Feasibility study for the surface prediction and mapping of phytonutrients in minimally processed rocket leaves (*Diplotaxis tenuifolia*) during storage by hyperspectral imaging

Muahmmad M.A. Chaudhry[a], Maria L. Amodio[a], José M. Amigo[b,c,*], Maria L.V. de Chiara[a], Farahmand Babellahi[a], Giancarlo Colelli[a]

[a] *Dip.to di Scienze Agrarie, degli Alimenti e dell'Ambiente, Universita di Foggia, Via Napoli, 25, 71122 Foggia, Italy*
[b] *IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain*
[c] *Department of Analytical Chemistry, University of the Basque Country UPV/EHU, P.O. Box 644, 48080 Bilbao, Basque Country, Spain*

ABSTRACT

A comprehensive study of the feasibility of hyperspectral imaging in visible (400–1000 nm) and near infrared (900–1700 nm) regions was investigated for prediction and concentration mapping of Vitamin C, ascorbic acid (AA), dehydroascorbic acid (DHAA) and phenols in wild rocket (*Diplotaxis tenuifolia*) over a storage span of 12 days at 5 °C. Partial least squares regression (PLSR) with different data pretreatments and wavelength selection resulted in satisfactory predictions for all parameters in the NIR range except DHAA. Prediction models were used for concentration mapping to follow changes over time. The prediction maps will be comprehensively study to assess the pixel to pixel variation within the rocket leaves. The PLSR models for Vitamin C, AA and phenols yielded an R^2 of 0.76, 0.73 and 0.78, respectively in external prediction with root mean square errors approximately equivalent to those of reference analysis. Conclusively, hyperspectral imaging, with the correct mapping approach, can be a useful tool for the prediction and mapping of phytonutrients in wild rocket (*Diplotaxis tenuifolia*) over time.

1. Introduction

Leafy vegetables have always served as a significant source of health promoting elements in human diet as they are an enormous reserve of active chemical compounds and are the cheapest and widely available source of fiber, proteins, vitamins, phenolic compounds antioxidants and minerals (Gibson et al., 2012). Along with aiding the consumers in meeting their optimum nutritional requirements they also act in the prevention of various morbid conditions (Lampe, 1999; Mann, 2001; He et al., 2006; Webb and Villamor, 2008). Therefore, the consumption of minimally processed ready-to-eat fruit and vegetable has significantly boosted in the last decades (Artés et al., 2009), since they are perceived as healthy, convenient, highly nutritive and appetizing (Oliveira et al., 2015; Ma et al., 2017).

In the Mediterranean countries, rocket leaves (*Diplotaxis tenuifolia*) with its pungent smell and strong flavour, lies among the most popular leafy vegetables, mostly consumed as stand-alone salads or as a part of mixed salad products. The rocket leaves are a rich source of phytonutrients such as fiber, Vitamin C, flavonoids and glucosinolates which are widely known for their positive impacts on human health (Cavaiuolo and Ferrante, 2014; Nurzyńska-Wierdak, 2015). The nutritional value of the wild rocket leaves and its degradation with the passage of the shelf life depends on the pre-harvest practices, postharvest handing, processing and storage conditions (Toivonen and Brummell, 2008; Cefola and Pace, 2015). After minimal processing operations (most commonly in this case, washing and drying), the rocket leaves are available packaged in plastic bags in the retail stores. Particularly, yellowing caused by chlorophyll degradation, wilting, and the production of off-odors are the main sources of deterioration for this product (Koukounaras et al., 2009; Løkke et al., 2012; Chaudhry et al., 2018).

The degradation process in fruits and vegetables also results in the degradation of the phytonutrients and many studies have taken into account the changes in the vitamin C, ascorbic acid (AA), phenols and anthocyanins (Amodio et al., 2015; Derossi et al., 2016). On the other hand, there are certain phytonutrients that increase with the passage pf storage time in rocket leaves such as glucosinolates, isothiocyanates, and amino acids (Bell et al., 2017). The kinetics of AA degradation is

effected by temperature, pH, enzymes, oxygen, metallic catalysers and light (Santos and Silva, 2008; Pérez-Balibrea et al., 2008). The effect of storage time and temperature on vitamin C degradation of rocket leaves has been reported by (Kim and Ishii, 2007; Spadafora et al., 2016; Mastrandrea et al., 2017). Kim and Ishii, 2007 observed that the vitamin C content was significantly affected during the storage of the rocket leaves at both 4 °C and 15 °C and also it was reported that the vitamin C content was higher in the leaves with and without roots in case of 4 °C as compared to 15 °C regardless of the storage time. In the same study it was also observed that the rocket leaves stored without roots showed better results during storage at 4 °C in terms of freshness and weight loss. Moreover, Spadafora et al., 2016 and Mastrandrea et al. (2017) also reported a rapid decrease in the vitamin C content of rocket leaves stored at high temperatures. Particularly, Mastrandrea et al. (2017) reported that the leaves stored at 0 °C both in air and in modified atmosphere packaging (MAP) did not show any changes in the AA content but those stored at 5 °C air portrayed slight decrease in the AA content with the passage of storage time while a rapid decrease was observed in the AA content degradation at 15 °C both in case of samples stored in air and in MAP. On the other side, other authors reported that AA remained unaltered in rocket leaves stored in controlled atmosphere (CA) at 4 °C, while it degraded in the leaves stored in air (Martínez-Sánchez et al., 2006b). The phenolic content decreased with the passage of storage time regardless of the storage atmosphere. Changes in AA content during storage of rocket leaves were also studied by Cavaiuolo et al. (2015) which observed a slight increase in the AA content in the initial days of storage.

All the above studies rely on chemical methods for the quantification of the phytonutrients; these methods are time consuming, require skilled personnel and are expensive to conduct. As an alternative, rapid and cheaper means of measuring nutritional quality may facilitate the access to this information also to processing companies and finally to the consumer. Hyperspectral imaging is a technique that integrates imaging and spectroscopy for the quantification and prediction of physical attributes and chemical compounds in food along with the mapping of their spatial distribution in the sample (Elmasry et al., 2012; Pu et al., 2015b; Huang et al., 2014). Every food product has, in fact, a specific spectral fingerprint, depending on the sample structure, the moisture content, the particle size, the temperature of the sample and most importantly of its chemical composition (Osorio et al., 2014). Most commonly, during the storage period, these spectral profiles are collected from the hyperspectral images which later integrated with multivariate tools can be used as a powerful tool for the estimation of the quality and shelf life of the food products during storage (Løkke et al., 2013a,b).

Particularly, hyperspectral imaging has been widely recognized for the prediction of various chemical constituents, contaminant, detection of defects, safety inspection, in fresh fruits and vegetables such as strawberries (Nagata et al., 2005; Tallada et al., 2006), apples (ElMasry et al., 2008), cucumbers (Liu et al., 2004), lychee fruits (Pu et al., 2015a), wheat grains; (Vigneau et al., 2011), and spinach leaves (Everard et al., 2014a,b). Many research works have concentrated on the non-destructive evaluation of leafy vegetables particularly spinaches (Diezma et al., 2013; Tewey et al., 2017; Lara et al., 2013; Zhang et al., 2017; Cho et al., 2017; Everard et al., 2014a,b; Lunadei et al., 2012; Yang et al., 2017), lettuce (Derossi et al., 2016; Xue and Yang, 2009), and rocket leaves (Toledo-Martín et al., 2017; Løkke et al., 2013a,b; Giovenzana et al., 2015; Chaudhry et al., 2018). Moreover, for rocket leaves (Løkke et al., 2013a,b) used CIELAB parameters obtained from multispectral imaging to predict color and texture. It was observed that a more reliable color evaluation was achieved using the CIELAB multispectral image data whereas the selective wavelengths in the NIR region were more reliable for prediction of textural variations. Kokalj et al. (2016) used FTIR spectroscopy for detecting the rocket contamination with the common groundsel leaves. Moreover, Chaudhry et al. (2018) used hyperspectral imaging in the Vis-NIR range applying

multivariate accelerated shelf life testing approach (MASLT) for the non-destructive shelf life estimation of stored rocket leaves. It was concluded by the study, that wavelength range between 550 and 700 nm significantly contributed towards the shelf life estimation based on appearance scores. NIRS was also employed by Villatoro-Pulido et al. (2012) for the prediction of mineral composition of the rocket leaves in the wavelength range of 400–2500 nm. Villatoro-Pulido et al. (2012) also used the Vis-NIR spectroscopy for the quantification of total phenolic content (TPC) and glucosinolates in the rocket leaves using MPLS regression with R^2 values ranging between 0.59 and 0.84 depicting reliable quantification results. Nevertheless, none of the studies have attempted to quantify and predict the phytonutrient changes over time in leafy vegetables particularly rocket leaves using hyperspectral imaging.

