Unravelling Plant-Pathogen Interactions: Proximal Optical Sensing as an Effective Tool for Early Detect Plant Diseases

Mafalda Reis-Pereira, Rui C. Martins, Anibal Filipe Silva, Fernando Tavares, Filipe Santos, and Mário Cunha

Abstract: This study analyzed the potential of proximal optical sensing as an effective approach for early disease detection. A compact, modular sensing system, combining direct UV–Vis spectroscopy with optical fibers, supported by a principal component analysis (PCA), was applied to evaluate the modifications promoted by the bacteria Xanthomonas euvesicatoria in tomato leaves (cv. cherry). Plant infection was achieved by spraying a bacterial suspension (10^8 CFU mL^{-1}) until run-off occurred, and a similar approach was followed for the control group, where only water was applied. A total of 270 spectral measurements were performed on leaves, on five different time instances, including pre- and post-inoculation measurements. PCA was then applied to the acquired data from both healthy and inoculated leaves, which allowed their distinction and differentiation, three days after inoculation, when unhealthy plants were still asymptomatic.

Keywords: plant disease detection; plant pathology; proximal sensing; spectroscopy; precision agriculture; principal component analysis

1. Introduction

Biotic agents, specifically pests and pathogens, cause significant losses in crop yields, with levels that can range between 20% and 40% [1]. Chemical phytosanitary products are usually applied to prevent and combat these organisms. However, their usage can negatively impact the environment, mainly when applied to treat plant diseases that appear suddenly and spread to large scales [2].

Nowadays, phytopathology methods are considered major challenges because, to be implemented, they often rely on the presence of indicator visible signs of the infection (disease symptoms), which frequently only manifest themselves at the middle to late stages of the process, compromising the effectiveness of phytosanitary measures [3]. An example is the scouting technique, which involves inspecting a crop field to detect and identify infected plant through disease symptoms [4]. Despite being extremely useful, this approach requires specialized trained observers (who must be capable of identifying disease symptoms and distinguishing them from those caused by other abiotic stresses (e.g., nutritional and physiological disorders)), and can be labor-intensive, time-consuming, and expensive [5–11]. Moreover, this approach can be an inefficient in the early stages of the infection and on large areas. Other strategies consist of laboratory-based techniques,
namely serological and molecular tests, largely used due to their sensitivity, accuracy, and effectiveness. They include enzyme-linked immunosorbent assay (ELISA) and polymerase chain reaction (PCR) methods, being the first serological approach based on protein in the detection of causative diseases and the second molecular technique based on the DNA sequence of the pathogen. Their development boosted plant disease diagnosis, since they allow the simultaneous processing of several samples and perform a precise pathogen identification. Furthermore, PCR enables the detection of pathogens that have not been cultured. Nevertheless, these procedures present some limitations, especially in the early phase of the infection process, due to the uneven spread of pathogens inside plants, compromising their effectiveness in analyzing asymptomatic samples \[9,12–14\]. Other drawbacks can also be enumerated. They require several hours to be completed, require the realization of detailed sampling procedures, and destructive sample preparation, not allowing a follow-up of the disease progression \[12,13\].

Therefore, the necessity of developing fast, accurate, and selective in vivo techniques for plant disease detection arises. These innovative approaches must provide complementary information to the current methods applied in the phytopathology field and combine with them. Several non-invasive methods have been developed in the last decade, which proved to be sensitive, consistent, standardize, rapid, cost-effective, and have high-throughput \[15\]. Hyperspectral spectroscopy (HS) is one of them and seems to be effective in estimating a wide variety of plant chemical, biophysical, and metabolic traits in living tissue \[16–22\], namely foliar structure, plant chemical composition, water concentration, and metabolic status \[23\]. Through spectral measurements in the visible (Vis, 400–700 nm), near-infrared (NIR, 700–1100 nm), and shortwave infrared wavelengths (SWIR, 1100–2500 nm), this approach assesses changes in optical properties of leaves, which derive from interactions between light, chemical bonds, and cellular structure \[24\]. Briefly, modifications in plants’ reflectance in the Vis range are mostly related to pigment concentration and physiological processes, such as photosynthesis. In turn, changes in the NIR are correlated with leaf structure and internal scattering processes. The SWIR region is affected by leaf structural and chemical composition (including lignins and proteins), as well as water content \[25–29\].

