AP6

DEVELOPING NON-DESTRUCTIVE TOOLS FOR THE DIAGNOSTIC OF GANODERMA ATTACK-LEVEL ON OIL PALMS TREES: POTENTIALITIES OF REFLECTANCE SPECTROSCOPY.

Camille C.D. Lelong[(1)], Nurul Amin Sitorus[(2)], Doni Artanto Raharja[(3)], Fabrice Dubertret[(1)], Simon Brégand[(1)], Mathieu Lanore[(1)], Achmad Wahyu Sulistyanto[(3)], Prasetya[(2)], Jean-Michel Roger[(4)], Jean-Pierre Caliman[(5)]

(1) CIRAD / UMR TETIS, Maison de la Télédétection, Montpellier, France
(2) P.T. SMART / SMARTRI, Padang Halaban, North Sumatra, Indonesia
(3) P.T. SMART / SMARTRI, Pekanbaru, Riau, Sumatra, Indonesia
(4) Cemagref / UMR ITAP, Montpellier, France
(5) CIRAD / UR34 & P.T. SMART / SMARTRI, Pekanbaru, Riau, Sumatra, Indonesia

Correspondence author: Camille Lelong; camille.lelong@cirad.fr

ABSTRACT

Ganoderma fungal disease is a plague suffered by most of the oil-palm plantations in South-East Asia. Its detection is a major issue in estate management and production. However, diagnostic is today far from reliable when done only by visual symptom observation, and very expansive and damaging when obtained by root or stem tissue chemical analysis. As an alternative, we propose in this study to evaluate the potential of hyperspectral reflectance data to help detecting efficiently the disease without destruction of tissues. This study focuses on the calibration of a statistical model of discrimination between several stages of Ganoderma attack on oil palm, based on field hyperspectral measurements at the canopy scale. Field protocol and measurements are first described. Then, combinations of pre-processing, partial least square regression and factorial discriminant analysis are tested on a hundred of samples to prove the efficiency of canopy reflectance to provide information about the plant sanitary status. A robust algorithm is thus derived, allowing classifying oil palm in a 4-level typology, based on disease severity levels from the sane to the critically sick tree with a global performance of more than 92%. Basic discrimination of trees in two classes: "sane" and "sick", is efficient at 100%. Applications and further improvements of this experiment are finally discussed.

INTRODUCTION

Early and non-destructive diagnostic of crop disease is a major issue in precision farming and sustainable agriculture in general. Oil-palm plantations, in particular, strongly suffer of fungi (eg. Ganoderma) but lack of efficient tools to manage properly this threat without great losses in production or the large use of chemicals (Wood, 2007).

Ganoderma here is used as the common word designating the basal stem rot induced by fungi belonging to the genius Ganoderma, such as Ganoderma boninense (Breton et al., 2006). This disease can cause considerable damage and is one of the major issues in oil-palm crop management, especially in South-East Asia. Together with the appearance of fruiting bodies at the base of the stem, several symptoms can indicate its contamination, like unopened spears, more or less yellowing of the crown, and appearance of dip cracks at the base of the stem (Flood et al., 2000). However, most of the time, only sampling of stem tissues and chemical analysis can evaluate with confidence the level of attack by the Ganoderma (Utomo & Nielpold, 2000; Bridge et al., 2000).

Hyperspectral reflectance spectroscopy theoretically meets the requirements of non-destructive detection at large scales, thanks to strong relationships existing between the plant optical properties on one hand, and leaf pigment concentration, and foliar and canopy structures on the other hand (Chaerle & VanDerStraeten, 2000; Larsolle & Muhammed, 2007; Liew et al., 2008 Muhammed, 2005; Wang et al., 2008). Some authors have even shown that hyperspectral data acquired by satellite or airborne remote sensing might be actually relevant to detect crop diseases or to assess crop damage severity (Nilsson, 1995; Apan et al., 2004; Goodwin et al., 2005). However, these studies are always crop and/or disease-dedicated and new experiments need to be performed to validate the detectability of a different pathology of another crop.