One of the major drawbacks of using vis-NIR in the quantitation of minor compounds in vegetables is the fact that the spectra is mostly determined by water. Consequently, the contributions of less concentrated compounds to the signal are normally very poor. This is normally increased when measuring hyperspectral images, where the effect of the surface roughness is added to the low signal effect. This manuscript studies, then, the feasibility of predicting changes in vitamin C content, AA, DHAA and phenols of the surface of rocket leaves over time by using hyperspectral images in the visible and NIR ranges using Partial Least Squares Regression (PLSR). The results will be mapped and a comprehensive study of the obtained results will be done. The pixel to pixel prediction maps will show the areas in which it is more prone to contain the compounds and how these areas evolve with time.

2. Materials and methods

2.1. Experimental design and spectral acquisition

Washed and dried conventionally grown rocket leaves (*Diplotaxis tenuifolia*) were harvested in the month of September in Puglia region of Italy and were received in the postharvest laboratory of University of Foggia, Italy. Representative samples were weighted, distributed into 100 g batches, packed into fifteen plastic clamshells and stored at 5 °C under humidified air flow. Fifteen replicates (each replicate comprising ~20 leaves) were acquired on each acquisition interval over a span of 12 days of storage. Hyperspectral image acquisition and reference analysis of the samples was done on 0, 2, 5, 7, 9 and 12 days of storage. Prior to the hyperspectral image acquisition, the samples were kept at room temperature for temperature regulation.

For the acquisition of the hyperspectral images, a hyperspectral line scan scanner (Version 1.4, DV srl, Padova, Italy) equipped with two spectrographs, one in the visible near infrared (Vis-NIR) region and the other in the near infrared region (NIR). The spatial resolution of the Vis-NIR spectrograph was 2000×1000 pixels with a spectral resolution of 5 nm over a wavelength range of 400–1000 nm, while the spatial resolution of NIR spectrograph was 623×320 pixels with a spectral resolution of 5 nm over a wavelength range of 900–1700 nm. The HSI cameras used were equipped with a CCD detector in case of the Vis-NIR, and a CMOS detector for NIR line scan camera with 50 frames per second with C-mount lenses on the cameras. Spatial resolution was 0.08 mm/pixels. The lighting system was comprised of a cooled halogen lamp with stabilized power source. The camera interface was GigE vision and a field of view (FOV) of 37°. The Full Width at Half Maximum of the camera was of 5 nm.

Self-developed MATLAB codes were used for image thresholding and the extraction of the average spectra of each replicate based on the best contrast between the object and the background followed by masking. A total of 90 spectra both in the NIR and Vis-NIR ranges were collected to formulate the dataset.

2.2. Chemical analysis

2.2.1. Total phenolic content evaluation

Total phenolic content was determined according to Singleton and Rossi (1965) with minor modifications. Three grams of leafy tissues representing the leaves of one replicate (one image) were homogenized in 2 mM sodium fluoride methanol:water solution (80:20) for 1 min and centrifuged at 5 °C and 12,000 rpm for 5 min. The total phenol content was expressed as mg of gallic acid equivalent (GAE) 100 g^{-1} fresh weight (f.w).

2.2.2. Vitamin C analysis

Three grams of fresh rocket tissues representing the leaves of one replicate (one image) were homogenized with 10 mL of MeOH/H_2O (5:95) plus citric acid (21 g L^{-1}) with EDTA (0.5 g L^{-1}). The homogenate was filtered through cheesecloth and a C18 Bakerbond SPE column (Waters, Milford, MA, USA). AA and DHAA contents were determined as described by Zapata and Dufour (1992), with some modifications. The HPLC analysis was achieved after derivatization of DHAA into the fluorophore 3-(1,2-dihydroxyethyl) furol [3,4-b]quinoxaline-1-one (DFQ), with 1,2-phenylenediamine dihydrochloride (OPDA). Samples of 20 μl were analyzed with an Agilent 1200 Series HPLC. The HPLC system consisted of a G1312A binary pump, a G1329A auto-sampler, a G1315B photodiode array detector from Agilent Technologies (Waldbronn, Germany). Separations of DFQ and AA were achieved on a Zorbax Eclipse XDB- C18 column (150 mm × 4.6 mm; 5 μm particle size; Agilent Technologies, Santa Clara, CA, USA). The mobile phase was MeOH/H_2O (5:95 v/v) containing 5 mM cetrimide and 50 mM potassium dihydrogen phosphate at pH 4.5. The flow rate was 1 mL min^{-1}. AA and DHAA contents were expressed as mg of ascorbic or dehydroascorbic acid 100 g^{-1} of f.w (mg 100 g^{-1}).

2.2.3. Partial least squares regression (PLSR)

Prediction models for the desired parameters were developed using the PLS algorithm in the PLS toolbox (Eigenvector Research Inc., version 7.2.5) working under MATLAB 2012b (version 8.0.0.783, MathWorks, MA, USA) as well as in HYPER-Tools (Version 2.0). The spectral dataset was divided into calibration set and validation set based on the 70/30 ratio with 70% of the samples in the calibration dataset and 30% of the samples reserved for external validation from the replicates of each acquisition interval. For the development of the PLSR calibration models random subset internal cross-validation was applied. The accuracy of the calibration models was accessed by visualizing the R^2 in calibration, R^2 cross-validation and the root mean square error for calibration (RMSEC) and cross-validation (RMSECV). As first approach all the wavelengths were used; then, after the formulation of the best calibration models, the loading plots were used for the selection of the most important variables in order to reduce the number of variables used to build the prediction models. All models were finally tested on the external data set to assess prediction performance. Moreover, for the best prediction models the slope, bias and the residual prediction deviation (RPD) were also assessed. RPD calculates the ratio of the standard deviation of the response variable to the root mean square error of prediction (RMSEP), and is considered as a parameter that gives an indication of robustness. On the other hand, bias shows the deviation of the predicted values form the observed which in ideal scenarios, should not exceed above or below 0.6 times the standard error of calibration (SEC). For a satisfactory model the minimum slope is considered to be 0.90 (Sánchez et al., 2011).

Mapping of the internal constituents was done by introducing a brand-new approach based on the selection of pixels in the image according to the calibration range. PLSR models were developed within the calibration range based on the spectra acquired from the average of the leaves in an image, standing to the consideration that the mean spectrum of that particular image corresponds to the averages of the vitamin C content of all the leaves in that image. Absorbance values of

Table 1
Range values and statistical distribution of Chemical Parameters.

Chemical Parameter	Min	Max	Mean	Standard Deviation
Vitamin C	15.80	123.33	66.39	22.94
Ascorbic Acid (AA)	7.97	109.38	55.83	20.53
Dehydroascorbic Acid (DHAA)	5.01	19.45	11.23	3.43
Phenols	83.53	200.92	136.06	29.79

the spectra of the leaf pixels within the calibration range and those outside the calibration range (which included both above and below the calibration range) were first plotted against the spectra of the calibration dataset to confirm their homogeneity to the spectra used to build the model. Nonetheless, it may be expected that PLSR calibration model developed with the mean spectra of all the leaves of each replicate when applied to a new image for pixel by pixel prediction, give some pixel values outside the calibration range. Therefore, since the prediction can be certain only if the obtained values are within the calibration range, the pixels with concentrations that fall out of the calibration range were excluded from the prediction map. Particularly, prediction values were accepted starting from a value corresponding to the minimum of the calibration range minus the RMSEC until the maximum of the calibration range plus the RMSEC.

