USE OF AIRBORNE HYPERSPECTRAL IMAGERY TO DETERMINE QUALITY OF SORGHUM CROPS

Natasha Wells[A], Rob Kelly[A] , Stuart Phinn[B] , Armando Apan[C], Troy Jensen[A], John Cooper[A] and Wayne Strong[A]

[A]Department of Primary Industries & Fisheries, Toowoomba, QLD 4350; Email natasha.wells@dpi.qld.gov.au
[B]School of Geography, Planning and Architecture, University of Queensland, Brisbane, QLD 4072; Email s.phinn@uq.edu.au
[C]Geospatial Information and Remote Sensing Group, University of Southern Queensland, Toowoomba, QLD 4350; Email apana@usq.edu.au

Abstract

Remote sensing has shown promise for predicting grain protein content for winter cereals, with satellite image data acquired near flowering, being significantly correlated with grain protein of wheat and barley crops. The use of commercially available satellite or airborne imagery to map grain protein content would allow growers and marketers the ability to evaluate crop performance and provide useful guidelines on harvest logistics and potential segregation to maximise returns. Sorghum is exposed to the air during ripening and this may provide an opportunity to identify grain protein content, a key agronomic indicator of the success or otherwise of nitrogen application to the crop. Our aim was to assess whether airborne hyperspectral imagery could be used to determine grain protein content of sorghum in the northern grains region (Darling Downs) of Australia. The first stage in this process is to determine if variations in the grain crop's protein content produce detectable variations in image data of the grain crop. The Hymap^{TM} sensor was used to acquire a 126 band, 3m pixels data set on 16 April 2004 for several sorghum fields at different growth stages. Availability of concurrent grain protein data restricted the analysis to one of these fields, which was at the end of the grain-filling stage. Grain protein was mapped within four and eight weeks of the Hymap^{TM} image by interpolating point samples collected from a near-infrared (NIR) protein sensor mounted on a combine harvester. Preliminary analysis of the image spectral reflectance and field data revealed grain protein content in sorghum was moderately correlated (r=-0.57) with red to near-infrared band (750nm) reflectance. Principal component bands derived from the Hymap^{TM} data were weakly correlated (r=0.43) with grain protein. Grain protein content was moderately (r<-0.5) correlated with variations in image spectral reflectance in bands falling between 730nm–1135nm. This information was then used to develop an inverse model, to predict grain protein content from Hymap^{TM} image data. A stepwise regression indicated that five bands in the red-edge and NIR regions (750-1150nm) explained the maximum variation in grain protein content (adjusted r^2=0.36). The results of this study are exploratory and will be refined in future papers.

Introduction

Although much attention has been given to predicting grain yields using remote sensed data (e.g. Staggenborg *et al*., 2000), there has been relatively little focus on forecasting crop quality such as grain protein content using spectral reflectance properties of the crop (Basnet *et al.,* 2003). Recent research by Basnet *et al*. (2003) suggested that certain spectral bands and indices extracted from satellite images could be used to predict grain protein in wheat and barley crops. Although the crop growth stage at image capture seems key in establishing a strong correlation between protein levels and spectral reflectance data, the optimal acquisition window has not been clearly defined for different grain crops. Basnet *et al.* (2003) found protein content for wheat and barley strongly correlated to reflectance measured from satellite imagery at crop flowering, while sorghum protein was poorly correlated to satellite bands captured three weeks before harvest.

Although Australian grain growers are not paid a premium for grain quality in sorghum crops, knowledge of grain protein is important in improving the management of soil nitrogen (N) for the optimisation of yield. In the northern grains region (i.e. north of Dubbo, NSW or 32°S), soil N and soil moisture are the primary factors limiting crop production (Strong and Holford, 1997). Yield and protein are known to vary spatially within a field (Stewart *et* al., 2002; Strong *et al*., 2003). Growers need reliable tools for acquiring broad scale information on in-season crop quality without having to rely on intensive ground sampling and traditional laboratory analysis. Remote sensed imagery has been used to predict crop yields and crop nitrogen content and may provide an opportunity for characterising grain quality (Yang and Everitt, 2002).

