensions of 256 rows × 256 columns for a given distance and orientation. Then, the GLCM descriptors obtained for each colour

channel can be collected together to obtain the final vector of texture features corresponding to a single RGB image of the dataset.

3.4.4 Texturegrams

In this PhD Thesis, the *texturegrams* approach is presented for the first time. It is an image-level data reduction method that converts each RGB image into a signal, called the *texturegram*, which encodes the texture information of the image. The rationale behind texturegrams is the same as for colourgrams and the other feature extraction methods presented in this Chapter: enabling the simultaneous analysis of large datasets of images by encoding the information of interest into feature vectors. These vectors can be further elaborated using common chemometric methods.

Briefly, the procedure used to compute the *texturegrams* consists of two main steps. Firstly, each RGB image is converted into a multichannel texture image using a modified version of the approach presented by Bharati and MacGregor (2001) and later extended to RGB images by Prats-Montalbán and Ferrer (2007). Then, the multichannel texture images are converted into the corresponding *texturegrams* using a data dimensionality reduction algorithm initially proposed for hyperspectral images (Calvini et al., 2016).

A detailed description of the algorithm used to calculate texturegrams is provided in Section 5.2.

3.5 References

Abdullah, M.Z., Aziz, S.A., DOS MOHAMED, A.M., 2000. Quality inspection of bakery products using a color-based machine vision system. Journal of Food Quality 23, 39–50.

Albregtsen, F., 2008. Statistical texture measures computed from gray level coocurrence matrices. Image processing laboratory, department of informatics, university of oslo 5.

Alsberg, B.K., Woodward, A.M., Kell, D.B., 1997. An introduction to wavelet transforms for chemometricians: A time-frequency approach. Chemometrics and intelligent laboratory systems 37, 215–239.

Antonelli, A., Cocchi, M., Fava, P., Foca, G., Franchini, G.C., Manzini, D., Ulrici, A., 2004. Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm. Analytica Chimica Acta 515, 3–13. https://doi.org/10.1016/j.aca.2004.01.005

Aquino, A., Millan, B., Gaston, D., Diago, M.-P., Tardaguila, J., 2015. vitisFlower®: Development and Testing of a Novel Android-Smartphone Application for Assessing the Number of Grapevine Flowers per Inflorescence Using Artificial Vision Techniques. Sensors 15, 21204–21218. https://doi.org/10.3390/s150921204

Bachelor, B., 1985. Lighting and viewing techniques in automated visual inspection. IFS Publication Ltd., Bedford, UK.

Barni, M., Cappellini, V., Mecocci, A., 1997. Colour-based detection of defects on chicken meat. Image and vision computing 15, 549–556.

Bevk, M., Kononenko, I., 2002. A statistical approach to texture description of medical images: a preliminary study, in: Proceedings of 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002). Presented at the 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002), IEEE Computer. Soc, Maribor, Slovenia, pp. 239–244. https://doi.org/10.1109/CBMS.2002.1011383

Beyaz, A., Martínez Gila, D.M., Gómez Ortega, J., Gámez García, J., 2019. Olive fly sting detection based on computer vision. Postharvest Biology and Technology 150, 129–136. https://doi.org/10.1016/j.postharvbio.2019.01.003

Bharati, M.H., Liu, J.J., MacGregor, J.F., 2004. Image texture analysis: methods and comparisons. Chemometrics and Intelligent Laboratory Systems 72, 57–71. https://doi.org/10.1016/j.chemolab.2004.02.005

Bharati, M.H., MacGregor, J.F., 2001. Texture analysis of images using principal component analysis, in: McCann, H., Scott, D.M. (Eds.), . Presented at the Intelligent Systems and Smart Manufacturing, Boston, MA, p. 27. https://doi.org/10.1117/12.417179

Blasco, J., Cubero, S., Gómez-Sanchís, J., Mira, P., Moltó, E., 2009. Development of a machine for the automatic sorting of pomegranate (Punica granatum) arils based on computer vision. Journal of Food Engineering 90, 27–34. https://doi.org/10.1016/j.jfoodeng.2008.05.035

Boukouvalas, C., De Natale, F., De Toni, G., Kittler, J., Marik, R., Mirmehdi, M., Petrou, M., Le Roy, P., Salgari, R., Vernazza, G., 1998. ASSIST: automatic system for surface inspection and sorting of tiles. Journal of Materials Processing Technology 82, 179–188.

Bovik, A.C., Clark, M., Geisler, W.S., 1990. Multichannel texture analysis using localized spatial filters. IEEE transactions on pattern analysis and machine intelligence 12, 55–73.

Brosnan, T., Sun, D.-W., 2004. Improving quality inspection of food products by computer vision—a review. Journal of Food Engineering 61, 3–16. https://doi.org/10.1016/S0260-8774(03)00183-3

Brosnan, T., Sun, D.-W., 2002. Inspection and grading of agricultural and food products by computer vision systems—a review. Computers and Electronics in Agriculture 36, 193–213. https://doi.org/10.1016/S0168-1699(02)00101-1

Calvini, R., Foca, G., Ulrici, A., 2016. Data dimensionality reduction and data fusion for fast characterization of green coffee samples using hyperspectral sensors. Anal Bioanal Chem 408, 7351–7366. https://doi.org/10.1007/s00216-016-9713-7

Calvini, R., Orlandi, G., Foca, G., Ulrici, A., 2020. Colourgrams GUI: A graphical user-friendly interface for the analysis of large datasets of RGB images. Chemometrics and Intelligent Laboratory Systems 196, 103915. https://doi.org/10.1016/j.chemolab.2019.103915

Calvini, R., Pigani, L., 2022. Toward the Development of Combined Artificial Sensing Systems for Food Quality Evaluation: A Review on the Application of Data Fusion of Electronic Noses, Electronic Tongues and Electronic Eyes. Sensors 22, 577. https://doi.org/10.3390/s22020577

Chang, Y.-C., Reid, J.F., 1996. RGB calibration for color image analysis in machine vision. IEEE Transactions on image processing 5, 1414–1422.

Chmiel, M., Słowiński, M., 2016. The use of computer vision system to detect pork defects. LWT 73, 473–480. https://doi.org/10.1016/j.lwt.2016.06.054

Cross, G.R., Jain, A.K., 1983. Markov Random Field Texture Models. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-5, 25–39. https://doi.org/10.1109/TPAMI.1983.4767341

Cruz-Fernández, M., Luque-Cobija, M.J., Cervera, M.L., Morales-Rubio, A., De La Guardia, M., 2017. Smartphone determination of fat in cured meat products. Microchemical Journal 132, 8–14. https://doi.org/10.1016/j.microc.2016.12.020

Cubero, S., Diago, M.P., Blasco, J., Tardaguila, J., Prats-Montalbán, J.M., Ibáñez, J., Tello, J., Aleixos, N., 2015. A new method for assessment of bunch compactness using automated image analysis: Bunch compactness assessment using image analysis. Australian Journal of Grape and Wine Research 21, 101–109. https://doi.org/10.1111/ajgw.12118

Davidson, V.J., Ryks, J., Chu, T., 2001. Fuzzy models to predict consumer ratings for biscuits based on digital image features. IEEE Transactions on Fuzzy Systems 9, 62–67.

Dutta, M.K., Singh, A., Ghosal, S., 2015. A computer vision based technique for identification of acrylamide in potato chips. Computers and Electronics in Agriculture 119, 40–50. https://doi.org/10.1016/j.compag.2015.10.007

ElMasry, G., Cubero, S., Moltó, E., Blasco, J., 2012. In-line sorting of irregular potatoes by using automated computer-based machine vision system. Journal of Food Engineering 112, 60–68. https://doi.org/10.1016/j.jfoodeng.2012.03.027

Esbensen, K., Geladi, P., 1989. Strategy of multivariate image analysis (MIA). Chemometrics and Intelligent Laboratory Systems 7, 67–86.

Fioravanti, S., Fioravanti, R., De Natale, F.G., Marik, R., Mirmehdi, M., Kittler, J., Petrou, M., 1995. Spectral and rank order approaches to texture analysis. European Transactions on Telecommunications 6, 287–299.

Fongaro, L., Kvaal, K., 2013. Surface texture characterization of an Italian pasta by means of univariate and multivariate feature extraction from their texture images. Food research international 51, 693–705.

Freeman, W.T., Adelson, E.H., 1991. The design and use of steerable filters. IEEE Transactions on Pattern analysis and machine intelligence 13, 891–906.

Giraudo, A., Calvini, R., Orlandi, G., Ulrici, A., Geobaldo, F., Savorani, F., 2018. Development of an automated method for the identification of defective hazelnuts based on RGB image analysis and colourgrams. Food Control 94, 233–240. https://doi.org/10.1016/j.foodcont.2018.07.018

Girolami, A., Napolitano, F., Faraone, D., Braghieri, A., 2013. Measurement of meat color using a computer vision system. Meat Science 93, 111–118. https://doi.org/10.1016/j.meatsci.2012.08.010

Gonzalez, R.C., Woods, R.E., 1992. Digital image processing addison-wesley. Reading, Ma.

Grahn, H., Geladi, P., 2007. Techniques and applications of hyperspectral image analysis. John Wiley & Sons.

Gunasekaran, S., 1996. Computer vision technology for food quality assurance. Trends in Food Science & Technology 7, 245–256. https://doi.org/10.1016/0924-2244(96)10028-5

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. IEEE Trans. Syst., Man, Cybern. SMC-3, 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Hosseinpour, S., Ilkhchi, A.H., Aghbashlo, M., 2019. An intelligent machine vision-based smartphone app for beef quality evaluation. Journal of Food Engineering 248, 9–22. https://doi.org/10.1016/j.jfoodeng.2018.12.009

Hunt, R.W.G., 2005. The Reproduction of Colour. John Wiley & Sons.

Jayasundara, J.M.V.D.B., Ramanayake, R.M.L.S., Senarath, H.M.N.B., Herath, H.M.S.L., Godaliyadda, G.M.R.I., Ekanayake, M.P.B., Herath, H.M.V.R., Ariyawansa, S., 2023. Deep learning for automated fish grading. Journal of Agriculture and Food Research 14, 100711. https://doi.org/10.1016/j.jafr.2023.100711

Keller, J.M., Chen, S., Crownover, R.M., 1989. Texture description and segmentation through fractal geometry. Computer Vision, Graphics, and Image Processing 45, 150–166. https://doi.org/10.1016/0734-189X(89)90130-8

Kononenko, I., Kukar, M., 2007. Data Preprocessing, in: Machine Learning and Data Mining. Elsevier, pp. 181–211. https://doi.org/10.1533/9780857099440.181

Lawless, H., 1995. Dimensions of sensory quality: A critique. Food Quality and Preference 6, 191–199. https://doi.org/10.1016/0950-3293(94)00023-O

León, K., Mery, D., Pedreschi, F., León, J., 2006. Color measurement in L∗a∗b∗ units from RGB digital images. Food Research International 39, 1084–1091. https://doi.org/10.1016/j.foodres.2006.03.006

López, F., 2008. Multivariate statistical projection methods to perform robust feature extraction and classification in surface grading. J. Electron. Imaging 17, 031106. https://doi.org/10.1117/1.2957886

López, F., Valiente, J.-M., Prats, J.-M., 2005. Surface Grading Using Soft Colour-Texture Descriptors, in: Sanfeliu, A., Cortés, M.L. (Eds.), Progress in Pattern Recognition, Image Analysis and Applications, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 13–23. https://doi.org/10.1007/11578079_2

Malegori, C., Franzetti, L., Guidetti, R., Casiraghi, E., Rossi, R., 2016. GLCM, an image analysis technique for early detection of biofilm. Journal of Food Engineering 185, 48–55. https://doi.org/10.1016/j.jfoodeng.2016.04.001

Marszalec, E., Pietikäinen, M., 1996. SOME ASPECTS OF RGB VISION AND ITS APPLICATIONS IN INDUSTRY. Int. J. Patt. Recogn. Artif. Intell. 10, 55–72. https://doi.org/10.1142/S0218001496000062

Materka, A., Strzelecki, M., 1998. Texture analysis methods–a review. Technical university of lodz, institute of electronics, COST B11 report, Brussels 10, 4968.

McCaig, T.N., 2002. Extending the use of visible/near-infrared reflectance spectrophotometers to measure colour of food and agricultural products. Food Research International 35, 731–736. https://doi.org/10.1016/S0963-9969(02)00068-6

Menesatti, P., Angelini, C., Pallottino, F., Antonucci, F., Aguzzi, J., Costa, C., 2012. RGB Color Calibration for Quantitative Image Analysis: The "3D Thin-Plate Spline" Warping Approach. Sensors 12, 7063–7079. https://doi.org/10.3390/s120607063

Molnár, P.J., 1995. A model for overall description of food quality. Food Quality and Preference 6, 185–190. https://doi.org/10.1016/0950-3293(94)00037-V

Narendra, V., Hareesh, K., 2010. Prospects of computer vision automated grading and sorting systems in agricultural and food products for quality evaluation. International Journal of Computer Applications 1, 1–9.

Narendra, V.G., Hareesh, K.S., 2010. Quality Inspection and Grading of Agricultural and Food Products by Computer Vision-a Review. IJCA 2, 43–65. https://doi.org/10.5120/612-863

Ornberg, R.L., Woerner, B.M., Edwards, D.A., 1999. Analysis of Stained Objects in Histological Sections by Spectral Imaging and Differential Absorption. J Histochem Cytochem. 47, 1307–1313. https://doi.org/10.1177/002215549904701010

Pace, B., Cefola, M., Renna, F., Renna, M., Serio, F., Attolico, G., 2013. Multiple regression models and Computer Vision Systems to predict antioxidant activity and total phenols in pigmented carrots. Journal of Food Engineering 117, 74–81. https://doi.org/10.1016/j.jfoodeng.2013.02.005

Pedreschi, F., León, J., Mery, D., Moyano, P., 2006. Development of a computer vision system to measure the color of potato chips. Food Research International 39, 1092–1098. https://doi.org/10.1016/j.foodres.2006.03.009

Pentland, A.P., 1984. Fractal-based description of natural scenes. IEEE transactions on pattern analysis and machine intelligence 661–674.

Pereira, A.C., Reis, M.S., Saraiva, P.M., 2009. Quality Control of Food Products using Image Analysis and Multivariate Statistical Tools. Ind. Eng. Chem. Res. 48, 988–998. https://doi.org/10.1021/ie071610b

Perez De Vargas-Sansalvador, I.M., Erenas, M.M., Martínez-Olmos, A., Mirza-Montoro, F., Diamond, D., Capitan-Vallvey, L.F., 2020. Smartphone based meat freshness detection. Talanta 216, 120985. https://doi.org/10.1016/j.talanta.2020.120985

Prats-Montalbán, J.M., de Juan, A., Ferrer, A., 2011a. Multivariate image analysis: A review with applications. Chemometrics and Intelligent Laboratory Systems 107, 1–23. https://doi.org/10.1016/j.chemolab.2011.03.002

Prats-Montalbán, J.M., De Juan, A., Ferrer, A., 2011b. Multivariate image analysis: A review with applications. Chemometrics and Intelligent Laboratory Systems 107, 1–23. https://doi.org/10.1016/j.chemolab.2011.03.002

Prats-Montalbán, J.M., Ferrer, A., 2007. Integration of colour and textural information in multivariate image analysis: defect detection and classification issues. J. Chemometrics 21, 10–23. https://doi.org/10.1002/cem.1026

Rolando, P.L., Calvini, R., Foca, G., Ulrici, A., 2023. Mixture design and multivariate image analysis to monitor the colour of strawberry yoghurt purée. Microchemical Journal 194, 109222. https://doi.org/10.1016/j.microc.2023.109222

Rosenfeld, A., Lipkin, B.S., 1970. Texture synthesis, in: Picture Processing and Psychopictorics. Academic Press New York, pp. 309–345.

Saha, K.K., Rahman, A., Moniruzzaman, M., Syduzzaman, M., Uddin, M.Z., Rahman, M.M., Ali, M.A., Al Riza, D.F., Oliver, M.M.H., 2023. Classification of starfruit maturity using smartphone-image and multivariate analysis. Journal of Agriculture and Food Research 11, 100473. https://doi.org/10.1016/j.jafr.2022.100473

Sarkar, A., Sharma, K.M.S., Sonak, R.V., 1997. A new approach for subset 2-D AR model identification for describing textures. IEEE Trans. on Image Process. 6, 407–413. https://doi.org/10.1109/83.557348

Simoncelli, E.P., Freeman, W.T., Adelson, E.H., Heeger, D.J., 1992. Shiftable multiscale transforms. IEEE transactions on Information Theory 38, 587–607.

Singh, S., Srivastava, D., Agarwal, S., 2017. GLCM and its application in pattern recognition. Presented at the 2017 5th International Symposium on Computational and Business Intelligence (ISCBI), IEEE, pp. 20–25.

Song, K.Y., Petrou, M., Kittler, J., 1995. Texture crack detection. Machine Vision and Applications 8, 63–75.

Srinivasan, G.N., Shobha, G., 2008. Statistical Texture Analysis. Proceedings of world academy of science, engineering and technology 36, 1264–1269.

Taghadomi-Saberi, S., Masoumi, A.A., Sadeghi, M., Zekri, M., 2019. Integration of wavelet network and image processing for determination of total pigments in bitter orange (*Citrus aurantium* L.) peel during ripening. J Food Process Engineering 42, e13120. https://doi.org/10.1111/jfpe.13120

Tomita, F., Tsuji, S., 1990. Statistical Texture Analysis, in: Computer Analysis of Visual Textures. Springer US, Boston, MA, pp. 13–36. https://doi.org/10.1007/978-1-4613-1553-7_2

Tzagkarakis, G., Beferull-Lozano, B., Tsakalides, P., 2006. Rotation-invariant texture retrieval with gaussianized steerable pyramids. IEEE Transactions on Image Processing 15, 2702–2718.

Ulrici, A., Foca, G., Ielo, M.C., Volpelli, L.A., Lo Fiego, D.P., 2012. Automated identification and visualization of food defects using RGB imaging: Application to the detection of red skin defect of raw hams. Innovative Food Science & Emerging Technologies 16, 417–426. https://doi.org/10.1016/j.ifset.2012.09.008

Ulrici, A., Serranti, S., Ferrari, C., Cesare, D., Foca, G., Bonifazi, G., 2013. Efficient chemometric strategies for PET–PLA discrimination in recycling plants using hyperspectral imaging. Chemometrics and Intelligent Laboratory Systems 122, 31–39. https://doi.org/10.1016/j.chemolab.2013.01.001

Van Poucke, S., Haeghen, Y.V., Vissers, K., Meert, T., Jorens, P., 2010. Automatic colorimetric calibration of human wounds. BMC Med Imaging 10, 7. https://doi.org/10.1186/1471-2342-10-7

Vidal, A., Talens, P., Prats-Montalbán, J.M., Cubero, S., Albert, F., Blasco, J., 2013. In-Line Estimation of the Standard Colour Index of Citrus Fruits Using a Computer Vision System Developed For a Mobile Platform. Food Bioprocess Technol 6, 3412–3419. https://doi.org/10.1007/s11947-012-1015-2

Viscarra Rossel, R.A., Minasny, B., Roudier, P., McBratney, A.B., 2006. Colour space models for soil science. Geoderma 133, 320–337. https://doi.org/10.1016/j.geoderma.2005.07.017

Walczak, B., 2000. Wavelets in chemistry. Elsevier.

Williams, P., Manley, M., Fox, G., Geladi, P., 2010. Indirect detection of Fusarium verticillioides in maize (Zea mays L.) kernels by near infrared hyperspectral imaging. Journal of Near Infrared Spectroscopy 18, 49–58.

Winkler, G., 2012. Image analysis, random fields and Markov chain Monte Carlo methods: a mathematical introduction. Springer Science & Business Media.

Wu, D., Sun, D.-W., 2013. Colour measurements by computer vision for food quality control – A review. Trends in Food Science & Technology 29, 5–20. https://doi.org/10.1016/j.tifs.2012.08.004

Zheng, C., Sun, D.-W., Zheng, L., 2006. Recent developments and applications of image features for food quality evaluation and inspection – a review. Trends in Food Science & Technology 17, 642–655. https://doi.org/10.1016/j.tifs.2006.06.005

Chapter Four

NIR spectroscopy for food analysis

4.1 Theory of infrared spectroscopy

Infrared (IR) spectroscopy is based on the analysis of the interaction between matter and electromagnetic radiation in the infrared region, which promotes energy transitions between the vibrational levels of molecules. In this case the radiation serves as a "probe" to investigate the composition of matter throughout spectroscopic analysis. Before delving deeper into the physical phenomena underlying infrared spectroscopy, it is necessary to clarify some concepts about the nature of light, i.e., electromagnetic radiation.

Electromagnetic radiation consists of oscillating electric and magnetic fields that are sinusoidally in phase and orthogonal to each other and to the direction of light propagation. Light exhibits a dual nature: a wave-like behaviour characterized by frequency (ν) and wavelength (λ), and a particle-like behaviour represented by photons, each carrying a discrete amount of energy.

In vacuum, light travels at a constant speed (c) of approximately 2.998×10^8 m s^{-1}. This wave-like behaviour can be mathematically expressed by means of the following equation:

$$c = \lambda \times \nu \tag{4.1}$$

Regarding the particle nature of light, it is understood that photons carry an amount of energy E, which can be expressed as:

$$E = h\nu \tag{4.2}$$

where h is the Planck's constant ($= 6.626 \times 10^{-34}$ Js).

According to Equation 4.1, Equation 4.2 can also be written as:

$$E = h\frac{c}{\lambda} = h\tilde{\nu} \tag{4.3}$$

Where $\tilde{\nu}$ (cm^{-1}) is the wavenumber, defined as the reciprocal of the wavelength.

In spectroscopic analysis, the electromagnetic radiation can be transferred to molecules or atoms that constitute the sample, resulting in absorption phenomena. When absorption occurs, particles temporarily gain energy, transitioning from an initial energy level ($E_{initial}$) to a higher, excited energy

state (E_{final}). However, to make this transition effectively take place, the energy of the absorbed photon must precisely match the energy gap between these two states. In other words, the energy gap between the higher and lower energy state must be in accordance with Planck's law:

$$E_{final} - E_{initial} = \Delta E = h\nu = h\frac{c}{\lambda} \tag{4.4}$$

In the infrared region, this energy is sufficient to promote transitions between the vibrational quantum levels of molecules. Since the energy gaps between vibrational states are unique to each substance, IR spectroscopy is a powerful tool for identifying and characterizing chemical species.

IR spectroscopy is divided into three main spectral regions, each with distinct applications (Coates, 2000; Miller, 2001) (Figure 4.1):

- *Near infrared (NIR)*, primarily used for a quantitative indirect analysis. Wavelength range: 800–2500 nm (12500–4000 cm^{-1});
- *Mid infrared (MIR)*, commonly used to identify fundamental vibrations of functional groups in organic molecules, aiding in structural analysis. Wavelength range: 2500–25000 nm (4000–400 cm^{-1});
- *Far infrared (FIR)*, mainly used to characterize inorganic and organo-metallic compounds. Wavelength range: 25000–1×10^6 nm (400–10 cm^{-1}).

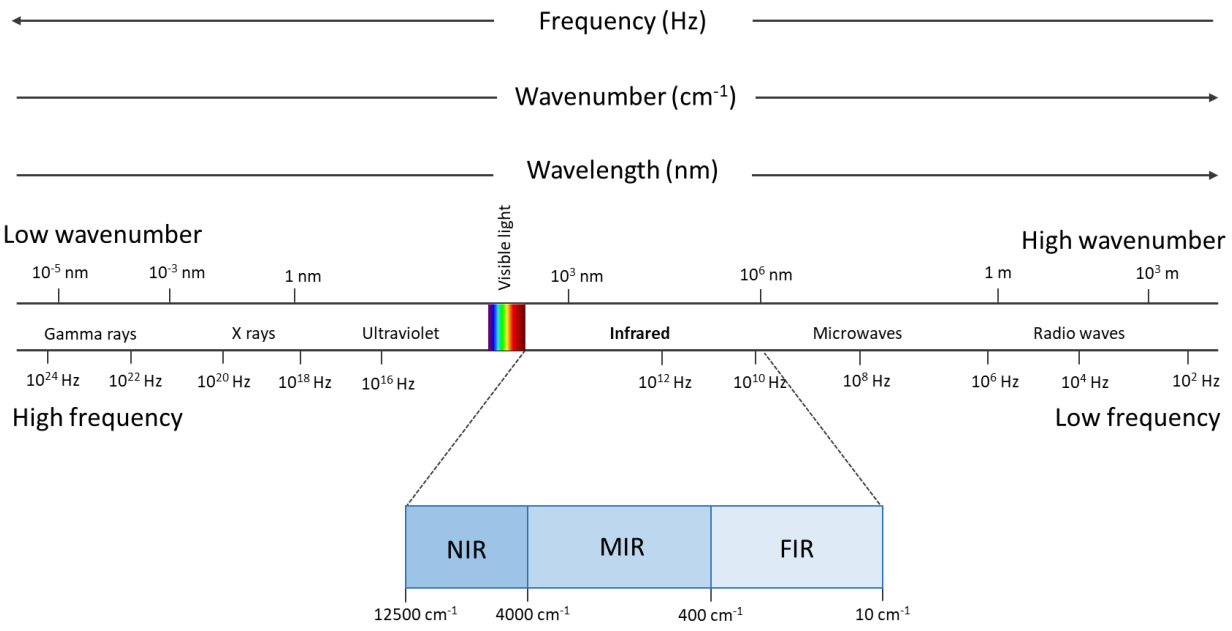


Figure 4.1 *Infrared (IR) region in the electromagnetic spectrum and highlight of near (NIR), mid (MIR) and far (FIR) infrared regions.*

The output of IR analysis is a spectrum, i.e., a graphical representation of IR radiation absorption as a function of wavenumber or wavelength. Each peak in the spectrum corresponds to energy absorption due to specific vibrational transitions of molecular bonds, providing insights into molecular structure and functional groups present.

To understand the origins of IR spectra, consider a diatomic molecule, modelled according to a classical harmonic oscillator composed of two masses (atoms) connected by a spring (chemical bond). The system oscillates with a frequency (ν) determined by the bond strength (spring's force constant, k) and the masses of the atoms:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{4.5}$$

Where μ is the reduced mass, calculated as:

$$\mu = \frac{(m_1 \times m_2)}{(m_1 + m_2)} \tag{4.6}$$

Stronger bonds (larger k) lead to higher frequencies, while heavier atoms (larger μ) result in lower frequencies. The system's potential energy follows a parabolic curve (harmonic potential), which increases as bond stretches or compresses (Figure 4.2).

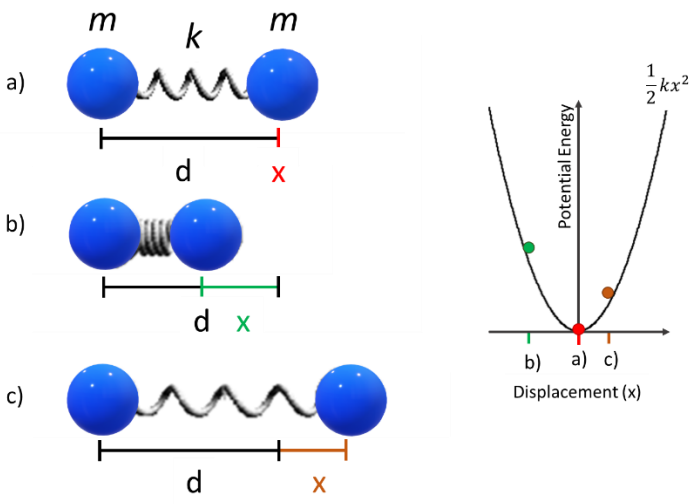


Figure 4.2 *Classical harmonic oscillator model.*

Actually, real molecules exhibit quantized vibrational energy levels. According to quantum mechanics (Figure 4.3), the energy of a vibrating molecule is given by:

$$E_n = \left(n + \frac{1}{2}\right) h\nu \tag{4.7}$$

where n (n= 0, 1, 2, 3, etc.) is the vibrational quantum number. Transitions occur between adjacent energy levels ($\Delta n = \pm 1$), absorbing or emitting light with energy equal to $h\nu$. Hence, for such a diatomic molecule the IR spectrum should have a single peak at the frequency corresponding to that energy.

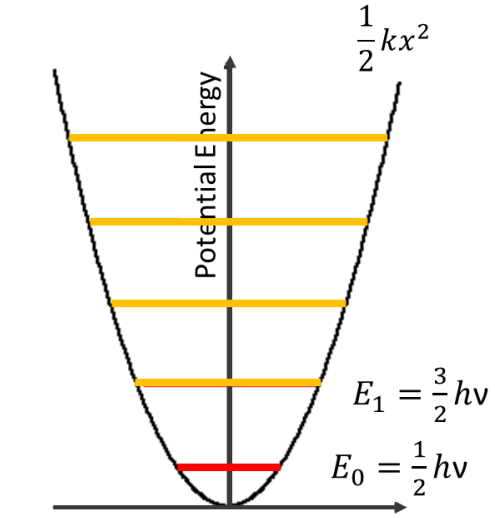


Figure 4.3 *Quantum mechanical model of a molecule.*

In reality, molecules deviate from harmonic oscillator model due to electrostatic forces between atoms, leading to an anharmonic potential (dark line in Figure 4.4). This allows for transitions with $\Delta n = \pm 2$ or $\Delta n = \pm 3$, known as overtones (light blue lines in Figure 4.4). In the IR spectrum, overtones appear at approximately twice or three times the fundamental vibrational frequency but are less intense. The farthest the vibration of the system from the harmonic behaviour (e.g., when dealing with bonds involving hydrogen, the lightest atom, and a heavier one like carbon, nitrogen or oxygen), the higher the intensity of the overtone bands.

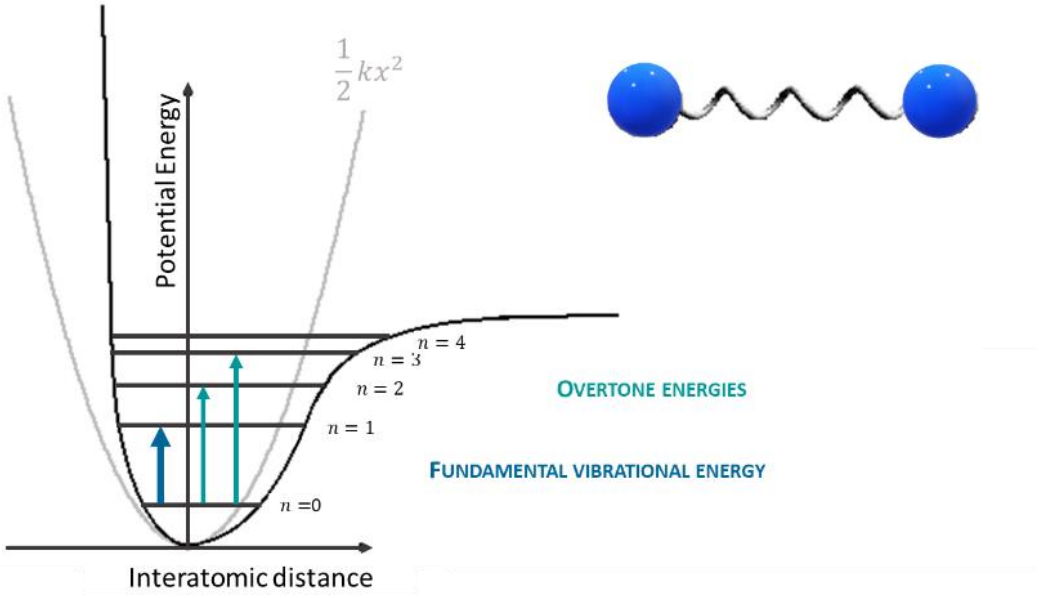


Figure 4.4 *Anharmonic oscillator model (dark line), where fundamental vibration and overtone bands are highlighted in blue and light blue, respectively, in comparison with a classical harmonic oscillator (grey line).*

Indeed, another aspect complicating the IR spectrum is owed to the different ways atoms constituting molecules are constantly moving and vibrating. There are two primary types of vibrations for a covalent bond:

- Stretching vibrations (symmetric or antisymmetric), when the chemical bond is stretched or compressed;
- Bending vibrations (rocking, twisting, wagging, scissoring), when the angle between atoms is tightened or expanded.

The number of vibrational modes for a given molecule depends on its composition and structure. Complex molecules with multiple atoms exhibit intricate IR spectra due to the overlapping of numerous vibrational modes.

From a dynamical point of view, IR radiation can – in specific conditions – increase the vibrational energy of the molecules and induce energy transitions which are manifested into an increase in the amplitude of the natural oscillations of the interatomic distances and of the bond angles of molecules.

As a final consideration, for a molecule to undergo IR absorption, it must have a dipole moment that changes during vibration. This change in dipole moment allows the molecule to interact with the oscillating electric field of the electromagnetic radiation. Consequently, electromagnetic radiation

with a frequency matching the natural vibrational frequency of the molecule can resonate with the system, leading to absorption. Homonuclear diatomic molecules (e.g., H_2, O_2 or N_2), however, lack a permanent dipole moment and do not experience a change in dipole moment during vibration. As a result, they do not absorb IR radiation.

In conclusion, IR spectroscopy is based on the principle that matter absorbs specific energy levels corresponding to different vibrational frequencies, known as modes, resulting in distinct spectral signatures. An IR spectrum is generated by recording the frequencies of light absorbed by the sample, i.e., the vibrations excited as the IR light passes through the sample. Each chemical species is characterized by vibrations at different frequencies, making the spectrum of each compound unique. Therefore, the spectrum can be considered as a "chemical fingerprint" for the identification and quantification of almost any chemical species absorbing in the IR region. Regarding quantification, higher concentration of an individual component in a sample leads to greater IR light absorption. In general, for a given IR spectrum, peak position provides qualitative information (frequencies are related to the type of functional groups involved), while peak intensity provides quantitative information.

As stated above, molecules typically absorb IR radiation with energy equal to $h\nu$ (fundamental vibrations) and, less commonly, at $2h\nu$ (first overtone) or $3h\nu$ (second overtone). Fundamental vibrations, which involve transitions between the vibrational ground state and the first excited state, are associated with the MIR region. In contrast, overtone vibrations, which involve transitions between higher excited states, are characteristic of the NIR region. In the NIR region, a further type of absorption bands, named combination bands, are frequently observed. These results from interactions among different fundamental vibrational frequencies in polyatomic molecules, where multiple vibrations occur simultaneously. A typical NIR spectrum includes both overtone bands and combinations bands corresponding to molecular vibrations initially detected in the MIR region (Barton, 2002; Burns and Ciurczak, 2007).

NIR electromagnetic radiation was the firstly discovered by William Herschel in 1800 (Pasquini, 2018). Herschel observed a rise in temperature, which he referred to as "radiant heat", by placing a conventional blackened bulb thermometer beyond the red end of the visible spectrum (dispersed sunlight through a glass prism). This allowed him to effectively detect the projection of NIR radiation (Herschel, 1800).

The increasing interest towards the advantages provided by NIR spectroscopy led to a notable spreading of this technique in many industrial and academic applications over time. All these aspects will be further detailed in the following sections.

4.2 Near infrared spectroscopy

NIR active modes are often associated with highly anharmonic vibrational modes of molecular functional groups composed of a relatively heavy atom attached to hydrogen (e.g., C-H, O-H and N-H groups). Vibrations associated to strong chemical bonds occurring between heavier atoms (e.g., C=O) can also be detected.

Figure 4.5 (*Bruker®*, Andriazzi et al., 2023) shows spectral bands in the NIR region corresponding to overtones and combinations of fundamental vibrational transitions, principally related to C-H, N-H, O-H and S-H groups.

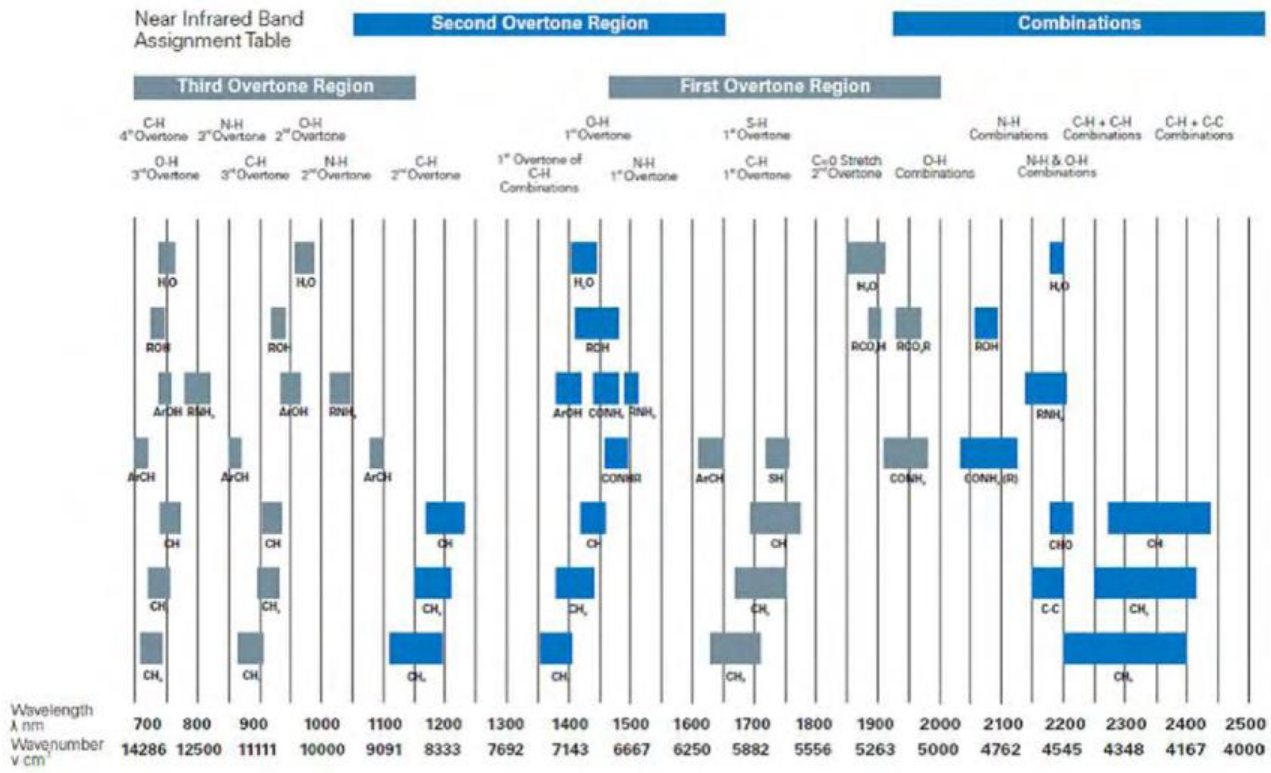


Figure 4.5 *NIR bands assignment table (Bruker®, Andriazzi et al., 2023).*

The intensity of absorption in the NIR region is related to the degree of anharmonicity of the vibrational mode and the resulting change in the dipole moment. However, NIR spectroscopy generally has lower sensitivity compared to MIR, indeed the absorption coefficients in the NIR region are 10-100 times lower than MIR region. Despite this limitation, the relatively brighter radiations sources and efficient detectors available in NIR spectroscopy help to mitigate this issue (Pasquini,

2018). For these reasons, NIR spectroscopy stands out as a valuable technique for the analysis of organic materials in the food, feed and agricultural sectors.

The characteristics of the final NIR signal depend on the chemical and physical properties of the specific sample. In addition, the NIR spectrum of a sample can be influenced by different environmental factors, ranging from microscopic (e.g., substitution of atoms at the molecular level) to macroscopic level (e.g., mechanical stress, temperature changes or alteration in the macroscopic arrangement). This versatility enhances the broad applicability of NIR spectroscopy and its potential for NIR-based analytical methods (Williams and Norris, 2001). NIR can thus be considered as a sensitive source of analytical information. As the complexity of the sample composition increases, especially due to micro- and macroscopic interactions among the constituent species, the corresponding NIR spectrum also becomes more complex (Pasquini, 2018).

Historically, after Herschel findings, the application of NIR spectroscopy for analytical purposes had been almost ignored for more than 150 years. Many absorptions overlap in the NIR region, resulting in spectra that are typically composed by broad, highly overlapping and low-intensity overtone and combination peaks. This characteristic hindered the interpretation and attribution of NIR absorption bands for a long time, especially in comparison to MIR. Moreover, the notable delay in NIR's application can also be attributed to the lack of suitable commercial instruments and the necessary mathematical resources to extract the meaningful analytical information. Even if some application studies began in the early 20th century (Ellis and Bath, 1938; Evans et al., 1951; Spakowski et al., 1950), the analytical potential of this spectral region did not gain widespread attention until the 1960s. The pioneering work carried out by Karl Norris at the U.S. Department of Agriculture (USDA, Beltsville, USA) highlighted the practical utility of this approach in the industrial field (Ben-Gera and Norris, 1968; BEN-GERA and NORRIS, 1968; Norris, 1965; Norris and Hart, 1996). The application of NIR spectroscopy in quality control and industrial processes took further steps forward since the 1970s, with the development of effective chemometric techniques for extracting and processing analytical data, as well as innovations in instrumentation for signal acquisition. The widespread use of personal computers in the 1980s further accelerated the application of this technique.

Nowadays, modern multivariate data analysis techniques and advanced instrumentation allow for the extraction of considerable information from NIR spectra regarding molecular and physical structure of samples. Since then, NIR spectroscopy has been gaining an increasing popularity in many industrial sectors, especially in the agri-food industry, for applications related to food quality evaluation and process monitoring. In fact, this method allows to address both quantitative (i.e., determination of a specific analyte in a given matrix by regression analysis) and qualitative (i.e.,

classification of samples into different categories based on the specific qualitative chemical characteristics) issues (Liu et al., 2008; Urbano-Cuadrado et al., 2004).

It is important to note that NIR spectroscopy is a fingerprint technique, i.e., it is a relative method. This is of particular importance when developing calibration or classification models, whose accuracy and robustness can only be achieved by considering a large number of representative samples to capture all potential chemical and/or physical variations. Besides, due to the sensitivity of this technique, NIR spectroscopy is principally recommendable for determining major components in a sample.

4.2.1 Relevance of NIR spectroscopy in food analysis

Even if multiple techniques are available for the analysis of complex matrices, NIR spectroscopy often stands out as a compelling and recommendable choice due to the multiple advantages it offers. These include its speed, non-invasive and non-destructive nature, and versatility for the analysis of different agri-food products (Armenta et al., 2010; Barbin et al., 2014; Holroyd, 2013; Holroyd et al., 2013; Jha et al., 2010; López et al., 2013; Mlček et al., 2010; Prieto et al., 2017, 2009; Stella et al., 2015; Sun et al., 2010; Woodcock et al., 2008; Zábrodská and Vorlová, 2015). Its non-destructive nature represents a clear advantage when the amount of available sample is limited and/or requires further analysis. The versatility of NIR spectroscopy allows for the simultaneous determination of many chemical constituents in most food products. Besides, sample preparation is generally not required, or is quite simple and quick, independently of the solid or liquid state of the sample. This simplicity translates into time and costs savings, as no reagents, materials or extensive procedures are usually required for sample pretreatment, unlike traditional analytical techniques which often involve time-consuming and complex procedures. From a practical point of view, the acquisition of spectra is a relatively simple and fast procedure, requiring only minimal training for operators.

Regarding the applicability and the adoption of NIR spectroscopy in real-case scenarios, the automated and reliable extraction of useful information from NIR spectra represents a crucial point. This issue is addressed by applying proper multivariate analysis methods, which are now effectively supported by modern computing capabilities. Within this framework, the combination of NIR spectroscopy and chemometrics represents a powerful tool for analysing the chemical properties of many food samples, as well as to extract the relevant information for classification and calibration in food-related applications.

Thanks to the recent technological advancements, nowadays some food-related applications could also be successfully handled through cheaper, portable and commercially available devices working

in narrow spectral regions (*see* Section 4.3). This development is particularly relevant for industrial applications of NIR spectroscopy where on-site analysis is required.

The relevance of NIR spectroscopy in the agri-food sector is also reflected in the growing volume of research conducted in the last two decades. Figure 4.6 shows the trend of the production of publications (namely, articles, conference papers, reviews, book chapters, books, conference reviews and editorials) from 2000 to 2024. These publications include terms such as "NIR", "Near Infrared spectroscopy", "Near-infrared spectroscopy" or "NIR spectroscopy", alongside "food" or "beverage" in the article title, abstract and keywords. The plot shows the number of total publications recorded for each year, along with the corresponding number of citations. A total of 5237 documents has been found, with both production and citation of publications exhibiting a constant upward trend through 2024, reaching a maximum peak of 656 publications and 22326 citations in Scopus.

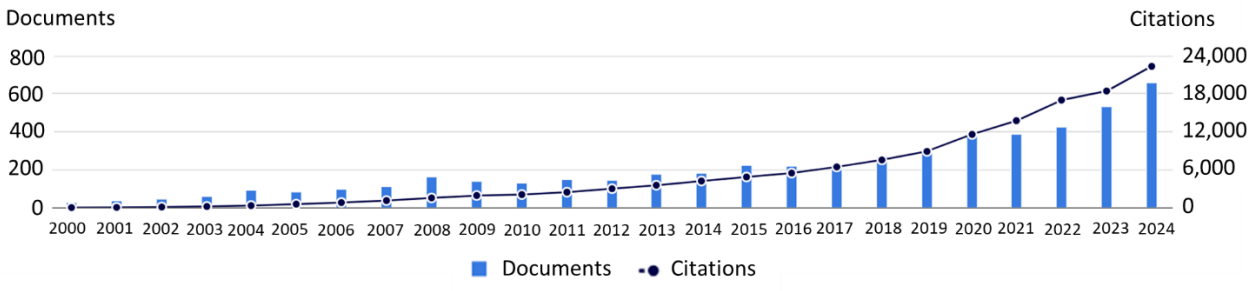


Figure 4.6 *Total number of publications related to NIR and food, and corresponding citations from 2000-2024. Source: Scopus.*

4.3 Instrumental settings for NIR spectroscopy

NIR spectrophotometers typically consist of the following components:

- *Light source*: Emits radiation within the spectral range of interest. In case of NIR spectroscopy, tungsten-halogen lamps with quartz windows (emitting a continuous spectrum in the 320-2500 nm range) and LEDs (Light Emitting Diodes) are commonly used;
- *Optical bench*: This includes the wavelength selector. NIR instrumentation can be categorized as follows based on the type of wavelength selector used:
    - *Dispersive spectrophotometers*: The wavelength selector consists traditionally of a diffraction grating or other optical devices that allow dispersing the input beam by separating the different wavelengths in different directions;

  - *Fourier Transform (FT) spectrophotometers*: The wavelength selector uses an interferometer instead of the dispersive device. FT-NIR instruments offer several advantages, including high speed, reliability, convenience and higher signal-to-noise ratio. For all these reasons, high-performance FT instruments have widely replaced dispersive spectrophotometers;
- *Sample holder*: The structure and characteristics of the sample holder strictly depend on the sampling procedure used for the analysis. Indeed, several methods can be used to acquire NIR spectra;
- *Light detector:* Measures the intensity of the radiation after interaction with the sample. NIR detectors are typically semiconductor-based devices:
  - *Lead sulphide (PbS) detectors*: Commonly used, thermoelectrically cooled (TE)-PbS photoconductors can detect radiation in the 780-2780 nm range (12800-3600 cm^{-1}).
  - *Indium Arsenide (InAs) and Indium Gallium Arsenide (InGaAs) detectors*: These are often used for transmission measurements of solid matrices at room temperature, detecting radiation in the 780-1725 nm range (12800-5800 cm^{-1}) (Workman and Burns, 2001). Thermoelectrically cooled (TE)-InGaAs detectors can collect the full NIR range, i.e., 780-2500 nm (12800-4000 cm^{-1}).
  - *Silicon (Si) Diodes*: Specialized detectors capable of collecting radiation in the 650-1100 nm range (15500-9000 cm-^{1}) (MPA user manual, 2007).
- *Computer*: Amplifies, converts in digital format and records the analytical signal, displaying the results for interpretation.