3. Results and discussion

In Table 1 are shown the mean values and respective interval range of the main chemical constituents analyzed in this study. The variation in the minimum and the maximum range of values for the chemical parameters were determined over a storage span of 12 days at the intervals of 0, 2, 5, 7, 9 and 12 days.

Mastrandrea et al. (2017) also found similar values of vitamin C in fresh rocket leaves with a much higher amount of AA and a low amount of DHAA, hence showing that the rocket leaves in the start of the storage period possess a significantly higher amount of AA as compared to DHAA. Moreover, in a similar study regarding vitamin C content Martínez-Sánchez et al. (2006b) also showed that the predominant form of vitamin C at the beginning of the storage time was AA in rocket leaves as compared to DHAA. They also confirmed that a significant decrease in the AA content was recorded after six days of storage ultimately decreasing the vitamin C content of the rocket leaves stored both in air and in controlled atmosphere.

Fig. 1 shows the spectra obtained in the measurements. Fig. 1a and b, represent the pre-processed (1st derivative) Vis-NIR and NIR spectra, respectively.

The Vis-NIR peaks correspond to the color related properties of rocket leaves and the leaf reflectance as affected by chlorophyll *a*, chlorophyll *b* and β-carotene, in the region of 400–700 nm (Chaudhry et al., 2018; Mishra et al., 2017). On the other hand, in case of NIR spectra, the reflectance signal is dominated by the leaf water content. Peaks observed between 900 and 1000 nm and 1400–1500 nm are located in the third overtone region and the beginning of the first overtone region, respectively, both predominantly affected by water absorption (Workman, 2003; Sasic and Ozaki, 2011). In these regions most of the peaks identified by Yang and Irudayaraj (2002) can be found for vitamin C in powdered mixtures and solutions (1000, 1210, 1360, 1457, 1579 and 1651 nm); the same author also reported one peak at 840 nm related to vitamin C. Therefore, in this regard, it can be inferred that only one peak for vitamin C can be observed in the Vis-NIR region while all the other peaks are located in the NIR region from 900 to 1700 nm. Moreover, the region between 1300 and 1700 nm held great significance for the other compounds analyzed in this study, since the correlations between the spectra and the corresponding concentration in this wavelength region were higher than in the Vis-NIR region.
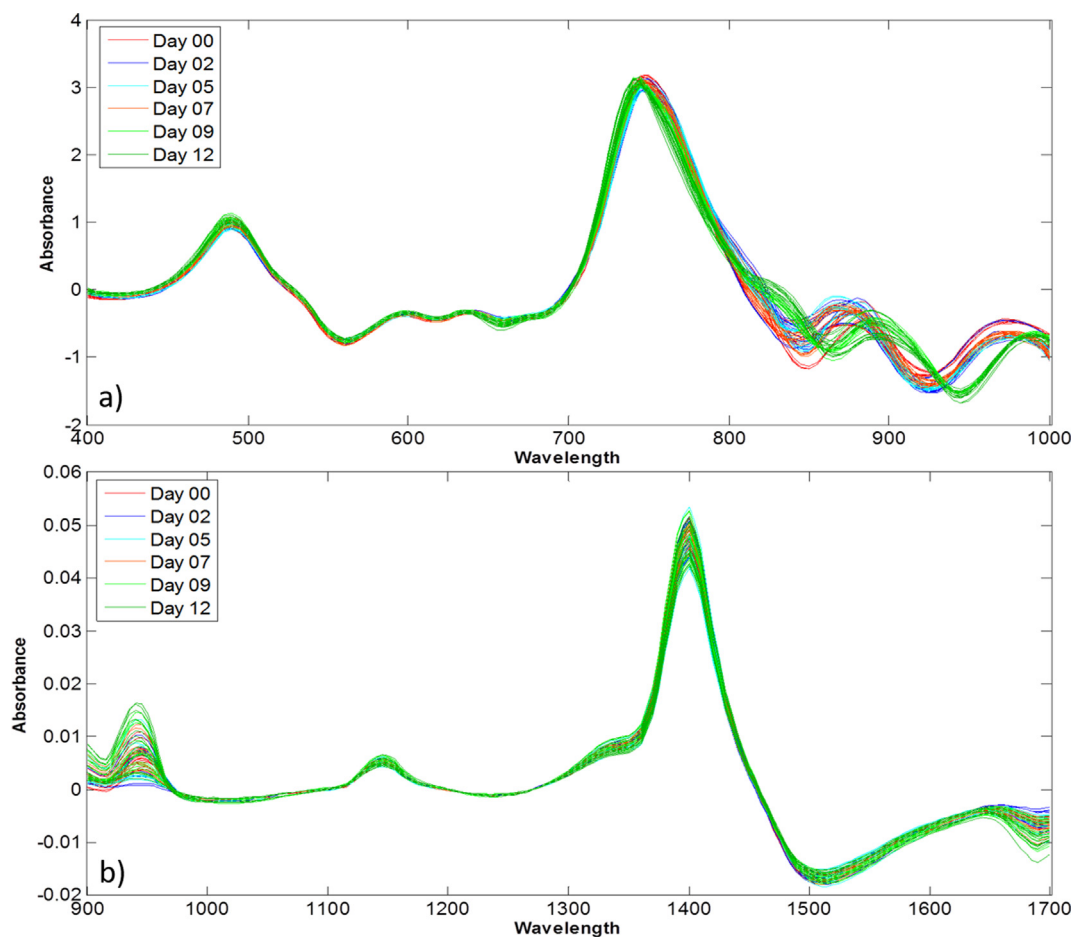
Fig. 1. a) 1st Derivative of Vis-NIR spectra b) 1st Derivative of NIR spectra Day 0 (red), Day 2 (blue), Day 5 (cyan), Day 7 (pink), Day 9 (light green) and Day 12 (green).

The PLSR models yielded reliable results for the Vitamin C content, AA, and phenols in NIR range while results obtained for DHAA were not satisfying. After the development of these models, the most significant variables were selected based on the loading weights for each parameter and simplified PLSR models were developed.

Table 2 shows the calibration results for the parameters in the NIR range in terms of R^2, root mean square error of calibration (RMSEC) and random block cross-validation. Individual calibration models were developed for each parameter. Different preprocessing techniques were attempted including mean centering, derivatization, SNV, MSC and their combinations. The best models obtained resulted from the

combination of 1st derivative followed by data mean centering in the NIR range which was also reported by Pissard et al., 2013 while measuring similar parameters in apples.

In case of vitamin C, the calibration dataset comprised of 70 samples and the external validation was done with 12 samples. Figure SM 1a (see supplementary material (SM) section) shows the PLSR calibration results for the vitamin C over a storage period of 12 days. The calibration model developed with a total of 161 variables in the NIR range yielded R^2_{cal} of 0.80 and R^2_{cv} of 0.71 with the RMSEC and RMSECV of 10.129 and 12.184 mg 100 g^{-1}f.w, respectively which was very similar to the laboratory error (9.179 mg 100 g^{-1}f.w) hence confirming the reliability of the calibration model. Optimal wavelength selection not only simplified the calibration model but also resulted in enhancing the performance in cross-validation with R^2_{cv} of 0.74 and RMSECV of 11.727 mg 100 g^{-1}f.w. It can be observed that the wavelength regions from 900 to 1000 nm, 1300–1500 nm and 1650–1700 nm held the most significant weight in the model for quantification of vitamin C and therefore, the variable reduction resulted in substantial improvement in the model performance.

