Computers and Electronics in Agriculture 211 (2023) 107947

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Including measurement effects and temporal variations in VIS-NIRS models to improve early detection of plant disease: Application to *Alternaria solani* in potatoes

Florent Abdelghafour [a,*,1], Sajeevan Radha Sivarajan [b,1], Ingi Abdelmeguid [d], Maxime Ryckewaert [a,c], Jean-Michel Roger [a,c], Ryad Bendoula [a], Erik Alexandersson [b]

[a] ITAP, University Montpellier, INRAE, Institut Agro, 34196 Montpellier, France
[b] Department of Plant Protection Biology, Swedish University of Agricultural Sciences, SE-23422, Lomma, Sweden
[c] ChemHouse Research Group, 34196 Montpellier, France
[d] Department of Botany and Microbiology, Faculty of Science, Helwan University, EG-11795, Cairo, Egypt

ARTICLE INFO

Keywords:
NIRS
Chemometrics
EPO-PLS
Plant disease
Early detection
Measurement effect
Potato
Alternaria solani

ABSTRACT

Early detection of plant diseases with automated, non destructive and high-throughput techniques is a major objective in plant breeding and crop protection. Near infrared spectroscopy and hyperspectral imaging are proven to be particularly relevant technologies. However, robust discriminant models remains a challenge because of the many uncontrolled sources of variability during the experiment. Indeed, at early stages of most diseases, the temporal variations due to environment and measurement effects can induce signal shifts of greater magnitude than the infection itself, masking the information of interest. Excluding the variations of the measurement environment and the temporal fluctuation of the plant-pathogen interaction can depreciate the model robustness. Here, the problem is addressed in a study of the seven potato cultivars monitored for the presence of early blight disease at 0, 18, 36, and 96 h after inoculation. Three practical corrections are proposed regarding the effect of temporal fluctuations. (i) subclass effect, (ii) kinetic effect of the disease, and (iii) measurement effect. Eventually, the application of EPO-PLSDA to orthogonalise the model regarding temporal variation to produce invariant models proved to be the only suitable and well-performing of the tested solutions. With this approach the disease can be detected from 36 h after inoculation for 6 of the 7 tested cultivars. Classification errors differ among the cultivars but on average are below 25% of errors.

1. Introduction

Spectroscopy and hyperspectral imaging (HSI), especially in the visible and near infrared (VIS–NIR) domain, are very popular automated and non-destructive techniques for the detection of numerous plant diseases (Mahlein et al., 2018). When these technologies are combined with appropriate analysis and modelling methods, it is possible to discriminate infected plants from healthy ones, to predict infection levels and stages, or even to discriminate between different diseases (Mas Garcia et al., 2021; Lowe et al., 2017; Mahlein et al., 2012; Couture et al., 2013). In addition, spectral data can be used to deduce changes in leaf chemistry or structure (Jacquemoud and Ustin, 2019; Bos and Parlevliet, 1995). As a matter of fact, some characteristic spectral regions are well-known to be related to the concentration of pigments (chlorophyll, anthocyanins and carotenoids), nitrogen, sugar and water content or even leaf surface and tissue structure (Gold et al., 2020b). The emerging need for explainable models, consistent with domain knowledge (*e.g.* plant sciences) makes spectral data and its modelling tools (chemometrics) (Wold, 1995) especially relevant within the new trend of explainable artificial intelligence (Confalonieri et al., 2021; Streich et al., 2020).

Being a growing threat for tomato and potato, early blight disease has been recently more and more investigated, including by means of spectral measurements and hyperspectral imaging. Usually, such studies use *in-vivo* foliar measurements to monitor the disease during its progression. Studies use either spectral vegetative indexes (Atherton et al., 2017), multivariate analysis of continuous spectra (Gold et al., 2020a) or both (Van De Vijver et al., 2020), to discriminate the healthy plants from the infected ones at specific disease progression stages. The studies report a discriminant accuracy above 90% from two to four days post infection depending on the growing and inoculation

* Corresponding author.
E-mail addresses: florent.abdelghafour@inrae.fr (F. Abdelghafour), sajeevan.radha.sivarajan@slu.se (S.R. Sivarajan).
[1] Authors contributed equally to this work.

https://doi.org/10.1016/j.compag.2023.107947
Received 11 August 2022; Received in revised form 2 May 2023; Accepted 20 May 2023
Available online 13 June 2023

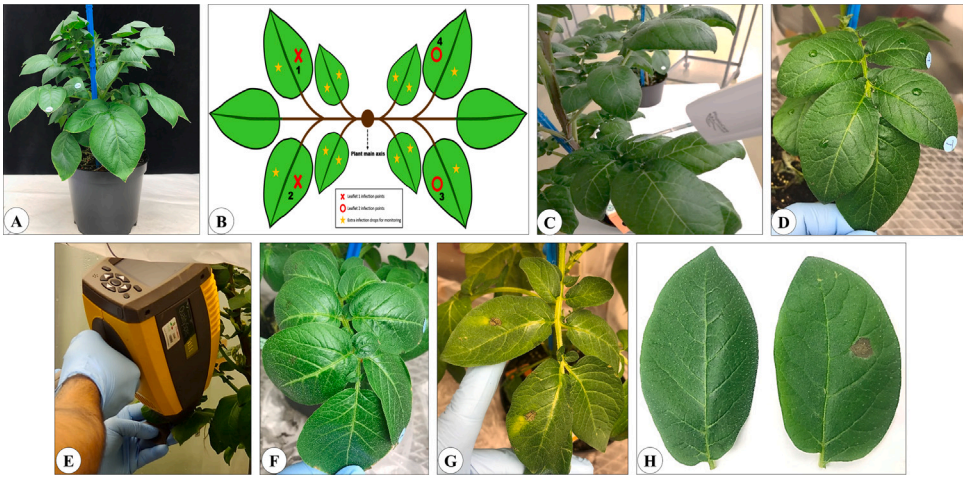


Fig. 1. Experimental setup: plant material, inoculation protocol, symptoms development and spectral measurement: a. photograph showing the six-week-old potato plant; b. schematic representation of the position of the inoculation in the leaflets (marked with red 1–4); c & d. photograph showing the inoculation and intact inoculum droplet post-inoculation, respectively; E. data measurement using the ASD QualitySpec Trek Handheld Spectrometer (NIR); f. leaflets showing the initial stage of development of early blight infection; g. leaflets showing the clear necrotic lesions of early blight infection; h. phenotype of the control and inoculated leaflets after 4 days.

conditions and the sensitivity of the cultivars. In the case of Van De Vijver et al. (2020), the results were obtained with field measurement, which can explain later detection threshold due to slower infection of the fungus. First signs of infection by *A. solani* are detected from differential changes in pigments, water content and cell damage (Bao and Zhang, 2023).

However, when monitoring individual plants or evaluating a field for the presence of a disease, measured signals evolve with time. The plant-pathogen interaction is a dynamic process, and even healthy plants are subject to uncontrolled natural physiological fluctuations. In addition, measured signals also depend on numerous fluctuations of the local environments. Therefore, temporal variations represent both a useful information and a source of detrimental effects for the modelling process. It provides information regarding the evolution or the kinetic of a disease and cultivar-specific interaction but it also represents a source of variability that can be assimilated to a "measurement effect". In this context, changes in signals induced by the infection, especially at early stages, can be masked by variabilities of greater magnitudes. Indeed, if some variations, due to light, temperature, or humidity, can be reduced or corrected by ensuring stable measurement environments and controlled growing conditions, measurement effects and temporal variation remain major obstacles to achieve reliable models have to be taken. Early detection of infection is crucial for plant phenotyping applications, for breeding, and crop monitoring. Although there are some examples in the existing literature, early detection of infections remains a challenge (Terentev et al., 2022; der Waals et al., 2001). Here, we investigated the interactions of *A. solani* with seven potato cultivars at early time points after inoculation. These cultivars have varying degree of susceptibility, and we know different types of transcriptome response to *A. solani* (Sajeevan et al., 2023).