Modern benchtop NIR spectrophotometers often rely on cooled InGaAs detectors (Pasquini, 2018). However, in the last decades recent technological advancements have enabled the use of less expensive sensors (like an uncooled InGaAs detector or silicon detectors) that work within narrower spectral regions. These advancements have facilitated the development of portable and cost-effective devices, now commercially available on the market (Table 4.1).

Many portable instruments have been evaluated for applications traditionally handled by conventional benchtop spectrophotometers. While portable devices sometimes exhibit slightly lower performance, they have shown satisfactory results for many applications, particularly in the agri-food sector (Basri et al., 2017; Bellincontro et al., 2012; Cayuela and Weiland, 2010; Giussani et al., 2021; Gorla et al., 2022; Malegori et al., 2017; Pasquini, 2018; Riu et al., 2020; Tardaguila et al., 2017).

Table 4.1 *NIR spectral regions considered in this PhD Thesis and corresponding types of sensors and instrumentation available on the market.*

Spectral range	**Detector**	**Type of available instrumentation**
800-2500 nm (12500-4000 cm^{-1})	cooled InGaAs detectors	expensive bench-top instruments
950-1650 nm (10526-6060 cm^{-1})	Cooled and uncooled InGaAs detectors	high-end portable instruments
800-1070 nm (12500-9346 cm^{-1})	silicon and (cooled and uncooled) InGaAs detectors	economical handheld instruments

4.3.1 FT spectrophotometers

The development of Fourier-transform infrared (FT-IR) spectrometers in the 1970s revolutionized infrared spectroscopy. Unlike dispersive instruments, FT-IR spectrometers measure the entire IR spectrum simultaneously using an interferometer. The raw signal collected by FT-IR instruments is an interferogram, which represents the intensity of IR light as a function of the position of the moving mirror in the interferometer. To produce the IR spectrum, Fourier Transform is applied.

The advent of FT-IR brought significant improvements to IR spectroscopy, including speed, accuracy and reproducibility of the analysis, and an improved signal-to-noise ratio. These advantages have made FT-IR a dominant technology in IR spectroscopy, effectively replacing traditional dispersive instruments in most applications.

Fourier-transform infrared (FT-IR) spectroscopy is based on the principle of interferometry. This technique takes advantage of waves properties and their interaction to achieve information about a sample.

Like all electromagnetic radiation, IR light travels in waves characterized by peaks and troughs, whose position in space define the wave's phase. Two waves may have different phases, resulting in a phase difference (Figure 4.7a) which determines how the waves interact, a concept known as interference. When peaks and troughs of two waves align, they are in phase, producing constructive interference (Figure 4.7b). When peaks of one wave coincide with the troughs of another, they are out of phase, leading to destructive interference (Figure 4.7d). For all other cases, the interaction produces an interference pattern depending on the phase difference (Figure 4.7c).

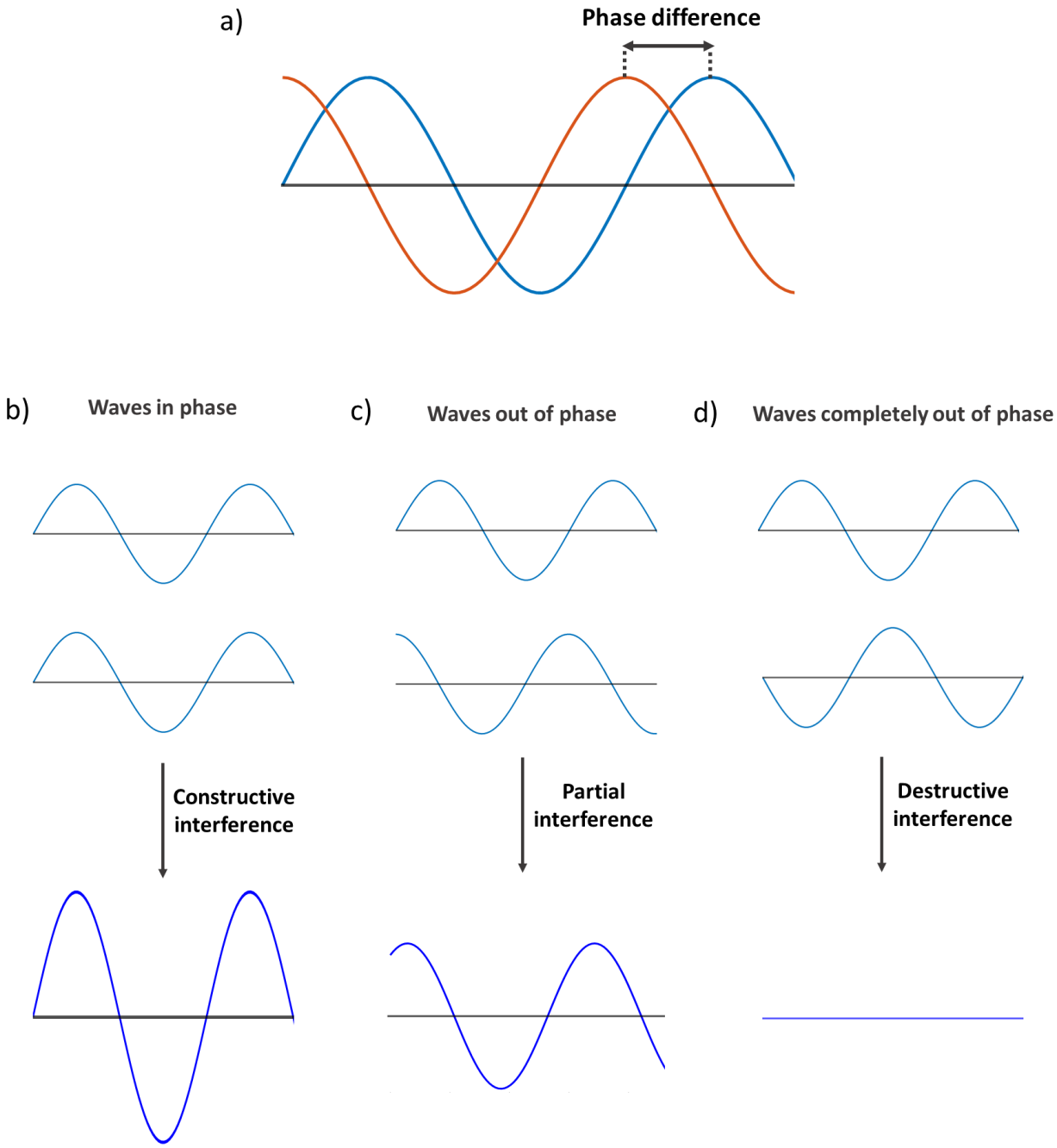


Figure 4.7 *Example of a) phase difference and b)-d) types of interference patterns between two waves.*

FT-IR spectroscopy uses an interferometer, such as the Michelson interferometer, which causes IR light to interfere with itself.

The Michelson interferometer consists of key components (Harris, 2017): a beamsplitter splits the incoming light into two beams directed towards two mirrors; one mirror is stationary, while the other is movable; the beams reflect back to the beamsplitter and recombine (Figure 4.8). If the two beams travel the same distance, they recombine in phase, generating a constructive interference. However,

by moving the movable mirror, the path length of one beam is altered relative to the other. This creates a phase difference, or delay (δ), between the beams. The delay modulates the interference pattern, which can vary from constructive interference (if $\delta = k\,\lambda$, where λ is the wavelength and k is an integer number) to destructive interference (if $\delta = \frac{1}{2}\,k\,\lambda$).

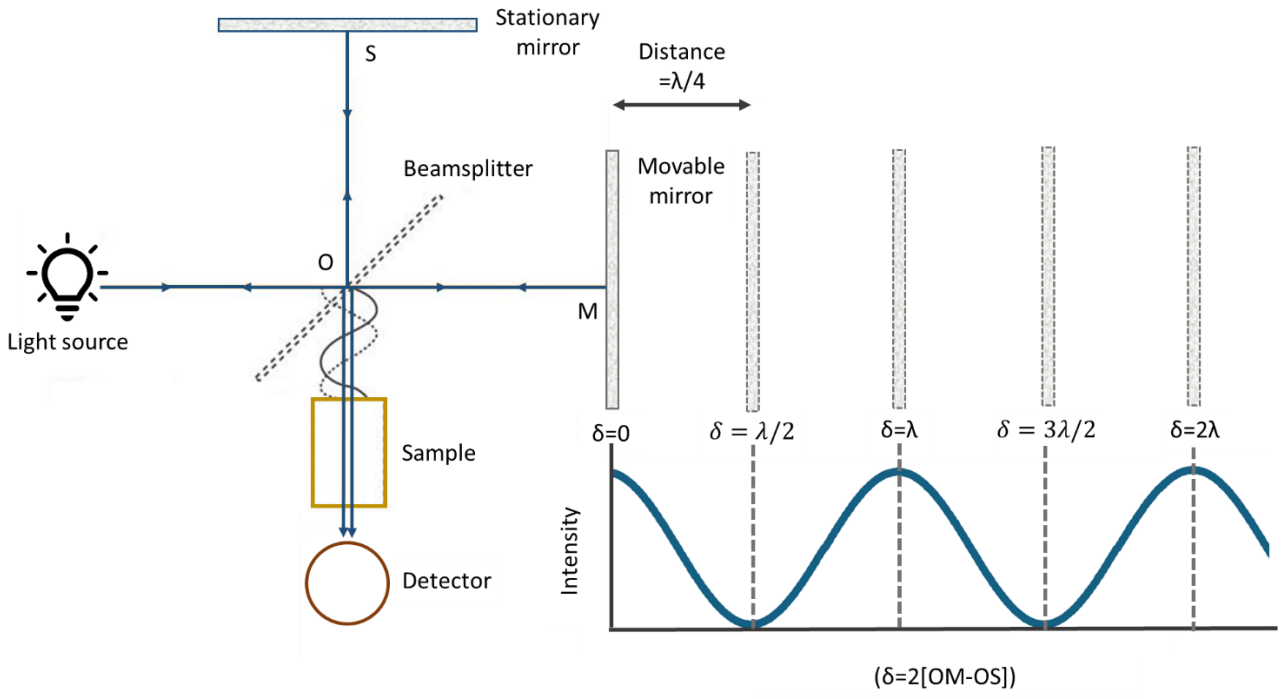


Figure 4.8 *Michelson interferometer; light waves from mirrors (M and S) are usually out of phase with each other; beams from M and S are shown separated, but they are actually superimposed.*

As the mirror moves at constant speed, the emerging light directed at the detector undergoes alternating constructive and destructive interference, producing a tuning of the radiation intensity. For each position of the moveable mirror, every wavelength in the IR region exhibits a unique interference pattern. The recombined light, which contains information about all wavelength simultaneously, is directed at the sample which absorbs specific wavelengths based on its chemical characteristics. The transmitted (or reflected, when performing reflectance measurements) light is detected, producing data on the variation in intensity of transmitted or reflected radiation as a function of the delay for all wavelengths simultaneously.

The interferometer produces an interferogram, which reports the intensity of recombined light against the delay (in the time domain). Fourier-transform (FT) allows then to convert the signal from the time domain to the frequency domain. This transform decomposes the interferogram in its frequency

components, extracting individual frequencies from the interference patterns. The result is a standard IR spectrum, reporting the absorbance values at each wavenumber.

FT-IR spectroscopy offers significant advantages over dispersive IR spectroscopy:

- Speed: rapid data acquisition due to simultaneous analysis of all wavelengths;
- High-quality spectra: excellent signal-to-noise ratio;

Interferometry is the technology embodied in many of modern commercial NIR instruments. FT-NIR spectroscopy has widely replaced traditional wet chemical analysis methods due to its versatility, non-destructive nature, speed and cost-effectiveness. Applications are widespread in the agri-food sector, including process monitoring and quality control of various food products (e.g., edible oils, dairy, meat, beverages, bakery ingredients, condiments, grains, seeds, feed and forage). FT-NIR enables both qualitative and quantitative analysis, making it an indispensable tool in food science and industry.

4.4 References

Andriazzi, C.V.G., Rocha, D.K., Custódio, C.C., 2023. Determination of the physiological quality of corn seeds by infrared equipment. Journal of Seed Science 45, e202345002.

Armenta, S., Moros, J., Garrigues, S., Guardia, M.D.L., 2010. The use of near-infrared spectrometry in the olive oil industry. Critical Reviews in Food Science and Nutrition 50, 567–582.

Barbin, D.F., Felicio, A.L.D.S.M., Sun, D.-W., Nixdorf, S.L., Hirooka, E.Y., 2014. Application of infrared spectral techniques on quality and compositional attributes of coffee: An overview. Food Research International 61, 23–32. https://doi.org/10.1016/j.foodres.2014.01.005

Barton, F.E., 2002. Theory and principles of near infrared spectroscopy. Spectroscopy Europe 14(1), 12–18.

Basri, K.N., Hussain, M.N., Bakar, J., Sharif, Z., Khir, M.F.A., Zoolfakar, A.S., 2017. Classification and quantification of palm oil adulteration via portable NIR spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 173, 335–342. https://doi.org/10.1016/j.saa.2016.09.028

Bellincontro, A., Taticchi, A., Servili, M., Esposto, S., Farinelli, D., Mencarelli, F., 2012. Feasible application of a portable NIR-AOTF tool for on-field prediction of phenolic compounds during the ripening of olives for oil production. Journal of agricultural and food chemistry 60, 2665–2673.

Ben-Gera, I., Norris, K.H., 1968. Determination of moisture content in soyabeans by direct spectrophotometry.

BEN-GERA, I., NORRIS, K.H., 1968. Direct spectrophotometric determination of fat and moisture in meat products. Journal of Food Science 33, 64–67.

Burns, D.A., Ciurczak, E.W. (Eds.), 2007. Handbook of Near-Infrared Analysis, 0 ed. CRC Press. https://doi.org/10.1201/9781420007374

Cayuela, J.A., Weiland, C., 2010. Intact orange quality prediction with two portable NIR spectrometers. Postharvest Biology and Technology 58, 113–120. https://doi.org/10.1016/j.postharvbio.2010.06.001

Coates, J., 2000. Interpretation of infrared spectra, a practical approach. Encyclopedia of analytical chemistry 12, 10815–10837.
Ellis, J.W., Bath, J., 1938. Alterations in the infrared absorption spectrum of water in gelatin. The Journal of Chemical Physics 6, 108–108.
Evans, A., Hibbard, R., Powell, A., 1951. Detrmination of Carbon-Hydrogen Groups in High Molecular Weight Hydrocarbons by Near-Infrared Absortion. Analytical Chemistry 23, 1604–1610.
Giussani, B., Escalante-Quiceno, A.T., Boqué, R., Riu, J., 2021. Measurement strategies for the classification of edible oils using low-cost miniaturised portable NIR instruments. Foods 10, 2856.
Gorla, G., Fumagalli, S., Jansen, J.J., Giussani, B., 2022. Acquisition strategies for fermentation processes with a low-cost miniaturized NIR-spectrometer from scratch: Issues and challenges. Microchemical Journal 183, 108035. https://doi.org/10.1016/j.microc.2022.108035
Harris, D.C., 2017. Chimica analitica quantitativa, Terza edizione italiana condotta sulla nona edizione americana. ed. Zanichelli.
Herschel, W., 1800. Experiments on the Refrangibility of the Invisible Rays of the Sun.
Holroyd, S.E., 2013. The use of near infrared spectroscopy on milk and milk products. Journal of Near Infrared Spectroscopy 21, 311–322.
Holroyd, S.E., Prescott, B., McLean, A., 2013. The use of in-and on-line near infrared spectroscopy for milk powder measurement. Journal of Near Infrared Spectroscopy 21, 441–443.
Jha, S., Narsaiah, K., Sharma, A., Singh, M., Bansal, S., Kumar, R., 2010. Quality parameters of mango and potential of non-destructive techniques for their measurement—a review. Journal of food science and technology 47, 1–14.
Liu, L., Cozzolino, D., Cynkar, W., Dambergs, R., Janik, L., O'neill, B., Colby, C., Gishen, M., 2008. Preliminary study on the application of visible–near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. Food chemistry 106, 781–786.
López, A., Arazuri, S., García, I., Mangado, J., Jarén, C., 2013. A review of the application of near-infrared spectroscopy for the analysis of potatoes. Journal of agricultural and food chemistry 61, 5413–5424.
Malegori, C., Nascimento Marques, E.J., De Freitas, S.T., Pimentel, M.F., Pasquini, C., Casiraghi, E., 2017. Comparing the analytical performances of Micro-NIR and FT-NIR spectrometers in the evaluation of acerola fruit quality, using PLS and SVM regression algorithms. Talanta 165, 112–116. https://doi.org/10.1016/j.talanta.2016.12.035
Miller, C.E., 2001. Chemical principles of near-infrared technology. Near-infrared technology in the agricultural and food industries 2.
Mlček, J., Rop, O., Šustová, K., Simeonovova, J., Gal, R., 2010. Potentials of NIR spectroscopy in meat industry. Chemické listy 104.
MPA user manual, 2007. Bruker Optics FT-NIR spectrophotometer user manual, 2nd updated edition.
Norris, K., 1965. Direct spectrophotometric determination of moisture content of grains and seeds. Principles and methods of measuring moisture content in liquides and solids 4, 19–25.
Norris, K.H., Hart, J.R., 1996. 4. Direct spectrophotometric determination of moisture content of grain and seeds. Journal of Near Infrared Spectroscopy 4, 23–30.
Pasquini, C., 2018. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. Analytica Chimica Acta 1026, 8–36. https://doi.org/10.1016/j.aca.2018.04.004
Prieto, N., Pawluczyk, O., Dugan, M.E.R., Aalhus, J.L., 2017. A review of the principles and applications of near-infrared spectroscopy to characterize meat, fat, and meat products. Applied spectroscopy 71, 1403–1426.
Prieto, N., Roehe, R., Lavín, P., Batten, G., Andrés, S., 2009. Application of near infrared reflectance spectroscopy to predict meat and meat products quality: A review. Meat science 83, 175–186.

Riu, J., Gorla, G., Chakif, D., Boqué, R., Giussani, B., 2020. Rapid Analysis of Milk Using Low-Cost Pocket-Size NIR Spectrometers and Multivariate Analysis. Foods 9, 1090. https://doi.org/10.3390/foods9081090

Spakowski, A., Evans, A., Hibbard, R., 1950. Determination of Aromatics and Olefins in Wide Boiling Petroleum Fractions. Analytical Chemistry 22, 1419–1422.

Stella, E., Moscetti, R., Haff, R.P., Monarca, D., Cecchini, M., Contini, M., Massantini, R., 2015. Recent advances in the use of non-destructive near infrared spectroscopy for intact olive fruits. Journal of Near Infrared Spectroscopy 23, 197–208.

Sun, T., Huang, K., Xu, H., Ying, Y., 2010. Research advances in nondestructive determination of internal quality in watermelon/melon: A review. Journal of Food Engineering 100, 569–577. https://doi.org/10.1016/j.jfoodeng.2010.05.019

Tardaguila, J., Fernández-Novales, J., Gutiérrez, S., Diago, M.P., 2017. Non-destructive assessment of grapevine water status in the field using a portable NIR spectrophotometer. J Sci Food Agric 97, 3772–3780. https://doi.org/10.1002/jsfa.8241

Urbano-Cuadrado, M., De Castro, M.L., Pérez-Juan, P., García-Olmo, J., Gómez-Nieto, M., 2004. Near infrared reflectance spectroscopy and multivariate analysis in enology: Determination or screening of fifteen parameters in different types of wines. Analytica Chimica Acta 527, 81–88.

Williams, P., Norris, K., 2001. Near infrared technology in the agricultural and food industries. St. Paul, Minnesota, USA: American Association of Cereal Chemists.

Woodcock, T., Fagan, C.C., O'Donnell, C.P., Downey, G., 2008. Application of near and mid-infrared spectroscopy to determine cheese quality and authenticity. Food and Bioprocess Technology 1, 117–129.

Workman, J., Burns, D.A., 2001. Commercial NIR instrumentation. In Burns, D.A., and Ciurczak, E.W., (Eds) Handbook of Near Infrared analysis (2nd Ed.), Marcel Dekker Inc, New York 53–70.

Zábrodská, B., Vorlová, L., 2015. Adulteration of honey and available methods for detection–a review. Acta Veterinaria Brno 83, 85–102.

Chapter Five

Application of chemometrics to image and spectral data in the agri-food sector

This chapter reports the application of analytical methods based on multivariate statistics for the analysis of RGB images and NIR spectral data to address different challenges related to the wine and vinegar production sector.

More in detail, this chapter is organized as follows:

- **Section 5.1** reports the application of the *colourgrams* approach on RGB images of red grape samples to predict the corresponding anthocyanins content by means of multivariate calibration models. Images were acquired using a smartphone-based portable device, conceived for the analysis of RGB images of red grapes directly in the field, that was optimized during the current PhD;
- **Section 5.2** reports the development and application of the *texturegrams* approach, an innovative method to extract and codify the texture-related information contained in RGB images. The effectiveness of this approach was evaluated and compared with state-of-the-art methods to codify the texture-related features of images. A subset of the RGB images reported in Section 5.1 was used as benchmark dataset, considering that texture-related information could help to improve the predictive ability of the calibration models that were previously developed based solely on colour-related information;
- **Section 5.3** reports the application of NIR spectroscopy coupled with chemometrics to monitor alcoholic and acetic fermentation in red grape musts, comparing the results that could be potentially obtained with both benchtop and portable devices available on the market and operating in different NIR spectral regions.

5.1 Design and application of a smartphone-based device for in vineyard determination of anthocyanins content in red grapes

What follows is the integral content of: Menozzi, C., Calvini, R., Nigro, G., Tessarin, P., Bossio, D., Calderisi, M., Ferrari V., Foca G., Ulrici, A. (2023). Design and application of a smartphone-based device for in vineyard determination of anthocyanins content in red grapes. *Microchemical Journal*, *191*, 108811. doi: 10.1016/j.microc.2023.108811

Design and application of a smartphone-based device for in vineyard determination of anthocyanins content in red grapes

Camilla Menozzi [a,], Rosalba Calvini [a,], Giovanni Nigro [b], Paola Tessarin [b], Domenico Bossio [b], Marco Calderisi [c], Veronica Ferrari [a], Giorgia Foca [a], Alessandro Ulrici [a]

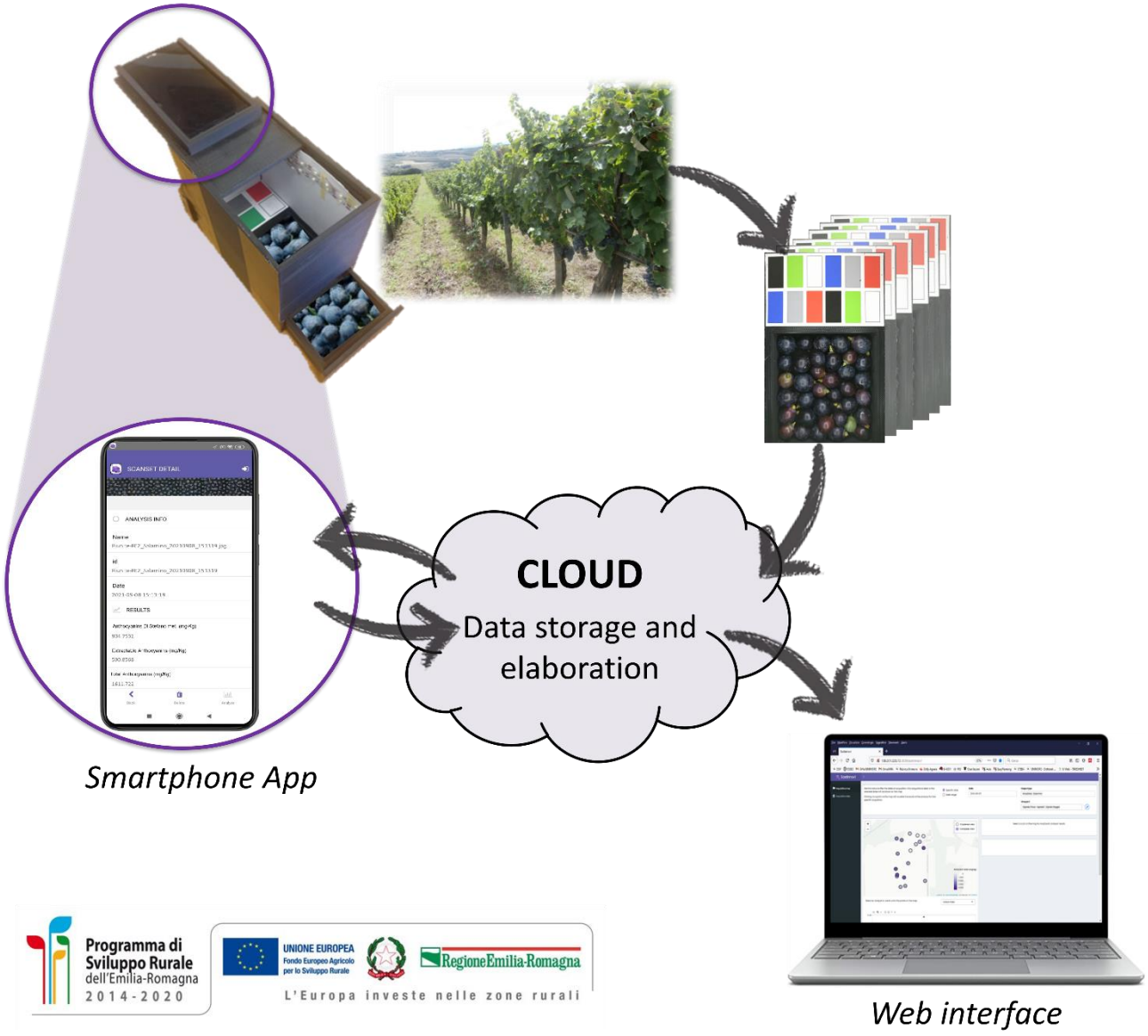
[a] *Dipartimento di Scienze della Vita, Università di Modena e Reggio Emilia, Via Amendola, 2, 42122, Reggio Emilia, Italy*
[b] *Ri.Nova – Filiera Vitivinicola ed Olivo-Oleicola, Via Tebano 45, 48018, Faenza, Italy*
[c] *Kode Solutions, Lungarno Galileo Galilei, 1, 56125, Pisa, Italy*

Abstract

The choice of the proper moment for harvesting is a crucial aspect in winemaking process, since the chemical attributes of grape berries strongly influence red wine quality. In particular, phenolic composition of red grapes plays a significant role in many sensory properties of wine related to colour and taste. Anthocyanins are the most important phenolic compounds for red grapes: they accumulate in the grape skin during ripening, and they are responsible for the purple colour of ripe berries. Routine analysis for the determination of grapes phenolic maturity includes chromatographic and spectroscopic techniques, that are time-consuming and expensive. In this work, we propose an innovative device conceived for the determination of grape phenolic maturity based on RGB images of grape berries acquired with a smartphone. The device has been designed to be used directly in the vineyard thanks to its small size and to the possibility of acquiring geolocated images of the berries under controlled lighting conditions. In this study, grape samples of three different varieties (*Ancellotta*, *Lambrusco Salamino* and *Sangiovese*) were collected at different harvest times from veraison to maturity and imaged by means of a common smartphone using the device. The RGB images were then converted into one-dimensional signals, named colourgrams, which codify the colour properties of the images. The dataset of colourgrams was then used to calculate calibration models using Partial Least Squares (PLS) regression in order to relate colour information with

chemical parameters generally employed to evaluate grape phenolic maturity, such as total anthocyanins content and extractable anthocyanins content. The calibration models were implemented in a software interface that allows to acquire geolocated images of the grape samples, visualize the outcomes of the analysis, visualize maps and plots related to phenolic maturity, store data and share relevant information.

Graphical abstract



Keywords: RGB imaging, grape ripening, multivariate image analysis, anthocyanins, smartphone.

1 Introduction

Grape harvesting at the optimal maturity stage is fundamental in wine production, since wine quality is strongly influenced by the amount of chemical compounds contained in the berries and the extractability of these compounds during the winemaking process (Bautista-Ortín et al., 2006; Vrochidou et al., 2021). The simple assessment of the so-called technological maturity, i.e., the ratio between sugars and acids content, is not sufficient to fully predict the oenological potential of grapes, since wine properties also depend on phenolic and aromatic maturity. Phenolic maturity refers to the potential of phenolic compounds, in particular anthocyanins present in grapes, while aromatic maturity corresponds to the maximum concentration of volatile compounds affecting wine aroma. Technological, phenolic and aromatic ripening processes have different behaviours and each one needs to be evaluated in order to properly assess oenological maturity of grapes and identify the optimal harvest time (Ribéreau-Gayon et al., 2006).

In particular, with regard to red grapes the concept of phenolic maturity has gained an increasing importance since it is related to the presence and abundance of phenolic compounds, which have a significant role in many sensory properties of red wines, like colour stability, bitterness, astringency, body, and structure (Chen et al., 2015; García-Beneytez et al., 2002; Kontoudakis et al., 2011; Pérez-Magariño and González-San José, 2006, 2004; Ribéreau-Gayon et al., 2006).

Anthocyanins are one of the most relevant groups of phenolic compounds. They are the red pigments in grapes which are responsible for the colour of ripe berries and red wine (García-Beneytez et al., 2002; Kennedy et al., 2002; Mattivi et al., 2002; Ribéreau-Gayon et al., 2006; Rodríguez-Pulido et al., 2022). The initial colour of red wine is due predominantly to anthocyanins extracted from grape skin during crushing, pressing and fermentation, while final colour depends on the formation of condensed pigments between free anthocyanins and colourless grape phenols, specifically, monomeric and polymeric flavanols (catechins and procyanidins), flavonols and phenolic acids during storage and ageing (Garciafalcon et al., 2007; Rodríguez-Pulido et al., 2022).

Anthocyanins gradually appear at veraison and accumulate primarily in the grape skin throughout the ripening process, while they tend to decrease during over-ripening (Kennedy et al., 2002; Ribéreau-Gayon et al., 2006). Grapes do not ripen homogeneously and even each berry of the same bunch ripens at different rates due to multiple causes (Kontoudakis et al., 2011). Indeed, different factors, like soil characteristics, climatic conditions, grape variety, agronomic practices and grape maturity level could affect the phenolic composition of grapes (Pérez-Magariño and González-San José, 2004; Rodríguez-Pulido et al., 2022).

Furthermore, the ongoing climatic change is an aspect not to be neglected (Drappier et al., 2019). Unusual seasonal trends and extreme events are becoming increasingly frequent, and they have a significant impact on rural conditions. Consequently, the amount of phenolic compounds in red grapes could be heterogeneous even within the same physiological stage (Nogales-Bueno et al., 2015). At the same time, having the optimal anthocyanins content at harvest time is important for wineries to obtain a quality product (Chen et al., 2015). In this context, vineyards belonging to the same winery but located in different geographical positions require to be frequently inspected in order to monitor grape ripening and real-time feedback from the field are needed to make proper oenological choices.

Currently, qualitative and quantitative anthocyanins analysis is performed using chromatographic and spectroscopic techniques, which are time-consuming, invasive and expensive (Gutiérrez et al., 2019; Ribera-Fonseca et al., 2016). Indeed, these methods require a long and complex extraction procedure performed by trained personnel, and the use of dedicated laboratory instrumentation. For this reason, alternative approaches have been proposed with the aim of providing a rapid and comprehensive overview of the ripening process using non-destructive methods and portable devices (Baca-Bocanegra et al., 2019; dos Santos Costa et al., 2019; Pampuri et al., 2021; Ribera-Fonseca et al., 2016). Since the colour of the berries is strongly related to the anthocyanins content, computer vision coupled with advanced techniques of image elaboration could be a rapid, non-destructive, cheap and environment-friendly alternative for the evaluation of phenolic maturity, thanks to the ease of use and flexibility of this tool (Vrochidou et al., 2021).

More in general, RGB image analysis techniques have gained an increasing attention in the field of food analysis, since the progress in image acquisition technology has made nowadays possible a detailed non-destructive colour evaluation even in case of matrices with heterogeneous aspect, like most of the food products (Antonelli et al., 2004; Brosnan and Sun, 2004, 2002; Calvini and Pigani, 2022; Ulrici et al., 2012). Furthermore, one of the main advantages of computer vision systems consists in the possibility of acquiring RGB images of the considered samples using dedicated instrumentation as well as objects of common use such as digital cameras, flatbed scanners or smartphones.

The high complexity of RGB images data arrays makes essential to identify proper strategies for the extraction of the useful information, in particular when multiple images have to be simultaneously analysed (Calvini and Pigani, 2022; Gonçalves Dias Diniz, 2020; Prats-Montalbán et al., 2011; Prats-Montalbán and Ferrer, 2007). To solve this issue, the colourgrams approach has been successfully applied in several studies related to the analysis of food matrices by means of RGB imaging (Antonelli et al., 2004; Borin et al., 2007; Corti et al., 2019; Foca et al., 2011; Giraudo et al., 2018; Giorgia

Orlandi et al., 2018; G. Orlandi et al., 2018; Ulrici et al., 2012). Colourgrams are one-dimensional signals codifying the colour-related information of each RGB image and these signals can be analysed by means of common multivariate analysis techniques (Foca et al., 2011; Giraudo et al., 2018). In two recent papers, some of us successfully proposed the use of colourgrams, both alone and in combination with electrochemical signals, to quantitatively assess parameters related to grape ripening by analysing grape must (G. Orlandi et al., 2018; Orlandi et al., 2019).

In the present study, we conceived a portable prototype in order to provide a fast and non-destructive support in the phenolic ripening assessment of grapes, by directly acquiring RGB images of grape berries using a common smartphone. The proposed system can be used directly in the vineyard with limited manipulation of the samples. The output of the analysis allows a quantitative determination of chemical parameters usually employed to assess phenolic ripening, i.e., total anthocyanins and extractable anthocyanins content. Furthermore, the use of a smartphone also allows to geolocate the images and share the results in real-time.

The prototype consists of a hardware for image acquisition (i.e., an acquisition box with a sample holder, illumination system and smartphone case) and proper software (smartphone app and web interface) for image elaboration and data storage. To our knowledge, this is the first time that RGB images of grape berries acquired with a smartphone are used to quantitatively estimate their anthocyanins content without any sample pretreatment.

In this work, grape samples of three different grape varieties (*Ancellotta*, *Lambrusco Salamino* and *Sangiovese*) were collected at different harvest times from veraison to maturity and imaged using the device. RGB images were converted into colourgrams and the dataset of colourgrams was used to build Partial Least Squares (PLS) calibration models to relate the colour information of the images with the chemical parameters used to assess their anthocyanins content.

The following Materials and methods section is structured as follows: at first, we will describe hardware and software functionalities of the device, which was conceived to be as user-friendly as possible (*see* Section 2.1), and then we will describe in detail the development of all the data elaboration steps (*see* Sections 2.2-2.5) that are implemented in the software back-end. Finally, in the Results section (*see* Section 3) we will mainly focus on the calibration models obtained in the present study.

2 Materials and methods

2.1 The device

The device developed in this study to determine phenolic maturity of grape berries combines both hardware and software components. The hardware was designed to be compact in size and to enable the acquisition of grape samples images under controlled lighting conditions; in this manner it can be easily used by the winegrower directly in the vineyard. The software includes both a smartphone app and an interface accessible by a web browser, and it allows to acquire the geolocated images, display the results in real time, perform data sharing and storage.

2.1.1 The hardware

The hardware consists of an acquisition box of size 16 cm × 10 cm × 15.5 cm, a drawer used as sample holder and a sliding lid with a smartphone case (Figure 1). All the hardware components were designed using the SketchUp software (Trimble) and printed with a 3D printer in black acrylonitrile-butadiene-styrene (ABS). For image acquisition, the smartphone is placed inside its case in the sliding lid located in the upper part of the acquisition box, while the grape samples are placed in the sample holder. The smartphone used for this study was a Xiaomi Redmi Note 8 (Xiaomi, China), therefore the case was specifically designed based on the geometry (i.e., smartphone size and camera location) of this model.

A strip of white light emitting diode (LED) lights was fixed in the internal upper edge of the acquisition box, to ensure constant and homogeneous lighting conditions. The LED lights are powered by a power bank fixed outside the acquisition case, as shown in Figure 1.

In addition, the lower inner part of the acquisition case contains a cardboard of colour references, which are imaged with the sample and used for image correction (*see* Section 2.4.2).

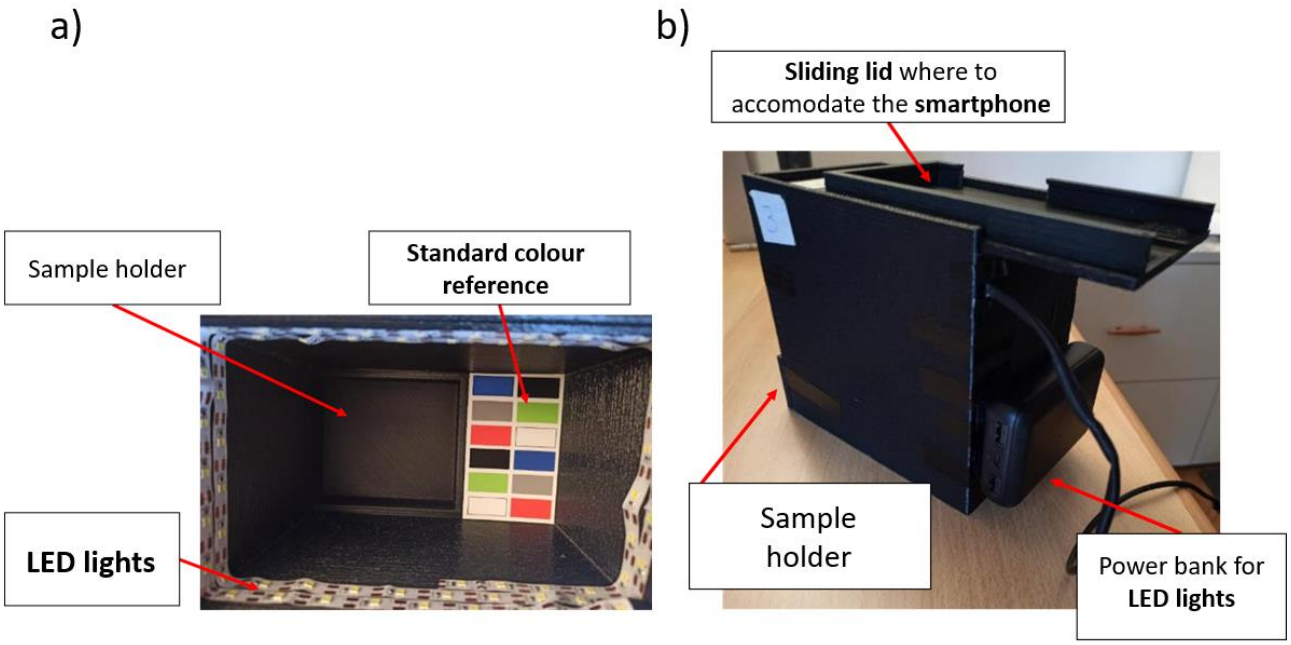


Figure 1 *Inner (a) and (b) outer structure of the acquisition device.*

2.1.2 The software

The dedicated software consists of a smartphone app and a web interface, which are connected to a web server where data are elaborated and stored.

Before providing a detailed description of the software features and for ease of comprehension, it is necessary to define some key terms/concepts, which will be used throughout the whole manuscript. In this work a *sample* consists of about 500 g of berries collected from different bunches located in different plants of the vineyard (*see* Section 2.2). For each sample, 6 *aliquots* were randomly collected (*see* Section 2.4.1) and a single *image* was acquired for each aliquot. Then, the 6 images corresponding to the 6 aliquots of each sample were merged together into a *scanset,* in order to better represent the intrinsic variability of the berries colour within the considered sample (*see* Section 2.4.3). Therefore, a *scanset* is an image representation of a whole sample, while a single image represents an individual aliquot.

The smartphone app (Figure 2) was designed to be used with ease, thanks to a friendly graphical user interface, which shows only the relevant information to the user.

From the left side of the screen an interactive list appears, where the user can select different menus (Figure 2a). The “Scan” menu allows to acquire the geolocated images of the grape samples and to set additional sample information such as grape type and vineyard (Figure 2b). According to the

experimental procedure elaborated in this study (*see* Section 2.4), for each grape sample six different images of six aliquots of berries have to be acquired and then concatenated into a *scanset* for further elaboration. To make image acquisition as much user-friendly as possible, the smartphone app was developed to acquire and handle a *scanset* for each grape sample using the "Multiple Acquisition" mode. When using the "Multiple Acquisition" mode, the app displays the current *scanset* name, the currently acquired image number (between 1 and 6), and a preview of the acquired image to make it possible to decide whether to accept or discard it in case of a wrong acquisition (Figure 2c).

Once acquired, the *scanset* is sent to the server together with the relevant metadata (geolocation, date and time) for image elaboration and storage, and the outcomes of the analysis can be displayed directly in the smartphone app using the "Analysis" menu. More in detail, as shown in Figure 2d for each analysed *scanset* it is possible to visualize sample name, time and date of acquisition, and the predicted content of the parameters of interest related to phenolic maturity: total anthocyanins and extractable anthocyanins, both measured according to Glories method (Glories, 1993; Saint Cricq de Gaulejac et al., 1998), and total anthocyanins measured according to Di Stefano method (Di Stefano and Cravero, 1991; Di Stefano and Maggiorotto, 1995).

The web interface (Figure 3) allows to access to the entire database of acquired images, considering also different grape varieties and different vineyards. In particular, it is possible to select a specific time range, one or more grape varieties and one or more vineyards, and visualize the corresponding results (Figure 3a). The results are displayed as ripening maps, where each point corresponds to a *scanset* located in a specific geographical position (coded by the GPS coordinates of the *scanset*) and the point colour indicates the corresponding value of the selected chemical parameter (Figure 3b). In this manner, it is possible to easily evaluate the ripening stage of different parts of the same vineyard and/or of different vineyards. In addition, it is also possible to observe the ripening trend over the selected time period (Figure 3c).

Finally, it is possible to select a specific sample in the ripening map, by clicking on the corresponding point, and visualize the corresponding *scanset* image together with the sample name, time and date of acquisition, and the predicted values of the chemical parameters (Figure 3d).

Both smartphone app and web interface have been designed to be easily used by winegrowers to employ the proposed device to monitor grape phenolic maturity. Therefore, the whole data elaboration workflow that goes from image acquisition to the visualization of the outcomes of the analysis is automatically handled in back-end by server computing. In the following Sections (from Section 2.2 to Section 2.5) we will provide a detailed description of the procedure that we have followed to develop the multivariate calibration models implemented in the software.