In this model for the prediction of vitamin C in the NIR region a total of 55 variables highlighted in green in the loadings plot were utilized ranging from 900 to 1000 nm, 1315–1435 nm and 1650–1700 nm (see figure SM 1b in SM section). External validation allowed to obtain R^2_{pred} of 0.76 with a RMSEP of 10.905 mg 100 g^{-1}f.w. The RPD for the model was calculated to be 2.1, the cross-validation bias was 0.13 and the calibration slope of the model was 0.76. The best prediction results are summed in Table 3 for all the parameters.

Fig. 2 shows five random leaves chosen from the replicates of each storage time in order to visualize the changes in Vitamin C over storage

Table 2
Calibration statistics for the PLSR modelling of the internal constituents in fresh cut rocket leaves (Dev = derivative, MC = mean centering).

NIR range (900–1700 nm)							
Parameter	Pretreatment	No. of variables	LVs	R^2_{cal}	RMSEC	R^2_{cv}	RMSECV
Vitamin C	1st Dev + MC	161	8	0.80	10.129	0.71	12.184
	1st Dev + MC	75	6	0.80	10.263	0.73	11.903
	1st Dev + MC	**55**	**5**	**0.80**	**10.149**	**0.74**	**11.727**
AA	1st Dev + MC	161	6	0.82	8.301	0.77	9.615
	1st Dev + MC	68	6	0.82	8.312	0.77	9.636
	1st Dev + MC	**39**	**6**	**0.81**	**8.377**	**0.77**	**9.438**
DHAA	2nd Dev + MC	161	2	0.17	2.978	0.10	3.119
	1st Dev + MC	93	3	0.22	2.887	0.13	3.060
Phenols	1st Dev + MC	**161**	**5**	**0.80**	**12.909**	**0.75**	**14.488**
	1st Dev + MC	80	7	0.79	13.261	0.72	15.344

Table 3
Prediction statistics for the PLSR modelling of fresh cut rocket leaves.

| Parameter	Preprocessing	Wavelength Range (nm)	LVs	R^2_{cal}	RMSEC	R^2_{cv}	RMSECV	R^2_{pred}	RMSEP
Vitamin C	1st Dev + MC	900–1000 1300–1500 1650–1700	5	0.80	10.149 mg/100 g	0.74	11.727	0.76	10.905
AA	1st Dev + MC	900–1000 1295–1480 1655–1700	6	0.81	8.377 mg/100 g	0.77	9.438	0.73	10.249
Phenols	1st Dev + MC	900–1700	5	0.80	12.909 mg gallic acid/100 g	0.75	14.488	0.78	13.816

period, where each row represents leaves from each acquisition interval. As can be observed, prediction values within the calibration range ranged between 15.99 and 133.33 mg/100 g f.w. Each leaf in this case is accompanied by two numbers the first one being the percentage of pixels predicted within the calibration range and the second representing the average vitamin C concentration over these pixels. In this case, the pixels above and below the calibration range were removed. It is quite normal to have pixels predicted below or above the calibration range when using hyperspectral images. The first reason for this fact is the development of the PLS model itself. It is normally developed by using the mean spectrum of each replicate, which might not contain all the expected variability of the individual leaf in the image. The second reason can be the textural features affecting each leaf. A leaf is far for being a homogeneous surface; it has different internal structure as well as a rough surface. Martínez-Sánchez et al. (2006a) studied the sensory quality of stored rocket leaves and reported that the variations in the texture of the rocket leaves lead to the quality losses. The texture is usually related to the freshness perception of the rocket leaves but in this case, it also affects the overall spectral signals at different acquisition intervals over the storage period. For this purpose, the spectra of the pixels of each leaf in the prediction image were plotted against the spectra in the calibration dataset (not shown) and were found to have similar absorbance values for a majority of pixels but it was also observed that some of the pixels of the leaves in the prediction image had spectra that differed substantially from the spectra of the calibration dataset (most probably due to scattering effect).

Fig. 3a, b and c shows the pixels predicted within the calibration range, percentage of pixels within calibration range and standard deviation per leaf w.r.t days of storage, vitamin C on an average over the six image acquisition intervals, respectively. As can be observed, with the passage of storage time the pixels within the calibration range are decreasing as represented in Fig. 3b. It can be hypothesized that this may be a result of changes in the leaf structure or more simply of not having these values in the calibration model due to the fact that average values were higher. Fig. 3c represents the variation in the average concentration of vitamin C over the pixels within the calibration range during the storage period for the leaves of each acquisition interval. This trend of decrease in vitamin C content was also reported by Mastrandrea et al. (2017) also for rocket leaves stored in air at 5 °C.

In Fig. 4 the histograms represent the number of pixels within the calibration range associated to a certain amount of vitamin C content in each leaf in the concentration map. On the x-axis is shown the vitamin C concentration in mg of vitamin C $100\ g^{-1}$f.w. and the y-axis represents the number of pixels representing a certain quantity of vitamin C. It can be observed that at the start of the storage period the number of pixels representing higher values of vitamin C are higher and at the end of the storage period the number of the pixels representing vitamin C are lower because of the fact that vitamin C deteriorated with time and also because a larger number of pixels at this point were located below the calibration range.

Calibration models were also developed for AA with 70 samples in the calibration set and 12 in case of the external validation as well. The model performance was more reliable in the NIR range with R^2_{cal} of 0.81 and R^2_{cv} of 0.77 with RMSEC of 8.377 mg AA 100^{-1} g f.w and RMSECV
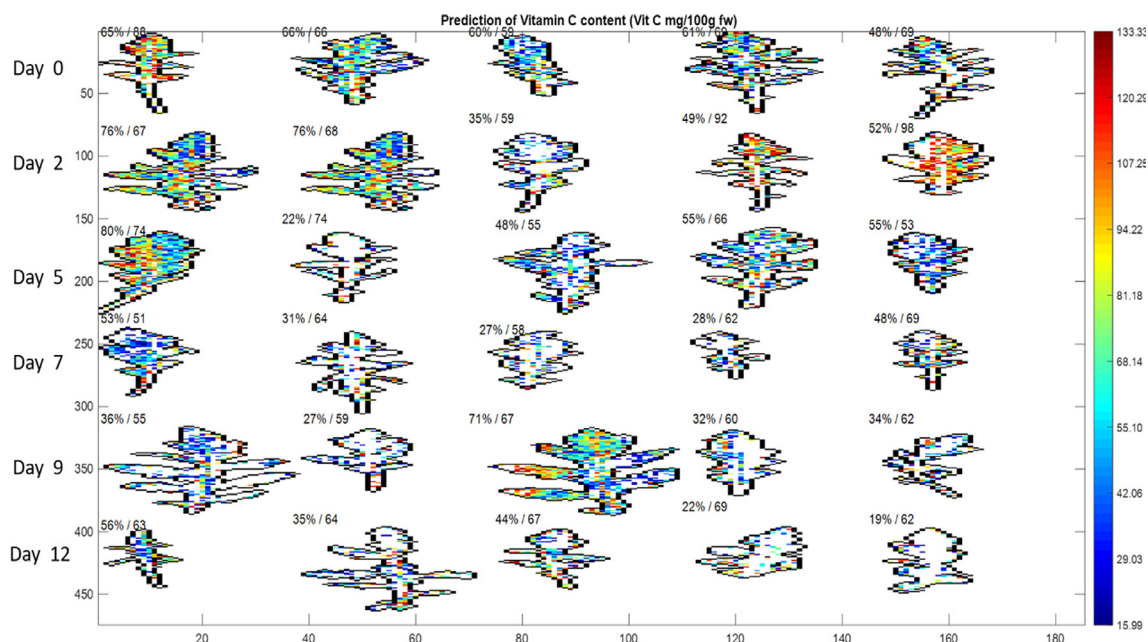
Fig. 2. PLS prediction map for Vitamin C after removal of pixels above and below the calibration range where each row corresponds to Day 0, Day 2, Day 5, Day 7, Day 9 and Day 12. For each leaf is indicated the percentage of pixels predicted within the calibration range and the average vitamin C concentration over these pixels.
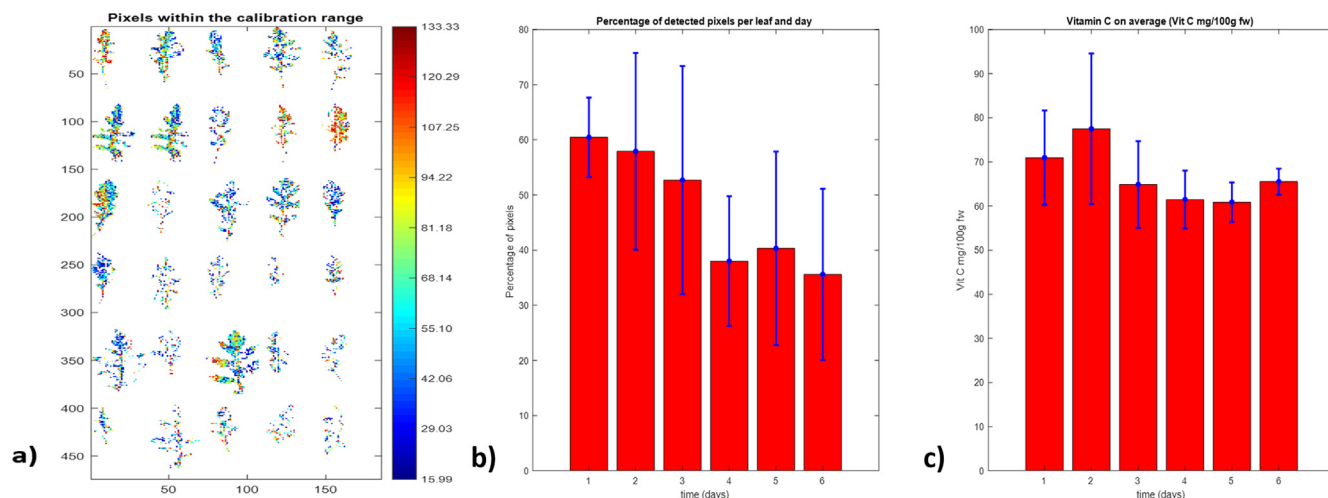
Fig. 3. a) Pixels within calibration range, b) percentage of pixels within calibration range and standard deviation per leaf w.r.t days of storage, c) Vitamin C on average (time in days refers to the 6 acquisition intervals i.e. Days 0, 2, 5, 7, 9, 12).

