Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 302 (2023) 123157

Contents lists available at ScienceDirect

Spectrochimica Acta Part A:
Molecular and Biomolecular Spectroscopy

journal homepage: www.journals.elsevier.com/spectrochimica-acta-part-a-molecular-and-biomolecular-spectroscopy

Multi-level color classification of post-consumer plastic packaging flakes by hyperspectral imaging for optimizing the recycling process

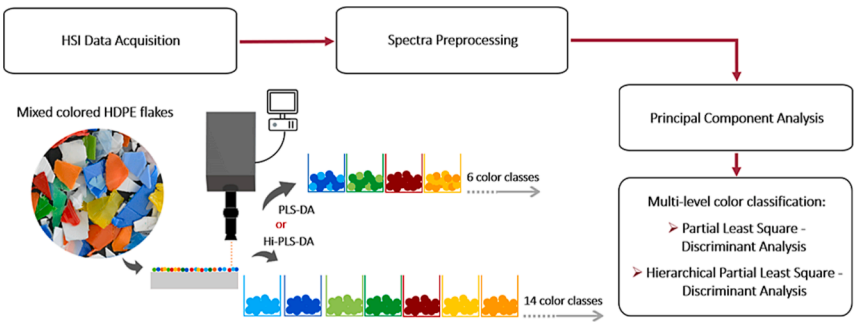
Paola Cucuzza, Silvia Serranti *, Giuseppe Capobianco, Giuseppe Bonifazi

Department of Chemical Engineering, Materials & Environment, Sapienza University of Rome, Rome, Italy

HIGHLIGHTS

- Hierarchical PLS-DA for color classification of plastic flakes.
- High-quality recycled HDPE by HSI working in the visible range.
- **Efficient** HSI and hierarchical machine learning based approach for plastic waste color sorting.
- Recycling-oriented classification by color of post-consumer HDPE flakes.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Hyperspectral imaging
HDPE
Hierarchical model
PLS-DA
Plastic waste
Recycling
Color sorting
Machine learning
Circular economy

ABSTRACT

In a circular economy perspective, the development of fast and efficient sensor-based recognition strategies of plastic waste, not only by polymer but also by color, plays a crucial role for the production of high quality secondary raw materials in recycling plants. In this work, mixed colored flakes of high-density polyethylene (HDPE) from packaging waste were simultaneously classified by hyperspectral imaging working in the visible range (400–750 nm), combined with machine learning. Two classification models were built and compared: (1) Partial Least Square-Discriminant Analysis (PLS-DA) for 6 HDPE macro-color classes identification (i.e., white, blue, green, red, orange and yellow) and (2) hierarchical PLS-DA for a more accurate discrimination of the different HDPE color tones, providing as output 14 color classes. The obtained classification results were excellent for both models, with values of *Recall*, *Specificity*, *Accuracy*, and *F-score* in prediction close to 1. The proposed methodological approach can be utilized as sensor-based sorting logic in plastic recycling plants, tuning the output based on the required needs of the recycling plant, allowing to obtain a high-quality recycled HDPE of different colors, optimizing the plastic recycling process, in agreement with the principles of circular economy.

1. Introduction

The upcycling of recycled plastics as secondary raw materials, reducing the use of virgin polymers, is still today an open challenge. The way of thinking about the life cycle of plastics, from product design to recycling, is currently focused on converting more and more waste into high-quality recycled products, increasing resources efficiency and reducing greenhouse gas emissions [1]. In this context, the development

* Corresponding author.
E-mail address: silvia.serranti@uniroma1.it (S. Serranti).

https://doi.org/10.1016/j.saa.2023.123157
Received 31 March 2023; Received in revised form 25 June 2023; Accepted 13 July 2023
Available online 14 July 2023

of effective, fast and robust sensor-based sorting procedures plays an important role for the improvement of the secondary raw material quality, contributing to reduce the presence of contaminants, in terms of quantity and types. In recycling plants, plastic waste can be separated not only by polymer but also by color, being characterized by different market values [2–7]. Sorting by polymer is essential to eliminate traces of impurities inside the recycled stream of a single polymer, i.e., other materials or other polymers, as they can create immiscibility and incompatibility problems, in terms of rheological and mechanical characteristics [8–9]. Equally important is sorting by color, as the use of colorants in plastic packaging is a determining factor strongly interfering with plastic recycling processes, influencing the final aesthetic properties of the recycled product and its possible applications [10]. More in detail, polymer identification is mainly achieved using optical sensing technologies working in the near or short-wave infrared (NIR-SWIR) range [11–16]. Sorting process by color can be carried out manually (laborious and difficult) or by automated selection through different optical systems equipped with RGB cameras and LED arrays [15–20–21]. Such devices allow to recognize plastic waste based on their color, depending on the recycling plant efficiency and the market needs. Hyperspectral imaging (HSI) is an optical-based approach widely used in many sectors thanks to the advantage of combining imaging with spectroscopy techniques [21–26]. In the plastic recycling industry, HSI can help to increase the speed and efficiency of automated plastic waste sorting devices, reducing contamination, improving secondary raw materials purity and profits [27–28]. In this context, HSI working in the visible (VIS) range has many advantages over other imaging methods for color analysis, such as those based on standard RGB cameras. These optical technologies (i.e., RGB cameras) are cost-effective and suitable for rapidly characterizing objects based on their shape and color. However, since only three visible bands are provided, their identification ability may not be accurate enough for specific industrial online-sorting and quality control processes, as it is the case of post-consumer plastic packaging flakes. In contrast, a hyperspectral camera is actually more expensive than an RGB camera but currently represents the most accurate imaging technique to adopt for recycling purposes, as it is able to acquire the spectrum, embedding materials characteristics, associated to each pixel of the image [29–30]. Therefore, with reference to colors, HSI provides a larger and more reliable number of information even when slight differences in tone occur [21,31–32]. In addition, spectroscopy combined with chemometric logics and machine learning techniques represents a valid and emerging tool useful to increase plastic recycling rates, automating the plastic sorting and increasing the quality of recycling processes [33]. However, the high discriminative capacity depends not only on the sensor but also on the use of processing approaches that must allow the management of a large data flow [34]. Currently, among the innovative methodologies, helpful to manage a large flow of highly variable data, ensemble methods are used, such as multi-level hierarchical modeling [35]. In statistics and machine learning, the ensemble methods use multiple learning algorithms to achieve better predictive performance than those possible with an individual learning algorithm. The fields of application of this methodological approach are numerous, showing considerable potential [36–38].

The largest destination market for plastics is the packaging sector, reaching 39.1% of European demand in 2021, in which the most demanded polymers are polyolefins, i.e., low-density polyethylene (LDPE), high-density polyethylene (HDPE) and polypropylene (PP), followed by polyethylene terephthalate (PET) [1]. At the present time, post-consumer PET bottles are the recycled packaging waste with the highest level of purity, thanks to efficient polymer and color sorting procedures, necessary to meet the beverage industrial needs and the ambitious European recycling targets [39–40]. Following the successful strategy developed for PET recycling, allowing to obtain a high-purity recycled product both in terms of polymer and color, this work is focused on end-of-life HDPE packaging recycling. Due to its characteristics and versatility, HDPE is one of the most utilized plastics, reaching an annual production of 6.3 million tons, corresponding to 12.6% of the total European plastics demand in 2021 [1]. Despite the widespread use of HDPE, the study of color characteristics to enhance the commercial quality of recycled HDPE is still not well discussed [41]. Thus, it is crucial to develop new and effective recycling strategies, to ensure successful recovery for specific applications, fully in line with the principles of circular economy.

For these reasons, this study was carried out to test the potential of the HSI technique in the VIS range combined with machine learning, using an ensemble method (i.e., hierarchical model based on Partial Least Square-Discriminant Analysis - PLS-DA) and an individual method (i.e., a single model based on PLS-DA), in order to build fast and efficient automated color classification strategies for the production of high-quality recycled HDPE flakes from packaging of various colors.

2. Materials and methods

2.1. The studied plastic samples

The studied plastic samples were provided by Ecosistem recycling plant, located at Lamezia Terme (Italy). They are representative of a recycled product, i.e., HDPE flakes from bottles, characterized by different types of color (Fig. 1a). A specific set of 180 flakes was selected as representative of the different types of colors and related shades constituting the plastic product (Fig. 1b). Six main classes of colors (white, blue, green, red, orange and yellow) were identified and further subdivided in 14 sub-classes based on color tones (i.e., two types of white, four types of blue, three of green, one of red, two of orange, and two of yellow). In order to build the classification models, plastic flakes were divided into a calibration and a validation dataset, constituted by about 70% and 30% of flakes, respectively (Fig. 1c).

2.2. Data acquisition

Hyperspectral images were acquired in the visible range (VIS: 400–750 nm) by the push-broom HSI sensor ImSpector V10E (Specim Ltd, Oulu, Finland), equipped with a VIS-NIR (400–1000 nm) Blaser A312F CCD camera, with 12 mm/s acquisition velocity and 850 frames. Samples were placed on a conveyor belt and scanned line by line simulating, at laboratory scale, the continuous transport occurring at recycling industrial plant scale.

2.3. Preprocessing strategies

All the acquired hyperspectral images were processed through the PLS_toolbox (version 8.8 Eigenvector Research, Inc.) working in Matlab environment (version R2020a, The MathWorks, Inc.).