At least, some robust methodologies have proven to be efficient whatever the context, to classify spectra into different groups as long as a good sample helps training a statistical model of discrimination. Cluster analyses, for instance, allow good classifications of plant stress levels when combined with Partial Least Square Regression (PLS) (Jorgensen, 2007; Huang & Apan, 2006) or Principal Component Analysis (PCA) (Zhang et al., 2002).

In this paper, we propose to apply these approaches to validate the efficiency of hyperspectral reflectance spectroscopy to discriminate several levels of Ganoderma fungus contamination on oil palm trees. Indeed, this actual plague in oil-palm estates disease will largely benefit from an appropriate remotely sensed diagnosis tool. We will thus evaluate different statistical models for the classification of spectra acquired at the canopy level depending on the number of attack degrees. Then, we will analyze the possibility of developing a remote sensing tool on this basis, in the aim of precision farming applications.

MATERIAL AND METHOD

TEST SITE AND GROUND-TRUTH

Field measurements were achieved in an oil palm plantation located in North Sumatra, Indonesia: Padang Halaban Estate, which has been drastically attacked by the Ganoderma fungus for years. It thus provides a wide variety of disease severity. We have surveyed more specifically a hundred of oil palm trees, geo-localized and spotted in the plantation grid for easy subsequent identification, and we have assigned them a score in a four-level disease typology: 0 for sane (not sick) trees, 1 for a light attack, 2 for a medium one, and 3 for a severe (close to death) infestation.

Among the different sampled trees, even the sane ones, some showed symptoms of nutritional stress like nitrogen, iron, bore or magnesium deficiencies.

HYPERSPECTRAL DATA

Then, we performed hyperspectral reflectance measurements above the canopy of these trees with a Unispec (http://www.ppsystems.com/Literature/EDSUniSpec-SC.pdf) from PP-SYSTEMS, equipped with a Cosine Receptor and fibre optics with 20° of field of view. This spectroradiometer covers 256 spectral bands in the range 310-1130nm.

Climbing on scaffoldings to reach the top of each tree (up to six to ten meters high), we made six to ten radiance acquisitions distributed around the crown (about nine meters in diameter), each one integrating a surface of about nine square-meters. Each canopy reflected radiance measurement was directly followed by a diffuse incident light radiance acquisition for a scaling in reflectance. Then, we averaged these intermediate reflectance values to derive the mean reflectance of the whole tree crown taking into account the directional effects and the canopy asymetry.

Due to high level of noise in the resulting spectra in the two extreme domains (310-450nm) and (1100-1130nm), only the range 450-1100nm was actually analyzed in this study (Figure 2).

At the end of the campaign, the data base contained the canopy reflectance spectra of 36 palm trees belonging to score 0, 20 to score 1, 36 to score 2, and only 3 to score 3, for a total of 95 trees.