of 9.438 mg AA 100^{-1} g f.w with 3 LVs explaining a total co-variance of 96.91% in the data (see Figure SM 2a SM section). The loading scores carrying the maximum weight in the model corresponded to wavelength ranges of 900–1000 nm, 1300–1500 nm and 1650–1700 nm. After the development of the calibration model in the wavelength range of 900–1000 nm, 1295–1480 nm and 1655–1700 nm (shown in green highlighted regions in figure SM 2b SM section), its performance was evaluated with external validation yielding R^2_{pred} of 0.73 with a RMSEP of 10.249 mg $100 \ g^{-1}$ f.w. (Table 3) and with a cross-validation bias of −0.08. The slope and the residual prediction deviation (RPD) for the model were 0.81 and 2.1, respectively.

Fig. 5 shows the prediction map of the variation of ascorbic acid content in the rocket leaves during storage based on the pixel to pixel

information within the calibration range after the removal of all the pixels predicted outside the calibration range now represented by white pixels. Also, in this case it can be clearly observed that the percentage of the pixels within the calibration range decreased with the passage of time and at the same time a decrease in the average concentration of AA on those pixels was observed which was also a major cause for the overall vitamin C concentration decrease. This could be expected since vitamin C content is composed mainly of AA and a minor part of it is oxidized by DHAA (Martínez-Sánchez et al., 2006b).

Figure SM 3a (see SM section), depicts the map of the pixels lying within (7.18 to 120.73 mg AA $100 \ g^{-1}$ f.w.) the calibration range for AA. Moreover, the percentage of pixels within the calibration range for AA has been shown in figure SM 3b (see SM section) which also depicts
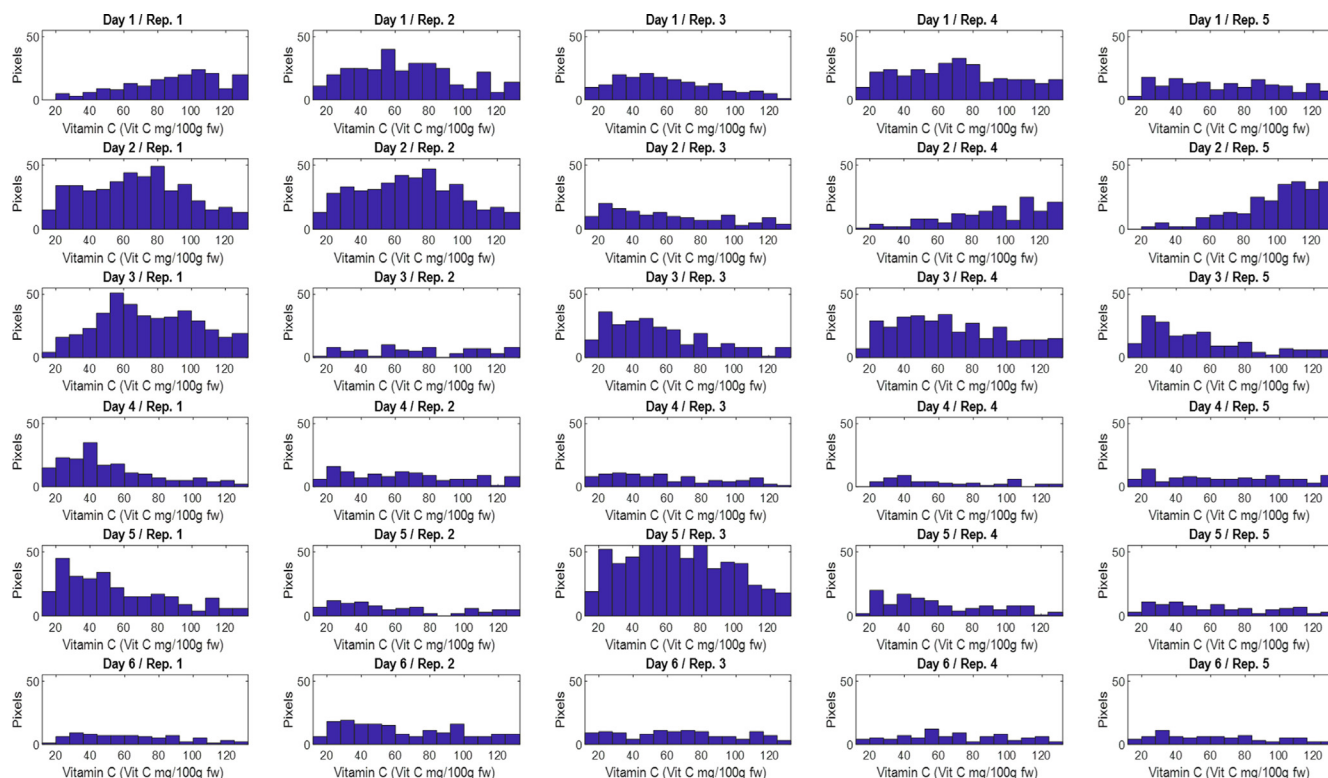