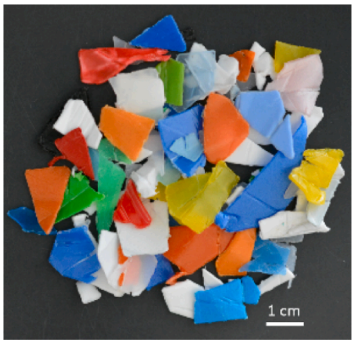
In order to optimize the two classification models, several pre-processing strategies were adopted to reduce noise (i.e., light scattering) and highlight the spectral differences between the studied HDPE color classes, following those widely adopted in the literature [42–45].

2.3.1. Standard Normal Variate (SNV)

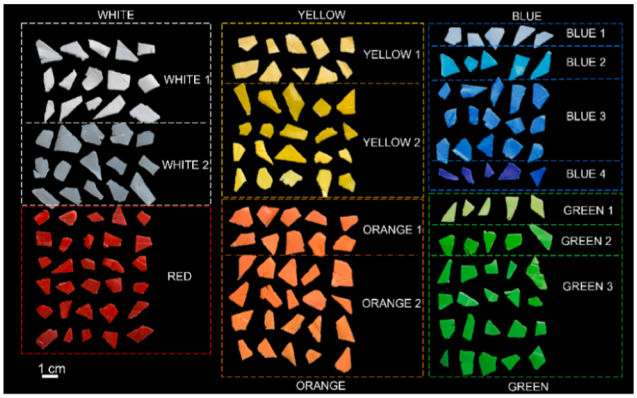
Standard Normal Variate (SNV) pre-processing was used to lessen the impact of light scattering on the collected hyperspectral pictures [46]. SNV converts each observed spectrum into a signal with zero mean and unitary variance, making it one of the most widely used techniques for scatter correction of NIR and SWIR data [47]. This algorithm also modifies the data such that signal to concentration relationships are often more linear.

2.3.2. Savitzky-Golay (SG) Derivative and Smoothing

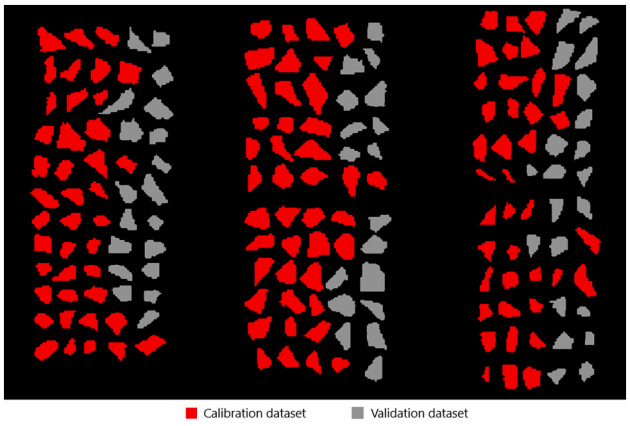
Spectral differences among the plastic colors have been emphasized using derivative approach. The Savitzky-Golay (SG) polynomial fitting algorithm was applied, due to its popularity and simplicity. Derivatives



(a)



(b)



(c)

Fig. 1. HDPE recycled flakes from plastic bottles (a); source image of the selected HDPE flakes of different colors (b); "image map" outlining flakes subdivision in calibration (red) and validation (grey) datasets (c).

have been employed in analytical spectroscopy thanks to the capability to reduce both additive and multiplicative influences in the spectra [47–48]. SG routine can be also utilized to reduce the noise in the spectra [49]. In this case routine can be used for smoothing/noise reduction, in order to avoid amplification of high-frequency noise during the derivation process, as it happens in the case of finite difference derivation [50]. In particular, the first derivative with different window points was used in this study.

2.3.3. Mean center (MC)

All data were preprocessed with a centering method. There are several additional preparation techniques, the best one relies usually on the type of data under investigation [51]. Mean centering (MC) was selected in this study, being appropriate for continuous data, such as HSI spectral data. In multivariate models, this technique has the effect of adding an adjustable intercept [52]. For example, mean centering both X and Y blocks in a regression model effectively allows for a non-zero intercept of the regression line [53].

2.4. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a chemometric technique for exploratory data analysis commonly and widely used for the treatment of hyperspectral data, allowing the visualization of the variability of a dataset with many variables (wavelengths) [49,52,54]. This visualization is achieved by transforming a set of covariates into a smaller set of principal components [55–57]. Considering a matrix X the PCA model is based on the decomposition of the dataset matrix (Equation (1)):

$$X = T\hat{A}\cdot P + E \tag{1}$$

where T was the score matrix, P was the loading matrix and E was the residual matrix. The utilized decomposition method was Singular Value Decomposition (SVD) [58].

The first two or three PCs, produced by PCA, are generally used to assess the shared characteristics between samples and their grouping; in fact, samples with comparable spectral signatures tend to cluster in the scores plot of the first two or three components. Thus, the reflectance value at each wavelength, or PC score at each PC (in the PC space), could be used to describe a reflectance spectrum. Samples of the same class and with comparable spectra will cluster in the same place of the PCA scores plot, while samples with different spectral characteristics will cluster in other areas of this space. In this work, PCA was used to evaluate the spectral variance of the data and set classes for the classification models.

2.5. Partial Least Squares-Discriminant Analysis (PLS-DA)

Partial Least Squares-Discriminant Analysis (PLS-DA) is a supervised technique combining Partial Least Squares (PLS) regression characteristics with classification techniques ability to Discriminant Analysis (DA) [59–61]. PLS-DA is based on the PLS regression algorithm, which looks for latent variables having a maximal correlation with the Y-variables (PLS1 when dealing with a single dependent Y variable and PLS2 when dealing with multiple dependent Y variables). The primary advantage of PLS-DA is that it models the relevant sources of data variability using the so-called Latent Variables (LVs), which are a linear combination of the original variables. As a result, it enables graphical visualization and understanding of the various data patterns and relationships by LV scores and loadings. Scores reflect the coordinates of samples in the LV projection hyperspace, whereas loadings are the coefficients of variables in the linear combinations that generate the LVs and are therefore interpretable as the influence of each variable on each LV. The chosen algorithm was the Nonlinear Iterative Partial Least Squares (NIPALS) [62]. In this case study, PLS-DA models were built starting from the class set in the PCA environment [49].

2.6. Color classification model building

Two different PLS-DA models were built, with different levels of plastic color discrimination:

- a 6 color classes PLS-DA model for identification of the main color types of HDPE flakes: white, blue, green, red, orange and yellow;
- a 14 color classes hierarchical PLS-DA model for identification of the different shades of the 6 main identified color classes: two types of white, four types of blue, three of green, one of red, two of orange, and two of yellow.

A hierarchical model is useful when the dataset is composed by many

classes and some of them have similar spectral signature. In fact, with this method is possible to preliminarily divide objects into subsets and then subdivide them into further subsets, until each subset contains a single object [63–64]. Following this multi-level strategy, the number of branches of each node is defined by the number of classes defined in the classification model [28,65]. In this case, the adopted method for each rule was based on PLS-DA.

The preprocessing algorithms selected for each rule of the 14 color classes Hi-PLS-DA model are shown in Table 1. The dendrogram of Hi-PLS-DA model, composed by 14 classes and the additional NC (Not Classified) branch, is reported in Fig. 2. The test sample will be assigned to the NC decision branch if its residuals (Q) or Hotelling's (T2) value (reduced) exceeds the assigned threshold value [65]. In order to evaluate the model complexity and to select the appropriate number of LVs, each PLS-DA model was cross validated using Venetian blinds method on calibration dataset, assuming a number of data splits equal to 10 with one sample per spectrum.

2.7. Classification performance metrics

In a pixel-based logic, the classification performances of PLS-DA and Hi-PLS-DA models were evaluated by the confusion matrix (where the diagonal and off-diagonal cells correspond to the number of correctly and incorrectly classified pixels, respectively), and the commonly used performance metrics, i.e., *Recall*, *Specificity* (Spec), *Accuracy* (Acc), and *F-score* defined by Eqs. (2), (3), (4), and (5). *Recall* (or *Sensitivity* in binary classification) describes the model capability to correctly identify positive labels (classes). *Specificity* shows the model ability to correctly reject samples belonging to all the other classes. These parameters can assume values between 0 and 1, the latter being the ideal value for a prediction model. Starting from the two previous mentioned parameters, *Accuracy* can be calculated, as the ratio of the well-recognized samples on the total number of all samples. At last, *F-score* measures the relation between data positive labels and those given by the predictor [66–67].

$$Recall = \frac{TruePositive}{(TruePositive + FalseNegative)} \tag{2}$$

$$Spec = \frac{TrueNegative}{(TrueNegative + FalsePositive)} \tag{3}$$

Table 1
Preprocessing algorithms and LVs selected for the rules of the Hi-PLS-DA model.

