SHORT COMMUNICATION

USING SPECTROSCOPIC DATA SETS TO PREDICT NUMBERS OF THRIPS (*FULMEKIOLA SERRATA*) IN SUGARCANE

ABDEL-RAHMAN EM[1], VAN DEN BERG M[2], AHMED FB[1], WAY MJ[2] and SEWPERSAD C[2]

[1]*School of Environmental Sciences, University of KwaZulu-Natal, Durban, 4041, South Africa*
[2]*South African Sugarcane Research Institute, Private Bag X02, Mount Edgecombe, 4300, South Africa*
alfatih100@hotmail.com Maurits.Vandenberg@sugar.org.za Ahmed@ukzn.ac.za
Mike.Way@sugar.org.za Nikki.Sewpersad@sugar.org.za

Abstract

Sugarcane thrips, *Fulmekiola serrata* (Kobus) (Thysanoptera: Thripidae), has recently been perceived as a constraint to southern African sugarcane production. Spectroscopic readings at leaf level in sugarcane variety N19 were recorded with the aim of predicting numbers of thrips (nymphs, adults and nymphs+adults) using partial least squares (PLS) regression. Correlations were high, *viz.* R^2=0.75 for predicting nymphs numbers, 0.72 for adults and 0.75 for nymphs+adults. Predictions were reliable for nymphs since root mean square error (RMSE) of prediction was 1.25 (47%). Higher RMSE values were obtained for adults (4.8; 66%) and combined life stages (6.2; 63%).

Keywords: sugarcane, thrips, *Fulmekiola serrata*, spectroscopy, partial least squares regression

Introduction

Sugarcane thrips, *Fulmekiola serrata* (Kobus) (Thysanoptera: Thripidae), has recently been perceived as a constraint to southern African sugarcane production. Monitoring techniques are required for this pest for various research purposes.

Abdel-Rahman *et al*. (2008a) demonstrated the potential of using hyperspectral remote sensing to detect sugarcane thrips damage at leaf level. The technique has been used for monitoring other crop pests (e.g. Sudbrink *et al*., 2003; Mirik *et al*., 2007). This study was aimed at investigating possible correlations between sugarcane thrips numbers and spectroscopic readings at leaf level.

Materials and Methods

During December 2007, 13 fields of 4-5 month old sugarcane variety N19 at Umfolozi in the KwaZulu-Natal province of South Africa were sampled for thrips. Twenty-seven samples were taken at random by taking five leaf spindles within a five metre radius. Samples were stored at –18ºC for one week before spectral measurements were taken. Reflectance measurements were taken in the laboratory from these samples, using a spectroradiometer (FieldSpec® 3) in the full range of the electromagnetic spectrum (350-2500 nm) at a room temperature of 23-25ºC. The spectroradiometer fibre-optic cable was directed 10 cm above the sample, which was laid on a black platform. A 50 watt halogen lamp was placed 40 cm and 45º from the sample. The reflectance spectra were transformed to the first-order derivative reflectance to enhance the absorption features on the spectra, as suggested by Abdel-Rahman *et al*. (2008b). The numbers of nymph and adult thrips per spindle were determined according to Way (2008).

A so-called random forest algorithm (Breiman, 1996 in Liaw and Wiener, 2002) was used to reduce the number of variables in the spectroscopic data set while preserving maximum relevant information for thrips prediction. This method ranks the 'importance' of each variable by comparing how much the prediction error increases when this variable is removed, while all other variables are left unchanged (Archer and Kimes, 2008). The portions between 1355-1450 nm, 1800-1950 nm and 2420-2500 nm were excluded from the analysis as suggested in the user's guide to the spectroradiometer (ASD, 2005). Partial least squares (PLS) regression was then used to predict thrips numbers with the 2.5% most important wavelengths as ranked by the random forest algorithm. The PLS models developed were validated using a jack-knifing (leave-one-out) method. The influence of each important wavelength into the thrips numbers prediction was evaluated by normalisation, as suggested by Martin *et al*. (2008). The analyses were done in R software for statistical analysis ([1]R Development Core Team, 2008).

