Precision Agriculture (2019) 20:398–422
https://doi.org/10.1007/s11119-018-9594-1

Assessment of the optimal spectral bands for designing a sensor for vineyard disease detection: the case of '*Flavescence dorée*'

H. Al-Saddik[1] · **J. C. Simon**[1] · **F. Cointault**[1]

Published online: 10 August 2018

Abstract
Disease detection and control is one of the main objectives of vineyard research in France. Manual monitoring of diseases is a time consuming operation especially when fields are large. Current research studies aim to develop automatic in-field systems to detect diseases. This study investigated the optimal spectral bands for the design of a dedicated high-resolution multispectral camera embedded on an unmanned aerial vehicle for identifying infected zones in a grapevine field. The target disease was *Flavescence dorée*, which is infectious, incurable and can result in considerable yield loss. An in-field spectrometry study was performed on four grapevine varieties in the Provence Alpes Côte d'Azur region in France. Two spectral analysis techniques were proposed for choosing the best spectral bands capable of discriminating healthy from diseased leaves. The first novel approach is a feature selection technique based on the successive projection algorithm (SPA), some spectral pre-processing techniques were jointly investigated. The second approach examines a set of traditional vegetation indices (VI). Support vector machine (SVM) and discriminant analysis (DA) are the two classifiers used in this paper and the accuracy of the results is compared for the two methods of analysis. The best models were computed as a function of the grapevine variety considered. The SPA technique performed better in general with respect to common VIs, the overall classification accuracy was more than 96%. Results demonstrated that employing a feature selection technique based on the SPA algorithm can provide a valid tool for determining the optimal bands that are sensitive to *Flavescence dorée* grapevine disease and assist in its identification. The benefit behind the presented procedure relies on the possibility of generalizing it for other infections and stresses or even for different crops.

Keywords Multispectral sensor · Feature selection · Vegetation indices · Successive projection algorithms · Classification · Vineyards · Diseases

✉ H. Al-Saddik
hania.al-saddik@agrosupdijon.fr

[1] INRA, UMR 1347 Agroecology, 21000 Dijon, France

Introduction

France has a reputation for producing wine of high quality. Over 800 000 ha of vineyards can be found in over seventy departments and thirteen regions within France. The economic and territorial development of the country highly relies on the wine and vine sector. In fact, the sector provides employment to around 100 000 people directly and 500 000 people indirectly. Furthermore, among the seven billion bottles of wine produced every year, more than 147 million are exported with a total value of seven billion euros (Maverick 2015).

Since the main losses in this sector are due to disease, a continuous protection approach is required, which means that fungicides/pesticides are uniformly sprayed in vineyards at regular and frequent intervals. More than ten treatments are carried out per season in several of the main wine-producing regions. It is important to detect initial symptoms of diseases in order to target their treatment selectively, preventing and controlling the formation of infection and its epidemic spread to other patches or even to the whole vineyard.

Currently there are two widely-used techniques for in-field disease detection: naked eye observations and biological approaches. Visual inspection is the most used technique for grapevine monitoring. It is a relatively easy approach when visual symptoms characterize a certain disease. Nevertheless, the accuracy of the diagnosis is always subject to an individual's experience and can be affected by temporal variation. Moreover, an expert should be available for permanent monitoring which is both expensive and impractical in larger fields. Although symptoms provide important information on the kind of disease, one cannot make a decision upon appearance only. Usually, other confirmatory tests are performed after visual inspection to ensure an accurate diagnosis (Alemu 2015).

Additional, biological tests can be applied for disease diagnosis. They have revolutionized the identification and quantification of pathogens and diseases. Under this category, are serological and molecular tests. The enzyme-linked immunosorbent assay is a serological method that relies on the proteins produced by a pathogen. The polymerase chain reaction (PCR) is a molecular technique that depends on the specific DNA sequences of a pathogen. There are many variants of PCR and the enzyme-linked immunosorbent assay techniques that can be used for detecting different pathogenic viruses and bacteria in plants.

The above-mentioned biological tests have their limitations. To obtain reliable and accurate results, a definite procedure must be used, especially during sample preparation (collection and extraction), which implies more labor and more time. In order to be more effective, the biology-based techniques must directly be used in the field, therefore, more advanced tests are required for real-time detection of pathogens in their natural environment without the need for culturing or amplification (Lopez et al. 2003).

However, these traditional techniques are incapable of detecting early stages of disease development. An ideal system should locate the initial symptoms and manage the treatments to prevent the infection from developing in the vineyards. The control of grapevine diseases would greatly benefit from an innovative, rapid, non-invasive sensing of vegetation status. Remote sensing enables the gathering of information about an object, area or phenomenon at a distance. Remote sensing can be adapted for the identification of crop diseases based on the assumption that the stress induced by the pest modifies the physiological structure of the plant and affects the absorption of light energy and reflectance spectrum. Many studies have confirmed that remote sensing tools can be used to quantify damage in crops. Dubbini et al. (2017) presented a new multispectral instrument, MAIA, for calculating many VI in the context of agricultural multispectral analysis. MAIA holds

9 different sensors allowing a high and well-distributed number of wavelength bands (1 RGB and 8 monochromatic) covering a range of 390–950 nm. The instrument has a high resolution and a frame rate frequency and it is light-weight making it ideal for monitoring vegetation. A light-weight hyperspectral mapping system was also introduced by Suomalainen et al. (2014) for a rotor-based unmanned aerial vehicle. The improvement made in the processing chain of the system is explained and shows the potential of the instrument in agricultural mapping and tracking applications. Several experimental flights related to habitat monitoring were undertaken. Lu et al. (2017) tested another hyperspectral imaging technique with the aim of discriminating yellow leaf curl on tomatoes. They mainly studied the reflectance spectra and its first derivative along with the absolute reflectance difference spectra in the range of 500 to1000 nm. Sensitive bands and ratios are selected, they were compared to 4 common VI. Texture features were also investigated. A total of 24 features produced by a grey level co-occurrence matrix were calculated and their performance was assessed. The spectral bands at 560, 575 and 720 nm were interesting to use to detect tomato yellow leaf curl virus infection. Vanegas et al. (2018) investigated the merging of different kinds of data. Their paper explored the design of a predictive model for pest detection in grapevine fields by processing airborne-collected Red–Green–Blue, multispectral and hyperspectral data at two different levels of phylloxera infection in vineyards. The collection, analysis and integration of different types of data is detailed in order to evaluate both old and new VI.

The purpose of the project was to define a novel on-board multispectral sensor dedicated to *Flavescence dorée* disease detection. *Flavescence dorée* is a contagious and incurable disease of grapevines, which is spreading throughout France and southern European. *Flavescence dorée* is declared as a reason for quarantine by the European Union. Today, more than half of the French vineyard area is in compulsory control zones. This has raised the attention of many researchers, in particular Chuche and Thiéry (2014) who reviewed the biology and the ecology of *Flavescence dorée* in their paper. Phytoplasmas colonize plant cells and are the micro-organisms behind *Flavescence dorée*. The disease is transmitted from one field to another by the spread of contaminated vegetative material. However, in the same field, the contamination is transmitted from one grapevine to another by the vector *Scaphoideus titanus*, a sucking insect that spreads the disease by feeding on an infected plant and then feeding on another plant with infected saliva. The following symptoms (Fig. 1) should be present simultaneously and on the same shoot to conclude the presence of Phytoplasma: discoloration of leaves depending on the type of grapevines (yellow for white grapevines and red for red grapevines), lack of lignification of new shoots, mortality of inflorescences and berries, non-awning of branches. The difficulty of detecting *Flavescence dorée* is that the symptoms do not appear until 1 year after contamination and can be limited to a single shoot. In addition, *Flavescence dorée* is just one form of the yellowing of the grapevine; its symptoms are similar to those of another disease belonging to the same category: the Stolbur of vines or ‘Bois noir’. Only a polymerase chain reaction can distinguish between the two diseases.

The sensor is supposed to capture a stack of spectral images of a grapevine field each at different narrow spectral bands. Later, an adequate analysis of the images will enable the spatial detection of *Flavescence dorée* symptoms. In this work, the optimal spectral bands that are sensitive to the disease were assessed.

