Early detection of the fungal disease "apple scab" using SWIR hyperspectral imaging

Nathalie Gorretta[1], Maroua Nouri[2], Ana Herrero[3], Aoife Gowen[3], Jean-Michel Roger[1]

[1] ITAP, Univ Montpellier, Irstea, Montpellier SupAgro, Montpellier, France
[2] UMR1208 IATE, INRA, CIRAD, Montpellier SupAgro, Univ Montpellier, Montpellier, France
[3] UCD School of Biosystems and Food Engineering, University College of Dublin (UCD), Belfield, Dublin 4 Ireland

ABSTRACT

The aim of our study is to examine the potential of SWIR hyperspectral imaging to early detect apple scab infection. Close range hyperspectral images of healthy and infected leaves were acquired daily under laboratory conditions from 2 days to 11 days after inoculation using a push-broom SWIR camera. A PLS-DA classification model was built at the advanced infection stage D11 and was applied on the infected and healthy leaves images acquired at others infection stages. This study showed that good predictions can be achieved when classifying infected leaf regions based on hyperspectral data using PLS-DA. Results suggest that the spectral domain between 1000 - 2500 nm is suited to early differentiation between infected and healthy leaves. At early infection stages, the water absorption band at 1940 nm has the major discriminatory effect.

Index Terms— Disease detection, hyperspectral imaging

1. INTRODUCTION

Many invasive plant disease and pests have spread to different parts of the world causing serious losses in yield and affecting the quality of cultivated plants [1]. Their treatment is most often based on repeated fungicide applications which not only causes potential risks to the environment and ecosystems but also is a costly approach [2]. Traditionally, assessment of diseases and stresses are based on visual inspection and diagnostics methods. These destructive methods are however time consuming and labor intensive and not allow early disease detection.

Light attenuation within the plant results from complex absorption and diffusion processes influenced by the biochemical composition and morphological characteristics of leaf tissues. Biophysical and biochemical modifications induced by stress will therefore directly influence the reflectance of the leaf obtained using spectral technologies.

The present study focused on the early detection of a serious fungal apple disease "apple scab infection" on leaves using hyperspectral imaging in SWIR spectral range.

2. MATERIAL AND METHODS

2.1. Samples

Ten seedlings of apple grown in pots, were used, i.e. eight infected and two untreated plants. The cultivar selected was Golden Delicious. It is a cultivar moderately susceptible to apple scab infection. Infected plants were inoculated through spray-inoculation with a suspension of 5.10^{-5} conidiospores per milliliter using a commercial hand sprayer. The suspension was obtained by rinsing spores of infected leaves. Untreated plants were kept healthy and sprayed with water. Plants were then placed in optimal incubation conditions with 100% relative humidity, 18°C temperature and total obscurity for 48 hours similarly to [3] and [4] to boost fungus development. Afterwards, the plants were placed in a greenhouse controlled environment (RH= 80% and T= 18°C) (Figure 1). Plants were irrigated and fertilized as required. Inoculated plants were assessed daily by visually rating the development of scab symptoms by an expert. On the last day of the experiment, inoculated leaves were detached from the plant. Then, they were examined by the expert under the magnifying glass in order to identify the scab positions.

2.2. Imaging system and acquisition

Close range hyperspectral images of healthy and infected leaves were acquired daily from D02 to D11 after inoculation. The leaf was held over a magnetized plate and fixed using thin iron bars. The acquisitions were made on intact leaves, attached to the plant. Images of leaves were acquired under laboratory conditions using a HySpex SWIR-320m-e camera (Norsk Elektro Optikk, Norway). This camera sampled the reflected radiation in 256 bands from 960 to 2490 nm, with a 6 nm spectral sampling interval. The

camera was placed at 30 cm above the samples and acquired successive scans of 356 pixels, with a resulting spatial resolution of approximately 0.287 mm. A halogen lamp focused on the viewing zone of the sensor was used to illuminate the sample with an incident angle of 40° relatively to the vertical direction. A 99 % diffuse reflectance reference panel (Spectralon®, Labsphere) was placed horizontally at the same distance from the sensor and next to the leaves to estimate the incoming irradiance while limiting possible saturation of the sensor.