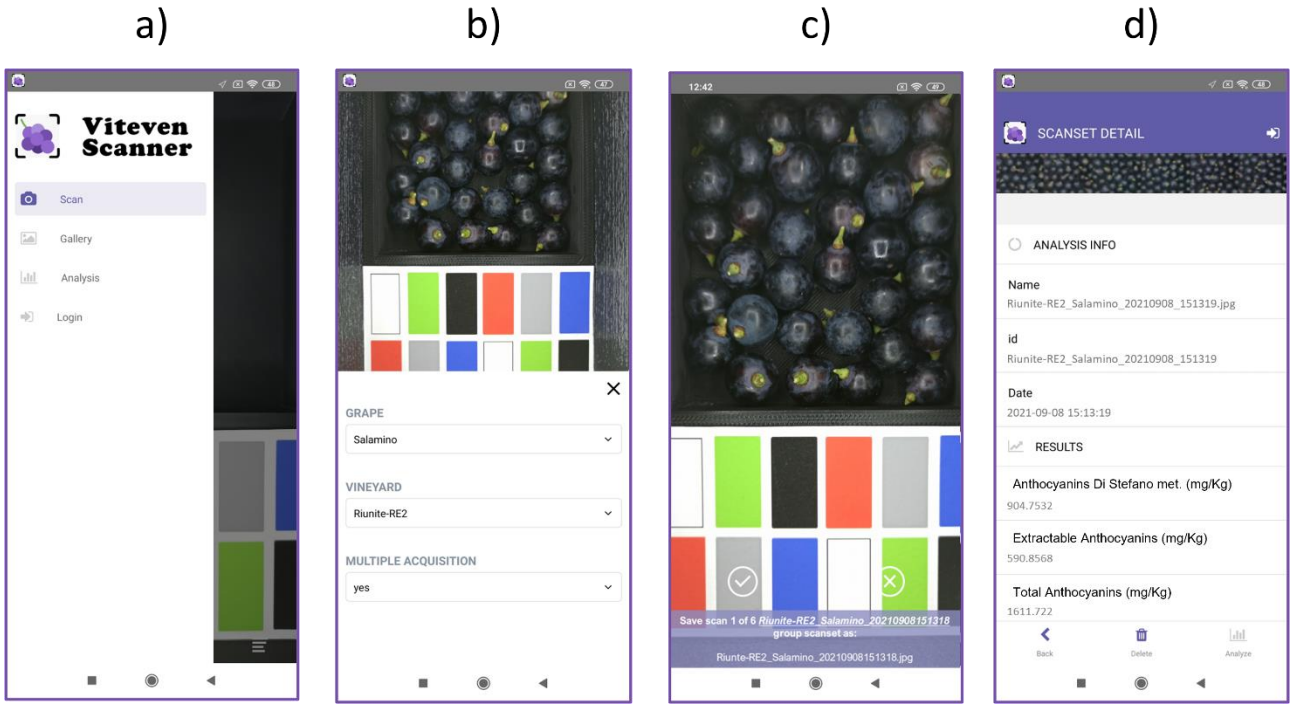


Figure 2 *Screenshots of the main features of smartphone app: a) selection of the "Scan" menu to acquire geolocated images of the grape samples; b) selection of the corresponding grape variety, vineyard and "Multiple Acquisition" mode to acquire the six images of a single sample that will be concatenated into a scanset; c) preview of the acquired image, with the possibility to discard or accept it; d) visualization of scanset details in the "Analysis" menu, showing sample name, time and date of acquisition, and predicted values of the parameters of interest for the selected scanset.*

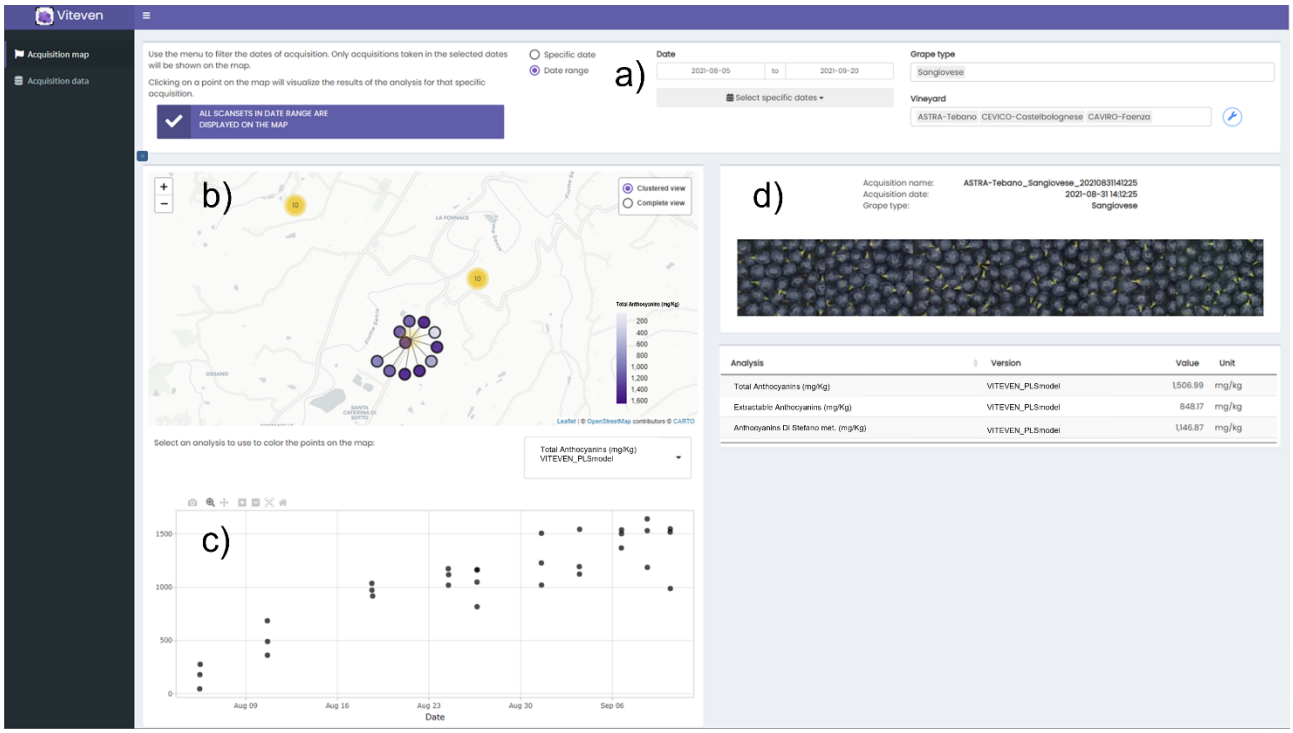


Figure 3 *Screenshot of the main features of the web interface: a) selection of date range, grape varieties and vineyards of interest; b) visualization of the ripening maps and c) ripening trend over the selected time period; d) visualization of a scanset image together with the predicted values of the chemical parameters.*

2.2 Grape samples

To evaluate the effectiveness of the proposed device, three Italian purple grape varieties were considered: *Ancellotta* (A), *Lambrusco Salamino* (L) and *Sangiovese* (S), which correspond to the most common varieties used for wine production in Emilia-Romagna region (Italy).

The grape samples were harvested during vintages 2020 and 2021 in different vineyards located in the province of Reggio Emilia for *Ancellotta* and *Lambrusco Salamino* varieties, and in the province of Ravenna for *Sangiovese* variety. In particular, two vineyards were considered for each grape variety in vintage 2020, while three vineyards for each grape variety were considered in vintage 2021. In addition, the different vineyards considered for each grape variety were located both in hilly and flat areas in order to account for pedoclimatic variability.

For each variety and each vintage, grape samples were collected at subsequent harvest times, from veraison (30% of the berries of the bunches starting to change colour) to complete ripening. In vintage

2020, a total of 8 harvest times were considered for *Ancellotta* and *Lambrusco Salamino* varieties, while 10 harvest times were considered for *Sangiovese* samples. Conversely, 10 subsequent harvest times were considered for all the three grape varieties in vintage 2021.

Each sample consisted of about 500 g of berries collected from different bunches located in different plants of the vineyard. The berries were collected in order to be as much representative as possible of vineyard variability of the ripening stage at the considered harvest time. Furthermore, the berries were collected paying attention to leave the fruit pedicel, in order to avoid the leakage of juice from the point of insertion.

RGB images of the collected samples were acquired using the device previously described in Section 2.1. In particular, for each sample six images of different randomly taken aliquots were acquired, in order to better represent the intrinsic variability of the berries colour within each sample (*see* Section 2.4.1). More detailed information about the collected samples is reported in Table 1.

Table 1 *Information about the grape samples of the three grape varieties collected in 2020 and 2021 vintages.*

Grape Variety	**Year**	**Number of vineyards**	**Number of harvest times**	**Number of samples**		**Number of acquired images**	
Ancellotta	2020	2	8	16	46	96	276
	2021	3	10	30		180	
Lambrusco Salamino	2020	2	8	16	46	96	276
	2021	3	10	30		180	
Sangiovese	2020	2	10	20	50	120	300
	2021	3	10	30		180	

After harvesting, the collected grape samples were immediately carried to the laboratory under refrigerated conditions, where they were analysed using the reference analytical methods that are commonly used by the wineries to determine the parameters of interest, i.e., total anthocyanins content (TAnt), extractable anthocyanins content (ExAnt) and total anthocyanins determined according to Di Stefano method (DS-TAnt).

2.3 Reference analysis for the determination of grape phenolic maturity

2.3.1 Total and extractable anthocyanins content

TAnt and ExAnt parameters were determined according to Glories method (Glories, 1993; Saint Cricq de Gaulejac et al., 1998). For each sample, grape berries were put in a blender (Kenwood CH-58) and mashed in order to homogenize pulp and skin. The so-obtained grape mash was divided into two aliquots of equal weight (Gibertini TM1600 Balance, Gibertini Elettronica, Milan, Italy): one aliquot was macerated in a pH 1 solution to have complete extraction of anthocyanins, while the second aliquot was macerated in a pH 3.2 solution to evaluate anthocyanins extractability at pH conditions similar to those of the winemaking process. After 4 h of maceration at room temperature under low agitation (BICASA BE39, BICASA srl, Monza Brianza, Italy), the two solutions were filtered on glass wool and spectrophotometric determination of anthocyanins (JASCO V-530, Tokyo, Japan) was performed on both clear juices using the methodology based on bleaching with sulfur dioxide (Ribéreau-Gayon, 1965; Río Segade et al., 2008; Sommer and Cohen, 2018). The total anthocyanins content (TAnt, expressed as mg/Kg of malvidin-3-glucoside) and extractable anthocyanins content (ExtAnt, expressed as mg/Kg of malvidin-3-glucoside) were determined from the pH 1 and pH 3.2 solutions, respectively.

2.3.2 Total anthocyanins according to Di Stefano method

Total anthocyanins were also determined according to the method proposed by Di Stefano (Di Stefano and Cravero, 1991; Di Stefano and Maggiorotto, 1995). This method consists in measuring the content of anthocyanins present in the grape peel extracted in an alcoholic solution at pH of 3.2 and determined using spectrophotometric techniques. For each sample, grape berries were frozen in the shortest possible time and were immersed in a hydro-alcoholic buffer at pH of 3.2 containing tartaric acid, ethyl alcohol, sodium metabisulphite and sodium hydroxide. The berries where then homogenized with a Silverson homogenizer (SILVERSON Machine ltd, Chesham, UK) and the homogenate was subsequently centrifuged (ALC 3229 Centrifuge). The supernatant was then used for the spectrophotometric determination of anthocyanins (JASCO V-530, Tokyo, Japan). In particular, the concentration of total anthocyanins (expressed as mg/Kg of malvidin-3-glucoside) was calculated from the maximum absorbance value registered in the visible range, which falls at about 540 nm.

2.4 Image acquisition and elaboration

Figure 4 provides a schematic representation of the different steps implemented in the device software from image acquisition of the grape samples to visualization of the predicted chemical parameters of interest related to phenolic ripening. These steps include:

1. Image acquisition using the smartphone app;
2. Image correction;
3. Concatenation of corrected sample images to create the *scansets*;
4. Conversion of *scansets* into colourgrams;
5. Prediction of chemical parameters by means of PLS models;
6. Display of the results on the smartphone app (and web interface).

In this study, the elaboration steps from image correction to the development and validation of the PLS models (steps 2-5) have been developed in MATLAB environment (The MathWorks Inc., USA) using both functions written *ad hoc* and the PLS Toolbox software (ver. 8.8.1, Eigenvector Research Inc., USA). Then, the corresponding algorithms and calibration models were implemented in the software so that the end user can simply acquire images and get the predictions with no further intervention needed.

All these steps are described in more detail in the following sections.

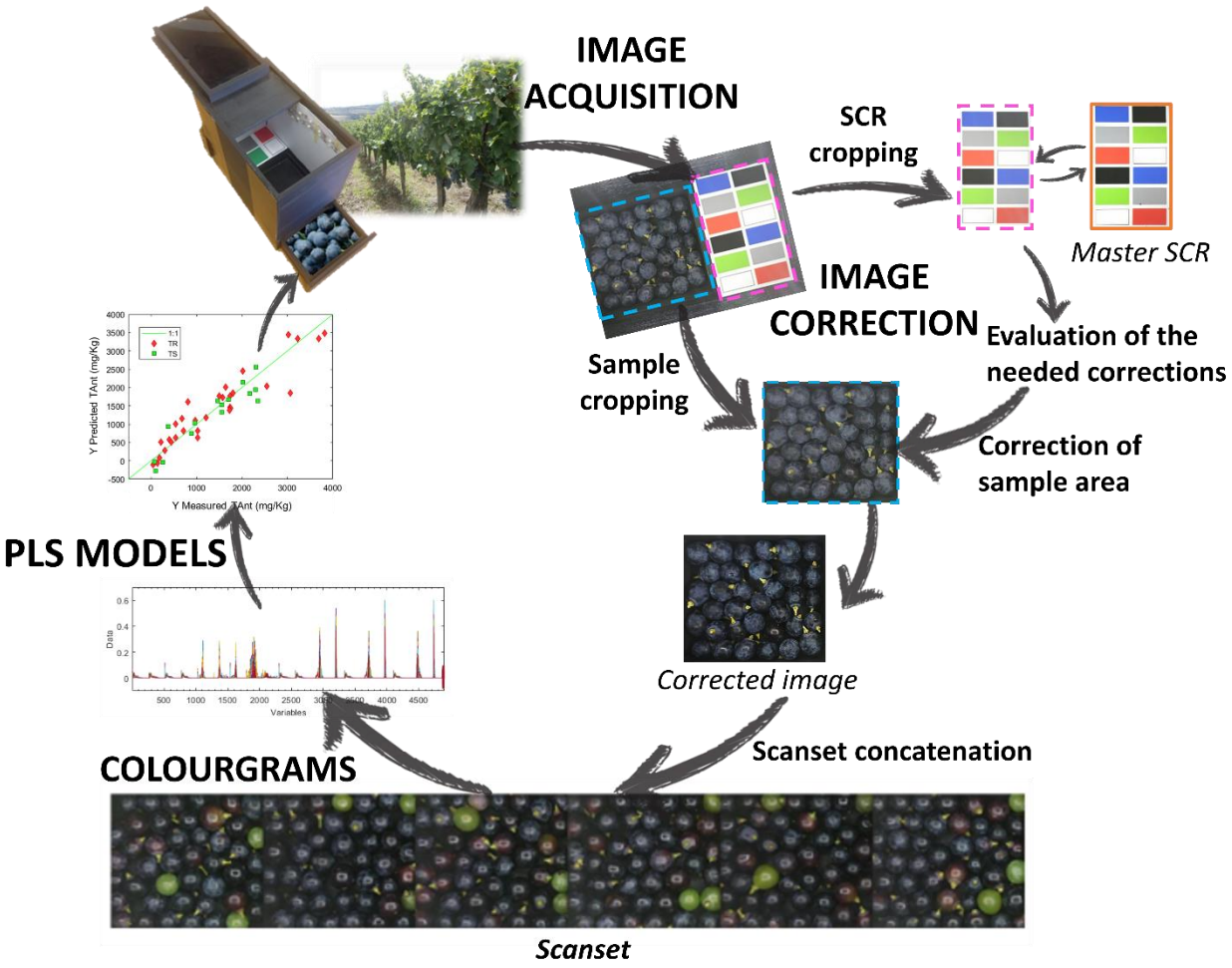


Figure 4 *From image acquisition of the grape samples to determination of anthocyanins content: main steps.*

2.4.1 Image acquisition

For each sample, the collected berries were split into 6 aliquots and each aliquot was imaged using the device described in Section 2.1, equipped with a Xiaomi Redmi Note 8 smartphone embedding a 48 megapixels camera with a spatial resolution of 8000 × 6000 pixels, and considering a shutter speed of 1/50 s, f /1.9 lens aperture and ISO 200. The image scene also included a cardboard of standard colour references used for the subsequent image standardization step. The RGB images were saved in JPEG file format with an average file size of 6 MB.

On the whole, 852 images were acquired, corresponding to 276 images for *Ancellotta* samples, 276 images for *Lambrusco Salamino* samples and 300 images for *Sangiovese* samples (*see* Table 1 for further details).

2.4.2 Image correction

A standardization procedure is highly recommended to reduce a possible variation among images, which can be caused by instrumental instability over time.

Each acquired image was composed by two different areas: the area with the cardboard of Standard Colour References (SCR), that contains 12 coloured rectangles, and the sample area with the grape berries. These two areas were automatically selected and cropped from the original image. The SCR area was then used to perform image standardization using the procedure previously described in detail by Orlandi et al., 2018 (G. Orlandi et al., 2018).

Basically, image correction consists in comparing the SCR of an image considered as reference (*master* image), usually the first acquired image, with the SCR of each one of the remainder images (*slave* images). If there are significant differences between the two references areas, the SCR of the slave image are corrected to be as much similar as possible to the SCR of the master image, and the same correction procedure is also applied to the whole slave image, including the sample area.

In particular, each one of the 12 colour patches depicted in the colour reference was divided into 4 pixels groups, according to a 2 × 2 grid. The median values of red (R), green (G) and blue (B) channels were calculated for each pixel group. These median values were subsequently stored in matrix of size {4×12, 3}, where 4 is the number of pixel groups, 12 is the number of the colour reference patches and 3 corresponds to the R, G and B channels.

Then, an internal calibration was performed by computing, channel by channel, a regression model between the vector of median values of the master image (y) and the vector of median values of the i-th slave image (x_i). Based on some preliminary tests, a first-order polynomial was considered, which can be expressed as follows:

$$y(c) = b_0(i,c) + b_1(i,c) \times S_i(c) \quad (1)$$

where b_0 and b_1 are the regression coefficients calculated for the channel c of the i-th slave image, S_i. Furthermore, the algorithm also estimates the statistical significance of the model coefficients in order to select only the significant coefficients.

Then, the regression coefficients calculated as described above were used to correct each pixel of the corresponding sample area according to the following equation:

$$img_{corr}(i,c) = b_0(i,c) + b_1(i,c) \times img_{orig}(i,c) \quad (2)$$

where $img_{corr}(i,c)$ is the channel c of the i-th corrected image, $img_{orig}(i,c)$ is the channel c of the corresponding original image, and $b_0(i,c)$ and $b_1(i,c)$ are the model coefficients calculated using Equation 1.

Image correction was performed using RGB Correction GUI, a graphical user-friendly interface working under MATLAB (The MathWorks) environment and freely downloadable from https://www.chimslab.unimore.it/downloads/. The standardization procedure described above was also implemented in the software of the device in order to automatically correct the acquired images when necessary.

2.4.3 Conversion to colourgrams

After correction, the sample area of each image (i.e., the area containing the grape berries) was cropped obtaining a new image of size 2310 × 2623 pixels.

As mentioned in Section 2.1.2, each grape sample was randomly split into 6 aliquots that were separately imaged, thus obtaining 6 different images for each sample. However, for further analysis the corrected and cropped images of the 6 aliquots of the same sample were merged together into a *scanset*, therefore obtaining one merged image with size 2310 × 15738 pixels for each sample. The dataset of merged sample images consisted therefore in 46 *scansets* for both *Ancellotta* and *Lambrusco Salamino* varieties, and in 50 *scansets* for *Sangiovese* variety.

This procedure was carried out in view of the development of the calibration models, where the determination of the chemical parameters was performed considering each sample as a whole and not considering the single aliquots. In addition, performing the evaluation considering more aliquots altogether allows to better account for sample variability and thus obtain more robust and reliable results.

Finally, each *scanset* was converted into the corresponding colourgram, a one-dimensional signal that codifies the colour-related information content of the RGB image.

At first, *scanset* image is unfolded into a two-dimensional matrix with as many rows as the number of pixels and three columns corresponding to the R, G, and B channels (Giraudo et al., 2018).

Such 2D matrix is then expanded by adding new columns, corresponding to the following 16 additional parameters calculated from the R, G, B values for each image pixel:

- lightness (L), defined as the sum of R, G, B values;

- relative colours: relative Red (rR), relative Green (rG), relative Blue (rB), calculated as the ratio between each channel value and L;
- hue (H), saturation (S) and intensity (I), obtained by converting the RGB data into HSI colour space;
- the three score vectors of the PCA model calculated from raw RGB data matrix (SC1R, SC2R, SC3R);
- the three score vectors of the PCA model calculated from mean centred RGB data matrix (SC1M, SC2M, SC3M);
- the three score vectors of the PCA model calculated from autoscaled RGB data matrix (SC1A, SC2A, SC3A).

Finally, the colourgram of each *scanset* is obtained by merging in sequence the frequency distribution curves of each one of the 19 considered colour parameters (i.e., 3 RGB channels + 16 additional parameters). Each frequency distribution curve (i.e., a histogram with 256 bins) was calculated accounting for the entire range of variability of the considered parameter. In addition, at the end of the signal also the loading vectors and the eigenvalues of the three PCA models calculated on the RGB matrix are added, obtaining a 4900 points long signal. An example of the colourgrams obtained from the *scanset* images of *Sangiovese* samples is provided in Figure 5. For a deeper insight of the algorithm developed to create the colourgrams, the reader is referred to Antonelli et al., 2004 and Calvini et al., 2020 (Antonelli et al., 2004; Calvini et al., 2020).

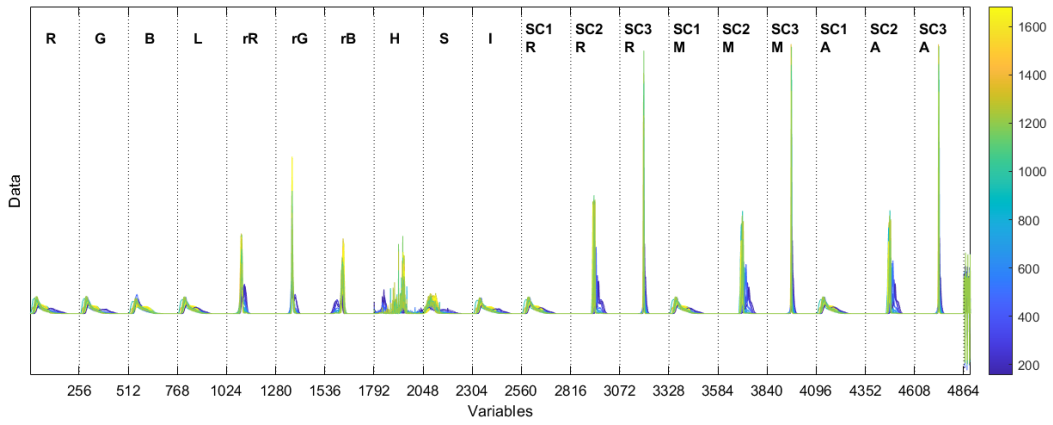


Figure 5 *Example of colourgrams calculated from RGB sample images of Sangiovese samples. The signals are coloured according to the corresponding measured content of total anthocyanins (mg/Kg).*

Colourgrams are subsequently collected into a data matrix in which each row corresponds to the signal derived from a specific *scanset*, and such matrix can then be analysed by means of common multivariate statistical methods for data exploration or for the development of regression models.

A preliminary evaluation of the colourgrams data matrix performed by means of PCA on mean centred signals allowed to observe that the correlation existing between the berry colour and the corresponding anthocyanins content is specific for each single variety. For this reason, the images obtained from the different grape varieties were analysed separately from each other and three different matrices of colourgrams were obtained: one matrix with size {46×4900} for *Ancellotta* samples, one matrix with size {46×4900} for *Lambrusco Salamino* samples and one matrix with size {50×4900} for *Sangiovese* samples.

The colourgrams were calculated using Colourgrams GUI (Calvini et al., 2020), a graphical user-friendly interface developed under MATLAB environment (The MathWorks Inc., USA) and freely downloadable from https://www.chimslab.unimore.it/downloads/. As for the standardization procedure, also the conversion of *scansets* into the corresponding colourgrams was implemented in the software of the device to be automatically performed.

2.5 Calculation of calibration models

Multivariate calibration models were calculated separately for each grape variety in order to predict the three chemical parameters of interest (**y** variables), i.e., total anthocyanins, extractable anthocyanins and total anthocyanins determined with Di Stefano method, using the colourgrams obtained from the *scansets* of the grape samples as predictors (**X** matrix).

To this aim, each one of the three colourgram matrices, corresponding to the three grape varieties, was split into a training set and a test set. The subdivision of the samples between training and test sets was done in a way to attribute to the test set different harvest times for each vineyard. The training set was used to calculate the calibration models using Partial Least Squares regression (de Jong, 1993; Geladi and Kowalski, 1986), while the test set was used to perform an external validation of the obtained models. In particular, for both *Ancellotta* and *Lambrusco Salamino* colourgrams matrices 31 signals were considered in the training set and 15 signals were considered in the test set. Conversely, for *Sangiovese* colourgrams matrix, one sample was detected as outlier and the remainder samples were split into training and test sets as follows: 34 signals in the training set and 15 signals in the test set.

For the calculation of the PLS models, the colourgrams matrices were pre-processed using mean center, and the optimal number of Latent Variables (LVs) was selected by minimizing the Root Mean Square Error in Cross-Validation (RMSECV), considering a venetian blinds cross-validation scheme with 5 deletion groups.

The performance of the calibration models was expressed in terms of Root Mean Square Error (RMSE) and coefficient of determination (R^2) statistics, both calculated in calibration (RMSEC, R^2_{Cal}), cross-validation (RMSECV, R^2_{CV}) and prediction (RMSEP, R^2_P). The PLS models were calculated using the PLS Toolbox (ver. 8.8.1, Eigenvector Research Inc., USA) running under MATLAB environment (The Mathworks Inc., USA). The calibration models developed in this study were then implemented in the software architecture of the device.

As mentioned in Section 2.4.3, the PLS models implemented into the device were calculated separately for each grape variety, since for each variety a specific relationship between berries colour and anthocyanins content was observed. This behaviour is likely ascribable to the peculiar physiology and ripening behaviour of each plant variety. However, for the sake of completeness, we also calculated global PLS models, i.e., considering the entire set of samples of the three varieties altogether.

3 Results and discussion

3.1 Evaluation of the chemical parameters related to phenolic maturity

In this work, the calibration models were calculated considering grape samples belonging to different grape varieties and harvested in three different vineyards. In particular, in the case of *Ancellotta* and *Lambrusco Salamino* varieties, one vineyard was located in hilly areas while the other two vineyards were located in flat areas. Conversely, for *Sangiovese* variety two vineyards were located in hilly areas and one vineyard was located in a flat area.

The different geographical position of the vineyards is a relevant aspect to be considered for the development of the calibration models, since the differences in pedoclimatic conditions can affect the grape ripening trend. In this context, we observed some differences in the ripening trends between vineyards with different geographical position. For example, Figure 6 shows the trends of total anthocyanins content for the three varieties in 2020 and 2021 vintage years according to the different vineyards. In the case of *Lambrusco Salamino* variety, it is particularly evident that the samples harvested in the vineyard located in a foot-hill region (Sal-B) tend to reach a higher anthocyanins amount compared to samples harvested in the vineyards located in flat areas (Sal-A and Sal-C).

Similar results were also obtained considering *Ancellotta* samples. As shown in Figure 6, for *Sangiovese* variety the variability among vineyards located in different geographical positions is less evident compared to the other grape varieties; this could be due to the fact that, compared with the vineyards of *Ancellotta* and *Lambrusco Salamino*, for *Sangiovese* variety the vineyards located in hilly areas were much closer to those located in flat areas.

Similar trends were also observed considering the evolution of ExAnt and DSAnt trends over time for the three grape varieties (Figures 7 and 8).

For a more comprehensive evaluation of the oenological characteristics of the collected samples, Figures 9-11 report the box plots of TAnt, ExAnt and DSAnt grouped according to grape varieties for both 2020 and 2021 vintage years. It is possible to observe that *Ancellotta* samples reach higher values for all the three parameters, while *Sangiovese* variety is generally characterized by a lower anthocyanins content with respect to *Ancellotta* and *Lambrusco Salamino* varieties. Considering vintage years, the differences are more evident considering TAnt parameter and *Ancellotta* samples, indeed in 2021 the total anthocyanins content for this grape variety reached higher values compared to 2020 vintage year.

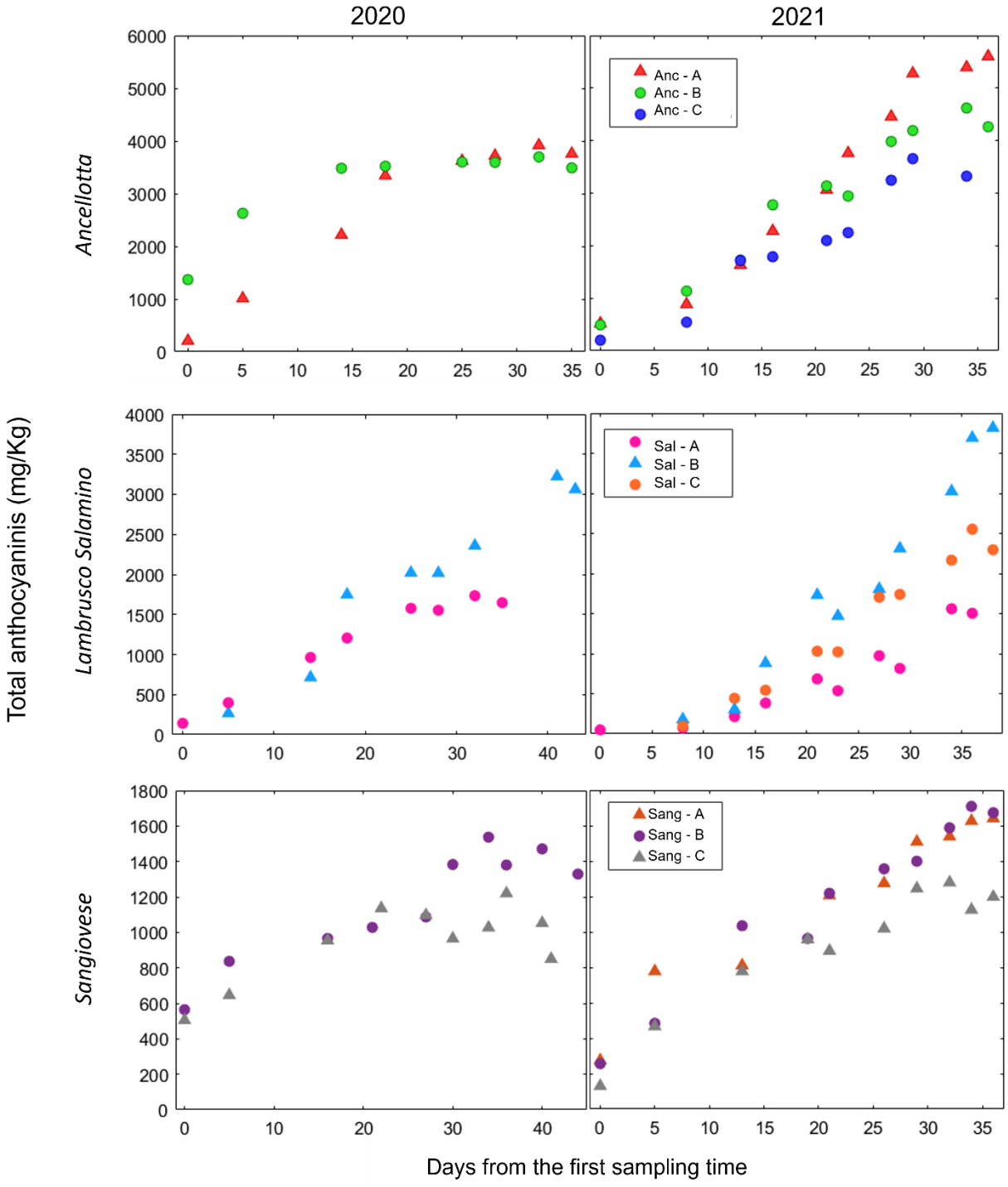


Figure 6 *Total anthocyanins trends for Ancellotta, Lambrusco Salamino and Sangiovese in 2020 and 2021 vintages; samples are coloured according to the corresponding vineyard; triangles are referred to vineyards located in hilly areas, while circles to vineyards located in flat areas.*

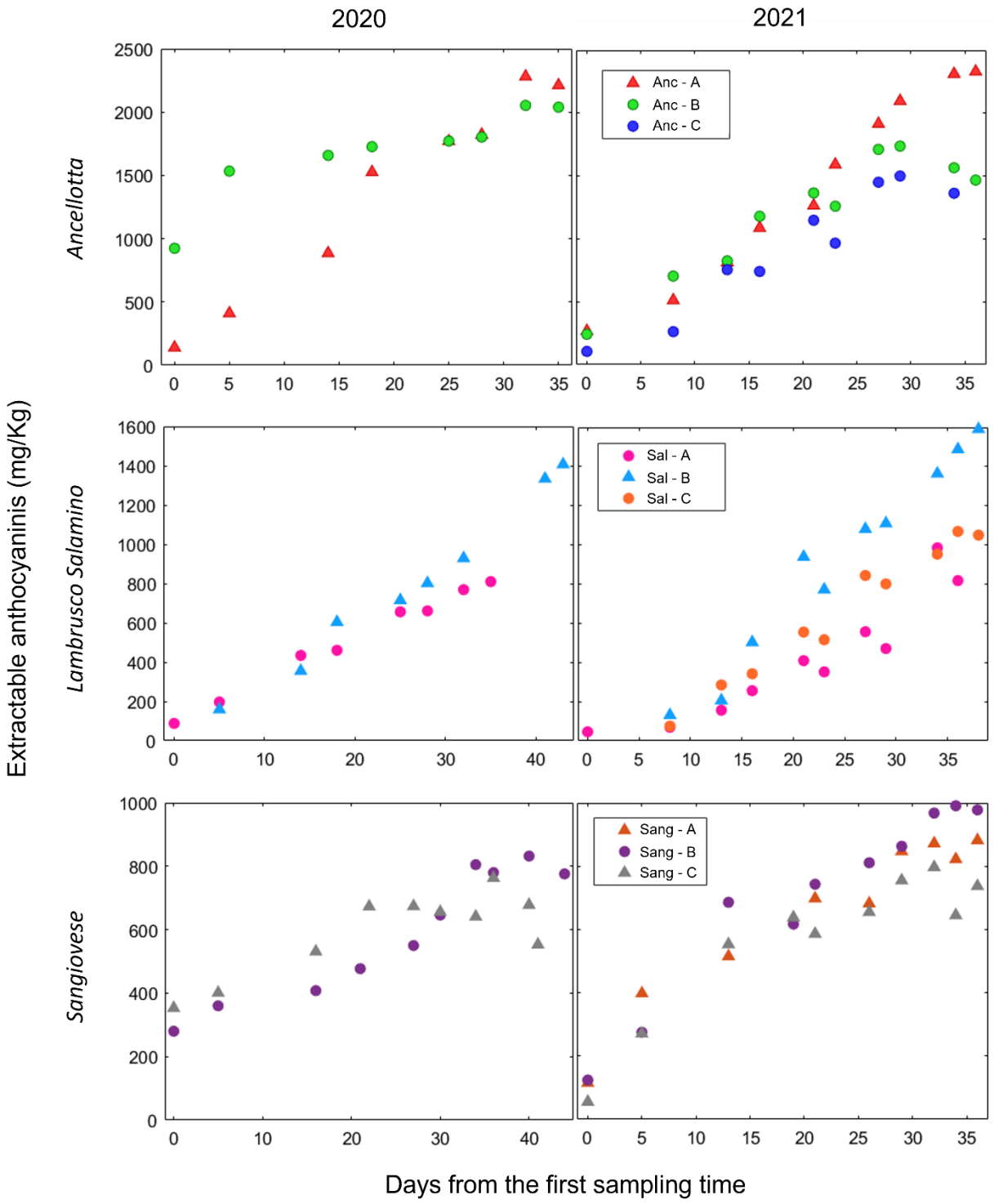


Figure 7 *Extractable anthocyanins trends for Ancellotta, Lambrusco Salamino and Sangiovese in 2020 and 2021 vintages; samples are coloured according to the corresponding vineyard; triangles are referred to vineyards located in hilly areas, while circles to vineyards located in flat areas.*

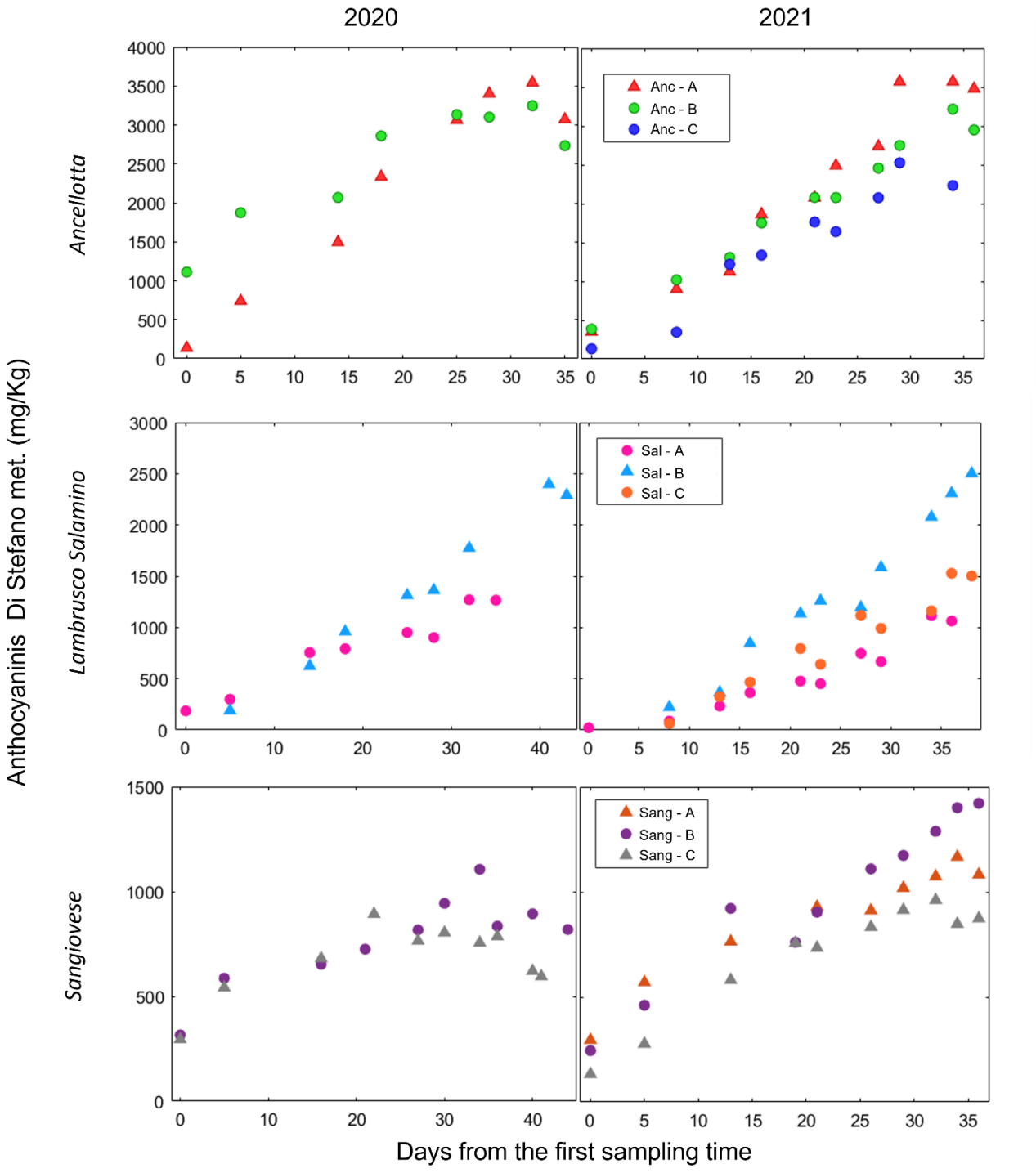


Figure 8 *Trends of anthocyanins determined according to Di Stefano method for Ancellotta, Lambrusco Salamino and Sangiovese in 2020 and 2021 vintages; samples are coloured according to the corresponding vineyard; triangles are referred to vineyards located in hilly areas, while circles to vineyards located in flat areas.*

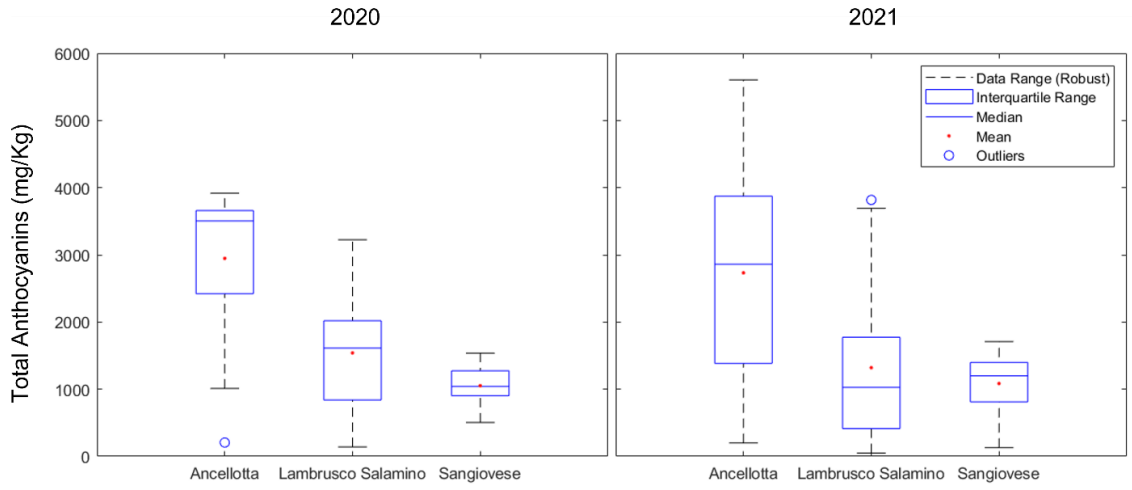


Figure 9 *Boxplots of total anthocyanins for Ancellotta, Lambrusco Salamino and Sangiovese in 2020 (on the left) and 2021 (on the right) vintages.*

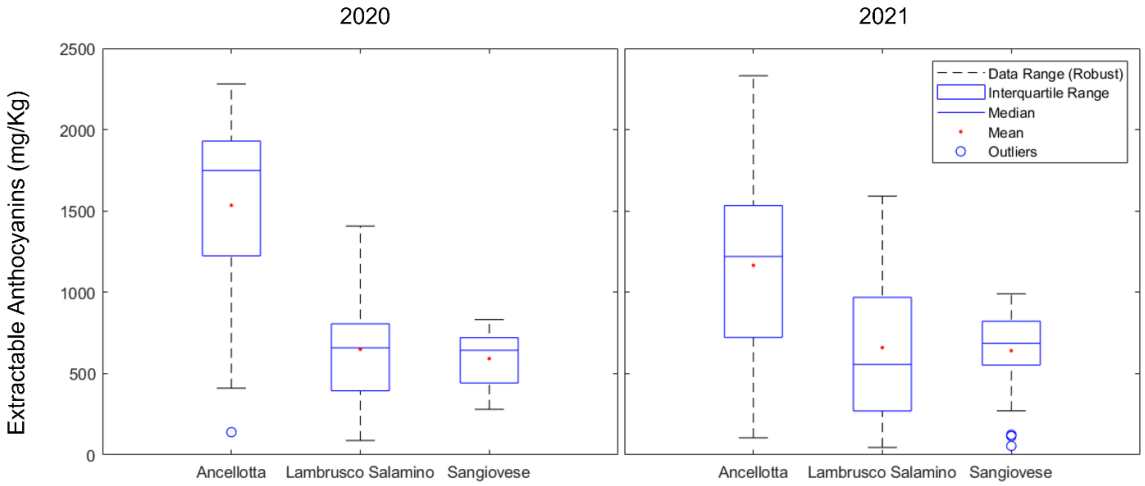


Figure 10 *Boxplots of extractable anthocyanins for Ancellotta, Lambrusco Salamino and Sangiovese in 2020 (on the left) and 2021 (on the right) vintages.*

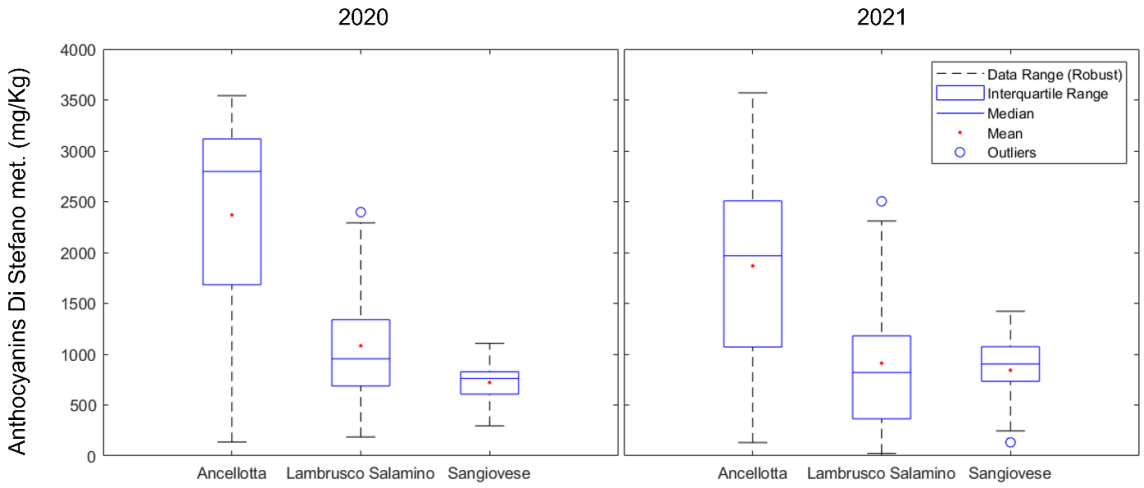


Figure 11 *Boxplots of total anthocyanins determined according to Di Stefano method for Ancellotta, Lambrusco Salamino and Sangiovese in 2020 (on the left) and 2021 (on the right) vintages.*

3.2 PLS models

Before the calculation of the PLS models, a preliminary exploratory analysis was performed by means of PCA considering the colourgrams of the images acquired from the samples belonging to all the three grape varieties. The corresponding PC1 and PC2 score plot is reported in Figure 12, where the samples are coloured according to grape variety (Figure 12a), vintage year (Figure 12b), vineyard location (Figure 12c) and harvest day (Figure 12d). It is possible to observe that PC1 mainly describes the colour evolution of grape samples during ripening and that the samples belonging to the different varieties follow different ripening trends. *Ancellotta* samples are generally located at higher PC1 score values, suggesting that this variety has a more pronounced colour variation during ripening. To better evaluate the relationship between the colour of the berries and total anthocyanins content, Figure 13 shows the plot of TAnt values *vs.* PC1 scores of the global PCA model calculated considering the colourgrams of the samples belonging to the three grape varieties. It is evident that each grape variety has a peculiar ripening trend and that a similar colour of the berries may correspond to a very different TAnt value.