The high spectral resolution, and the broad spectral range, of imaging and laboratory spectrometers allow plant compounds such as protein to be measured (Sims and Gamon, 2002). Hence, the use of commercially available and appropriately corrected airborne or satellite hyperspectral imagery may improve our ability to identify differences in grain weight and quality, the latter by mapping grain protein and N content. The aim of this study was to assess the use of hyperspectral imagery to predict grain protein content in sorghum, by examining the extent to which image spectral reflectance variations were controlled by variations in crop protein content.

Study Site, Data and Methods

Sorghum (*Sorghum bicolor* cv Buster and cv MR43) was sown at Well Park (Figure 1) near Kaimkillenbun, on the Darling Downs in southern Queensland (27.105°S, 151.345°W), in late December 2003 and early January 2004. Fertiliser nitrogen (N) was added as urea (46% N) to provide 100kg N/ha, in addition to 40kg of starter-Z (10.3/20.6/1.9/1.4% equivalent of N/P/S/Zn) which added an additional 4kg N/ha.

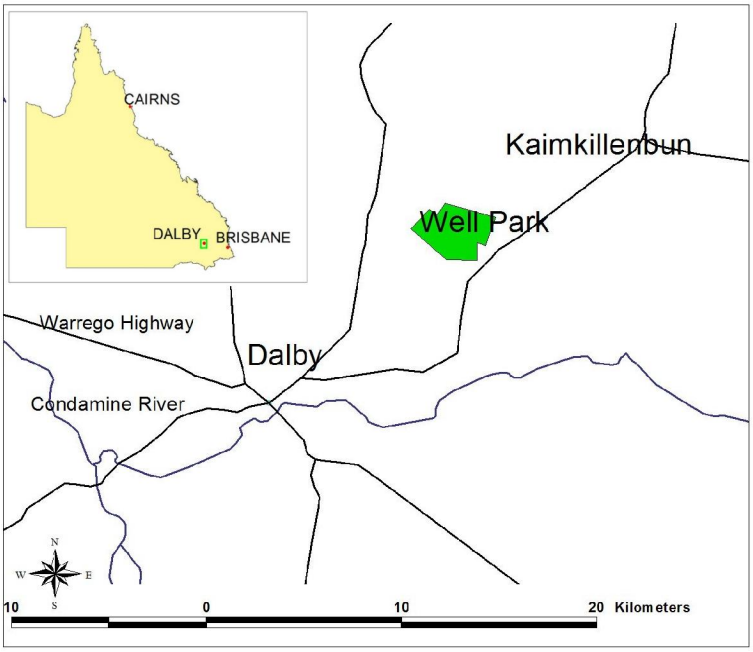

Figure 1. Well Park field study site, near Dalby, Queensland.

An aerial hyperspectral image was acquired on 16 April by HyVISTA Corporation using the Hyperspectral Mapper (HyMap^{TM}) (www.hymap.com). Spatial resolution was ca. 3m. In-house spectral and radiometric calibration was performed to convert digital numbers into radiance values (μW/cm^2 nm sr), and then to apparent surface reflectance (scaled by 10000) using an atmospheric correction model. Geometric corrections were performed following the radiometric correction using the navigation parameters measured by an on-board inertial measurement unit (IMU).

Prior to harvest, a prototype *Cropscan 2000H* on-the-go near-infrared (NIR) protein sensor designed by NIR Technology Australia (www.lineart.zip.com.au), was fitted to the bubble-up auger of the contract harvester's *New Holland TR66* combine. The protein and moisture values, as well as the associated spectral information, were recorded to the disk-on-chip of the instrument. A differentially corrected global positioning system (DGPS) receiver was connected to the NIR sensor via an RS232 connection, providing a geographic coordinate (latitude and longitude) for each sample. At harvest (22-24 May, 15-21 June), protein samples were captured about 30-40 m apart throughout the field. A total of 5334 samples were taken within the 282 ha sample area. Additionally, 120 grab samples were collected from the bubble-up auger for laboratory analysis of protein, although results were not forthcoming for inclusion in this paper. Thus all grain protein data used in the analysis was not calibrated.