2.3. Data processing

2.3.1. Image Processing

Reflectance images were computed by dividing the signal measured over the target by that measured over the reference panel and multiplying by the reflectance value provided by Labsphere. The image background was removed by image processing in order to keep only vegetation-related pixels for further analysis. Images acquired over the same leaf and over whole acquisition period (i.e. day 2 to day 11 after inoculation) were then registered. This registration was performed using a registration estimator application implemented in MATLAB (MATLAB R2017a, the MathWorks Inc., Natick, MA).

2.3.2 Discriminant model

In order to build a discriminant model a calibration set and a test set were set up: among the eight infected leaves, seven ones were selected randomly for the calibration set and one for the test set. Among the two healthy leaves, one was used for the calibration set and one for the test set. On each infected leaf, at stage D11, infection spots were extracted and the mean spectrum was computed as the average of all the pixels of the spots. On each healthy leaf, regions were randomly chosen. The spectra were preprocessed by using a logarithmic transformation. This procedure produced a calibration set (Xcal, Ycal) and a test set (Xtest, Ytest) with respectively 34 and 16 spectra (Figure 1). The matrices Ycal and Ytest contained the membership degree for each class: [1 0] for healthy, [0 1] for infected.

A discriminant model was constructed by using partial least squares-discriminant analysis (PLS-DA) as defined in [5]. The number of latent variables of the PLS-DA was chosen according to the leave-one-out cross-validation error.

The best PLS-DA model was then applied on the images acquired from D02 to D11 after inoculation.

Analyses were performed using MATLAB R2017a (the MathWorks Inc., Natick, MA).

2.3.3 ANOVA analysis

In order to identify the characteristics of the spectra most involved in the separation of the two classes (healthy, infected), an ANOVA analysis was carried out on 3 parameters extracted from spectra: slope between 1000 nm - 1400 nm and between 1400 nm - 2500 nm and peak height at 1940 nm (corresponding to the main features in the PLS-DA regression vector). The height of the peak was measured as the value of the second derivative of the spectra, computed by means of the Savitsky and Golay algorithm. For each day, an ANOVA was calculated on each parameter independently, regarding the separation of infected and healthy spectra.

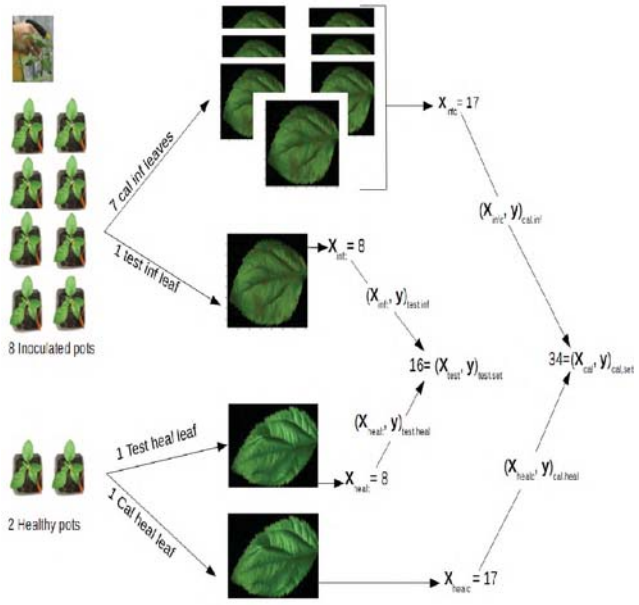

Figure 1: set up of calibration and test sets (heal: healthy leaf, inf : infected leaf).

2. RESULTS

3.1 Spectra analysis

Figure 2 shows healthy (green) and infected (red) mean class spectrum. Healthy and infected spectra have a similar overall shape. However, spectra show different slopes between 1400 and 2500 nm. On the one hand, the mean infected class spectrum has a more strongly negative slope than the healthy one in the 1000 - 1400 nm spectral domain. On the other hand, the spectrum of the healthy class has a more positive slope than the spectrum of the infected class in 1400-2500 nm spectral domain. Both mean spectra are

characterized by two peaks located respectively at 1450 nm and 1940 nm.

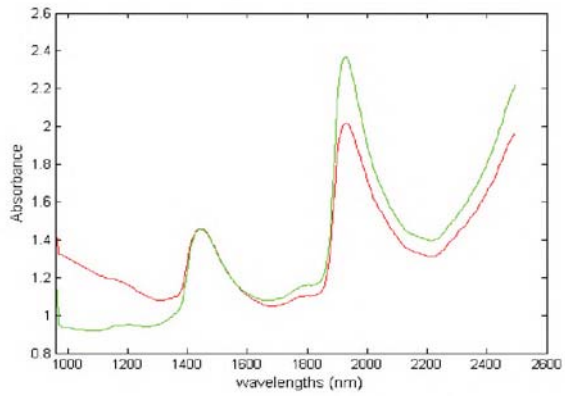

Figure 2: Healthy (green) and infected (red) mean spectrum.