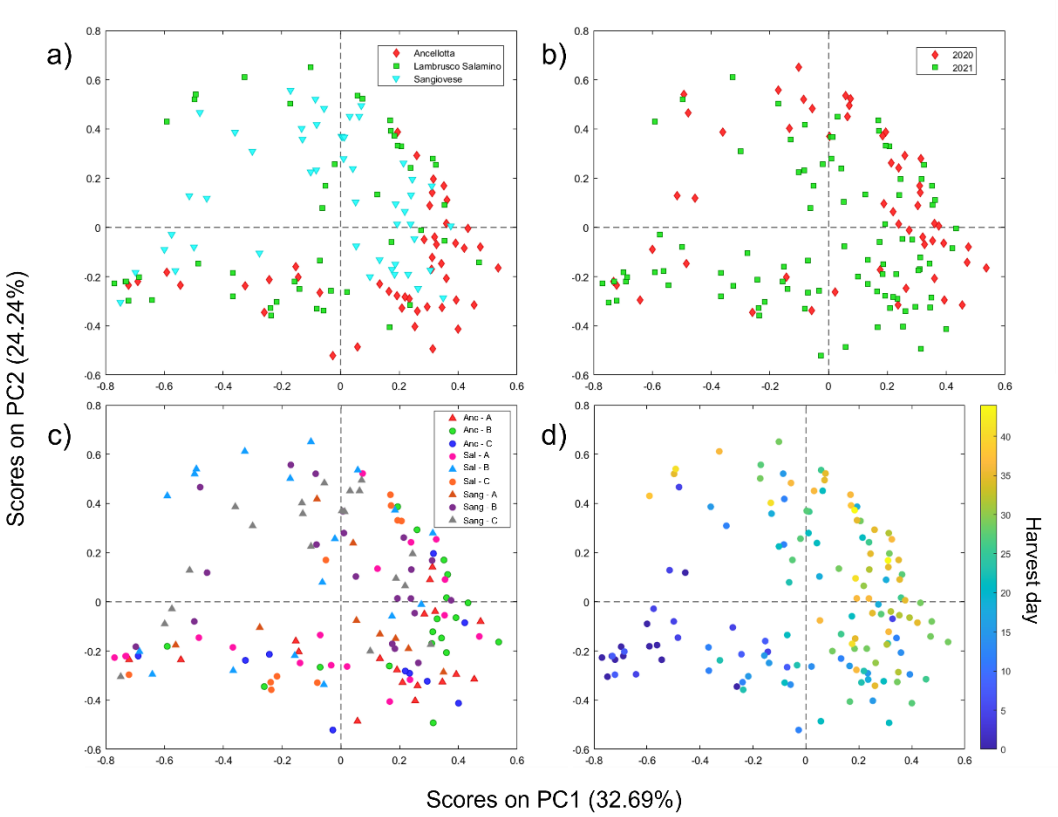


Figure 12 *PC1 and PC2 score plots of the global PCA model calculated considering the colourgrams dataset obtained from the images of grape samples belonging to the three varieties. In a) samples are coloured according to grape variety, in b) the samples are coloured according to the corresponding harvest year, in c) the samples are coloured according to the corresponding vineyard where triangles are referred to vineyards located in hilly areas, while circles to vineyards located in flat areas, and in d) the samples are coloured according to the harvest day with respect to the first harvest time for each year.*

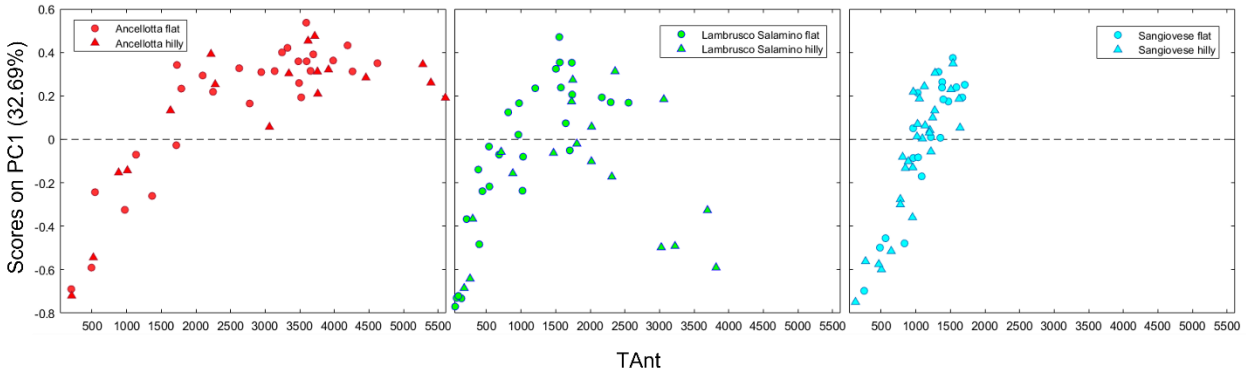


Figure 13 *Total anthocyanins content (TAnt) vs. PC1 scores of the global PCA model: samples are coloured according to grape variety; triangles are referred to vineyards located in hilly areas, while circles to vineyards located in flat areas.*

In order to relate the colour information contained in the RGB images of grape samples with the corresponding parameters of interest to estimate phenolic maturity, a total of 9 PLS1 models (= 3 grape varieties × 3 chemical parameters) were calculated using the colourgrams approach. In particular, as shown from the global PCA results discussed above, each grape variety has a peculiar maturation trend, therefore it is necessary to calculate separate calibration models for each grape variety. The results of the corresponding PLS models are reported in Table 2.

Generally, for all the three grape varieties the best results were obtained in the prediction of the total anthocyanins content (TAnt), with R^2_P values equal to 0.60, 0.85 and 0.81 for *Ancellotta*, *Lambrusco Salamino* and *Sangiovese*, respectively. For both *Lambrusco Salamino* and *Sangiovese* varieties the calibration models for the prediction of extractable anthocyanins content led to similar or slightly lower results compared to those obtained for TAnt, while for *Ancellotta* variety ExAnt provided lower prediction results and DSAnt model performed better.

Comparing the model performances obtained for the three grape varieties, the overall best results for all the three phenolic ripening parameters were obtained for *Lambrusco Salamino* variety followed by *Sangiovese* variety, with R^2_P values higher than 0.81 and 0.71, respectively. In general, for these two varieties the obtained results can be considered very satisfactory, in view of the practical application of the device. Indeed, the possibility of an extensive and frequent sampling to get information in real-time about grape ripening, although with lower precision, is preferrable compared to using only the reference analytical methods. In fact, the latter require to send the samples to specialized analytical laboratories, which implies a significant time delay between sampling and

obtaining the analysis result; moreover, this also drastically limits the number of determinations that can be done during grape ripening.

Table 2 *Results of the PLS models.*

Grape Variety	**Y Variable**	**LVs**	**RMSEC**	**RMSECV**	**RMSEP**	**R^2_{Cal}**	**R^2_{CV}**	**R^2_{P}**
Ancellotta	TAnt	2	710	821	867	0.76	0.67	0.60
	ExAnt	1	346	382	427	0.70	0.64	0.51
	DSAnt	1	524	567	653	0.74	0.69	0.58
Lambrusco Salamino	TAnt	2	374	465	314	0.88	0.81	0.85
	ExAnt	2	159	188	137	0.87	0.81	0.85
	DSAnt	2	258	322	233	0.86	0.79	0.81
Sangiovese	TAnt	4	100	144	158	0.94	0.87	0.81
	ExAnt	4	71	95	103	0.91	0.83	0.77
	DSAnt	4	97	149	137	0.90	0.76	0.71

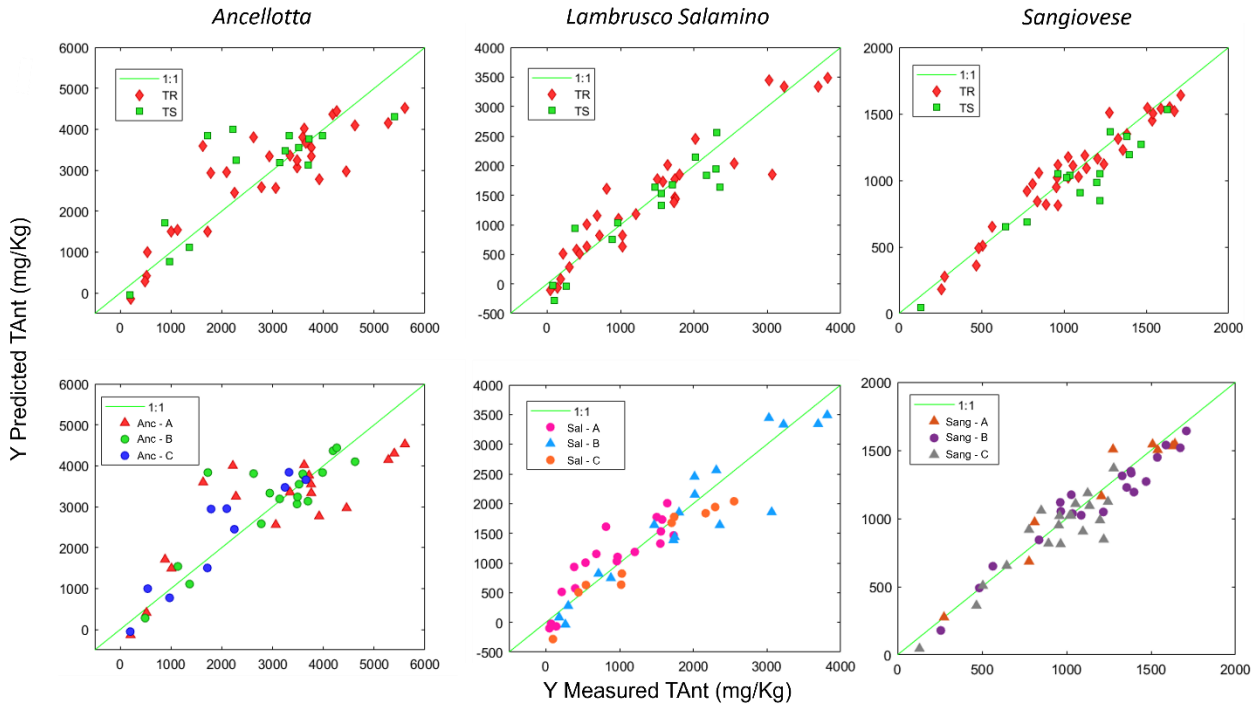


Figure 14 *Experimentally measured amount of total anthocyanins content (TAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the PLS models calculated for Ancellotta, Lambrusco Salamino and Sangiovese varieties. For each model, in the upper part of the figure the samples are coloured according to training (TR) and test (TS) sets, while on the lower part of the figure the samples are coloured according to the corresponding vineyard; triangles are referred to vineyards located in hilly areas, while circles to vineyards located in flat areas.*

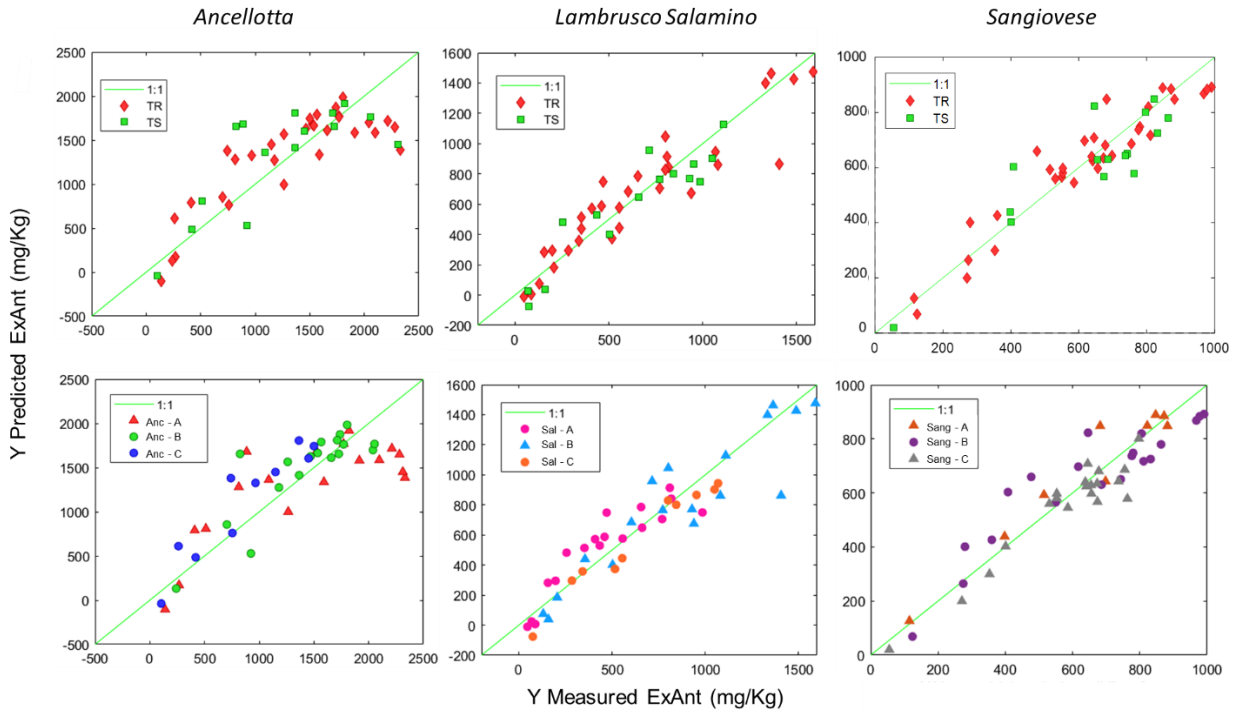


Figure 15 *Experimentally measured amount of extractable anthocyanins content (ExAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the PLS models calculated for Ancellotta, Lambrusco Salamino and Sangiovese varieties. For each model, in the upper part of the figure the samples are coloured according to training (TR) and test (TS) sets, while on the lower part of the figure the samples are coloured according to the corresponding vineyard; triangles are referred to vineyards located in hilly areas, while circles to vineyards located in flat areas.*

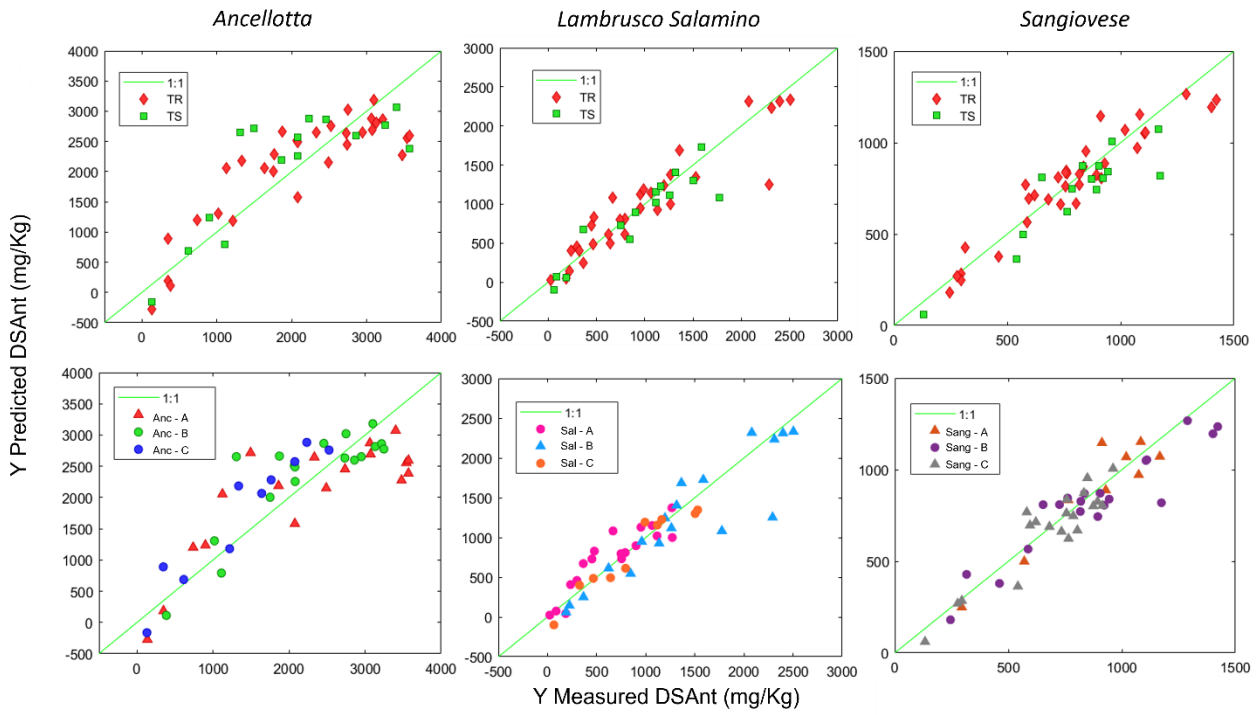


Figure 16 *Experimentally measured amount of total anthocyanins content determined according to the method proposed by Di Stefano (DSAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the PLS models calculated for Ancellotta, Lambrusco Salamino and Sangiovese varieties. For each model, in the upper part of the figure the samples are coloured according to training (TR) and test (TS) sets, while on the lower part of the figure the samples are coloured according to the corresponding vineyard; triangles are referred to vineyards located in hilly areas, while circles to vineyards located in flat areas.*

Figure 14 reports the measured values *vs.* the predicted values of the PLS models calculated for TAnt parameter for all the three grape varieties (the same plots for ExAnt and DSAnt are reported as in Figures 15 and 16). These plots confirm the satisfactory calibration performances obtained for both *Lambrusco Salamino* and *Sangiovese* varieties; indeed, the samples are generally close to the bisector.

Conversely, for *Ancellotta* grape variety the calibration models (Table 2) led to not fully satisfactory results: the R^2 values calculated in cross-validation (R^2_{CV}) were between 0.64 and 0.69, while in prediction the R^2 values (R^2_P) were between 0.51 and 0.60.

Observing the plots of the measured TAnt values *vs.* the predicted values obtained for the PLS model of *Ancellotta* variety reported in Figure 14, it is possible to understand the reason behind the low performances obtained for this grape variety. As a matter of fact, the samples with measured TAnt values lower than 2000 mg/Kg are close to the bisector, thus correctly predicted by the model. Conversely, the samples with measured TAnt values higher than 2000 mg/Kg are all predicted with values falling in the range 2500 – 4500 mg/Kg but with low accordance with the corresponding values determined with the reference analytical methods. In particular, the samples with actual TAnt values between 2000 mg/Kg and 3000 mg/Kg tend to be overestimated, while those samples with actual TAnt values higher than 4000 mg/Kg tend to be underestimated. These findings suggest that for *Ancellotta* variety at high TAnt values there is a low correlation between the colour of the berries and the total amount of anthocyanins. As a matter of fact, *Ancellotta* variety has a peculiar behaviour compared to the other grape varieties considered in this study: *Ancellotta* berries tend to change colour and become totally dark violet at the beginning of the maturation even if the ripening phenomena continue to evolve, also including the increase of anthocyanins content. In fact, compared to the other red grape varieties commonly used in Emilia Romagna region (Italy), *Ancellotta* is characterized by a high concentration of anthocyanins in the skin of the berries leading to their strong pigmentation, which makes this variety ideal as a colorant for red wines. A similar ripening behaviour was also observed in previous studies for other grape varieties such as *Tempranillo* and *Syrah* (Rodríguez-Pulido et al., 2022). A possible solution to overcome this limitation could be the acquisition of overexposed images of ripe berries, in order to better detect subtle colour changes that are hardly distinguishable at the naked eye when the berries colour is dark purple.

Another factor that could negatively affect the prediction models is the presence of grape bloom, a waxy compound of white colour that covers the berries. In this study, the grape samples were imaged without a prior removal of grape bloom to minimize their manipulation and, thus, to facilitate the use of the device directly in the vineyard. However, for grape varieties with peculiar ripening trends such

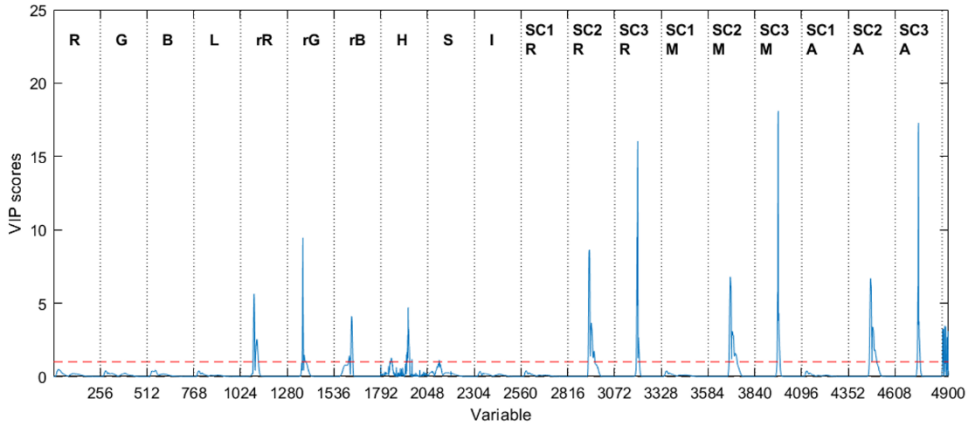
as *Ancellotta*, the removal of the grape bloom from the berries prior to image acquisition could help in removing non-informative variability sources and to improve the results of the calibration models.

Finally, in order to evaluate the colour parameters included in the colourgrams that are more relevant for the calibration models, the Variable Importance in Projection (VIP) scores were considered. As a general rule, variables with VIP score values greater than one are assumed to provide a substantial contribution to the corresponding model.

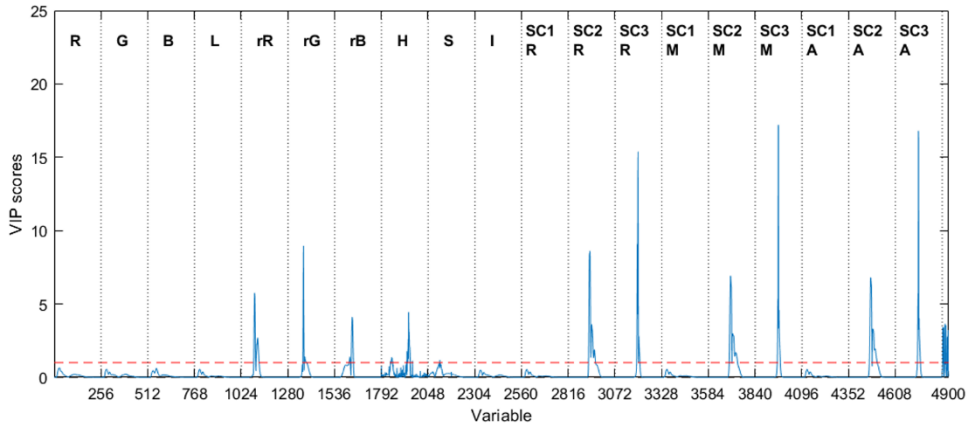
A comparison of the VIP scores of the three models for TAnt (available in Figures 17-19) shows that the most important variables, even if not precisely overlapping, always correspond to the same colourgram parameters for every variety. The colourgrams intervals arising as significant fall in the regions corresponding to the relative colours (rR, rG and rB), hue (H), the score vectors of PC2 and PC3 applied to both raw (SC2 R, SC3 R), mean centred (SC2 M, SC3 M) and autoscaled (SC2 A, SC3 A) RGB data matrix and a portion the loadings, which slightly diverges between the three models.

The same colour parameters were observed as significant also considering the VIP scores of the PLS models calculated to predict ExAnt and DSAnt for all the varieties (Figures 17-19).

TAnt



ExAnt



DSAnt

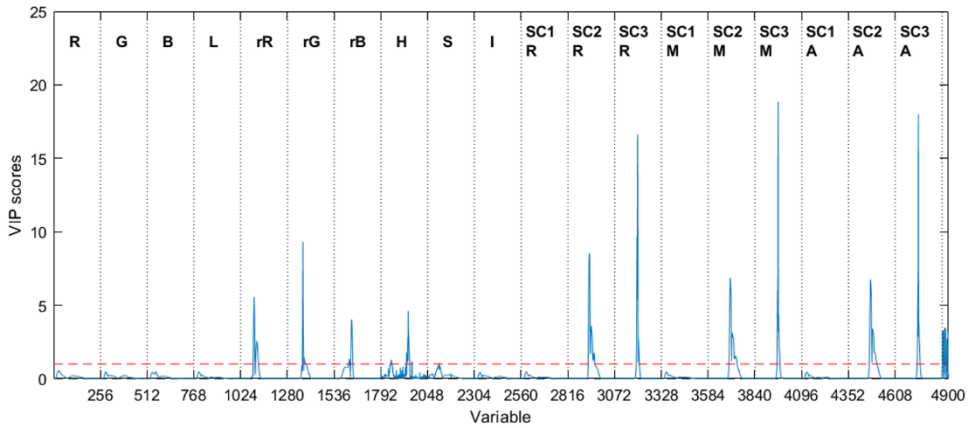


Figure 17 *VIP scores of the models for Lambrusco Salamino grape variety.*

TAnt

ExAnt

DSAnt

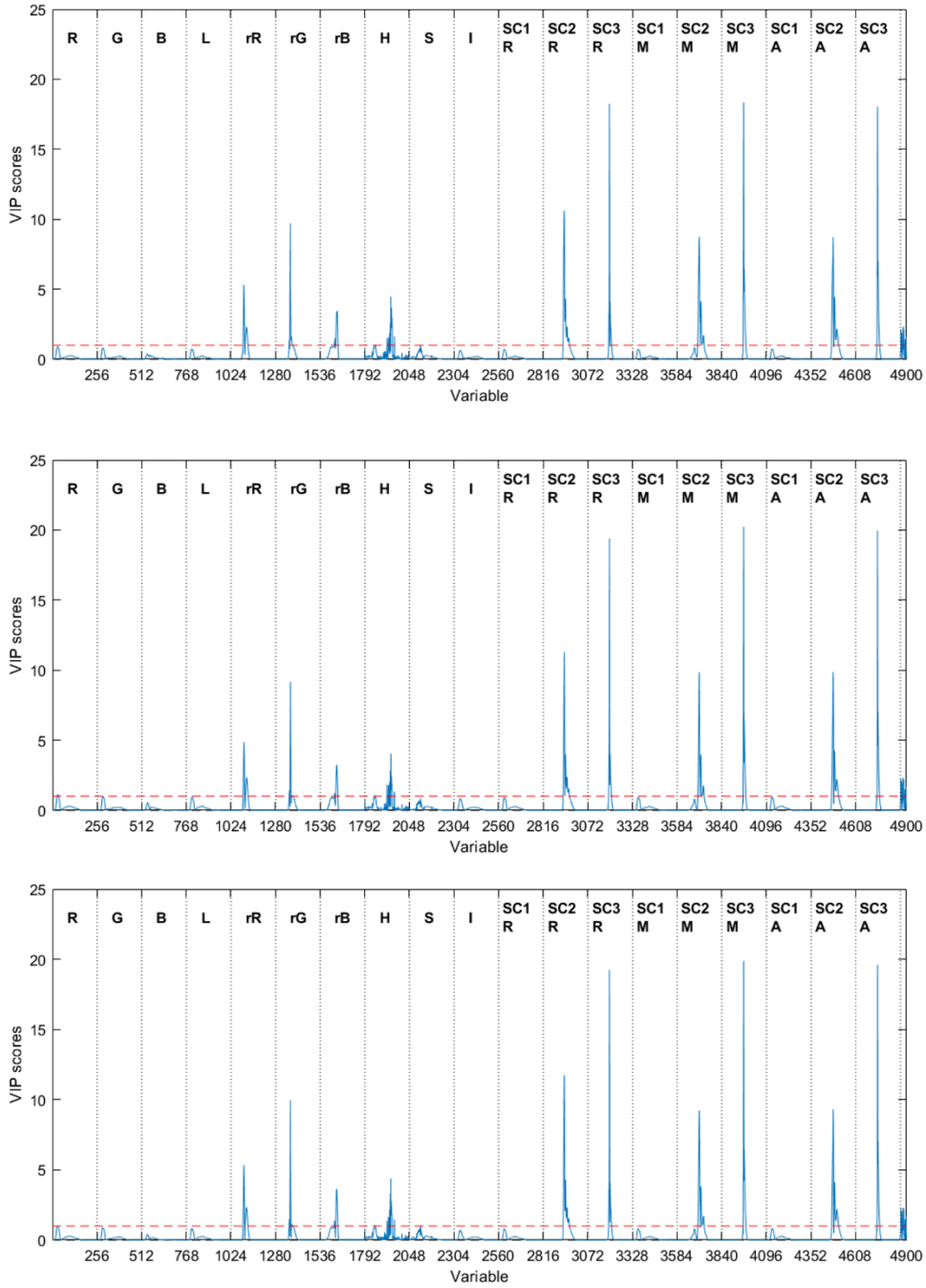


Figure 18 *VIP scores of the models for Ancellotta grape variety.*

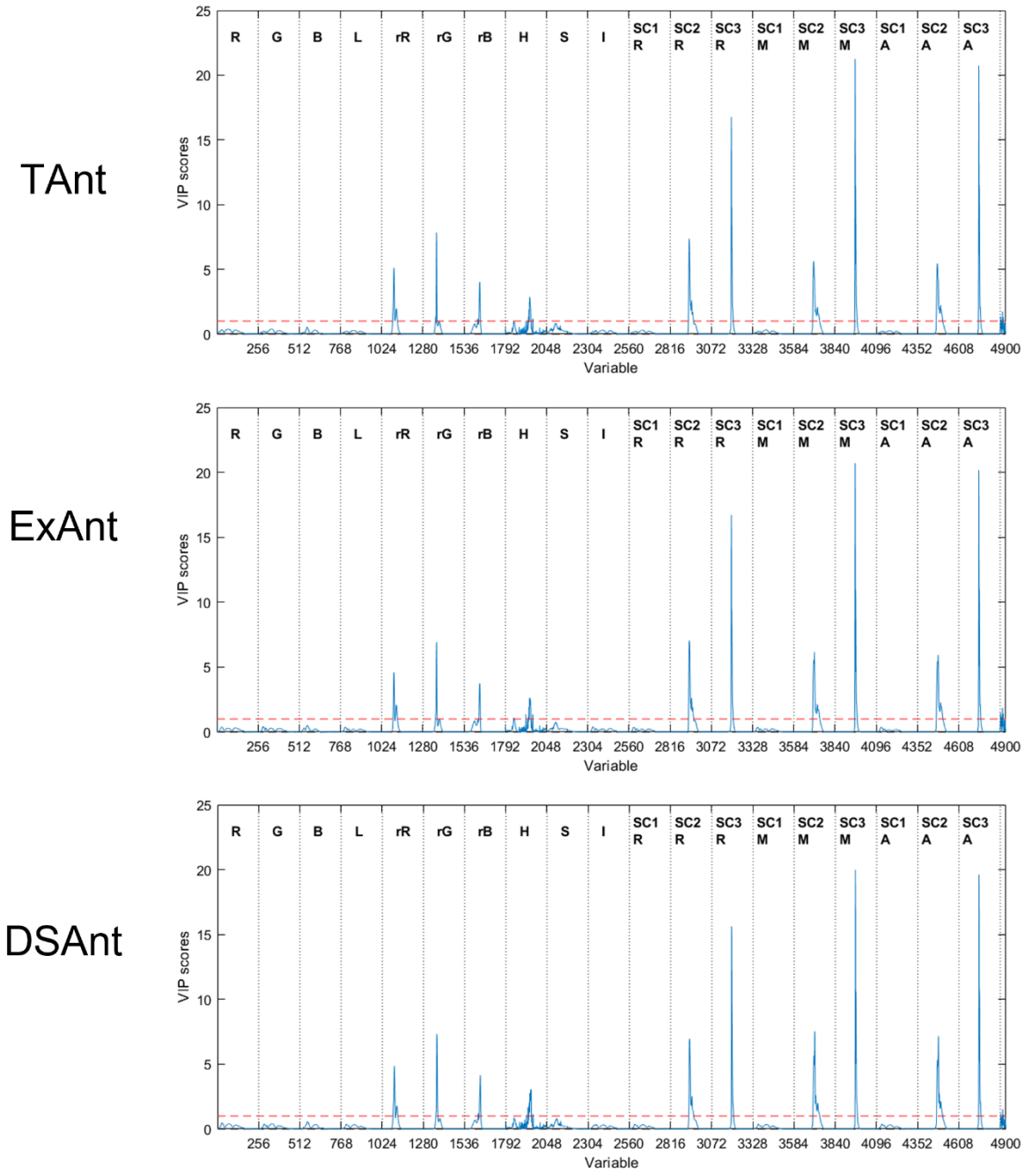


Figure 19 *VIP scores of the models for Sangiovese grape variety.*

For the sake of completeness, also global PLS models were calculated, considering all the three grape varieties altogether. However, the corresponding results were much lower than those obtained for the PLS models specific for each variety (Table 3 and Figures 20-22). As discussed above, *Ancellotta* samples have a very peculiar ripening trend which is different from the behaviour of the samples of the other varieties, therefore we also calculated PLS models considering only the *Lambrusco Salamino* and *Sangiovese* samples altogether (Table 4 and Figures 23-25). However, also in this case

the model performances resulted much lower than those of the PLS models calculated for each single variety. These findings further confirmed that the relationship between colour of the berries and anthocyanins profile is peculiar for each grape variety, therefore it is necessary to model such relationship considering specific PLS models for each considered variety.

Table 3 *Results of the global PLS models calculated considering Ancellotta, Lambrusco Salamino and Sangiovese varieties altogether.*

Y Variable	LVs	RMSEC	RMSECV	RMSEP	R^2_{Cal}	R^2_{CV}	R^2_{P}
Tant	4	834	953	823	0.59	0.47	0.50
ExAnt	4	340	408	373	0.63	0.47	0.47
DSAnt	3	679	679	602	0.46	0.46	0.49

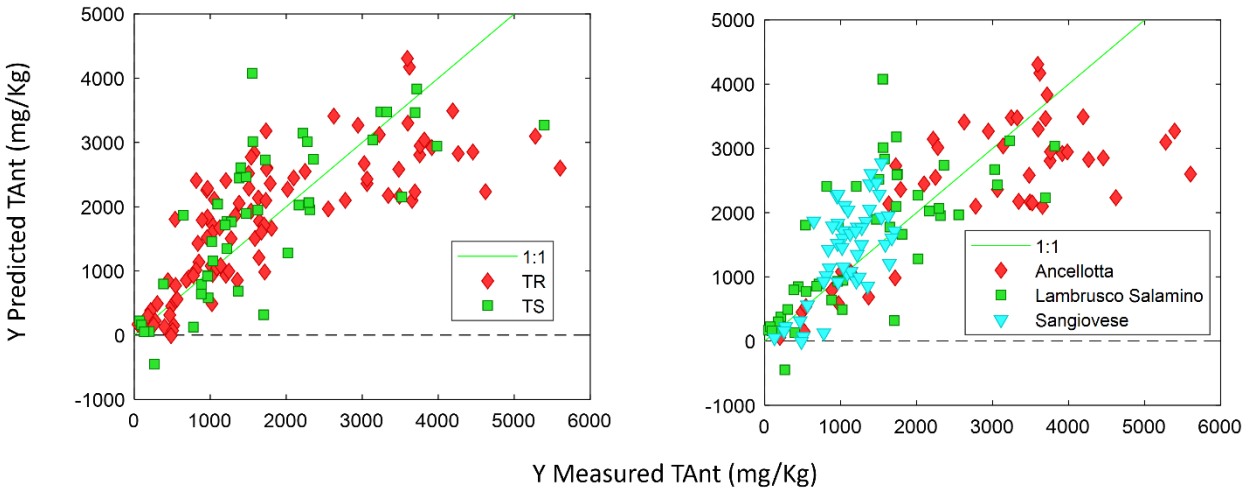


Figure 20 *Experimentally measured amount of total anthocyanins content (TAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the global PLS models calculated considering Ancellotta, Lambrusco Salamino and Sangiovese varieties altogether. In the left part of the figure the samples are coloured according to training (TR) and test (TS) sets, while in the right part of the figure the samples are coloured according to the corresponding variety.*

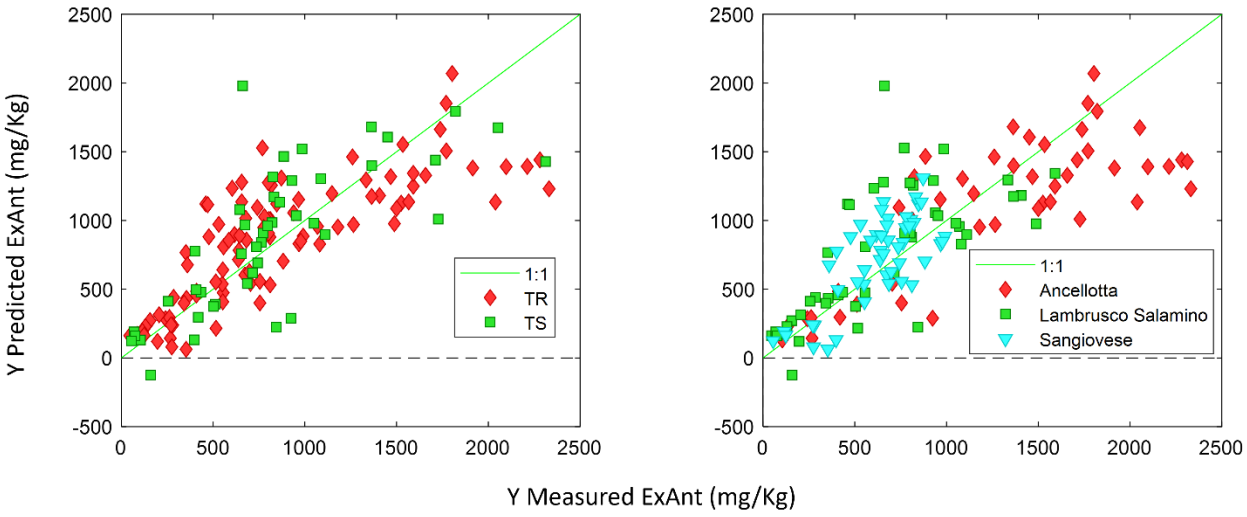


Figure 21 *Experimentally measured amount of extractable anthocyanins content (ExAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the global PLS models calculated considering Ancellotta, Lambrusco Salamino and Sangiovese varieties altogether. In the left part of the figure the samples are coloured according to training (TR) and test (TS) sets, while in the right part of the figure the samples are coloured according to the corresponding variety.*

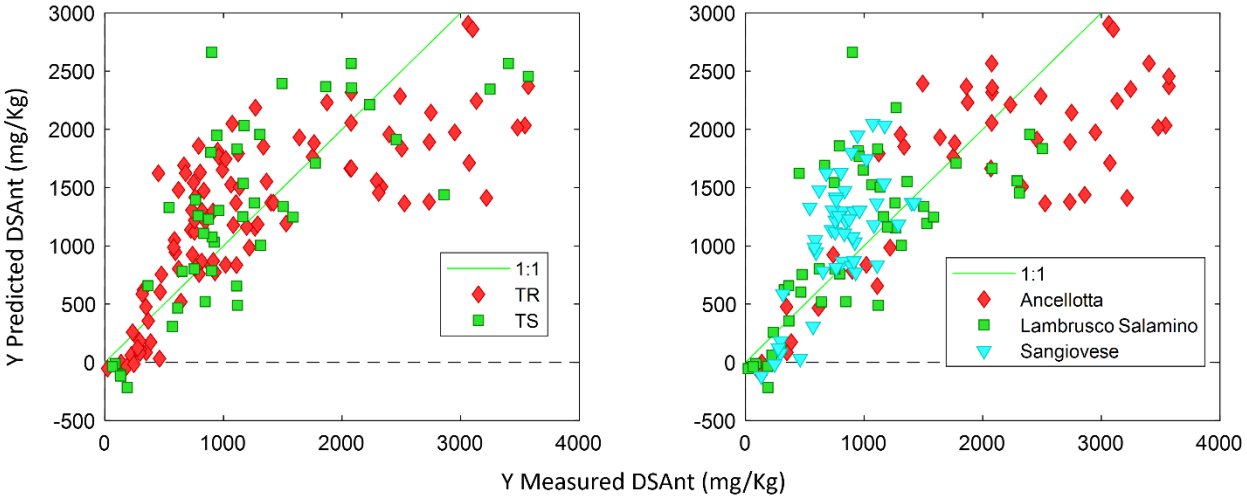


Figure 22 *Experimentally measured amount of total anthocyanins content determined according to the method proposed by Di Stefano (DSAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the global PLS models calculated considering Ancellotta, Lambrusco Salamino and Sangiovese varieties altogether. In the left part of the figure the samples are coloured according to training (TR) and test (TS) sets, while in the right part of the figure the samples are coloured according to the corresponding variety.*

Table 4 *Results of the PLS models calculated considering Lambrusco Salamino and Sangiovese varieties altogether.*

Y Variable	LVs	RMSEC	RMSECV	RMSEP	R^2_{Cal}	R^2_{CV}	R^2_{P}
Tant	4	428	580	423	0.72	0.48	0.56
ExAnt	3	197	240	167	0.67	0.51	0.67
DSAnt	4	300	396	292	0.69	0.45	0.52

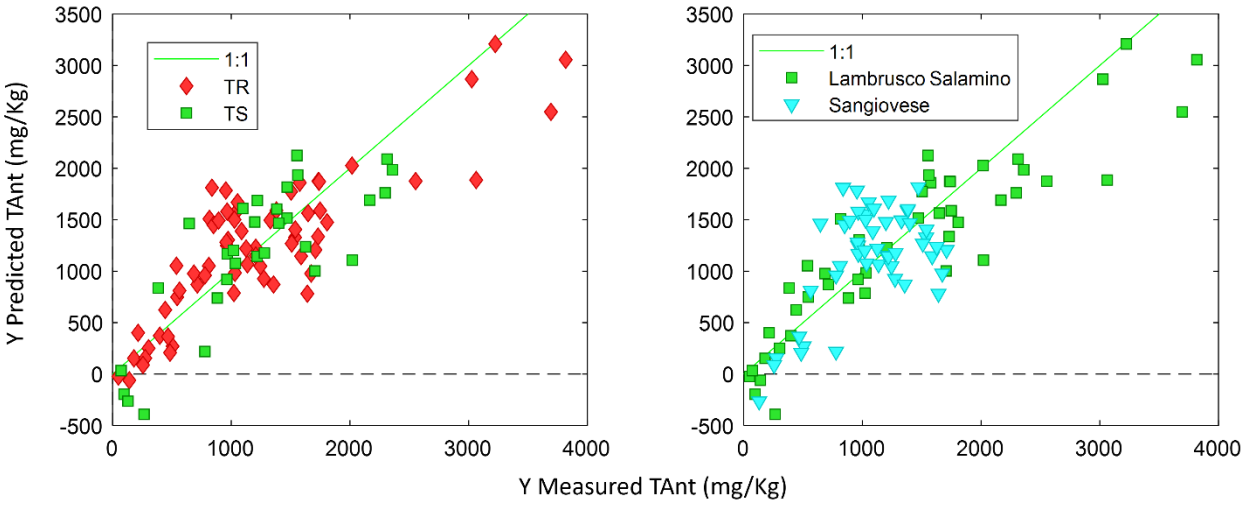


Figure 23 *Experimentally measured amount of total anthocyanins content (TAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the PLS models calculated considering Lambrusco Salamino and Sangiovese varieties altogether. In the left part of the figure the samples are coloured according to training (TR) and test (TS) sets, while in the right part of the figure the samples are coloured according to the corresponding variety.*

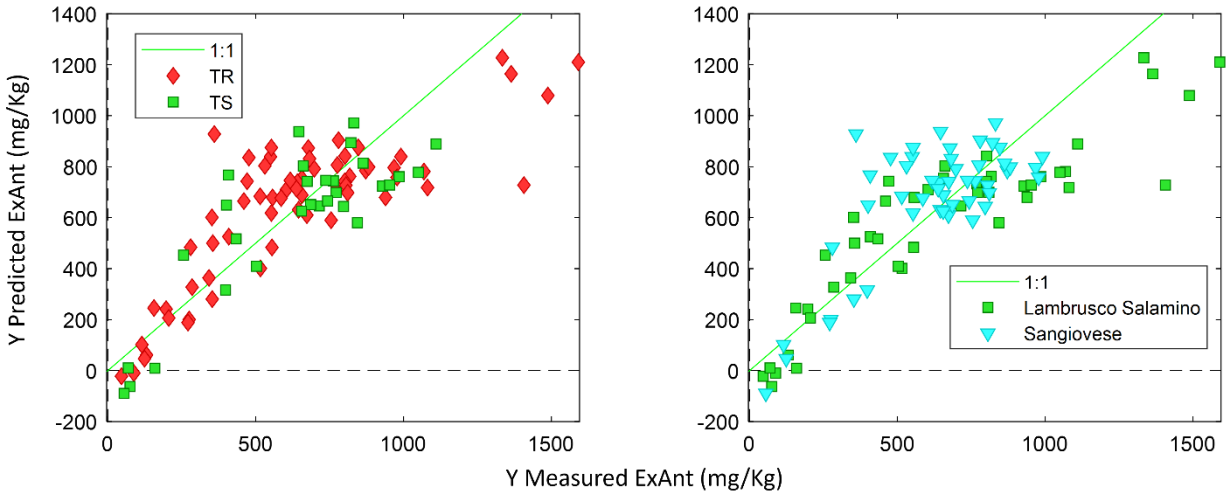


Figure 24 *Experimentally measured amount of extractable anthocyanins content (ExAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the PLS models calculated considering Lambrusco Salamino and Sangiovese varieties altogether. In the left part of the figure the samples are coloured according to training (TR) and test (TS) sets, while in the right part of the figure the samples are coloured according to the corresponding variety.*

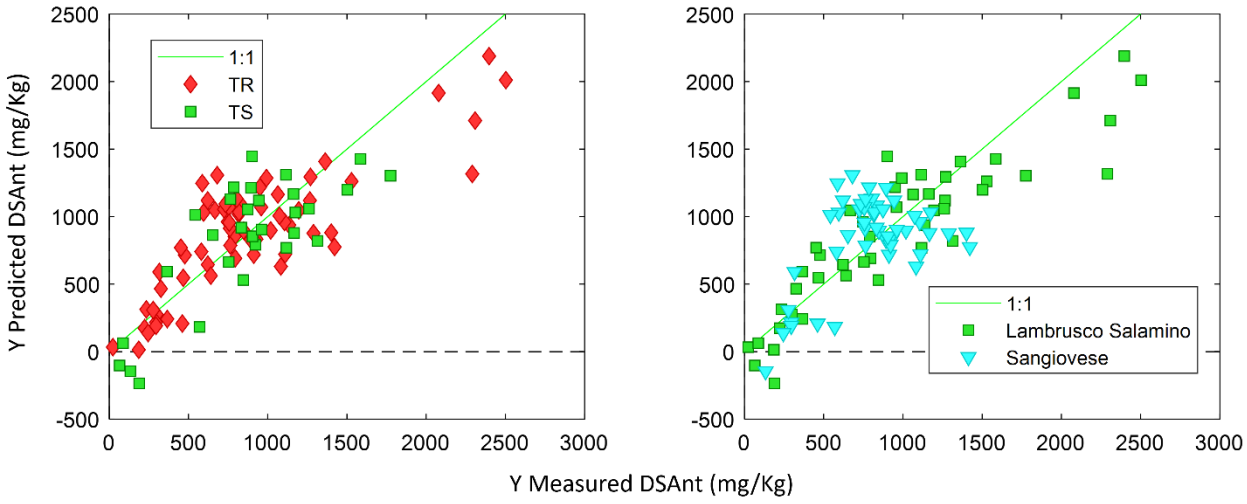


Figure 25 *Experimentally measured amount of total anthocyanins content determined according to the method proposed by Di Stefano (DSAnt, expressed as mg/Kg of malvidin-3-glucoside) vs. predicted amount of the PLS models calculated considering Lambrusco Salamino and Sangiovese varieties altogether. In the left part of the figure the samples are coloured according to training (TR) and test (TS) sets, while on right part of the figure the samples are coloured according to the corresponding variety.*

4 Conclusions

In the present study, we discussed the design and application of a portable device conceived to be a fast and user-friendly tool to support winegrowers in the assessment of anthocyanins content of red grapes directly in the vineyard. The hardware part of the device is composed of a portable acquisition box that allows to acquire RGB images of the grape samples under controlled lighting conditions using a smartphone. The core idea behind the device consists in relating the colour information of the grape samples contained in RGB images with the chemical parameters of interest using a data dimensionality reduction algorithm and multivariate calibration. All the data elaboration steps required to obtain the final prediction of the parameters of interest starting from the RGB images of the grape samples were implemented in the user-friendly software interface of the device. Through this software interface, the end user can easily acquire geolocated images of grape samples, view and share the results of the analysis in real time, without the need for any particular skills other than to carry out correct sampling.