Rules	Applied preprocessing
Rule 1: White 1–2 vs others (LVs: 2)	SNV, Smoothing (15 points), MC
Rule 2: White 1 vs White 2 (LVs: 2)	SNV, 1st Derivative (order: 2, window: 33 points), MC
Rule 3: Blue 1–2-3–4 vs others (LVs: 2)	SNV, MC
Rule 4: Blue 1 vs Blue 2 vs Blue 3 vs Blue 4 (LVs: 2)	SNV, 1st Derivative (order: 2, window: 15 points), MC
Rule 5: Green 1–2-3 vs others (LVs: 3)	1st Derivative (order: 2, window: 15 points), MC
Rule 6: Green 1 vs Green 2 vs Green 3 (LVs: 2)	SNV, 1st Derivative (order: 2, window: 15 points), MC
Rule 7: Red vs others (LVs: 2)	1st Derivative (order: 2, window: 15 points), MC
Rule 8: Orange 1–2 vs Yellow 1–2 (LVs: 2)	1st Derivative (order: 2, window: 15 point), MC
Rule 9: Orange 1 vs Orange 2 (LVs: 2)	1st Derivative (order: 2, window: 15 point), MC
Rule 10: Yellow 1 vs Yellow 2 (LVs: 2)	SNV, MC

$$Acc = \frac{Correctlyidentifiedsamples}{Totalnumberofallsamples} \tag{4}$$

$$F - score = \frac{2TruePositive}{(2TruePositive + FalsePositive + FalseNegative)} \tag{5}$$

3 Results and Discussion

3.1. Spectral signatures of HDPE samples

The average reflectance spectra acquired in the VIS range and used for the calibration dataset of the 6 main classes of color (i.e., 950 spectra of White, 935 spectra of Blue, 960 spectra of Green, 927 spectra of Red, 960 spectra of Orange and 950 spectra of Yellow) and 14 sub-classes of color tones (i.e., 150 spectra of White 1, 152 spectra of White 2, 163 spectra of Blue 1, 152 spectra of Blue 2, 153 spectra of Blue 3, 147 spectra of Blue 4, 144 spectra of Green 1, 154 spectra of Green 2, 160 spectra of Green 3, 155 spectra of Red, 165 spectra of Orange 1, 169 spectra of Orange 2, 153 spectra of Yellow 1 and 152 spectra of Yellow 2) are shown in Fig. 3.

Information on the minimum and maximum reflectance values and on the inflection points (i.e., points of the curve where the sign of the curvature changes [68], useful to explain the color variability of the studied plastic flakes) are reported in the following.

3.2. Reflectance spectra of the 6 main color classes

As shown in Fig. 3a, a low inflection point around 430 nm and an overall flat trend from 400 to 450 nm are observed in the average reflectance spectrum of the White class. The average reflectance spectrum of the Blue class is characterized by two inflection points around 500 and 700 nm. The average reflectance spectrum of Green class is characterized by 3 inflection points mainly localized around 500, 550 and 620 nm, respectively. The average spectrum of the Red class is characterized by a continuous absorption up to 550 nm and an inflection point around 600 nm. Finally, the average spectra of the Orange and Yellow classes show a similar trend to the Red class spectrum, with evident inflection points around 500 and 650 nm, respectively.

3.3. Reflectance spectra of the 14 sub-classes based on color tones

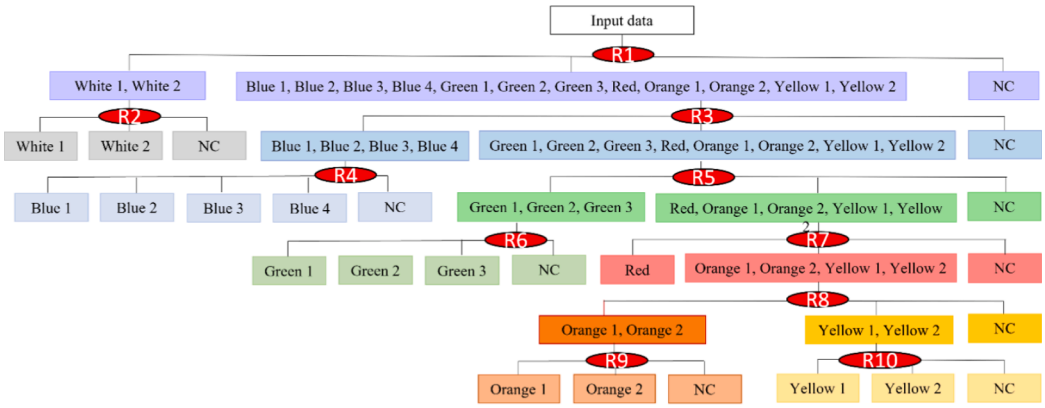
As shown in Fig. 3b, the average reflectance spectra of the white classes (i.e., White 1 and White 2) have an almost similar shape trend (i.e., mainly horizontal), but a clear distinction in the reflectance values (i.e., around 0.88 for White 1 and 0.3 for White 2). Since the White 2 class is slightly transparent, its reflectance spectra were also influenced by the black background, resulting in lower reflectance values than those of the White 1 class.

The average reflectance spectra of the blue classes (i.e., Blue 1, Blue 2, Blue 3, and Blue 4) are characterized by a similar shape trend. A significant curve variation of the Blue 1 class spectrum, with an inflection point at around 700 nm, is observed. Furthermore, a marked difference in reflectance values is shown by all blue classes spectra, with values increasing according to the sequence: Blue 1, Blue 2, Blue 3 and Blue 4.

In the average reflectance spectra of the green classes (i.e., Green1, Green 2, and Green 3), the Green 1 spectrum shows two inflection points localized around 500 and 580 nm; the Green 2 and Green 3 classes show similar fingerprints and spectral variations mainly localized around 480 nm.

Concerning the average spectrum of the Red class, its description is the same as previously reported, as there is a homogeneous tone of red color in the examined samples.

In the average spectra of the orange classes (Orange 1 and Orange 2) an almost similar trend is observed. It is characterized by a continuous



(NC: Not Classified)

Fig. 2. Dendrogram of Hi-PLS-DA model built for the classification of 14 HDPE color classes: White 1, White 2, Blue 1, Blue 2, Blue 3, Blue 4, Green 1, Green 2, Green 3, Red, Orange 1, Orange 2, Yellow 1 and Yellow 2.

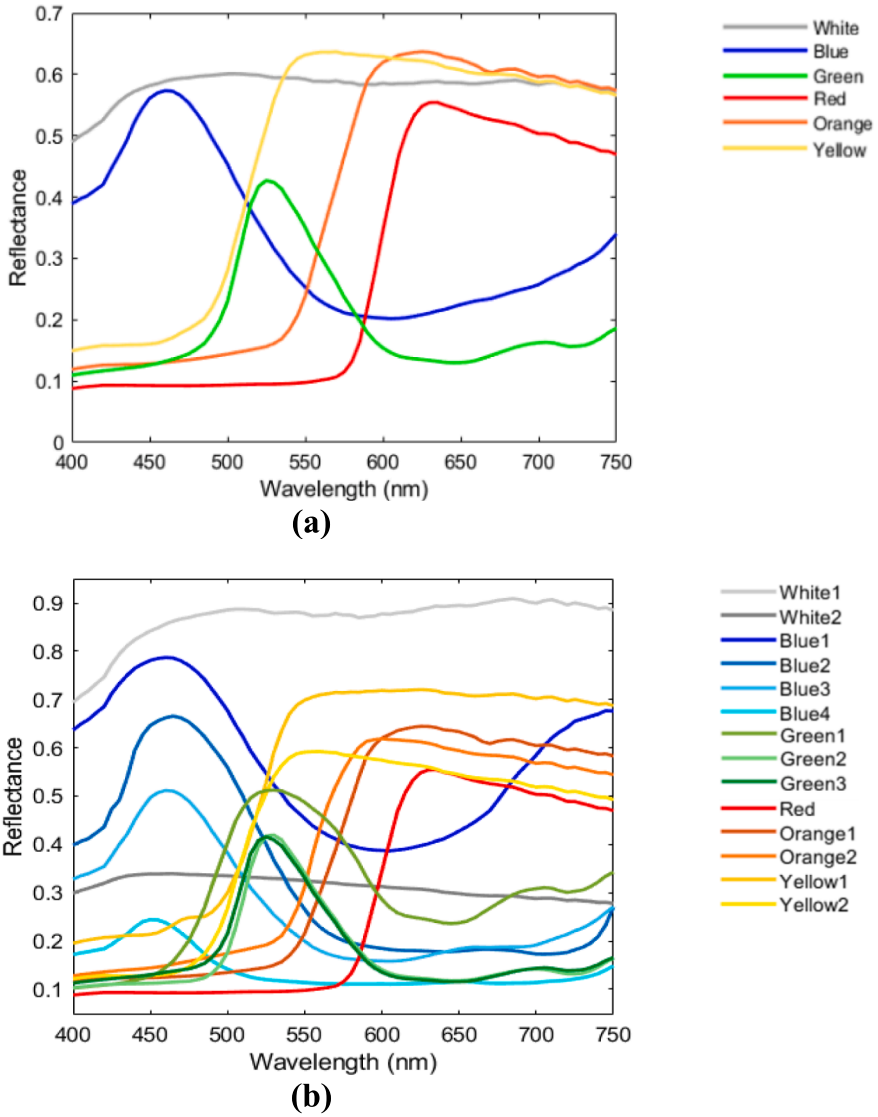


Fig. 3. Average raw reflectance spectra acquired in the VIS range (400-750 nm) of the 6 main color classes (a) and 14 color sub-classes (b) of studied HDPE flakes.

absorption up to 570 nm and two different inflection points at 500 nm and 570 nm for the Orange 2 and for the Orange 1, respectively.

A similar trend was observed in the average spectra of the yellow classes (Yellow 1 and Yellow 2) with variations mainly concentrated around 500 nm and in the reflectance values of spectral regions from 400 to 450 nm and from 550 to 750 nm.