Since phytopathogens induce physiological, biochemical, and structural changes in host plants, HS seems to be promising in plant disease detection, identification, and quantification \[30–38\]. Hyperspectral sensors can be used alone or mounted in different platforms, allowing the performance of mapping, monitoring, scouting, and application tasks \[2\]. Their flexibility allows them to assess leaf, single-plant, canopy (proximal sensing), and even plot and regional scales (remote sensing) \[2\]. Some examples, sorted by measurement scale, include handheld sensors, rail systems, vehicle, and tractor-mounted systems, drones UAVs, as well as aircrafts and satellites \[39\].

Despite the possibilities provided by these optical devices for simple, rapid, non-destructive disease detection and identification, its application is still very limited, due to the scarcity of extensive agronomic and phytopathological studies aiming to explore their full potential. Their technology readiness levels (TRL) are close to TRL3 (analytical and experimental critical function, and/or characteristic proof-of-concept) \[40\]. Hence, this study aimed to evaluate the potential of UV–Vis spectroscopy to detect diseased tomato leaves and discriminate between healthy and infected leaves, through a multi-temporal approach. Furthermore, the capability of this technology in detecting changes in the reflectance spectrum of infected leaves was analyzed, before the first symptoms became visible.

2. Materials and Methods

2.1. Experimental Design

Tomato (Solanum lycopersicum L.) plants, of the cultivar cherry, were grown in 200 mL pots containing a commercial potting substrate, in a walk-in plant growth chamber under controlled conditions (temperature of 25–27 °C, humidity of approximately 60%, and
photoperiod of 12/12 h). Plants were divided into two groups, one of them being inoculated with *Xanthomonas euvesicatoria* LMG 905 (Xeu) bacteria and the other being treated with sterile distilled water only (control group, Con). Plants were inoculated in the laboratory, at the growth stage of 5–6 fully expanded leaves, by spraying until they became fully wet, and run-off occurred. The bacterial suspensions used for these inoculation assays consisted of $1 \times 10^8$ cells/mL. They were prepared from a 48 hour-old culture, grown on YDC medium (yeast extract, 10.0 g; dextrose, 20.0 g; CaCO$_3$, 20.0 g; agar, 15.0 g; distilled water up to 1.0 L). The inoculated plants were then covered with transparent polythene bags for 48 h to increase the relative humidity that fosters bacterial entry into plant tissues through natural openings, such as stomata [41]. Plants were monitored daily for symptom development for 5 days.

At the same time, to verify if the bacteria cultures used in these inoculation tests were viable, 20 µL of Xeu solution were cultured in different Petri dishes containing YDC media. After 48 h, it was possible to observe the bacteria growth in both nutrient media, proving that bacteria were viable at inoculation.

2.2. Spectral Measurements

Hyperspectral data were collected in vivo from the adaxial side of healthy and infected tomato plant leaves, using a compact benchtop system consisting of a D2 (deuterium) light source (Ocean Optics, model DH-2000-BAL, Ostfildern, Germany), spectrometer (Ocean Optics, model HR4000, Ostfildern, Germany), transmission optical fiber bundle (UV), and stainless-steel slitted reflection probe for sample measurement. The spectrometer operated in the 195–1100 nm wavelength range, with a high spectral response and optical resolution of 0.025 nm (full width at half maximum—FWHM). The measurements were carried out using an experimental setup in the laboratory. A LED light source was placed beneath the leaf and provided homogeneous illumination to its entire surface. The light signal from the sample analyzed was guided to the entrance lens of the spectrometer by the fiber-optic cable placed perpendicularly 1 cm above the measured surface. Specialized software was used for data acquisition and processing. Data acquisition was performed with 10 scans for an integration period of 60 ms, in three leaves per plant, on nine locations on each leaf.

2.3. Data Pre-Processing

Spectral pre-processing was performed, in order to remove possible artifacts, e.g., baseline shifts, Mie and Rayleigh scattering, and stray light. Also, a pretreatment with a fast fourier transform (FFT) was carried out on spectral data to smooth/denoise it. FFT is an algorithm that computes the discrete Fourier transform (DFT) of a sequence or its inverse (IDFT). Fourier analysis converts a signal from its original domain (often time or space) to a representation in the frequency domain and vice versa. The DFT is obtained by decomposing a sequence of values into components of different frequencies [42]. Spectral data pre-processing was performed with RStudio software.