Since the Hymap^{TM} image was acquired at a late stage of grain maturation, the grain protein content at the time of image capture was expected to have varied little from the grain protein content measured at harvest, one and two months post image acquisition. Grain moisture however is expected to have varied

significantly from image capture to crop harvest as the sorghum grain fully dried down.

Protein values that appeared to be errors (i.e. protein <5% and >15%, or ca. 3 standard deviations from the mean) were deleted, reducing the data set by 3%. In ArcView 3.1 (ESRI, 1998), the protein samples were overlayed on the aerial image and a polygon of the image area corresponding to the cloud-free area of the sampled field was created (Figure 2). The polygon was then used to clip the protein data removing a further 30% of the sample points which corresponded to an area of the image obscured by a cloud and its shadow, as well as removing samples near the edge of the field boundary (to avoid a "mixed pixel" effect when merged with the pixel data). A total of 3635 protein samples remained. A "point" file corresponding to the location of the protein samples was used as the criteria for the "pixel to ASCII" export in Erdas Imagine® 8.7 (Leica Geosystems, 2004). This generated a text file containing the pixel value for each band at each protein sample site. In a spreadsheet, protein values were then added to the matrix of pixel data for statistical analysis.

Data were then compared using correlation analysis to determine the extent to which variations in protein content co-varied with spectral reflectance in each Hymap^{TM} band. Only individual bands and transformed bands (principal components analysis) were used in this preliminary assessment. The assumption of this analysis was that co-variation between grain protein and spectral reflectance may indicate a spectral band where reflectance is associated with the protein content of the grain crop. A stepwise regression procedure in Genstat (LAT, 2000) was then used to develop a form of inverse model, where the reflectance bands that were found to co-vary with protein, were used to develop a model to map protein content from reflectance data.

Figure 2. True colour HyMap^{TM} image (blue = 455nm, green = 558nm, red = 675nm) of the sorghum field at Well Park, Kaimkillenbun acquired at ca. 0930 on April 16, 2004. Protein sample points are indicated as black dots. Red outline corresponds to the polygon used to clip the protein samples mapped from the field data in Figure 3.

Results and Discussion

Field sampled and interpolated grain protein content varied considerably within the field, and ranged from <5% to >15%. Mean protein content was 8.64% (uncorrected), with a standard deviation of 1.8% (Figure 4). Higher values were observed in the earlier sampled portion (northern section) than in the later sampled portion (Figure 3). Around half of the grain samples had a protein content that fell between 7.3 and 10.2% (Figure 3).

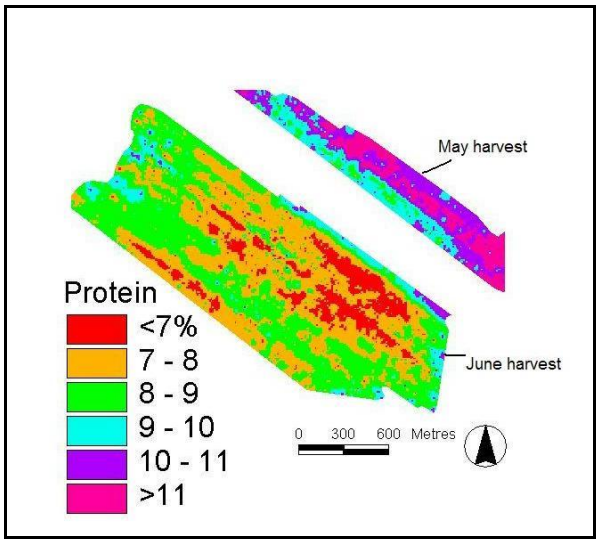

Figure 3. Protein content map derived from the protein sensor at Well Park, Kaimkillenbun, in 2004. (The surface was created using an inverse-distance weighted interpolation routine in ArcView 3.1).