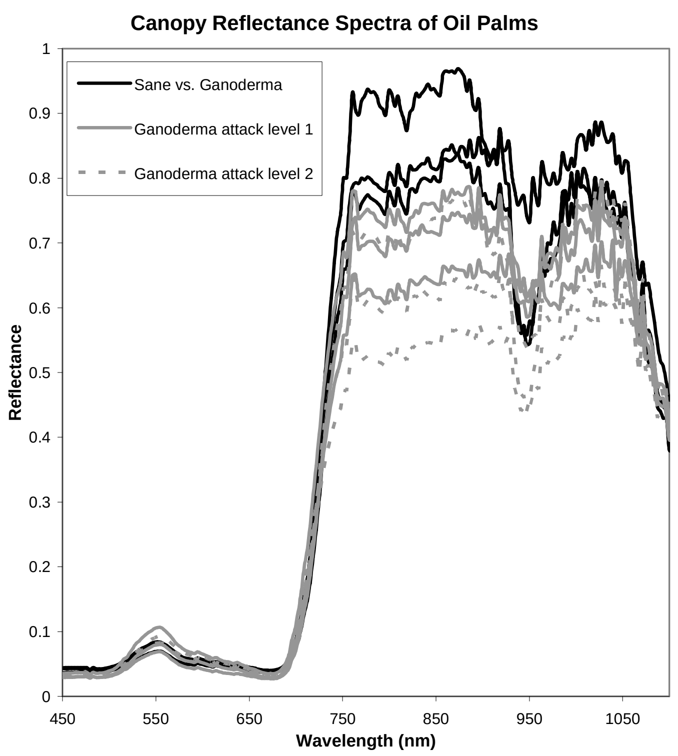

Figure 1: Examples of reflectance spectra for oil palm canopies when the tree is sane (black solid lines) or attacked by the Ganoderma disease at the respective levels 1 (grey solid lines) and 2 (grey dotted lines).

SPECTRA PRE-PROCESSING

As spectral signatures associated with Ganoderma disease symptoms might be very faint, it seems necessary to avoid any noise source or signal contamination due, for instance, to variations of sunlight and skylight illumination, soil and other backgrounds reflectance or even instrumentation itself. The Savitzky–Golay filtering (Savistky & golay, 1964) consisting in a polynomial fitting followed by a derivative computation is commonly performed to meet this requirement (Tsai & Philpot, 1998; Estep & Carter, 2005; Ruffin et al., 2008).

However, its major constraint is the choice of, on one hand, the smoothing window size and, on the other hand, the degree of the polynomial fit (Browne et al., 2007) along with the derivative order (Tsai & Philpot, 1998). We have thus chosen to test a large set of combinations of these parameters, calculating derivative-spectra at the null, first, and second order of derivation for polynomials of second and third degrees, each smoothed at nine different window sizes selected to broom the spectral bins from 10 nm to 160 nm. It thus results in 54 databases of derivative spectra. Higher orders of derivation were not tested because the canopy architecture effects on reflectance is blurred at higher orders, and ganoderma symptoms largely appear on the tree canopy structure. The original unprocessed reflectance data was also tested to evaluate the actual gain of preprocessing.