In the phase of development and validation of the calibration models that have been implemented in the device, satisfactory results were obtained in the prediction of the chemical parameters of interest for the assessment of grape phenolic maturity of *Lambrusco Salamino* and *Sangiovese* varieties. Compared to the traditional methods used to evaluate grape phenolic ripening, the proposed device offers the possibility to perform much more frequent measurements and to get information in real-time about grape ripening. Furthermore, it must be considered that traditional analytical methods are quite expensive and time consuming, since they require a laborious sample preparation step to extract anthocyanins from the skin of grape berries. In addition, also these reference analytical methods are affected by a significant degree of uncertainty due to the low stability of the extracted grape must.

Conversely, lower calibration performances were obtained for *Ancellotta* grape variety, probably due to its exceptionally high concentration of anthocyanins in the skin of the berries. To improve the prediction performances for grape varieties with an early colour change and particularly high anthocyanin contents, like *Ancellotta*, it will be necessary to optimize *ad hoc* the conditions for image acquisition.

In general, in order to further increase the robustness of the calibration models, it will be necessary to acquire a larger number of samples, and to consider both the application of variable selection algorithms and the possibility of developing local models for specific vineyards or restricted geographic areas.

The proposed device provides winegrowers and all the stakeholders (like, e.g., winemakers) with relevant information about grape ripening in real-time, allowing to support a proper planning of the optimal harvest time and the organization of the entire winemaking process. In fact, it is possible not only to monitor time evolution and spatial variability of the parameters of interest, but also to make comparisons with stored data of previous vintages. In this manner, it is possible to have a real-time screening of grape phenolic ripening in a specific vineyard and move to laboratory analysis only if necessary.

For a practical application of the proposed device, it must be taken into account the possible variability due to the acquisition of the images using different smartphones. In this regard, the standardization procedure described in this study could be helpful to minimize colour variations and ensure the stability of the prediction models. In addition, it is also possible to consider the use of calibration transfer strategies, which are commonly used in Near Infrared (NIR) spectroscopy applications to transfer the calibration models from different devices. These aspects will be evaluated in future studies.

Acknowledgements

The present study was carried out in the frame of the Regional Rural Development Program 2014-2020 - Type of Operation 16.1.01 - Operational Groups of the European Innovation Partnership: "Agricultural productivity and sustainability" Focus Area 3A - Project "Technological innovation for efficient harvest forecasting - VITEVEN" (Programma Regionale di Sviluppo Rurale 2014-2020 – Tipo di Operazione 16.1.01 – Gruppi operativi del Partenariato Europeo per l'Innovazione: "produttività e sostenibilità dell'agricoltura" Focus Area 3A – Progetto "Innovazione tecnologica per una efficiente previsione vendemmiale - VITEVEN", CUP E48H19000830005).

The authors wish also to thank the team of KODE Solutions for the development of the software interface of the device and Dr. Paolo Piscolla (ASTRA) for the support with reference analyses on grape samples.

Author contributions

Camilla Menozzi: Methodology; Software; Formal analysis; Validation; Investigation; Data curation; Writing – Original Draft; Writing – Review & Editing. **Rosalba Calvini:** Methodology; Software; Formal analysis; Validation; Investigation; Data curation; Writing – Original Draft; Writing

– Review & Editing. **Giovanni Nigro:** Project administration; Funding acquisition Resources; Writing – Review & Editing. **Paola Tessarin:** Resources; Investigation; Writing – Review & Editing. **Domenico Bossio:** Resources; Investigation; Writing – Review & Editing. **Marco Calderisi:** Software; Data Curation; Resources; Visualization; Writing – Review & Editing. **Veronica Ferrari:** Investigation; Writing – Review & Editing. **Giorgia Foca:** Investigation; Writing – Review & Editing. **Alessandro Ulrici:** Conceptualization; Methodology; Validation; Investigation; Writing – Review & Editing; Supervision; Project administration; Funding acquisition.

5 References

Antonelli, A., Cocchi, M., Fava, P., Foca, G., Franchini, G.C., Manzini, D., Ulrici, A., 2004. Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm. Analytica Chimica Acta 515, 3–13. https://doi.org/10.1016/j.aca.2004.01.005

Baca-Bocanegra, B., Hernández-Hierro, J.M., Nogales-Bueno, J., Heredia, F.J., 2019. Feasibility study on the use of a portable micro near infrared spectroscopy device for the “in vineyard” screening of extractable polyphenols in red grape skins. Talanta 192, 353–359. https://doi.org/10.1016/j.talanta.2018.09.057

Bautista-Ortín, A.B., Fernández-Fernández, J.I., López-Roca, J.M., Gómez-Plaza, E., 2006. The effect of grape ripening stage on red wine color. OENO One 40, 15. https://doi.org/10.20870/oeno-one.2006.40.1.879

Borin, A., Ferrão, M.F., Mello, C., Cordi, L., Pataca, L.C.M., Durán, N., Poppi, R.J., 2007. Quantification of Lactobacillus in fermented milk by multivariate image analysis with least-squares support-vector machines. Anal Bioanal Chem 387, 1105–1112. https://doi.org/10.1007/s00216-006-0971-7

Brosnan, T., Sun, D.-W., 2004. Improving quality inspection of food products by computer vision—a review. Journal of Food Engineering 61, 3–16. https://doi.org/10.1016/S0260-8774(03)00183-3

Brosnan, T., Sun, D.-W., 2002. Inspection and grading of agricultural and food products by computer vision systems—a review. Computers and Electronics in Agriculture 36, 193–213. https://doi.org/10.1016/S0168-1699(02)00101-1

Calvini, R., Orlandi, G., Foca, G., Ulrici, A., 2020. Colourgrams GUI: A graphical user-friendly interface for the analysis of large datasets of RGB images. Chemometrics and Intelligent Laboratory Systems 196, 103915. https://doi.org/10.1016/j.chemolab.2019.103915

Calvini, R., Pigani, L., 2022. Toward the Development of Combined Artificial Sensing Systems for Food Quality Evaluation: A Review on the Application of Data Fusion of Electronic Noses, Electronic Tongues and Electronic Eyes. Sensors 22, 577. https://doi.org/10.3390/s22020577

Chen, S., Zhang, F., Ning, J., Liu, X., Zhang, Z., Yang, S., 2015. Predicting the anthocyanin content of wine grapes by NIR hyperspectral imaging. Food Chemistry 172, 788–793. https://doi.org/10.1016/j.foodchem.2014.09.119

Corti, M., Cavalli, D., Cabassi, G., Vigoni, A., Degano, L., Marino Gallina, P., 2019. Application of a low-cost camera on a UAV to estimate maize nitrogen-related variables. Precision Agric 20, 675–696. https://doi.org/10.1007/s11119-018-9609-y

de Jong, S., 1993. SIMPLS: An alternative approach to partial least squares regression. Chemometrics and Intelligent Laboratory Systems 18, 251–263. https://doi.org/10.1016/0169-7439(93)85002-X

Di Stefano, R., Cravero, M.C., 1991. Metodi per lo studio dei polifenoli dell'uva. Riv. Vitic. Enol 44, 37–45.

Di Stefano, R., Maggiorotto, G., 1995. Antociani, acidi idrossicinnamici e flavonoli del frutto, delle foglie, dei raspi e dei tralci della vite. Riv. Vit. Enol. 51–65.

dos Santos Costa, D., Oliveros Mesa, N.F., Santos Freire, M., Pereira Ramos, R., Teruel Mederos, B.J., 2019. Development of predictive models for quality and maturation stage attributes of wine grapes using vis-nir reflectance spectroscopy. Postharvest Biology and Technology 150, 166–178. https://doi.org/10.1016/j.postharvbio.2018.12.010

Drappier, J., Thibon, C., Rabot, A., Geny-Denis, L., 2019. Relationship between wine composition and temperature: Impact on Bordeaux wine typicity in the context of global warming—Review. Critical Reviews in Food Science and Nutrition 59, 14–30. https://doi.org/10.1080/10408398.2017.1355776

Foca, G., Masino, F., Antonelli, A., Ulrici, A., 2011. Prediction of compositional and sensory characteristics using RGB digital images and multivariate calibration techniques. Analytica Chimica Acta 706, 238–245. https://doi.org/10.1016/j.aca.2011.08.046

García-Beneytez, E., Revilla, E., Cabello, F., 2002. Anthocyanin pattern of several red grape cultivars and wines made from them. European Food Research and Technology 215, 32–37. https://doi.org/10.1007/s00217-002-0526-x

Garciafalcon, M., Perezlamela, C., Martinezcarballo, E., Simalgandara, J., 2007. Determination of phenolic compounds in wines: Influence of bottle storage of young red wines on their evolution. Food Chemistry 105, 248–259. https://doi.org/10.1016/j.foodchem.2006.11.006

Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Analytica Chimica Acta 185, 1–17. https://doi.org/10.1016/0003-2670(86)80028-9

Giraudo, A., Calvini, R., Orlandi, G., Ulrici, A., Geobaldo, F., Savorani, F., 2018. Development of an automated method for the identification of defective hazelnuts based on RGB image analysis and colourgrams. Food Control 94, 233–240. https://doi.org/10.1016/j.foodcont.2018.07.018

Glories, Y., 1993. Maturité phénolique du raisin, conséquences technologiques: application aux millésimes 1991 et 1992, in: Journée Technique Du CIVB: Actes Du Colloque. pp. 56–61.

Gonçalves Dias Diniz, P.H., 2020. Chemometrics-assisted color histogram-based analytical systems. Journal of Chemometrics 34. https://doi.org/10.1002/cem.3242

Gutiérrez, S., Tardaguila, J., Fernández-Novales, J., Diago, M.P., 2019. On-the-go hyperspectral imaging for the in-field estimation of grape berry soluble solids and anthocyanin concentration: Hyperspectral imaging for in-field grape estimation. Australian Journal of Grape and Wine Research 25, 127–133. https://doi.org/10.1111/ajgw.12376

Kennedy, J.A., Matthews, M.A., Waterhouse, A.L., 2002. Effect of Maturity and Vine Water Status on Grape Skin and Wine Flavonoids. Am J Enol Vitic. 53, 268. https://doi.org/10.5344/ajev.2002.53.4.268

Kontoudakis, N., Esteruelas, M., Fort, F., Canals, J.M., De Freitas, V., Zamora, F., 2011. Influence of the heterogeneity of grape phenolic maturity on wine composition and quality. Food Chemistry 124, 767–774. https://doi.org/10.1016/j.foodchem.2010.06.093

Mattivi, F., Zulian, C., Nicolini, G., Valenti, L., 2002. Wine, Biodiversity, Technology, and Antioxidants. Annals of the New York Academy of Sciences 957, 37–56. https://doi.org/10.1111/j.1749-6632.2002.tb02904.x

Nogales-Bueno, J., Baca-Bocanegra, B., Rodríguez-Pulido, F.J., Heredia, F.J., Hernández-Hierro, J.M., 2015. Use of near infrared hyperspectral tools for the screening of extractable polyphenols in red grape skins. Food Chemistry 172, 559–564. https://doi.org/10.1016/j.foodchem.2014.09.112

Orlandi, G., Calvini, R., Foca, G., Pigani, L., Vasile Simone, G., Ulrici, A., 2019. Data fusion of electronic eye and electronic tongue signals to monitor grape ripening. Talanta 195, 181–189. https://doi.org/10.1016/j.talanta.2018.11.046

Orlandi, Giorgia, Calvini, R., Foca, G., Ulrici, A., 2018. Automated quantification of defective maize kernels by means of Multivariate Image Analysis. Food Control 85, 259–268. https://doi.org/10.1016/j.foodcont.2017.10.008

Orlandi, G., Calvini, R., Pigani, L., Foca, G., Vasile Simone, G., Antonelli, A., Ulrici, A., 2018. Electronic eye for the prediction of parameters related to grape ripening. Talanta 186, 381–388. https://doi.org/10.1016/j.talanta.2018.04.076

Pampuri, A., Tugnolo, A., Giovenzana, V., Casson, A., Guidetti, R., Beghi, R., 2021. Design of cost-effective LED based prototypes for the evaluation of grape (Vitis vinifera L.) ripeness. Computers and Electronics in Agriculture 189, 106381. https://doi.org/10.1016/j.compag.2021.106381

Pérez-Magariño, S., González-San José, M.L., 2006. Polyphenols and colour variability of red wines made from grapes harvested at different ripeness grade. Food Chemistry 96, 197–208. https://doi.org/10.1016/j.foodchem.2005.02.021

Pérez-Magariño, S., González-San José, M.L., 2004. Evolution of Flavanols, Anthocyanins, and Their Derivatives during the Aging of Red Wines Elaborated from Grapes Harvested at Different Stages of Ripening. J. Agric. Food Chem. 52, 1181–1189. https://doi.org/10.1021/jf035099i

Prats-Montalbán, J.M., de Juan, A., Ferrer, A., 2011. Multivariate image analysis: A review with applications. Chemometrics and Intelligent Laboratory Systems 107, 1–23. https://doi.org/10.1016/j.chemolab.2011.03.002

Prats-Montalbán, J.M., Ferrer, A., 2007. Integration of colour and textural information in multivariate image analysis: defect detection and classification issues. J. Chemometrics 21, 10–23. https://doi.org/10.1002/cem.1026

Ribera-Fonseca, A., Noferini, M., Jorquera-Fontena, E., Rombolà, A.D., 2016. Assessment of technological maturity parameters and anthocyanins in berries of cv. Sangiovese (Vitis vinifera L.) by a portable vis/NIR device. Scientia Horticulturae 209, 229–235. https://doi.org/10.1016/j.scienta.2016.06.004

Ribéreau-Gayon, P., 1965. Le dosage des anthocyanes dans le vins rouge. Bull. Soc. Chim. 9, 2649.

Ribéreau-Gayon, P., Dubourdieu, D., Donèche, B., 2006. Handbook of enology, 2nd ed. ed. John Wiley, Chichester, West Sussex, England ; Hoboken, NJ.

Río Segade, S., Soto Vázquez, E., Díaz Losada, E., 2008. Influence of ripeness grade on accumulation and extractability of grape skin anthocyanins in different cultivars. Journal of Food Composition and Analysis 21, 599–607. https://doi.org/10.1016/j.jfca.2008.04.006

Rodríguez-Pulido, F.J., Mora-Garrido, A.B., González-Miret, M.L., Heredia, F.J., 2022. Research Progress in Imaging Technology for Assessing Quality in Wine Grapes and Seeds. Foods 11, 254. https://doi.org/10.3390/foods11030254

Saint Cricq de Gaulejac, N., Vivas, N., Glories, Y., 1998. Maturation phénolique des raisines rouges. Relation avec la qualité des vins. Comparison des cépages Merlot et Tempranillo. Progrès Agricole et Viticole 115, 306–318.

Sommer, S., Cohen, S., 2018. Comparison of Different Extraction Methods to Predict Anthocyanin Concentration and Color Characteristics of Red Wines. Fermentation 4, 39. https://doi.org/10.3390/fermentation4020039

Ulrici, A., Foca, G., Ielo, M.C., Volpelli, L.A., Lo Fiego, D.P., 2012. Automated identification and visualization of food defects using RGB imaging: Application to the detection of red skin defect of raw hams. Innovative Food Science & Emerging Technologies 16, 417–426. https://doi.org/10.1016/j.ifset.2012.09.008

Vrochidou, E., Bazinas, C., Papakostas, G.A., Pachidis, T., Kaburlasos, V.G., 2021. A Review of the State-of-Art, Limitations, and Perspectives of Machine Vision for Grape Ripening Estimation. Engineering Proceedings 9, 2. https://doi.org/10.3390/engproc2021009002

5.2 Comparison of colour and texture feature extraction methods to predict anthocyanins content in grapes

What follows is the integral content of: Menozzi, C., Prats-Montalbán, J. M., Calvini, R., Ulrici, A. Comparison of colour and texture feature extraction methods to predict anthocyanins content in grapes, submitted to *Chemometrics and Intelligent Laboratory Systems*

Comparison of colour and texture feature extraction methods to predict anthocyanins content in grapes

Camilla Menozzi[a], José Manuel Prats-Montalbán[b], Rosalba Calvini[a,c,], Alessandro Ulrici[a,c]

[a] *Department of Life Sciences, University of Modena and Reggio Emilia, Via Amendola, 2, Reggio Emilia, Italy*
[b] *Multivariate Statistical Engineering Group, Department of Applied Statistics, Operations Research and Quality, Universitat Politècnica de València. Camino Vera s/n, València, Spain*
[c] *Consorzio Interuniversitario Nazionale per la Scienza e Tecnologia dei Materiali (INSTM), Via Giusti, 9, 50121, Firenze, Italy*

Abstract

Colour and texture are the two main sources of information that can be retrieved from RGB images of food products. Different approaches are available to analyse the image properties based on the extraction of colour and texture features, and the selection of the most appropriate method is a critical point, since it could significantly impact the outcomes. Furthermore, when dealing with large datasets of images that have to be analysed simultaneously it is necessary to work at the image-level by applying data dimensionality reduction methods that allow to code the colour or texture properties of interest into a feature vector. In this context, the present study has three main objectives. Firstly, we propose an innovative data dimensionality reduction method to extract and codify the texture features of an RGB image into a one-dimensional signal, named texturegram (TXG). Then, TXG approach is compared with different image-level feature extraction methods, such as colourgrams (CLG), Soft Colour Texture Descriptors (SCTD) and Grey Level Co-occurrence Matrices (GLCM). These different image-level techniques were used to analyse a benchmark dataset of RGB images already considered in a previous study to build Partial Least Squares (PLS) models and relate the image features with anthocyanins content of red grape samples. Furthermore, we also investigated the possible advantages of combining the colour and texture information brought by the different image-level techniques using both low-level and mid-level data fusion. The PLS models were calculated considering different partitions of the RGB image dataset into training and test set, and the outcomes

of the different models were statistically evaluated by means of Analysis of Variance (ANOVA) and Principal Component Analysis (PCA).

Keywords: RGB imaging, colour, texture, data reduction, data fusion

1 Introduction

The use of Red, Green, Blue (RGB) digital imaging has rapidly increased in the field of analytical chemistry to inspect food quality (Cubero et al., 2015; Meenu et al., 2021; Wu and Sun, 2013; Zheng et al., 2006). As a matter of fact, the visual aspect of food plays a significant role in quality assessment and, at the same time, automated analyses are highly needed for food monitoring. The attractiveness of RGB imaging lies in its ability to provide an objective evaluation of food products through technically sophisticated systems that are accessible at relatively low cost.

Colour and texture are two of the most important sources of information contained in RGB images. They are inextricably linked and are always present in an image, even if one property can sometimes prevail over the other. In a greyscale image, the concept of tone is based on the varying shades of grey of the pixels in an image, while texture is concerned with the spatial distribution of grey tones (Haralick et al., 1973). The same principles can be extended to RGB images, where colour corresponds to pixels intensity values in the red, green and blue channels, while texture can be defined as the spatial variation of pixels colour intensities, i.e., the dependence between pixels and their neighbouring ones in terms of colour variation (Zheng et al., 2006).

RGB image analysis involves processing colour-related and texture-related information of RGB images, and it can be performed at two different levels, i.e., at the pixel-level and at the image-level.

Pixel-level approaches consider each pixel of an image as a separate sample that can be characterized by its colour, intensity or location, and these properties can be simultaneously evaluated according to Multivariate Image Analysis (MIA) (Prats-Montalbán et al., 2011).

On the other hand, image-level approaches consist in extracting different types of colour or texture features from images, to provide an appropriate image characterisation and to consider the whole image as a single object. In this way, the information of interest contained in each image is encoded into a feature vector containing colour and/or texture descriptors. This approach is particularly useful when it is necessary to simultaneously compare a large number of images and to gain a deep understanding of the whole dataset structure. As a matter of fact, in practical applications, since each image is often supposed to enclose the properties of a given sample, the possibility to get an

immediate and overall estimate of a sample is preferable, focusing on the single pixels only when necessary (Calvini et al., 2016a).

Colour features can generally be extracted from different colour spaces, of which RGB (red, green, blue) and HSI (hue, saturation, intensity) are the most common (Rolando et al., 2023). In this context, some of the authors have previously proposed the colourgrams method as an image-level approach accounting for colour information. Based on this method, each image is transformed into a *colourgram* (CLG), which is a one-dimensional signal of 4900 points encoding the colour information of the corresponding image using the frequency distribution curves of several colour-related descriptors, including R, G and B values, the relative colours, H, S and I values, and the score values of different Principal Component Analysis (PCA) models calculated on RGB data (Antonelli et al., 2004; Biancolillo et al., 2023; Borin et al., 2007; Lopes et al., 2020; Orlandi et al., 2019a). Hence, the colourgrams approach allows a comprehensive investigation of many colour-related properties of an image together with a significant reduction of image data dimensionality, enabling the possibility to handle large datasets of RGB images with a relatively low computational effort. This represents a great advantage of colourgrams approach, since several colour-related parameters are simultaneously evaluated without the need of defining *a priori* which are the most suitable for the problem at hand. It is then possible to select the most relevant colour descriptors *a posteriori* using variable selection methods. A simplified version of colourgrams has been recently proposed in literature based only on some colour histograms (i.e., R, G, B, L, H, S, I or a subset of these parameters) and excluding the part of the signal obtained through the data elaboration by PCA (Gonçalves Dias Diniz, 2020). However, this simplified approach requires available information about the best possible combination of colour parameters for a given problem (Gonçalves Dias Diniz, 2020). Furthermore, different applications of the colourgrams approach demonstrated the relevance and usefulness of the PCA part of the signal (Giraudo et al., 2018; Orlandi et al., 2018).

Concerning the extraction of image texture features, the available approaches can be grouped into four main categories: statistical methods, structural methods, transform-based methods and model-based methods (Bharati et al., 2004).

Statistical methods are the most common ones due to the relative simplicity and immediacy associated with these techniques. Based on the assumption that the spatial distribution of grey levels in an image is one of the qualities defining texture, statistical methods involve the extraction of a set of statistical parameters from such distributions. Statistical methods can be further classified into first-order, second-order and higher-order statistics (Tomita and Tsuji, 1990).

First-order statistics describe the distribution of individual pixel values, independently of the spatial relationship between pixels (Bevk and Kononenko, 2002; Kononenko and Kukar, 2007; Srinivasan and Shobha, 2008). This category encompasses the calculation of Soft Colour-Texture Descriptors (SCTD), a set of statistical features describing colour and texture in a defined colour space. As an example, SCTD can include the mean and 2^{nd} to 5^{th} order mathematical central moments of the single R, G and B channels or of colour parameters of colour spaces different from RGB. SCTD resulted in a successful approach to describe well-spread random textures (López, 2008; López et al., 2005).

Second-order statistics are more suitable to describe non-random textures since they account for the spatial interaction between pairs of pixels in an image (Aggarwal and K. Agrawal, 2012; Prats-Montalbán et al., 2011). The most common second-order statistics are the texture features that can be extracted from Grey Level Co-occurrence Matrices (GLCM), such as contrast, correlation, energy, homogeneity and entropy among others (Haralick et al., 1973; Chandini and Maheswari B., 2018; Malegori et al., 2016; Paliwal et al., 2003).

Nevertheless, the search and the development of other approaches for texture evaluation has not yet come to an end. In this work, we propose an alternative statistical method to extract texture features from RGB images, which consists in codifying the image texture properties into a one-dimensional signal, hereinafter named *texturegram* (TXG). The proposed approach is based on a combination between a modified version of the approach proposed by (Bharati and MacGregor, 2001) and lately extended to RGB images by (Prats-Montalbán and Ferrer, 2007), and a data dimensionality reduction algorithm of hyperspectral images proposed by (Calvini et al., 2016b).

According to Bharati and MacGregor (Bharati and MacGregor, 2001) each pixel of a grey level image can be characterized not only by its intensity values but also by the intensity values of neighbouring pixels. In this manner it is possible to capture the spatial information associated with each pixel by linking its grey level value with the grey level values of neighbouring pixels. Prats-Montalban and Ferrer (Prats-Montalbán and Ferrer, 2007) applied the same approach to the three R, G and B channels composing an RGB image in order to obtain textural information from such data.

In this study we applied a modified version of the above-mentioned methods, since for each channel we account for texture information by calculating for each pixel the differences between its intensity values and the intensity values of neighbouring pixels in some specific direction. The resulting outcome can be seen as a multichannel image where each channel corresponds to the pixel intensity differences calculated for a given spatial direction in one of the original RGB channels. In order to retrieve useful information from such multichannel images we applied the Common Space Hyperspectrograms (CSH) approach, an algorithm initially proposed to perform data dimensionality

reduction of hyperspectral images (Calvini et al., 2016b). The principle behind CSH consists in converting each hyperspectral image into a one-dimensional signal, the hyperspectrogram, obtained by merging in sequence the frequency distribution curves of quantities (i.e., scores, Hotelling T^2 and Q residual values) obtained from a PCA model common to all the images of the dataset. Since the multichannel images obtained by computing the differences between neighbouring pixels in the original RGB domain account for texture information, in this case the signals obtained from CSH approach are referred to as *texturegrams*. Therefore, each texturegram acts like a fingerprint of texture properties of the corresponding RGB image.

When evaluating food quality, it is difficult to establish the most appropriate method for image feature extraction given the very wide variety of food products and issues to deal with. The selection of the most relevant features to be analysed is a critical point, since it could have a significant influence on the outcomes. Different techniques allow to focus on different features and one or several methods might be employed together to acquire image features for food quality evaluation (Zheng et al., 2006).

In this work we applied different image-level techniques to analyse colour and texture features of a benchmark dataset of RGB images already considered in a previous study (Menozzi et al., 2023), where the colourgrams approach has been successfully applied to predict anthocyanins content of red grape samples from RGB images of grape berries acquired with a smartphone-based device. Although in this type of dataset colour is the primary aspect to consider when assessing anthocyanins content, we believed that texture-related information could also help to improve the predictive ability of the calibration models. For example, the presence of the green pedicels of the grape berries, which have the same colour as unripe berries but different shapes, and/or the image background, could negatively affect the model performance if only colour is considered.

The aim of this work is to compare the results obtained from different feature extraction methodologies, namely based on colourgrams (CLG), Soft Colour-Texture Descriptors (SCTD), second order statistics extracted from co-occurrence matrices (GLCM) and texturegrams (TXG), and to investigate the possible advantages and usefulness of combining the colour and texture information of the dataset by low-level and mid-level data fusion.

2 Materials and methods

2.1 Samples

In this work, we considered grape samples of *Sangiovese* grape variety, as it is one the most representative grape varieties used for wine-production in Emilia-Romagna region (Italy).

Samples were collected in the province of Ravenna (Italy). Grape samples were collected at different ripening times, from about 30% veraison up to complete ripeness during vintages 2020 and 2021. Two different vineyards were involved in vintage 2020 (Vin A and Vin B), while an additional vineyard (Vin C) was included in 2021 vintage for a total of 3 vineyards. The three vineyards were located in different areas and 10 harvest times were considered for each of them in each vintage, for a total of 50 grape samples (i.e., 20 samples in vintage 2020 + 30 samples in vintage 2021).

For each grape sample six field aliquots were collected and separately imaged (*see* Section 2.3).

2.2 Measurement of total anthocyanins

Total anthocyanins content (TAnt) was determined according to Glories method (Glories, 1993; Saint Cricq de Gaulejac et al., 1998; Menozzi et al., 2023). For each sample, grape berries were put in a blender (Kenwood CH-58) and mashed to homogenize pulp and skin. One half of the so-obtained grape mash was macerated in a pH 1 solution to have complete extraction of anthocyanins. After 4 hours of maceration at room temperature under low agitation, the solution was filtered on glass wool. Then, a spectrophotometric determination of anthocyanins (JASCO V-530, Tokyo, Japan) was performed on the clear juice using the methodology based on bleaching with sulphur dioxide (Ribéreau-Gayon, 1965; Río Segade et al., 2008; Sommer and Cohen, 2018) after which the total anthocyanins content (TAnt, expressed as mg/kg of malvidin-3-glucoside) was determined.

2.3 Image acquisition

The RGB images of the grape samples were acquired with a portable smartphone-based device specifically developed for on-field determination of grape phenolic maturity (Menozzi et al., 2023). Briefly, the acquisition device consists in a black case made of acrylonitrile butadiene styrene (ABS) with size 16 cm × 10 cm × 15.5 cm. The grape berries were placed inside a drawer (sample holder) positioned in the lower part of the acquisition device while the smartphone was positioned in a specific case located in the sliding lid on the top of the device. Controlled illumination conditions were ensured by a strip of white light emitting diodes (LED) placed inside the case and powered by a power bank fixed outside the device.

Image acquisition was performed using a dedicated app installed on a Xiaomi Redmi Note 8 smartphone (Xiaomi, Cina), equipped with a 48 megapixels camera with CCD technology, a spatial resolution of 8000 × 6000 pixels, and considering a shutter speed of 1/50 seconds, f/1.9 lens aperture

and ISO 200. A cardboard of standard colour references was included in each image scene, to perform image correction (*see* Section 2.4).

Each one of the 6 aliquots of a sample was separately imaged, for a total of 300 acquired images (= 6 aliquots × 50 grape samples). The previous analysis of this dataset (Menozzi et al., 2023) allowed to identify as outlier one grape sample that was removed. Therefore, in this study we excluded from further elaborations the 6 aliquot images obtained from the outlier sample and only the remainder 294 images were kept (= 49 samples × 6 aliquots).

For more details about the image acquisition system and the smartphone app, the reader is referred to (Menozzi et al., 2023).

2.4 Image correction

As described in the previous section, the scene of each image was composed by two different areas: the area with the cardboard of Standard Colour References (SCR) and the sample area with the grape berries, as shown in Figure 1. The SCR consisted of a white paper sheet including 12 square patches of various colours and it was used, after being properly selected and cropped from the original image, for image correction. This procedure is highly recommended to minimize the possible variability among images, principally owed to instrumental instability over time.

The first step of image correction consisted in the selection of an image referred to as *master image*, which is an image of the dataset whose SCR area was considered as reference for the further correction steps. In fact, the SCR of the master image was used to correct all the other images, referred to as *slave images*.

The correction algorithm was based on an internal calibration performed channel by channel, between median vectors calculated for red (R), green (G) and blue (B) channels of the SCR of the *master* image and the median vectors of the SCR of each remainder *slave* image. The SCR of each slave image was corrected to be as similar as possible to the SCR of the master image only when significant differences were detected; in this case, the same correction function was applied to correct each pixel of the corresponding sample area. After some preliminary tests, a first-order polynomial was considered for image correction.

The standardization procedure was implemented in the software of the device used for image acquisition, to automatically correct the acquired images when necessary.

For a more detailed description about the algorithm employed to standardize the images, the reader is referred to (Orlandi et al., 2018); the image correction algorithm is available in RGB Correction GUI, a graphical user-friendly interface working under MATLAB environment (The MathWorks Inc., USA) and downloadable for free from https://www.chimslab.unimore.it/downloads/.

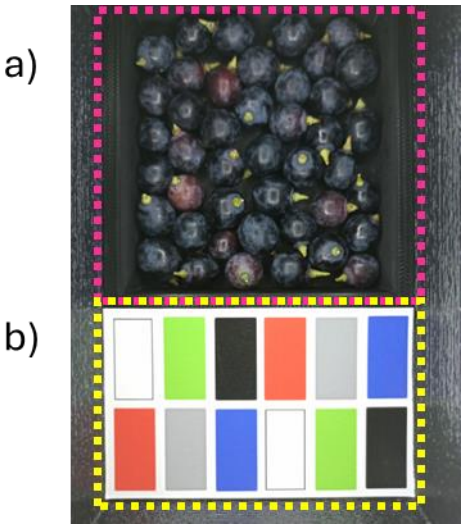


Figure 1 *Example of an RGB image of an aliquot of grape samples. In a) sample area is highlighted with a magenta dotted rectangle and in b) SCR area is highlighted with a yellow dotted rectangle.*

3 Image-level extraction of colour and texture features

After image correction, the sample area (i.e., the area containing the grape berries) was cropped from the corrected images obtaining smaller images of size 2310 × 2623 pixels. Only the cropped images containing the grape berries were considered for further elaboration, which consists in the extraction of colour and texture features using different techniques, i.e., colourgrams (CLG, *see* Section 3.1), texturegrams (TXG, *see* Section 3.2), Soft Colour Texture Descriptors (SCTD, *see* Section 3.3) and Grey Level Co-occurrence Matrices (GLCM, *see* Section 2.5.4). All the considered approaches perform an image-level extraction of the features of interest, which consists in calculating for each image a feature vector accounting for colour or texture descriptors. These feature vectors obtained from all the images of the dataset can be collected into a data matrix, which in turn can be analysed using unsupervised or supervised chemometric methods.

3.1 Colourgrams

The cropped images of the 6 aliquots of the same sample were concatenated into a single image with size 2310 × 15738 pixels (hereinafter named *scanset*, *see* Figure 2) to better account for sample

variability within the same ripening stage and thus obtain more robust and reliable results when developing the calibration models. Overall, a dataset including 49 *scanset* images was obtained.

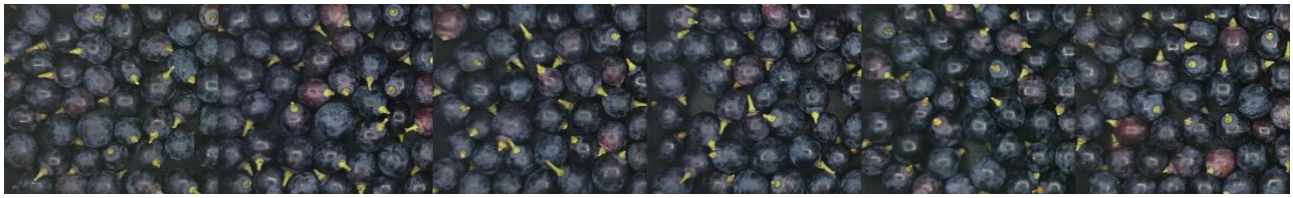


Figure 2 *Example of a scanset, obtained by concatenating the 6 corrected and cropped images of the aliquots of the same sample.*

Each *scanset* can be converted into the corresponding colourgram, i.e. a one-dimensional 4900 points-long signal, codifying the colour-related information of an RGB image and allowing significant data compression.

At first, each RGB image is unfolded in a two-dimensional matrix which contains all the pixels in the rows and the R, G, B channels in the columns (Giraudo et al., 2018). Then, 16 additional colour-related parameters are calculated for each pixel and derived from the R, G and B values: lightness (L = R + G + B), relative colours (rR = R / L; rG = G / L; rB = B / L), the hue, saturation and intensity values of the corresponding HSI colour space, and the 3 PCA score vectors calculated on the RGB matrix considering raw ($SC1_R$, $SC2_R$, $SC3_R$), mean centred ($SC1_M$, $SC2_M$, $SC3_M$) and autoscaled data ($SC1_A$, $SC2_A$, $SC3_A$), as there is no *a priori* knowledge of which pretreatment performs best. For each PCA model, loading vectors and eigenvalues are calculated too. Therefore, 19 colour-related parameters are considered: the 3 R, G and B channels and the 16 additional parameters mentioned above.

For each colour-related parameter, the corresponding frequency distribution curve (i.e., histogram) is calculated considering 256 bins. Then, after normalising by the number of pixels, the frequency distribution curves of each *scanset* image are merged in sequence to obtain a one-dimensional signal which composes the first part of the colourgram. Finally, the loading vectors and eigenvalues of the 3 PCs for the 3 PCA models calculated on RGB data are added at the end of the colourgram to complete the signal.

Following the procedure described above, each *scanset* image was converted into the corresponding colourgram and the signals were collected into a data matrix with size 49 × 4900, where each row corresponds to the colourgram derived from a specific *scanset* image. The colourgrams matrix can be analysed by multivariate analysis techniques enabling to evaluate all the images together and consider

the colour-related information of the dataset as a whole. Even if the computation of colourgrams implies to calculate multiple parameters from R, G and B values, the procedure is easily automatable and the computational time required is quite low, allowing an effective and comprehensive characterisation of food samples without requiring *a priori* knowledge of the properties of interest for the given task.

Figure 3 shows the colourgrams matrix obtained in this study, where each signal is coloured according to the anthocyanins content of the corresponding grape samples, and the vertical dashed lines define the colourgram regions related to the histograms of the 19 colour-related parameters, identified with the corresponding abbreviations.

For a deeper insight of the algorithm developed to create the colourgrams, the reader is referred to (Antonelli et al., 2004; Calvini et al., 2020).

Colourgrams were calculated using Colourgrams GUI (Calvini et al., 2020), a graphical user interface written in MATALAB language (The Mathworks Inc., USA) and freely downloadable from https://www.chimslab.unimore.it/downloads/.

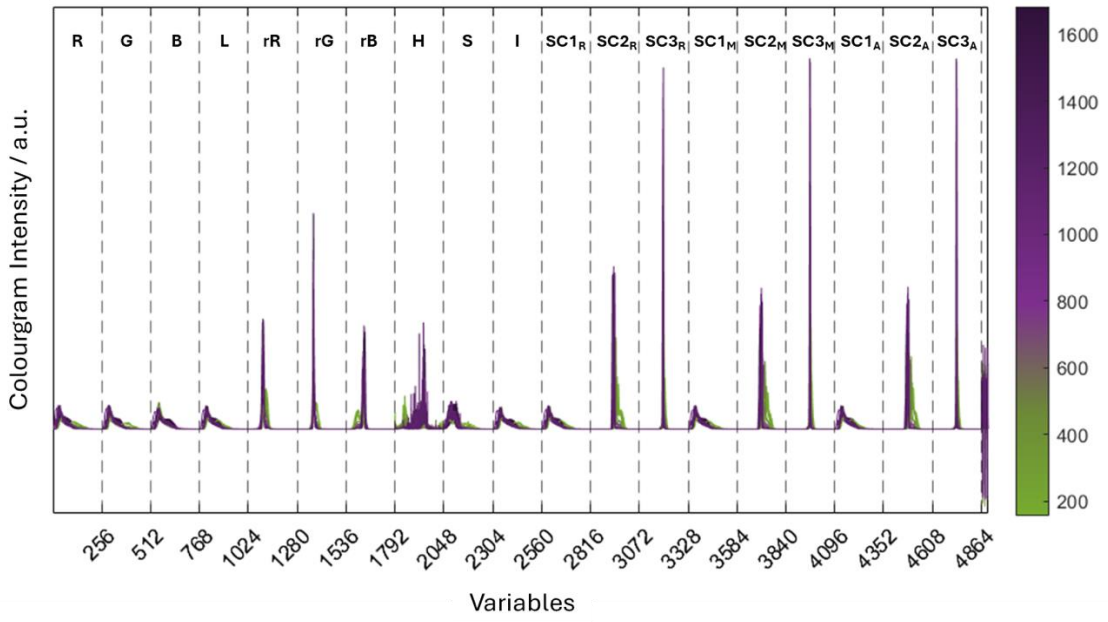


Figure 3 *Colourgrams calculated starting from the scansets; each signal is coloured according to the total anthocyanins content of the corresponding sample.*

3.2 Texturegrams

In this work, we propose an innovative method to codify the texture features of an image, which is based on the same idea of colourgrams: converting each image into a one-dimensional signal, i.e., the *texturegram*, retaining the relevant information contained in the corresponding original image. As suggested by their name, the texturegrams take into account the texture-related properties of RGB images. The procedure used for the calculation of texturegrams involves two main steps: the conversion of each RGB image into a multichannel texture image accounting for texture properties and the application of a data dimensionality reduction algorithm to convert the multichannel texture images into one-dimensional signals.

3.2.1 Conversion of RGB images into multichannel texture images

In the first step of texturegram calculation, each original RGB image is converted into a multichannel image accounting for texture information using a modified version of the approach initially proposed by (Bharati and MacGregor, 2001) and lately extended to RGB images by (Prats-Montalbán and Ferrer, 2007).

More in detail, for each channel of a single RGB image, the spatial-textural information associated with each pixel is obtained by linking its intensity value with those of neighbouring pixels at a given distance d (Bharati and MacGregor, 2001). Considering for example the red channel, the corresponding grey scale image can be unfolded into a column vector with as many elements as the number of pixels and each element corresponds to the red value of a single pixel. For each pixel it is possible to register also the red values of neighbouring pixels at distance d selected according to an established order, i.e., the pixel above, then in the right-up corner, then on the right and so on, following the clockwise direction. In this manner, each pixel is characterized by 9 values: its red value and the red values of the 8 neighbouring pixels. This information can be calculated across all the image pixels obtaining a textural matrix with 9 columns, corresponding to the intensity value of each pixel and the intensity values of neighbouring pixels in the 8 considered directions.

The same procedure can also be extended to the remainder G and B channels of an RGB image (Prats-Montalbán and Ferrer, 2007) and the textural matrices obtained for each colour channel can be subsequently stacked one beside the other, obtaining a colour-textural RGB matrix with 27 columns (= 9 columns for each channel × 3 channels).

In this study, to extract the texture features from each RGB image we used the same principle described above but in a modified version. In fact, after unfolding, for each channel instead of

reporting the values of the considered pixel together with those of the neighbouring pixels, we described the texture properties of each pixel by calculating the *difference* between its intensity value and the intensity values of neighbouring pixels in all the 8 possible directions, following a clockwise order starting from the pixel above the considered pixel. As a result, for each channel we obtained a two-dimensional matrix with 8 columns, corresponding to the 8 spatial directions, and a number of rows which is lower than the number of pixels of the original image and depends on the considered distance between neighbouring pixels. Indeed, for a given distance *d*, we decided to discard the information related to nominal pixels located at a distance lower than *d* from the edges of the original image since they lack some neighbours. The number of rows (*r)* of this bidimensional matrix can be calculated as follows:

$$r = [n_r - (2 \times d)] \times [n_c - (2 \times d)] \quad (1)$$

where n_r and n_c represent the number of row and column pixels of the original image, respectively.

The bidimensional matrix accounting for pixel differences can be refolded back into the original image domain obtaining a multichannel texture image with $n_r - (2 \times d)$ rows, $n_c - (2 \times d)$ columns, and 8 channels (Figure 4b).

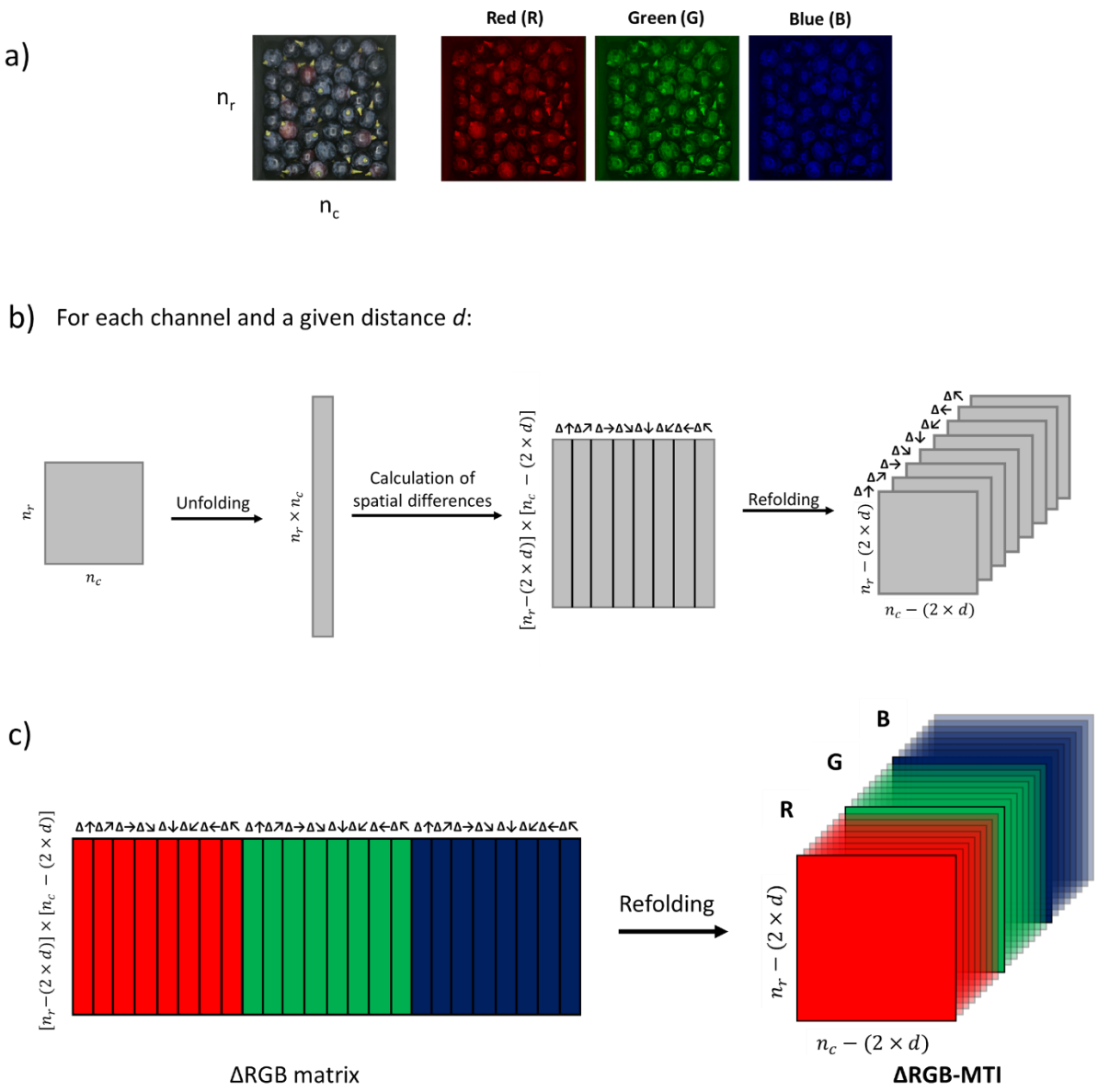


Figure 4 *Creation of a ΔRGB-MTI from an image of an aliquot: a) extraction of the red (R) green (G) and blue (B) channels; b) for each colour channel, calculation of the corresponding multichannel texture image accounting for the intensity differences between nominal pixels and the corresponding neighbours in the 8 different directions at a given distance d and c) ΔRGB-MTI obtained from the concatenation of the matrices obtained in b).*

This procedure was applied separately to all the three R, G and B channels of an image and the corresponding bidimensional difference matrices were staked one beside the other obtaining a new matrix, from here onwards named as ΔRGB matrix. Each ΔRGB matrix is a two-dimensional matrix with *r* rows (*see* Eq. 1) and 24 columns, corresponding to 8 directions ×3 channels.