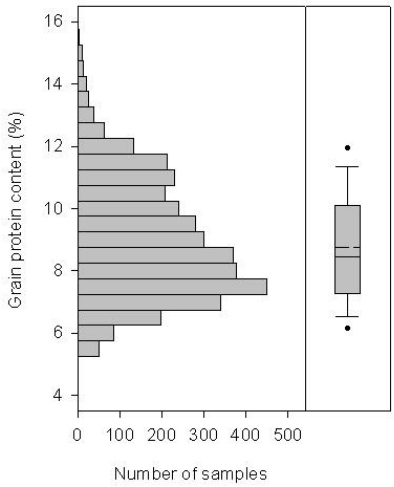

Figure 4. Frequency distribution and box plot of grain protein content at harvest from Well Park, Kaimkillenbun, in 2004.

A visual inspection of the HyMap™ image displayed in true colour (Figure 5a) revealed considerable variability in the crop canopy across the scene, with

slightly less variation evident in the false colour image (Figure 5b). Prior to the 2004 sorghum crop, the field was farmed in strips with varying crops and rotations. Differences in maturity and canopy density, presumably due to spatial variations in moisture supply, can be identified in the image.

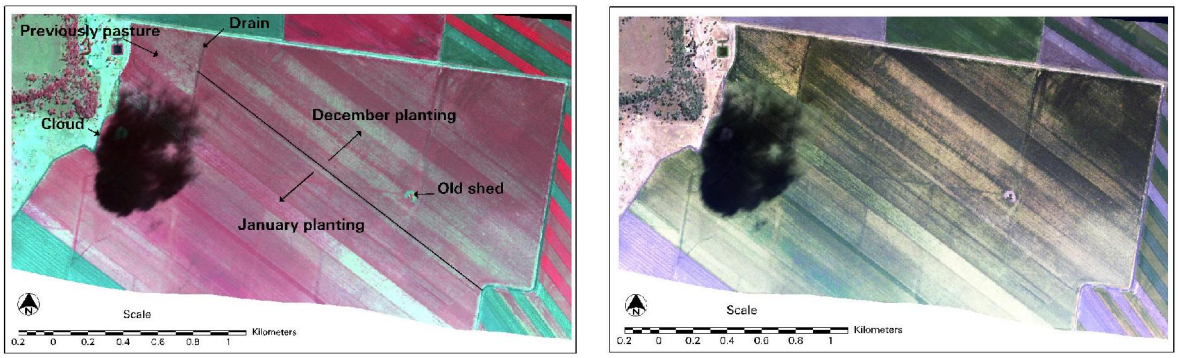

Figure 5. (a) False colour HyMap^{TM} image (red = 800nm, green = 675nm, blue = 545nm) and (b) true colour image (blue = 455nm, green = 558nm, red = 675nm) showing a late season sorghum crop at Well Park, Kaimkillenbun in 2004.

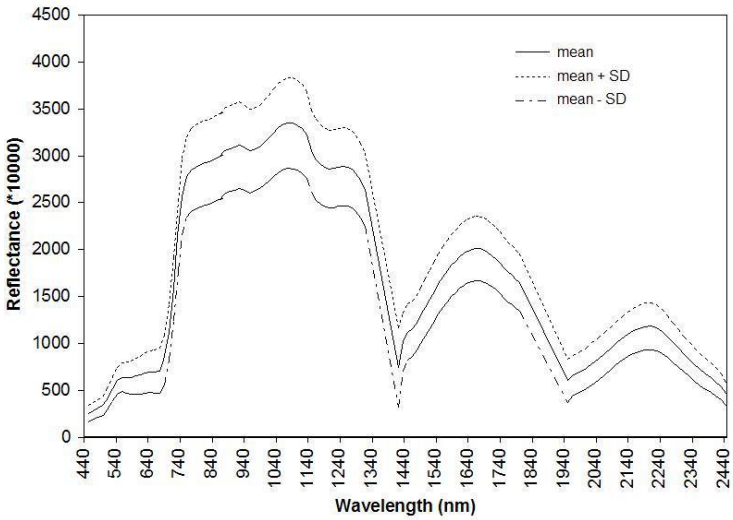

Figure 6. Mean spectral reflectance of all pixels (n=3635) matched to protein samples, with mean reflectance +/- 1 standard deviation superimposed.

