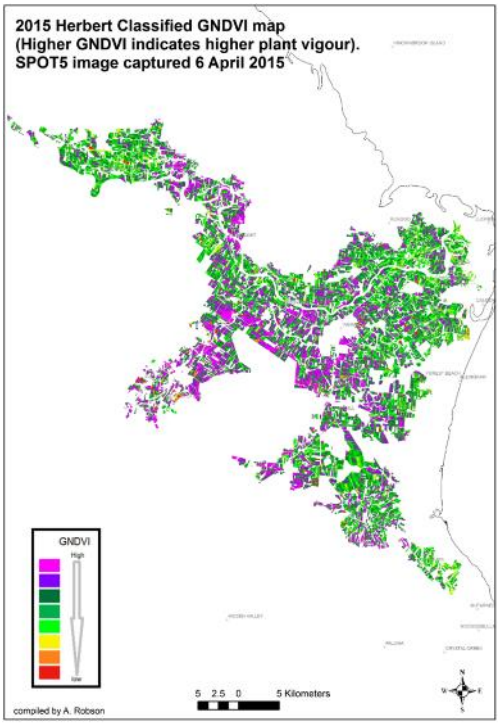
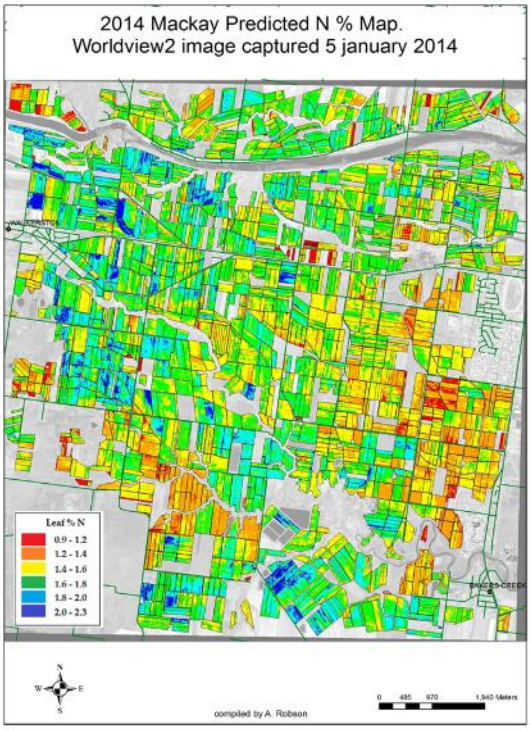
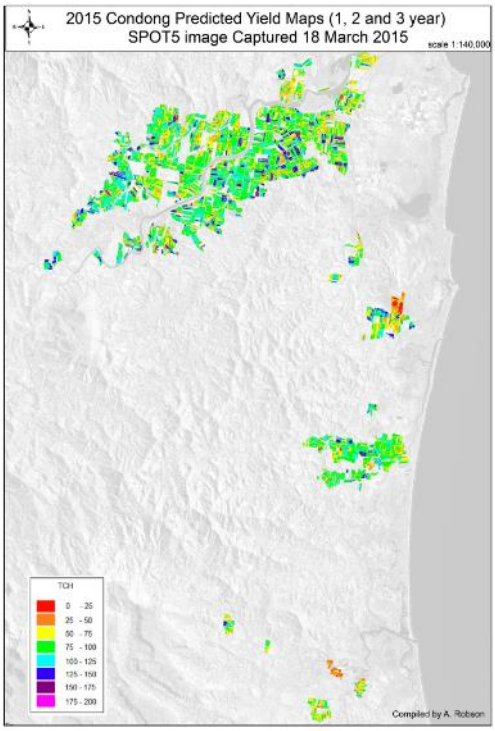
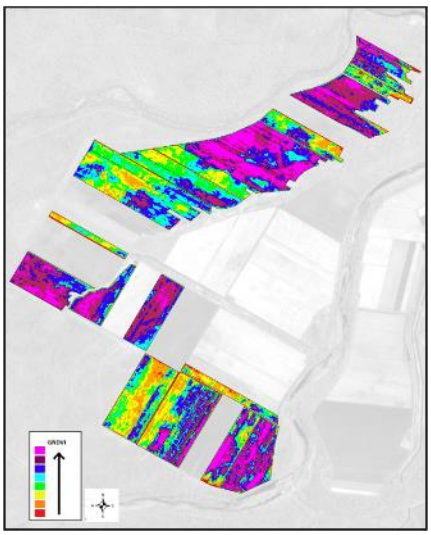