In an earlier work (AL-Saddik et al. 2017b), special spectral indices were developed which were sensitive to the *Flavescence dorée* disease using a genetic algorithms. However, the procedure was somewhat computationally intensive. In contrast, in this article, the SPA algorithm acts as a faster technique that could select interesting

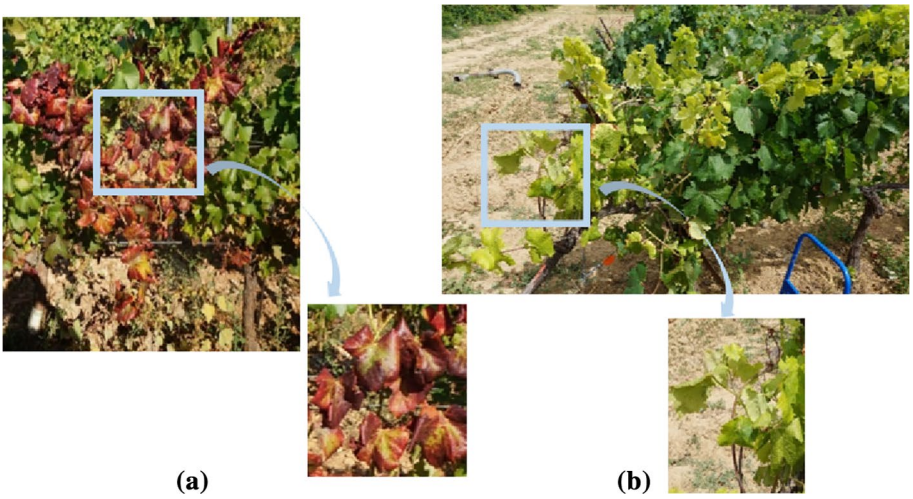

Fig. 1 Foliar symptoms of *Flavescence dorée*, a red discoloration on a red grapevine variety (**a**) and a yellow discoloration on a white grapevine variety (**b**), a twisting of leaves can also be observed (Color figure online)

spectral bands for *Flavescence dorée* detection with a low computational load. Two, instead of one, classifier along with a new multiclass classification approach are also presented in the current study. A detailed comparison between healthy and diseased spectral data is also presented in this article.

In this work, the use of the SPA algorithm, was extended with respect to what was already presented in (AL-Saddik et al. 2017a). SPA has the advantage of reducing data dimensionality and selecting discriminating bands for *Flavescence dorée* disease detection. The SPA approach in AL-Saddik et al. (2017a) was compared to using complete hyperspectral data and gave better results. This is somewhat predictable since the collinearity in the full spectral data will highly degrade results. In this paper, the efficiency of the SPA technique was investigated with respect to another common dimension reduction technique: VI. These are extensively used in remote sensing applied to agricultural fields. Besides, comparing the accuracy of SPA and VI in choosing the most discriminative multispectral bands, in this research a new grouping of the data depending on the grapevine variety was proposed along with two binary and multiclass configurations.

The main advantage of this work is that the methodology can be used for detecting other types of infection on other crops. Existing studies dealing with spectral signature analysis for plant disease detection, give promising results in general but they must be improved. In fact, most of them were performed under controlled conditions; while the measurements in this study were acquired under natural conditions, directly in the field. In addition, many applications use a binary classification between infested and healthy plants. In this work, a multi-class approach was examined which assessed the impact of the disease and its severity on spectral characteristics of plants. Moreover, in existing scientific literature, a specific disease is only tested in a specific cultivar.

However, in this study, four fields were investigated with four different grapevine varieties (2 red-grape and 2 white-grape).

Materials and methods

Sample set-up

Various fields were tested in the French Provence-Alpes Côte d'Azur region. Four varieties of grapevine were studied, two red: Grenache and Marselan, and two white: Vermentino and Chardonnay.

Two tests were performed in 2016: the first one took place at the beginning of August which meant that the symptoms were not yet clearly visible and the second one occurred in September, late in the season, with symptoms being clearly visible. The red grapevine variety tests were performed in the morning (10:00–12:00) and the white grapevine variety tests in the afternoon (14:00–16:00).

From each variety, four diseased and four healthy grapevines were selected. Samples included about two-four leaves per grapevine and two-four measurements were undertaken per leaf. The infested leaves were chosen in order to obtain a complete and representative range of diseases symptoms at the end of the experiment. For the healthy group, the selected leaves were of different ages.

All the tested leaves were inspected by a plant pathologist from the Regional Federation of Defense against Pests of Provence Alpes Côte d'Azur, and they were classified according to disease symptoms presence and intensity. The tests were completed with a polymerase chain reaction analysis on the healthy and diseased grapevines to verify the expert's results. In total, 213 diseased and 201 healthy leaves were assessed (63 Diseased Grenache and 64 Healthy Grenache; 63 Diseased Marselan and 64 Healthy Marselan; 47 Diseased Vermentino and 40 Healthy Vermentino, 42 Diseased Chardonnay and 34 Healthy Chardonnay). In order to ensure a timely follow-up, the grapevines were located using a GPS and the leaves were also labeled.

Reflectance measurements

A portable Spectroradiometer (FieldSpec 3, Analytical Spectral Devices, Boulder, CO, USA) was used to obtain spectral reflectance measurements of the leaf surfaces. Measurements were made on each leaf using a plant probe attachment, which is essentially a closed chamber with an internal light source specially made for vegetative surfaces. The instrument uses a spectral resolution of 3 nm at 700 nm wavelengths and of 10 nm for wavelengths above 1400 nm; however, the software eventually interpolates the spectra to 1 nm intervals. Therefore, each measurement generated a spectrum ranging between 350 nm and 2500 nm at 1 nm increments. The instrument was warmed up for at least 20 min prior to the tests. Absolute reflectance was obtained using a Teflon calibration disk. The number of samples for the spectrum was set to 30, the number of samples for dark current and white reference were set to 100. The measurements were completed within approximatively 4 h.

Spectral data analysis for disease detection

The symptoms which appear on an infected plant are the result of the physiological alteration made by the pathogen. The amount of water, the concentration of pigments and the functionality of the tissue, are all factors that may vary depending on the interaction between the pathogen and the host and may also modify the spectral signatures of the plants.

The objective of the spectral analysis in this paper was to select a subset of wavelengths with the most useful spectral information; wavelengths were included or excluded based on a good discrimination classification efficiency. Once the optimum wavelengths are selected, a proper model can be designed using only the optimum wavelengths instead of the whole spectral data.

In the present paper, some pre-processing techniques were tested and combined with SPA. Another technique for spectral analysis and dimension reduction was further investigated: vegetation indices, since these are extensively used in the remote sensing in the agricultural field. The flowchart of the approach used is detailed in Fig. 2.

Vegetation indices (VI)

Vegetation indices are, by definition, combinations of reflectance at two or many wavelengths, enabling the dimension reduction of the hyperspectral data. They aim was to highlight a particular property of the vegetation. More than 150 VI have been published in the scientific literature. Some VI correlate with the biochemical constituent concentrations of vegetation (chlorophyll, carotenoids, water, cellulose, lignin, dry matter…), thus, associating the physiological status of crops to other hyperspectral data. Other VI have not been systematically tested or do not have a biophysical base. Vegetation Indices are essential in crop management since they can be applied to monitors, maps and can analyze variations in vegetation (spatial/temporal). Furthermore, pigment-specific VI may form an effective

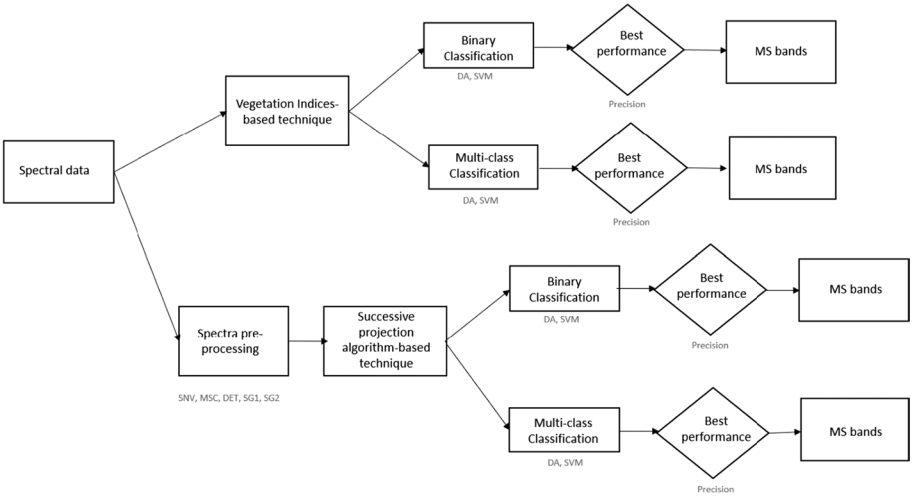

Fig. 2 Flowchart of the methodology used for band selection for *Flavescence dorée* disease detection

