IMAGE-BASED TECHNIQUE FOR ESTIMATING PHOSPHORUS LEVELS IN COTTON (*GOSSYPIUM HIRSUTUM* L.)

Mahdi Ali[1], Ahmed AL-Ani[1], Daniel Tan[2], Derek Eamus[3], Ian Rochester[4]

[1] Faculty of Engineering and IT, University of Technology, Sydney, NSW 2007, Australia

[2] Faculty of Agriculture and Environment, the University of Sydney, NSW 2006, Australia

[3] School of the Environment, University of Technology, Sydney, NSW 2007, Australia

[4] CSIRO Plant Industry, Cotton Catchment Communities Cooperative Research Centre, Narrabri, NSW, Australia

Abstract

In this glasshouse study, we proposed a new image-based non-destructive technique for estimating P levels in cotton leaves. The plants were grown on a nutrient medium containing various P concentrations, i.e. 0%, 50% and 100% of recommended P levels for 10 weeks, and then leaf P contents were analysed using a destructive method. For comparison, we collected leaf images of the leaves using a handheld crop sensor. The RGB (red, green and blue) values of the collected images were used for calculating leaf area and leaf perimeter. This data on leaf growth parameters was used in a linear discriminant analysis (LDA) to estimate leaf P contents. Using LDA, we successfully classified cotton plants into different group based on their leaf P contents indicating that P deficiency in cotton can be estimated using leaf morphological data. Our proposed non-destructive method was efficient in estimating P requirements for cotton.

Key words: *Image-based techniques, cotton, leaf P contents, linear discriminant analysis*

1. INTRODUCTION

Crop plants require sufficient phosphorus (P) supply in the soils for their proper growth and functioning (Batten 1992). Due to its crucial role in cellular division and expansion in plants, P deficiency in plant tissues can inhibit leaf size, light interception and overall carbohydrate assimilation resulting in stunted plant growth (Rodríguez et al. 1998, Lloyd et al. 1995). P deficiency severely affects the early vegetative growth of cotton (Hearn, 1981), which results in stunted foliage, restricted carbohydrate accumulation and consequently development of new nodes (Sawan et al. 2001). Since final yield in cotton is associated with the total number of fruiting nodes, this P-induced reduced plant growth can ultimately influence cotton yield. On the other hand, higher P concentration in plant tissues can cause toxicity leading towards growth inhibition, senescence and development of chlorotic or necrotic region on leaves (Shane et al. 2004).

The P status of field-grown cotton is generally estimated by analysing the soil or plant samples, and P deficiency is corrected by applying additional fertilisers. These destructive techniques are generally accurate but time consuming and expensive (Sui et al. 2005). Thus there is a need to develop a cheap and rapid method for estimating cotton P requirements. As compared with plant N status, which can be easily estimated using non-destructive techniques (Ali et al. 2012), limited information are available on non-destructive techniques for estimating crop P status (Christensen & Jørgensen, 2004). Although canopy reflectance data is also used for estimating P contents of pastures i.e. sainfoin (Albayrak 2008) and pasture grass (Mutanga et al. 2004), limited information is available on estimating leaf P contents of crops.

Limited P supply can restrict the expansion of cotton leaves by reducing root hydraulic conductivity and cell turgor in the leaf cell (Radin & Eidenbock1984). Impaired canopy development in P deficient cotton plants inhibits light interception, carbohydrate assimilation, overall shoot development and consequently the production of fruiting nodes (Sawan et al. 2001). Since leaf development (leaf area) is influenced by P supply (Sawan et al. 2001, Shane et al. 2004, Radin & Eidenbock 1984), we

proposed that data on leaf morphological parameters can be used for estimating P status of cotton. Using linear discriminant analysis (LDA) and the related Fisher's linear discriminant methods, we separated cotton plants in different groups on the basis of their leaf P contents.