The reflectance values for pixels that coincided with grain protein readings were variable and scattered from the mean reflectance value in the NIR region corresponding to 740-1340nm (Figure 6). This region of the spectrum may offer the potential for discriminating differences in grain protein content.

Linear regression analysis, examining the relationship between spectral reflectance and protein content, indicated that a NIR band (747nm) had the highest Pearson correlation coefficient (r=-0.566) (Figure 7). Bands 20-49

corresponding to the red-edge and NIR portions of the spectrum (730-1135nm) also produced very similar correlation coefficient values (r<-0.5) (Figure 8). Principal component analysis (PCA) was performed on the 126 band reflectance data set to obtain a reduced number of independent bands to use in correlation analysis. PC1 and PC2 were not strongly correlated to grain protein content with Pearson correlation coefficients of 0.429 and -0.332 respectively.

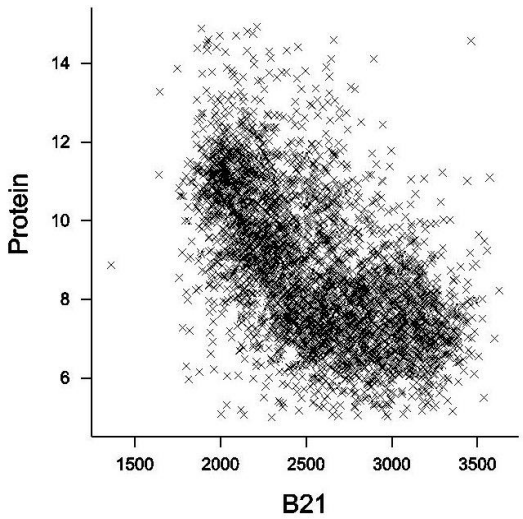

Figure 7. Scatter plot of grain protein content (%) and reflectance in NIR band 21 (747nm).

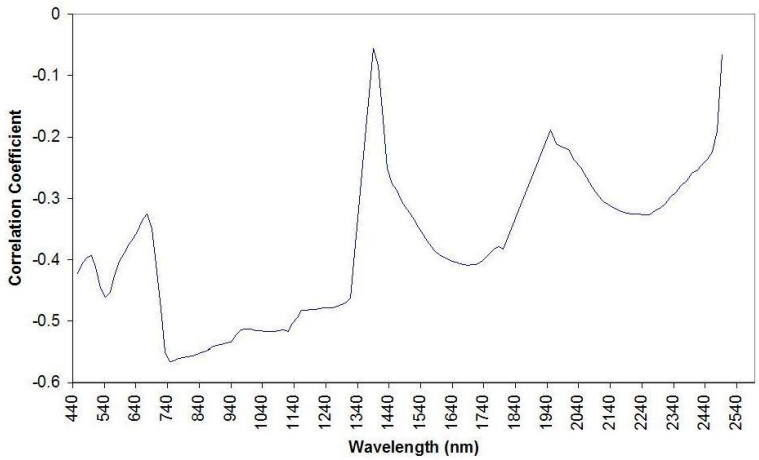

Figure 8. Absolute value of the Pearson correlation coefficient (r) for each of the 126 spectral bands.

An inverse approach was then taken to use the most appropriate Hymap^{TM} bands to estimate grain protein content by applying a stepwise regression using bands in the 20-49 range which corresponded to the region of the spectrum most highly correlated to protein (r<-0.5). The stepwise regression was used to determine the relative importance of each band in predicting grain protein. Table 1 shows the result of changing the number of variables for each stepwise procedure and the corresponding adjusted r^2.