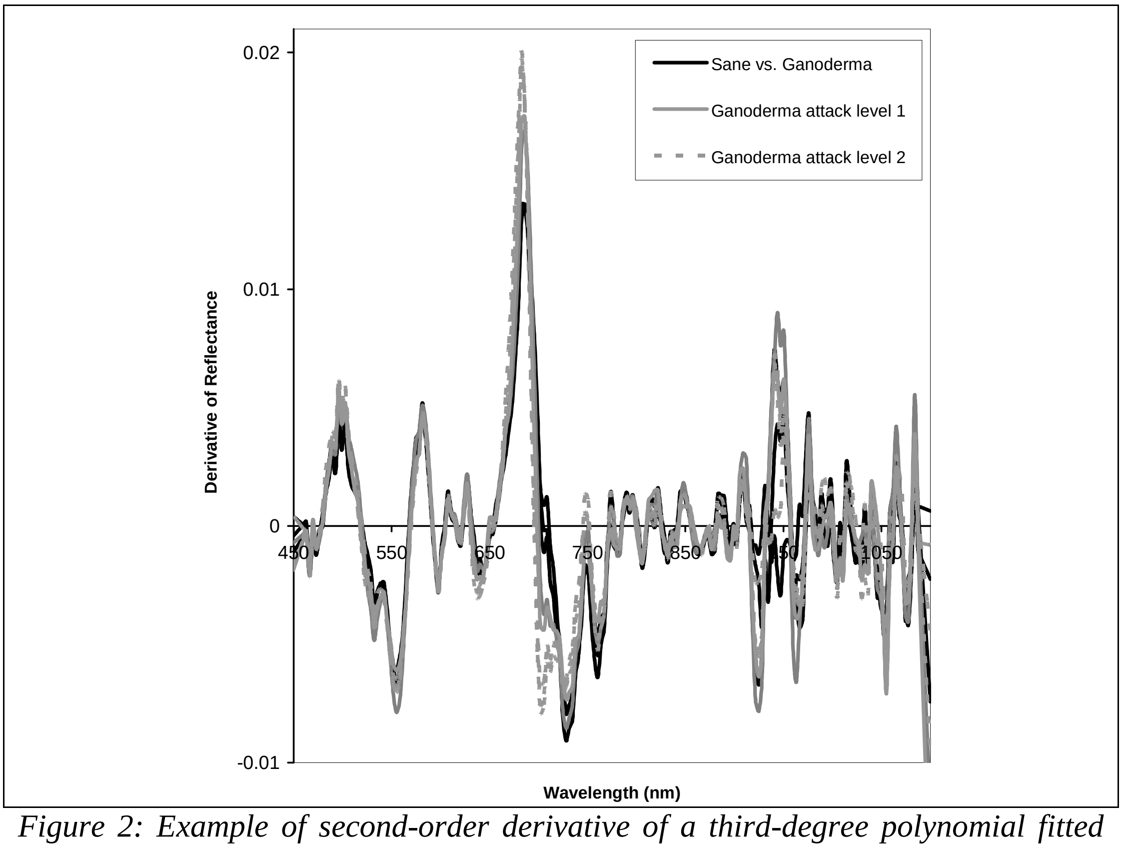

Figure 2: Example of second-order derivative of a third-degree polynomial fitted spectra (smoothing window of 26 nm).

PARTIAL-LEAST-SQUARE DISCRIMINATION

We have applied on each database the Partial-Least-Square Discrimination Analysis (**PLS-DA**) (Roger et al., 2005; Gorretta et al., 2006). It consists in:

1. Partial Least Square Regression (PLS) applied to the preprocessed derivative spectra, reducing the spectral data to little or not correlated latent variables that are fewer than the number of variables in the sample. To perform a simple PLS, we have chosen to set a specific scale fixing a predictable value for each analyzed class in the bin [0,1].
2. Discriminant Analysis (DA) applied to the most significant latent variables, enhancing the interclass variability while minimizing the intraclass variability of the sample to build a classification model. The selection of the number of PLS-variables is guided by the compromise between minimization of Root Mean Square Error of Prediction (RMSEP) and gain in correlation coefficient (R^2) between predicted and reference values, on one hand, and stability of the model thanks to the fewer number of implied variables, on the other hand.

PLS-DA was achieved on the entire sample of 95 individuals by cross-validation based on the "leave-one-out" method. The potential of the method was tested following two objectives independently:

1. potential discrimination between the healthy ("score 0"; 36 individuals in the sample) and the sick ("scores 1", "2", and "3" together; 56 individuals in the sample) trees,
2. potential of classification of a given tree in the 4-level scoring of disease severity.

Classification results were then compared on the basis of the confusion matrix and the global precision values.

RESULTS AND DISCUSSION

The first discrimination objective, aiming at detection wether an oil-palm tree is sick or not whatever its level of attack, was perfectly met (100% of good classification for each of the two classes) with a PLS-DA applied on the first-order derivative of a second-degree polynomial fitted on a smoothing window of 32 nm. PLS-predictable values were set to 0 for "sane" and 1 for "sick". DA was applied on the ten first latent variables derived from the PLS with a root mean square error of prediction of 0.27 and a correlation coefficient of 70%.

The second objective was best met, corresponding to a global accuracy of 92.6%, while using the second-order derivative of a third-degree polynomial fitted on a smoothing window of 26 nm. PLS-predictable values were set on the basis of a simple unmixing of mean spectra of each class between the two endmembers "score 0" and "score 4": 0 for "score 0", 0.4 for "score 1", 0.6 for "score 2", and 1.0 for "score 3". DA was applied on the seven first latent variables derived from the PLS with a root mean square error of prediction of 0.13 and a correlation coefficient of 80%. The two first discriminant factors are then able to split the space into four clusters clearly separated (Figure 3). The corresponding confusion matrix is given Table1.

As a factor of comparison, the best result obtained on the original (not filtered and not derived) reflectance spectra only gave a global precision of 63%, with strong confusion between sane and sick trees and bad assignation of individuals inside distant classes.