Table 1 Typical VI in the literature and specifically applied in this study

Index name	Formula	Association with relevant plant pigment	Reference example
Normalized difference vegetation index (NDVI)	NDVI705 = (R750 − R_705)./(R750 + R705)	NDVI is a very typical index. Positive values suggest vegetated areas	(Tucker 1979; Tucker et al. 1981)
Photochemical reflectance index (PRI)	PRI = (R570 − R531)/(R570 + R531)	PRI index is a function of the reflectance at the 531 nm, this reflectance is related to xanthophyll. When the xanthophyll activity is high, the light use efficiency is low, meaning a possible stress occurred	(Penuelas et al. 1995b; Penuelas et al. 1994; Trotter et al. 2002)
Anthocyanin reflectance index (ARI)	ARI = (1/R550) − (1/R700)	ARI index is designed to estimate the stack of anthocyanin in senescing and stressed leaves	(Gitelson et al. 2001)
The structure insensitive pigment index (SIPI)	SIPI = (R800 − R445)/(R800 + R680)	The SIPI index is responsive to the ratio of carotenoids to chlorophyll. It is very practical to use when the canopy structure or leaf area index are inconsistent	(Blackburn 1998; Penuelas et al. 1995a)
Modified chlorophyll absorption integral (mCAI)	mCAI = (R545 + R752)/2 × ((752 − 545) × ∑ (1.158 × R))	The mCAI is sensitive to chlorophyll content. It calculates the area between a straight line connecting two points (the green peak at 545 nm and 752 nm) and the curve itself	(Laudien et al. 2005)
Pigment specific simple ratio chlorophyll a (PSSRa)	PSSRa = R800/R680	The pigment specific ratio indices were suggested to estimate the pigment's content at the leaf level. Samples from trees at different senescence stages were studied aiming to empirically determine the best individual wavebands for pigment assessment (680 nm for chlorophyll a, 635 nm for chlorophyll b, 470 nm for the carotenoids)	(Blackburn 1998; Richardson et al. 2004; Sims and Gamon 2002)
Pigment specific simple ratio chlorophyll b (PSSRb)	PSSRb = R800/R635		(Blackburn 1998; Apan et al. 2003; Kooistra et al. 2003)
Pigment specific simple ratio carotenoids (PSSRc)	PSSRc = R800/R470		(Blackburn 1998; Kooistra et al. 2003)

Table 1 (continued)

Index name	Formula	Association with relevant plant pigment	Reference example
Gitelson and Merzylak 1 (GM1)	GM1 = R750/R550	GM1 and GM2 were created to measure the chlorophyll content in vegetation leaves	(Gitelson et al. 1996a; Gitelson and Merzylak 1997)
Gitelson and Merzylak 2 (GM2)	GM2 = R750/R700		(Main et al. 2011)
Water index (WI)	WI = R1300/R1450	WI is related to water deficit	(Seelig et al. 2009)
Zarco-Tejada Miller (ZTM)	ZTM = R750/R710	ZTM is a Red edge index highly correlated to chlorophyll content. At the canopy level, it has the advantage of minimizing shadow effects	(Zarco-Tejada et al. 2001; Underwood et al. 2003)

data analysis tool for disease discrimination. In this study, 12 of the most common VI were tested, the list is detailed in Table 1.

Successive projection algorithm (SPA)

The feasibility of choosing some spectral bands to describe the full data while maintaining its efficiency will be proven by using a SPA-based spectral analysis technique.

Prior to SPA, some pre-processing procedures were tested. The most common, for near-infra-red spectra, can be divided into 2 categories: Scatter-Correction (SC) methods and spectral derivatives.

The SC methods reduce the physical variability between samples due to scatter and adjust the baseline shifts between samples. The Multiplicative Scatter Correction (MSC) which is probably the most widely used, is capable of removing artifacts and imperfections. It estimates the correction coefficients (additive and multiplicative contributions). The Standard Normal Variate (SNV) is the second most applied method for scatter correction and it is very similar to normalization (object-wise standardization of the spectra). Removing the trend from the data means the fluctuations can be analysed, and therefore the detrending (DET) technique subtracts the mean or the best-fit line (in the least-squares sense) from the data.

The spectral derivatives can remove both additive and multiplicative effects in spectra. Savitzky-Golay (SG) is one of the methods for numerical derivation of a vector including a smoothing step. To find the derivative at the center point i, a polynomial is fitted in a symmetric window on the raw data. When the parameters for this polynomial are calculated, the derivative of any order of this function can easily be found analytically; this value is subsequently used as the derivative estimate of this center point. Two pre-processing SG techniques were used in this paper: SG1 and SG2 which represent the first and the second SG derivatives, respectively. More details about the above cited techniques can be found in Rinnan et al. (2009).

In the present case, as the data objects (reflectance spectra) are described by a large number of features, dimension reduction seems essential to improve computational efficiency and precision of the analysis. This was done using the SPA algorithm (Araujo et al. 2001). The technique was investigated in many other studies such as by Zhang et al. (2013) or Yang et al. (2015). For this, three main steps are needed. Firstly, the instrumental response data are used to create chains of variables according to a sequence of operations concerning vector projection. These operations are designed to minimize multi-collinearity among the variables of the chain. Secondly, the algorithm creates a model for each of the candidate subsets of variables extracted from the generated chains. Finally, each model is evaluated and the optimal candidate subset is chosen according to its performance.

The number of chosen bands was limited to eight. The maximal area under the curve (AUC) was chosen as a criterion for assessing the performance of a defined SPA subset. The ideal case is to have an AUC close to one.

Classification

Choosing the most appropriate classifier for an application is a difficult task because hundreds of classifiers can be found in the literature. The empirical approach is usually used, which means that researchers try several classifiers and then adopt the one having the

highest accuracy for their application. In this research, the two most widely used classifiers, Discriminant Analysis (DA) and Support Vector Machine (SVM) were applied.

A tenfold cross-validation technique was used where the original sample is randomly partitioned into 10 equal sized subsamples. Of the 10 subsamples, a single subsample is retained as the validation data for testing the model, and the remaining 9 subsamples are used as training data. The model accuracy defines the percentage of correctly classified test set samples; the false negative rate is the percentage of negative results that are, in fact, positive; and the, false positive rate is the percentage of positive results that are, in fact, negative. Note that, in this case the positive class is considered as the diseased class and the negative class as the healthy one.

Discriminant analysis (DA)

The original discriminant analysis was first developed by *R. Fisher,* however, Welling (2014) detailed the theory of linear discriminant analysis that is a generalization of Fisher's linear discriminant.

Discriminant analysis derives from discriminant functions which are linear combinations of the independent variables providing the best discrimination between the groups of the dependent variables. Each function is tested with a discriminant score to evaluate its capacity of predicting groups.

Support vector machine (SVM)

Support vector machines (SVM) find the best separating hyperplane between classes according to the hyperplane with the maximal margin. A cost parameter defines the penalty for misclassifying objects. This can be used to introduce a soft margin allowing objects to lie between the margins or on the wrong side of the plane. This is mainly suitable when classes are not fully separable.

When non-linear decision boundaries are needed in the feature space, kernels might be applied in SVM modeling. There are different kernels in the literature. Due to its flexibility, the radial basis function kernel is the most used, therefore it was also applied in the current study. For more details about SVM, refer to Ben-Hur and Weston (2010) who discussed the concept and the application of SVM in detail.

Data configuration

Two dimensions were used in the analysis. The first dimension is the severity of infestation which implies the number of classes used. The latter can be 2 for binary (shown in Fig. 3) or 3 for multi-class (shown in Fig. 4); in fact, it is possible to either study the healthy group versus the diseased group as a whole (slightly infested measurements from the August acquisition campaign + highly infested measurements from the September acquisition campaign); or to study the healthy group versus the slightly infested group versus the highly infested group as different classes.

The second dimension is the type of measurements. In the testing campaigns, four varieties were taken into consideration: Marselan, Grenache, Vermentino and Chardonnay. They each could have been studied separately or combined which means that the analysis could be conducted based on the grapevine color; so, 2 groups were considered: Red (measurements from Marselan + measurements from Grenache) vs White (measurements from