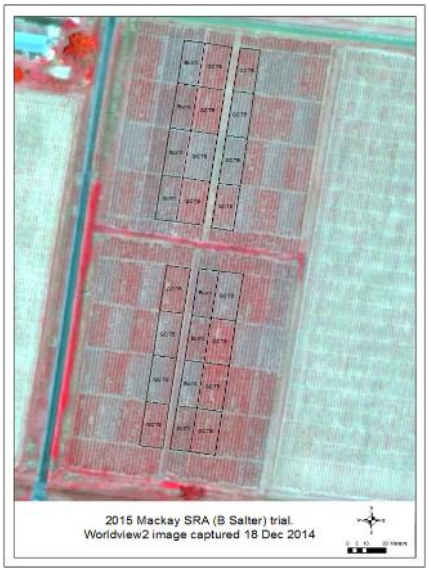


FINAL REPORT- SRA PROJECT DPI025
DEVELOPING REMOTE SENSING AS AN INDUSTRY WIDE YIELD FORECASTING, NITROGEN MAPPING AND RESEARCH AID

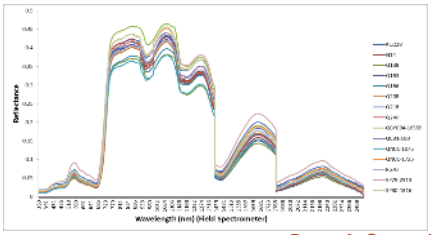
ASSOCIATE PROFESSOR ANDREW ROBSON, DR MOSHIUR RAHMAN AND JASMINE MUIR.

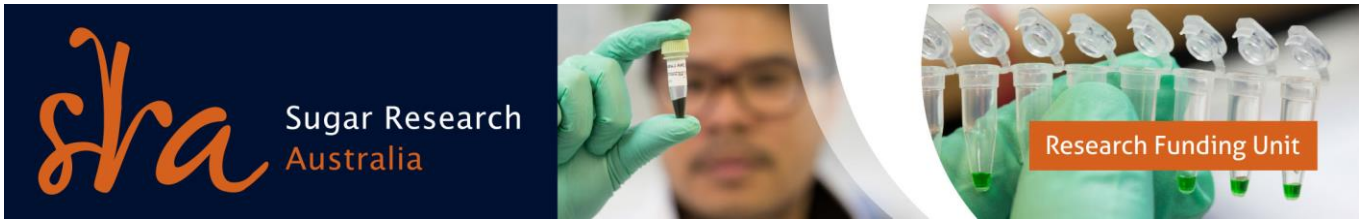


University of New England — Precision Agriculture Research Group



Dr Moshiur Rahman, A/Prof Andrew Robson, Dr Niva Verma and Dr Greg Falzon





SRA Research Project Final Report

SRA Project Code	**2013/025 (DPI025)**		
Project Title	Developing remote sensing as an industry wide yield forecasting, nitrogen mapping and research aid.		
Key Focus Area in SRA Strategic Plan	4		
Research Organisation(s)	University of New England		
Chief Investigator(s)	Assoc. Prof. Andrew Robson		
Project Objectives	This project has three main objectives: 1) To refine the accuracies and delivery of crop, farm and regional scale yield forecasts derived from satellite imagery; 2) To evaluate multispectral and hyper-spectral tools as a method for screening research and breeding trials and; 3) Evaluate multispectral and hyper-spectral tools as a method for measuring canopy nitrogen status.		
Milestone Number	9 (Final Report)		
Milestone Due Date	1 October 2016	**Date submitted**	22 November 2016
Reason for delay (if relevant)	Awaiting final analysis and writing the final report.		
Milestone Title	Final Report		
Success in achieving the objectives	☒ Completely Achieved ☐ Partially Achieved ☐ Not Achieved		
SRA measures of success for Key Focus Area (from SRA Strategic Plan)	Adoption of PA technology and techniques.		

Please cite as: Robson A. 2016. Developing remote sensing as an industry wide yield forecasting, nitrogen mapping and research aid: final Report Project 2013/025. Sugar Research Australia Limited.

PART A

Section 1: Executive Summary

The science of Earth Observation (EO) is a rapidly developing discipline that has seen an unprecedented rise in remote sensing technologies and application development, including those in agriculture. The Australian sugarcane industry has seen a steady increase in the development and adoption of remote sensing applications over the last decade, predominantly as a result of investment by SRDC and then Sugar Research Australia (SRA). SRA project (DPI025), with collaborative support from Australia's sugar mills, grower's, research institutions and extension agencies, has been at the forefront of this evolution, evaluating modern remote sensing technologies and novel analysis methodologies for improved in-season yield forecasting and Nitrogen management, both issues identified as priorities by the industry.