In this study, a method to correct time induced variabilities is proposed to improve the robustness and performances of discrimination between healthy and infected plants. The problem is addressed with a case study of the monitoring of early blight disease during the initial hours after inoculation (HAI). *A. solani*, the agent responsible for the early blight disease, is a fungal pathogen of the phylum *Ascomycota*, infecting leaves, stems, fruits or tubers. It is a necrotroph that causes lesions and characteristic "bullseye" spots symptoms on the leaves. Eventually, leaves sustain yellowing (due to chlorosis and nitrogen deficiency) and then dry out, leading to premature defoliation. The tubers also sustain surface dryness and dark lesions while the flesh turns "leathery" and can even rot. *A. solani* can infect numerous species of the *Solanaceae* family other than potato, including tomato (*Solanum lycopersicum L.*), eggplant (*Solanum melongena*) and, bell pepper (*Capsicum spp.*) (Weir et al., 1998).

The purpose of this study is to investigate three practical correction methods of the time induced measurement effects and to compare them with the reference method PLSDA (Partial Least Square Discriminant Analysis) (Ruiz Perez et al., 2020):

(a) predicting the health status and the HAI altogether within a PLS2 model

(b) predicting the health status after internal linear correction (Wülfert et al., 2000) with a time constrained PLS2 model

(c) predicting the health status with EPO-PLSDA considering time (external parameter orthogonalisation — partial least square (Roger et al., 2003))

2. Material and methods

2.1. Plant growth conditions

In vitro grown potato (*Solanum tuberosum*) cultivars Aracy, Bintje, Désirée, King Edward, Kuras, Matilda, and Magnum Bonum were initially transferred in 0.5 L pots (9×9×9.5 cm) filled with potting mixture (Exklusiv Blom & Plantjord, Emmaljunga Torvmull AB, Sweden) for two weeks and later transferred to 2 L pots and allowed to grow for four more weeks (a representative picture of a six weeks old potato plant is presented in Fig. 1A). The plants were grown in an artificial light chamber of 160 µmol s^{-1} m^{-2} with 16 h light and 8 h dark regime (Biotron, SLU Alnarp, Sweden). The chamber was maintained at 25 and 22 °C day and night temperature, respectively, with relative humidity (RH) of 55%–60%.

2.2. Fungal culture preparation and inoculation

Actively growing mycelial discs (5 × 5 mm) of *A. solani* As112 strain Odilbekov et al. (2014) was placed on the centre of a 20% potato dextrose agar (PDA) medium plate and allow it to grow in dark at room temperature for 3 days. The plates were subsequently transferred for 8 h to an UV-C light supplying incubator (OSRAM HNS15G13 with dominant wavelength of 254 nm) with a temperature of 18 °C for inducing sporulation. The plates were incubated for 10 to 12 days and

Table 1
Datasplit: repartition of samples within potato cultivars, health status and time points (HAI).

Cultivar	Total	Health status		HAI			
		Control	Inoculated	0	18	36	96
Aracy	191	119	72	48	48	48	47
Bintje	271	172	99	68	68	67	68
Desiree	288	180	108	72	72	72	72
King Edwards	281	175	106	70	70	70	71
Kuras	271	169	102	68	68	67	68
Magnum Bonum	361	222	139	92	93	92	84
Matilda	345	219	126	86	86	86	87
Total	2008	1256	756	504	505	502	497

the conidia were harvested by flooding the plates with 10 to 15 mL of sterile tap water containing 0.01% (v/v) Tween 20 (Sigma Aldrich). The concentration of conidia was measured using a Fuchs Rosenthal haemocytometer and was adjusted to 25,000 conidia/mL for inoculation. Three different potato plants from each cultivar with two leaves each from the centre (22 to 25 days old) was used for inoculation. Fig. 1b presents the inoculation scheme and 1c shows the inoculation in practice and 1d shows the resulting intact droplets. For each infection, a 15 µl inoculum carrying 25,000 conidia/mL of As112 was placed on the adaxial side on either sides of mid rib towards the centre part of each leaflets leaving the first two and centre leaflets. For control plants (mock inoculation), 15 µL each of sterile water containing 0.01% (v/v) Tween 20 was used. Four different potato cultivars were kept in each trolley and covered with transparent plastic to maintain high humidity (>95%) and the lights were turned off for efficient infection. Measurements were taken at 0, 18, 36, and 96 HAI. Fig. 1f, g and h show respectively the initial state of the infection after 18 h, the evolution after 96 h and the comparison with the control samples.

2.3. Spectral acquisition

The spectral measurements were conducted with the handheld contact spectrometer ASD Qualityspec Trek®. Measures were directly performed on leaves, targeting inoculated areas for the infected plants (see Fig. 1e. The resulting data are a collection of reflectance spectra (Lindon et al., 2017) (one for each measured point in the spectral domain of VIS–NIR–SWIR (short wave infrared) [350–2500 nm], with a spectral resolution of ~1 nm. Then each spectrum is described by $p = 2151$ wavelengths

2.4. Database description

In total, the constructed database contains 2008 spectral samples measured on leaflets from 42 plants (7 cultivars × 3 replicates × 2 groups). For each plant, three, four or five leaflets were monitored depending on the cultivar. For example, for Aracy fewer leaflets were measured, hence the difference in the number of samples per cultivars that can be seen in Table 1 that details the number of measurements according to the cultivars, HAI and health status.

2.5. Data processing

2.5.1. Data split and models' evaluation protocol

The database is iteratively divided into calibration and validation sets. For each iteration, all the samples of one cultivar is picked out for validation while the remaining data is used for calibration. The cultivar split ensures the independence of the validation data by dissociating biological replicates. In addition, it enables to test the robustness of models calibrated on different cultivars and to compare models properties and performances of the different cultivars. Essentially, it is the same principle as "k-fold" cross-validation, where each fold is determined by the cultivar split.

2.5.2. PLSDA and EPO-PLSDA

Chemometrics is a reference approach in spectral analysis and integrative "omics" analysis (Worley and Powers, 2013; Worley et al., 2013; Rohart et al., 2017), particularly Partial least square (PLS) (Wold et al., 2001), which is used to model the relationship between explanatory variables **X** (*e.g.* spectral data) and variable of interests **Y** (*e.g.* health status of samples). PLS is essentially a multilinear analysis, producing predictive models. It is particularly adapted to strongly collinear (correlated), noisy and high dimensional **X** data (especially when the dimension of the variables p is superior to the number N of samples) (Jobson, 1991). PLS enables to achieve at once dimensionality reduction, variable selection and regression. PLS can also be applied for discriminant problems, such as predicting semantic classes (*e.g.* healthy, infected, and response to abiotic stress), in which case it is called PLS-DA (discriminant analysis). In practice it consists in combining a discriminant analysis, generally the Fischer's linear discriminant analysis (LDA, also known as FDA).

External Parameters Orthogonalisation PLS (EPO-PLS) was proposed by Roger et al. (2003) to calibrate robust models that are invariant to a given detrimental variation effect. Originally it was proposed to deal with the effect of the temperature of measurement for the prediction of sugar content in apples. In the context of discriminating between control and inoculated samples, it is proposed to summarise the various "measurement effects" using the HAI as a combined proxy. Then EPO-PLS is applied by considering the HAI similarly as it is conventionally done for temperature in Roger et al. (2003). Appendix presents a pseudo-algorithm detailing the practical implementation of EPO-PLS in this context.

The data processing framework then consists in comparing EPO-PLSDA with PLSDA as the reference method in the domain, in terms of discriminant performances, robustness and sensitivity regarding cultivars and time-points.

2.5.3. Evaluation criteria

Confusion Matrix and classification error:

A confusion matrix is a k by k array, that compares the predicted class and the actual class of samples, where k denotes the number of semantic classes or label of interests to discriminate. In this application $k = 2$ ["Control", "Inoculated"].

From this matrix, several performance metrics can be derived, such as the classification error, which represent for each class, the number of misclassified samples. (*i.e.* the number of control samples classified as inoculated and *vice versa*. It is equivalent to $1 - Recall$).