3.4. Spectra preprocessing and Principal Component Analysis of 6 color classes

The preprocessed spectra and the corresponding PCA score and loadings plots are reported in Fig. 4. The preprocessed spectra, derived from the combination of SNV, Smoothing (15 points) and MC algorithms, are reported in Fig. 4a.

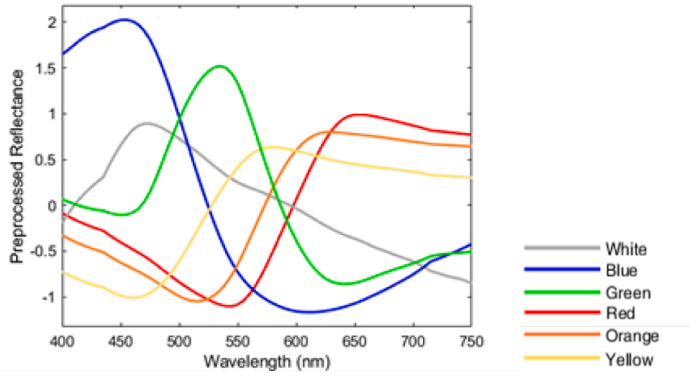
The PCA scores plot showed that most of the variance was captured by the first two PCs, where PC1 and PC2 explained 72.82% and 22.29% of the total variance, respectively. As shown in the PCA scores plot (Fig. 4b), the spectral data of the 6 color class samples are clustered into 6 distinct groups according to their spectral signatures, i.e., Blue, Green, Orange, Red, White, and Yellow. In detail, White class scores cloud shows high variability and is distributed in a fairly central area of the plot, especially in the first and second quadrants. Green class scores cloud is mainly localized in the first quadrant and a small part in the second. Yellow class scores cloud is identified in the second quadrant. The scores clouds of Orange and Red classes are located in the third quadrant and, finally, the Blue scores cloud is located in the fourth quadrant. The loadings plot of the first 2 PCs (Fig. 4c) highlighted the main spectral regions contributing to the identification of spectral differences between the examined classes. More in detail, the loadings of PC1 were mainly given, for positive values, by the wavelengths from 450 to 530 nm and for negative values by the wavelengths from 600 to 650 nm. The loadings of PC2 were mainly influenced for positive values by wavelengths around 550 nm and for negative values by wavelengths around 450 and 650 nm.

3.5. Spectra preprocessing and Principal Component Analysis of 14 color classes

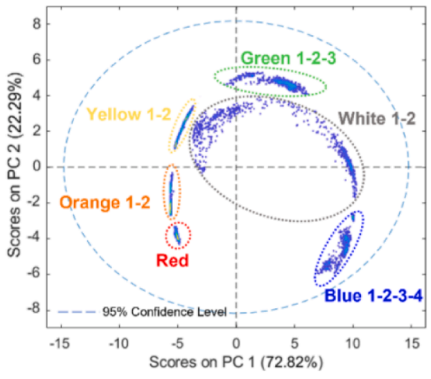
The results obtained for the different preprocessing combinations and the corresponding PCA, selected for the definition of the hierarchical PLS-DA rules, are presented and discussed in the following.

Concerning the 14 color classes Hi-PLS-DA model, PCA scores plot shown in Fig. 4 was useful to build Rule 1 of the hierarchical classifier, considering two clusters of pixels: a first group related to the white flakes and a second one to the other color classes. As a consequence, Rule 1 separates two classes: white flakes and colored flakes.

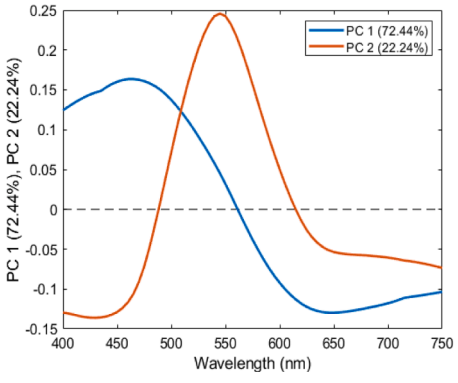
The preprocessed reflectance spectra, PCA scores plots and loadings plots related to the Rules from 2 to 10 are reported in Fig. 5. In order to distinguish White 1 from White 2 color classes, Rule 2 was created. The corresponding PCA scores plot (Fig. 5a) shows that the first two PCs are useful to separate the White classes, with PC1 and PC2 representing 71.40% and 16.87% of the total variance, respectively. Moreover, White 1 and White 2 scores (respectively second and fourth quadrants) appear quite distinct in two clouds. The loadings plot, shown in Fig. 5a, highlights as PC1 variance was mainly given for positive values by the wavelengths around 410 nm and for negative values by the wavelengths



(a)



(b)



(c)

Fig. 4. Average preprocessed spectra of investigated HDPE flakes resulting from the sequential application of SNV, Smoothing (15 points) and MC (a) algorithms. PC1-PC2 scores plot of the 6 HDPE color class spectral signatures (b), and PCs loadings plot of macro-color classes (c) for the calibration dataset.

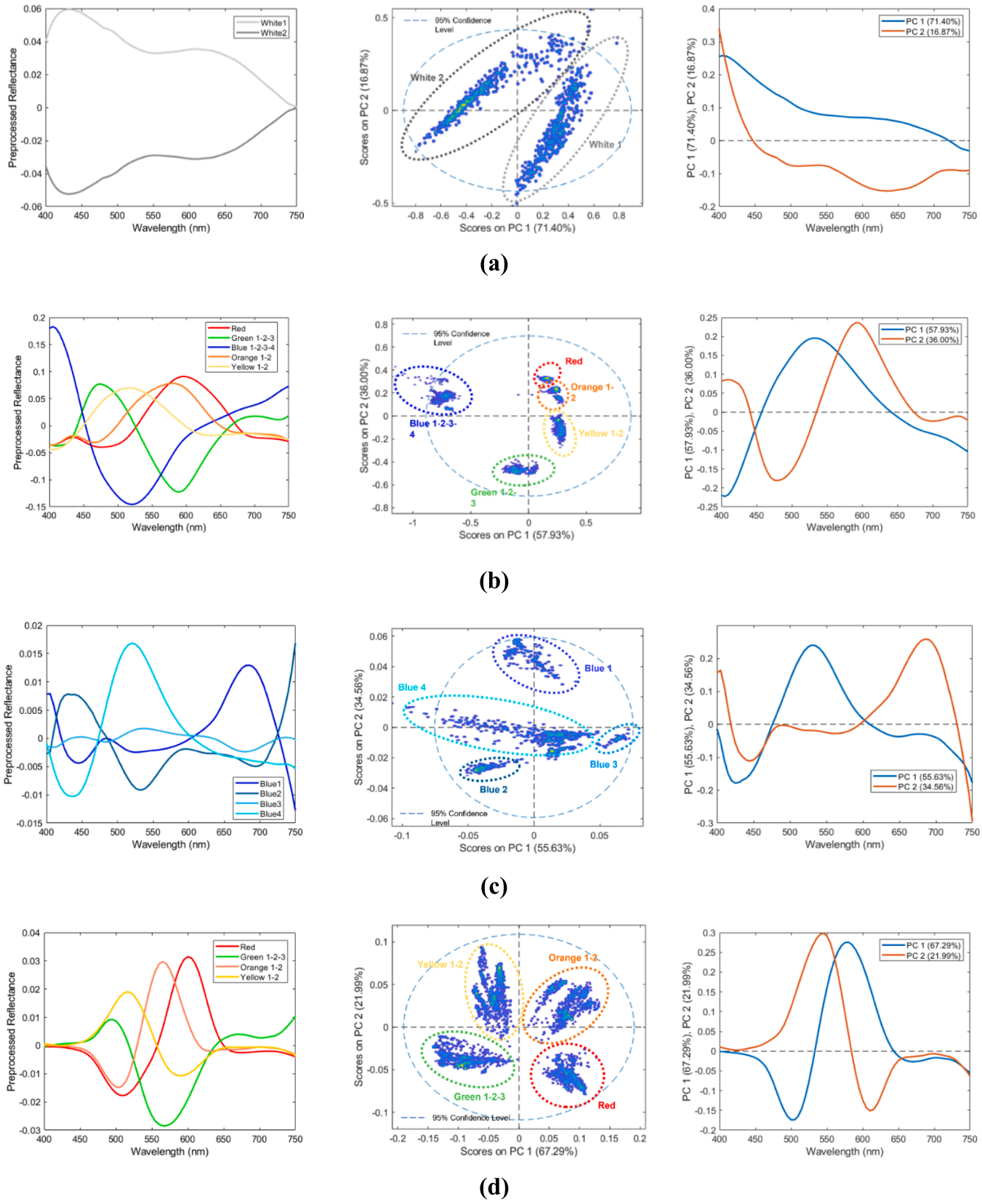
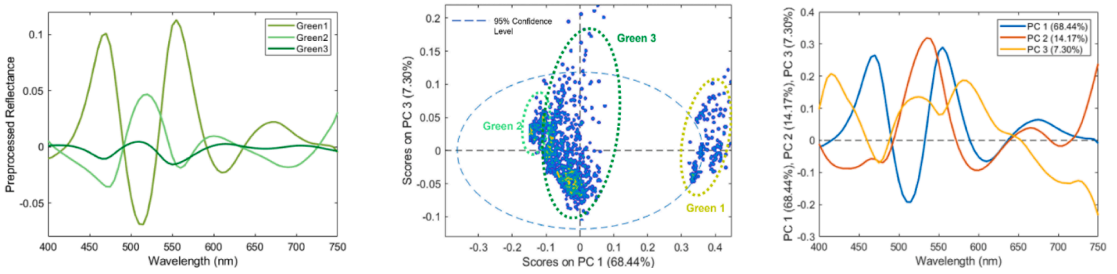
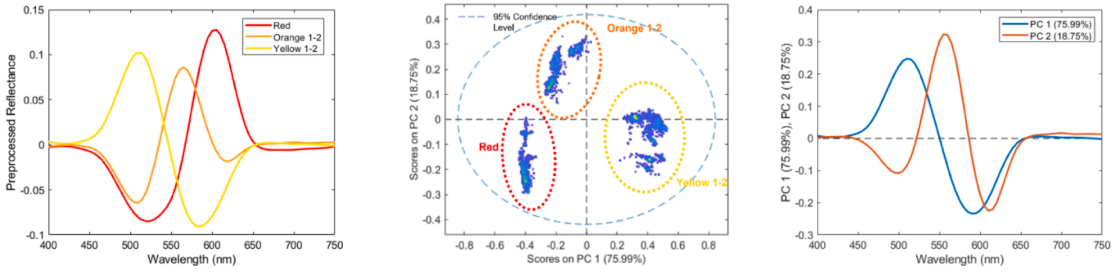