2. MATERIAL AND METHODS

Seeds of a cotton cultivar Sicot 71BRF were planted in 33 plastic pots (11 replicates for each treatment) at the Faculty of Science, University of Technology Sydney, Australia. All pots were filled with vermiculite, and were irrigated using a nutrient solution containing 5.4 mM of NH_4NO_3, 1.6 mM of K_2HPO_4, 0.3 mM of K_2SO_4, 4 mM of $CaCl_2.2H_2O$, 1.4 mM of $MgSO_4.H_2O$, 5 µM of Fe-EDDHA, 2 µM of $MnSO_4.H_2O$, 1 µM of $ZnSO_4.7H_2O$, 0.25 µM of $CuSO_4.5H_2O$, 0.3 µM of $Na_2MoO_4.2H_2O$, and 0.5 µM of H_3BO_3. For the first seven days, P level in the nutrient solution was kept constant for all the plants (2.5 mL / 10 L), and the nutrient solution was renewed every three days. After seven days, three different P treatments in the form of NaH_2PO_4 were applied for a period of seven weeks (P0 = no P, L; P1 = 2.5 mL / 10 L of P and P2 = 5 mL / 10 L of P). The P levels applied to each pot represent three different levels of recommended P application, i.e. P0 = 0 %, P1 = 50 % and P2=100 % of recommended P concentration.

2.1. Data collection

After 8 weeks of growth on variable P levels, data on leaf growth parameters such as leaf area (LA) and perimeter were collected from the uppermost fully plant expanded leaves of individual plants using a portable leaf scanner (Pico Life). The images collected by leaf scanner were processed to get RGB values. After scanning, the same leaves were oven dried at 80^{o}C for 24 hours, and ground to powder form. These samples were used for tissue P content analysis using inductively coupled plasma mass spectrometer (ICPMS). Plants were harvested after 8 weeks of treatment, and the samples were used for collecting data on shoot length and shoot dry weight.

2.2.Linear discriminant analysis (LDA)

We used linear discriminant analysis (LDA) and related Fisher's linear discriminant methods to find the linear combination of features, which best separate two or more classes of objects or events. These combinations were then used as a linear classifier. Considering that we have two classes or categories, which can be related to a certain plant condition (e.g P sufficient vs. P deficient), as shown in Fig. 1A), we used LDA to find a projection matrix, when this matrix is multiplied by our original data matrix would increase the distance between the two classes and minimise the distance between samples of the same class. For this, the LDA constructs two scatter matrices denoted as: within-class scatter (*Sw*) and between-class scatter (*Sb*):

When considering the between-class scatter (S_b), the main task was to maximise the distance between the centers of different classes or simply maximise the distance between different centers and the mean (center) of the whole data as per the equation below:

$$S_B = \sum_{i=1}^{c} (v_i - \bar{x})\,(v_i - \bar{x})^T \quad (5)$$

Where v_i represents the center of each of the classes and $\bar{x}$ the mean of the whole data set. An example of what S_b attempts to achieve is shown in Fig. 1B, as the main task here was to push the centers of the two classes far away from each other.

On the other hand, considering the between-class scatter (S_w), the main task is to minimize the distance between the samples of each class and their corresponding centre as per the equation below:

$$S_W = \sum_{i=1}^{c} * \sum_{k=1}^{li} (xk - vi)(xk - vi)^T \quad (6)$$

Where x_k represents the samples of each of the classes and v_i the mean or the center of same class. An example of what S_w attempts to achieve is shown in Fig. 1C, LDA reduces the distance between the samples of same color simply pushing the points toward their centers.

After constructing these scatter matrices, LDA attempts to project the data in a manner that maximises between-class distance and minimizes within-class distance i.e., the discriminant criterion in mathematical formulation is given as:

$$agr_G max \frac{Trace(G^T S_b G)}{Trace(G^T S_w G)} \quad (7)$$

Where the optimal transformation "*G*" is given by solving a generalized eigenvalue problem $S_W^{-1} S_B$. Once the matrix "*G*" is found, then the next step is to multiply the original data by *G* and submit the result for classification.