Score	0	1	2	3	% of good classification
0	34	2	0	0	94 %
1	0	17	3	0	85 %
2	0	2	34	0	94 %
3	0	0	0	3	100 %
Global precision					92.6%

Table 1: Confusion matrix obtained for the classification in four levels of disease severity.

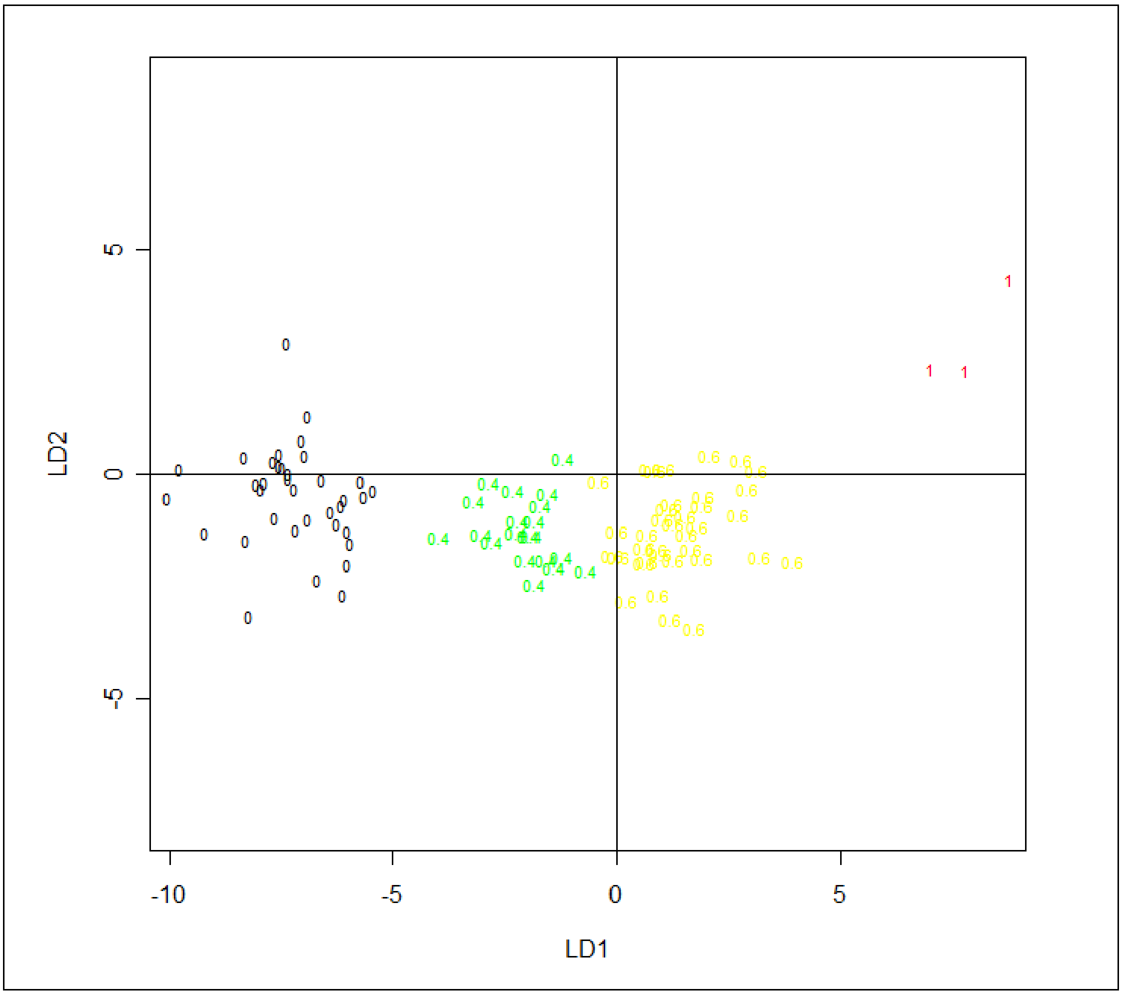

Figure 3: Representation of oil-palm trees individuals in the new two-dimensionnal space defined by the discriminant analysis applied on the PLS: sane palms are displayed in black "0" symbols, level 1 in green "0.4" symbols, level 2 in yellow "0.6" symbols, and level 3 in red "1" symbol. The four populations of trees are clearly split in this space, defining sickness severity classes' contours.