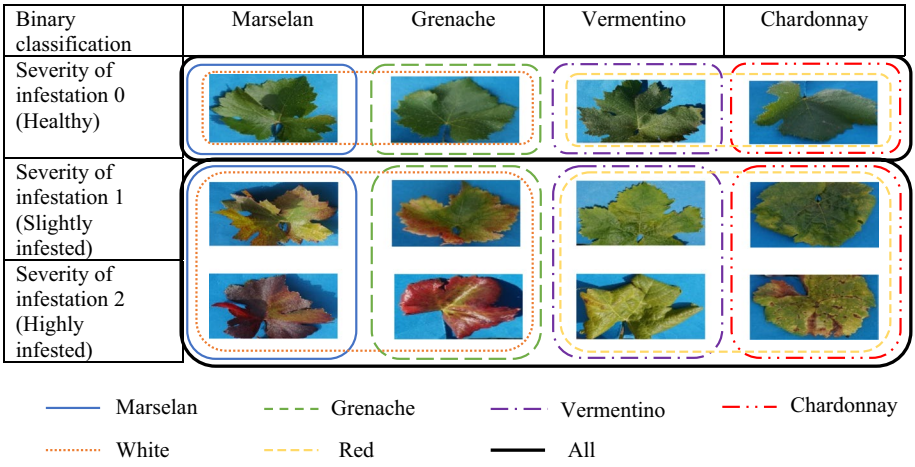

Fig. 3 Configuration relative to a binary classification

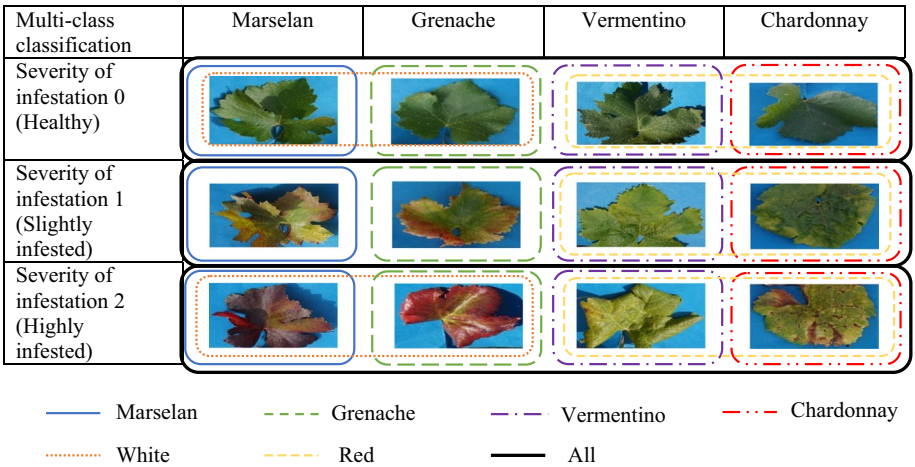

Fig. 4 Configuration relative to a Multi-class classification

Vermentino + measurements from Chardonnay). In the last case, all the grapevine measurements were combined together (measurements from Marselan + measurements from Grenache + measurements from Vermentino + measurements from Chardonnay).

Results

The ultimate aim of band selection is to determine the design of the multispectral sensor. The spectra were assessed starting from 400 nm to 2300 nm due to a high signal to noise ratio near the ends of the spectra.

Reflectance spectra of diseased grapevine leaves

In the normal case, the spectra have a strong absorption in the visible region due to photochemical pigments (chlorophylls and carotenoids). However, in the near infra-red region, the reflectance is high, this is the result of multiple internal light scattering that is a function of the complexity of the leaf's structure.

When compared to healthy spectral signatures, it can be clearly seen from Fig. 5 that the reflectance for both diseased red and white berried leaves were somewhat different. Thus, this clearly shows that the spectral response was affected by the infestation.

For the Marselan variety (red-grape), the peak in the green range appearing in the visible part of the infected spectra moved towards longer wavelengths. Moreover, the healthy spectra were higher than the infested one in the visible region (mainly between 500 and 700 nm), but the opposite occurred in the near infra-red region (800–1300 nm) and in the infra-red region (> 1300 nm). It would seem that when the infestation level increases, the spectral signature decreases in the visible region and increases in the near infra-red region. The same trend was also observed for Grenache white grape leaves (spectra not shown).

On the other hand, for the Chardonnay variety (white-grape), the peak in the green line in the visible part of the infected leaf remained in position for infected leaves. Furthermore, the healthy spectra were lower than the infested one in the visible region (mainly between 500 and 700 nm), but the opposite occurred in the near infra-red region (800–1300 nm) and in the infra-red region (> 1300 nm). It would therefore seem that when the infestation level increases, the spectral signature increases in the visible region and decreases in the near infra-red region. The same trend was also being observed for Vermentino red grape leaves (spectra not shown).

Above 1400 nm, the difference between the healthy and diseased spectra was the same in the red and white varieties

These changes suggest that the spectral signature depends on the pathogen-host interaction. In other words, grape cultivars do not show the same patterns of spectral response when the same infestation occurs.

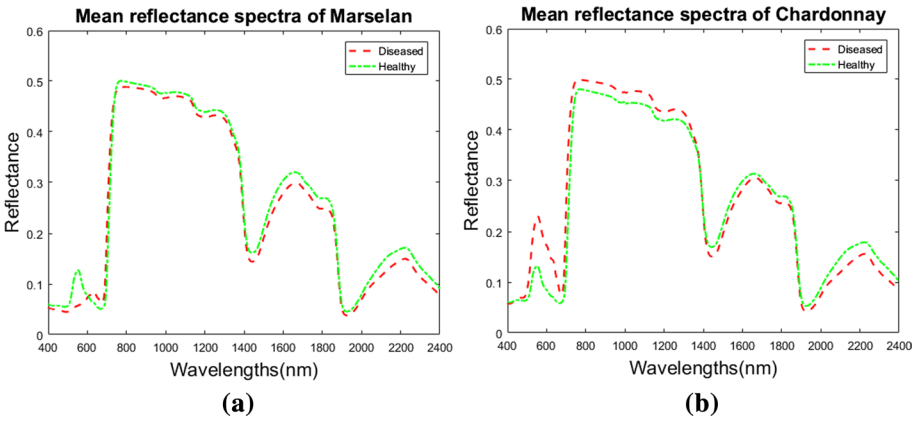

Fig. 5 Mean reflectance of highly infected leaves (severity level = 2; Red), healthy (severity level = 0; green) for Marselan grapevine leaves (**a**) and Chardonnay grapevine leaves (**b**) (Color figure online)

Binary approach

In this part of the study, each group contained measurements from both the August and September acquisitions. This means that there was no distinction between the infestation level and that only healthy and infested groups exist.

SPA-based technique

For the SPA-based technique, all the pre-processing techniques were assessed but only the best one was kept and compared to the case where no pre-processing was applied.

From Table 2, when no pre-processing is applied, the accuracy of SVM was slightly better than DA in all cases. Consider as an illustration, the case where no pre-processing is applied and all data are gathered, the DA classifier scored 92.0% while the SVM scored 95.4%. The results were generally close when each type of grapevine data is studied separately and when grouped data are used. For example, when no pre-processing is applied and the Vermentino data are considered, the SVM classifier scored 96.5%, similar to the white grape data which scored 96.3% with the SVM classifier.

When pre-processing was used, DA and SVM performed similarly when each grapevine type was studied alone. Indeed, when pre-processing is applied and when Vermentino data are considered, both the DA and SVM classifiers scored 98.8%. However, when the data were grouped, SVM was more robust than DA. Notably, where pre-processing is

Table 2 Classification results from the SPA-based technique with the binary approach applied and with no pre-processing and with pre-processing

Technique		Data	Accuracy (%)		False Negative Rate (%)		False Positive Rate (%)	
			DA	SVM	DA	SVM	DA	SVM
No Pre-Processing		Marselan	96.06	**97.63**	4.76	1.63	3.12	3.03
		Grenache	93.54	**95.96**	7.57	3.17	5.17	4.91
		Vermentino	95.40	**96.55**	6.97	4.76	2.27	2.22
		Chardonnay	93.42	**93.42**	5.88	5.88	7.14	7.14
		Red	94.42	**95.61**	5.46	5.34	5.69	3.33
		White	95.09	**96.31**	6.49	4.00	3.48	3.40
		All	92.02	**95.41**	10.69	8.21	5.02	0.51
Best Pre-Processing	MSC/**MSC**	Marselan	**99.21**	**99.21**	0.00	0.00	1.58	1.58
	SG2/SG2 or **MSC** or SG1	Grenache	**98.38**	**99.19**	0.00	0.00	0.98	1.63
	MSC/**MSC**	Vermentino	**98.85**	**98.85**	0.00	0.00	2.12	2.12
	MSC/**MSC**	Chardonnay	**98.68**	**98.68**	0.00	0.00	2.38	2.38
	DET/**SG2**	Red	98.00	**99.60**	1.57	0.00	2.41	0.80
	SG2/SG1 or SNV	White	95.70	**97.54**	6.41	2.66	2.35	2.27
	SG2/**SG2**	All	94.20	**97.82**	7.61	3.82	3.92	0.48

The values in bold represent the highest precision

NB. Only the best pre-processing technique is displayed

DA discriminant analysis, *SVM* support vector machines, *MSC* multiplicative scatter correction, *SG1* Savitzy-Golay first derivative, *SG2* Savitzy-Golay second derivative

applied and all data are considered, the DA classifier scored 94.2% while the SVM classifier resulted in a higher accuracy of 97.8%.

The results show that using pre-processing with SPA was more efficient in terms of accuracy no matter what classifier is used. For instance, when no pre-processing is applied and the Grenache data are considered, the SVM classifier scored a high result (95.9%). This was also the case when pre-processing was used with the Grenache data, the SVM classifier resulted in a good accuracy of 99.1%.

The best pre-processing technique was MSC when each grapevine type was studied separately and SG2 when grapevine data were grouped.

VI based technique

For the VI-based technique, firstly, the efficiency of each index alone was tested and only the best one was displayed. Secondly, all VIs were combined to form a feature vector for each spectral signature and the results were compared.

From Table 3, when only one VI was studied, SVM and DA gave comparable results except for Chardonnay, where SVM proved to be better. In fact, when the best VI is considered, the SVM resulted in a higher accuracy: 97.3% than DA: 93.4%. Furthermore, for all grouped data, the accuracy was worse than ungrouped data but for the grouped data, it was better for SVM (88.1%) than DA (85.5%).