The most parsimonious model for predicting grain protein using HyMap^{TM} spectral data was described by the equation:

Protein=13.334-0.014206(B21)+0.00826(B24)+0.003663(B30)-0.00770(B33)+0.006197(B45)

Table 1. Adjusted Pearson correlation coefficients (r^2_{adj}) between HyMap^{TM} bands and protein values (n = 3635) using a stepwise regression procedure from Well Park, near Kaimkillenbun, in 2004.

Number of terms	Stepwise bands	r^2_{adj} (%)	Gain in r^2_{adj} (%)
1	B21	32.03	-
2	B21, B30	35.28	3.25
3	B21, B24, B30	35.39	0.11
4	B21, B24, B33, B39	36.12	0.73
5	B21, B24, B30, B33, B45	36.29	0.17

Conclusions and Future Research

The results of this study suggest that variations in harvest grain protein content for sorghum did not adequately explain spectral reflectance variations from any one spectral band in a concurrently acquired airborne hyperspectral image.

A continuous portion of the spectrum in the red-edge to NIR region (730-1135nm) exhibited moderate co-variation with the protein content measured at harvest time.

The approach applied in this paper was an exploratory assessment of the relationships between airborne hyperspectral image data and grain protein content. The factors responsible for the observed variations will be examined in more detail, along with more advanced spectral band combinations and stratification of both image and field data. Protein data from the NIR sensor were one area of concern, as these data were not yet calibrated. Another factor, which may have affected the results, was the timing of the image acquisition. Since sorghum is an exposed grain, it was postulated that a near end of season image, after the grain had finished filling, might be an ideal time to remotely sense grain yield and protein. Previous success in linking remotely sensed data with grain protein has focused on acquisition at crop flowering. The effect of crop stage at image acquisition will be investigated using field spectrometry and manual protein samples. The approach used to match protein samples with spectral data in this investigation relies on a high level of accuracy in ground control between image and field data. Any errors in the rectification of the image or the correct tagging of geographic position (as provided by the DGPS) to the protein sensor samples may make the results unreliable. The

extent of this mis-registration may have been reduced by the use interpolated protein data in place of individual point samples.

References

Basnet, B., Apan, A., Kelly, R., Jensen, T., Strong, W., and Butler, D., 2003, Relating Satellite Imagery with Grain Protein Content. *In* 'Proceedings of the Spatial Sciences Conference', Canberra.

Environmental Systems Research Institute Redlands (ESRI), 1998, ArcView Version 3.1, CA, USA.

Lawes Agricultural Trust, 2000, Genstat v.6. (Rothamsted Experimental Station, U.K.)

Leica Geosystems, 2004, Erdas Imagine ® Version 8.7, Atlanta, GA, USA.

Sims, D. A., and Gamon, J. A., 2002, Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sensing of Environment* **84**: 526-537.

Steggenborg, S., and Taylor, R., 2000, Predicting Grain Yield Variability with Infrared Images. *In* 'Proceedings of the 5th Annual Precision Agriculture Conference.' Bloomington, Minnesota, USA.

Stewart, C.M,, McBratney, A.B., and Skerritt, J.H., 2002, Site-specific durum wheat quality and its relationship to soil properties in a single field in northern New South Wales. *Precision Agriculture* **3**: 155-168.

Strong, W.M., and Holford, I.C., 1997, Fertiliser and manures. *In* 'Sustainable Crop Production in the Subtropics' (Ed. A.L. Clarke, P.B. Wylie), pp.214-34 (QDPI: Brisbane).

Strong, W., Kelly, R., Jensen, T., Butler, D., and Town, B., 2003, Within-field protein variation in the northern grains region. *In* 'Proceedings of the 11th Australian Agronomy Conference, Geelong, Victoria,' (Australian Society of Agronomy). www.regional.org.au/au/asa/2003/p/7/kelly.htm

Yang, C., and Everitt, J.H., 2002, Relationships between Yield Monitor Data and Airborne Multi-date Multispectral Digital Imagery for Grain Sorghum. *Precision Agriculture*, **3**: 373–388.