Both models thus allow a very good discrimination between the sane and the sick trees, even if 6% of false alarms can be expected in the 4-level classification process. Even though, the remaining false alarms concern only two individuals that are classified as lightly attacked by the disease; it is still possible that the visual symptoms on which was based the ground-truth diagnostic were not yet observed while the reflectance spectrum already features some changes compared to sane individuals.

Moreover, errors occurring in this classification all shift the trees from one class to the direct next one. Considering that the limits between one score evaluated in-situ and the closer one is very fuzzy, these errors can be either due to the classification or to the field diagnosis without clue to conclude. Even if it is purely a classification error, they are very rare and allow a good confidence in the overall results. The cross validation process also allows a good confidence in the robustness and stability of the model.

The tests also show that the preprocessing of the data has a considerable impact on the detectability of the spectral features associated with the disease presence, and its level of severity too.

Considering the loadings of the PLS, no privileged spectral range seems to contribute more than others do, and no range seems uninformative. This proves the essential need of the entire spectral richness to detect discriminating features in the canopy reflectance. Applications in remote sensing seems thus limited to hyperspectral sensors only, but this still has to be checked by dedicated studies. Indeed, airborne acquisitions would be of great help for fast mapping of the Ganoderma infestation in plantations, compared to field measurements that are still long, difficult, and dangerous to set up at this canopy height.

CONCLUSION

Statistical algorithms like PLS-DA applied to preprocessed hyperspectral reflectance data acquired in the fields over oil palm canopies are thus efficient to detect the Ganoderma fungal disease attack with a very high confidence. They can even classify it into four levels of severity from sane to highly damaged trees with more than 92% accuracy. Even sane trees that present nutritional deficiencies are not misclassified as sick. It proves the potential of hyperspectral reflectance spectroscopy for oil palm crop sanitary status evaluation and pushes for further improvements towards remote sensing applications.

Indeed, present measurements using field spectroradiometer on top of oil palm canopies is still very hard to set up and somehow dangerous, especially for mature and older trees. It might also be long to perform with a good quality. Acquiring such hyperspectral data on board an aircraft, for instance, or even a satellite, would be of major interest to cover a larger area in less time and better conditions. Moreover, hyperspectral imagery would add the spatial information, and so the opportunity to map quickly the location of attacked trees and thus the disease focus, and, furthermore, to analyze the epidemiology inside a palm block, the plantation, or even the planting region depending of the width of the survey. Nevertheless, new protocols would then to be fitted to airborne or satellite-borne hyperspectral images, to calibrate a dedicated model that would take into account the imaging specificities (e.g. radiometric noise and sensibility, transfer of scales, spectral contribution of neighbors, effects of non-palm objects and of background vegetation, etc…).

ACKNOWLEDGEMENTS

This project was funded by SMART-TBK Company in Indonesia, and the French Center of International Cooperation in Agronomical Research for the Development (CIRAD). The authors would also like to thanks Dr. Prasetya, Head of SMART-RI antenna in Padang Halaban Estate, and Dr. Wahyu, phytopathologist at SMART-RI office in Pekanbaru, for their help in the ganoderma diagnostic survey and operations in the fields.

REFERENCES

APAN, A; HELD, A; PHINN, S; MARKLEY, J (2004). Detecting sugarcane 'orange rust' disease using EO-1 hyperion hyperspectral imagery. *International Journal of Remote Sensing,* 25, p489-498.

BRIDGE, PD; O'GRADY, EB; PILOTTI, C.A; et al. (2000). Development of molecular diagnostics for the detection of *Ganoderma* isolates pathogenic to oil palm. In *Ganoderma diseases of perennial crops*, CABInternational Ed.- Wallingford Press, p225-234**.**

BROWNE, M; MAYER, N; CUTMORE, TRH (2007). A multiscale polynomial filter for adaptive smoothing. *Digital Signal Processing,* 17,p 69-75.