3. Results and discussion

3.1. Data visualisation

Fig. 2 shows balanced subsets of the spectral database which are coloured according to their health status (a), HAI (b) and cultivar (c). The reflectance spectra present the typical features of leaves in the VIS–NIR–SWIR domain Jacquemoud and Ustin (2019). Notable spectral bands can be observed, first for the pigments, around 480 nm for the carotenoids, 680 nm for chlorophyll then the "red-edge" after 780 nm

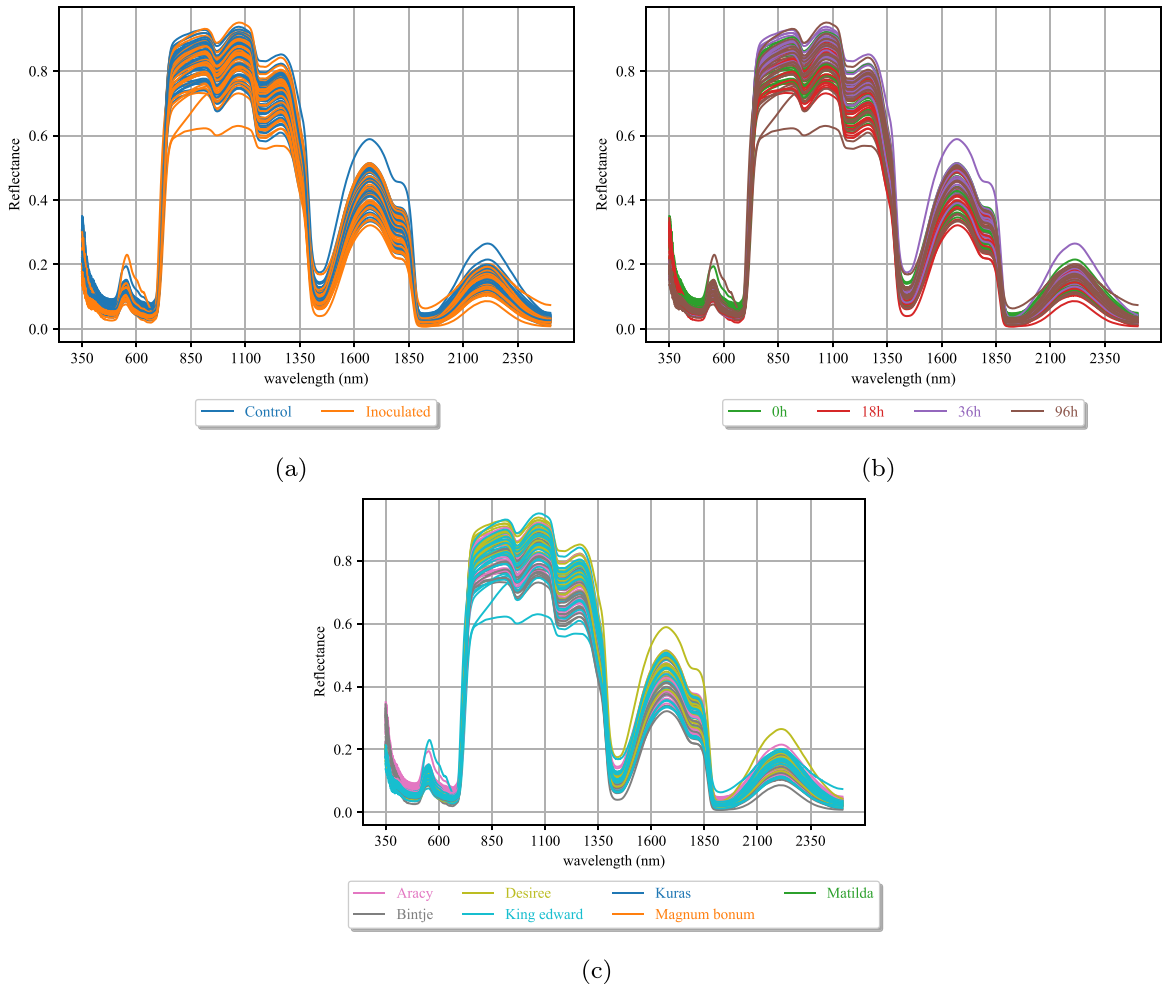


Fig. 2. Distribution of spectra according to the labels of interest: health status (a), HAI time-points (b) and cultivar (c).

and the water absorption bands around 980, 1480 and 1870 nm. There are no visible difference between the spectra of the different health statuses, HAI and cultivars.

Fig. 3 presents the score plot of the dataset projected on the 3rd and 4th component of a PCA (Principal Component Analysis), representing 11 and 9% of the explained variability, respectively. Spectral samples are coloured according to their HAI (a) and health status (b). The projection shows a relative separability of the data according to HAI on the 3rd and 4th PC's. No combination of axes can, however underline separability between the control and inoculated groups (example Fig. 3.b). Therefore, temporal (assimilated to a "measurement effect") variations represent here the largest variance and occludes the health status which represents differences of lesser magnitude between the spectra. The following parts aim at stating and illustrating the problem of the measurement effect that increases the complexity of the target problem: predicting health status.

3.2. The reference method: PLSDA for the discrimination of health status

Fig. 4 presents the discrimination performances achieved by the PLSDA method. Fig. 4.a present the classification errors (CE in %) cumulated for each of the seven cultivar splits, depending on the number of latent variables (LV). It illustrates a representative and generic behaviour for all the models calibrated for each of the targeted cultivars. Results are presented for both the "control" and "inoculated" group. A discriminant model is achieved when both CE curves for the two classes, reach a low value. At the first LV, all samples are classified in the control group (results are shown from the second LV). Then, the error curves decrease toward the joint minimum CE reached for 7 LVs, with respectively 31 and 32% of cumulated CE for the control and inoculated groups. After, the CE curves start to be unstable and diverge after the 13th LV, where the model is depreciated (form of overfitting) and tend to classify samples predominantly into the "inoculated" group. Fig. 4.b presents the confusion matrix for the selected model at 7 LV. This confusion matrix is also the cumulated classifications for each cultivar split. Overall, the performances are insufficient for applicative perspectives but still shows that the model can learn from a relevant information.

Table 2 shows the CE values per cultivar. Each cultivar is iteratively chosen as a validation set, when the remaining are used for the calibration. For most cultivars, the value of the evaluation criteria are very poor. Indeed, in most cases, the samples are predominantly attributed to one of the two classes and the cumulative assessment is only balanced because iterations tend to compensate each other. Only the performances of the models calibrated for King Edward and Kuras seem to bear some relevant information. Both King Edward and Kuras are early blight susceptible cultivars. This might have enhance faster changes in the spectral differences between control and inoculated compared to others cultivars considered moderately susceptible or partially resistant.

Table 3 presents the CE detailed according to the HAI. The performances are resulting from the same data split and models as before, but

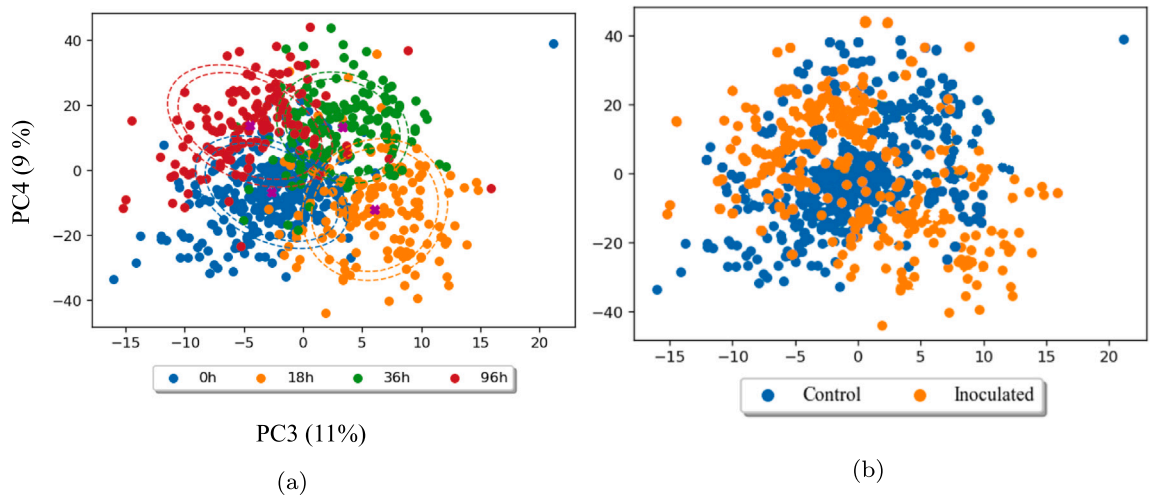


Fig. 3. PCA analysis: separability of the data according to HAI (a) and health status effects (b). The samples are projected on PC3 and PC4 representing respectively 11 and 9% of the explained variance.