Table 3 Classification results from VI-based technique: each VI was tested alone and the best VI displayed, then VI were combined (Binary approach applied)

Technique	Data		Accuracy (%)		False Negative Rate (%)		False Positive Rate (%)	
			DA	SVM	DA	SVM	DA	SVM
Best VI	mCAI/**ARI**	Marselan	92.12	**93.70**	10.14	8.69	5.17	3.44
	ARI/**ARI**	Grenache	**91.93**	**91.12**	13.69	13.69	1.96	1.96
	GM1/**GM1**	Vermentino	**95.40**	**95.40**	9.52	9.52	6.66	6.66
	NDVI-PRI-mCAI-PSSRb-GM1-GM2-ZTM/**ZTM**	Chardonnay	93.42	**97.36**	10.52	20.00	2.63	8.33
	ARI/**ARI**	Red	**93.62**	**93.22**	11.26	11.18	1.83	0.92
	GM1/**ARI**	White	**95.09**	**95.09**	28.00	8.64	23.86	1.21
	mCAI/**mCAI**	All	85.50	**88.16**	21.37	16.14	4.21	7.85
VI combination	Marselan		92.91	**92.12**	10.14	7.81	5.17	9.52
	Grenache		95.96	**97.58**	4.55	3.08	1.72	1.69
	Vermentino		94.25	**94.25**	9.09	6.98	2.33	2.27
	Chardonnay		94.73	**97.36**	10.53	2.86	2.63	2.44
	Red		95.21	**95.61**	4.58	5.30	2.50	2.52
	White		94.47	**96.31**	8.64	6.33	1.22	1.19
	All		91.54	**94.44**	12.83	8.76	2.66	2.03

The values in bold represent the highest precision

DA discriminant analysis, *SVM* support vector machines, *mCAI* modified chlorophyll absorption integral, *ARI* anthocyanin reflectance index, *GM1* Gitelson and Merzylak 1, *NDVI* normalized difference vegetation index, *PRI* photochemical reflectance index, *PSSRb* pigment specific simple ratio carotenoids, *GM2* Gitelson and Merzylak 2, *ZTM* Zarco Tejada Miller

ARI was the best VI for Marselan, Grenache and for both red and white grapes, GM1 was best for Vermentino, ZTM was the most suitable VI for Chardonnay grapes and mCAI was the most robust when all white and red-berried data were grouped together.

When VIs were combined, the accuracy increased overall for both classifiers with respect to using only one VI, especially for the case of grouped data. Take for example, the case where the best VI is considered and all data are grouped, the DA classifier gave a lower accuracy: 85.5% than the case where all VIs are combined: 91.5%. In addition, the SVM classifier gave better classification accuracy than DA. This was illustrated by the case where all VIs are combined for the Chardonnay data, the SVM classifier achieved higher accuracy (97.3%) than DA (94.7%).

Multi-class approach

In this part of the research, a distinction is made in the measurements from August and September between infestation levels and hence healthy, slightly infested and highly infested groups exist.

SPA based technique

Referring to Table 4, when no pre-processing was undertaken, the accuracy of SVM was slightly better than DA except for white grape data. As proof, when all data are grouped together and no pre-processing is used, the DA gave an accuracy of 87.9% slightly less than SVM: 89.9%.

Table 4 Classification results from the SPA-based technique with no pre-processing and with pre-processing

Technique		Data	Accuracy (%)		False Negative Rate (%)		False Positive Rate (%)	
			DA	SVM	DA	SVM	DA	SVM
No pre-processing		Marselan	93.30	**94.86**	9.09	6.92	0.00	0.00
		Grenache	96.60	**98.21**	4.76	2.38	0.00	0.00
		Vermentino	91.25	**94.70**	7.26	6.92	12.50	0.00
		Chardonnay	88.41	**88.41**	10.98	10.98	12.50	12.50
		Red	88.98	**95.63**	6.47	5.76	19.76	0.00
		White	**89.99**	87.95	8.82	9.00	12.54	17.89
		All	87.91	**92.05**	10.13	9.14	16.23	3.66
Best pre-processing	**SG1**/SG1	Marselan	**98.43**	96.65	2.38	4.76	0.00	0.00
	SG1/SG2	Grenache	**100.00**	98.21	0.00	2.38	0.00	0.00
	SG2/SG2	Vermentino	98.21	**98.27**	2.38	2.38	0.00	0.00
	SG1/SG1 or MSC	Chardonnay	97.61	**97.72**	3.57	3.57	0.00	0.00
	SG2/**SG1**	Red	94.94	**98.41**	4.87	2.38	5.60	0.00
	SG1/**SG2**	White	92.93	**96.03**	4.41	5.63	11.66	0.00
	SG1/**SG1**	All	93.92	**96.37**	5.75	3.98	6.22	2.92

The values in bold represent the highest precision

NB only the best pre-processing technique is displayed (Multi-class approach applied)

When pre-processing was applied, the accuracy for both classifiers was enhanced with respect to the case where no pre-processing was used. For Marselan and Grenache, DA performed better than SVM. For example, when pre-processing is applied for Grenache data, the DA classifier gave a 100% accuracy which was better than SVM (98.2%). However, for other grapevine types and when the grouping of data occurred, SVM was more robust than DA. Actually, when pre-processing is used for all gathered data, the SVM classifier scored a 96.3% accuracy which was better than DA (93.9%). It can be seen that when data is taken individually the accuracy was a bit better than the case where grapevine varieties are combined for the DA classifier. Consider for instance each type of grapevine varieties individually, when pre-processing is used together with the DA classifier, the accuracy was higher than 97%. This accuracy was also higher than the case where data are grouped and pre-processing is also used with the DA classifier (93.9%), but the results were comparable for SVM.

The results show that applying pre-processing with SPA was more efficient in terms of accuracy no matter what classifier was used. As a matter of fact, for the Chardonnay data and when no preprocessing was used, the accuracy for the DA classifier was 88.4%, however, when pre-processing was applied, DA achieved a higher result of 97.6%.

The best pre-processing techniques were SG1 and SG2 when each grapevine type was studied separately and when spectral grapevine data were grouped.

VI based technique

Referring to Table 5, when the best VI is considered, the DA classifier was slightly more accurate than SVM especially for the red-grape varieties. In fact, when the best VI was used for the red variety data, the DA classifier resulted in an accuracy of 98.7% which is higher than that given by the SVM classifier (92.4%). Furthermore, for all grouped data, the accuracy was worse. For example, the best VI with the DA classifier gave an accuracy of 88.3%.

ARI was the best for Grenache and Red-berried data. GM1 was chosen for Vermentino, Chardonnay and for white-berried data, the same goes for ZTM. NDVI performed well for Vermentino, Chardonnay, white varieties and for all combined data. mCAI was the most robust when all white and red-berried data were grouped together and for Marselan data.

In the case of the multi-class approach, unlike the binary one, when VIs were combined, the overall accuracy decreased slightly for both classifiers. Like when the best VI is applied for Vermentino data along with the DA classifier, the accuracy was 94.5%. However, when the VIs were combined also for the Vermentino data, a lower accuracy (91.8%) was found. When all data are grouped, the trend was the opposite. In fact, when the best VI was used with the DA classifier, the accuracy was 88.3% while when VIs were combined, the same classifier scored 91.0%. Similar to using only one VI, the DA classifier also gave better classification accuracy than SVM. This can be seen in the case where VIs are combined, the Marselan with the DA classifier resulted in 90.0% accuracy which is better than the case where SVM was used: 87.5%.

Table 5 Classification results from VI-based technique: each VI was tested alone and the best VI results displayed, then VIs were combined to give a multi-class approach

Technique	Data		Accuracy (%)		False Negative Rate (%)		False Positive Rate (%)	
			DA	SVM	DA	SVM	DA	SVM
Best VI	**SIPI-mCAI**/SIPI-mCAI	Marselan	**90.00**	87.50	16.00	19.23	5.88	5.88
	ARI-GM1/ARI-GM1	Grenache	**89.74**	89.74	11.53	14.28	7.69	0.00
	NDVI-PSSRb-GM1-GM2-WI-ZTM/same	Vermentino	**94.59**	94.59	4.76	4.76	6.25	6.25
	NDVI-PRI-mCAI-PSSRb-GM1-GM2-ZTM/PRI	Chardonnay	**93.33**	93.33	7.14	7.14	6.25	6.25
	ARI/mCAI	Red	**98.73**	92.40	4.65	13.04	0.00	3.03
	NDVI-PSSRb-GM1-GM2-ZTM/same	White	**95.52**	95.52	5.71	5.71	3.12	3.12
	NDVI-mCAI/NDVI-ARI	All	**88.35**	89.04	9.72	12.50	12.16	6.15
VI Combination	Marselan		**90.00**	87.50	0.13	0.13	0.05	0.11
	Grenache		**84.61**	84.61	0.04	0.04	0.12	0.06
	Vermentino		**91.89**	94.59	0.13	0.04	0.07	0.06
	Chardonnay		**93.33**	93.33	0.07	0.07	0.06	0.06
	Red		**97.46**	96.20	0.04	0.04	0	0
	White		**95.52**	95.52	0.05	0.05	0.03	0.03
	All		**91.09**	91.09	0.09	0.13	0.07	0.03

The values in bold represent the highest precision

DA discriminant analysis, *SVM* support vector machines, *SIPI* structure insensitive pigment index, *mCAI* modified chlorophyll absorption integral, *ARI* anthocyanin reflectance index, *GM1* Gitelson and Merzylak 1, *NDVI* normalized difference vegetation index, *PRI* photochemical reflectance index, *PSSRb* pigment specific simple ratio carotenoids, *GM2* Gitelson and Merzylak 2, *ZTM* Zarco Tejada Miller, *WI* water index

Discussion

The spectral reflectance characteristics change when the cellular leaf structure is altered due to the presence of a pathogen (Jacquemoud and Ustin 2001) which is in accordance with the results. In fact, some differences in leaf reflectance between *Flavescence dorée*-infected and healthy grapevines were distinguished. This was also compatible with the results found by Naidu et al. (2009), where grapevine leaf-roll infection caused metabolic and pigment changes and the leaf spectral properties were not the same in healthy and infected leaves.