Then, the ΔRGB matrix is refolded back into the original image domain obtaining a ΔRGB Multichannel Texture Image (ΔRGB-MTI), with $n_r - (2 \times d)$ rows, $n_c - (2 \times d)$ columns and 24 channels (Figure 4c). The procedure used to obtain the ΔRGB-MTI from RGB images is summarized in Figure 4.

In this manner, each RGB image is converted into the corresponding ΔRGB-MTI, a multichannel image accounting for texture properties by considering the colour differences between each pixel and its neighbours. These ΔRGB-MTIs can be analysed using the common MIA methods available for pixel-level exploratory analysis of multispectral and hyperspectral images.

However, when dealing which a large number of images altogether that need to be analysed simultaneously, it is necessary to work at the image level and apply proper data dimensionality reduction algorithms, which is the second step in texturegrams calculation.

3.2.2 Conversion of ΔRGB Multichannel Texture Images into signals

In the second step, the ΔRGB-MTIs are converted into the corresponding texturegrams using the Common Space Hyperspectrograms (CSH) algorithm, a data dimensionality reduction method originally proposed for the simultaneous analysis of hyperspectral images (Calvini et al., 2016b). The main idea behind the CSH approach is to convert each hyperspectral image into a hyperspectrogram, a one-dimensional signal accounting for the relevant information obtained by merging in sequence the frequency distributions of the score vectors, Q residuals and Hotelling T^2 values obtained from a PCA space common to all the images of the dataset.

The different steps in the computation of texturegrams using the CSH approach can be summarized as follows:

1. Pixel-wise mean centring of the ΔRGB-MTIs according to the grand mean, i.e. to the mean calculated for each variable (thus, for each of the 8 directional differences of intensity in each of the three colour bands) on all the retained pixels;
2. Calculation of the kernel variance-covariance matrix (Z) of the whole dataset. Z has size $\{24 \times 24\}$, where 24 is the number of channels of the ΔRGB-MTIs. To make this step computationally feasible, the variance-covariance matrices were separately calculated for each ΔRGB-MTI and subsequently summed entry-wise to obtain Z;
3. Decomposition of Z by singular value decomposition (SVD) to obtain the loading vectors of the common PC space. After some preliminary tests, in this work we decided to keep 16 PCs;

4. Calculation of the scores, Q residuals and Hotelling T^2 vectors of each ΔRGB-MTI by projecting the ΔRGB-MTIs onto the common PC space;
5. For each ΔRGB-MTI, computation of the frequency distribution curves for each score, Q residuals and Hotelling T^2 vector. For each vector, the corresponding frequency distribution curve was calculated considering a common range corresponding to the interval between the global minimum and the global maximum of the outputs of the PCA model obtained for all the images. After some preliminary tests, we considered 100 bins to obtain the frequency distribution curves (i.e., histograms) and each histogram was normalized according to the number of pixels of the corresponding image;
6. Creation of the texturegrams by joining in sequence the frequency distribution curves calculated in step 5.

In this work, each texturegram resulted in a 1800 points-long signal, resulting from 18 frequency distribution curves (16 PCs + Q residuals + Hotelling T^2) × 100 bins.

The procedure used to convert the ΔRGB-MTIs into texturegrams is schematically represented in Figure 5 where, for the sake of conciseness, only 2 PCs are considered to calculate the common PC space.

According to the approach proposed in this paper, a whole dataset of RGB images can be converted into a matrix of texturegrams, in which each row represents a sort of texture-fingerprint of the imaged sample. The texturegrams matrix can be easily processed by means of different multivariate statistical methods for exploratory analysis or for the calculation of supervised regression or classification models. In other words, each signal codifies a significant part of the texture features of interest contained in a specific image and it is possible to make quick comparisons of a high number of images at the same time.

As will be better detailed later in Section 4, in this study the PLS models were obtained after splitting the RGB image dataset into training and test images using different partitions. To ensure a real external validation of the models, the computation of the grand mean (step 1) and of the common PC space (step 2 and step 3) were performed considering only the ΔRGB-MTIs obtained from the training set images. When dealing with the test set images, the corresponding ΔRGB-MTIs were firstly mean centred according to grand mean of training data and then projected into the PCs space common to the whole training dataset (step 4). Then the texturegrams of the test set were calculated following steps 5-6.

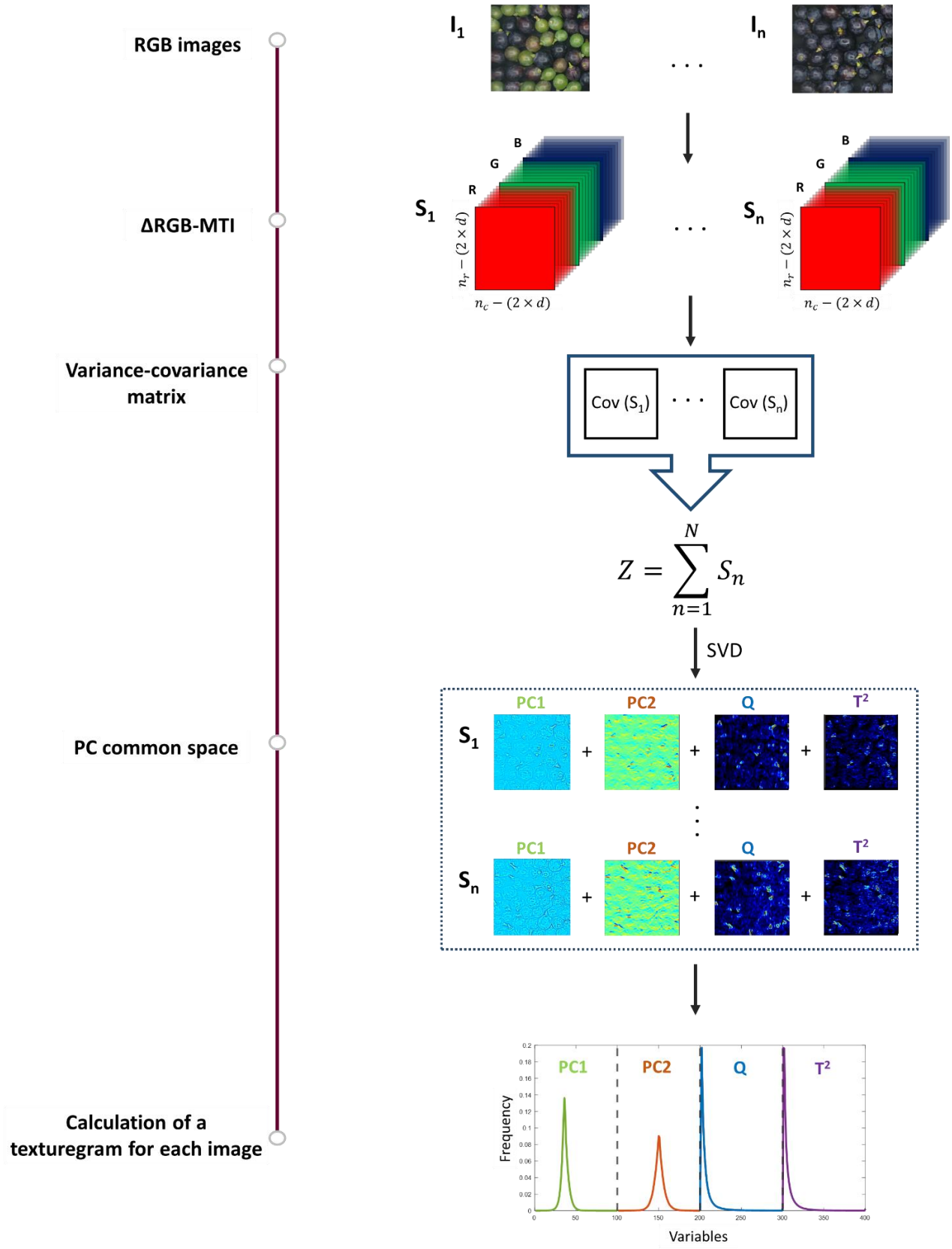


Figure 5 *Schematic overview of the whole process to calculate the corresponding texturegrams from RGB images; for the sake of conciseness, in this scheme the common PC space was calculated considering only 2 PCs by way of illustration.*

In this study, the texturegrams were firstly calculated separately for each one of the 6 aliquot images of a *scanset*. Then, the 6 texturegrams of the aliquot images of the same sample were averaged in order to obtain a single texturegram for each *scanset*.

Furthermore, in the computation of texturegrams it is necessary to define the distance *d* of neighbouring pixels used to convert the original RGB images into the corresponding ΔRGB-MTIs. In this study we considered three different distance values, corresponding to 5, 10 and 15 pixels respectively, and for each distance a texturegrams matrix with size {49 × 1800} was obtained.

As an example, Figure 6 shows the texturegrams matrix obtained in this study considering a distance of 15 pixels, where each signal is coloured according to the anthocyanins content of the corresponding grape sample.

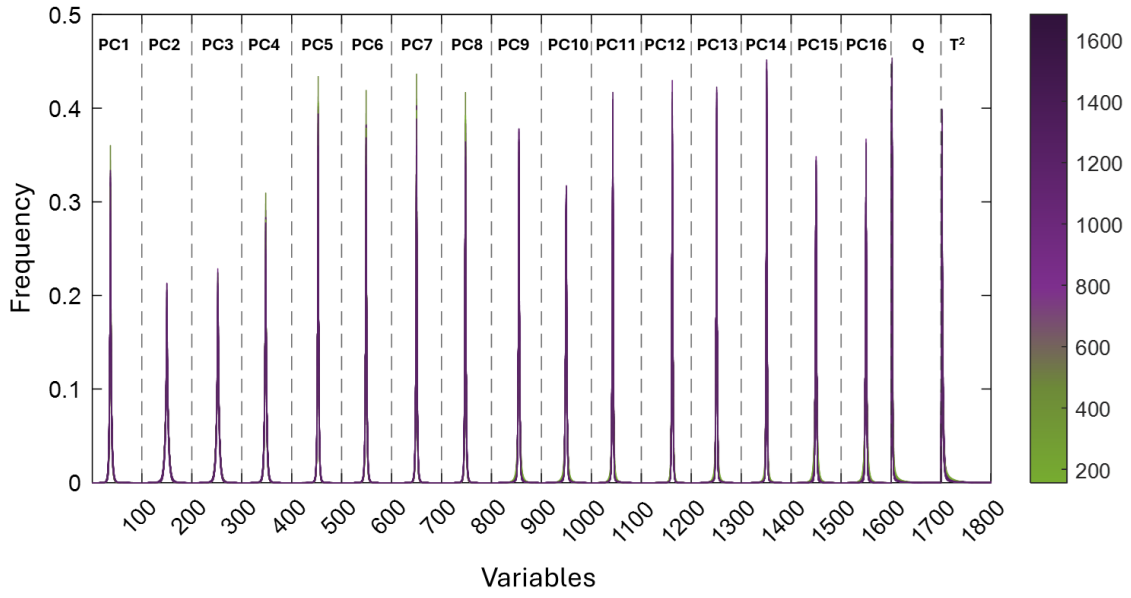


Figure 6 *Example of texturegrams calculated starting from the RGB images considered in this study and considering a distance of 15 pixels; each signal is coloured according to the total anthocyanins content of the corresponding sample.*

The conversion of RGB images into ΔRGB-MTIs was performed using a script written *ad hoc* in MATLAB language, while the conversion of ΔRGB-MTIs into signals was performed using Hyperspectrograms GUI, a graphical user interface working in MATLAB environment and available for the conversion of multispectral or hyperspectral images into CSH. Hyperspectrograms GUI is freely downloadable from https://www.chimslab.unimore.it/downloads/.

3.3 Soft Colour-Texture Descriptors

Soft Colour-Texture Descriptors (SCTD) (López, 2008; López et al., 2005) are a set of statistics requiring a relatively low computational demand to be extracted and elaborated when approaching colour and texture description in RGB images.

In fact, these features can be extracted by computing the mean and the different mathematical statistical central moments from 2nd to 5th (i.e., variance, skewness, kurtosis and 5th order) related to each vectorized image channel, for a total of 15 descriptors in an RGB image, i.e., (1 + 4) statistical features × 3 colour channels.

In this work, the SCTD were firstly calculated separately for each aliquot image. Then, for the 6 aliquot images of a *scanset* the mean and variance of each descriptor were calculated to consider also the internal variability of each descriptor within the same sample. In this manner, for each *scanset* a feature vector of 30 parameters was obtained (15 mean values of SCTD + 15 variances of SCTD). The so-obtained SCTD matrix has dimensions of {49 × 30}.

3.4 Grey Level Co-occurrence Matrices (GLCM)

A Grey Level Co-occurrence Matrix (GLCM) (Haralick et al., 1973) is a matrix containing texture features of each image and it is obtained using second order statistics. For a grey-scale image, the GLCM has a number of both rows and columns equal to the discrete grey level intensities of the image in order to embrace all the possible combinations of intensities that can appear between two image pixels at a given relative position. The relative pixel position has to be defined *a priori* before calculating GLCM, usually by setting a certain distance (d) and a certain angle (θ).

A GLCM can be also seen as a cross tabulation of a bidimensional histogram associated to the random distribution of the bidimensional variable (Prats-Montalbán et al., 2009). In fact, each element (i, j) of the GLCM expresses the total number of occurrences (i.e., the relative frequencies) for which an image pixel with intensity value i is coupled to another image pixel with intensity value j, at the given spatial relationship depending on the considered distance and angle. Several features can be extracted from the GLCM, although only a few are widely used and, in this study, we focused on five of the most common ones, namely contrast, correlation, energy (also known as Angular Second Moment), homogeneity and entropy (Clausi, 2002).

When defining the spatial relationship among pixels to be considered, the distance of 1 pixel is the most commonly used, while four directions are typically explored and correspond to angles of 0°, 45°, 90° and 135°, corresponding to four distinct GLCMs for each original greyscale image. Usually,

when the features sought are expected to be invariant to the direction, it is a common practice to average the statistical parameters extracted from the four GLCMs corresponding to the four angles.

All these principles can be easily extended to RGB images, by calculating GLCM features for each single channel.

In this work, GLCMs were firstly calculated separately for each aliquot image considering for each channel 256 intensity levels and the four directions listed above. For each channel, the values of the 5 statistical features (i.e., contrast, correlation, energy, homogeneity and entropy) extracted from the four GLCMs corresponding to the four directions were computed. Hence, for each aliquot image we obtained a final vector of 15 features (5 GLCM features × 3 colour channels).

Similarly to SCTD, for the 6 aliquot images of a single *scanset* the mean and variance of every GLCM parameter were calculated to consider the internal variability of every parameter within the same sample. In this manner, a final row feature vector of 30 parameters (15 average GLCM features + 15 variances of GLCM features) was obtained for each *scanset*.

This procedure was performed considering four distance values in GLCM computation corresponding to 1, 5, 10 and 15 pixels, which led to four separate datasets of GLCM descriptors with size {49 × 30}.

4 PLS calibration models

PLS calibration models were calculated in order to relate the colour and/or texture features extracted from the RGB images of grapes to the anthocyanin content of the corresponding samples. More in detail, the PLS models were calculated not only considering the different image-level feature extraction methods as separate data blocks (*see* Section 4.1) but also combining the colourgram approach with each one of the texture extraction methods (*see* Section 4.2).

For the computation of the PLS models, multiple partitions of the acquired RGB images into training and test images were considered. In fact, in order to make a detailed and accurate evaluation of strengths and weaknesses of the feature extraction methods considered in this study, the PLS calibration models were calculated by iteratively including in the test set only the samples collected in a specific vineyard and in a specific year. For example, the partition “Vin_A_2020” in Table 1 refers to the training/test set partition in which the *scanset* images of grape samples collected in vineyard A in 2020 were used as test images, while the remainder images were used as training images. In this manner, 5 different training/test set partitions were considered. More detailed information about the considered training/test set partitions is reported in Table 1. It is interesting to

highlight that in the partition "Vin_C_2021" the external validation considers as test set grape samples harvested in a vineyard not considered during calibration.

In addition, the same training/test set partition approach adopted in the previous work (Menozzi et al., 2023) was also considered for the sake of comparison. This training/test set partition is referred to as "Original" in Table 1 and from here onwards in the manuscript. It consists in attributing to the test set the *scansets* of different harvest times for each vineyard as follows: 34 *scansets* in the training set and 15 *scansets* in the test set. Originally, this training/test set partition was carried out with the purpose of encompassing in the model the possible different trends expected due to the sampling procedure: even within the same grape variety, the different geographical position of the vineyards and the climatic variability from one vintage to another could significantly impact the accumulation of anthocyanins in grape berries during ripening.

Table 1 *Information about the different training/test set partitions considered in this work. The partitioning names stand for the samples collected in a specific vineyard and year iteratively left in the external test set, while "Original" refers to the training/test partition considered in the previous work.*

TR/TS partition	N° training set samples	N° test set samples
Vin_A_2020	39	10
Vin_B_2020	39	10
Vin_A_2021	39	10
Vin_B_2021	39	10
Vin_C_2021	40	9
Original	34	15

The proper number of Latent Variables (LVs) was selected by minimizing the Root Mean Square Error in Cross-Validation (RMSECV).

According to the different training/test set partitions, specific cross-validation approaches were applied. When considering the "Original" training/test set partition, venetian blinds cross-validation with 5 deletions groups was performed. For the remainder partitions, a "one-vineyard-one-year-out" cross-validation was followed. In other words, for each training set, cross-validation was performed by considering 4 deletion groups, each one corresponding to the samples collected in only one vineyard in a specific year.

The results of the calibration models were expressed in terms of Root Mean Square Error calculated in cross-validation (RMSECV) and prediction of the external test set (RMSEP). Finally, the evaluation of the performance of the calibration models obtained using the different feature extraction techniques and their combinations was done by using both Analysis of Variance (ANOVA) and Principal Component Analysis (PCA).

All calculations were performed using the PLS Toolbox ver. 8.8.1 (Eigenvector Research Inc.,USA), the Statistics and Machine Learning Toolbox ver. 12.2 (The MathWorks Inc., USA), and ad hoc written functions running in the MATLAB ver. 9.11 environment (The Mathworks Inc., USA).

4.1 Single blocks

Initially, the PLS calibration models were calculated on each individual dataset accounting for colour or texture features of the original RGB images, i.e., colourgrams (CLG), texturegrams (TXG), Soft Colour Texture Descriptors (SCTD) and GLCM features (GLCM).

The dataset of colourgrams was pre-processed using mean centering, as in the previous study. As for texturegrams, three different datasets were obtained based on the distance value used for neighbouring pixels and corresponding to 5 pixels (TXG D5), 10 pixels (TXG D10) and 15 pixels (TXG D15). The PLS models were calculated for each TXG dataset using Pareto scaling and mean centring as preprocessing methods. Only one dataset of SCTD was obtained and it was pre-processed using autoscaling. Similarly to TXG, also for GLCM different distance values were considered, resulting in four GLCM datasets corresponding to distances equal to 1 pixel (GLCM D1), 5 pixels (GLCM D5), 10 pixels (GLCM D10) and 15 pixels (GLCM D15). The four GLCM datasets were pre-processed using autoscaling before the calculation of PLS models.

On the whole, 54 PLS models were calculated considering the single blocks, corresponding to 6 training/test set partitions × 9 single data blocks (1 CLG + 3 TXG + 1 SCTD + 4 GLCM).

4.2 Data fusion

In order to evaluate the possible advantages of integrating together the colour and texture information obtained from the different approaches starting from the same images, data fusion strategies were used to iteratively combine the information brought by CLG with each one of the texture analysis approaches described in the previous sections. Both low-level data fusion, which consists in merging

together the different blocks of data, and mid-level data fusion, which consist in combining features extracted from the original data blocks, were considered (Calvini and Pigani, 2022).

4.2.1 Low-level data fusion

Considering low-level data fusion, the single blocks were separately pre-processed as described in Section 4.1 and then concatenated into a unique data matrix, with a number of rows equal to the number of samples, and a number of columns equal to the sum of the variables from each block. Consequently, the size of the low-level fused matrices was equal to 49 rows (samples) and:

- 6700 columns when fusing CLG and TXG data blocks (= 4900 CLG variables + 1800 TXG variables); three CLG + TXG fused matrices were obtained, corresponding to the three distance values used to calculate TXG;
- 4930 columns when fusing CLG with SCTD data block (=4900 CLG variables + 30 SCTD variables);
- 4930 columns when fusing CLG with GLCM data blocks (=4900 CLG variables + 30 GLCM variables); four CLG + GLCM fused matrices were obtained, corresponding to the four distance values used to calculate GLCM.

For each combination of colour- and texture- related data blocks, after low-level data fusion the resulting data matrix was pre-processed using block scaling before the calculation of the corresponding PLS model.

A total of 48 low-level data fusion models were calculated, corresponding to 6 training/test set partitions × 8 combinations between CLG and each one of the remainder 8 datasets accounting for texture features.

4.2.2 Mid-level data fusion

Mid-level data fusion consists in concatenating features of interest extracted from the different data blocks by a preliminary data analysis step. To this aim, the outcomes of the calibration models previously calculated on the single blocks (*see* Section 4.1) have been combined to perform mid-level data fusion, by concatenating the PLS score vectors. In particular, the score vectors obtained from the PLS model calculated on CLG were combined with the score vectors of each PLS model calculated on a data block of texture features. Therefore, the size of each mid-level data fusion matrix was equal to 49 rows (samples) and as many columns as the sum of the number of LVs of the PLS model

calculated on the CLG block and the number of LVs of the PLS model of the other texture-related data block.

Considering mid-level data fusion, two different preprocessing strategies were tested on the fused datasets of scores vectors: autoscaling and mean centring followed by block scaling. When performing autoscaling all the single block PLS score vectors used as descriptors have the same scale, while with block scaling each block of data is scaled to unit variance. In other words, each block of data brings the same contribution to the calibration models regardless of the number of latent variables. At the same time, the relative scales of the score vectors within each block are kept (Orlandi et al., 2019b).

On the whole, 96 mid-level data fusion models were calculated, corresponding to 6 training/test set partitions × 8 combinations between CLG and texture data × 2 preprocessing methods.

4.3 Statistical analysis of calibration performances

ANOVA was performed on the RMSEP values of all externally validated models with the samples collected in a specific vineyard and in a specific year, *i.e.*, not considering the models calculated using the "Original" training/test set partition (*see* Section 4). This was aimed at better focusing on the significance of the possible variability in the results due to the different vineyards in each vintage.

In our case, an ANOVA without interaction was performed on the RMSEP values obtained by the PLS models calculated on the single blocks, in order to highlight any statistically significant difference between the data coming from the single feature extraction techniques and the vineyards in each vintage. Regarding the models obtained using data fusion approaches, a new ANOVA with interaction was used to evaluate whether the effects of the techniques, the data fusion approaches, the vineyards and/or their interactions were significant.

Moreover, to obtain an overall comparison of the outcomes, PCA was carried out to compare the results coming from the different techniques and data fusion approaches, taking into account all the training/test set partitions applied in this study. PCA was calculated on the whole set of RMSECV and RMSEP values obtained considering both the single feature extraction techniques and their combinations through data fusion, and all the 6 training/test set partitions. Since the RMSE values have the same scale, mean centring was chosen as preprocessing.

5 Results and discussion

5.1 PLS results

To relate the colour and textural information contained in the RGB images of grape samples with the corresponding amount of total anthocyanins, 198 PLS models were computed on the whole (= 54 single block models + 48 low-level data fusion models + 96 mid-level data fusion models) applying the different techniques and approaches described in the previous sections. The results of the corresponding PLS models are reported in Table 2 and expressed as RMSECV and RMSEP values associated to each model. For a more immediate interpretation of the results, the cells are coloured according to the corresponding RMSE values, from green for lowest RMSE values to red for highest RMSE values. For a more detailed visualization of the outcomes, the results of all the 198 PLS models are reported in Table 3-8 including also the corresponding number of LVs, RMSEC, and R^2 (calculated in calibration, cross-validation and prediction of external test set).

This section collects some general comments from a qualitative point of view of the results reported in Table 2. A more detailed and quantitative evaluation of the statistically significant differences between the RMSEP values of the different models is presented in Section 5.2.1, while Section 5.2.2 reports an overall evaluation of the RMSECV and RMSEP values by PCA.

Regarding the feature extraction technique, the highest RMSECV and RMSEP values are generally reached when considering as single blocks the GLCM features at all distances and SCTD. As a matter of fact, GLCM principally focuses on the textural features in an image, indeed too weak information for this specific application, since the progressive accumulation of anthocyanins in red grape has a main impact on the colour of berries rather than texture. On the other side, the immediate and overall estimate of the colour and texture properties of an image returned by SCTD does not appear representative enough to account for the intrinsic variability of the berries within the same grape sample/ripening stage and a more detailed image characterisation is required for the issue at hand. This latter aspect could explain the generally better RMSECV and RMSEP values reached when calculating the models considering CLG and TXG as single blocks. Indeed, these two types of signals seem to allow a more thorough representation of the image properties compared to GLCM and SCTD.

When moving from the PLS models calculated on single data blocks to those calculated after data fusion, there seems to be a general improvement of the results, especially when considering mid-level data fusion in cross-validation. However, it has to be mentioned that, interestingly, CLG and TXG as single blocks reach relatively low errors for some training/test set partitions even with respect to data fusion. In addition, the different pretreatments applied for mid-level data fusion seem not to have a notable impact on cross-validation and prediction performances. As already mentioned, the prediction

of anthocyanins from RGB images of grape samples is mainly a colour-related problem and CLG approach allows to account for this aspect. However, the calibration models seem to generally benefit, albeit modestly, from the inclusion of texture features.

With regard to the different partitioning in training and test set, generally lower prediction performances were achieved when considering samples belonging to Vin A 2021 and Vin B 2020 in the test set, but these aspects will be further detailed in the following section.

Table 2 *Results of the PLS models. The cells are coloured according to the corresponding RMSE values, from green for lowest RMSE values to red for highest RMSE values.*

		Vin A 2020		Vin B 2020		Vin A 2021		Vin B 2021		Vin C 2021		Original	
	Technique	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**	**RMSECV**	**RMSEP**
Single Blocks	**CLG**	164	122	160	186	135	195	151	120	174	130	144	158
	TXG D5	159	158	198	204	232	141	201	181	174	209	200	179
	TXG D10	129	222	178	181	191	131	212	159	167	198	175	142
	TXG D15	123	250	165	132	190	116	217	134	153	192	158	126
	GLCM D1	282	306	250	503	305	308	305	266	307	245	256	249
	GLCM D5	268	261	241	357	291	235	219	306	283	230	241	224
	GLCM D10	251	258	246	288	267	376	233	286	254	260	232	194
	GLCM D15	246	250	242	245	244	227	254	275	229	251	223	181
	SCTD	243	259	262	289	265	190	235	243	215	354	234	238
Low-Level Data Fusion	**CLG + TXG D5**	201	140	168	246	140	260	176	160	199	154	163	149
	CLG + TXG D10	192	124	156	218	140	232	187	144	182	136	154	154
	CLG + TXG D15	173	123	159	202	134	209	168	138	164	127	138	141
	CLG + GLCM D1	209	141	167	275	145	283	208	188	207	143	160	189
	CLG + GLCM D5	183	128	146	209	146	260	167	114	189	187	152	178
	CLG + GLCM D10	153	134	147	177	143	208	156	119	172	183	156	156
	CLG + GLCM D15	141	136	149	163	141	172	144	119	159	184	159	139
	CLG + SCTD	143	147	142	134	138	149	134	124	141	161	151	142
Mid-Level Data Fusion (autoscaling)	**CLG + TXG D5**	99	104	111	174	90	158	117	137	100	130	106	148
	CLG + TXG D10	119	145	95	182	80	158	116	127	118	126	99	136
	CLG + TXG D15	113	155	103	154	83	155	101	119	111	122	95	125
	CLG + GLCM D1	184	161	167	295	115	224	159	199	162	141	118	161
	CLG + GLCM D5	149	132	148	216	94	292	126	205	144	130	114	164
	CLG + GLCM D10	119	130	142	177	102	306	110	188	136	121	121	158
	CLG + GLCM D15	115	132	140	157	100	285	108	192	129	124	122	146
	CLG + SCTD	115	141	133	202	124	176	120	140	131	144	123	170
Mid-Level Data Fusion (mean centering & Block Scaling)	**CLG + TXG D5**	121	105	115	201	83	199	132	126	142	125	112	153
	CLG + TXG D10	133	113	103	165	84	229	124	133	124	123	104	158
	CLG + TXG D15	121	113	112	157	81	201	104	127	103	123	93	148
	CLG + GLCM D1	158	139	152	287	110	240	149	154	163	174	120	166
	CLG + GLCM D5	152	106	143	220	97	289	129	175	139	125	120	169
	CLG + GLCM D10	129	124	141	173	93	307	120	159	147	173	122	159
	CLG + GLCM D15	115	132	138	152	102	284	120	159	139	171	124	152
	CLG + SCTD	116	140	134	174	112	174	120	128	130	135	114	157

Table 3 *Results of the PLS models considering the Vin A 2020 partition.*

	Technique	N° LVs	RMSEC	RMSECV	RMSEP	R2Cal	R2CV	R2P
Single Blocks	**CLG**	3	123	164	122	0.91	0.83	0.83
	TXG D5	6	103	159	158	0.93	0.84	0.72
	TXG D10	5	100	129	222	0.94	0.90	0.44
	TXG D15	5	99	123	250	0.94	0.91	0.30
	GLCM D1	6	163	282	306	0.83	0.50	-0.06
	GLCM D5	6	155	268	261	0.85	0.55	0.23
	GLCM D10	3	184	251	258	0.79	0.60	0.25
	GLCM D15	3	173	246	250	0.81	0.62	0.30
	SCTD	4	151	243	259	0.86	0.63	0.24
Low-Level Data Fusion	**CLG + TXG D5**	2	149	201	140	0.86	0.75	0.78
	CLG + TXG D10	6	65	192	124	0.97	0.77	0.83
	CLG + TXG D15	6	68	173	123	0.97	0.81	0.83
	CLG + GLCM D1	3	128	209	141	0.90	0.73	0.77
	CLG + GLCM D5	3	121	183	128	0.91	0.79	0.82
	CLG + GLCM D10	3	112	153	134	0.92	0.85	0.80
	CLG + GLCM D15	3	107	141	136	0.93	0.87	0.79
	CLG + SCTD	2	117	143	147	0.91	0.87	0.76
Mid-Level Data Fusion (autoscaling)	**CLG + TXG D5**	2	81	99	104	0.96	0.94	0.88
	CLG + TXG D10	2	86	119	145	0.95	0.91	0.76
	CLG + TXG D15	2	83	113	155	0.96	0.92	0.73
	CLG + GLCM D1	2	118	184	161	0.91	0.79	0.71
	CLG + GLCM D5	2	107	149	132	0.93	0.86	0.80
	CLG + GLCM D10	3	106	119	130	0.93	0.91	0.81
	CLG + GLCM D15	3	102	115	132	0.93	0.92	0.80
	CLG + SCTD	3	103	115	141	0.93	0.92	0.77
Mid-Level Data Fusion (mean centering & Block Scaling)	**CLG + TXG D5**	6	87	121	105	0.95	0.91	0.87
	CLG + TXG D10	4	96	133	113	0.94	0.89	0.85
	CLG + TXG D15	4	89	121	113	0.95	0.91	0.86
	CLG + GLCM D1	1	134	158	139	0.89	0.84	0.78
	CLG + GLCM D5	2	108	152	106	0.93	0.85	0.87
	CLG + GLCM D10	2	107	129	124	0.93	0.90	0.83
	CLG + GLCM D15	4	102	115	132	0.94	0.92	0.80
	CLG + SCTD	3	103	116	140	0.93	0.92	0.78

Table 4 *Results of the PLS models considering the Vin B 2020 partition.*

	Technique	N° LVs	RMSEC	RMSECV	RMSEP	R2Cal	R2CV	R2P
Single Blocks	**CLG**	2	133	160	186	0.90	0.85	0.22
	TXG D5	5	125	198	204	0.91	0.77	0.07
	TXG D10	6	93	178	181	0.95	0.81	0.27
	TXG D15	6	94	165	132	0.95	0.84	0.61
	GLCM D1	2	199	250	503	0.76	0.63	-4.67
	GLCM D5	2	195	241	357	0.77	0.65	-1.86
	GLCM D10	2	195	246	288	0.77	0.64	-0.87
	GLCM D15	2	194	242	245	0.78	0.65	-0.34
	SCTD	3	178	262	289	0.81	0.59	-0.87
Low-Level Data Fusion	**CLG + TXG D5**	2	136	168	246	0.89	0.83	-0.35
	CLG + TXG D10	1	138	156	218	0.89	0.86	-0.06
	CLG + TXG D15	1	138	159	202	0.89	0.85	0.08
	CLG + GLCM D1	2	120	167	275	0.91	0.83	-0.69
	CLG + GLCM D5	2	118	146	209	0.92	0.87	0.02
	CLG + GLCM D10	2	118	147	177	0.92	0.87	0.30
	CLG + GLCM D15	2	118	149	163	0.92	0.87	0.40
	CLG + SCTD	5	75	142	134	0.97	0.88	0.60
Mid-Level Data Fusion (autoscaling)	**CLG + TXG D5**	2	94	111	174	0.95	0.93	0.32
	CLG + TXG D10	2	78	95	182	0.96	0.95	0.26
	CLG + TXG D15	2	85	103	154	0.96	0.94	0.47
	CLG + GLCM D1	2	132	167	295	0.90	0.83	-0.96
	CLG + GLCM D5	2	128	148	216	0.90	0.87	-0.05
	CLG + GLCM D10	2	127	142	177	0.90	0.88	0.30
	CLG + GLCM D15	2	127	140	157	0.90	0.88	0.45
	CLG + SCTD	2	117	133	202	0.92	0.89	0.08
Mid-Level Data Fusion (mean centering & Block Scaling)	**CLG + TXG D5**	6	90	115	201	0.95	0.92	0.09
	CLG + TXG D10	5	86	103	165	0.96	0.94	0.39
	CLG + TXG D15	5	89	112	157	0.95	0.92	0.45
	CLG + GLCM D1	1	134	152	287	0.89	0.86	-0.84
	CLG + GLCM D5	1	132	143	220	0.90	0.88	-0.08
	CLG + GLCM D10	1	132	141	173	0.90	0.88	0.33
	CLG + GLCM D15	1	130	138	152	0.90	0.89	0.48
	CLG + SCTD	2	117	134	174	0.92	0.89	0.32

Table 5 *Results of the PLS models considering the Vin A 2021 partition.*

	Technique	N° LVs	RMSEC	RMSECV	RMSEP	R2Cal	R2CV	R2P
Single Blocks	**CLG**	4	97	135	195	0.92	0.85	0.83
	TXG D5	6	105	232	141	0.91	0.57	0.91
	TXG D10	6	100	191	131	0.92	0.71	0.92
	TXG D15	5	104	190	116	0.91	0.71	0.94
	GLCM D1	3	207	305	308	0.66	0.26	0.56
	GLCM D5	3	203	291	235	0.67	0.33	0.75
	GLCM D10	1	230	267	376	0.58	0.43	0.35
	GLCM D15	3	183	244	227	0.73	0.53	0.76
	SCTD	1	228	265	190	0.59	0.44	0.83
Low-Level Data Fusion	**CLG + TXG D5**	6	68	140	260	0.96	0.84	0.69
	CLG + TXG D10	6	62	140	232	0.97	0.85	0.75
	CLG + TXG D15	6	58	134	209	0.97	0.86	0.80
	CLG + GLCM D1	3	111	145	283	0.90	0.83	0.63
	CLG + GLCM D5	4	95	146	260	0.93	0.83	0.69
	CLG + GLCM D10	4	94	143	208	0.93	0.84	0.80
	CLG + GLCM D15	4	92	141	172	0.93	0.84	0.86
	CLG + SCTD	5	67	138	149	0.96	0.85	0.90
Mid-Level Data Fusion (autoscaling)	**CLG + TXG D5**	3	73	90	158	0.96	0.94	0.88
	CLG + TXG D10	2	64	80	158	0.97	0.95	0.88
	CLG + TXG D15	3	70	83	155	0.96	0.94	0.89
	CLG + GLCM D1	4	86	115	224	0.94	0.89	0.77
	CLG + GLCM D5	6	82	94	292	0.95	0.93	0.61
	CLG + GLCM D10	3	87	102	306	0.94	0.92	0.57
	CLG + GLCM D15	6	90	100	285	0.94	0.92	0.63
	CLG + SCTD	3	95	124	176	0.93	0.88	0.86
Mid-Level Data Fusion (mean centering & Block Scaling)	**CLG + TXG D5**	6	69	83	199	0.96	0.95	0.82
	CLG + TXG D10	4	69	84	229	0.96	0.94	0.76
	CLG + TXG D15	3	72	81	201	0.96	0.95	0.81
	CLG + GLCM D1	5	88	110	240	0.94	0.90	0.74
	CLG + GLCM D5	6	83	97	289	0.95	0.92	0.62
	CLG + GLCM D10	4	87	93	307	0.94	0.93	0.57
	CLG + GLCM D15	6	90	102	284	0.94	0.92	0.63
	CLG + SCTD	4	96	112	174	0.93	0.90	0.86

Table 6 *Results of the PLS models considering the Vin B 2021 partition.*

	Technique	N° LVs	RMSEC	RMSECV	RMSEP	R2Cal	R2CV	R2P
Single Blocks	**CLG**	3	122	151	120	0.90	0.84	0.88
	TXG D5	6	118	201	181	0.90	0.72	0.73
	TXG D10	5	117	212	159	0.91	0.69	0.79
	TXG D15	5	107	217	134	0.92	0.67	0.85
	GLCM D1	5	200	305	266	0.72	0.36	0.42
	GLCM D5	5	173	219	306	0.79	0.67	0.23
	GLCM D10	6	149	233	286	0.85	0.63	0.32
	GLCM D15	5	158	254	275	0.83	0.56	0.38
	SCTD	3	161	235	243	0.82	0.62	0.51
Low-Level Data Fusion	**CLG + TXG D5**	5	86	176	160	0.95	0.79	0.79
	CLG + TXG D10	6	73	187	144	0.96	0.76	0.83
	CLG + TXG D15	6	70	168	138	0.97	0.80	0.84
	CLG + GLCM D1	4	100	208	188	0.93	0.70	0.71
	CLG + GLCM D5	2	142	167	114	0.86	0.81	0.89
	CLG + GLCM D10	3	107	156	119	0.92	0.83	0.88
	CLG + GLCM D15	3	102	144	119	0.93	0.86	0.88
	CLG + SCTD	5	69	134	124	0.97	0.88	0.87
Mid-Level Data Fusion (autoscaling)	**CLG + TXG D5**	2	93	117	137	0.94	0.91	0.85
	CLG + TXG D10	2	92	116	127	0.94	0.91	0.87
	CLG + TXG D15	2	83	101	119	0.95	0.93	0.88
	CLG + GLCM D1	2	110	159	199	0.92	0.83	0.67
	CLG + GLCM D5	3	96	126	205	0.94	0.89	0.65
	CLG + GLCM D10	3	87	110	188	0.95	0.92	0.71
	CLG + GLCM D15	3	89	108	192	0.95	0.92	0.70
	CLG + SCTD	3	102	120	140	0.93	0.90	0.84
Mid-Level Data Fusion (mean centering & Block Scaling)	**CLG + TXG D5**	4	96	132	126	0.94	0.88	0.87
	CLG + TXG D10	4	96	124	133	0.94	0.89	0.85
	CLG + TXG D15	4	90	104	127	0.94	0.93	0.87
	CLG + GLCM D1	2	109	149	154	0.92	0.85	0.80
	CLG + GLCM D5	3	96	129	175	0.94	0.88	0.75
	CLG + GLCM D10	3	95	120	159	0.94	0.90	0.79
	CLG + GLCM D15	3	95	120	159	0.94	0.90	0.79
	CLG + SCTD	3	102	120	128	0.93	0.90	0.87

***Table** 7 Results of the PLS models considering the Vin C 2021 partition.*

	Technique	N° LVs	RMSEC	RMSECV	RMSEP	R2Cal	R2CV	R2P
Single Blocks	**CLG**	3	116	174	130	0.90	0.77	0.91
	TXG D5	6	92	174	209	0.94	0.77	0.78
	TXG D10	6	87	167	198	0.94	0.79	0.80
	TXG D15	6	81	153	192	0.95	0.82	0.81
	GLCM D1	4	192	307	245	0.72	0.28	0.70
	GLCM D5	5	168	283	230	0.79	0.39	0.73
	GLCM D10	3	181	254	260	0.75	0.51	0.66
	GLCM D15	3	174	229	251	0.77	0.60	0.68
	SCTD	1	191	215	354	0.72	0.65	0.37
Low-Level Data Fusion	**CLG + TXG D5**	6	75	199	154	0.96	0.70	0.88
	CLG + TXG D10	6	69	182	136	0.96	0.75	0.91
	CLG + TXG D15	6	65	164	127	0.97	0.80	0.92
	CLG + GLCM D1	3	126	207	143	0.88	0.68	0.90
	CLG + GLCM D5	1	155	189	187	0.82	0.73	0.82
	CLG + GLCM D10	2	130	172	183	0.87	0.78	0.83
	CLG + GLCM D15	2	127	159	184	0.88	0.81	0.83
	CLG + SCTD	5	64	141	161	0.97	0.85	0.87
Mid-Level Data Fusion (autoscaling)	**CLG + TXG D5**	2	78	100	130	0.95	0.92	0.91
	CLG + TXG D10	2	77	118	126	0.96	0.89	0.92
	CLG + TXG D15	2	73	111	122	0.96	0.91	0.92
	CLG + GLCM D1	6	102	162	141	0.92	0.80	0.90
	CLG + GLCM D5	2	105	144	130	0.92	0.84	0.91
	CLG + GLCM D10	2	115	136	121	0.90	0.86	0.93
	CLG + GLCM D15	2	111	129	124	0.91	0.87	0.92
	CLG + SCTD	2	111	131	144	0.91	0.87	0.90
Mid-Level Data Fusion (mean centering & Block Scaling)	**CLG + TXG D5**	4	94	142	125	0.93	0.85	0.92
	CLG + TXG D10	4	85	124	123	0.95	0.88	0.92
	CLG + TXG D15	4	79	103	123	0.95	0.92	0.92
	CLG + GLCM D1	1	139	163	174	0.85	0.80	0.85
	CLG + GLCM D5	4	96	139	125	0.93	0.85	0.92
	CLG + GLCM D10	1	126	147	173	0.88	0.84	0.85
	CLG + GLCM D15	1	122	139	171	0.89	0.85	0.85
	CLG + SCTD	3	111	130	135	0.91	0.87	0.91

Table 8 *Results of the PLS models considering the Original partition.*

	Technique	N° LVs	RMSEC	RMSECV	RMSEP	R2Cal	R2CV	R2P
Single Blocks	**CLG**	4	100	144	158	0.93	0.87	0.81
	TXG D5	6	103	200	179	0.93	0.74	0.75
	TXG D10	5	112	175	142	0.92	0.80	0.84
	TXG D15	5	102	158	126	0.93	0.84	0.88
	GLCM D1	2	218	256	249	0.69	0.58	0.52
	GLCM D5	2	210	241	224	0.72	0.62	0.61
	GLCM D10	2	205	232	194	0.73	0.65	0.71
	GLCM D15	2	199	223	181	0.74	0.68	0.74
	SCTD	1	215	234	238	0.70	0.64	0.56
Low-Level Data Fusion	**CLG + TXG D5**	4	102	163	149	0.93	0.83	0.83
	CLG + TXG D10	6	68	154	154	0.97	0.85	0.82
	CLG + TXG D15	5	66	138	141	0.97	0.88	0.84
	CLG + GLCM D1	4	94	160	189	0.94	0.83	0.72
	CLG + GLCM D5	4	92	152	178	0.95	0.85	0.75
	CLG + GLCM D10	4	89	156	156	0.95	0.84	0.81
	CLG + GLCM D15	4	88	159	139	0.95	0.84	0.85
	CLG + SCTD	1	134	151	142	0.88	0.85	0.84
Mid-Level Data Fusion (autoscaling)	**CLG + TXG D5**	2	77	106	148	0.96	0.93	0.83
	CLG + TXG D10	2	81	99	136	0.96	0.94	0.86
	CLG + TXG D15	2	73	95	125	0.97	0.94	0.88
	CLG + GLCM D1	5	97	118	161	0.94	0.91	0.80
	CLG + GLCM D5	5	96	114	164	0.94	0.92	0.79
	CLG + GLCM D10	4	99	121	158	0.94	0.91	0.81
	CLG + GLCM D15	4	100	122	146	0.93	0.90	0.83
	CLG + SCTD	2	105	123	170	0.93	0.90	0.77
Mid-Level Data Fusion (mean centering & Block Scaling)	**CLG + TXG D5**	5	78	112	153	0.96	0.92	0.82
	CLG + TXG D10	3	88	104	158	0.95	0.93	0.81
	CLG + TXG D15	3	78	93	148	0.96	0.94	0.83
	CLG + GLCM D1	5	98	120	166	0.94	0.91	0.79
	CLG + GLCM D5	4	98	120	169	0.94	0.91	0.78
	CLG + GLCM D10	4	96	122	159	0.94	0.90	0.80
	CLG + GLCM D15	4	97	124	152	0.94	0.90	0.82
	CLG + SCTD	4	97	114	157	0.94	0.92	0.81

5.2 Statistical evaluation of the results

5.2.1 ANOVA

As described in Section 4.3, ANOVA was performed on the RMSEP values of all the models externally validated with only the samples collected in a specific vineyard and in a specific year, to have a more accurate estimate of the variability possibly due to the different vineyards in each vintage.

The RMSEP values obtained from the models calculated using the single blocks, i.e., considering the performances of the different feature extraction techniques individually, were evaluated using two-way ANOVA without interaction. The results, reported in Table 9, show that significant effects are ascribable only to the different feature extraction techniques. Based on this outcome, the least significant difference (LSD) multiple comparison test (P = 95%) was performed to compare each feature extraction method based on the RMSEP values across the 5 training/test set partitionings (Figure 7). The RMSEP values obtained considering both CLG and TXG (at all the tested distances)

were significantly lower with respect to those obtained by the PLS models calculated with GLCM and SCTD. As already outlined on Section 5.1, the best performances of both CLG and TXG suggest that in the problem under investigation these two feature extraction methods allow a better representation of image properties of interest for the problem at hand. Considering CLG and TXG, no statistically significant differences were found in the RMSEP values of the two feature extraction techniques. This result suggests that the TXG approach may represent a hybrid method that can account for both colour and texture features, since the signals are based on the spatial differences of pixel intensities in the three colour channels.