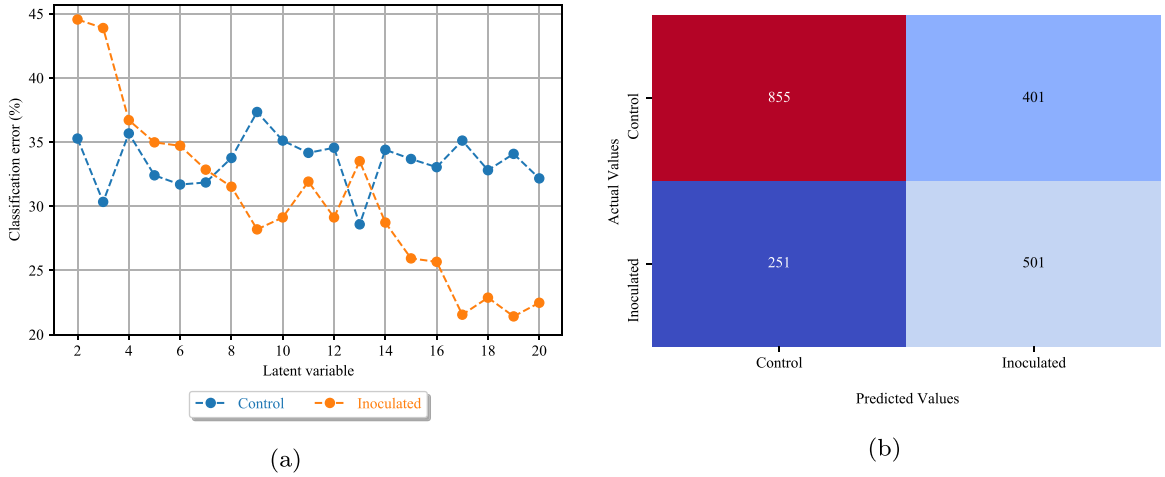


Fig. 4. Classification errors of the PLSDA method for the discrimination of control and inoculated samples, all cultivars and HAI considered: (a) cumulative classification error (%) per class for PLSDA, and (b) confusion matrix for 7 LV cumulated.

Table 2
Classification errors (%) for control and inoculated of PLSDA detailed by cultivar.

Cultivar	Control	Inoculated
Aracy	50	15
Bintje	29	38
Desiree	54	15
King Edwards	27	26
Kuras	23	30
Magnum Bonum	23	41
Matilda	24	54

Table 3
Classification errors (%) of PLSDA detailed by HAI.

Treatment	0 HAI	18 HAI	36 HAI	96 HAI
Control	4	81	48	20
Inoculated		64	20	15

are here detailed according to the HAI of samples. The control samples at 0 h and both classes for 96 h are fairly well-modelled. However, in the case of HAI = 18 or 36 h, the model is completely off-target.

3.2.1. *Conclusion about the discriminant potential of PLSDA*

The proposed problem is quite complex because it mixes different cultivars and measurement acquire at different key HAI. This situation is chosen to illustrate a case close to the real application. Indeed, when evaluating a field or greenhouse, models are calibrated in advanced with a subset of cultivars and measurement conditions. In addition, when testing a plant for the presence of a pathogen, it is impossible to have prior knowledge of the time since primary infection. In that context, the PLSDA is limited by the detrimental source of variabilities that are unrelated to the health status and achieves insufficient performances. However, some encouraging results for some cultivars or HAI lead to consider that PLS is a relevant "backbone" as modelling system that can integrate the time factor as a solving constraint.

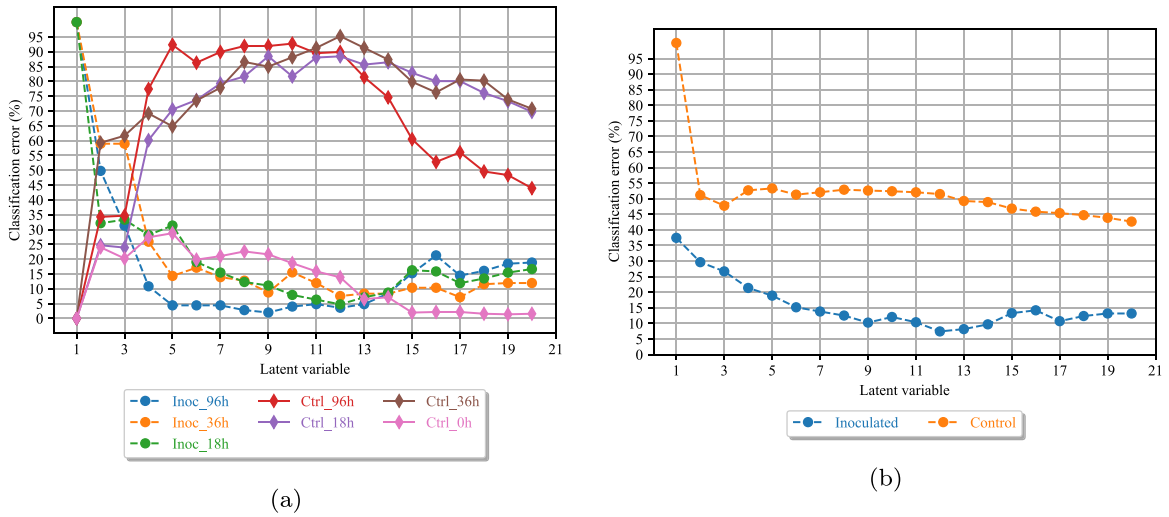


Fig. 5. Classification performances with the PLSDA-2 method considering time-points: (a) mixed classes classification cumulated CE, and (b) summarised binary classification cumulated CE.

3.3. PLSDA-2, predicting HAI and health status altogether

A first hypothesis regarding the limits of the PLSDA is the heterogeneity of the classes. Namely, each time point would correspond to a subclass of the inoculated group with varying stages and severity of the disease. Under this assumption, the approach to solve the problem, consists in modelling and predicting altogether health status and HAI with a PLSDA-2 (Stocchero et al., 2019).

Fig. 5.a presents the cumulative CE depending on the LV for crossed classes (health status and HAI) with the same cultivar split as presented in Section 3.2. The continuous curves with "diamond" markers are used for the control group, when the dashed curves with dot markers are used for the inoculated group. Fig. 5.b presents the evolution of CE considering a summarised binary classification where errors on HAI can be tolerated and only the health status is considered to compute performances. For example a control sample measured at 18 h that would be classified in the 36 h control group is considered as a valid classification. The PLSDA-2 method results in a fair modelling of inoculated samples and of the control group at 0 h. For the other HAI, control samples are predominantly classified as inoculated. This results in CE above 50% for the control group in the summarised binary classification.

Fig. 6 presents the confusion matrices obtained for the optimal model (*i.e.* with the joint minimum CE) at 13 LV (see Fig. 5.a). Fig. 6.a is the "semi-binary" confusion matrix, *i.e.* it shows the confusion matrix when predicting only the health status, but detailing the actual HAI and 6.b is the summarised binary confusion matrix. Fig. 6.c presents the binary CE detailed per HAI. Globally, the inoculated samples are well classified with CE below 10% while more than half of the control samples are mistakenly classified in the other group. This approach is completely ineffective and results in lower performance than the reference PLSDA approach. The conclusion is then that time represents indeed more a factor of disturbance than a subdivision of stages.

Fig. 7 presents the performances for the time-constrained method: tc-PLSDA. This method is derived from Wülfert et al. (2000) and consists in calibrating a model with both reference values for health status and HAI, but realising predictions only considering the regression coefficients regarding the health status part of the model. This process is essentially an implicit orthogonalisation performed within the PLS algorithm. Fig. 7.a is a comparison of the MCE between tc-PLSDA and the reference PLSDA method, depending on the number of LV. Fig. 7.b is the confusion matrix for 7LV (optimal model for PLSDA: see Fig. 8).

Table 4
Cultivar wise comparison of the classification errors for 7 LV PLSDA and tc-PLSDA models.