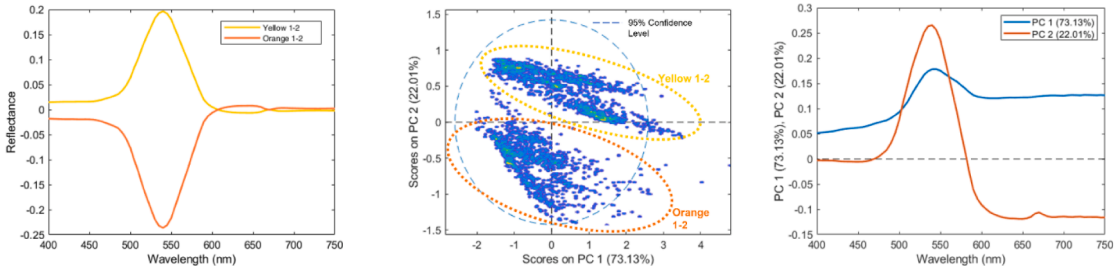
Fig. 5. Preprocessed average spectra, PCA scores plots, where the studied color classes are highlighted with colored circles, and PCs loadings plots for Rule 2 (a), Rule 3 (b), Rule 4 (c), Rule 5 (d), Rule 6 (e), Rule 7 (f), Rule 8 (g), Rule 9 (h) and Rule 10 (i).



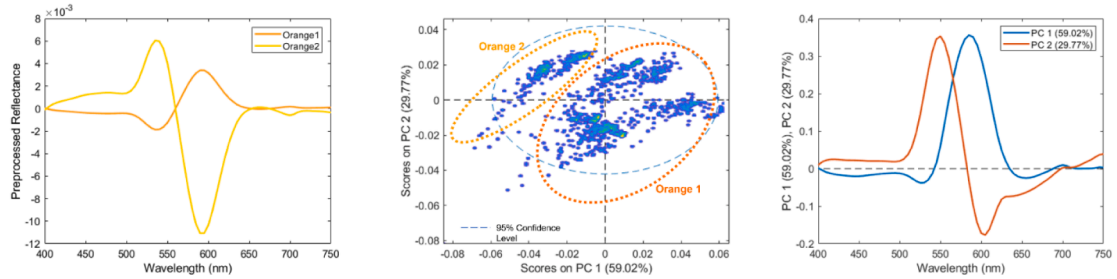
(e)



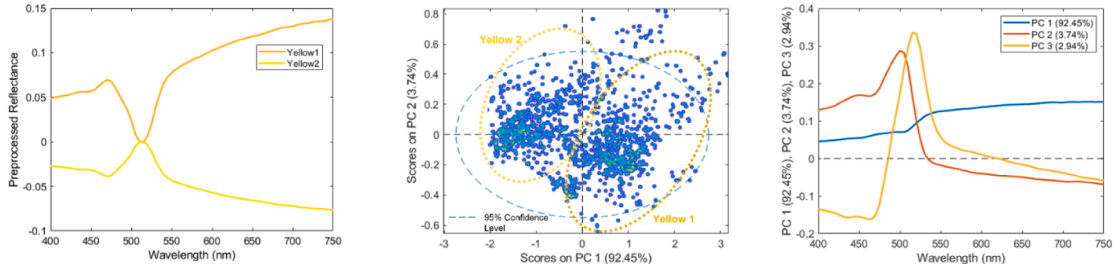
(f)



(g)



(h)



(i)

Fig. 5. (*continued*).

around 750 nm. PC2 was mostly marked for positive values by wavelengths around 400 nm and for negative values by wavelengths from 600 and 660 nm.

Rule 3 allows to distinguish the macro-group of blue (Blue 1–2-3–4) from the other classes (Green 1–2-3, Red, Orange 1–2 and Yellow 1–2). The PCA scores plot of the first two PCs is shown in Fig. 5b, where PC1 and PC2 explained 57.93% and 36.00% of the total variance, respectively. A significant separation of Blue macro-color class from the other classes in the fourth quadrant is obtained. The loadings plot, shown in Fig. 5b, highlights as PC1 variance was mainly given for positive values by the wavelengths around 530 nm and for negative values by the wavelengths around 410 and 750 nm. PC2 was mostly marked for positive values by wavelengths around 420 and 590 nm and for negative values by wavelengths around 475 and 700 nm.

The spectral separation between HDPE flakes, characterized by different tones of blue, is achieved by Rule 4. The scores plot of the first two PCs is shown in Fig. 5c, in which PC1 and PC2 explained 55.63% and 34.56% of the variance, respectively. As shown in PCA scores plot, a good separation between the four clusters is achieved. In more detail, Blue 1 is placed among the first and the second quadrants, Blue 2 and Blue 3 are located in the third and fourth quadrants, respectively, and Blue 4 is placed in the central zone of the scores plot with a main cluster in the fourth quadrant. The loadings plot, shown in Fig. 5c, highlights as PC1 variance was mainly given for positive values by the wavelengths around 525 nm and for negative values by the wavelengths around 425, 645 and 750 nm. PC2 was mostly marked for positive values by wavelengths around 405 and 690 nm and for negative values by wavelengths around 445 and 750 nm.

The separation between Green 1–2-3 macro-color class from other classes (Red, Orange 1–2 and Yellow 1–2) is obtained by Rule 5. PC1 and PC2 explained 67.29% and 21.99% of the total variance, respectively, as shown in Fig. 5d. PCA scores plot shows that a separation between the HDPE color classes involved in this rule can be easily performed. In fact, Orange1-2 cloud is located in the first quadrant, Yellow 1–2 cloud in the second quadrant, Green 1–2-3 cloud in the third quadrant and Red cloud in the fourth quadrant. The loadings plot, shown in Fig. 5d, highlights as PC1 variance was mainly given for positive values by the wavelengths around 575 nm and for negative values by the wavelengths around 500, 660 and 750 nm. PC2 was mostly marked for positive values by wavelengths around 550 nm and for negative values by wavelengths around 610 and 750 nm.

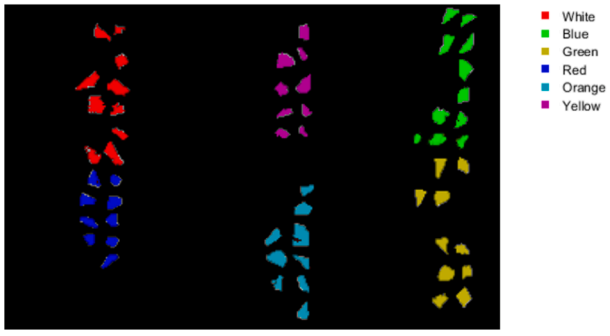
Rule 6 was created to classify the 3 different green color classes The first three PCs obtained by PCA are the most significant in order to perform the separation between green classes. PC1-PC3 scores plot is shown in Fig. 5e, explaining 68.44% and 7.30% of the total variance, respectively. Green 1 scores are well isolated in the first and fourth quadrants of the plot, Green 2 scores are located in the second quadrant, while Green 3 scores showed a more central and larger distribution between the second and third quadrants. The loadings plot, shown in Fig. 5e, highlights as PC1 variance was mainly given for positive values by the wavelengths around 460, 550, and 675 nm and for negative values by the wavelengths around 510 and 600 nm. PC2 was mostly marked for positive values by wavelengths around 545, 655, and 750 nm and for negative values by wavelengths around 425 and 600 nm. PC3 variance was mainly given for positive values by the wavelengths around 410, 525 and 585 nm and for negative values by the wavelengths around 475 and 750 nm.