Results and Discussion

Thrips numbers averaged 2.64 nymphs, 7.31 adults and 9.95 nymphs+adults per spindle. Figure 1 shows the mean first-order derivative reflectance of sugarcane leaves, the 2.5% most important wavelengths for predicting the numbers of nymphs, adults and nymphs+adults. Most of the important wavelengths for predicting nymphs, adults and nymphs+adults were found in the near infrared region (700-1300 nm) of the electromagnetic spectrum. The PLS regression analysis showed R^2 of 0.75 (one component) for predicting nymphs; 0.72 (two components) for adults and 0.75 (2 components) for nymphs+adults. The root mean square error of prediction (RMSEP) was 1.25 (47.3%) nymphs, 4.8 (65.7%) adults and 6.2 (62.8%) nymphs+adults (Figure 2). R^2 values for validating the PLS regression models with a jack-knifing are shown in Figure 2. These results suggest that nymph numbers can be reliably estimated using spectroscopic analysis. This finding is an interesting topic for further research. Figure 2a,b,c shows PLS normalised regression coefficients which reflect the relative influence of each of the 2.5% most important wavelength in the prediction of thrips numbers. The negative relations of some wavebands in the near infrared region of the PLS models were expected since the reflectance in this region is mainly controlled by leaf water content (Kumar *et al*., 2003), and the leaf ruptures or lesions caused by thrips could induce water loss.

Conclusions

Results show the potentiality of spectroscopy in detecting thrips numbers at leaf level. Although R^2 values were high for all life stages, predicting thrips numbers based on RMSE was possible for the nymph only. The next steps should be validating the predictive model under field conditions at leaf and canopy levels and with other varieties and different ages.