2.4. Data Processing—Analytical Techniques

Spectral data was subjected to a principal component analysis (PCA), a multivariate data analysis technique was used to reduce the dimensionality, while preserving its structure by projecting it into a new coordinate system. This technique allows the preservation of the total variance of the dataset and minimizes the mean square approximate errors. PCA uses eigenvectors and eigenvalues to define the reduced subspace (representing the original coordinate system). It originates principal components (PC), which are linear combinations of interrelated variables. PC1 accounts for the maximum possible proportion of the variance information of the original dataset (explained by the eigenvalue), and subsequent principal components (PC2, PC3, . . . ) account for the maximum proportion of the unexplained residual variance, and so forth [43,44].

Contiguous hyperspectral wavebands present redundant information [45]. The application of a PCA allows the transformation of this type of high-dimensional data into a few
wavebands that contain most of the information in the original bands. The importance of these hyperspectral bands in each PC is then established based on the magnitude of eigenvectors or factor loadings for crop biophysical and biochemical traits, being that the higher the eigenvector, the higher is the importance of the band. So, PCA allows the selection of the best wavebands to model biophysical and biochemical quantities and the elimination of redundant bands (by highlighting the main bands) [46].

3. Results

The spectral response properties of tomato leaves to the stress caused by *Xanthomonas euvesicatoria* LMG 905 is very important for discriminating bacterial infection levels in precise pest management using hyperspectral proximal sensing data. The averaged raw spectral curves of healthy and diseased tomato leaves were slightly different in some spectral ranges, namely through the visible region of the wavelength spectrum (~480–680 nm) (please refer to Figure 1) Similarly, the spectral measurements assessed on infected leaf tissue presented a decrease in signal intensity throughout the sampling period (24–144 h), which accompanied the appearance of the first visual symptoms of the disease after 72 h.

![Figure 1](image-url)

**Figure 1.** Spectral measurement curve evolution for tomato leaves inoculated with bacteria *Xanthomonas euvesicatoria* LMG 905, within the sampling period (24–144 h). Leaf spectral curves were assessed in vivo on the adaxial side of fully expanded leaves, on the spectral region from 195 to 1100 nm.

Figure 2 presents the principal components (PC) Gabriel plot for the healthy (Con) and diseased (Xeu) leaves spectra, three days after inoculation (before the appearance of the first symptoms). The PCA algorithm has obtained two PCs, accounting for 99.6% of the total variance. PC1 (94.3%) discriminates the effects on the variance of these two types of tomato leaves, which is more evident in PC2 (5.3%).
Figure 2 presents the principal components (PC) Gabriel plot for the healthy (Con) and diseased (Xeu) leaves spectra, three days after inoculation (before the appearance of the first symptoms). The PCA algorithm has obtained two PCs, accounting for 99.6% of the total variance. PC1 (94.3%) discriminates the effects on the variance of these two types of tomato leaves, which is more evident in PC2 (5.3%).

The wavelengths that have a higher contribution in these PC are in the interval of ~454–654 nm (visible range of the wavelength spectrum). The ones between ~492–510 nm (essentially the blue region of the electromagnetic spectrum) explain 30% of the variance of the PC1, whereas ~454–461 nm (blue region) explain 40% of the variance of the PC2 and 50% of the PC3. In all the first four dimensions of this analysis, the wavelengths ranging from approximately 445–480 nm (blue) and 580–700 nm (red) were the ones that explain most of the variance of the data.

This evidence can be related to the symptoms caused by Xeu, since these bacteria cause small, brown, angular lesions on leaves (which can be surrounded by a yellow halo with time), affecting the levels of photosynthetic pigments (contributing especially to the reduction of the chlorophyll levels, whose absorption features are more evident in the blue and red ranges of the Vis spectral region), cellular content, and structural arrangement.

4. Discussion

The spectral behavior of tomato plants depends on their biochemical and structural profile. In brief, plants’ spectral signature in the visible spectral region (400–700 nm) depends mainly on the content of photosynthetic pigments. These compounds are good absorbers of red and blue wavelengths. Of the major pigments, Chlorophyll a (Chl a) has maximum absorption in the 410–430 and 600–690 nm regions, whereas Chlorophyll b
(Chl b) has maximum absorption in the 450–470 nm range. In healthy plants, chlorophyll concentration is approximately ten times higher than that of other pigments, thus masking out the specific absorption features of these compounds. The green part of the spectrum, on the other hand, is less strongly absorbed, resulting in a reflectance peak in the green domain (at about 550 nm) [25]. Hence, when a light source illuminates healthy plants, they will preferentially absorb red and blue wavelengths, being the green part of the incident light less absorbed and, consequently, more reflected, leading to their green appearance [26].