Cultivar	PLSDA		tc-PLSDA	
	Control	Inoculated	Control	Inoculated
Aracy	50	15	29	37
Bintje	29	38	23	17
Desiree	54	15	30	12
King Edwards	27	26	24	9
Kuras	23	30	25	21
Magnum Bonum	23	41	31	30
Matilda	24	54	27	29

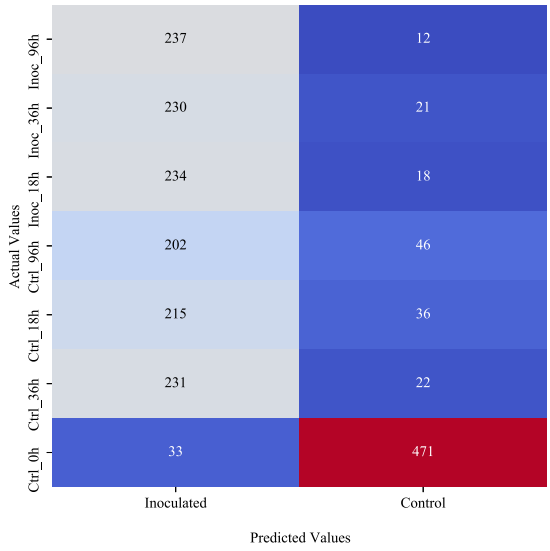
Table 5
Classification error of tc-PLSDA (%) regarding HAI for the 7 LV model.

Treatment	HAI			
	0	18	36	96
Control	12	54	33	25
Inoculated		22	31	13

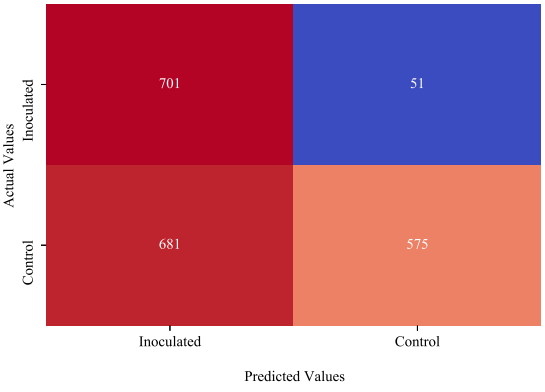
The tc-PLSDA does not seem to represent a considerable improvement compared to the PLSDA reference method. The MCE of tc-PLSDA is a little lower for the optimal model at 7 LV (see Fig. 8), where the MCE is 30% for tc-PLSDA compared to 32% for PLSDA. However, when looking at the results per cultivar, presented in Table 4, the model is actually relevant in this case, with the exception of the Aracy cultivar. Similarly, when looking at the detail per HAI presented in Table 5, the model is relevant from 36 h where the CE is respectively of 33 and 31% for the control and inoculated groups. Globally, the performances are still insufficient before 96 h for an application. However, it shows a major interest of constraining the PLS model with information related to the time, that acts here as a proxy for the measurement effect.

3.4. EPO-PLSDA: Explicit constraint

Fig. 8 presents a comparison of the discrimination performances between PLSDA and EPO-PLSDA. (a) and (b) present the evolution of CE for both classes according to the number of LV and (c) summarises the MCE for both classes. First of all, EPO-PLSDA seems to improve significantly the performances with a first minimum CE below 25% at 7 LV (optimum for PLSDA) and a second minimum CE below 20% at 19 LV. Secondly, the EPO-PLSDA seems to be able to learn a more complex



(a)

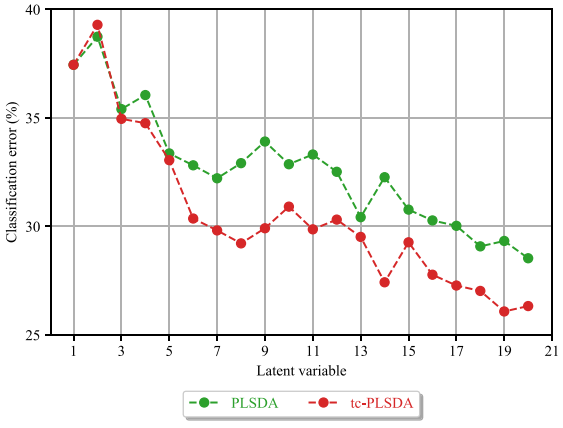


(b)

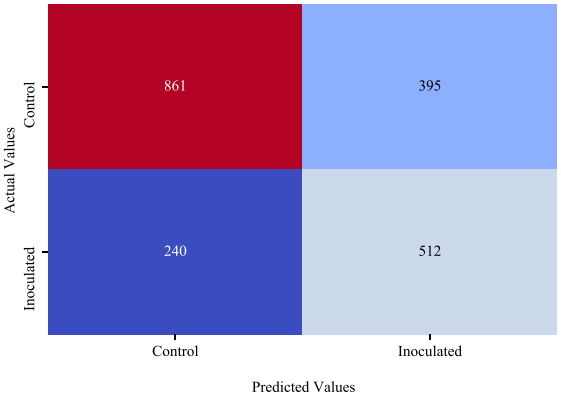
Treatment	HAI 0	18	36	96	all
Control	6	86	91	81	54
Inoculated		7	8	5	7

(c)

Fig. 6. Confusion matrix and classification errors (%) of PLSDA for 13 LV model. Semi-binary (a), binary (b), and binary classification error per HAI (c).



(a)



(b)

Fig. 7. Comparison of mean classification error between PLSDA and tc-PLSDA (a) and confusion matrix of tc-PLSDA at 7 LV (b).

model from the data, as shows Fig. 8.b, the two curves start to diverge after the 15th LV.

Fig. 9 compares the cumulated confusion matrix for the PLSDA (a) and EPO-PLSDA (b) at 7 LV and the resulting CE for both class (c). For the PLSDA model, the cumulative CE for both classes are over 30%, which is insufficient for applicative perspectives but still shows the relevance of both the information and the modelling method. In the case of the EPO-PLSDA, the results are significantly better with