Rule 7 was built in order to discriminate Red color class from other classes (Orange 1–2 and Yellow 1–2). PCA shows how the first two PCs are enough to obtain an efficient separation of the color classes, with PC1 and PC2 representing 75.99% and 18.75% of the captured variance, respectively. The scores plot (Fig. 5f) shows Red cloud located in the third quadrant, Orange 1–2 cloud in the second quadrant, Yellow 1–2 cloud in the first and fourth quadrant. The loadings plot, shown in Fig. 5f, highlights as PC1 variance was mainly given for positive values by the wavelengths around 515 nm and for negative values by the wavelengths around 590 nm. PC2 was mostly marked for positive values by wavelengths around 555 nm and for negative values by wavelengths around 500 and 610 nm.

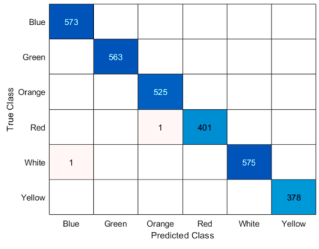
The separation between Orange 1–2 and Yellow 1–2 classes is achieved by Rule 8. The scores plot of the first two PCs is shown in Fig. 5g, where PC1 and PC2 explained 73.13% and 22.01% of the total variance, respectively. The PCA scores plot shows a considerable separation of Orange and Yellow macro-color classes, in which Yellow 1–2 scores are mainly located among the first and the second quadrants, while Orange 1–2 scores among the third and fourth quadrant. The loadings plot, shown in Fig. 5g, highlights as PC1 variance was mainly given for positive values by the wavelengths around 545 nm. PC2 was mostly marked for positive values by wavelengths around 540 nm and for negative values by wavelengths from 600 to 750 nm.

Rule 9 was built to obtain separation between Orange 1 and Orange 2 color classes. The scores plot of the first two PCs is shown in Fig. 5h, where PC1 and PC2 explained 59.02% and 29.77% of the captured variance, respectively. Orange 1 cloud shows a larger variability, and it is located in the central plot area among the first, third and fourth quadrants, while Orange 2 cloud is mostly located in the second quadrant. The loadings plot shown in Fig. 5h, highlights as PC1 variance was mainly given for positive values by the wavelengths around 580 nm and for negative values by the wavelengths around 525 and 650 nm. PC2 was mostly marked for positive values by wavelengths around 550 nm and for negative values by wavelengths around 600 nm.

Finally, Rule 10 is applied for the identification of Yellow 1 and Yellow 2 color classes. PC1-PC2 scores plot is shown in Fig. 5i, in which PC1 and PC2 describe 92.45%, 3.74% of the total captured variance, respectively. Yellow classes show a wide distribution in the scores plot, in particular Yellow 1 scores are mainly located among the first and fourth quadrants, while Yellow 2 scores are mainly concentrated between the second and third quadrants. The loadings plot, shown in Fig. 5i, highlights as PC1 variance was mainly given for positive values by the wavelengths from 550 to 750 nm. PC2 was mostly marked for positive values by wavelengths around 500 nm and for negative values by wavelengths around 750 nm. PC3 was mostly marked for positive



(a)



(b)

Fig. 6. HDPE macro-color classes prediction image (in false colors) of the validation dataset resulting from PLS-DA modelling (a); confusion matrix of true and predicted HDPE macro-color classes resulting from the application of the PLS-DA model to the validation dataset (number LVs: 3) (b).

values by wavelengths around 520 nm and for negative values by wavelengths around 460 and 750 nm.

3.6. Classification results of HDPE color typologies

3.6.1. 6 Color classes PLS-DA model

Fig. 6a shows the prediction results for the 6 HDPE macro-color classes of the validation dataset. An excellent classification is reached for all classes, except for some pixels of Red flakes incorrectly assigned to the Orange class and few incorrectly classified pixels related to "border effect", as also shown in the confusion matrix (Fig. 6b). However, these few misclassified pixels do not substantially affect the overall recognition of color classes. *Recall* and *Specificity* values in calibration (Cal), and cross-validation (CV) phases are shown in Table 2, while *Recall*, *Specificity, Accuracy* and *F-score* values in prediction (Pred) phase are shown in Table 3. The results indicate that the model has excellent classification performance metrics, ranging from 0.95 to 1.00 for all the measured parameters.

3.6.2. 14-classes hierarchical PLS-DA model

The prediction results obtained by the application of the 14 classes Hi-PLS-DA model to the HDPE flakes validation dataset are shown in Fig. 7a. An efficient classification is achieved for each color tone, with the exception of some pixels incorrectly assigned. These few pixels are mainly due to residual "border effect" or light reflectance from the preprocessing phase. However, even in this case, these incorrectly classified pixels do not damage the overall detection of color classes. The excellent classification results are confirmed by the confusion matrix (Fig. 7b), and by the model performance parameters values in calibration, cross-validation (Table 4) and in prediction phases (Table 5), showing values higher than 0.90 in all cases.

4. Economic impact considerations

Effective solutions are needed to improve the recovery and recycling of plastic waste enabling to obtain high-quality products, and to dramatically reduce the amount of waste ending up in landfills or incinerators. Better recycling practices would allow to decrease the demand for virgin plastics, with both industrial and environmental benefits, according to the principles of circular economy and AGENDA 2030 Sustainable Development Goals (SDGs), with particular reference to SDG12 on efficient use of resources. The proposed automated color sorting of HDPE packaging waste, based on HSI combined with machine learning approach, can support promising solutions to achieve highly efficient labor savings and an increase value of recycled plastics. Further developments and investments in advanced recycling technologies are fundamental to meet the high-quality requirements of industry. Moreover, in contrast to traditional RGB imaging, HSI allows to acquire and handle for each pixel a larger number of information (i.e., spectra related to the investigated wavelength range, instead of the common RGB components. Thanks to these spectral signatures, different materials, including plastics, can be more efficiently characterized. Therefore, HSI can sort materials quickly and effectively, reducing time and costs of sorting architectures and allowing a more accurate material identification, thus reducing contamination and increasing, at the same time, the purity of recycled materials and profit [28,30].

Table 2
Classification performances of the 6 color PLS-DA model in the calibration and cross-validation phases.

Class	Recall (Cal)	Spec (Cal)	Recall (CV)	Spec (CV)
White	1.00	0.99	1.00	0.99
Blue	1.00	1.00	1.00	1.00
Green	0.99	1.00	0.99	1.00
Red	0.99	0.99	0.99	0.99
Orange	0.99	0.95	0.99	0.95
Yellow	0.99	0.99	0.99	0.99

Table 3
Classification performances of the 6 color PLS-DA model in the prediction phase.

Class	Recall (Pred)	Spec (Pred)	Acc (Pred)	F-score (Pred)
White	1.00	0.99	0.99	0.99
Blue	0.99	1.00	0.99	1.00
Green	1.00	1.00	1.00	0.99
Red	0.99	1.00	0.99	0.99
Orange	0.98	0.98	0.98	0.99
Yellow	1.00	0.99	0.99	1.00

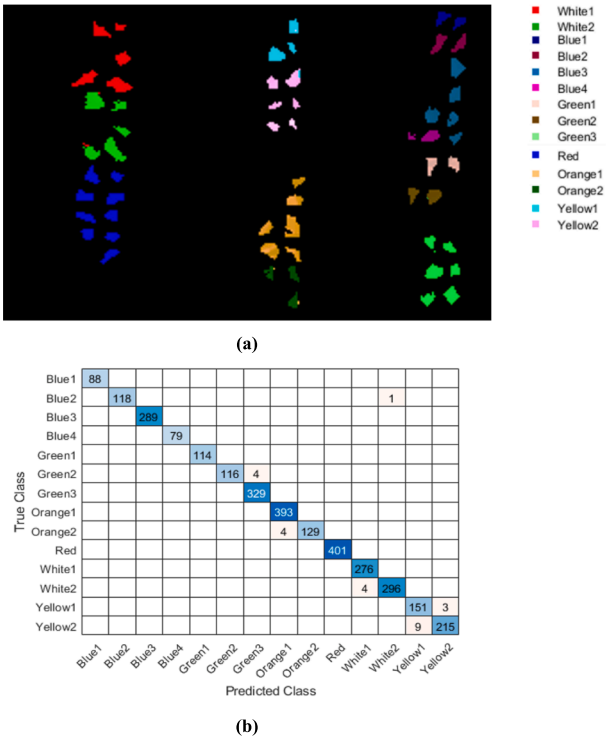


Fig. 7. HDPE 14 color classes prediction image (in false colors) of the validation dataset obtained by the application of the Hi-PLS-DA model (a); confusion matrix of true and predicted HDPE 14 color classes resulting from the application of the Hi-PLS-DA model to the validation dataset (b).

5. Conclusions

Two classification models were developed to obtain efficient plastic color sorting strategies using HSI in the VIS range combined with machine learning logics:

- PLS-DA model for the macro-colors identification;
- Hierarchical PLS-DA (Hi-PLS-DA) model able to detect different shades of the same color.

The results showed high prediction accuracy for both models, highlighting the possibility to apply the proposed methodological approaches for automated color recognition, to increase the quality and quantity of recycled plastic packaging, i.e., HDPE flakes from bottles.

The described approaches can represent attractive and flexible solutions for a sorting system that meets the challenges of a dynamic market scenario, such as plastic packaging recycling, according to a sustainable use of resources.

In conclusion, the hierarchical approach could be used to modulate the output results based on the needs of the plastic market (i.e., greater or lesser variety of colors) compared to a single classification model.

Table 4
Classification performances of the Hi-PLS-DA model in the calibration and cross-validation phases.