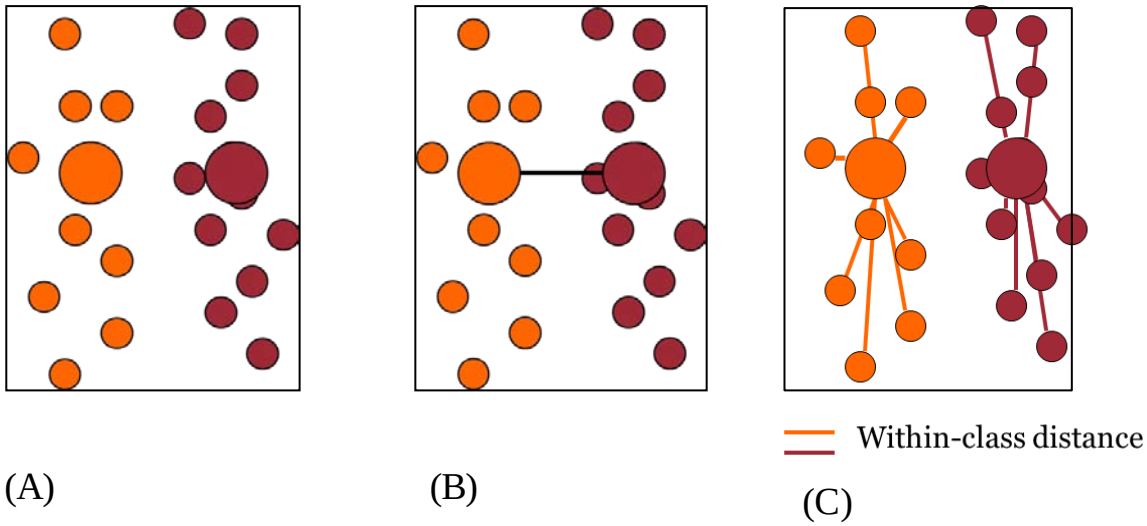


Figure 1: Classification mechanism in linear distribution analysis. (A) set of reading or measurements that belong to two classes, orange and brown, (B) example of what *Sb* attempts to achieve by pushing away the centers of the two classes far away from each other and (C) example of what *Sw* attempts to achieve by bringing closer the samples of each class together.

Sw: within-class scatter; *Sb*: between-class scatter

2.3. The proposed algorithm for estimating leaf P contents

Data collected on RGB, LA and leaf perimeter were used for classifying cotton plants on P level using the LDA classifier. LDA detects the effect of various P application rates on leaf growth using one outcross validation scheme and classifies plants into groups on the basis of tissue P levels. In addition, using this data on leaf growth, we proposed the following algorithm that can estimate P contents in cotton leaves. We used an optimization algorithm to identify the weight that needs to be assigned to each of the four variables of R, G, B and LA:

R*-0.6466-G*0.0203+B*1.4837-LA*0.3758 (1)

In this testing scheme, one sample is used for testing at a time, while all other samples, excluding the testing sample, are utilised for training. The error rate is then computed by observing the ability of the classifier to correctly classify all the testing samples.

3. RESULTS AND DISCUSSIONS

3.1. Cotton growth under limited P supply

In this experiment, data on various attributes of cotton such as shoot length, dry biomass, leaf area (LA) and P concentrations in leaf tissues indicated that cotton plants can be divided into distinct groups based on the rate of applied P i.e. P0 (P-deficient), P1 (partially P-deficient) or P2 (P-sufficient), as shown in Table 1. Although a degree of overlapping in LA and shoot dry biomass of P1 and P2 plants was observed, suggesting that these parameters were less sensitive to partial P deficiency, whereas, shoot length could be used as an indicator of plant P status.

Table Classification of cotton plants on the basis of leaf P concentration

	Shoot length (cm)	Shoot dry biomass (g)	Leaf area (cm^2)	Leaf P concentration (g / kg)	P class
Average	19.5	0.8	31.1	0.9	
Min values	16.7	0.5	21.9	0.8	P0
Max Values	21.3	1.1	37.9	1.1	
Average	32.5	4.6	54.9	1.9	
Min values	27.5	3.5	37.3	1.4	P1
Max Values	37.0	6.2	70.6	2.6	
Average	41.9	4.9	67.7	4.5	
Min values	39.2	3.3	59.1	3.4	P2
Max Values	47.0	7.3	79.7	5.6	

3.2. Estimating leaf P contents using weight optimization algorithm method

Using RGB, LA and leaf perimeter values in weight optimization algorithm (equation 1), P contents of cotton leaves were estimated. To validate the efficiency of our proposed method, we compared the P estimated (our proposed method) against the original leaf P contents (determined by destructive laboratory technique). Data showed a strong correlation between the estimated and true P values (R^2 = 0.8067) suggesting high accuracy of our proposed method (Fig. 2).