Both red and white grape varieties have green leaves when healthy and discoloured leaves when infected. However, when a disease occurs, the leaf discoloration is different. For a red grapevine variety, infected leaves become reddish and for white grapevine varieties, infected leaves become yellowish. This is indeed evidence that the concentration of biochemical pigmentation after infection varies differently for the two types of variety. This is also highlighted by the alteration of the diseased signatures compared to the healthy ones in the visible part of the spectrum (Fig. 5).

The Chlorophyll a + b absorption bands at 430 and 450 nm and then at 670 and 630 nm (Blackburn 1998; Richardson et al. 2004; Sims and Gamon 2002; Apan et al. 2003; Kooistra et al. 2003) decreased in both varieties for *Flavescence dorée* infested leaves (Fig. 3a, b). This shows that chlorophyll pigments were significantly reduced in both varieties after infection. Carotenoid absorption bands at 450 and 480 nm (Blackburn 1998; Kooistra et al. 2003) similarly decreased for *Flavescence dorée* infested leaves from both varieties (Fig. 3a, b). This indicates that Carotenoids were also significantly reduced in both varieties after infection. The result is supported by Bertamini et al. (2003) who investigated a field of grown apple (Malus pumila Mill) leaves infected by apple Phytoplasma proliferation and found that the contents of Chlorophyll a + b and carotenoids distinctly decreased in infected leaves. In general, the reduction in the concentration of photosynthetic pigments (mainly chlorophylls), leads to a reduction of the photosynthetic rate (Richardson et al. 2004; Riedell and Blackmer 1999; Penuelas and Filella 1998).

The anthocyanin absorption band at 550 nm (Gitelson et al. 2001) increased in red-grapevine varieties and decreased in white grapevine varieties (Fig. 5a, b). This means that the anthocyanin concentration in red grapevine varieties is raised when *Flavescence dorée* occurs while it is reduced in white grapevine varieties. This is in accordance with the findings of Himeno et al. (2014). They demonstrated that anthocyanin accumulation is directly responsible for the purple discoloration symptoms. The Phytoplasma infection caused a significant activation of the anthocyanin biosynthetic pathway and a dramatic accumulation of sucrose.

Tan et al. (2015) stated that Phytoplasmas reduce stomatal conductance making it hard for water vapor to exit through the stomata of a leaf. Their finding was validated in this study. In fact, the lines for absorption bands at 1400 and 1900 nm (Fig. 5a, b) decreased after infection in both varieties.

Between 800 and 1400 nm, the reflectance changed depending on the structure of the leaf. For red varieties, the absorbance increased, but for white varieties, the opposite occurred. This is basically due to modified leaf tissue and the stunting of plants due to the *Flavescence dorée* infection. The results obtained for the Chardonnay white-berried case, were also found for Russian wheat aphid damage in winter wheat (Mirik et al., 2007). For the Marselan red-berried case, damaged leaves reflected more energy than the control leaves in the near infra-red wavelengths. This was also found in Russian aphid-damaged

leaves compared to healthy ones by Riedell and Blackmer (1999). As mentioned in (Vitali et al., 2013), the contents of soluble proteins were decreased in phytoplasma-infected leaves of field-grown grapevines, but the contents of soluble sugars and total saccharides significantly increased. Leaves were unable to produce sufficient carbohydrates for their own needs. The starch content of leaves was further significantly increased. Nevertheless, in the scientific literature, the real difference of the effects of *Flavescence dorée* on the 2 grapevine varieties wasn't clear.

It is not possible to tell if the spectral difference function of varieties is constant and repeatable or also if it differs between regions. In fact, the grapevine varieties responded to the same virus in different manners. This is most likely due to other specific nuances in addition to the variety itself. MacDonald et al. (2016) noted that factors such as leaf inclination, illumination and surface texture (Gitelson et al. 2002) that differ among vineyard sites due to plant height and health, vegetative growth stages, trellis system, soil properties (type, moisture and available nutrients) may result in spectral variations that affect sensor data and detection sensitivity. Moreover, the development of symptoms may be affected by scion-rootstock combinations and environmental factors and can therefore vary between vineyards and growing seasons (Naidu et al. 2014). All these factors were uncontrollable and they may very well have influenced the results. Hence, more tests in other regions must be undertaken with other varieties of grapevines to validate all the results obtained during the research.

Mirik et al. (2007), confirmed that NDVI was correlated to the presence of Russian wheat aphid in four out of six tested fields. In addition to that, in their study, Yuan et al. (2014) identified the NDVI as the most discriminating feature since it was capable of discriminating between 3 kinds of wheat stressors (yellow rust, powdery mildew, wheat aphid). In this study, the popular NDVI performed well in only a few cases: in the multi-class approach and only for the Vermentino, Chardonnay, white varieties and all data combined. This observation was also reported by Devadas et al. (2009), where NDVI was not capable of discriminating rust infection in wheat leaves.

ARI tracks changes in photosynthetic efficiency. It was chosen in the binary and multi-class approaches for red-grape data and the index was also shown to be significant in the study of Devadas et al. (2009). ARI was identified and considered by Zhang et al. (2012) as an ideal candidate to diagnose yellow rust disease in winter wheat owing to its sustained sensitivity to the disease presence. The index was also found to be efficient in further studies such as that of Huang et al. (2007).

The PRI was used to assess the infestation of the grapevine leaf roll-associated virus (Naidu et al. 2009), and pest damage (Luedeling et al. 2009). PRI was also proven to be sensitive to different infestation scales of rice leaves damaged by the rice leaf holder pest (Huang et al. 2012) but, in the present study, it was only an indicator of *Flavescence dorée* for the Chardonnay grapevine variety.

Vegetation stress can be assessed by measuring the chlorophyll because it is directly related to the photosynthesis process of light harvesting, initiation of electron transport and it is responsive to a range of stresses (Zarco-Tejada et al. 2000). This was confirmed in this study as the ZTM was chosen as the best VI for white-grape data. A reduction in chlorophyll levels as a reaction to an infestation induced by sap feeding insects like aphids (Khawas and Khawas 2008; Lisa et al. 2007; John et al. 2007), phylloxera (Blanchfield et al. 2006) and leafhoppers (Murugesan and Kavitha 2010) have been reported. GM1 was selected also for white-grape data in this study. In fact, differences in reflectance between healthy and stressed vegetation due to changes in Chlorophyll a + b levels have

been detected previously in the green peak and along the red-edge spectral region of 690–750 nm (Gitelson and Merzlyak 1996b).

Other VIs such as GM2, WI and PSSR were not relevant for *Flavescence dorée* detection in this study.

In the binary approach, the VI combination increased the accuracy with respect to using only one VI. The analysis, conducted by Mahlein et al. (2010) to discriminate between three fungal leaf diseases of sugar beet, confirmed that combinations of VIs are efficient for discriminating between healthy and diseased vegetation and even between the 3 diseases in early stages of development. In addition, the classification accuracies of grapevine leaf roll disease infection improved when using multiple variables (indices and individual reflectance bands) with respect to using only a single VI (Naidu et al. 2009). Rumph et al. (2010) mentioned that the accuracies of automatic classification of diseases can be enhanced when applying a combination of spectral VIs that use different wavelengths and, hence, describe different physiological parameters. In general, it can be concluded that there is a noticeable potential for VI combinations in disease detection and identification. This was, however, not the case for the multi-class approach in this study, where the accuracies of using one VI and a combination of VIs were comparable and slightly lower for the VI combination.

The SVM classifier was better than DA in almost all cases. This is in accordance with the findings of Rumph et al. (2010), who found that by comparing different classifiers, SVMs use the inherent information of the vegetation indices in an optimal way. With SVMs, the identification of diseased leaves (binary approach) and the differentiation between distinct disease levels (multi-class approach) can both be achieved.

The optimal performing models with the best accuracies are shown in Table 6. In conclusion, the wavelengths selected are related to characteristic points (peaks, valleys, shoulders, inflections) in the spectra.