These two peaks are characteristic of the water absorption regions. The mean spectrum of the healthy class shows a higher absorbance peak than the spectrum of the infected class at 1940 nm.

3.2 Discriminant model

According to the cross-validation error, a two latent variables PLS-DA model was built and used. By applying this model on the test set (D11) a perfect prediction with zero percent of prediction error was yielded. This first result indicates that good predictions can be obtained using a simple model with few latent variables. This also confirms the potential of the NIR spectroscopy as an interesting tool to detect plant diseases.

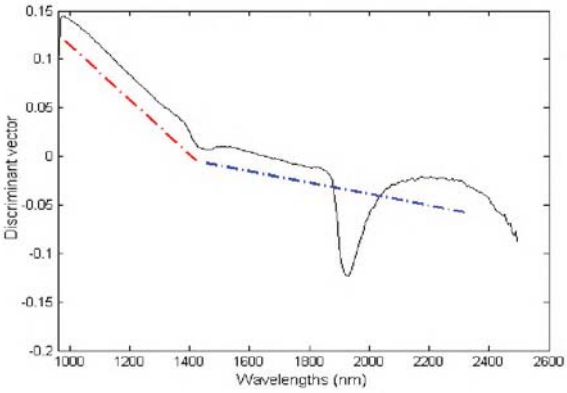

Figure 3: Discriminant vector of PLS-DA model

The discriminant vector is shown in Figure 3. Discriminant features depicted in section 3.2 to discriminate healthy and infected spectra are clearly visible: a slope between 1000-1400 nm (dash red line) and between 1400-2500 nm (dash blue line), a deep absorption peak at 1950 nm.

Changes in slope are known to be related to the physical structure of constituents (particle size, refractive index, roughness). Thus, it can be put forward that the differences in slopes observed are related to differences in physical structure between the cells of the infected and healthy zones. Peaks are linked to the chemistry of constituents. The 1940 nm spectral band is known as a water absorption band. Therefore, it can be put forward that there is a difference of water content or of water status between the infected zones and the healthy ones. The mean spectra in Figure 2 indicate a lower absorption of NIR light for infected leaves, indicating there is a lower water content.

3.3 PLS-DA model on the hyperspectral image series

The regression vector from the PLS-DA model was applied to the images acquired from D02 to D11 belonging to the image test set. Figures 4 and 5 show the scores images produced by the PLS-DA model respectively on one infected and one healthy leaf.

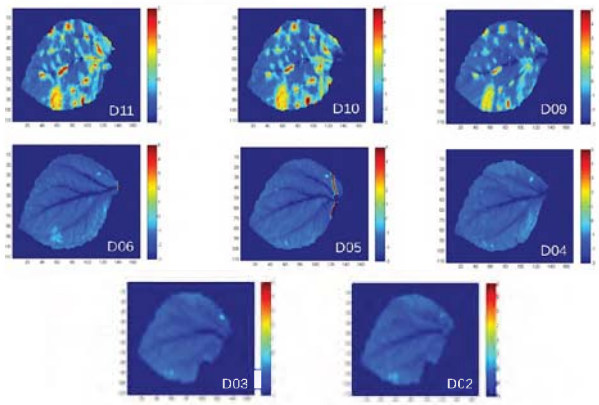

Figure 4: scores images of one infected leaf from D02 to D11 after inoculation. Higher values of scores (ie -1 to 3 or light blue to red color) correspond to infected pixels and negative values (-3 to -1 or dark blue color) to healthy pixels.

In Figure 4 and particularly in the score image at D11, sharp spots can be observed. The positions of the spots predicted by the model correspond to the positions of scab spots identified by the expert. In Figure 4, two disease development phases can be distinguished: a development phase between D02 and D06 and another one between D09 and D11, where the disease spreads rapidly.

It can be noted that the scab spots appear suddenly during the second stage of development (D09) whereas only a few

small spots are visible during the first phase. On day 9 (D09) after inoculation, scab spots also become visible to the naked eye. This result could be explained by the development cycle of the fungus: the first phase corresponds to the primary infection phase where the fungus is invisible to the naked eye. It grows under the cuticle without killing the host cells until sporulation. Sporulation triggers the second phase of infection known as the secondary infection. During the secondary infection, the scab spots become visible to the naked eye thanks to the conidiophores produced by the secondary fungus stromata [3], [4].