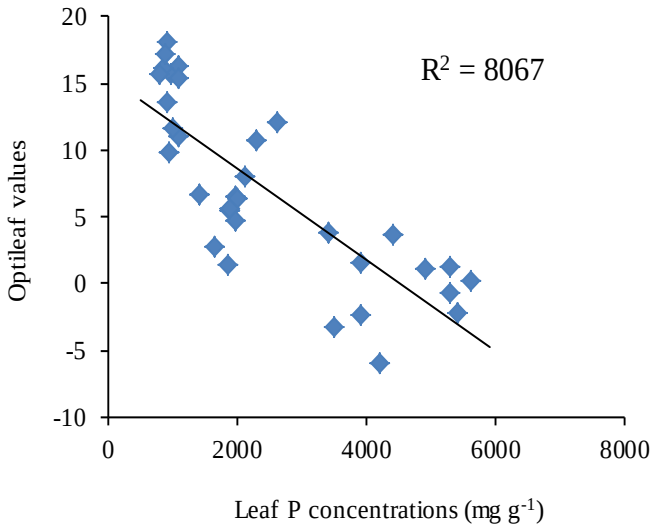


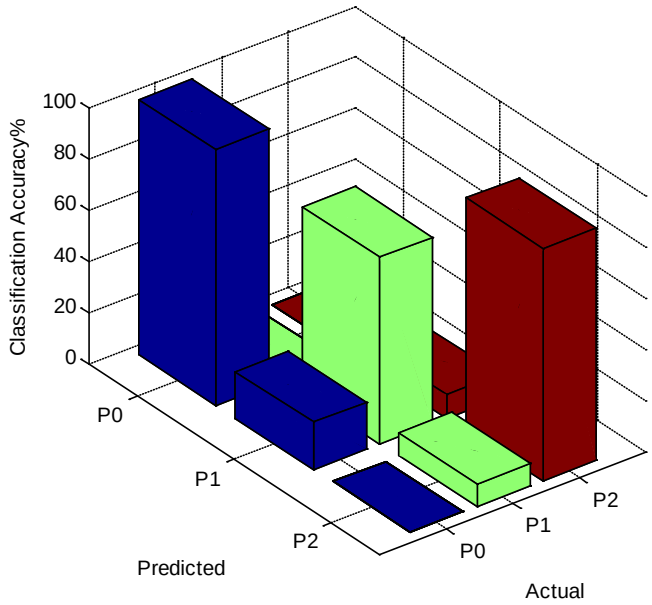
Figure 2: Correlation between leaf P contents (laboratory based analysis) and values estimated by our proposed method in cotton.

3.3. Linear discriminant analysis (LDA)

In the present experiment, we used Fisher's linear discriminant analysis (LDA) method for finding the linear combination of features, which best separate two or more classes, objects or events. When an LDA classifier is used to classify data samples of leaf RGB, LA and perimeter based on their leaf P levels (3 classes P0, P1 and P2), the achieved error rates were 18%, 9% and 9% for the three classes of P0, P1 and P2 respectively (Fig. 3). This low error rate suggests that the leaf growth features in combination with the LDA classifier could be used to estimate crop P deficiency. Data obtained with weight optimization algorithm suggests that P contents in cotton can be estimated using our proposed non-destructive method.

Estimation of P status of cotton crop using image-based technique offers a quick alternative of predicting crop P requirement without destructive sampling. Once in the plant tissues, P is used for division and expansion of cells. Thus, leaf development is largely affected by P supply (Rodríguez et al. 1998, Lloyd et al. 1995). Changes in leaf size and colour are detected by spectral images, while, LDA used these variations to classify the plants into different group. A high accuracy in group plants on the basis of leaf P contents indicated the potential of LDA for estimating cotton P status. Zhang and Lei (2011) also suggested potential of LDA for classifying vegetation on the basis of leaf sizes. Similarly, Casanova et al. (2009) suggested use of a plant classification model (LDA) for classifying plants on the basis of leaf characteristics such as shape, contour and colour.

Compared with other method used for estimating leaf P contents (Albayrak 2008, Mutanga et al. 2004), our proposed method, is more accurate, simple and straightforward, which not only can estimate the P contents of cotton tissues, but also classifies cotton plants on the basis of P application rates and assist in scheduling P application rates and time.