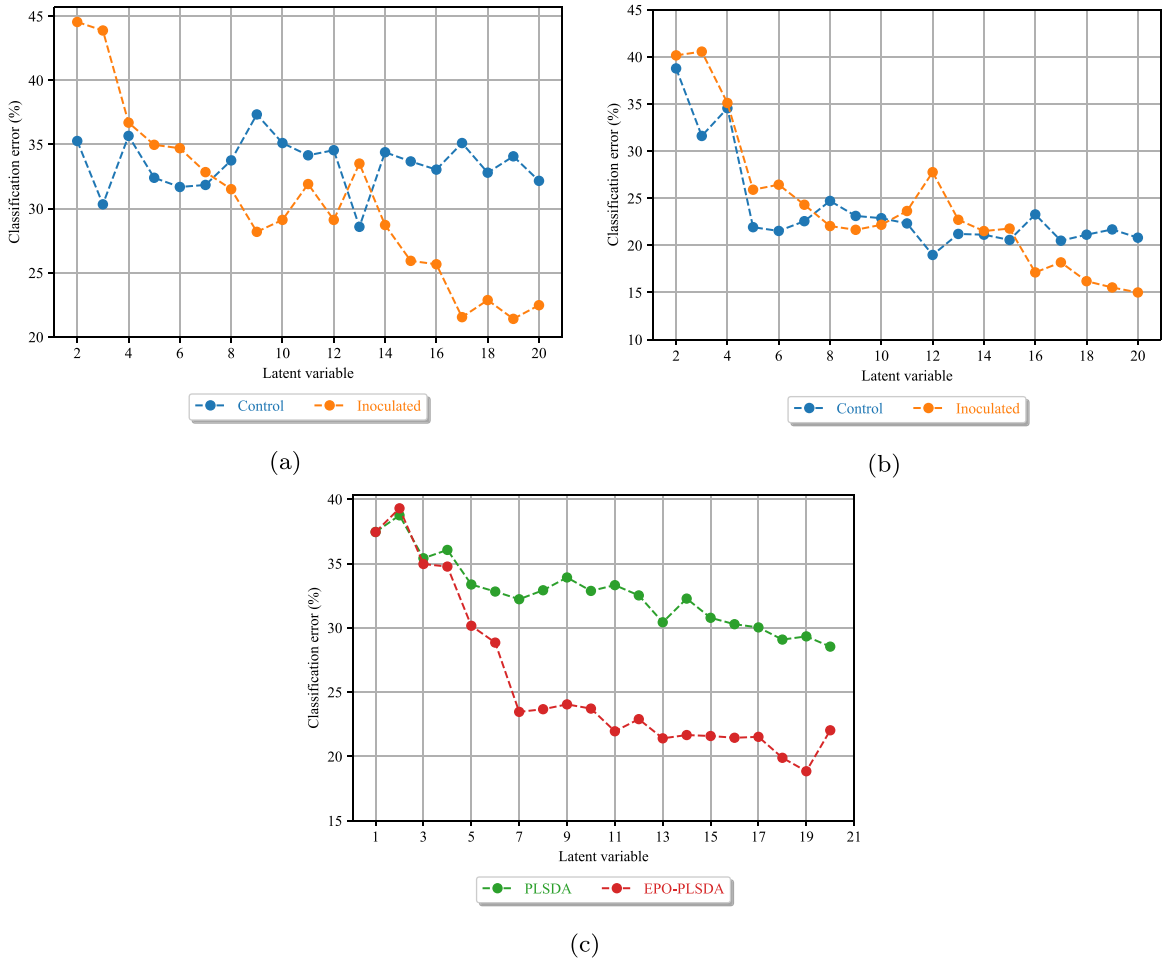


Fig. 8. Comparison of the evolution of the cumulative classification error (%) for PLSDA (a), EPO-PLSDA (b), models applied iteratively to each cultivar, according to number of LV. (c) presents the comparison of the mean classification errors for PLSDA and EPO-PLSDA.

26 and 20% of cumulative CE for the control and inoculated groups, respectively. It means that 80% of the inoculated samples are retrieved, while 26% of control samples are mistakenly classified as inoculated. This level of performances might remain limited for applications such as plant nursery sanitation, but shows a significant improvement with EPO-PLSDA compared to the reference PLSDA.

Table 6 presents the comparison of CE between both methods, regarding the targeted cultivar. It is essentially the detail of the cumulative CE for each iteration over the different calibration sets and their respective cultivar used for validation. For the EPO-PLSDA method, performances are quite satisfying for all cultivars except Aracy. As one of the most early blight tolerant cultivar known, there might be only minor differences between the inoculated and control plants at the very early time points. In the case of Desiree, for which the inoculated samples are well recalled ($CE = 11\%$) but 31% of the control samples are misclassified as inoculated, the model could be interesting for applications more tolerant to false positives.

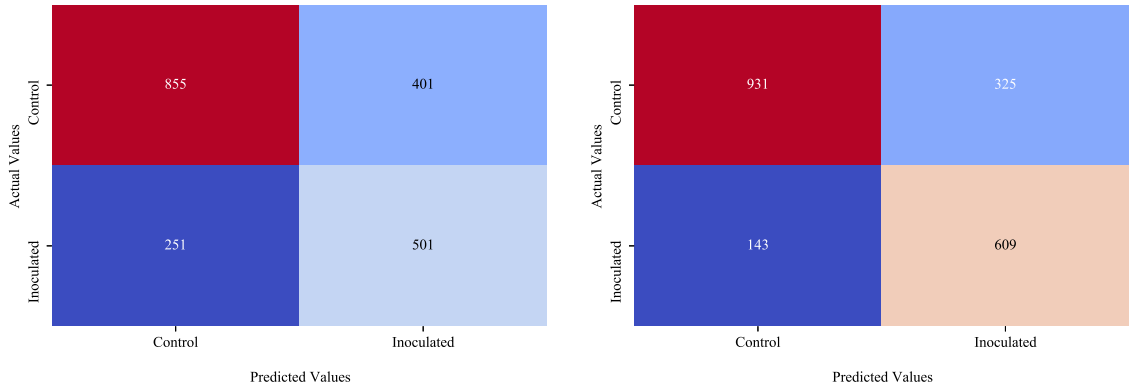
Table 7 presents the comparison of CE regarding the HAI. The performances are resulting from the same data split and models as before, but are here detailed according to the HAI of samples. In the case of the EPO-PLSDA method, CE are lower for 0 h and 96 h, However, compared to the PLSDA method, the model is still relevant at 36 h where the inoculated are well retrieved ($CE = 10\%$) but with a high false positive rate for control samples ($CE = 30\%$). It is the sensitivity — detection limit of the model which is not relevant at 18 h.

Table 6
Cultivar wise comparison of the classification errors for 7 LV PLSDA (a) and EPO-PLSDA (b) models.

Cultivar	PLSDA		EPO-PLSDA	
	Control	Inoculated	Control	Inoculated
Aracy	50	15	25	38
Bintje	29	38	21	11
Desiree	54	15	31	7
King Edwards	27	26	24	12
Kuras	23	30	23	19
Magnum Bonum	23	41	25	25
Matilda	24	54	25	23

3.5. Robustness and stability of the models calibrated with PLSDA or EPO-PLS: variability of the "b-coefficient"

Fig. 10 compares the stability — variability of the "b-coefficient" (regression coefficient) regarding the different cultivars considered for calibration with respectively (a) PLS and (b) EPO-PLS. These coefficients are estimated before the DA step of the process. The blue curves represent the mean b-coefficients, *i.e.* the average between the coefficients obtained for each iteration over cultivars. The orange areas represent the variability magnitude expressed with the standard deviation $\pm \sigma(b_\lambda)$. The EPO-PLS method present a substantially higher stability than the reference PLS method. It means that the produced model are more robust to changes in calibration data.

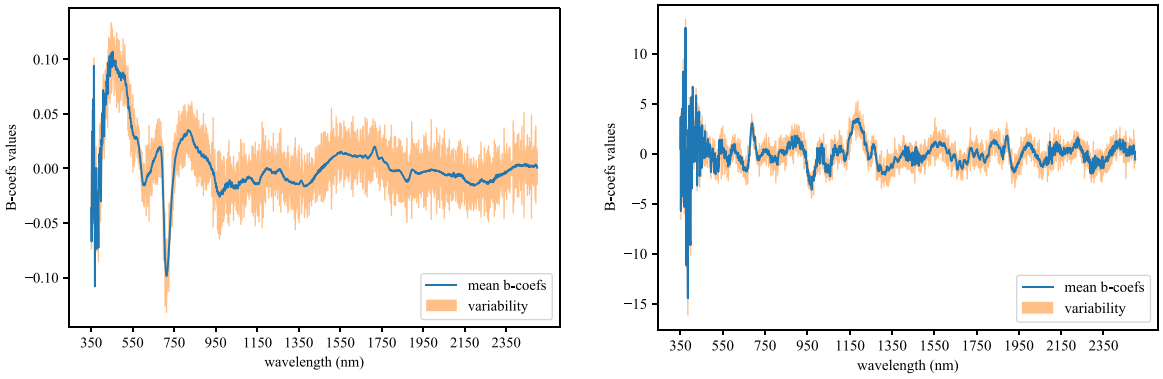


(a) (b)

CE per class (%)	PLSDA	EPO-PLSDA
Control	32	26
Inoculated	33	20

(c)

Fig. 9. Comparison of the cumulative confusion matrices resulting from 7 LV models applied iteratively to each cultivar, respectively PLSDA (a), and EPO-PLSDA (b), corresponding cumulative classification error per class (c).



(a) (b)

Fig. 10. Comparison of the variabilities of b-coefficients between PLS (a) and EPO-PLS (b) models. The blue curves represent the average b coefficients computed for all models. The orange areas mark the variance between the mean and each of the models specifically calibrated to predict one cultivar.

Table 7
HAI wise comparison of the classification errors for 7 LV PLSDA and EPO-PLSDA models.