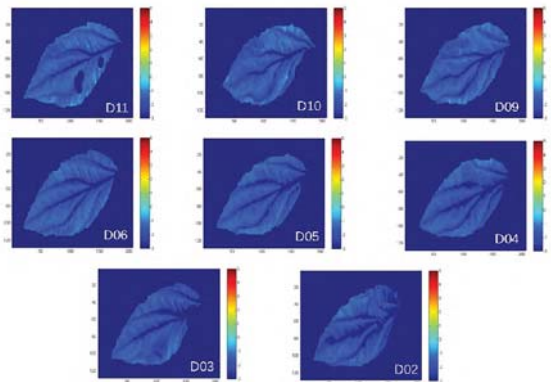

Figure 5: scores images of one healthy leaf from D02 to D11 after inoculation.

Score images obtained on a healthy leaves between D02 to D11 do not show any scab spot detected which confirms the good selectivity of the PLS-DA model (Figure 5).

The results mentioned above on an infected leave show that a model built at advanced infection stage D11 gives good results on early infection stages (D02 to D06). This result is somewhat surprising. As explained above, the discrimination mainly relies on three features (two slopes and one peak). To explain why the D11 model works on early stages, an ANOVA was calculated on each of the three features.

3.4. ANOVA exploration

According to the ANOVA results (Table 1), for each advanced infection stage (D09 to D11), the slope differences and the peak height differences between infected and healthy spectra are significant (p-value<0.05). For early infection stages (D02 to D06), only the difference in peak height at 1940 nm is always significant (p<0.05). The difference in slopes between the infected and healthy spectra is not always significant.

Statistical results prove that at early infection stages, apple scab infection affects the water content, or the water status of infected zones but not necessary the physical cell structure of the leaves.

DAI	Slope 1000-1400 nm	Slope 1400-2500 nm	Peak Heigh
D11	$8.05\ 10^{-6}$	$5.90\ 10^{-3}$	$8.25\ 10^{-5}$
D10	0.0183	0.0293	$8.25\ 10^{-5}$
D09	$2.99\ 10^{-5}$	$3.26\ 10^{-8}$	$1.17\ 10^{-7}$
D06	0.647	0.0135	$0.7\ 10^{-3}$
D05	2.1010^{-4}	$0.8\ 10^{-3}$	$0.3\ 10^{-3}$
D04	$0.9\ 10^{-3}$	0.6364	10^{-4}
D03	$5.99\ 10^{-5}$	$1.80\ 10^{-4}$	$1.50\ 10^{-4}$
D02	0.2564	$1.1\ 10^{-4}$	10^{-4}

Table 1: ANOVA results – P-value for each selected features

6. CONCLUSION

The present research is a baseline study which emphasizes the potential of SWIR hyperspectral imaging in early apple scab detection. Results suggest that the spectral domains between 1000- 2500 nm proved to be an effective region to separate infected and healthy leaves. Apple scab infection affects the physical structure and the water of infected leaves at advanced infection stages. At early infection stages, apple scab infection primarily affects the water content or the water status of infected leaves. The spectral band at 1940 nm has the most discriminatory effect a few days after infection.

7. REFERENCES

[1] E. Ballini, E., Nguyen and J.B. Morel. Diversity and genetics of nitrogen-induced susceptibility to the blast fungus in rice and wheat. Rice, 6, 32. 2013.

[2] ZHANG, M., QIN, Z., LIU, X. & USTIN, S. L. 2003. Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing. International Journal of Applied Earth Observation and Geoinformation, 4, 295-310.

[3]E.C. Oerke, P. Rohling, and U. Steiner. Thermographic assessment of scab disease on apple leaves. Precision Agriculture, 12, 699-715, 2011.

[4] S. Delalieux, W.J. Van AardtVAN AARDT, W., Keuleulemans, E. Schrevens and P. Copping. Detection of biotic stress (Venturia inaequalis) in apple trees using hyperspectral. dataEuropean Journal of Agronomy, 27, 130-143, 2007.

[5] M. Barker, and W.Rayens, Partial least squares for discrimination. Journal of Chemometrics: A Journal of the Chemometrics Society, 17(3), 166-173, 2003.