Fig 4. Histograms of vitamin C content of the leaves representing number of pixels with respect to vitamin C concentration (Day1, 2, 3, 4, 5, 6 refers to the 6 acquisition intervals i.e. Days 0, 2, 5, 7, 9, 12).
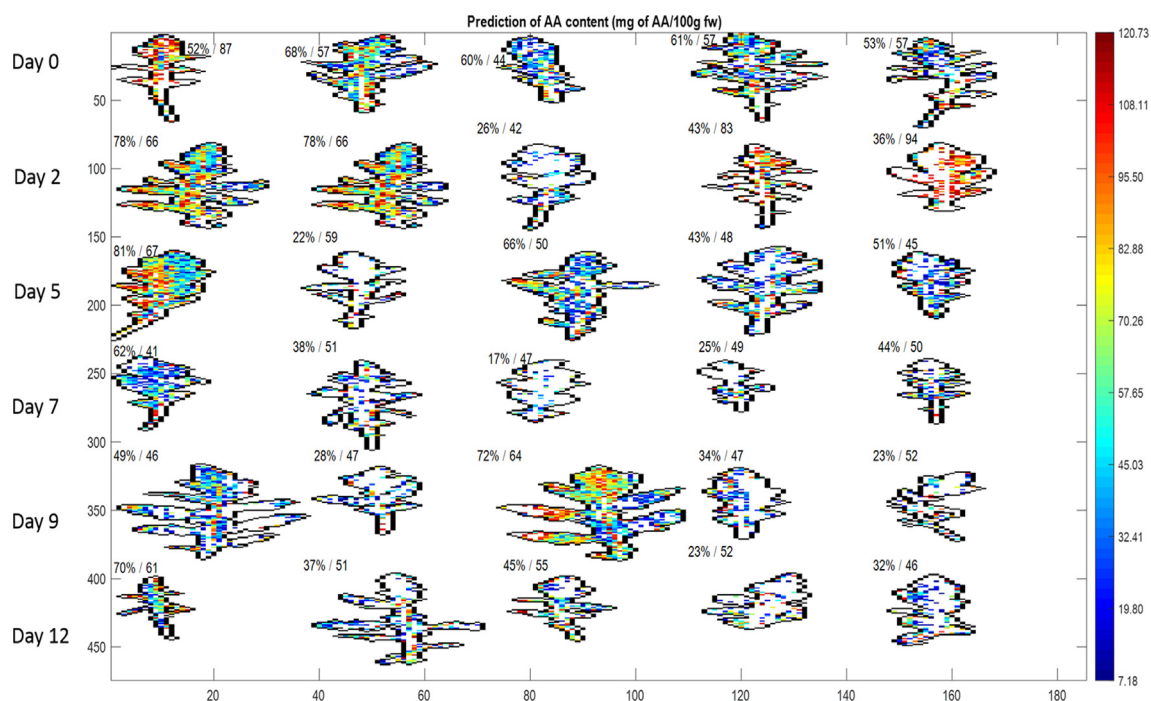
Fig. 5. PLS prediction map for AA with each row corresponding to Day 0, Day 2, Day 5, Day 7, Day 9 and Day 12 / For each leaf, first number: percent pixels predicted within calibration range; second number: average concentration of AA on predicted pixels.

a decrease in the average percentage of the pixels within the calibration range over time followed by a decrease in the average AA content represented by these pixels as depicted in figure SM 3c (see SM section). Furthermore, the number of pixels in each leaf associated with the certain AA content is shown in the histograms in figure SM 4 (see SM section) which also confirms the fact that a higher number of pixels were associated to a higher AA content in the initial days of storage while at the end of the storage period the number of pixels dropped due to the fact that a larger number of pixels were below the calibration range and also due to a drop in the AA content.

It can be observed that both average vitamin C and AA contents slightly increased in the initial days, especially at the second acquisition interval and then decreased with time; similar results were reported by Cavaiuolo et al. (2015) in which the AA content increased at initial days of storage and then decreased later during the entire storage span. These results are also compliant to those reported by Martínez-Sánchez et al. (2006a) and Amodio et al., 2015 which revealed that the rocket leaves when stored in the continuous air flow tend to decrease the AA and vitamin C content even at lower temperatures during storage time. Moreover, the calibration models for DHAA were also developed but the correlation between the spectra and the DHAA values was not encouraging which led to the conclusion that the calibration and prediction models for vitamin C possessed a slightly lower accuracy as compared to the PLSR models for AA because of the fact that vitamin C was determined using both AA and DHAA.

For phenolic content the PLSR models were developed yielding an R_{cal}^2 of 0.80 with RMSEC of 12.909 mg of gallic acid 100 g^{-1} and R_{cv}^2 of 0.75 with RMSECV 14.488 mg of gallic acid 100 g^{-1} (see figure SM 5a in SM section) with all 161 variables (see figure SM 5b in SM section) without any reduction in model variables. The reliability of the calibration model was accessed with external validation yielding R_{pred}^2 of 0.78 and RMSEP of 13.816 mg of gallic acid 100 g^{-1}. Figure SM 6 (see SM section) shows the changes in the phenolic content of the rocket leaves over the 12-day storage period within the calibration range (77.37 to 229.89 mg of gallic acid 100 g^{-1}). Moreover, in case of phenols also the percentage of pixels detected within the calibration range decreased with the passage of storage time and so did the average

phenolic content. In case of the best model for phenol prediction, the bias, slope and RPD were 2.27 $\times$ 10^{-13}, 0.70 and 2.2, respectively.

The map for the pixels within the calibration range (77.37 to 229.89 mg of gallic acid 100 g^{-1}) for the phenolic content is represented by figure SM 7a (see SM section). Additionally, the percentage of the pixels within the calibration range for the phenols is shown in figure SM 7b (see SM section) which decreases over the storage period followed by a decrease in the average phenol content depicted by figure SM 7c (see SM section). Furthermore, figure SM 8 (see SM section) depicts the number of pixels in each leaf representing the concentration of predicted phenols by each pixel in the map.

This approach can be useful for the processors to take logistic decisions based not only on the visual scores but also based on the nutritional values which would help better labelling the packaging in order to enhance consumer interest and trust.

4. Conclusions

Hyperspectral images in the NIR region, together with multivariate data analysis, has proven to be potentially useful in the mapping of phytonutrients in fresh-cut rocket leaves. Bearing in mind the difficulty of quantifying minor compounds in a signal mostly composed by the water signal, this manuscript has shown how, under the correct premises of spectra homogeneity and calibration range, acceptable results can be obtained. It was shown that the number of pixels detected within the calibration range decreased with the passage of storage time with a simultaneous increase in the pixels below the calibration range hence reducing the capability of the PLSR model to predict the corresponding phytonutrient. In this way, hyperspectral images revealed that the central part of the leaves lose vitamin C content faster as compared to the leaf edges or in other words the vitamin C starts degrading from the center of the leaf. Furthermore, the PLSR models in this case were very sensitive to the number of samples and their performance can be further enhanced with increasing the number of the samples in the calibration sets. Conclusively, NIR combined with hyperspectral imaging can be a very informative tool for studying the changes in the phytonutrient content during storage.

CRediT authorship contribution statement

Muahmmad M.A. Chaudhry: Conceptualization, Methodology, Software, Validation, Investigation, Writing - original draft, Writing - review & editing. **Maria L. Amodio:** Conceptualization, Methodology, Validation, Investigation, Writing - review & editing, Supervision, Project administration, Funding acquisition. **José M. Amigo Rubio:** Methodology, Software, Investigation, Writing - review & editing, Supervision. **Maria L.V. de Chiara:** Investigation, Data curation. **Farahmand Babellahi:** Investigation, Data curation. **Giancarlo Colelli:** Conceptualization, Methodology, Validation, Investigation, Writing - review & editing, Supervision, Project administration, Funding acquisition.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2020.105575.

References

Amodio, M.L., Derossi, A., Mastrandrea, L., Colelli, G., 2015. A study of the estimated shelf life of fresh rocket using a non-linear model. J. Food Eng. 150, 19–28. https://doi.org/10.1016/J.JFOODENG.2014.10.030.

Artés, F., Gómez, P., Aguayo, E., Escalona, V., Artés-Hernández, F., 2009. Sustainable sanitation techniques for keeping quality and safety of fresh-cut plant commodities. Postharvest Biol. Technol. 51, 287–296. https://doi.org/10.1016/J.POSTHARVBIO.2008.10.003.

Bell, L., Yahya, H.N., Oloyede, O.O., Methven, L., Wagstaff, C., 2017. Changes in rocket salad phytochemicals within the commercial supply chain: glucosinolates, iso-thiocyanates, amino acids and bacterial load increase significantly after processing. Food Chem. 221, 521–534. https://doi.org/10.1016/j.foodchem.2016.11.154.