In the NIR region, the plants’ spectral response is related to their structure, structural components, and internal scattering processes. Likewise, the SWIR region is also affected by leaf structural and chemical composition (including the action of lignin’s and proteins) and water content [25–29].

Since phytopathogens cause changes in plants’ biochemical and structural composition, affecting the levels of photosynthetic pigments and structural elements, tracking changes in plants’ spectral behavior can allow an indirect analysis of their phytosanitary status. Generally, unhealthy plants have more reflection in the red region and lower reflectance in the NIR region. Briefly, stress usually promotes an increase of reflectance over the whole spectrum, since it causes a rapid decrease of chlorophylls, which increases reflectance in the Vis range and exposes the absorption characteristics of other pigments, such as carotenoids (responsible for the yellowing of the leaves) and xanthophylls (responsible for the reddening of the leaves). With continuing stress, leaf structures decompose, resulting in extra intra-leaf scattering and an increased NIR signal. At the same time, concentrations of brown pigments, which absorb radiance in the Vis and at the onset of the NIR, can increase leading to a flattening of the red edge. Absorption in the SWIR decreases, due to reduced leaf moisture. With a decay of the leaf tissue, the absorption features characteristics of healthy plants gradually disappear [47].

Our findings seem to be in accord with the previous information, showing evidence that UV–Vis spectroscopy can be suitable for plant disease assessment in laboratory conditions. Data collected in a randomized experimental design, combined with a PCA, allowed the discrimination of healthy and diseased tomato leaves, even at the third day after bacteria inoculation, when no visual symptoms were observable. Most of the variance of the data can be comprised with the first four PCs. In all of them, the wavelengths that explain most of the variance of the data ranged from approximately 445–480 nm (blue) and 580–700 nm (red), which was expected, since Xanthomonas euvesicatoria causes tissue lesions, degrading the chlorophylls levels and affecting their absorption features in these spectral regions.

Therefore, our results can be related to those obtained in different research, where sensor-based approaches proved to be capable of assessing modifications in plants’ spectral behavior, allowing the detection, identification, and quantification of different types of plant diseases [44,48–51]. They involve the capture and analysis of the optical properties of plants, within different regions of the electromagnetic spectrum and their relationship with modifications in plant physiology, namely alterations in tissue color, structural composition, and transpiration rate [19]. These non-invasive methods have been explored in the last decade, presenting the benefits of being sensitive, consistent, standard, high-throughput, rapid, and cost-effective [47], surpassing the limitations of the current methods used in plant disease detection.

5. Conclusions

The present study suggests that UV–Vis spectroscopy can be a potential tool for the early detection of plant diseases under laboratory conditions, even when unhealthy plants are asymptomatic. Despite these findings, its application is still very limited, due to the scarcity of comprehensive agronomic and phytopathological studies aiming to explore their full potential, as well as the development of applied advanced statistical approaches for data analysis. More research is necessary, especially in field conditions, where more external factors have to surpass, including atmospheric, edaphic, and biotic conditions.
Future research should also include more stress levels to discriminate not only healthy leaves from the diseased ones but also different levels of disease severity.


**Funding:** Mafalda Reis-Pereira and Aníbal Filipe Silva were supported by fellowships from Fundação para a Ciência e a Tecnologia (FCT), with the references SFRH/BD/146564/2019 and DFA/BD/9136/2020, respectively. Rui C. Martins acknowledges Fundação para a Ciência e Tecnologia (FCT) research contract grant (CEEIND/017801/2018). This research was supported by the project ‘SpecTOM—Metabolomics Tomography Spectroscopy System’, University of Porto, Fundação Amadeus Dias, and Santander-Universities Grant.

**Acknowledgments:** Mafalda Reis-Pereira and Aníbal Filipe Silva were supported by fellowships from Fundação para a Ciência e a Tecnologia (FCT), with the references SFRH/BD/146564/2019 and DFA/BD/9136/2020, respectively. Rui C. Martins acknowledges Fundação para a Ciência e Tecnologia (FCT) research contract grant (CEEIND/017801/2018). This research was supported by the project ‘SpecTOM—Metabolomics Tomography Spectroscopy System’, University of Porto, Fundação Amadeus Dias, and Santander-Universities Grant.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