The RMSEP values of the PLS models obtained using data fusion were evaluated using three-way ANOVA with interaction, and the results are reported in Table 10. ANOVA highlighted a statistically significant difference between the different vineyards-years and between the different feature extraction techniques, and also the technique × vineyard-year interaction resulted statistically significant. No statistically significant effects were ascribable to the different data fusion approaches tested in this study. Figures 8 and 9 report the LSD intervals (P = 95%) for each combination of feature extraction techniques and for each vineyard-year partition, respectively.

According to Figure 8 the fusion of CLG with TXG calculated considering a distance of 15 pixels led to significantly lower RMSEP values with respect to the models calculated by fusing CLG with GLCM, independently of the pixel distance considered and of the data fusion approach. Interestingly, the RMSEP values of the models calculated by combining CLG and SCTD are not significantly different from those of the models obtained from the fusion of CLG and TXG. Considering that SCTD alone provided significantly higher RMSEP values compared to CLG and TXG as single blocks, we can assume that the improved performances obtained after fusing CLG and SCTD are mainly ascribable to the contribution of CLG block.

Table 9 *Results of two-way ANOVA on the RMSEP of the models calculated using the single blocks.*

Source	Sum of squares	Degrees of freedom	Mean square	F	p- value*
Technique	165461.5	8	20682.68	7.323883	**1.7E-05**
Vineyard-year	14496.91	4	3624.228	1.283365	0.297108
Error	90368.15	32	2824.005		
Total	270326.5	44			

** statistically significant differences (p-value <0.05) are highlighted in bold.*

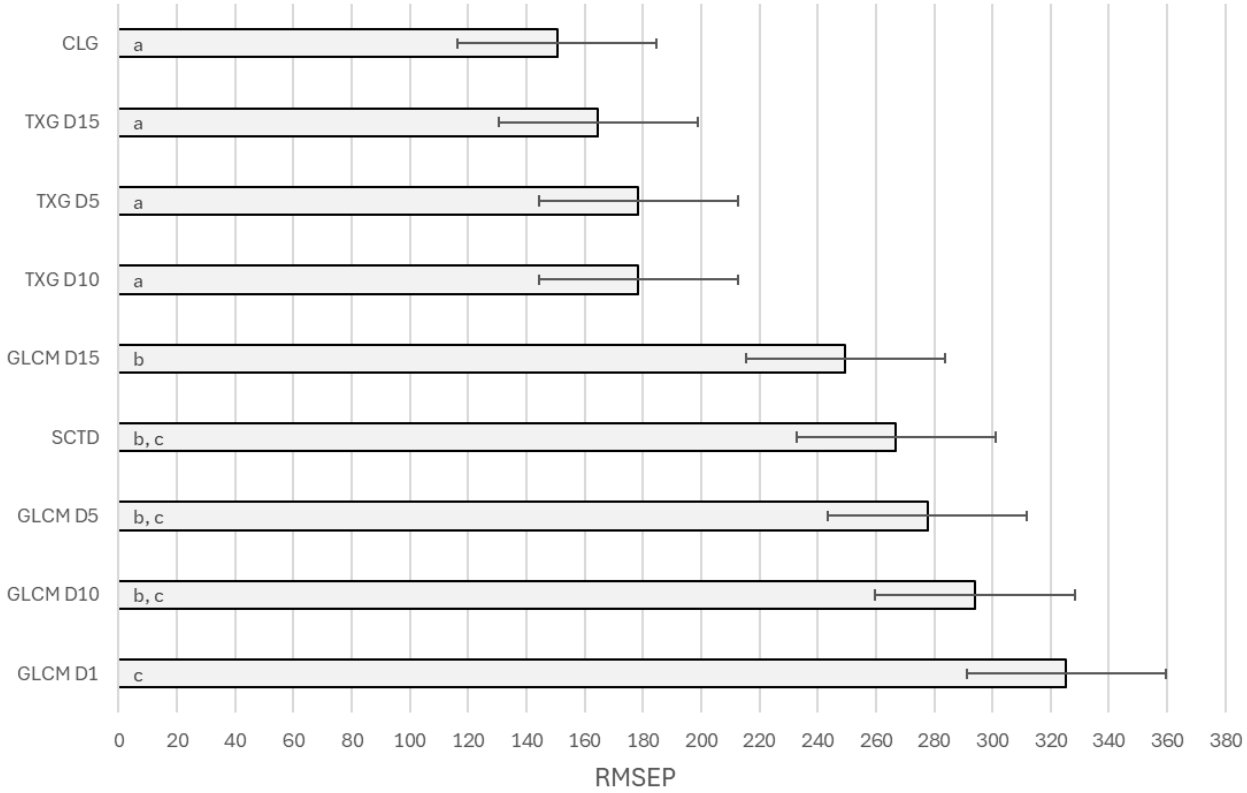


Figure 7 *Means and LSD intervals (P = 95%) of RMSEP values of the feature extraction techniques resulting from the ANOVA model reported in Table 9. Statistically significant differences are highlighted with different letters.*

Table 10 *Results of three-way ANOVA with interaction performed on the RMSEP of the models obtained using data fusion approaches.*

Source	Sum of squares	Degrees of freedom	Mean square	F	p- value*
Technique	37125.3	7	5303.6	8.63	**3.8829E-07**
Data fusion	192.3	2	96.1	0.16	0.8556
Vineyard-year	157502.9	4	39375.7	64.04	**3.0673E-20**
Technique × Data fusion	14698.4	14	1049.9	1.71	0.0800
Technique × Vineyard-year	47289.2	28	1688.9	2.75	**6.4497E-04**
Data fusion × Vineyard-year	9808.9	8	1226.1	1.99	0.0639
Error	34433.1	56	614.9		
Total	301050.1	119			

**statistically significant differences (p-value <0.05) are highlighted in bold.*

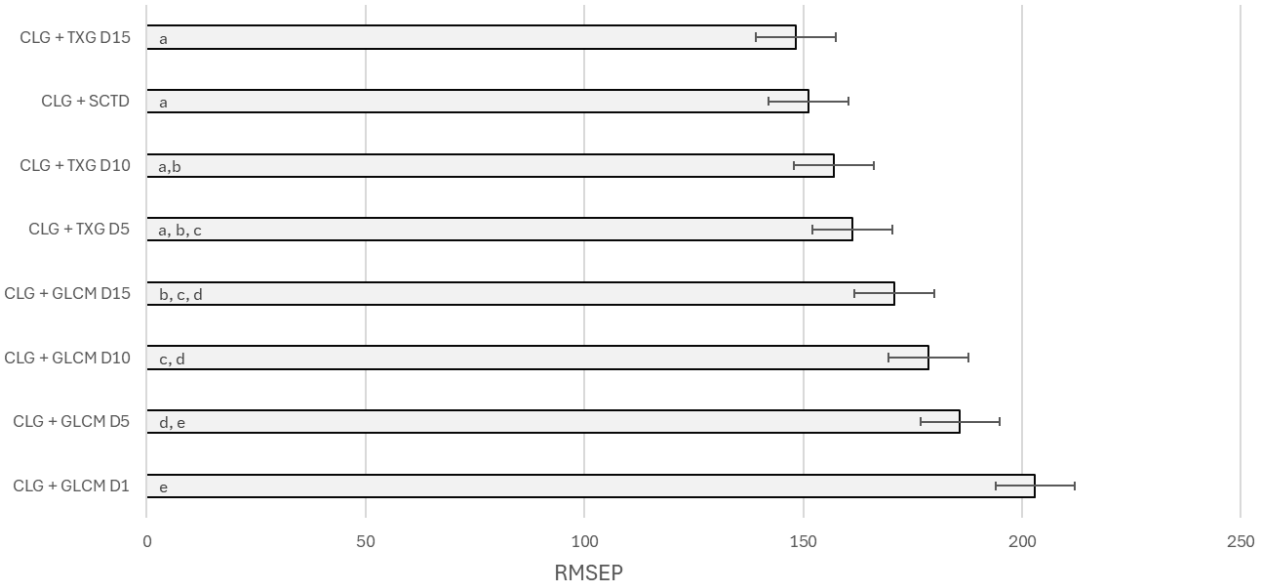


Figure 8 *Means and LSD intervals (P = 95%) of RMSEP values of the different combinations of feature extraction techniques resulting from the ANOVA model reported in Table 10. Statistically significant differences are highlighted with different letters.*

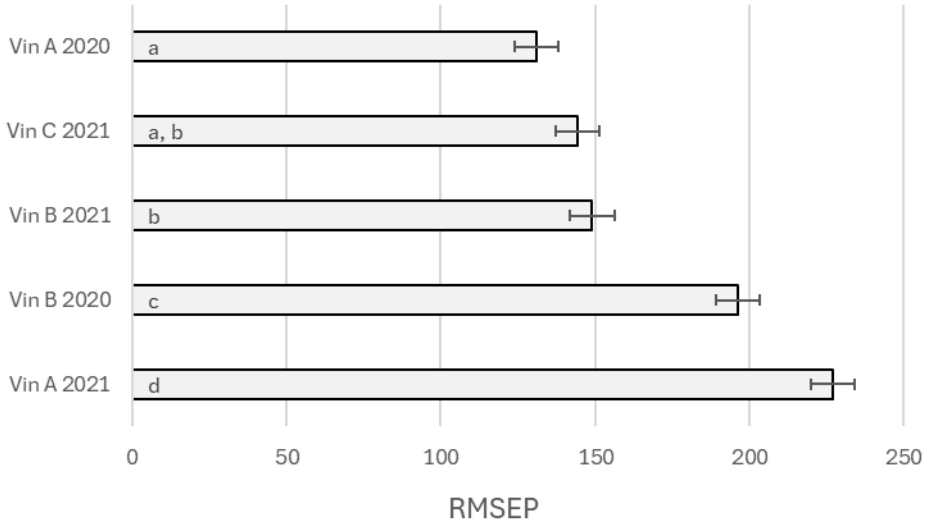


Figure 9 *Means and LSD intervals (P = 95%) of the RMSEP values of the different vineyard-year partitioning resulting from the ANOVA model reported in Table 10. Statistically significant differences are highlighted with different letters.*

Even if the vineyard-year effect resulted not significant in the ANOVA based on the results of single blocks, this effect was significant when considering the ANOVA of data fusion results. As shown in Figure 9, significantly higher RMSEP values are obtained when considering Vin A 2021 and Vin B 2020 as external test sets. In accordance with the results reported in Table 2, considering the different training/test partitioning based on vineyard and vintage year, the lowest prediction performances were obtained for Vin A 2021, especially when considering the data fusion approaches. However, the application of the texturegrams as single technique at all the pixel distances considered seem to reach relatively low RMSEP values when the samples of Vin A 2021 are in the external test set.

The lower performances observed for Vin A 2021 and Vin B 2020 with respect to the other training/test set partitions suggest that a larger dataset, including samples from more vineyards and more vintages, would be required to better cover geographical and time related variability sources.

On the other hand, Vin C 2021 partition was initially supposed to be the most critical case, since the external validation is performed on samples belonging to a vineyard not considered in the training set. However, the outcomes show a general good predictive capability for Vin C 2021 samples, which is not significantly different from Vin A 2020 and Vin B 2021 partitions.

5.2.2 Exploratory data analysis of the results

To have an overall comparison of the different model performances and a general overview of the contribution coming from feature extraction techniques, data fusion approaches and training/test set partitions to the characterisation of the models under investigation, an exploratory analysis was performed by means of PCA on the results achieved with both the single techniques and their combinations through data fusion. i.e., on the data reported in Table 2. The dataset used for PCA was therefore composed by a total of 33 sets of descriptors (9 single techniques + 8 technique combinations × 3 data fusion approaches) described by the corresponding RMSECV and RMSEP values calculated according to each training/test set partition tested, i.e. by a total of 12 variables. Figure 10a reports the score plot of the first two PCs of the PCA model, accounting for 87.14% of total data variance. In Figure 10a, the different objects (corresponding to the sets of descriptors) are coloured according to the image feature extraction technique (both as single block and fused with CLG), while their marker shape refers to the data block approach (i.e., single block, low-level data fusion, mid-level data fusion of autoscaled score vectors, mid-level data fusion of mean centered and block scaled score vectors).

Figure 10b reports the corresponding PC1 vs. PC2 loading plot, where the variables are coloured according to the cross validation (RMSECV) or prediction error (RMSEP) and labelled according to the respective training/test partition.

It is possible to observe that PC1 mainly describes the general error associated with each technique throughout all the training/test set partitions, since all the elements of PC1 loading vector are positive and with similar values. Therefore, along PC1 lower errors are associated to negative scores, and vice versa.

PC2 instead mainly reflects the differences between the RMSECV and the corresponding RMSEP values obtained for the various PLS models and is particularly influenced by the high RMSEP values generally observed for Vin A 2021 and Vin B 2020.

The models obtained with single blocks are more spread when compared with the ones obtained after data fusion. GLCM and SCTD models as single blocks are located at the highest positive score values of PC1 and distant from the remainder models. CLG on its own is close to the low- and mid-level data fusion approaches. According to the results reported in Table 2, CLG as single technique generally appears to have similar performance to low-level data fusion especially in cross-validation, but also to mid-level data fusion in prediction. At the same time, as shown in Figure 10a, also TXG approach shows interesting performances when considered individually, at all the pixel distances explored in this study. In accordance with this, Figure 7 and Table 2 give evidence of the performances

obtained when TXG are considered as single technique, which are comparable to those obtained with CLG, and for some training/test set partitions TXG perform even better than the colourgrams. For example, when Vin A 2021 is in the external test set, TXG provided lower RMSEP values when compared with the colourgrams, independently of the pixel distance considered. Indeed, the models related to TXG calculated at a distance of 5, 10 and 15 pixels are located at the lowest score values on PC2 which, according to Figure 10b, accounts for the difference between RMSEP and the corresponding RMSECV value in particular for Vin A 2021 and, at a minor extent, Vin B 2020 training/test set partitions. According to the score plot in Figure 10a, the highest differences between RMSEP and RMSECV values to predict these vineyards are associated to the models calculated by fusing at a mid-level the colourgrams with the GLCM at the different distances, by considering the GLCM D1 and GLCM D10 as single blocks and by fusing at a low-level the colourgrams with GLCM D1 and with TXG D5.

The overall higher errors were obtained when considering GLCM D1 alone; in fact, the models related to this single block are located at the extreme positive value of PC1, corresponding to generally high values of both RMSECV and RMSEP, and at high values of PC2, indicating RMSEP values much higher than the (already high) RMSECV ones. This is consistent with the ANOVA results reported in Figures 7 and 8, were GLCM D1 always shows the worst performance.

As shown in Figure 10a and in accordance with Table 2, the different preprocessing methods tested for mid-level data fusion do not appear to have a substantial impact on the model performances.

Concerning both GLCM and TXG, a trend along PC1 is observed in Figure 10a, where increasing pixel distance leads to lower PC1 values, i.e., to generally lower RMSECV and RMSEP values.

With reference to the scores on PC1, Figure 10a shows that the objects related to mid-level data fusion models are generally characterised by slightly lower PC1 score values, when compared to the corresponding models calculated with low-level data fusion. This suggests as a general trend that mid-level data fusion tends to provide overall best performances, as one might expect. In fact, mid-level data fusion implies a sort of “cleaning” of the useless information from the data. Conversely, low-level data fusion is a more conservative data fusion approach on one side, but it is less effective in selecting only the relevant information for the problem at hand. Comparing these results with those of three-way ANOVA on the RMSEP values reported in Table 10, where no significant differences were observed among the different data fusion approaches, the results of PCA suggest that this trend on PC1 on the one hand is due to the generally better performances in cross-validation of mid-level data fusion models, and on the other hand to the different performances of the considered feature extraction techniques.

With regard to the technique, the two models obtained by mid-level fusion of CLG and TXG D15 show the lowest PC1 scores, suggesting an interesting, even if slight, improvement of the model performances when combining the information brought by the colourgrams and by the texturegrams.

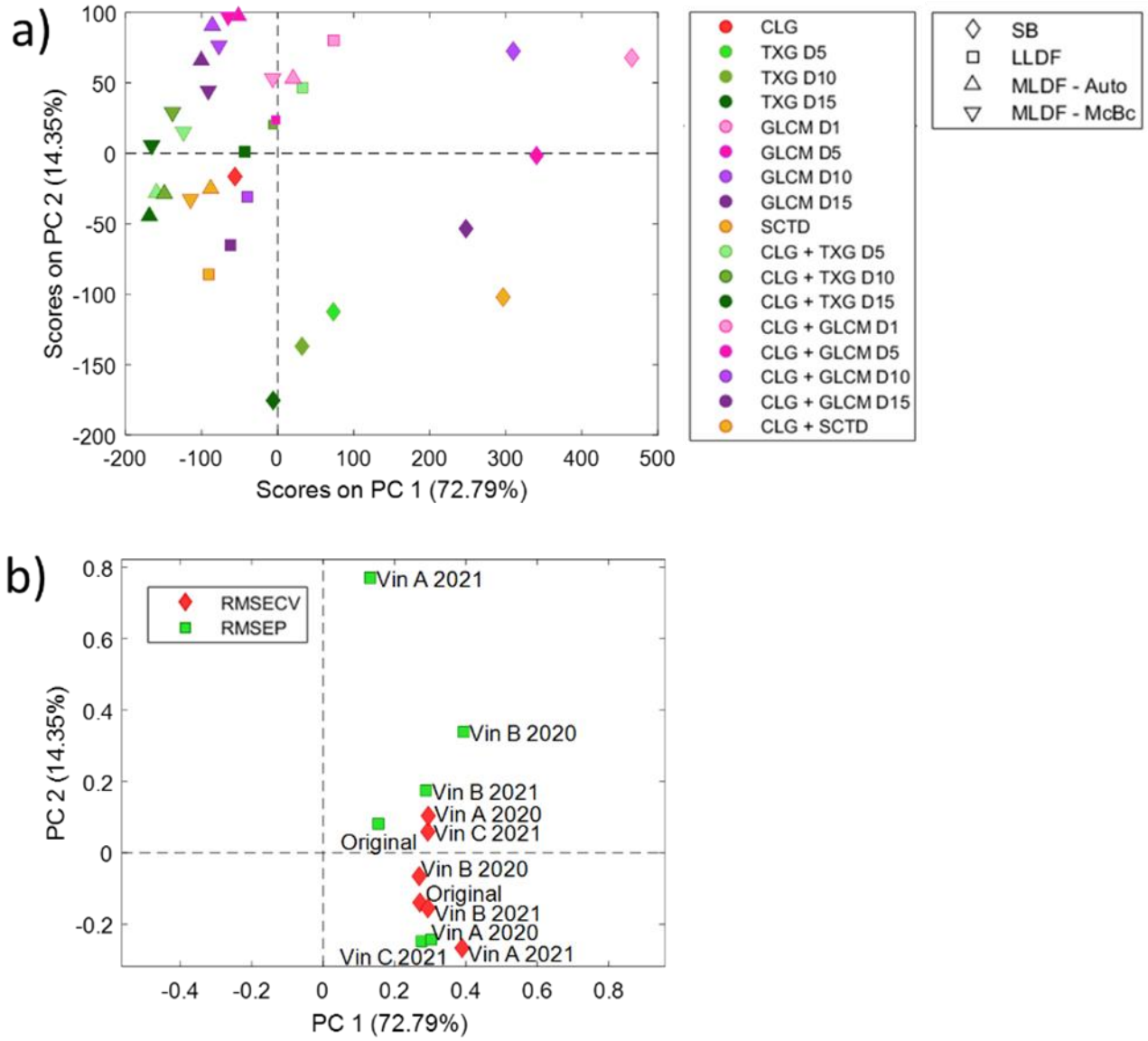


Figure 10 *Results of the PCA model calculated on the data reported in Table 2. In a) PC1 vs. PC2 score plot, where the different models are coloured according to feature extraction technique, while the marker shape represents single block modelling (diamond), low-level data fusion (square), mid-level data fusion with autoscaling (upward-pointing triangle) and mid-level data fusion with mean centering and block scaling (downward-pointing triangle); in b) PC1 and PC2 loading plot where the variables are coloured according to the cross-validation (red) or test set prediction (green) error, while labels refer to the type of training/test set partition.*

6 Conclusions

In the present study, we discussed the application of different image-level techniques able to extract and analyse colour and texture features of a benchmark dataset of RGB images of red grape samples to build PLS models and relate the image features with anthocyanins content. As a colour-related technique we considered the colourgrams approach, while for the extraction of texture features, we used literature methods such as SCTD and GLCM. In addition, we also proposed an alternative method to extract and codify the texture features of an RGB image into a one-dimensional signal, named texturegram.

The predictive abilities of the different colour and texture techniques were compared. Colourgrams and texturegrams allowed to achieve statistically better prediction results compared to GLCM and SCTD, probably due to the more detailed representation of the image properties allowed by these two types of signals and required for the issue at hand.

To evaluate the possible benefits of integrating colour and texture information, we also performed data fusion by combining colourgrams with each one of the considered texture methods. As a result, we obtained general slight improvement in model performances. Overall encouraging results were achieved by fusing at the mid-level the colourgrams with the texturegrams, suggesting an interesting, even if slight, improvement of the model performances when combining the information brought by the colourgrams and the texturegrams in a unique solution.

However, considering practical applications, both CLG and TXG as single blocks could represent a good compromise between the predictive capability and computational effort required for the issue explored in this work. In fact, the performances obtained with both approaches as single blocks appear to be comparable to those reached with data fusion, with the advantage of requiring the computation of only one type of feature.

As a matter of fact, the anthocyanins content in grape berries has a notable influence on colour properties rather than on texture ones. The interest in texture features in this work could represent a supplementary improvement of the model performances, by potentially characterizing the images with regard to elements like the background or the presence of berries pedicels.

Based on these considerations, for a more balanced evaluation of the potential of the different texture features and their possible combination with the colour-related ones, future studies are required to be carried out on a more systematically designed dataset. However, even if we conceived and proposed the texturegrams to be principally focused on texture properties of an image, the outcomes of this study suggest the potential use of this technique as a hybrid tool to simultaneously explore both colour

and texture properties. This outcome can be motivated considering that these signals rely on the spatial differences of pixel intensities in the three colour channels.

Acknowledgements

The authors wish also to thank Dr. Giovanni Nigro, Dr. Paola Tessarin and Dr. Domenico Bossio for technical assistance and support during grape harvesting and image acquisition.

CRediT Author Statement

Camilla Menozzi: Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing

José Manuel Prats-Montalbán: Conceptualization, Methodology, Software, Supervision, Writing - Review & Editing

Rosalba Calvini: Software, Conceptualization, Methodology, Writing - Review & Editing

Alessandro Ulrici: Conceptualization, Methodology, Supervision, Funding acquisition, Project administration, Writing - Review & Editing

7 References

Aggarwal, N., K. Agrawal, R., 2012. First and Second Order Statistics Features for Classification of Magnetic Resonance Brain Images. JSIP 03, 146–153. https://doi.org/10.4236/jsip.2012.32019

Antonelli, A., Cocchi, M., Fava, P., Foca, G., Franchini, G.C., Manzini, D., Ulrici, A., 2004. Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm. Analytica Chimica Acta 515, 3–13. https://doi.org/10.1016/j.aca.2004.01.005

Bevk, M., Kononenko, I., 2002. A statistical approach to texture description of medical images: a preliminary study, in: Proceedings of 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002). Presented at the 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002), IEEE Computer. Soc, Maribor, Slovenia, pp. 239–244. https://doi.org/10.1109/CBMS.2002.1011383

Bharati, M.H., Liu, J.J., MacGregor, J.F., 2004. Image texture analysis: methods and comparisons. Chemometrics and Intelligent Laboratory Systems 72, 57–71. https://doi.org/10.1016/j.chemolab.2004.02.005

Bharati, M.H., MacGregor, J.F., 2001. Texture analysis of images using principal component analysis, in: McCann, H., Scott, D.M. (Eds.), . Presented at the Intelligent Systems and Smart Manufacturing, Boston, MA, p. 27. https://doi.org/10.1117/12.417179

Biancolillo, A., Ferretti, R., Scappaticci, C., Foschi, M., D'Archivio, A.A., Di Santo, M., Di Martino, L., 2023. Development of a Non-Destructive Tool Based on E-Eye and Agro-Morphological Descriptors for the Characterization and Classification of Different Brassicaceae Landraces. Applied Sciences 13, 6591. https://doi.org/10.3390/app13116591

Borin, A., Ferrão, M.F., Mello, C., Cordi, L., Pataca, L.C.M., Durán, N., Poppi, R.J., 2007. Quantification of Lactobacillus in fermented milk by multivariate image analysis with least-squares support-vector machines. Anal Bioanal Chem 387, 1105–1112. https://doi.org/10.1007/s00216-006-0971-7

Calvini, R., Foca, G., Ulrici, A., 2016a. Data dimensionality reduction and data fusion for fast characterization of green coffee samples using hyperspectral sensors. Anal Bioanal Chem 408, 7351–7366. https://doi.org/10.1007/s00216-016-9713-7

Calvini, R., Foca, G., Ulrici, A., 2016b. Data dimensionality reduction and data fusion for fast characterization of green coffee samples using hyperspectral sensors. Anal Bioanal Chem 408, 7351–7366. https://doi.org/10.1007/s00216-016-9713-7

Calvini, R., Orlandi, G., Foca, G., Ulrici, A., 2020. Colourgrams GUI: A graphical user-friendly interface for the analysis of large datasets of RGB images. Chemometrics and Intelligent Laboratory Systems 196, 103915. https://doi.org/10.1016/j.chemolab.2019.103915

Calvini, R., Pigani, L., 2022. Toward the Development of Combined Artificial Sensing Systems for Food Quality Evaluation: A Review on the Application of Data Fusion of Electronic Noses, Electronic Tongues and Electronic Eyes. Sensors 22, 577. https://doi.org/10.3390/s22020577

Chandini, A.A., Maheswari B., U., 2018. Improved Quality Detection Technique for Fruits Using GLCM and MultiClass SVM, in: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI). Presented at the 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, Bangalore, pp. 150–155. https://doi.org/10.1109/ICACCI.2018.8554876

Clausi, D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. Canadian Journal of Remote Sensing 28, 45–62. https://doi.org/10.5589/m02-004

Cubero, S., Diago, M.P., Blasco, J., Tardaguila, J., Prats-Montalbán, J.M., Ibáñez, J., Tello, J., Aleixos, N., 2015. A new method for assessment of bunch compactness using automated image analysis: Bunch compactness assessment using image analysis. Australian Journal of Grape and Wine Research 21, 101–109. https://doi.org/10.1111/ajgw.12118

Giraudo, A., Calvini, R., Orlandi, G., Ulrici, A., Geobaldo, F., Savorani, F., 2018. Development of an automated method for the identification of defective hazelnuts based on RGB image analysis and colourgrams. Food Control 94, 233–240. https://doi.org/10.1016/j.foodcont.2018.07.018

Glories, Y., 1993. Maturité phénolique du raisin, conséquences technologiques: application aux millésimes 1991 et 1992, in: Journée Technique Du CIVB: Actes Du Colloque. pp. 56–61.

Gonçalves Dias Diniz, P.H., 2020. Chemometrics-assisted color histogram-based analytical systems. Journal of Chemometrics 34. https://doi.org/10.1002/cem.3242

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. IEEE Trans. Syst., Man, Cybern. SMC-3, 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Kononenko, I., Kukar, M., 2007. Data Preprocessing, in: Machine Learning and Data Mining. Elsevier, pp. 181–211. https://doi.org/10.1533/9780857099440.181

Lopes, J.F., Barbon, A.P.A.C., Orlandi, G., Calvini, R., Lo Fiego, D.P., Ulrici, A., Barbon, S., 2020. Dual Stage Image Analysis for a complex pattern classification task: Ham veining defect detection. Biosystems Engineering 191, 129–144. https://doi.org/10.1016/j.biosystemseng.2020.01.008

López, F., 2008. Multivariate statistical projection methods to perform robust feature extraction and classification in surface grading. J. Electron. Imaging 17, 031106. https://doi.org/10.1117/1.2957886

López, F., Valiente, J.-M., Prats, J.-M., 2005. Surface Grading Using Soft Colour-Texture Descriptors, in: Sanfeliu, A., Cortés, M.L. (Eds.), Progress in Pattern Recognition, Image Analysis and Applications, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 13–23. https://doi.org/10.1007/11578079_2

Malegori, C., Franzetti, L., Guidetti, R., Casiraghi, E., Rossi, R., 2016. GLCM, an image analysis technique for early detection of biofilm. Journal of Food Engineering 185, 48–55. https://doi.org/10.1016/j.jfoodeng.2016.04.001

Meenu, M., Kurade, C., Neelapu, B.C., Kalra, S., Ramaswamy, H.S., Yu, Y., 2021. A concise review on food quality assessment using digital image processing. Trends in Food Science & Technology 118, 106–124. https://doi.org/10.1016/j.tifs.2021.09.014

Menozzi, C., Calvini, R., Nigro, G., Tessarin, P., Bossio, D., Calderisi, M., Ferrari, V., Foca, G., Ulrici, A., 2023. Design and application of a smartphone-based device for in vineyard determination of anthocyanins content in red grapes. Microchemical Journal 191, 108811. https://doi.org/10.1016/j.microc.2023.108811

Orlandi, G., Calvini, R., Foca, G., Pigani, L., Vasile Simone, G., Ulrici, A., 2019a. Data fusion of electronic eye and electronic tongue signals to monitor grape ripening. Talanta 195, 181–189. https://doi.org/10.1016/j.talanta.2018.11.046

Orlandi, G., Calvini, R., Foca, G., Pigani, L., Vasile Simone, G., Ulrici, A., 2019b. Data fusion of electronic eye and electronic tongue signals to monitor grape ripening. Talanta 195, 181–189. https://doi.org/10.1016/j.talanta.2018.11.046

Orlandi, G., Calvini, R., Pigani, L., Foca, G., Vasile Simone, G., Antonelli, A., Ulrici, A., 2018. Electronic eye for the prediction of parameters related to grape ripening. Talanta 186, 381–388. https://doi.org/10.1016/j.talanta.2018.04.076

Paliwal, J., Visen, N.S., Jayas, D.S., White, N.D.G., 2003. Cereal Grain and Dockage Identification using Machine Vision. Biosystems Engineering 85, 51–57. https://doi.org/10.1016/S1537-5110(03)00034-5

Prats-Montalbán, J.M., De Juan, A., Ferrer, A., 2011. Multivariate image analysis: A review with applications. Chemometrics and Intelligent Laboratory Systems 107, 1–23. https://doi.org/10.1016/j.chemolab.2011.03.002

Prats-Montalbán, J.M., Ferrer, A., 2007. Integration of colour and textural information in multivariate image analysis: defect detection and classification issues. J. Chemometrics 21, 10–23. https://doi.org/10.1002/cem.1026

Prats-Montalbán, J.M., Ferrer, A., Bro, R., Hancewicz, T., 2009. Prediction of skin quality properties by different Multivariate Image Analysis methodologies. Chemometrics and Intelligent Laboratory Systems 96, 6–13. https://doi.org/10.1016/j.chemolab.2008.10.012

Ribéreau-Gayon, P., 1965. Le dosage des anthocyanes dans le vins rouge. Bull. Soc. Chim. 9, 2649.

Río Segade, S., Soto Vázquez, E., Díaz Losada, E., 2008. Influence of ripeness grade on accumulation and extractability of grape skin anthocyanins in different cultivars. Journal of Food Composition and Analysis 21, 599–607. https://doi.org/10.1016/j.jfca.2008.04.006

Rolando, P.L., Calvini, R., Foca, G., Ulrici, A., 2023. Mixture design and multivariate image analysis to monitor the colour of strawberry yoghurt purée. Microchemical Journal 194, 109222. https://doi.org/10.1016/j.microc.2023.109222

Saint Cricq de Gaulejac, N., Vivas, N., Glories, Y., 1998. Maturation phénolique des raisines rouges. Relation avec la qualité des vins. Comparison des cépages Merlot et Tempranillo. Progrès Agricole et Viticole 115, 306–318.

Sommer, S., Cohen, S., 2018. Comparison of Different Extraction Methods to Predict Anthocyanin Concentration and Color Characteristics of Red Wines. Fermentation 4, 39. https://doi.org/10.3390/fermentation4020039

Srinivasan, G.N., Shobha, G., 2008. Statistical Texture Analysis. Proceedings of world academy of science, engineering and technology 36, 1264–1269.

Tomita, F., Tsuji, S., 1990. Statistical Texture Analysis, in: Computer Analysis of Visual Textures. Springer US, Boston, MA, pp. 13–36. https://doi.org/10.1007/978-1-4613-1553-7_2

Wu, D., Sun, D.-W., 2013. Colour measurements by computer vision for food quality control – A review. Trends in Food Science & Technology 29, 5–20. https://doi.org/10.1016/j.tifs.2012.08.004

Zheng, C., Sun, D.-W., Zheng, L., 2006. Recent developments and applications of image features for food quality evaluation and inspection – a review. Trends in Food Science & Technology 17, 642–655. https://doi.org/10.1016/j.tifs.2006.06.005

5.3 Comparison of different spectral ranges to monitor alcoholic and acetic fermentation of red grape must using FT-NIR spectroscopy and PLS regression

What follows is the integral content of: Menozzi, C., Foca, G., Calvini, R., Catellani, L., Bezzecchi, A., & Ulrici, A. (2024). Comparison of Different Spectral Ranges to Monitor Alcoholic and Acetic Fermentation of Red Grape Must Using FT-NIR Spectroscopy and PLS Regression. *Food Analytical Methods*, 1-12. doi: 10.1007/s12161-024-02636-3

Comparison of different spectral ranges to monitor alcoholic and acetic fermentation of red grape must using FT-NIR spectroscopy and PLS regression

Camilla Menozzi[1], Giorgia Foca[1,2], Rosalba Calvini[1,2], Lisa Catellani[3], Andrea Bezzecchi[3], Alessandro Ulrici[1,2]

[1] *Dipartimento di Scienze della Vita, Università di Modena e Reggio Emilia, Padiglione Besta, Via Amendola, 2 – 42122 Reggio Emilia, Italy*

[2] *Centro Interdipartimentale BIOGEST-SITEIA, Università degli Studi di Modena e Reggio Emilia, Piazzale Europa, 1 – 42122 Reggio Emilia, Italy*

[3] *Acetaia San Giacomo, Strada Pennella, 1- 42017 Novellara (RE), Italy*

Abstract

Wine vinegar is produced through a two-phase fermentation of grape must: initially, yeast converts grape sugars into ethanol, and subsequently acetobacteria oxidize ethanol into acetic acid. This process, spanning weeks when conducted by surface fermentation, requires constant monitoring of ethanol and total acidity levels. To enhance the quality and efficiency of process monitoring, vinegar production is shifting to faster, environmentally sustainable methods. Near-Infrared (NIR) spectroscopy, recognized for its non-invasiveness and speed, is ideal for online implementation in process control. This study tracked dual fermentation in red grape must over an extended period, monitoring two different batches simultaneously to assess fermentation kinetics and reproducibility. Ethanol content and total acidity were analysed in fermenting musts throughout the whole fermentation process using both classical laboratory analyses and FT-NIR spectroscopy. Principal Component Analysis (PCA) was used to explore the spectral dataset, then Partial Least Squares (PLS) was used to develop calibration models for predicting ethanol and acidity. The models calculated considering the entire spectral range were compared with those obtained for two narrower zones, where more cost-effective and easily miniaturizable sensors are available on the market. FT-NIR

allowed to effectively determine ethanol content and acidity ($R^2_{Pred} > 0.98$), both over the entire range (12500-4000 cm^{-1}, corresponding to 800-2500 nm) and in the 10526-6060 cm^{-1} (950-1650 nm) region. Although less satisfactory, still acceptable results were obtained in the 12500-9346 cm^{-1} (800-1070 nm) region ($R^2_{Pred} > 0.81$), confirming the potential for cost-effective devices in real-time fermentation monitoring.

Keywords: Near-infrared spectroscopy; Red grape must; Fermentation; Process monitoring; Multivariate calibration.

1 Introduction

Wine vinegar is a food product derived from the dual fermentation of grape must. This process involves two primary phases: initially, yeast transforms the sugars in grapes into ethanol (Cozzolino, 2016), and subsequently, acetobacteria bio-oxidize ethanol produced in the first phase into acetic acid (Bhat et al., 2014). Achieving full completion of this process may span several weeks when the process is conducted through surface fermentation, requiring periodic monitoring to assess fermentation progression and determine the quantities of its byproducts, such as ethanol content and total acidity.

The increasing need for enhanced product quality and efficiency in vinegar production has induced the gradual shift from conventional, destructive analytical methods to faster, more environmentally sustainable alternatives. Vibrational spectroscopies emerge as a compelling choice to continuously monitor the fermentation stage of grape must, ensuring process control and enabling prompt intervention in case of anomalies during fermentation (Gishen et al., 2005). These techniques offer multiple advantages, being rapid and non-invasive, as they necessitate no prior sample preparation.

Among various vibrational spectroscopies, Near-Infrared (NIR) spectroscopy stands out for its remarkable versatility. It produces reliable results even when utilized by non-specialized personnel to analyse samples containing bonds such as C−H, N−H and O−H, with the condition that the analyte concentration is at least 0.1% of the total (Burns and Ciurczak, 2007). NIR spectroscopy is exceptionally fast, making it suitable for online implementation in process control, contributing to the ease of automating the control system. On the flip side, the application of this technique initially entails a price to pay, namely the preliminary work on signals to "train" the system to recognize the relevant features in the sample. NIR spectra serve as a sort of fingerprints of the sample that need to

be appropriately "interpreted"; this is done using chemometric methods for spectral analysis, which focus on eliminating irrelevant information and interferences and extracting information useful for solving the problem of interest (Barbon Junior et al., 2020; Burns and Ciurczak, 2007; Ferrari et al., 2011; Foca et al., 2016, 2011).

In the literature, numerous works have investigated the processes of alcoholic and acetic fermentation, employing NIR spectroscopy to quantify various analytes indicative of the fermentation stage. Table 1 provides a summary of these articles, highlighting their characteristics. Firstly, some of them specifically focus on the alcoholic fermentation phase (Buratti et al., 2011; Cozzolino et al., 2006; Di Egidio et al., 2010; Frohman and Mira De Orduña Heidinger, 2018; Jiménez-Márquez et al., 2016; Kasemsumran et al., 2022; Ouyang et al., 2016; Peng et al., 2016; Wu et al., 2015; Ye et al., 2014), while others concentrate on acetic fermentation (Chen et al., 2023; González-Sáiz et al., 2008; Phanomsophon et al., 2019; Sedjoah et al., 2021; Yano et al., 1997), depending on whether the final product of interest is wine or vinegar, respectively. The articles also differ based on the type of substrate involved in fermentation. Indeed, wine and vinegar can be produced not only from grapes but also from rice, apples, or other cereals or fruits – essentially from any vegetable matrix rich in sugars (Jagtap and Bapat, 2015; Solieri and Giudici, 2009). Additionally, different articles often consider various spectral ranges for NIR analysis, which can have important implications for the performance of the models.

Table 1 *Overview of literature articles addressing the monitoring of alcoholic and acetic fermentation through FT-NIR spectroscopy.*

Reference	Food matrix	Investigated process	Analytes of interest	Spectral range	Sampling time
(Buratti et al., 2011)	Red wine	Alcoholic fermentation	Glucose, fructose, ethanol and glycerol	12500-3600 cm^{-1}	Up to 20 days
(Cozzolino et al., 2006)	Red wine	Alcoholic fermentation	None (just spectra time-related changes)	400-2500 nm	Up to 90 days
(Di Egidio et al., 2010)	Red wine	Alcoholic fermentation	Glucose, fructose, ethanol, total phenolics, total anthocyanins and total flavonoids	12500-3600 cm^{-1}	Up to 35 days
(Frohman and Mira De Orduña Heidinger, 2018)	Red wine must	Alcoholic fermentation	Glucose, fructose, total sugars and ethanol	12500-4000 cm^{-1}	Not specified (fed-batch fermentation)
(Jiménez-Márquez et al., 2016)	Model solutions of grape must components and red wine must	Alcoholic fermentation	Ethanol	3 LEDs emitting at 1.20, 1.30 and 1.70 μm	Up to 180 hours
(Kasemsumran et al., 2022)	Pineapple fruit wine	Alcoholic fermentation	Ethanol, total soluble solids, total acidity and total volatile acids	11536–3952 cm^{-1}	Up to 10 days
(Ouyang et al., 2016)	Rice wine	Alcoholic fermentation	Total sugar content, ethanol and pH	300-1000 nm	Up to 35 days
(Peng et al., 2016)	Apple wine	Alcoholic fermentation	Alcohol strength and titratable acidity	12000-4000 cm^{-1}	Up to 10 days
(Ye et al., 2014)	Apple wine	Alcoholic fermentation	Soluble solid content, pH, total acidity and total ester content	12000-4000 cm^{-1}	Not specified
(Wu et al., 2015)	Chinese rice wine	Alcoholic fermentation	Ethanol and total acid	10000-4000 cm^{-1}	Up to 20 days
(Sedjoah et al., 2021)	Mulberry vinegar	Acetic fermentation	Total acids and total polyphenol contents	900–1700 nm	Up to 10 days
(Chen et al., 2023)	Vinegar	Acetic fermentation	Ethanol and reducing sugar content	900–1700 nm	Up to 2 weeks
(González-Sáiz et al., 2008)	Onion vinegar	Acetic fermentation	Biomass, ethanol and acetic acid content	1100–2500 nm	Up to 90 hours
(Phanomsophon et al., 2019)	Rice vinegar	Acetic fermentation	Acetic acid and ethanol	12500-4000 cm^{-1}	Up to 10 days
(Yano et al., 1997)	Rice vinegar	Acetic fermentation	Acetic acid and ethanol	400-2500 nm	Up to 8 hours

To the best of our knowledge, no study has been conducted so far to track the entire progression of the dual surface fermentation, encompassing both phases over an extended time period. In this work, we monitored the course of simultaneous fermentation for two different batches of red grape must, to assess the reproducibility of the process kinetics when musts from different red grape varieties are kept under the same environmental conditions. Fermentation was monitored by analysing ethanol content and total acidity for almost four months, evaluating the evolution over time of both parameters.

At the same time, grape must samples were analysed using FT-NIR spectroscopy over a broad spectral range between 12500 and 4000 cm^{-1}. The resulting spectra, after appropriate preprocessing, were initially explored by Principal Component Analysis (PCA) (Bro and Smilde, 2014), to visualize in the principal components space how they evolved in the different phases of fermentation.

Subsequently, multivariate calibration models were developed and validated to predict ethanol content and total acidity using Partial Least Squares (PLS) (Geladi and Kowalski, 1986). Firstly, a PLS model was calculated considering the entire spectral range (12500-4000 cm^{-1}, equivalent to 800-2500 nm). Then, the feasibility of achieving satisfactory calibration models using narrower spectral regions was also investigated, since this could allow the use of cheaper and miniaturizable sensors with potential applications in portable devices. Specifically, we focused on two spectral regions labelled for simplicity as "NIR1" and "NIR2" covering the ranges 10526-6060 cm^{-1} (950-1650 nm) and 12500-9346 cm^{-1} (800-1070 nm), respectively. The choice to shorten the spectra to obtain the corresponding signals in these spectral regions is driven by the need to simulate spectra acquired with commercially available devices working in these spectral ranges. The NIR1 region corresponds to the spectral range that could be acquired using an uncooled InGaAs detector, which is significantly more economical than its cooled counterpart. On the other hand, the NIR2 region is towards the NIR region closer to the visible one, which could also allow the use of much cheaper silicon detectors (Pu et al., 2021; Wang et al., 2020).

Monitoring of alcoholic and acetic fermentation in red grape musts through FT-NIR has proven effective in accurately determining ethanol content and total acidity values both considering the entire spectral range and focusing on the region labelled as NIR1. Less satisfactory but still acceptable results were obtained in the region referred to as NIR2, confirming the potential use of simple and cost-effective devices for real-time monitoring of fermenting must.

2 Materials and Methods

2.1 Must samples and fermentation conditions

In this study, two samples of red grape must, referred to as batch A and batch B, both produced in the agricultural year 2022 in Emilia Romagna Region, Italy, were subjected to fermentation. Batch A must was from *Ancellotta* grape variety, while batch B must was from a blend of *Lambrusco Salamino* and *Lambrusco Grasparossa* grape varieties.

After pressing the grapes, the grape musts were cooled and stabilized at a temperature between 0 and 4°C. Subsequently, as the experimentation began, they were brought to room temperature (ranging from 18 to 25 °C during the whole fermentation period) and inoculated with *Saccharomyces cerevisiae* yeast to initiate the alcoholic fermentation process. Throughout the alcoholic fermentation phase, two aliquots of grape must were collected almost daily for each batch. One was promptly analysed in the laboratory to determine ethanol content and total acidity, while the other was frozen for subsequent spectroscopic analysis. Starting around the 40th day of fermentation, the acetic bio-oxidation phase began spontaneously (acetobacteria were present in the air of the fermentation room, which is adjacent to a vinegar production area). During this phase, sampling of the two aliquots for each batch was conducted every 2-4 days. In total, 49 samples were collected for batch A and 40 samples for batch B, for which the fermentation process proceeded more rapidly. Monitoring ended after 116 days for batch A and after 80 days for batch B.

2.2 Laboratory analysis and spectra acquisition

Each sampling day, ethanol content (% v/v) was measured in triplicate using a Malligand ebulliometer (Ing. Bullio Castore, Milan, Italy). Total acidity (g/100 ml) was also measured in triplicate utilizing a potentiometric titrator (TitraLab AT1000, Hach-Crison, Barcelona, Spain). Additionally, at the time of aliquot collection, the temperature of the musts was recorded with a food thermometer (Tecna2, Tecnafood, Bomporto, Italy) to track the trend and identify potential fermentation anomalies.

Regarding spectroscopic analysis, after some preliminary tests, an appropriate sample treatment and signal acquisition procedure was developed. The aliquots of frozen must intended for analysis were transferred from the freezer to the refrigerator the evening before the analysis day to ensure complete thawing of the sample without restarting fermentation. Thirty minutes prior to analysis, the samples were extracted from the refrigerator and filtered through a 25 µm cellulose filter to eliminate suspended particles.