[1]R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

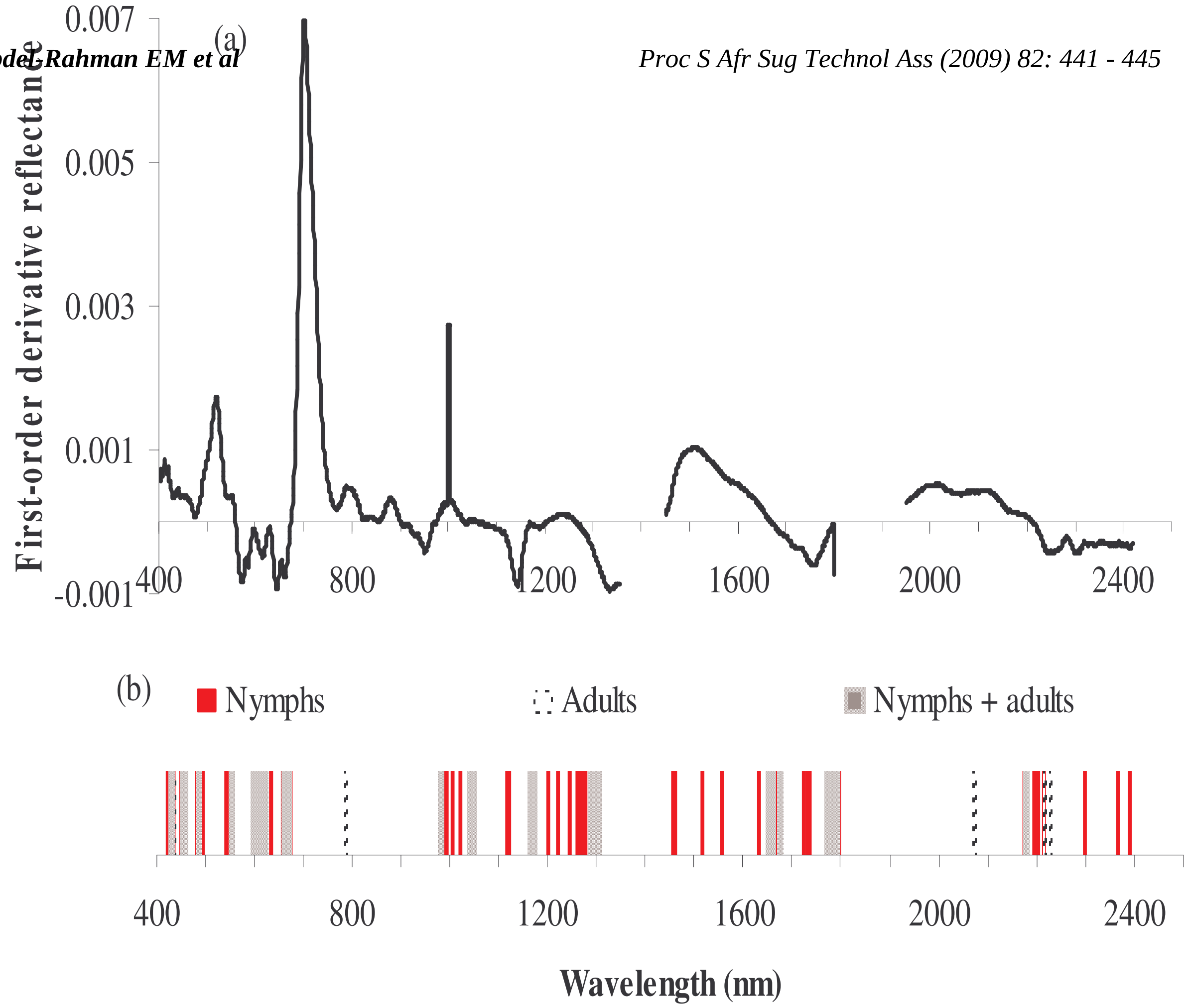

Figure 1. First-order derivative spectrum of sugarcane leaves (a) and the result of random forest algorithm showing the 2.5% most important wavelengths for predicting nymphs, thrips adults and nymphs+adults, and (b) spectral features at 1355-1450, 1800-1950 and 2420-2500 nm were removed due to excessive noise.