	Actual		
Predicted	P0	P1	P2
P0	100.00	0	0
P1	18.18	72.72	9.09
P2	0	9.29	90.90

Figure 3: Confusion matrix for the classification of cotton on basis of leaf P levels

P0; P deficient, P1, partially P sufficient and P2, P sufficient plants

4. CONCLUSIONS

The present study indicated that P requirements in cotton can be estimated from leaf morphological features, such as leaf color and dimensions. We proposed a an algorithm for measuring RGB values of cotton leaves. We used these RGB values in a linear discriminant analyser to classify cotton plants into different groups on the basis of leaf P contents. Our proposed algorithm for estimating leaf P showed a good relationship with the true leaf P contents indicating that this non-destructive method can efficiently estimate the P contents of crops growing under variable P levels.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support provided by Cotton Research and Development Corporation (CRDC) for this project.

REFERENCES

Albayrak, S 2008, 'Use of Reflectance Measurements for the Detection of N, P, K, ADF and NDF Contents in Sainfoin Pasture', *Sensors*, vol. 8, pp. 7275-7286.

Ali, M Al-Ani, A Eamus, D & Tan, D 2012, 'A new Image Processing based Technique to Determine Chlorophyll in Plants', *American-Eurasian Journal of Agricultural and Environmental Sciences*, vol. 12, pp. 1323-1328.

Batten, GD 1992, 'A review of phosphorus efficiency in wheat', *Plant and Soil*, vol. 146, pp. 163-168.

Casanova, D de Mesquita Sa Junior, J J & Bruno, OM 2009, 'Plant leaf identification using Gabor wavelets', *International Journal of Imaging Systems and Technology*, vol. 19, no. 3, pp. 236-243.

Christensen, LK & Jørgensen, RN 2004, 'Spectral NPK stress discrimination in barley plants by use of three wavebands', New directions for a diverse planet". Edited by RA Fischer. *Proceedings of the 4th International Crop Science Congres*s, Brisbane, Australia, 26 September - 1 October.

Hearn, A 1981, 'Cotton nutrition', *Field Crop Abstracts*, vol. 34, no. 1, pp. 11-34.

Hemming, J & Rath, T 2000, 'Computer-vision based weed identification under field condition using controlled lighting', *Journal of Agricultural Engineering Research*, vol.78, pp. 233–243.

Lloyd, J Grace, J Miranda, A Meir, P Wong, S Miranda, H Wright, I Gash, J & McIntyre, J 1995, 'A simple calibrated model of Amazon rainforest productivity based on leaf biochemical properties', *Plant, Cell & Environment*, vol. 18, pp. 1129-1145.

Mutanga, O Skidmore, AK & Prins, H 2004 'Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features', *Remote Sensing of Environment*, vol. 89, pp. 393-408.

Radin, JW & Eidenbock, MP 1984, 'Hydraulic conductance as a factor limiting leaf expansion of phosphorus-deficient cotton plants', *Plant Physiology*, vol.75, pp, 372-377.

Rodríguez, D Keltjens, W & Goudriaan, J 1998, 'Plant leaf area expansion and assimilate production in wheat (*Triticum aestivum* L.) growing under low phosphorus conditions', *Plant and Soil*, vol. 200, pp. 227-240.

Sawan, Z Hafez, S & Basyony, A 2001, 'Effect of phosphorus fertilization and foliar application of chelated zinc and calcium on seed, protein and oil yields and oil properties of cotton', *The Journal of Agricultural Science*, vol. 136, pp. 191-198.

Shane, MW McCully, ME & Lambers, H 2004, 'Tissue and cellular phosphorus storage during development of phosphorus toxicity in *Hakea prostrata* (Proteaceae)', *Journal of Experimental Botany*, vol. 55, pp. 1033-1044.

Sui, R Wilkerson, JB Hart, WE Wilhelm, LR Howard, DD 2005, 'Multi - spectral sensor for detection of nitrogen status in cotton', *Applied Engineering in Agriculture*, vol. 21, pp.167-172.

Wu, C Fan, Z & Wang, Z 2004, 'Effect of phosphorus stress on chlorophyll biosynthesis, photosynthesis and biomass partitioning pattern of *Fraxinus mandchurica* seedlings', *Ying Yong Sheng Tai Xue Bao*, vol. 15, pp. 935-40.

Zhang, SW & Lei, YK 2011, 'Modified locally linear discriminant embedding for plant leaf recognition', *Neurocomputing* vol.74, pp. 2284.