Cavaiuolo, M., Cocetta, G., Bulgari, R., Spinardi, A., Ferrante, A., 2015. Identification of innovative potential quality markers in rocket and melon fresh-cut produce. Food Chem. 188, 225–233. https://doi.org/10.1016/J.FOODCHEM.2015.04.143.

Cavaiuolo, M., Ferrante, A., 2014. Nitrates and glucosinolates as strong determinants of the nutritional quality in rocket leafy salads. Nutrients. https://doi.org/10.3390/nu6041519.

Cefola, M., Pace, B., 2015. Application of oxalic acid to preserve the overall quality of rocket and baby spinach leaves during storage. J. Food Process. Preserv. 39, 2523–2532. https://doi.org/10.1111/jfpp.12502.

Chaudhry, M.M.A., Amodio, M.L., Babellahi, F., de Chiara, M.L.V., Amigo Rubio, J.M., Colelli, G., 2018. Hyperspectral imaging and multivariate accelerated shelf life testing (MASLT) approach for determining shelf life of rocket leaves. J. Food Eng. 238, 122–133. https://doi.org/10.1016/j.jfoodeng.2018.06.017.

Cho, H., Baek, I., Oh, M., Kim, S., Lee, H., Kim, M.S., 2017. Characterization of E coli biofilm formations on baby spinach leaf surfaces using hyperspectral fluorescence imaging, in: Spiedigitallibrary.Org. p. 102170X. https://doi.org/10.1117/12.2264826.

Derossi, A., Mastrandrea, L., Amodio, M.L., De Chiara, M.L.V., Colelli, G., 2016. Application of multivariate accelerated test for the shelf life estimation of fresh-cut lettuce. J. Food Eng. 169, 122–130. https://doi.org/10.1016/j.jfoodeng.2015.08.010.

Diezma, B., Lleó, L., Roger, J.M., Herrero-Langreo, A., Lunadei, L., Ruiz-Altisent, M., 2013. Examination of the quality of spinach leaves using hyperspectral imaging. Postharvest Biol. Technol. 85, 8–17. https://doi.org/10.1016/j.postharvbio.2013.04.017.

Elmasry, G., Kamruzzaman, M., Sun, D.-W., Allen, P., 2012. Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. Crit. Rev. Food Sci. Nutr. 52, 999–1023. https://doi.org/10.1080/10408398.2010.543495.

ElMasry, G., Wang, N., Vigneault, C., Qiao, J., ElSayed, A., 2008. Early detection of apple bruises on different background colors using hyperspectral imaging. LWT – Food Sci. Technol. 41, 337–345. https://doi.org/10.1016/j.lwt.2007.02.022.

Everard, C., Kim, M., Lee, H., 2014a. A comparison of hyperspectral reflectance and fluorescence imaging techniques for detection of contaminants on spinach leaves. J. Food Eng. 143. https://doi.org/10.1016/j.jfoodeng.2014.06.042.

Everard, C.D., Kim, M.S., Lee, H., 2014b. A comparison of hyperspectral reflectance and fluorescence imaging techniques for detection of contaminants on spinach leaves. J. Food Eng. 143, 139–145. https://doi.org/10.1016/j.jfoodeng.2014.06.042.

Gibson, A., Edgar, J.D., Neville, C.E., Gilchrist, S.E., McKinley, M.C., Patterson, C.C.,

Young, I.S., Woodside, J.V., 2012. Effect of fruit and vegetable consumption on immune function in older people: a randomized controlled trial. Am. J. Clin. Nutr. 96, 1429–1436. https://doi.org/10.3945/ajcn.112.039057.

Giovenzana, V., Beghi, R., Civelli, R., Guidetti, R., 2015. Trends in Food Science & Technology Optical techniques for rapid quality monitoring along minimally processed fruit and vegetable chain. Trends Food Sci. Technol. 46, 331–338. https://doi.org/10.1016/j.tifs.2015.10.006.

He, F., Nowson, C., MacGregor, G., 2006. Fruit and vegetable consumption and stroke: meta-analysis of cohort studies. Lancet.

Huang, H., Liu, L., Ngadi, M., 2014. Recent developments in hyperspectral imaging for assessment of food quality and safety. Sensors 14, 7248–7276. https://doi.org/10.3390/s140407248.

Kim, S.J., Ishii, G., 2007. Effect of storage temperature and duration on glucosinolate, total vitamin C and nitrate contents in rocket salad (Eruca sativa Mill.). J. Sci. Food Agric. 87, 966–973. https://doi.org/10.1002/jsfa.2787.

Kokalj, M., Prikeržnik, M., Kreft, S., 2016. FTIR spectroscopy as a tool to detect contamination of rocket (Eruca sativa and diplotaxis tenuifolia) salad with common groundsel (senecio vulgaris) leaves. J. Sci. Food Agric. 97, 2238–2244. https://doi.org/10.1002/jsfa.8034.

Koukounaras, A., Siomos, A.S., Sfakiotakis, E., 2009. Impact of heat treatment on ethylene production and yellowing of modified atmosphere packaged rocket leaves. Postharvest Biol. Technol. 54, 172–176. https://doi.org/10.1016/J.POSTHARVBIO.2009.07.002.

Lampe, J.W., 1999. Health effects of vegetables and fruit: assessing mechanisms of action in human experimental studies. Am. J. Clin. Nutr. 70, 475s–490s. https://doi.org/10.1093/ajcn/70.3.475s.

Lara, M.A., Lleó, L., Diezma-Iglesias, B., Roger, J.M., Ruiz-Altisent, M., 2013. Monitoring spinach shelf-life with hyperspectral image through packaging films. J. Food Eng. 119, 353–361. https://doi.org/10.1016/j.jfoodeng.2013.06.005.

Liu, Y., Chen, Y.-R., Wang, C.Y., Chan, D.E., Kim, M.S., 2004. Development of hyperspectral imaging technique for the detection of chilling injury in cucumbers, in: Elibrary.Asabe.Org. p. 18. https://doi.org/10.1117/12.597550.

Løkke, M.M., Seefeldt, H.F., Edelenbos, M., 2012. Freshness and sensory quality of packaged wild rocket. Postharvest Biol. Technol. 73, 99–106. https://doi.org/10.1016/j.postharvbio.2012.06.004.

Løkke, M.M., Seefeldt, H.F., Skov, T., Edelenbos, M., 2013. Color and textural quality of packaged wild rocket measured by multispectral imaging. Postharvest Biol. Technol. 75, 86–95. https://doi.org/10.1016/j.postharvbio.2012.06.018.

Lunadei, L., Diezma, B., Lleó, L., Ruiz-Garcia, L., Cantalapiedra, S., Ruiz-Altisent, M., 2012. Monitoring of fresh-cut spinach leaves through a multispectral vision system. Postharvest Biol. Technol. 63, 74–84. https://doi.org/10.1016/J.POSTHARVBIO.2011.08.004.

Ma, L., Zhang, M., Bhandari, B., Gao, Z., 2017. Recent developments in novel shelf life extension technologies of fresh-cut fruits and vegetables. Trends Food Sci. Technol. 64, 23–38. https://doi.org/10.1016/J.TIFS.2017.03.005.

Mann, J.I., 2001. Diet and risk of coronary heart disease and type 2 diabetes. Lancet 360, 783–789.

Martínez-Sánchez, A., Allende, A., Bennett, R.N., Ferreres, F., Gil, M.I., 2006a. Microbial, nutritional and sensory quality of rocket leaves as affected by different sanitizers. Postharvest Biol. Technol. 42, 86–97. https://doi.org/10.1016/J.POSTHARVBIO.2006.05.010.

Martínez-Sánchez, A., Marín, A., Llorach, R., Ferreres, F., Gil, M.I., 2006b. Controlled atmosphere preserves quality and phytonutrients in wild rocket (Diplotaxis tenuifolia). Postharvest Biol. Technol. 40, 26–33. https://doi.org/10.1016/J.POSTHARVBIO.2005.12.015.