Rules	Classes	Recall (Cal)	Spec (Cal)	Recall (CV)	Spec (CV)
Rule 1	White 1–2	0.99	0.96	0.99	0.96
	Others	0.96	0.99	0.96	0.99
Rule2	White 1	1.00	0.98	1.00	0.98
	White 2	0.98	1.00	0.98	1.00
Rule 3	Blue 1–2-3–4	1.00	1.00	1.00	1.00
	Others	1.00	1.00	1.00	1.00
Rule 4	Blue 1	1.00	1.00	1.00	1.00
	Blue 2	1.00	1.00	1.00	1.00
	Blue 3	0.98	0.99	0.99	0.98
	Blue 4	1.00	1.00	1.00	1.00
Rule 5	Green 1–2-3	1.00	1.00	1.00	1.00
	Others	1.00	1.00	1.00	1.00
Rule 6	Green 1	1.00	1.00	1.00	1.00
	Green 2	0.98	0.99	0.99	0.98
	Green 3	0.98	0.99	0.99	0.98
Rule 7	Red	1.00	1.00	1.00	1.00
	Others	1.00	1.00	1.00	1.00
Rule 8	Orange 1–2	1.00	1.00	1.00	1.00
	Yellow 1–2	1.00	1.00	1.00	1.00
Rule 9	Orange 1	1.00	0.98	1.00	0.98
	Orange 2	0.98	1.00	0.98	1.00
Rule 10	Yellow 1	1.00	1.00	1.00	1.00
	Yellow 2	1.00	1.00	1.00	1.00

Table 5
Classification performances of the Hi-PLS-DA model in the prediction (Pred) phase.

Classes	Recall (Pred)	Spec (Pred)	Acc (Pred)	F-score (Pred)
White 1	1.00	1.00	1.00	1.00
White 2	0.99	1.00	1.00	0.99
Blue 1	1.00	1.00	1.00	1.00
Blue 2	0.97	1.00	0.98	1.00
Blue 3	1.00	1.00	1.00	1.00
Blue 4	1.00	1.00	1.00	1.00
Green 1	0.98	1.00	0.99	1.00
Green 2	0.98	1.00	0.99	0.98
Green 3	0.96	1.00	0.98	1.00
Red	1.00	1.00	1.00	1.00
Orange 1	0.97	1.00	1.00	1.00
Orange 2	1.00	1.00	0.98	0.98
Yellow 1	1.00	1.00	1.00	0.99
Yellow 2	1.00	1.00	1.00	0.98

Future studies may focus on combining the ensemble classification approach with multi-sensor data.

CRediT authorship contribution statement

Paola Cucuzza: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Silvia Serranti:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Writing – review & editing, Supervision. **Giuseppe Capobianco:** Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Giuseppe Bonifazi:** Methodology, Resources, Writing – review & editing, Supervision.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgments

The authors thank Ecosistem Srl (Lamezia Terme, Italy) for providing the investigated plastic samples.

References

[1] P. Europe, An analysis of European plastics production, demand, conversion and end-of-life management, Available online (2022). https://plasticseurope.org/knowledge-hub/plastics-the-facts-2022/.
[2] Y. Tachwali, Y. Al-Assaf, A.R. Al-Ali, Automatic multistage classification system for plastic bottles recycling, Resour. Conserv. Recycl. 52 (2) (2007) 266–285.
[3] T. Yoshioka, G. Grause, Recycling of Waste Plastics, in: Y. Tanaka, M. Norton, Y. Y. Li (Eds.), Topical Themes in Energy and Resources, Springer, Japan, 2015, pp. 195–214, https://doi.org/10.1007/978-4-431-55309-0_11.
[4] Y. Zheng, J. Bai, J. Xu, X. Li, Y. Zhang, A discrimination model in waste plastics sorting using NIR hyperspectral imaging system, Waste Manag. 72 (2017) 87–98, https://doi.org/10.1016/j.wasman.2017.10.015.
[5] G. Faraca, T. Astrup, Plastic waste from recycling centres: Characterisation and evaluation of plastic recyclability, Waste Manag. 95 (2019) 388–398, https://doi.org/10.1016/j.wasman.2019.06.038.
[6] Z. Wang, B. Peng, Y. Huang, G. Sun, Classification for plastic bottles recycling based on image recognition, Waste Manag. 88 (2019) 170–181, https://doi.org/10.1016/j.wasman.2019.03.032.
[7] D.J. Da Silva, H. Wiebeck, Current options for characterizing, sorting, and recycling polymeric waste, Prog. Rubber Plast. Recycl. Technol. 36 (4) (2020) 284–303, https://doi.org/10.1177/1477760620918603.
[8] K. Ragaert, L. Delva, K. Van Geem, Mechanical and chemical recycling of solid plastic waste, Waste Manag. 69 (2017) 24–58, https://doi.org/10.1016/j.wasman.2017.07.044.
[9] M. Larrain, S. Van Passel, G. Thomassen, B. Van Gorp, T.T. Nhu, S. Huysveld, P. Billen, Techno-economic assessment of mechanical recycling of challenging post-consumer plastic packaging waste, Resour. Conserv. Recycl. 170 (2021), 105607.
[10] B. Ruj, V. Pandey, P. Jash, V.K. Srivastava, Sorting of plastic waste for effective recycling, Journal of Applied Sciences and Engineering Research 4 (4) (2015) 564–571.
[11] S. Serranti, A. Gargiulo, G. Bonifazi, Characterization of post-consumer polyolefin wastes by hyperspectral imaging for quality control in recycling processes, Waste Manag. 31 (2011) 2217–2227, https://doi.org/10.1016/j.wasman.2011.06.007.
[12] S. Serranti, A. Gargiulo, G. Bonifazi, Classification of polyolefins from building and construction waste using NIR hyperspectral imaging system, Resour. Conserv. Recycl. 61 (2012) 52–58, https://doi.org/10.1016/j.resconrec.2012.01.007.
[13] G. Bonifazi, G. Capobianco, S. Serranti, A hierarchical classification approach for recognition of low-density (LDPE) and high-density polyethylene (HDPE) in mixed plastic waste based on short-wave infrared (SWIR) hyperspectral imaging, Spectrochim. Acta A Mol. Biomol. Spectrosc. 198 (2018) 115–122.
[14] X. Wu, J. Li, L. Yao, Z. Xu, Auto-sorting commonly recovered plastics from waste household appliances and electronics using near-infrared spectroscopy, J. Clean. Prod. 246 (2020), 118732.
[15] X. Xia, M. Wang, Y. Shi, Z. Huang, J. Liu, H. Men, H. Fang, Identification of white degradable and non-degradable plastics in food field: A dynamic residual network coupled with hyperspectral technology, Spectrochim. Acta A Mol. Biomol. Spectrosc. 296 (2023), 122686.
[16] A.A. Kamnev, Basics and applications of analytical molecular and biomolecular spectroscopy, Spectrochim. Acta A Mol. Biomol. Spectrosc. 204 (2018) 576–580.
[17] S.M. Safavi, H. Masoumi, S.S. Mirian, M. Tabrizchi, Sorting of polypropylene resins by color in MSW using visible reflectance spectroscopy, Waste Manag. 30 (11) (2010) 2216–2222, https://doi.org/10.1016/j.wasman.2010.06.023.
[18] P. Spiga, A. Bourely, Application of visible spectroscopy in waste sorting, In SPIE Optical Complex Systems OCS11, Proceedings 8172 (2011), 817212, https://doi.org/10.1117/12.899451.
[19] S. Brunner, P. Fomin, C. Kargel, Automated sorting of polymer flakes: Fluorescence labeling and development of a measurement system prototype, Waste Manag. 38 (2015) 49–60, https://doi.org/10.1016/j.wasman.2014.12.006.
[20] Chen, P., Gao, M., Huang, J., Yang, Y., Zeng, Y., 2018. High-Speed Color Sorting Algorithm Based on FPGA Implementation. IEEE 27th International Symposium on Industrial Electronics (ISIE), Cairns, QLD, Australia, 2018, pp. 235-239. https://doi: 10.1109/ISIE.2018.8433831.
[21] D.K. Chlebda, A. Rogulska, T. Łojewski, Assessment of hyperspectral imaging system for colour measurement, Spectrochim. Acta A Mol. Biomol. Spectrosc. 185 (2017) 55–62.
[22] H. Yin, B. Li, F. Zhang, C.T. Su, A.G. Ou-Yang, Detection of early bruises on loquat using hyperspectral imaging technology coupled with band ratio and improved Otsu method, Spectrochim. Acta A Mol. Biomol. Spectrosc. 283 (2022), 121775.
[23] X. Zhou, C. Zhao, J. Sun, K. Yao, M. Xu, J. Cheng, Nondestructive testing and visualization of compound heavy metals in lettuce leaves using fluorescence hyperspectral imaging, Molecular and Biomolecular Spectroscopy, Spectrochimica Acta Part A, 2023, p. 122337.