	PLSDA				EPO-PLSDA			
Treatment	0 HAI	18 HAI	36 HAI	96 HAI	0 HAI	18 HAI	36 HAI	96 HAI
Control	4	81	48	20	8	63	30	20
Inoculated		64	20	15		27	10	19

Table 8 presents a comparison of the angles (in degrees) between the different b-coefficients calibrated for each model, respectively with PLS and EPO-PLS. This angle express how similar the obtained regression coefficient (~the model) when a part of the calibration data is substituted. Angles close to 0° describe similar models while angles close to 90° describe orthogonal, *i.e.* independent models. This table shows a relative stability of the models for EPO-PLS, Only the model for Aracy present significant dissimilarities with the other models. It means that, the process is quite robust to the variabilities within the considered cultivar, and that it is possible to calibrate a generic model

Table 8
Comparison of the variability of the b coefficients for PLS and EPO-PLS models. The values represent angles (°) between the b coefficient of the model calibrated for the prediction of each cultivar and the average b-coefficients of each model.

	Aracy	Bintje	Desiree	King Edwards	Kuras	Magnum Bonum	Matilda
PLS	37°	26°	23°	35°	21°	31°	18°
EPO-PLS	25°	6°	4°	16°	7°	11°	3°

on several cultivars to be applied for the prediction on other cultivars. In the case of PLS, the angles show that the different models are more "oblique" to one another. It means that while they preserve some correlation and common properties they are still significantly different. These observation are consistent with the differences in performance and classification behaviours between the two methods. The natural variabilities within the dataset, both due to samples (cultivar range) and to the signal (measurement effect), are detrimental for the PLS method that cannot reach a stable and well-performing model. When applying corrected models with EPO-PLS, the resulting performances and the rather stability of the models show that it is relevant to try a robust approach and that it enable to solve the application.

3.5.1. Interpretation of the model from the regressions "b-coefficients"

Even though the coefficient of the PLS seem easier to interpret, the coefficients of EPO-PLS bear additional information. The coefficients from the PLS are related to changes in pigment content, notably at 480 nm for the carotenoids, 530 nm for the anthocyanins and 680 nm, 730 nm, 860 nm for chlorophyll and photosynthetic systems (Merzlyak et al., 2003). In addition there are notable spikes around 1725 and 1820 nm that correspond respectively to CH_2 content (lipids, wax from cuticles) and cellulose. The coefficients of the EPO-PLS model include the spectral regions of interest previously cited and some additional information. Notable spikes around 1200 nm confirms the importance of lipid content (Osborne and Fearn, 1993). Furthermore, some usual areas around 950, 980, and 1940 nm related to water content appeared. Otherwise, regions around 2080 and 2242 nm respectively related to hydroxide/alcohol content, sucrose and amino-acids seem to be discriminant.

4. Conclusion & perspectives

When dealing with complex NIR applications such as the detection of plant diseases at early stages, the reference PLSDA method seems limited. Indeed external sources of variations related to measurement or growth effects can be detrimental to the calibration process and can result in poorly performing models, instabilities and lack of robustness. In this article, it was shown that a interesting process to overcome this issue is to constraint the model with *a priori* knowledge regarding plant physiology, pathogenesis and there interaction with spectroscopy. Eventually, the method EPO-PLS performed better than the implicit linear correction method. This work also showed that reference of time, here in the form of HAI values can be used as proxies to correct growth and measurement effects. The resulting models significantly outperform PLSDA and are more robust for different cultivars which are not included in the model. With EPO-PLS it is possible to detect potato early blight 36 hours after the inoculation of *A. solani*. This demonstration covered only one generic type of effect with two example of corrections which essentially consist in orthogonalising a known detrimental effect. It would be interesting to extend this work to other limiting/detrimental sources of external variability and propose other correction methods or constraints. For example, it is known that the plant physiology as well as the pathogens physiology can be influenced by numerous environmental factors that lead to the susceptibility or on the contrary the resistance of the host. In that context, it could be considered to study the effects of plant age, leaf position, metabolic content before infection or even compounds inducing plant resistance. A recent transcriptomic study (Sajeevan et al., 2023) captured major transcription factor families whose differential expressed genes are associated with resistance or susceptibility to early blight within potato cultivars. Associating these findings with hyperspectral imaging could improve the robustness of statistical models aiming at predicting health status and infection stages. Providing a new target category to predict that is more consistent with the dynamic of plant pathogen interaction (or proxy), would allow to include a form of constraint to improve the accuracy of predictive models. In addition, by investigating more cultivars with different types of defence mechanisms or resistance levels against a specific pathogen it might also be possible to train future models to identify specific types of interactions and also help characterise these based on the spectral signature. Regarding correction methods, statistical tools and especially chemometrics are rich in robust, transfer or domain invariant methods. It would be interesting to adapt them and compare with EPO-PLS. Finally, punctual spectroscopy cannot capture the spatial patterns of the disease, which could represent a valuable information. Reproducing this work from hyperspectral images could improve further the discrimination performance and robustness

CRediT authorship contribution statement

Florent Abdelghafour: Conceptualization, Methodology, Writing – original draft, Software, Formal analysis. **Sajeevan Radha Sivarajan:** Conceptualization, Methodology, Data curation, Writing – original draft, Laboratory work. **Ingi Abdelmeguid:** Laboratory work. **Maxime Ryckewaert:** Sofware, Validation, Writing – review & editing. **Jean-Michel Roger:** Writing – review & editing, Validation. **Ryad Bendoula:** Writing – review & editing, Supervision. **Erik Alexandersson:** Conceptualization, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. This work was part of the SLU Grogrund project 'Resistance breeding for healthy crops' with financial support from NordPlant (NordForsk grant no. 84597)

Data availability

Data will be made available at the end of the project. It is a python panda ready database with the spectral collection plus metadata (cultivars, hai, health status).

Acknowledgements

This work was part of the SLU Grogrund project 'Resistance breeding for healthy crops' with financial support from NordPlant (NordForsk grant no. 84597). We thank Johanna Wetterlind (SLU), for lending us the ASD QualitySpec Trek and Qiang Wang (UiT The Arctic University of Norway) for the initial help with data analysis.

Appendix A

Algorithm 1 Practical implementation of the EPO-PLS algorithm

Estimation of G and computation of its orthogonal basis

1: Compute the difference matrix of the "external factors"

$$D[k=4, p=2151] = \begin{pmatrix} \bar{X}-\bar{X}_{t=0} \\ \bar{X}-\bar{X}_{t=18} \\ \bar{X}-\bar{X}_{t=36} \\ \bar{X}-\bar{X}_{t=96} \end{pmatrix},$$

where $\bar{X}$ denotes the column-wise mean of X, *i.e.* the mean spectrum.

2: Perform singular value decomposition (SVD) of the covariance matrix of D

$[U, s, V] = SVD(D^{\mathrm{T}}.D)$,

where "T" denotes the transposition operator and . denotes the matrix multiplication operator.

3: Compute the "detrimental space" G

$G = D^{\mathrm{T}}.V$

4: Determine G_{opt} the optimal subset of G

$\hat{G}_{opt} \subset G$

5: Compute the vector orthonormal basis $Proj_{\perp[p,p]}$ to project X orthogonally on G_{opt}

$Proj_{\perp[p,p]} = Id_{[p,p]} - (G_{opt}.({G_{opt}}^{\mathrm{T}}.G_{opt})^{-1}.{G_{opt}}^{\mathrm{T}})$,

6: where $Id_{[p,p]}$ denotes the identity matrix of dimension $Proj_{\perp} = Id - [\hat{G}_{opt}.(\hat{G}_{opt})^{\mathrm{T}}.(\hat{G}_{opt}^{-\frac{1}{2}})^{\mathrm{T}})].\hat{G}_{opt}.[\hat{G}_{opt}^{\mathrm{T}}.\hat{G}_{opt})^{-\frac{1}{2}}]$

7: Project X orthogonally to G_{opt}

$X_{orth} = X.Proj_{\perp}$

8: Calibrate a robust PLS model invariant to the factors in D, with X_{orth}

dv = PLSDA.calibration(X_{orth}), so that:

for any given measured X_{new}, regardless the HAI, $\hat{Y} = dv.X_{new}$,

$\hat{Y}$ denotes the estimation of the sample class, from the data X_{new} and dv denotes the discriminant vector of the PLSDA model.

References

Atherton, D., Choudhary, R., Watson, D., 2017. Hyperspectral remote sensing for advanced detection of early blight (*Alternaria solani*) disease in potato (*Solanum tuberosum*) plants prior to visual disease symptoms. http://dx.doi.org/10.13031/aim.201701010.