The results suggest that changes in spectral reflectance differ by cultivar and therefore the technique can be optimized for each variety in order to obtain good results. Since more than one variety can be grown in the same vineyard area, different measurements were combined to simulate the case where grapevine varieties of the same group were grown together and the approach was tested (Marselan and Grenache measurements were combined in the "Red" group/Vermentino and Chardonnay measurements were combined in the "White" group). In the case where different grapevine varieties were grown together (Marselan, Grenache, Vermentino and Chardonnay were combined in the "All" group), all the measurements are combined. More than 96% accuracy was still achieved using preprocessing and SPA, which means that the sets of spectral bands obtained in a binary or a multi-class context can be considered as generalized for designing the multispectral sensor.

Conclusions

Highly accurate, reliable and non-contact based systems are required to automatically examine plant diseases. Remote sensing devices can identify infected zones in a field by defining the spatial locations of the infection spots in an early manner. Thus, remote sensing may provide a proper tool to analyze large fields quickly and eradicate primary infections early, thus avoiding secondary spread.

In general, imaging technology has emerged as a powerful tool in agriculture but it is not exploited much in vineyards. For the moment, the effort to make the early detection of grapevine diseases automatic is minimal with respect to the major impact of those diseases

Table 6 Wavelengths selected according to the best classification results

Grapevine type	Classification type	Technique	Accuracy (%)	Wavelengths
Marselan	Binary	**MSC-SPA-**DA/**SVM**	**99.212**	499 539 679 759 1389 1659 2029 2299
	Multi-class	**SG1-SPA-DA**	**98.437**	599 689 749 1059 1119 1279 1539 1989
Grenache	Binary	**MSC**/SG1/SG2-**SPA-SVM**	**99.193**	459 559 609 669 769 979 1069 1999
	Multi-class	**SG1-SPA-DA**	**100.000**	409 519 689 1049 1099 1139 1649 1799
Vermentino	Binary	**MSC-SPA-**DA/**SVM**	**98.850**	429 519 689 739 999 1119 1889 2299
	Multi-class	**SG2-SPA-SVM**	**98.275**	419 559 869 999 1039 1389 1579 1669
Chardonnay	Binary	**MSC-SPA-**DA/**SVM**	**98.684**	479 509 679 729 739 1129 1389 1889
	Multi-class	MSC/**SG1-SPA-SVM**	**97.727**	459 679 809 1029 1109 1149 1519 1809
Red	Binary	**SG2-SPA-SVM**	**99.601**	449 689 839 889 1039 1219 1569 1899
	Multi-class	**SG1-SPA-SVM**	**98.412**	679 739 1049 1119 1559 1639 1799 1909
White	Binary	**SG1/**SNV-**SPA-SVM**	**97.546**	419 679 739 849 1099 1129 1809 1989
	Multi-class	**SG2-SPA-SVM**	**96.038**	499 569 789 879 1009 1209 1539 1889
All	Binary	**SG2-SPA-SVM**	**97.826**	409 489 839 879 1009 1219 1539 1889
	Multi-class	**SG1-SPA-SVM**	**96.378**	409 499 689 1039 1099 1129 1779 1809

The values in bold represent the highest precision

on grapes yield and wine quality. In this study, *Flavescence dorée* vine disease was studied under real production conditions and in many cultivars including those with white and red-grapes at different stages of severity.

This study showed that spectral reflectance can be a promising technique for in-field diagnosis of *Flavescence dorée*. The results from this article further suggest that changes in spectral reflectance are function of the cultivar type and therefore to get the best results, the technique should be refined for each variety. However, in the same field, many grapevine varieties could be grown together, so measurements were combined in a binary and multi-class classification context to replicate that case.

On the basis of the above analysis, it can be concluded that dimension reduction methods including the use of traditional VI and SPA employed here gave promising results. The advantages include the fact that they save time and are less complex. The technique based on pre-processing then SPA was however better and SG1, SG2 and MSC were the best pre-processing techniques. In this study, an accuracy higher than 96% was eventually demonstrated for *Flavescence dorée* detection.

Here, only one infection was investigated, other studies targeted the detection of powdery and downy mildew and grapevine leaf-roll in grapes and so it seems essential to explore other diseases and stresses. Moreover, damage need to be detected not only in symptomatic hosts in their natural production environment but also in asymptomatic hosts when the symptoms are not so evident.

The objective of the overall project was to develop a specific solution for the automatic foliar detection of *Flavescence dorée* disease using low altitude-micro-UAV imagery. In the near future, a high-resolution multispectral camera will be developed based on this analysis to identify the occurrence of infected leaves and additional work will address the processing of the associated multispectral images.

Acknowledgements We would like to thank Alice Dubois and Sylvain Bernard from the Provence Alpes Côte Azur Regional Federation of Defense against Pests, Corinne Trarieux from the Interprofessional Office of Burgundy Wine and Jocelyn Dureuil from the Chamber of Agriculture 71, for their expertise. We also would like to thank the funders of the DAMAV project: BPI, FEDER and Burgundy region and of course the winegrowers for their cooperation.

Compliance with ethical standards

Conflicts of interest The authors declare no conflict of interest.

References

Alemu, K. (2015). Detection of diseases, identification and diversity of viruses: a review. *Journal of Biology, Agriculture and Healthcare, 5*(1), 204–213.

Al-Saddik, H., Simon, J. C., Brousse, O., & Cointault, F. (2017a). Multispectral band selection for imaging sensor design for vineyard disease detection: case of Flavescence Dorée. In J. A. Taylor, D. Cammarano, A. Prashar, & A. Hamilton (Eds.), *Proceedings of the 11th European Conference on Precision Agriculture*. *Advances in Animal Biosciences ECPA 2017* (Vol. 8, no. 2, pp. 150–155).

Al-Saddik, H., Simon, J. C., & Cointault, F. (2017b). Development of spectral disease indices for 'Flavescence Dorée' grapevine disease indentification. *Journal of Sensors*, 17, 1–25.

Apan, A., Held, A., Phinn, S., & Markley, J. (2003). Formulation and assessment of narrow-band vegetation indices from EO-1 hyperion imagery for discriminating sugarcane disease. *Spatial Sciences*, 1–13.

Araujo, M., Kawakami, T., Galvao, R., Yoneyama, T., Chame, H., & Visani, V. (2001). The succesive projection algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems, 57*(2), 65–73.

Ben-Hur, A., & Weston, J. (2010). A user's guide to support vector machines. *Data Mining Techniques for the Life Sciences, Methods in Molecular Biology, 609,* 223–239.

Bertamini, M., Grando, M. S., & Nedunchezhian, N. (2003). Effect of phytoplasma infection on pigments, chlorophyll-protein complex and photosynthetic activities in field grown apple leaves. *Biologia Plantarum, 47*(2), 237–242.

Blackburn, G. A. (1998). Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *International Journal of Remote Sensing, 19*(4), 657–675.

Blanchfield, A. L., Robinson, S. A., Renzullo, L., & Powell, K. S. (2006). Can leaf pigment composition help us identify grapevines infested with phylloxera? *Functional Plant Biology, 33,* 507–517.

Chuche, J., & Thiéry, D. (2014). Biology and ecology of the *Flavescence dorée* vector scaphoideus titanus: A review. *Agronometry for Sustainble Development, 34,* 381–403.

Devadas, R., Lamb, D., Simpfendorfer, S., & Backhouse, D. (2009). Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. *Precision Agriculture, 10*(6), 459–470.

Dubbini, M., Pezzuolo, A., De Giglio, M., Gattelli, M., Curzio, L., Covi, D., et al. (2017). Last generation instrument for agriculture multispectral data collection. *AgricEngInt: CIGR Journal, 19*(1), 87–93.

Gitelson, A., Kaufman, J. & Merzlyak, N. (1996a). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment,* 58(3), 289–298.

Gitelson, A. & Merzlyak, N. (1996b). Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. *Journal of Plant Physiology,* 148, 494–500.

Gitelson, A., & Merzlyak, N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing, 18*(12), 2691–2697.

Gitelson, A., Merzlyak, N., & Chivkunova, O. B. (2001). Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochemistry and Photobiology, 74*(1), 38–45.

Gitelson, A., Stark, R., Grits, U., Rundquist, D., Kaufman, Y., & Derry, D. (2002). Vegetation and soil lines in visible spectral space: a concept and technique for remote estimation of vegetation fraction. *International Journal of Remote Sensing, 23*(13), 2537–2562.

Himeno, M., Kitazawa, Y., Yoshida, T., Maejima, K., Yamaji, Y., Oshima, K., et al. (2014). Purple top symptoms are associated with reduction of leaf cell death in phytoplasma-infected plants. *Scientific Reports, 4,* 1–7.

Huang, J., Liao, H., Zhu, Y., Sun, S., Sun, Q., & Liu, X. (2012). Hyperspectral detection of rice damaged by rice leaf folder (Cnaphalocrocis medinalis). *Computers and Electronics in Agriculture, 82,* 100–107.