[24] G. Capobianco, L. Calienno, C. Pelosi, M. Scacchi, G. Bonifazi, G. Agresti, R. Picchio, U. Santamaria, S. Serranti, A.L. Monaco, Protective behaviour monitoring on wood photo-degradation by spectroscopic techniques coupled with chemometrics, Spectrochim. Acta A Mol. Biomol. Spectrosc. 172 (2017) 34–42.
[25] T. Palomar, C. Grazia, I.P. Cardoso, M. Vilarigues, C. Miliani, A. Romani, Analysis of chromophores in stained-glass windows using Visible Hyperspectral Imaging in-situ, Spectrochim. Acta A Mol. Biomol. Spectrosc. 223 (2019), 117378.
[26] C. Pelosi, G. Capobianco, G. Agresti, G. Bonifazi, F. Morresi, S. Rossi, U. Santamaria, S. Serranti, A methodological approach to study the stability of selected watercolours for painting reintegration, through reflectance spectrophotometry, Fourier transform infrared spectroscopy and hyperspectral imaging, Spectrochim. Acta A Mol. Biomol. Spectrosc. 198 (2018) 92–106.
[27] G. Bonifazi, G. Capobianco, P. Cucuzza, S. Serranti, A. Uzzo, Recycling-oriented characterization of PET waste stream by SWIR hyperspectral imaging and variable selection methods, Detritus 18 (2022) 42–49.
[28] Hyperspectral imaging vs RGB. Available online: https://www.specim.com/hyperspectral-technology-vs-rgb/ (accessed on 22 May 2023).
[29] G. Bonifazi, G. Capobianco, S. Serranti, Fast and effective classification of plastic waste by pushbroom hyperspectral sensor coupled with hierarchical modelling and variable selection, Resour. Conserv. Recycl. 197 (2023), 107068, https://doi.org/10.1016/j.resconrec.2023.107068.
[30] J. Zhang, R. Su, Q. Fu, W. Ren, F. Heide, Y. Nie, A survey on computational spectral reconstruction methods from RGB to hyperspectral imaging, Sci. Rep. 12 (1) (2022) 11905.
[31] S. Serranti, P. Cucuzza, G. Bonifazi, Hyperspectral imaging for VIS-SWIR classification of post-consumer plastic packaging products by polymer and color, In SPIE Future Sensing Technologies, Proceedings 11525 (2020) 1152510, https://doi.org/10.1117/12.2580504.
[32] Q. Ouyang, Y. Liu, Q. Chen, Z. Zhang, J. Zhao, Z. Guo, H. Gu, Intelligent evaluation of color sensory quality of black tea by visible-near infrared spectroscopy technology: A comparison of spectra and color data information, Spectrochim. Acta A Mol. Biomol. Spectrosc. 180 (2017) 91–96.
[33] E.R.K. Neo, Z. Yeo, J.S.C. Low, V. Goodship, K. Debattista, A review on chemometric techniques with infrared, Raman and laser-induced breakdown spectroscopy for sorting plastic waste in the recycling industry, Resour. Conserv. Recycl. 180 (2022), 106217.
[34] M. Bansal, I. Chana, S. Clarke, A survey on iot big data: current status, 13 v's challenges, and future directions, ACM Computing Surveys (CSUR) 53 (6) (2020) 1–59.
[35] O. Sagi, L. Rokach, Ensemble learning: A survey, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 8 (4) (2018) e1249.
[36] H. Chen, C. Tan, Z. Lin, Ensemble of extreme learning machines for multivariate calibration of near-infrared spectroscopy, Spectrochim. Acta A Mol. Biomol. Spectrosc. 229 (2020), 117982.
[37] S. Yu, J. Liu, Ensemble calibration model of near-infrared spectroscopy based on functional data analysis, Spectrochim. Acta A Mol. Biomol. Spectrosc. 280 (2022), 121569.
[38] J. Wang, J. Xu, C. Zhao, Y. Peng, H. Wang, An ensemble feature selection method for high-dimensional data based on sort aggregation, Systems Science & Control Engineering 7 (2) (2019) 32–39.
[39] Directive (EU) 2018/851 of the European Parliament and of the Council of 30 May 2018 amending Directive 2008/98/EC on waste. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018L0851&from=EN.
[40] Directive (EU) 2018/852 of the European Parliament and of the Council of 30 May 2018 amending Directive 94/62/EC on packaging and packaging waste. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018L0852&from=EN.
[41] C. Lubongo, P. Alexandridis, Assessment of performance and challenges in use of commercial automated sorting technology for plastic waste, Recycling 7 (2) (2022) 11.
[42] J.M. Amigo, J. Cruz, M. Bautista, S. Maspoch, J. Coello, M. Blanco, Study of pharmaceutical samples by NIR chemical-image and multivariate analysis, TrAC-Trends in Analytical Chemistry 27 (2008) 696–713, https://doi.org/10.1016/j.trac.2008.05.010.
[43] J.M. Amigo, Practical issues of hyperspectral imaging analysis of solid dosage forms, Anal. Bioanal. Chem. 398 (2010) 93–109, https://doi.org/10.1007/s00216-010-3828-z.
[44] H. Martens, T. Næs, Pretreatment and linearization, John Wiley & Sons Ltd, Multivariate Calibration, 2011.
[45] M. Vidal, J.M. Amigo, Preprocessing of hyperspectral images. Essential steps before image analysis, Chemom. Intel. Lab. Syst. 117 (2012) 138–148, https://doi.org/10.1016/j.chemolab.2012.05.009.
[46] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra, Appl. Spectrosc. 43 (1989) 772–777, https://doi.org/10.1366/0003702894202201.
[47] Å. Rinnan, F. Van den Berg, S.B. Engelsen, Review of the most common preprocessing techniques for near-infrared spectra, TrAC - Trends Anal, Chem 28 (2009) 1201–1222, https://doi.org/10.1016/j.trac.2009.07.007.
[48] A. Savitzky, M.J.E. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Anal. Chem. 36 (1964) 1627–1639, https://doi.org/10.1021/ac60214a047.
[49] Amigo, J. M., H. Babamoradi, H., Elcoroaristizabal, S., 2015. Hyperspectral image analysis. A tutorial, Anal. Chim. Acta, 896, 34–51. https://doi.org/10.1016/j.aca.2015.09.030.
[50] Å. Rinnan, L. Nørgaard, F. van den Berg, J. Thygesen, R. Bro, S.B. Engelsen, in: Data Pre-processing. D.-W, Elsevier/Academic Press, 2009, pp. 29–50.
[51] R. Bro, A.K. Smilde, Centering and scaling in component analysis, J. Chemom. 17 (1) (2003) 16–33.
[52] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (2014) 2812, https://doi.org/10.1039/c3ay41907j.
[53] Eigenvector, 2012. Advanced Preprocessing: Variable Centering. http://wiki.eigenvector.com/index.php?title=Advanced_Preprocessing:_Variable_Centering (accessed 01 June 2023).
[54] J.M. Amigo, I. Martì, A. Gowen, Hyperspectral Imaging and Chemometrics. A Perfect Combination for the Analysis of Food Structure, Composition and Quality, Data Handl. Sci. Technol. 28 (2013) 343–370, https://doi.org/10.1016/B978-0-444-59528-7.00009-0.
[55] P. Geladi, H. Isaksson, L. Lindqvist, S. Wold, K. Esbensen, Principal component analysis of multivariate images, Chemom. Intel. Lab. Syst. 5 (3) (1989) 209–220.
[56] K.R. Beebe, R.J. Pell, M.B. Seasholtz, Chemometrics: a practical guide, Vol. 4, Wiley, New York, 1998.
[57] I.T. Jolliffe, Principal component analysis for special types of data, Springer, New York, 2002, pp. 338–372.
[58] M.E. Wall, A. Rechtsteiner, L.M. Rocha, Singular value decomposition and principal component analysis, A practical approach to microarray data analysis (2003) 91–109.
[59] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models, PLS-DA. Analytical Method 5 (2013) 3790, https://doi.org/10.1039/C3AY40582F.
[60] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemom. 17 (3) (2003) 166–173, https://doi.org/10.1002/cem.785.
[61] A. Höskuldsson, PLS regression methods, J. Chemom. 2 (1988) 211–228, https://doi.org/10.1002/cem.1180020306.
[62] Wold, H., 1973. Nonlinear iterative partial least squares (NIPALS) modelling: some current developments. In Multivariate analysis–III, 383-407. Academic Press.
[63] K. Sendin, F. Marini, P.J. Williams, Hierarchical classification pathway for white maize, defect and foreign material classification using spectral imaging, Microchem. J. 162 (2021), 105824, https://doi.org/10.1016/j.microc.2020.105824.
[64] Y.B. Monakhova, M. Hohmann, N. Christoph, H. Wachter, D.N. Rutledge, Improved classification of fused data: Synergetic effect of partial least squares discriminant analysis (PLS-DA) and common components and specific weights analysis (CCSWA) combination as applied to tomato profiles (NMR, IR and IRMS), Chemom. Intel. Lab. Syst. 156 (2016) 1–6.
[65] PLS_Toolbox. Hierarchical Model Builder. https://wiki.eigenvector.com/index.php?title=Hierarchical_Model_Builder (accessed 01 June 2023).
[66] M.M. Oliveira, B.V. Cerqueira, S. Barbon Jr, D.F. Barbin, Classification of fermented cocoa beans (cut test) using computer vision, J. Food Compos. Anal. 97 (2021), 103771.
[67] T. Kato, S.M. Mastelini, G.F.C. Campos, A.P.A. da Costa Barbon, S.H. Prudencio, M. Shimokomaki, S. Barbon Jr, White striping degree assessment using computer vision system and consumer acceptance test, Asian Australas. J. Anim. Sci. 32 (7) (2019) 1015.
[68] M. Picollo, M. Aceto, T. Vitorino, UV-Vis spectroscopy. Physical sciences reviews 4 (4) (2018) 20180008.