Mastrandrea, L., Amodio, M.L., de Chiara, M.L.V., Pati, S., Colelli, G., 2017. Effect of temperature abuse and improper atmosphere packaging on volatile profile and quality of rocket leaves. Food Packag. Shelf Life 14, 59–65. https://doi.org/10.1016/J.FPSL.2017.08.004.

Mishra, P., Shahrimie, M., Asaari, M., Herrero-Langreo, A., Lohumi, S., En Diezma, B., Scheunders, P., 2017. Close range hyperspectral imaging of plants: A review. Biosyst. Eng. 164, 49–67. https://doi.org/10.1016/j.biosystemseng.2017.09.009.

Nagata, M., Tallada, J.G., Kobayashi, T., Toyoda, H., 2005. NIR hyperspectral imaging for measurement of internal quality in strawberries, in: 2005 ASAE Annual Meeting. p. 1.

Nurzyńska-Wierdak, R., 2015. Protein nutritional value of rocket leaves and possibilities of its modification during plant growth. Turkish J. Agric. For. 39, 1023–1028. https://doi.org/10.3906/tar-1412-6.

Oliveira, M., Abadias, M., Usall, J., Torres, R., Teixidó, N., Viñas, I., 2015. Application of modified atmosphere packaging as a safety approach to fresh-cut fruits and vegetables – a review. Trends Food Sci. Technol. 46, 13–26. https://doi.org/10.1016/J.TIFS.2015.07.017.

Osorio, M.T., Haughey, S.A., Elliott, C.T., Koidis, A., 2014. Evaluation of methodologies to determine vegetable oil species present in oil mixtures: proposition of an approach to meet the EU legislation demands for correct vegetable oils labelling. Food Res. Int. 60, 66–75. https://doi.org/10.1016/j.foodres.2013.12.013.

Pérez-Balibrea, S., Moreno, D.A., García-Viguera, C., 2008. Influence of light on health-promoting phytochemicals of broccoli sprouts. J. Sci. Food Agric. 88, 904–910. https://doi.org/10.1002/jsfa.3169.

Pissard, A., Fernández Pierna, J.A., Baeten, V., Sinnaeve, G., Lognay, G., Mouteau, A., Dupont, P., Rondia, A., Lateur, M., 2013. Non-destructive measurement of vitamin C, total polyphenol and sugar content in apples using near-infrared spectroscopy. J. Sci. Food Agric. 93, 238–244. https://doi.org/10.1002/jsfa.5779.

Pu, H., Liu, D., Wang, L., Sun, D.W., 2015. Soluble solids content and pH prediction and maturity discrimination of lychee fruits using visible and near infrared hyperspectral imaging. Food Anal. Methods 235–244. https://doi.org/10.1007/s12161-015-0186-7.

Pu, Y.-Y., Feng, Y.-Z., Sun, D.-W., 2015b. Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: a review. Compr. Rev. Food Sci. Food Saf. 14, 176–188. https://doi.org/10.1111/1541-4337.12123.

Sánchez, M.T., De la Haba, M.J., Guerrero, J.E., Garrido-Varo, A., Pérez-Marín, D., 2011. Testing of a local approach for the prediction of quality parameters in intact nectarines using a portable NIRS instrument. Postharvest Biol. Technol. 60, 130–135. https://doi.org/10.1016/j.postharvbio.2010.12.006.

Santos, P.H.S., Silva, M.A., 2008. Retention of vitamin C in drying processes of fruits and vegetables – a review. Dry. Technol. https://doi.org/10.1080/07373930802458911.

Sasic, S., Ozaki, Y., 2011. Raman, infrared, and near-infrared chemical imaging.

Singleton, V.L., Rossi, J.A., Jr, J., 1965. Colorimetry of total phenolics with phosphomoybdic-phosphotungstic acid reagents. Am J Enol Viticult. 144–158.

Spadafora, N.D., Amaro, A.L., Pereira, M.J., Müller, C.T., Pintado, M., Rogers, H.J., 2016. Multi-trait analysis of post-harvest storage in rocket salad (Diplotaxis tenuifolia) links sensorial, volatile and nutritional data. Food Chem. 211, 114–123. https://doi.org/10.1016/j.foodchem.2016.04.107.

Tallada, J.G., Nagata, M., Kobayashi, T., 2006. Non-destructive estimation of firmness of strawberries (Fragaria x ananassa Duch.) using NIR hyperspectral imaging. Environ. Control Biol. 44, 245–255.

Tewey, K., Lefcourt, A., Tasch, U., Shilts, P., Kim, M., 2017. Hyperspectral, time-resolved, fluorescence imaging system for large sample sizes: Part II. Detection of fecal contamination on spinach.

Toivonen, P.M.A., Brummell, D.A., 2008. Biochemical bases of appearance and texture changes in fresh-cut fruit and vegetables. Postharvest Biol. Technol. 48, 1–14. https://doi.org/10.1016/J.POSTHARVBIO.2007.09.004.

Toledo-Martín, E.M., Font, R., Obregón-Cano, S., De Haro-Bailón, A., Villatoro-Pulido, M., Del Río-Celestino, M., 2017. Rapid and cost-effective quantification of glucosinolates and total phenolic content in rocket leaves by visible/near-infrared spectroscopy. Molecules 22. https://doi.org/10.3390/molecules22050851.

Vigneau, N., Ecarnot, M., Rabatel, G., Roumet, P., 2011. Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in Wheat. F. Crop. Res. 122, 25–31. https://doi.org/10.1016/j.fcr.2011.02.003.

Villatoro-Pulido, M., Moreno Rojas, R., Muñoz-Serrano, A., Cardeñosa, V., Amaro López, M.Á., Font, R., Del Río-Celestino, M., 2012. Characterization and prediction by near-infrared reflectance of mineral composition of rocket (Eruca vesicaria subsp. sativa and Eruca vesicaria subsp. vesicaria). J. Sci. Food Agric. 92, 1331–1340. https://doi.org/10.1002/jsfa.4694.

Webb, A.L., Villamor, E., 2008. Update: effects of antioxidant and non-antioxidant vitamin supplementation on immune function. Nutr. Rev. 65, 181–217. https://doi.org/10.1111/j.1753-4887.2007.tb00298.x.

Workman, J., 2003. NIR spectral characteristics. NIR newa 14, 10–21. https://doi.org/10.1255/nirn.710.

Xue, L., Yang, L., 2009. Deriving leaf chlorophyll content of green-leafy vegetables from hyperspectral reflectance. ISPRS J. Photogramm. Remote Sens. 64, 97–106. https://doi.org/10.1016/j.isprsjprs.2008.06.002.

Yang, H.-Y., Inagaki, T., Ma, T., Tsuchikawa, S., 2017. High-resolution and non-destructive evaluation of the spatial distribution of nitrate and its dynamics in spinach (Spinacia oleracea L.) leaves by near-infrared hyperspectral imaging. Front. Plant Sci. 8. https://doi.org/10.3389/fpls.2017.01937.

Yang, H., Irudayaraj, J., 2002. Rapid determination of vitamin C by NIR, MIR and FT-Raman techniques. J. Pharm. Pharmacol. 54, 1247–1255. https://doi.org/10.1211/0022235702320402099.

Zapata, S., Dufour, J.-P., 1992. Ascorbic, dehydroascorbic and isoascorbic acid simultaneous determinations by reverse phase ion interaction HPLC. J. Food Sci. 57, 506–511. https://doi.org/10.1111/j.1365-2621.1992.tb05527.x.

Zhang, C., Wang, Q., Liu, F., He, Y., Xiao, Y., 2017. Rapid and non-destructive measurement of spinach pigments content during storage using hyperspectral imaging with chemometrics. Meas. J. Int. Meas. Confed. 97, 149–155. https://doi.org/10.1016/j.measurement.2016.10.058.