Bao, H., Zhang, Y., 2023. A model for hyperspectral image monitoring and identification of tomato early blight incubation period. SSRN Electron. J. http://dx.doi.org/10.2139/ssrn.4360369.

Bos, L., Parlevliet, J.E., 1995. Concepts and terminology on plant/pest relationships: Toward consensus in plant pathology and crop protection. Annu. Rev. Phytopathol. 33 (1), 69–102. http://dx.doi.org/10.1146/annurev.py.33.090195.000441, PMID: 18288897.

Confalonieri, R., Coba, L., Wagner, B., Besold, T.R., 2021. A historical perspective of explainable artificial intelligence. WIREs Data Min. Knowl. Discov. 11 (1), e1391. http://dx.doi.org/10.1002/widm.1391, arXiv:https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/widm.1391.

Couture, J.J., Serbin, S.P., Townsend, P.A., 2013. Spectroscopic sensitivity of real-time, rapidly induced phytochemical change in response to damage. New Phytol. 198 (1), 311–319. http://dx.doi.org/10.1111/nph.12159.

Gold, K.M., Townsend, P.A., Chlus, A., Herrmann, I., Couture, J.J., Larson, E.R., Gevens, A.J., 2020a. Hyperspectral measurements enable pre-symptomatic detection and differentiation of contrasting physiological effects of late blight and early blight in potato. Remote Sens. 12 (2), http://dx.doi.org/10.3390/rs12020286, URL https://www.mdpi.com/2072-4292/12/2/286.

Gold, K.M., Townsend, P.A., Larson, E.R., Herrmann, I., Gevens, A.J., 2020b. Contact reflectance spectroscopy for rapid, accurate, and nondestructive phytophthora infestans clonal lineage discrimination. Phytopathology® 110 (4), 851–862. http://dx.doi.org/10.1094/PHYTO-08-19-0294-R, PMID: 31880984.

Jacquemoud, S., Ustin, S., 2019. Leaf Optical Properties. Cambridge University Press, http://dx.doi.org/10.1017/9781108686457.

Jobson, J.D., 1991. Multiple linear regression. In: Applied Multivariate Data Analysis: Regression and Experimental Design. Springer New York, New York, NY, pp. 219–398. http://dx.doi.org/10.1007/978-1-4612-0955-3_4.

Lindon, J., Tranter, G., Koppenaal, D., 2017. Index. In: Encyclopedia of Spectroscopy and Spectrometry. Elsevier, pp. 753–810.

Lowe, A., Harrison, N., French, A., 2017. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. Plant Methods 13, http://dx.doi.org/10.1186/s13007-017-0233-z.

Mahlein, A.-K., Kuska, M., Behmann, J., Polder, G., Walter, A., 2018. Hyperspectral sensors and imaging technologies in phytopathology: State of the art. Annu. Rev. Phytopathol. 56 (1), 535–558. http://dx.doi.org/10.1146/annurev-phyto-080417-050100, PMID: 30149790.

Mahlein, A.-K., Steiner, U., Hillnhütter, C., Dehne, H.-W., Oerke, E.-C., 2012. Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet disease. Plant Methods 8, 3. http://dx.doi.org/10.1186/1746-4811-8-3.

Mas Garcia, S., Ryckewaert, M., Abdelghafour, F., Metz, M., Moura, D., Feilhes, C., Prezman, F., Bendoula, R., 2021. Combination of multivariate curve resolution with factorial discriminant analysis for the detection of grapevine diseases using hyperspectral imaging. A case study: Flavescence dorée. Analyst 146, 7730–7739. http://dx.doi.org/10.1039/D1AN01735G.

Merzlyak, M.N., Solovchenko, A.E., Gitelson, A.A., 2003. Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. Postharvest Biol. Technol. 27 (2), 197–211. http://dx.doi.org/10.1016/s0925-5214(02)00066-2.

Odilbekov, F., Carlson-Nilsson, U., Liljeroth, E., 2014. Phenotyping early blight resistance in potato cultivars and breeding clones. Euphytica 197, http://dx.doi.org/10.1007/s10681-013-1054-4.

Osborne, B.G., Fearn, T., 1993. Practical NIR Spectroscopy, second ed. Prentice-Hall, London, England.

Roger, J.-M., Chauchard, F., Bellon-Maurel, V., 2003. EPO–PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. Chemometr. Intell. Lab. Syst. 66 (2), 191–204. http://dx.doi.org/10.1016/S0169-7439(03)00051-0.

Rohart, F., Gautier, B., Singh, A., Lê Cao, K.-A., 2017. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput. Biol. 13 (11), e1005752.

Ruiz Perez, D., Guan, H., Madhivanan, P., Mathee, K., Narasimhan, G., 2020. So you think you can PLS-DA? BMC Bioinformatics 21, http://dx.doi.org/10.1186/s12859-019-3310-7.

Sajeevan, R.S., Abdelmeguid, I., Saripella, G.V., Lenman, M., Alexandersson, E., 2023. Comprehensive transcriptome analysis of different potato cultivars provides insight into early blight disease caused by Alternaria solani. BMC Plant Biol. 23 (1), http://dx.doi.org/10.1186/s12870-023-04135-9.

Stocchero, M., Locci, E., d'Aloja, E., Nioi, M., Baraldi, E., Giordano, G., 2019. PLS2 in metabolomics. Metabolites 9 (3), 51. http://dx.doi.org/10.3390/metabo9030051.

Streich, J., Romero, J., Gazolla, J.G.F.M., Kainer, D., Cliff, A., Prates, E.T., Brown, J.B., Khoury, S., Tuskan, G.A., Garvin, M., Jacobson, D., Harfouche, A.L., 2020. Can exascale computing and explainable artificial intelligence applied to plant biology deliver on the united nations sustainable development goals? Curr. Opin. Biotechnol. 61, 217–225. http://dx.doi.org/10.1016/j.copbio.2020.01.010, URL https://www.sciencedirect.com/science/article/pii/S0958166920300100, Plant Biotechnology Food Biotechnology.

Terentev, A., Dolzhenko, V., Fedotov, A., Eremenko, D., 2022. Current state of hyperspectral remote sensing for early plant disease detection: A review. Sensors 22 (3), 757. http://dx.doi.org/10.3390/s22030757.

Van De Vijver, R., Mertens, K., Heungens, K., Somers, B., Nuyttens, D., Borra-Serrano, I., Lootens, P., Roldán-Ruiz, I., Vangeyte, J., Saeys, W., 2020. In-field detection of alternaria solani in potato crops using hyperspectral imaging. Comput. Electron. Agric. 168, 105106. http://dx.doi.org/10.1016/j.compag.2019.105106, URL https://www.sciencedirect.com/science/article/pii/S0168169919304582.

der Waals, J.V., Korsten, L., Aveling, T., 2001. A review of early blight of potato. African Plant Protection 7 (2), 91–102. http://dx.doi.org/10.10520/EJC87837.

Weir, T.L., Huff, D.R., Christ, B.J., Romaine, C.P., 1998. RAPD-PCR analysis of genetic variation among isolates of *Alternaria solani* and *Alternaria alternata* from potato and tomato. Mycologia 90 (5), 813–821. http://dx.doi.org/10.1080/00275514.1998.12026975.

Wold, S., 1995. Chemometrics; What do we mean with it, and what do we want from it? Chemometr. Intell. Lab. Syst. 30 (1), 109–115. http://dx.doi.org/10.1016/0169-7439(95)00042-9, InCINC '94 Selected papers from the First International Chemometrics Internet Conference.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: A basic tool of chemometrics. Chemometr. Intell. Lab. Syst. 58 (2), 109–130. http://dx.doi.org/10.1016/S0169-7439(01)00155-1, PLS Methods.

Worley, B., Halouska, S., Powers, R., 2013. Utilities for quantifying separation in PCA/PLS-DA scores plots. Anal. Biochem. 433 (2), 102–104.
Worley, B., Powers, R., 2013. Multivariate analysis in metabolomics. Curr. Metabol. 1 (1), 92–107.
Wülfert, F., Kok, W., De Noord, O., Smilde, A., 2000. Linear techniques to correct for temperature-induced spectral variation in multivariate calibration. Chemometr. Intell. Lab. Syst. 51, 189–200. http://dx.doi.org/10.1016/S0169-7439(00)00069-1.