Huang, W. J., David, W. L., Niu, Z., Zhang, Y. J., Liu, L. Y., & Wang, J. H. (2007). Identification of yellow rust in wheat using in situ spectral reflectance measurements and airborne hyperspectral imaging. *Precision Agriculture, 8*(4–5), 187–197.

Jacquemoud, S., & Ustin, S. (2001). Leaf optical properties: a state of the art. In *Proceedings of 8th International Symposium Physical Measurements & Signatures in Remote Sensing* (pp. 223–232). Aussois, France: CNES.

John, D. M., John, C. R., William, T. S., & Leslie, R. C. (2007). Chlorophyll loss caused by soybean aphid (Hemiptera: Aphididae) feeding on soybean. *Journal of Economic Entomology, 100,* 1657–1662.

Khawas, E. S. A. M., & Khawas, E. M. A. M. (2008). Interactions between Aphis gossypii (Glov.) and the common predators in eggplant and squash fields, with evaluating the physiological and biochemical aspects of biotic stress induced by two different aphid species, infesting squash and cabbage plants. *Australian Journal of Basic and Applied Sciences, 2,* 183–193.

Kooistra, L., Leuven, R. S. E. W., Wehrens, R., Nienhuis, P. H., & Buydens, L. M. C. (2003). A comparison of methods to relate grass reflectance to soil metal contamination. *International Journal of Remote Sensing, 24*(24), 4995–5010.

Laudien, R., Burcky, K., & Doluschitz, R. (2005). Development of a field-based management information system (MIS) for sugar beet. *Sugar Industry, 129*(8), 565–567.

Lisa, D. F., Andrea, R. G., Tiffany, M. H., Leon, G. H., Gautam, S., & John, D. B. (2007). Physiological and biochemical responses of resistant and susceptible wheat to injury by Russian wheat aphid. *Journal of Economic Entomology, 100,* 1692–1703.

Lopez, M., Bertolini, E., Olmos, A., Caruso, P., Gorris, M. T., Llop, P., et al. (2003). Innovative tools for detection of plant pathogenic viruses and bacteria. *International Microbiology, 6*(4), 233–243.

Lu, J., Zhou, M., Gao, Y., & Jiang, H. (2017). Using hyperspectral imaging to discriminate yellow leaf curl disease in tomato leaves. *Precision Agriculture, 19*(3), 1–16.

Luedeling, E., Hale, A., Zhang, M., Bentley, W. J., & Dharmasri, L. C. (2009). Remote sensing of spider mite damage in California peach orchards. *International Journal of Applied Earth Observation and Geoinformation, 11*(4), 244–255.

MacDonald, S., Staid, M., Staid, M., & Cooper, M. (2016). Remote hyperspectral imaging of grapevine leafroll-associated virus 3 in cabernet sauvignon vineyards. *Computers and Electronics in Agriculture, 130*(15), 109–117.

Mahlein, A. K., Steiner, U., Dehne, H. W., & Oerke, E. C. (2010). Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. *Precision Agriculture, 11*(4), 413–431.

Main, R., Cho, A., Mathieu, R., O'Kennedy, M., Ramoelo, A., & Koch, S. (2011). An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*(6), 751–761.

Maverick, J.B. (2015). *The 4 Countries that Produce the Most Wine*. Retrieved July 22, 2018, from https://www.investopedia.com/articles/investing/090915/4-countries-produce-most-wine.asp.

Mirik, M., Michels, G. J., Jr., Kassymzhanova-Mirik, S., & Elliott, N. C. (2007). Reflectance characteristics of Russian wheat aphid (Hemiptera: Aphididae) stress and abundance in winter wheat. *Computers and Electronics in Agriculture, 57*(2), 123–134.

Murugesan, N., & Kavitha, A. (2010). Host plant resistance in cotton accessions to the leaf hopper Amrasca devastans (Distant). *Journal of Biopesticides, 3,* 526–533.

Naidu, R., Perry, E., Pierce, F., & Mekuria, T. (2009). The potential of spectral reflectance technique for the detection of Grapevine leafroll-associated virus-3 in two red-berried wine grape cultivars. *Computers and Electronics in Agriculture, 66*(1), 38–45.

Naidu, R., Rowhani, A., Fuchs, M., Golino, D., & Martelli, G. P. (2014). Grapevine leafroll: a complex viral disease affecting a high-value fruit crop. *Plant Disease, 98,* 1172–1185.

Penuelas, J., Baret, F., & Filella, I. (1995a). Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica, 31*(2), 221–230.

Penuelas, J., & Filella, I. (1998). Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends in Plant Science, 3,* 151–156.

Penuelas, J., Filella, I., & Gamon, J. A. (1995b). Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytologist, 131*(3), 291–296.

Penuelas, J., Gamon, J. A., Fredeen, A. L., Merino, J., & Field, C. B. (1994). Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sensing of Environment, 48*(2), 135–146.

Richardson, A. D., Aikens, M., Berlyn, G. P., & Marshall, P. (2004). Drought stress and paper birch (*Betula papyrifera*) seedlings: effects of an organic biostimulant on plant health and stress tolerance, and detection of stress effects with instrument-based, noninvasive methods. *Journal of Arboriculture, 30,* 52–61.

Riedell, W. E., & Blackmer, T. M. (1999). Leaf reflectance spectra of cereal aphid-damaged wheat. *Crop Science, 39*(6), 1835–1840.

Rinnan, A., Berg, F., & Engelsen, S. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry, 28,* 1201–1222.

Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., & Plümer, L. (2010). Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture, 74*(1), 91–99.

Seelig, H. D., Hoehn, A., Stodieck, L., Klaus, D., Adams, W., III, & Emery, W. (2009). Plant water parameters and the remote sensing R_{1300}/R_{1450} leaf water index: controlled condition dynamics during the development of water deficit stress. *Irrigation Science, 27*(5), 357–365.

Sims, A., & Gamon, A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment, 81*(2–3), 337–354.

Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., et al. (2014). A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles. *Remote Sensing, 6,* 11014–11030.

Tan, Y., Wei, H. R., Wang, J. W., Zong, X. J., Zhu, D. Z., & Liu, Q. Z. (2015). Phytoplasmas change the source-sink relationship of field-grown sweet cherry by distributing leaf function. *Physiological and Molecular Plant Pathology, 92,* 22–27.

Trotter, G. M., Whitehead, D., & Pinkney, E. J. (2002). The photochemical reflectance index as a measure of photosynthetic light use efficiency for plants of varying foliar nitrogen contents. *International Journal of Remote Sensing, 23*(6), 1207–1212.

Tucker, C. J. (1979). Red and photographic infrared linear combination for monitoring vegetation. *Remote Sensing of Environment, 8*(2), 127–150.

Tucker, C. J., Holben, B. N., Elgin, J., James, H., McMurtrey, I., & James, E. (1981). Remote sensing of total dry-matter accumulation in winter wheat. *Remote Sensing of Environment, 11,* 171–189.

Underwood, E., Ustin, S., & DiPietro, D. (2003). Mapping nonnative plants using hyperspectral imagery. *Remote Sensing of Environment, 86*(2), 150–161.

Vanegas, F., Bratanov, D., Powell, K., Weiss, J., & Gonzalez, F. (2018). A novel methodology for improving plant pest surveillance in vineyards and crops using UAV-based Hyperspectral and spatial data. *Sensors, 18*(260), 1–21.

Vitali, M., Chitarra, W., Galetto, L., Bosco, D., Marzachi, C., Gullino, M. L., et al. (2013). Flavescence doree phytoplasma deregulates stomatal control of photosynthesis in *Vitis vinifera*. *Annals of Applied Biology, 162,* 335–346.

Welling, M. (2014). *Fisher Linear Discriminant Analysis*. Retrieved July 23, 2018, from https://www.ics.uci.edu/~welling/teaching/273ASpring09/Fisher-LDA.

Yang, X., Hong, H., You, Z., & Cheng, F. (2015). Spectral and image integrated analysis of hyperspectral data for waxy corn seed variety classification. *Sensors, 15*(7), 15578–15594.

Yuan, L., Huang, Y., Loraamm, R., Nie, C., Wang, J., & Zhang, J. (2014). Spectral analysis of winter wheat leaves for detection and differentiation of diseases and insects. *Field Crops Research, 156,* 199–207.

Zarco-Tejada, P. J., Miller, J. R., Mohammed, G. H., & Noland, T. L. (2000). Chlorophyll fluorescence effects on vegetation apparent reflectance: Leaf-level measurements and model simulation. *Remote Sensing of Environment, 74,* 582–595.

Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H., & Sampson, P. H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 39,* 1491–1507.

Zhang, J., Pu, R., Huang, W., Yuan, L., Luo, J., & Wang, J. (2012). Using in situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. *Field Crops Research, 134,* 165–174.

Zhang, Y., Tan, L., Shi, H., & He, Y. (2013). Successive projections algorithm for variable selection on the rapid and non-destructive classification of coolant. *International Journal of Digital Content Technology and its Applications, 7,* 386–394.