Fourier-transform near-infrared (FT-NIR) spectra were acquired in transmission mode using a Bruker Multi Purpose Analyzer (MPA) spectrophotometer equipped with a TE-InGaAs sensor, that is thermoelectrically cooled to ensure a high signal-to-noise ratio (Beć et al., 2022). Measurements were performed using a quartz Suprasil cuvette with an optical path length of 0.5 mm. A background spectrum was acquired on the empty cuvette before the analysis of each aliquot. To evaluate measurement repeatability and enhance the robustness of calibration models, three spectra were acquired for each sample, corresponding to three distinct sample aliquots. Each spectrum, derived from the average of 128 scans, was collected throughout the entire spectral range (12500-4000 cm^{-1}) excluding the low-wavenumber region between 4000 and 3600 cm^{-1} where the signal was too noisy. Considering a spectral resolution of 4 cm^{-1}, each spectrum was composed of a total of 4407 variables. Given that it is well-known from the literature that temperature variation causes changes in both the position and intensity of bands in the NIR spectrum, especially for aqueous and alcoholic samples (Cozzolino et al., 2007; Wülfert et al., 1998), the cuvette compartment was maintained at a constant temperature of 28°C during measurements. This was implemented to mitigate any potential variability associated with differences in the laboratory's ambient temperature throughout the entire period during which NIR analyses were conducted.

2.3 Data Analysis

The spectral dataset, including 267 spectra, i.e. (49 samples from batch A + 40 samples from batch B) × 3 replicates, was first subjected to exploratory data analysis using Principal Component Analysis (PCA). This analysis aimed to evaluate the reproducibility of measurements, potential variations between samples from different batches, identification of outliers, and exploration of trends related to ethanol content and total acidity. Based on some preliminary tests, spectra were pre-processed using Standard Normal Variate (SNV) and mean centering.

Subsequently, Partial Least Squares (PLS) was employed as a multivariate calibration method to determine the values of ethanol and acidity from NIR spectra. The calibration models were initially derived from full NIR spectra, encompassing a spectral range spanning the entire acquired region (12500-4000 cm^{-1}). Analogous PLS models were also developed for two additional spectral datasets, obtained by trimming the signals to keep only specific spectral ranges. In detail, the dataset abbreviated as NIR1 encompasses the spectral range from 10526 to 6060 cm^{-1} (equivalent to 950-1650 nm), while the dataset referred to as NIR2 covers the range from 12500 to 9346 cm^{-1} (corresponding to 800-1070 nm).

To calculate all the PLS models, the spectral datasets were randomly divided into a calibration set (training set) containing 180 spectra (60 samples × 3 replicates) and an external validation set (test set) containing the remaining 87 spectra (29 samples × 3 replicates). Based on some preliminary trials conducted on the training set, the optimal spectral preprocessing for predicting the parameters of interest resulted from a combination of smoothing (first-order, 15 points window), SNV, and mean centering. The optimal number of latent variables (LVs) was chosen by minimizing the cross-validation error, that was estimated using a venetian blinds procedure with 10 deletion groups, each including 18 spectra (corresponding to the 3 replicates of 6 samples). Subsequently, the models were validated using the external test set.

The performance of the calibration models was evaluated using Root Mean Square Error (RMSE) and coefficient of determination (R^2) metrics. These metrics were calculated on the training set for calibration (RMSEC, R^2_{Cal}) and cross-validation (RMSECV, R^2_{CV}), and then on the test set for validation (RMSEP, R^2_{Pred}).

The PLS models were calculated using the PLS Toolbox (version 8.8.1, Eigenvector Research Inc., USA) operating in the MATLAB environment (version 2022b, The Mathworks Inc., USA).

3 Results and discussion

3.1 Exploratory analysis of laboratory parameters

The laboratory-measured physicochemical parameters (ethanol content, total acidity and temperature) were plotted to evaluate their trends over time for the two batches. Figure 1a illustrates the variation in ethanol content and total acidity with respect to fermentation time, where each reported value is the average of three replicate measurements. From the graph it can be observed that the ethanol content and total acidity values followed a similar pattern in both batches, but each batch progressed at a different rate. Initially, a rapid increase in ethanol content was observed, reaching its peak around day 10 for batch B and around day 25 for batch A. Around day 40 total acidity started to increase, while alcohol content declined rapidly. For batch B, ethanol was depleted after 80 days, leading to the end of fermentation. In the case of batch A, the formation of acetic acid continued more gradually and at the time of sampling interruption, ethanol was not yet fully depleted.

For thoroughness, the graph illustrating the temperature changes over time is included as well (Figure 1b). Temperature demonstrated a more stable pattern for batch A, whereas it exhibited two distinct peaks for batch B. Specifically, for batch B, temperature reached around 30°C at the peak of alcoholic fermentation, then decreased and slightly increased again at the onset of acetic oxidation.

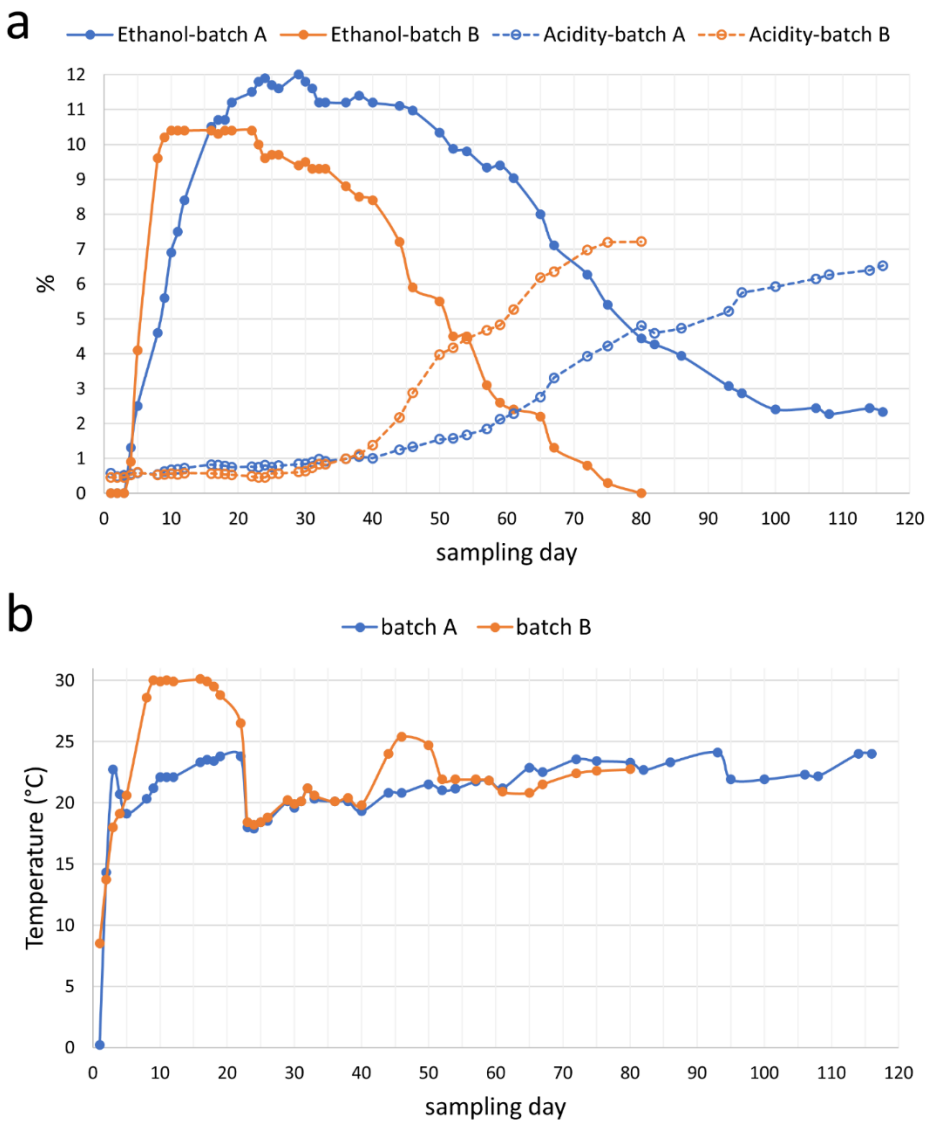


Figure 1 *Variation in ethanol content and total acidity (a) and temperature (b) on different sampling days during the fermentation period.*

3.2 Exploratory analysis of FT-NIR spectra

Figure 2 reports the score plots of the first two principal components (PCs) of the PCA model calculated on the whole dataset of FT-NIR spectra, accounting for 91.20% of total data variance. In Figure 2a, were the samples are color-coded by batch, it is observed that the spectra belonging to the two batches are distinguishable in the PC space, although the fermentation evolution essentially follows the same trend for both musts. This analogy becomes more evident when the samples are coloured based on the sampling day, as reported in Figure 2b. Up to approximately day 40, the values

of PC1 remain approximately constant, while the values of PC2 decrease. Then, the values of PC1 gradually increase, while the values of PC2 show a less pronounced rise.

In Figures 2c, d, the same score plot is depicted based on ethanol content and total acidity values, respectively. Figure 2c shows that, during alcoholic fermentation, the PC2 values gradually decrease while the PC1 values remain relatively constant. Subsequently, as acetic oxidation begins, the PC1 value progressively rises (Figure 2d) in tandem with the gradual consumption of ethanol (Figure 2c). Therefore, samples with high acidity exhibit high PC1 values, while samples with high ethanol content have low values for both PC1 and PC2.

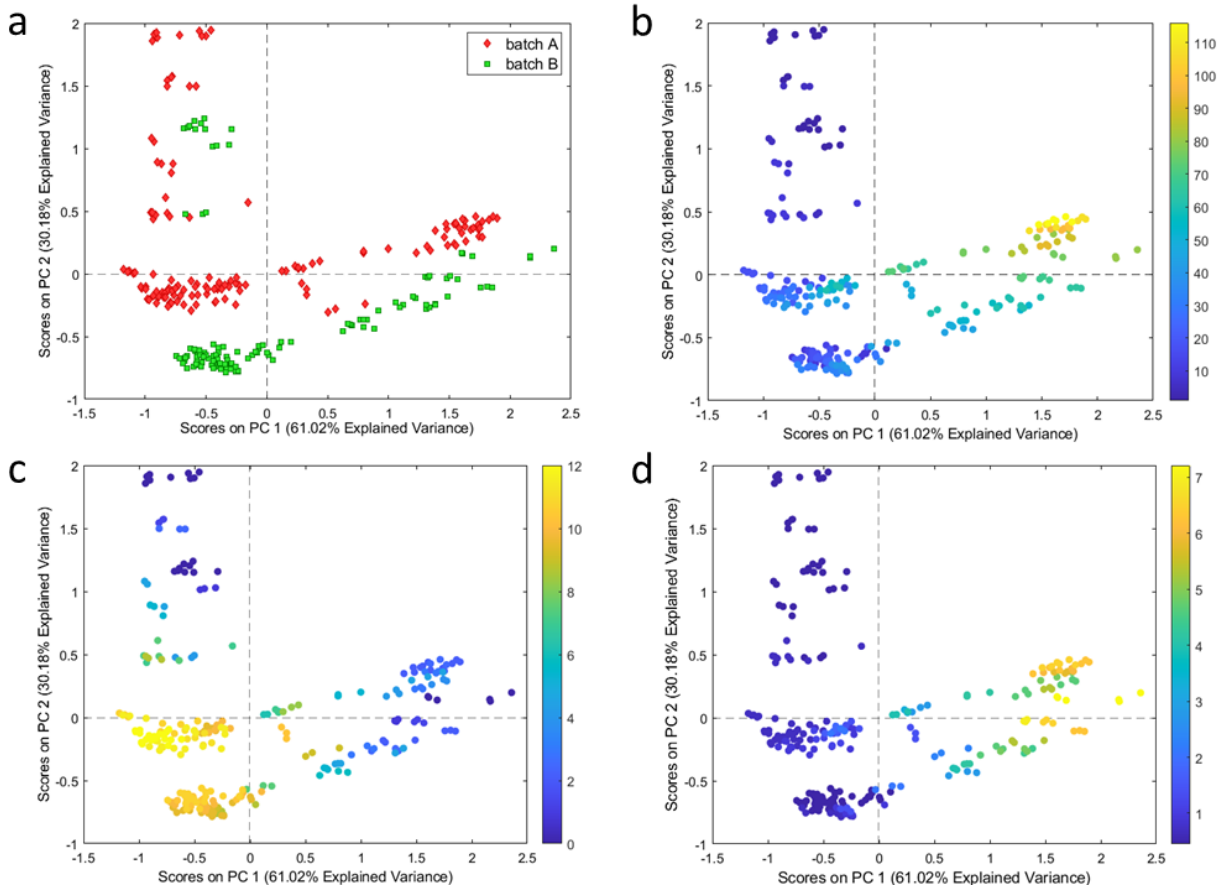


Figure 2 *Score plots of PC1 vs PC2 of the PCA model obtained on the complete FT-NIR spectral dataset. Samples are coloured based on batch (a), on fermentation day (b), on ethanol content (c), and on total acidity (d).*

Through the loading plots, it is possible to analyse the contribution of each spectral variable to the two PCs of the model and identify correlations. In Figure 3a, the original spectra are coloured based on the fermentation day, along with indications of the attributions of the main spectral bands according to the literature (Cozzolino et al., 2006; Di Egidio et al., 2010; Dong et al., 2019;

Kasemsumran et al., 2022; Nespeca et al., 2017; Ye et al., 2014). In Figure 3b, the loading plots of PC1 and PC2 are reported, in comparison with the pre-processed spectra, coloured based on both ethanol content (Figure 3c) and total acidity (Figure 3d). Firstly, it can be observed that from the raw spectra (Figure 3a), it is not possible to clearly discern any trend with fermentation time, as the signals are highly overlapped and exhibit a very similar profile. However, the subsequent application of preprocessing highlights the differences (Figures 3c and 3d).

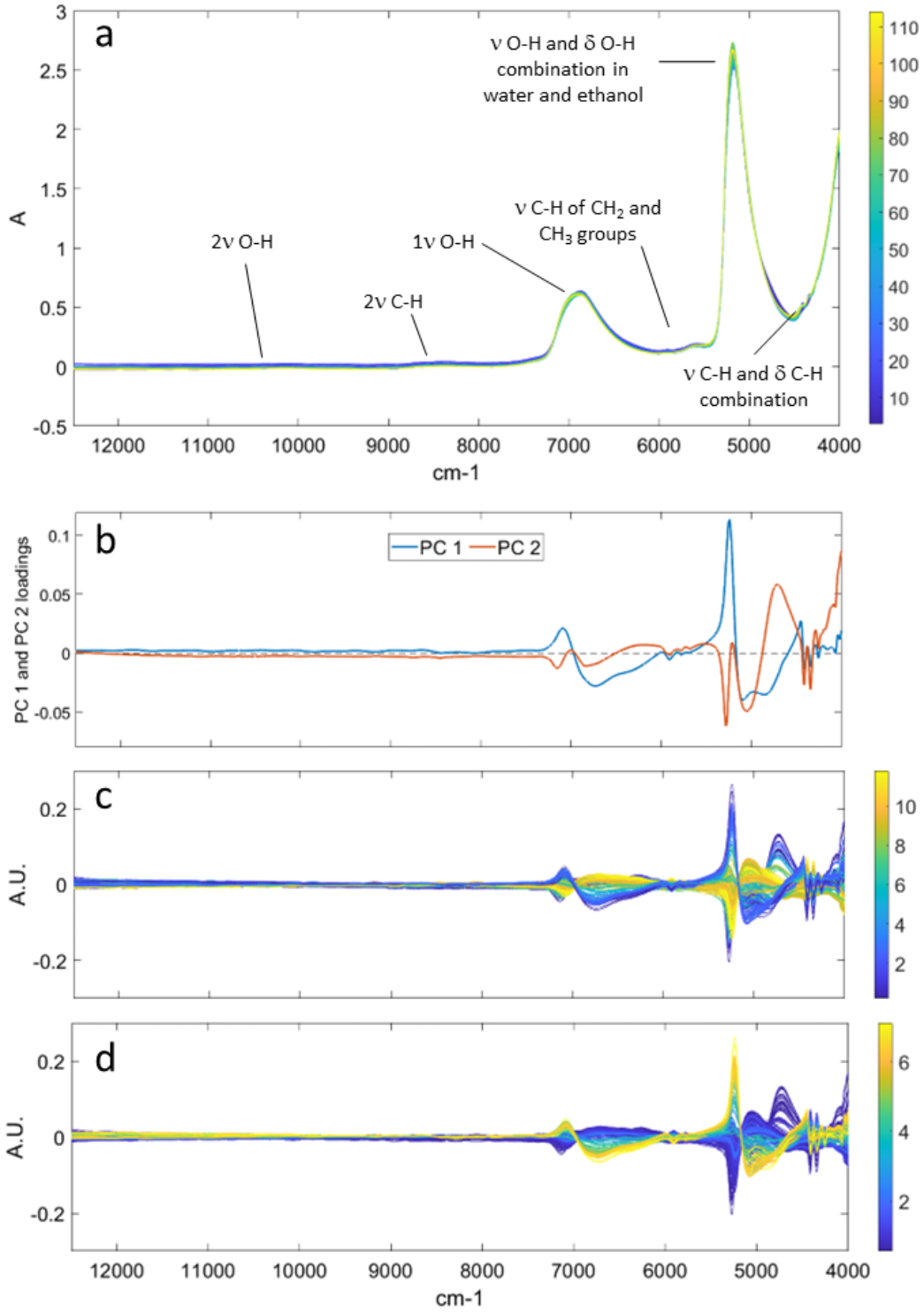


Figure 3 *Original spectra coloured according to the fermentation day with assignment of the main spectral bands (a) compared to the loadings plots of PC1 and PC2 (b) and with pre-processed spectra coloured based on ethanol content (c) and total acidity (d).*

According to existing literature, it is well-established that the absorption pattern in the NIR region during the fermentation of grape must is predominantly influenced by the water spectrum (Jiménez-Márquez et al., 2016). The presence of water gives rise to intense and broad bands attributed to the O-H bend stretching and the combination of O-H stretching and bending, as illustrated in Figure 3a. However, it is worth noting that ethanol and acetic acid absorb at approximately the same wavelengths, as these molecules partially share similar chemical bonds and functional groups, just like fructose, glucose, and glycerol always present in grape musts. Therefore, any differences in the spectra of samples at different fermentation stages can be attributed to subtle variations in the relative intensity of peaks and/or to the horizontal shift of these peaks due to their chemical surroundings.

Nonetheless, it is interesting to observe how the two adjacent peaks at around 4400-4300 cm^{-1} increase with the rise in alcohol percentage, while the loadings of PC2 decrease in this region, aligning with the patterns seen in the score plot of Figure 2c. These peaks are distinctive of ethanol, attributed to the combination band of C-H in the $-CH_2$ and $-CH_3$ groups (Cozzolino et al., 2006; Di Egidio et al., 2010; Dong et al., 2019). Conversely, the spectral variables that appear more correlated with acetic acid content (Figure 3d) are those associated with the peak at approximately 5170 cm^{-1}, linked to the combination of O-H stretching and bending vibrations. In addition to increasing in intensity (like the loadings of PC1), it undergoes a slight shift towards higher wavenumbers (Cozzolino et al., 2006; Di Egidio et al., 2010).

3.3 Calibration

As detailed in Section 2.3, calibration models were developed for three distinct datasets: the complete NIR spectra dataset (12500-4000 cm^{-1}), the dataset labelled as NIR1 (10526-6060 cm^{-1}), and the dataset labelled as NIR2, which includes only the spectral region adjacent to the visible range (12500-9346 cm^{-1}). In each case, the same sample partitioning into training set, test set, and deletion groups for cross-validation was used, along with identical signal preprocessing. This ensured complete comparability of the models.

Table 2 provides a comparison of the performance of all the obtained calibration models. It is worth noting that excellent models were achieved for both parameters when considering both the complete spectrum and the NIR1 region, with R^2_{Pred} values exceeding 0.98 in each case. However, performance in the NIR2 region is notably less satisfactory, albeit still of acceptable quality for screening analysis aimed at providing indicative results on ethanol and acidity content in grape musts.

Table 2 *Performance of PLS models obtained for ethanol content and total acidity considering the three spectral ranges.*

Spectral range	**Y**	**#LVs**	**RMSEC**	**RMSECV**	**RMSEP**	**R^2_{Cal}**	**R^2_{CV}**	**R^2_{Pred}**
Complete	Ethanol	5	0.399	0.435	0.412	0.990	0.988	0.990
	Acidity	6	0.114	0.135	0.218	0.997	0.996	0.990
NIR1	Ethanol	5	0.427	0.469	0.463	0.989	0.986	0.987
	Acidity	8	0.147	0.169	0.196	0.995	0.994	0.992
NIR2	Ethanol	8	1.625	1.981	1.561	0.834	0.755	0.849
	Acidity	5	0.907	1.040	0.909	0.818	0.762	0.813

Figure 4 reports the laboratory-measured parameter values against the values predicted by the FT-NIR models for the three spectral regions and the two investigated properties. Consistently with the model performance parameters outlined in Table 2, the measured values closely align with the predicted values for both the complete spectrum and the NIR1 region (Figures 4a-d). On the contrary, in the graphs regarding the NIR2 region (Figures 4e and 4f), a higher dispersion of samples around the fitted line (represented in red) is observed; furthermore, the bisector (represented in green) and the fitted line do not overlap, in contrast to the earlier models. Specifically, for batch B samples, both models predict ethanol content and total acidity values higher than the actual values at low levels of these parameters; for batch A, the total acidity model predicts lower values than the actual ones at low total acidity levels.

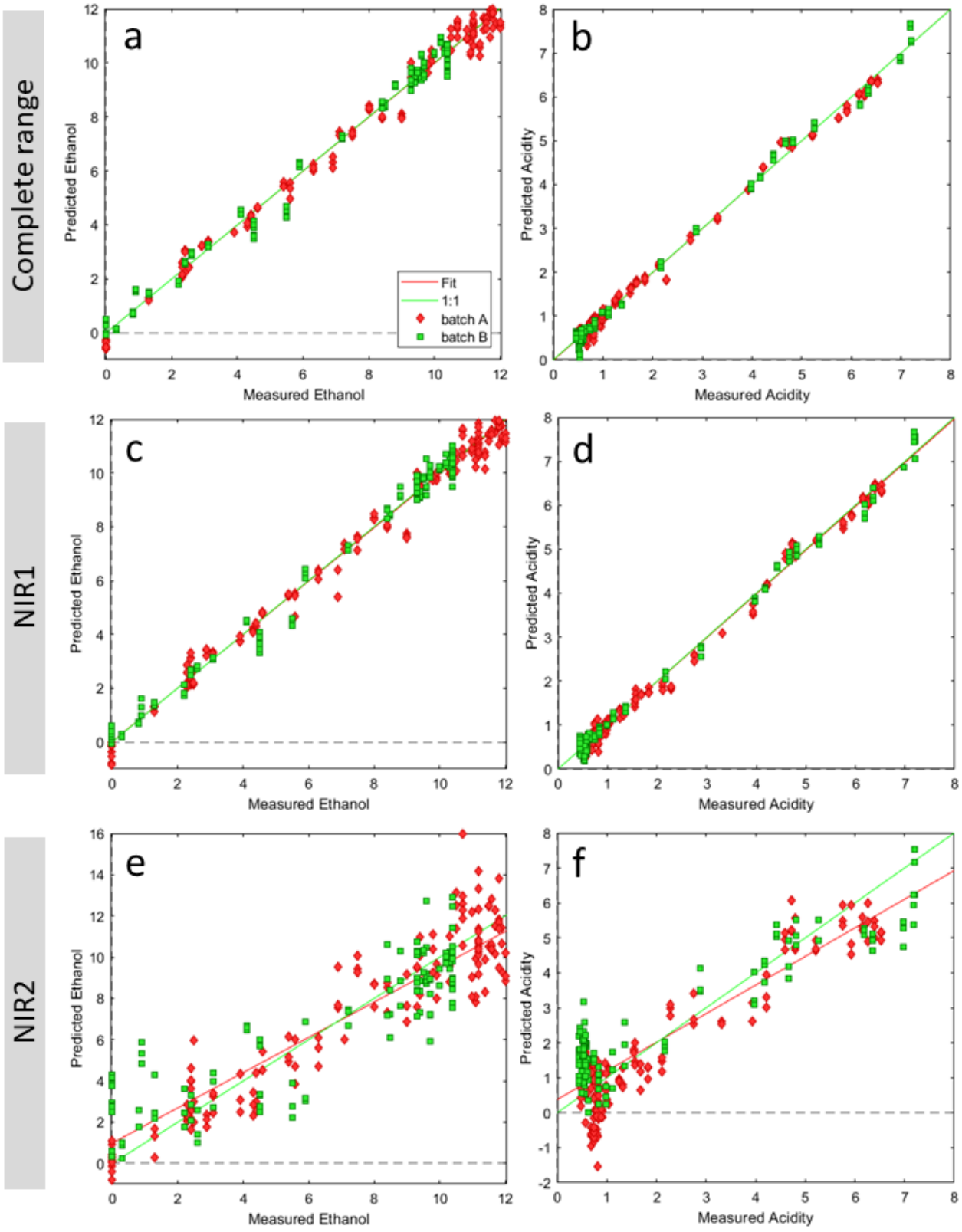


Figure 4 *Y measured vs. Y predicted plots of PLS models obtained on the complete spectra for ethanol content (a) and total acidity (b), on the NIR1 region for ethanol content (c) and total acidity (d), and for the NIR2 region for ethanol content (e) and total acidity (f).*

Finally, in Figure 5 the Variable Importance in Projections (VIP) scores of the models are presented. The spectral variables positioned above the threshold value set equal to 1 (red dashed line) in the VIP plots are the most valuable for modelling ethanol content and acidity (Gosselin et al., 2010). It is evident that, for the complete spectrum, the VIP scores of the ethanol content model (Figure 5a) differ from those of the acidity model (Figure 5b), especially in the region of the two adjacent peaks at around 4400-4300 cm^{-1}. As mentioned earlier, these peaks are characteristic of ethanol (C-H combination bands in the $-CH_2$ and $-CH_3$ groups). Crucial for constructing both models are the spectral variables associated with the combination of O-H stretching and bending vibrations (around 5170 cm^{-1}) and the stretching of the O-H bond (around 7000-6800 cm^{-1}).

For the models calculated considering the NIR1 region (Figures 5c and 5d), there is a slight alteration in the VIP scores graphs, given that these models cannot benefit from the information of bands below 6000 cm^{-1}. Consequently, the significant variables primarily involve the stretching of the O-H bond, a functional group found in both ethanol and acetic acid, as well as in the predominant aqueous matrix. Comparing the VIP scores with the shape of the original signals in Figure 3a, it emerges that, during fermentation, this peak decreases in intensity and gradually shifts towards higher wavenumbers.

For the models in the NIR2 region, the presence of noise due to the low signal-to-noise ratio in this spectral region is immediately noticeable. However, the application of smoothing preprocessing has partially mitigated this issue. Additionally, from the VIP plots it is observed how some spectral variables, not considered fundamental in the previous models, have now become significant. Specifically, the graphs suggest that the peak where the most pertinent information for the models is concentrated lies at approximately 10300 cm^{-1}, corresponding to the second overtone of the O-H stretching vibration (Jiménez-Márquez et al., 2016).

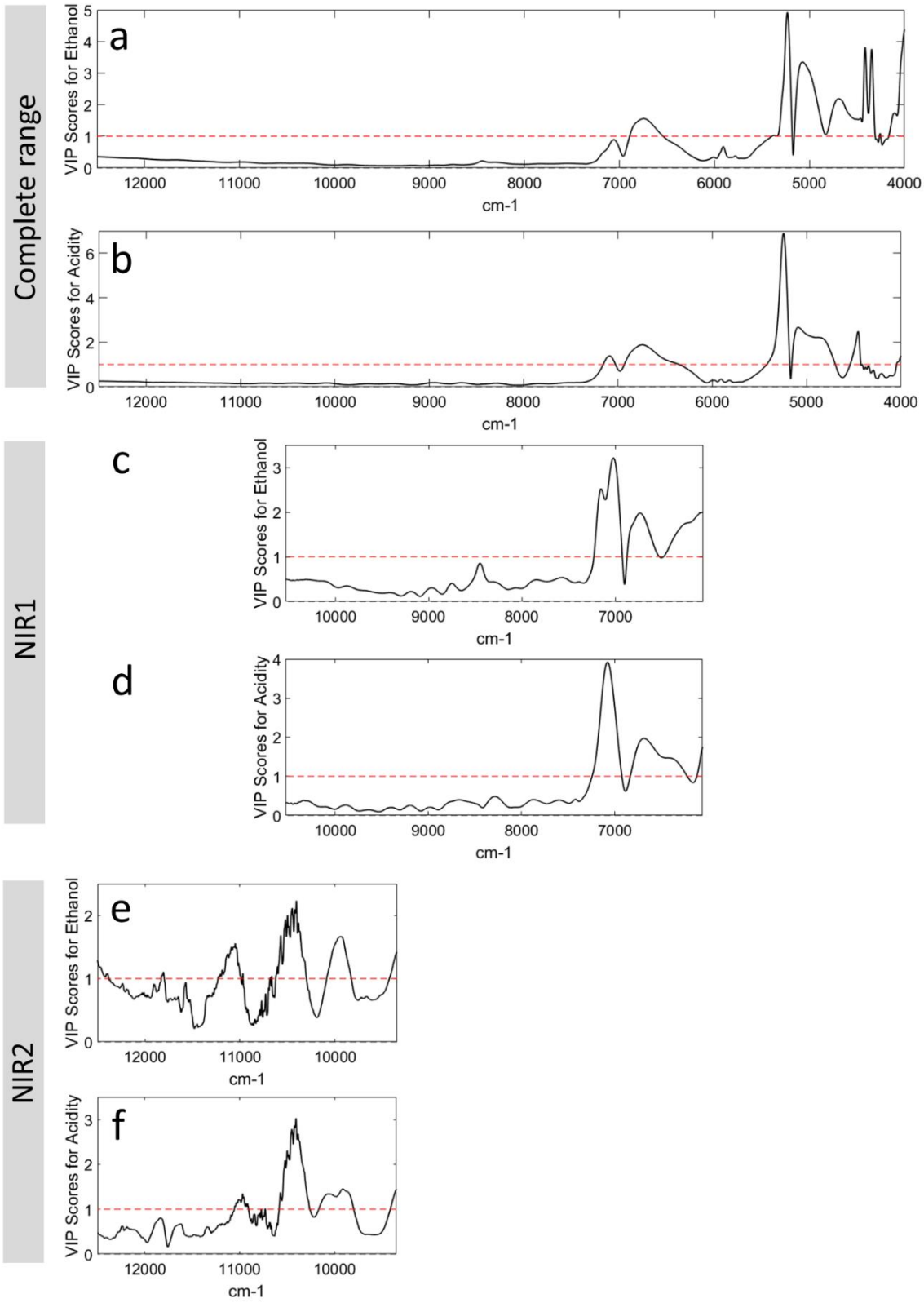


Figure 5 *VIP score plots of PLS models obtained on the complete spectra for ethanol content (**a**) and total acidity (b), on the NIR1 region for ethanol content (c) and total acidity (d), and for the NIR2 region for ethanol content (e) and total acidity (f).*

4 Conclusions

The aim of this study was to test the effectiveness of NIR spectroscopy to monitor the surface fermentation process of red grape must, comparing the results that could be obtained with devices operating in different NIR spectral regions. To this aim, we conducted a simultaneous analysis of two different batches of red grape must undergoing fermentation over an extended period up to about 4 months. During this time, the musts were subjected to both alcoholic and acetic fermentation stages. In particular, our goal was to simulate the use of devices operating in different spectral regions to assess the potential of various commercial instruments. Specifically, models were generated over the entire range from 12500-4000 cm^{-1}, corresponding to expensive bench-top instruments, in the 10526-6060 cm^{-1} range, corresponding to high-end portable instruments, and in the 12500-9346 cm^{-1} range, corresponding to economical handheld instruments.

Overall, the results were excellent for models obtained considering the entire spectrum and the 10526-6060 cm^{-1} region. Results obtained for the 12500-9346 cm^{-1} region were less satisfactory but still acceptable for fast screening. This indicates that FT-NIR spectroscopy, even in its simpler and more economical versions, remains an optimal choice for product monitoring in the field of vinegar production, as it combines speed and efficiency with the non-invasiveness of the analysis, preserving the integrity of the sample.

Acknowledgements

The authors sincerely thank Beatrice Guzzi and Paolo Tucci of Acetyca s.r.l. (Saronno, Italy) for their generous support and collaboration.

Author contributions

C. Menozzi: investigation, methodology, formal analysis, and writing—original draft. G. Foca: conceptualization, methodology, formal analysis, and writing—original draft. **R. Calvini:** methodology, formal analysis, and writing—review and editing. **L. Catellani:** investigation, methodology, and writing—review and editing. **A. Bezzecchi:** supervision, conceptualization, and writing—review and editing. **A. Ulrici:** supervision, project administration, conceptualization, and writing—review and editing.

Funding

The project was supported by Acetyca S.r.l. (Saronno, Italy).

5 References

Barbon Junior, S., Mastelini, S.M., Barbon, A.P.A.C., Barbin, D.F., Calvini, R., Lopes, J.F., Ulrici, A., 2020. Multi-target prediction of wheat flour quality parameters with near infrared spectroscopy. Information Processing in Agriculture 7, 342–354. https://doi.org/10.1016/j.inpa.2019.07.001

Beć, K.B., Grabska, J., Huck, C.W., 2022. Miniaturized NIR Spectroscopy in Food Analysis and Quality Control: Promises, Challenges, and Perspectives. Foods 11, 1465. https://doi.org/10.3390/foods11101465

Bhat, S.V., Akhtar, R., Amin, T., 2014. An Overview on the Biological Production of Vinegar. Internat. Jour. of Fermen. Foods 3, 139. https://doi.org/10.5958/2321-712X.2014.01315.5

Bro, R., Smilde, A.K., 2014. Principal component analysis. Anal. Methods 6, 2812–2831. https://doi.org/10.1039/C3AY41907J

Buratti, S., Ballabio, D., Giovanelli, G., Dominguez, C.M.Z., Moles, A., Benedetti, S., Sinelli, N., 2011. Monitoring of alcoholic fermentation using near infrared and mid infrared spectroscopies combined with electronic nose and electronic tongue. Analytica Chimica Acta 697, 67–74. https://doi.org/10.1016/j.aca.2011.04.020

Burns, D.A., Ciurczak, E.W. (Eds.), 2007. Handbook of Near-Infrared Analysis, 0 ed. CRC Press. https://doi.org/10.1201/9781420007374

Chen, C., Li, X., Zhu, S., Cui, P., Lei, H., Yan, H., 2023. Detection of the alcohol fermentation process in vinegar production with a digital micro-mirror based NIR spectra set-up and chemometrics. Journal of Food Composition and Analysis 115, 105036. https://doi.org/10.1016/j.jfca.2022.105036

Cozzolino, D., 2016. State-of-the-art advantages and drawbacks on the application of vibrational spectroscopy to monitor alcoholic fermentation (beer and wine). Applied Spectroscopy Reviews 51, 302–317. https://doi.org/10.1080/05704928.2015.1132721

Cozzolino, D., Liu, L., Cynkar, W.U., Dambergs, R.G., Janik, L., Colby, C.B., Gishen, M., 2007. Effect of temperature variation on the visible and near infrared spectra of wine and the consequences on the partial least square calibrations developed to measure chemical composition. Analytica Chimica Acta 588, 224–230. https://doi.org/10.1016/j.aca.2007.01.079

Cozzolino, D., Parker, M., Dambergs, R.G., Herderich, M., Gishen, M., 2006. Chemometrics and visible-near infrared spectroscopic monitoring of red wine fermentation in a pilot scale. Biotechnol. Bioeng. 95, 1101–1107. https://doi.org/10.1002/bit.21067

Di Egidio, V., Sinelli, N., Giovanelli, G., Moles, A., Casiraghi, E., 2010. NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation. Eur Food Res Technol 230, 947–955. https://doi.org/10.1007/s00217-010-1227-5

Dong, Q., Yu, C., Li, L., Nie, L., Li, D., Zang, H., 2019. Near-infrared spectroscopic study of molecular interaction in ethanol-water mixtures. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 222, 117183. https://doi.org/10.1016/j.saa.2019.117183

Ferrari, E., Foca, G., Vignali, M., Tassi, L., Ulrici, A., 2011. Adulteration of the anthocyanin content of red wines: Perspectives for authentication by Fourier Transform-Near InfraRed and 1H NMR spectroscopies. Analytica Chimica Acta 701, 139–151. https://doi.org/10.1016/j.aca.2011.05.053

Foca, G., Ferrari, C., Sinelli, N., Mariotti, M., Lucisano, M., Caramanico, R., Ulrici, A., 2011. Minimisation of instrumental noise in the acquisition of FT-NIR spectra of bread wheat using experimental design and signal processing techniques. Anal Bioanal Chem 399, 1965–1973. https://doi.org/10.1007/s00216-010-4431-z

Foca, G., Ferrari, C., Ulrici, A., Ielo, M.C., Minelli, G., Lo Fiego, D.P., 2016. Iodine Value and Fatty Acids Determination on Pig Fat Samples by FT-NIR Spectroscopy: Benefits of Variable Selection in the Perspective of Industrial Applications. Food Anal. Methods 9, 2791–2806. https://doi.org/10.1007/s12161-016-0478-6

Frohman, C.A., Mira De Orduña Heidinger, R., 2018. The substratostat - an automated near-infrared spectroscopy-based variable-feed system for the continuous alcoholic fed-batch fermentation of high gravity grape must at static sugar concentrations. OENO One 52. https://doi.org/10.20870/oeno-one.2018.52.4.2199

Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Analytica Chimica Acta 185, 1–17. https://doi.org/10.1016/0003-2670(86)80028-9

Gishen, M., Dambergs, R.G., Cozzolino, D., 2005. Grape and wine analysis - enhancing the power of spectroscopy with chemometrics..: A review of some applications in the Australian wine industry. Aust J Grape Wine Res 11, 296–305. https://doi.org/10.1111/j.1755-0238.2005.tb00029.x

González-Sáiz, J.M., Esteban-Díez, I., Sánchez-Gallardo, C., Pizarro, C., 2008. Monitoring of substrate and product concentrations in acetic fermentation processes for onion vinegar production by NIR spectroscopy: value addition to worthless onions. Anal Bioanal Chem 391, 2937–2947. https://doi.org/10.1007/s00216-008-2186-6

Gosselin, R., Rodrigue, D., Duchesne, C., 2010. A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications. Chemometrics and Intelligent Laboratory Systems 100, 12–21. https://doi.org/10.1016/j.chemolab.2009.09.005

Jagtap, U.B., Bapat, V.A., 2015. Wines from fruits other than grapes: Current status and future prospectus. Food Bioscience 9, 80–96. https://doi.org/10.1016/j.fbio.2014.12.002

Jiménez-Márquez, F., Vázquez, J., Úbeda, J., Sánchez-Rojas, J.L., 2016. Optoelectronic sensor for measuring ethanol content during grape must fermentation using NIR spectroscopy. Microsyst Technol 22, 1799–1809. https://doi.org/10.1007/s00542-016-2835-1

Kasemsumran, S., Boondaeng, A., Ngowsuwan, K., Jungtheerapanich, S., Apiwatanapiwat, W., Janchai, P., Meelaksana, J., Vaithanomsat, P., 2022. Simultaneous Monitoring of the Evolution of Chemical Parameters in the Fermentation Process of Pineapple Fruit Wine Using the Liquid Probe for Near-Infrared Coupled with Chemometrics. Foods 11, 377. https://doi.org/10.3390/foods11030377

Nespeca, M.G., Rodrigues, C.V., Santana, K.O., Maintinguer, S.I., De Oliveira, J.E., 2017. Determination of alcohols and volatile organic acids in anaerobic bioreactors for H 2 production by near infrared spectroscopy. International Journal of Hydrogen Energy 42, 20480–20493. https://doi.org/10.1016/j.ijhydene.2017.07.044

Ouyang, Q., Zhao, J., Pan, W., Chen, Q., 2016. Real-time monitoring of process parameters in rice wine fermentation by a portable spectral analytical system combined with multivariate analysis. Food Chemistry 190, 135–141. https://doi.org/10.1016/j.foodchem.2015.05.074

Peng, B., Ge, N., Cui, L., Zhao, H., 2016. Monitoring of alcohol strength and titratable acidity of apple wine during fermentation using near-infrared spectroscopy. LWT - Food Science and Technology 66, 86–92. https://doi.org/10.1016/j.lwt.2015.10.018

Phanomsophon, T., Sirisomboon, P., Lapcharoensuk, R., Shrestha, B., Krusong, W., 2019. Evaluation of acetic acid and ethanol concentration in a rice vinegar internal venturi injector bioreactor using Fourier transform near infrared spectroscopy. Journal of Near Infrared Spectroscopy 27, 416–423. https://doi.org/10.1177/0967033519870304

Pu, Y., Pérez-Marín, D., O'Shea, N., Garrido-Varo, A., 2021. Recent Advances in Portable and Handheld NIR Spectrometers and Applications in Milk, Cheese and Dairy Powders. Foods 10, 2377. https://doi.org/10.3390/foods10102377

Sedjoah, R.-C.A.-A., Ma, Y., Xiong, M., Yan, H., 2021. Fast monitoring total acids and total polyphenol contents in fermentation broth of mulberry vinegar using MEMS and optical fiber near-infrared spectrometers. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 260, 119938. https://doi.org/10.1016/j.saa.2021.119938

Solieri, L., Giudici, P. (Eds.), 2009. Vinegars of the World. Springer Milan, Milano. https://doi.org/10.1007/978-88-470-0866-3

Wang, B., Ranjan, R., Khot, L.R., Peters, R.T., 2020. Smartphone Application-Enabled Apple Fruit Surface Temperature Monitoring Tool for In-Field and Real-Time Sunburn Susceptibility Prediction. Sensors 20, 608. https://doi.org/10.3390/s20030608

Wu, Z., Xu, E., Wang, F., Long, J., Jiao, X.X.A., Jin, Z., 2015. Rapid Determination of Process Variables of Chinese Rice Wine Using FT-NIR Spectroscopy and Efficient Wavelengths Selection Methods. Food Anal. Methods 8, 1456–1467. https://doi.org/10.1007/s12161-014-0021-6

Wülfert, F., Kok, W.Th., Smilde, A.K., 1998. Influence of Temperature on Vibrational Spectra and Consequences for the Predictive Ability of Multivariate Models. Anal. Chem. 70, 1761–1767. https://doi.org/10.1021/ac9709920

Yano, T., Aimi, T., Nakano, Y., Tamai, M., 1997. Prediction of the concentrations of ethanol and acetic acid in the culture broth of a rice vinegar fermentation using near-infrared spectroscopy. Journal of Fermentation and Bioengineering 84, 461–465. https://doi.org/10.1016/S0922-338X(97)82008-9

Ye, M., Yue, T., Yuan, Y., Li, Z., 2014. Application of FT-NIR Spectroscopy to Apple Wine for Rapid Simultaneous Determination of Soluble Solids Content, pH, Total Acidity, and Total Ester Content. Food Bioprocess Technol 7, 3055–3062. https://doi.org/10.1007/s11947-014-1385-8

Chapter Six

General conclusions

The work presented in this PhD Thesis explored the development and application of advanced analytical systems based on RGB imaging and NIR spectroscopy, coupled with chemometrics, for the analysis of food matrices, with a special attention to the wine sector.

The fast, extensive and non-destructive characterization of sample properties that can be achieved by means of imaging and spectroscopic techniques comes with a challenge: the large amount of data generated in a very short time. Indeed, the resulting datasets meet the concept of "Big Data", which require a multivariate statistical approach for an efficient extraction of relevant information. In this context, the present PhD Thesis was focused on the implementation of chemometric strategies able to properly manage Big Data from RGB imaging and NIR spectroscopy.

Concerning RGB images, this Thesis investigated the application of data dimensionality reduction algorithms to extract colour-related and texture-related information at the image-level. These methods encode the information of interest into feature vectors, which can be further elaborated using common chemometric strategies, like e.g., PCA or PLS. In this way, it is possible to simultaneously analyse large datasets of images and address the practical needs of most real-case scenarios, where hundreds or even thousands of images have to be acquired to solve the problem at hand.

These techniques were applied to a dataset of RGB images of red grape samples acquired using a smartphone-based portable device to predict the corresponding anthocyanins content through PLS calibration models. Firstly, the RGB images of red grape samples were processed at the image-level using the *colourgrams* method, a data reduction technique able to convert the colour-related information brought by each image into a signal. The best overall results in predicting anthocyanins were achieved for *Lambrusco Salamino* and *Sangiovese* varieties, while less satisfactory results were obtained for *Ancellotta* grape variety, probably due to the specific ripening behaviour associated to this variety. In general, the *colourgrams* approach successfully related the colour features of RGB images of the grape samples to their corresponding anthocyanins content. This represents the key idea behind the development of the portable smartphone-based device presented in this thesis, providing a smart, efficient, and low-cost alternative compared to traditional analytical methods for the determination of anthocyanins content.

Although colour is the primary aspect to consider when assessing anthocyanins content, the calibration models were further refined by also considering texture features of RGB images. Indeed,

the presence of the green pedicels of the grape berries – which share the same colour as unripe berries but have different shapes – and/or the image background could negatively affect model performance if only colour is considered. In this context, the *texturegrams* approach was presented for the first time in this Thesis. Texturegrams are an alternative image-level data dimensionality reduction method to extract and codify texture-related information from each RGB image into a signal. Furthermore, the benefit of combining image colour- and texture- related features through data fusion was also explored. The outcomes showed some overall (even if slight) improvement in model performances when combining the information from colourgrams and texturegrams, especially when fusing the colourgrams with texturegrams calculated using a distance of 15 pixels at the mid-level.

Even if the texturegrams were originally proposed to consider the texture properties of an image, the obtained results highlighted the possibility of considering this method as a hybrid tool to simultaneously explore both colour and texture properties. A possible explanation consists in the fact that texturegrams account for spatial differences between pixel intensities across the three colour channels. These findings need a further investigation, and future studies carried out on a systematically designed dataset are required to assess the potential of different texture features and their possible combination with colour-related ones.

The present Thesis also investigated the application of NIR spectroscopy coupled with chemometrics to monitor alcoholic and acetic surface fermentation in red grape musts. PLS models were calculated considering both the entire spectral range of a benchtop instrument (i.e., 12500-4000 cm^{-1}) and narrower regions (i.e., 10526-6060 cm^{-1} and 12500-9346 cm^{-1} ranges) to simulate and compare the potential performance of different commercial benchtop and portable instruments. Excellent results were achieved considering both the entire spectral range and the 10526-6060 cm^{-1} region, while the outcomes based on the 12500-9346 cm^{-1} region were satisfactory for a fast screening. This work confirmed the advantages and effectiveness of NIR spectroscopy for product monitoring in the field of vinegar production, boasting speed and efficiency while preserving the integrity of the sample, even when relying on cheaper and simpler setups.