CHAERLE, L; and VAN DER STRAETEN, D (2000). Imaging techniques and the early detection of plant stress. *Trends in Plant Science,* 5, p295-500.

ESTEP, L; CARTER, GA (2005). Derivative analysis of AVIRIS data for crop stress detection. *Photogrammetric Engineering & Remote Sensing,* 71, p1417-1421.

FLOOD, J; BRIDGE, PD; HOLDERNESS, M (2000). *Ganoderma diseases of perennial crops*. CABInternational Ed. - Wallingford Press, 245 pages.

GOODWIN, N; COOPS, CC; STONE, C (2005). Assessing plantation canopy condition from airborne imagery using spectral mixture analysis and fractional abundances. *International Journal of Applied Earth Observation and Geoinformation,* 7, p11-28.

GORRETTA, N; ROGER, JM; AUBERT, M; BELLON-MAUREL, V; CAMPAN, F; ROUMET, P (2006). Determining vitreousness of durum wheat kernels using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy,* 14, p231-239.

HUANG, JF; APAN, A (2006). Detection of Sclerotinia rot disease on celery using hyperspectral data and partial least squares regression. *Journal of Space Science,* 51, p129-142.

JORGENSEN, RN; CHRISTENSEN, LK; BROS, R (2007). Spectral reflectance at sub-leaf scale including the spatial distribution discriminating NPK stress characteristics in barley using multiway partial least square regression. *International Journal of Remote Sensing*, 28, p943-962.

LARSOLLE, A; and MUHAMMED, HH (2007). Measuring crop status using multivariate analysis of hyperspectral field reflectance with application to disease severity and plant density. *Precision Agriculture,* 8, p1385-2256.

LIEW, OW; CHONG, PCJ; LI, G; ASUNDI, A (2008). Signature optical clues: emerging technologies for monitoring plant health. *Sensors* **2008**, 8, p3205-3239.

MUHAMMED, HH (2005). Hyperspectral Crop Reflectance Data for characterising and estimating Fungal Disease Severity in Wheat *Biosystems Engineering,* 91,p 9-20.

NILSSON, HE (1995). Remote sensing and image analysis in plant pathology. *Annual Review of Phytopathology,* 15, p489-527.

ROGER, JM; PALAGOS, B; GUILLAUME, S; BELLON-MAUREL, V (2005). Discriminating from highly multivariate data by Focal Eigen Function discriminant analysis; application to NIR spectra. *Chemiometrics and Intelligent Laboratory Systems,* 79, p31-41.

RUFFIN, C; KING, RL; YOUNANI, NH (2008). A combined derivative spectroscopy and Savitzky-Golay filtering method for the analysis of hyperspectral data. *Giscience &Remote Sensing,* 45, p1-15.

SAVITZKY, A; and GOLAY, MJE (1964). Smoothing and differentiation of data by simplified least-squares procedures. *Analytical Chemistry,* 64, p1627-1639.

TSAI, F; and PHILPOT, W (1998). Derivative analysis of hyperspectral data. *Remote Sensing of Environment,* 66, p41-51.

UTOMO, C; NIEPOLD, F (2000). Development of diagnostic methods for detecting Ganoderma-infected oil palms . *Journal Of Phytopathology*, 148 (9-10), p235-247.

WANG, W; ZHANG, M; ZHU, J; GENG, S (2008). Spectral prediction of Phytophtora infestans infection on tomatoes using artificial neural netword (ANN). *International Journal of Remote Sensing, 29,* p1693-1706.

WOOD, BJ (2007). Opportunities for oil palm R&D in further meeting the challenges of the new dynamics. *Planters*, 83(972), p155-177.

ZHANG, M; LIU, X; O'NEILL, M (2002). Spectral discrimination of Phytophtora infestans infection on tomatoes based on principal component and cluster analyses. *International Journal of Remote Sensing*, 23, p1095-1107.