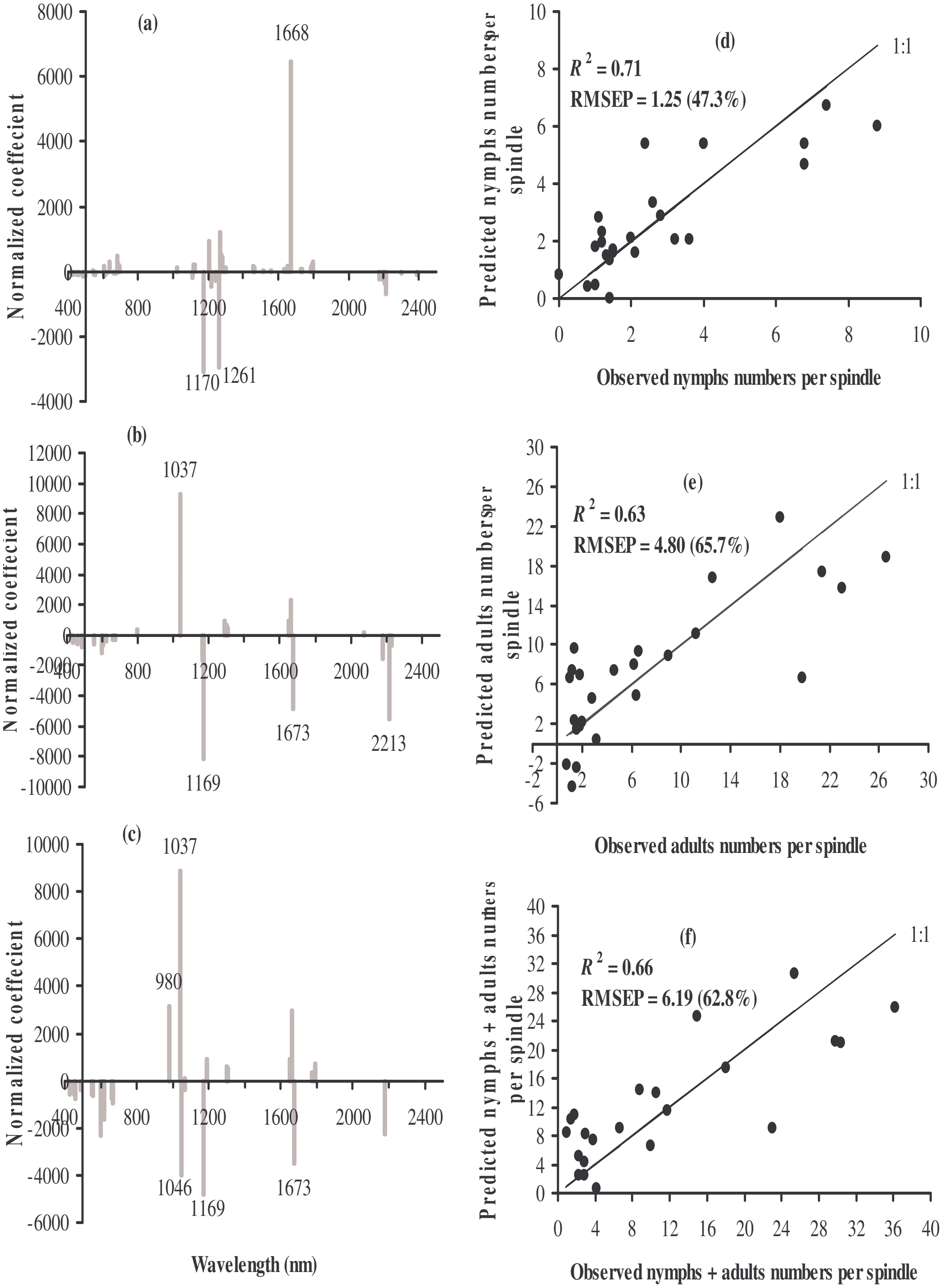

Figure 2. The normalised PLS regression coefficients for predicting nymphs (a), adults thrips (b) and nymphs + adults (c) numbers (the labels show the most influential wavelengths for predicting thrips numbers); d, e and f are 1:1 relationships between observed and predicted numbers of nymphs, adults and nymphs+adults, respectively, using a jack-knifing method.

Acknowledgments

The University of Khartoum (Sudan), the University of KwaZulu-Natal and the South African Sugarcane Research Institute are thanked for providing funding. Mr Marius Adendorff, Cyril Cele, Nitesh Poona, Innocent Shezi and Tholang Mokhele are thanked for their assistance in the field.

REFERENCES

Abdel-Rahman EM, Ahmed FB, van den Berg M. and Way MJ (2008a). Preliminary study on sugarcane thrips (*Fulmekiola serrata*) damage detection using imaging spectroscopy. *Proc S Afr Sug Technol Ass* 81: 287-289.

Abdel-Rahman EM, Ahmed FB and van den Berg M (2008b). Imaging spectroscopy for estimating sugarcane leaf nitrogen concentration. *Proc SPIE Remote Sensing for Agriculture, Ecosystems, and Hydrology X Conference*, Cardiff, United Kingdom, 7104: V1-V12.

Archer KJ and Kimes RV (2008). Empirical characterization of random forest variable importance measures. *Computational Statistics and Data Analysis* 52: 2249-2260.

ASD (2005). Handheld Spectroradiometer: User's guide version 4.05, Analytical Spectral Devices, Inc., Suite A, Boulder, USA. 136 pp.

Kumar L, Schmidt K, Dury S and Skidmore A (2003). Imaging spectrometry and vegetation science. pp 111-156 In: van der Meer FD and de Jong SM (Eds), *Image Spectrometry*, Vol 3, Kluwer Academic Publishers, London, UK.

Liaw A and Wiener M (2002). Classification and regression by random forest. *R News* 2(3): 18-22.

Martin ME, Plourde LC, Ollinger SV, Smith ML and McNeil BE (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment* 112(9): 3511-3519.

Mirik M, Michels Jr GJ, Kassymzhanova-Mirik S and Ellott NC (2007). Reflectance characteristics of Russian wheat aphid (Hemiptera: Aphididae) stress and abundance in winter wheat. *Computers and Electronics in Agriculture* 57(2): 123-134.

Sudbrink DL, Harris FA, Robbins JT, English PJ and Willers JL (2003). Evaluation of remote sensing to identify variability of cotton plant growth and correlation with larval densities of beet armyworm and cabbage looper (Lepidoptera: Noctuidae). *Florida Entomologist* 86(3): 290-294.

Way MJ (2008). Arthropods associated with sugarcane leaf spindles in South Africa. *Proc of S Afr Sug Technol Ass* 81: 362-364.