Accurate yield forecasting at the regional level is vital for the Australian sugar industry as it supports decision making processes including harvest scheduling, product handling and forward selling. At the farm scale, accurate yield mapping provides growers with a stronger understanding of in-crop variability, both spatially and temporally, thus supporting the adoption of precision agricultural practices to maximize productivity. Currently, yield forecasting within the Australian sugarcane industry is undertaken by visual inspection or destructive sampling by either growers or mill funded productivity officers. Although relatively accurate, these methods are labour intensive and are subject to the influences of varied seasonal climatic conditions, crop age and human error. Remote sensing technologies have evolved across many cropping systems as an accurate 'tool' for measuring in-season performance and for the prediction of yield, pre- harvest. This project, built on the initial findings of DPI021, further developed regional yield prediction algorithms derived from SPOT satellite imagery for 11 growing regions: Broadwater, Harwood, Condong, Isis, Bundaberg, Maryborough, Burdekin, Herbert, Tully, South Johnstone and Mulgrave); investigated novel statistical methods for improving prediction accuracies at the block level; and investigated time-series remote sensing based models for improved forecasting accuracies earlier in the growing season.

Improved Nitrogen (N) management is also an important consideration for the Australian sugarcane industry. N is an essential crop input, assisting plant growth, tillering and stalk elongation. However, excess N can have a detrimental effect on sucrose levels and sugar quality as well as environmentally, through soil acidification, surface runoff, leaching and volatilization. Currently, soil and plant tissue testing are the preferred methods of determining in-season crop nitrogen status. However, as these methods are destructive, time consuming and expensive to analyze, they are not widely used. Also, being point source, the results may not accurately represent the N status of the entire crop. This project evaluated remote sensing as a non-destructive alternative for measuring foliar nitrogen (N) concentration of sugarcane at the out- of -hand stage. The platforms evaluated included multispectral satellite sensors: Landsat, SPOT5, WorldView2, WorldView3 and GeoEYE as well as hyperspectral airborne and field based sensors. Through direct collaboration with SRA, University of Queensland and Farmacist, this evaluation included a number of growing season, cultivars, ratoons and regions (Tully, Burdekin and Mackay).

The main out comes of this project are summarised as:

- Continued development of yield forecasting algorithms from annual SPOT satellite captures tonow include 11 Australian growing regions (~60,000 individual crops).

- Evaluation of an alternative yield forecasting model based on a time-series approach. This method, using 15 years of Landsat (98 images) acquired over the Bundaberg region produced a number of encouraging outcomes including:
    - Identified the influence of seasonal variation on the annual rate of crop development and maximum crop vigour achieved. This variation, predominantly influenced by rainfall, provides an important crop growth rate 'benchmark' that can indicate the incidence of widespread abiotic or biotic constraints in future seasons.
    - Identified Bundaberg crops, on average, achieved peak growth early to mid- April, indicating this period as optimal for the acquisition of single image captures for the purpose of yield forecasting and deriving yield maps;
    - Enabled an improved quadratic yield forecasting model to be developed that produced a higher regional yield prediction accuracy for the Bundaberg region than that achieved from the single SPOT capture method. The model also allowed for predictions to be made earlier in the growing season (i.e. from December to April);
- Evaluation of univariate and multivariate statistical models that incorporate historic imagery (SPOT 5 GNDVI) and crop attribute data (yield, crop, class) for improved yield prediction accuracies at the individual block level (Bundaberg, Condong and Herbert).
- Identifying the vegetation index (N2RENDVI), a ratio of canopy reflectance in the red- edge and Near infrared spectral regions, as consistently producing a higher correlation to foliar Nitrogen concentration (%) than the commonly used NDVI.
- Demonstrating the potential of this technology for monitoring replicated trials, particularly treatment response, as well as those from abiotic and biotic constraints. Understanding the later provides crucial information when analysing and comparing individual plot results.

Section 2: Background

Project DPI025, is a progression of the SRDC funded project "*DPI021: Remote sensing- based precision agriculture tools for the sugar industry. Sugar Research Development Corporation (SRDC). Final Project Re*port" (Robson et al., 2013). DPI021 successfully identified satellite remote sensing technologies as an effective tool for a number of applications including:

- mapping the spatial and temporal variability of crop performance;
- derivation of surrogate yield maps and for regional yield forecasting; and
- monitoring of abiotic and biotic constraints.

More specifically, DPI021 identified SPOT 5 imagery as optimal for mapping Australian sugarcane due to its moderate spatial resolution (10 m); image extent (3600 km^2 per image scene); repeat capture time; 4 band spectral resolution and cost ~$1/$km^2$. Whilst, high resolution multispectral imagery (i.e. IKONOS with a 3.2 m) was required for identifying sub paddock constraints such as pest and disease outbreaks. The greenness normalised difference vegetation index (GNDVI) derived from imagery acquired between March – May produced stronger correlations to final crop yield than other vegetation indices and displayed less saturation than the commonly used index NDVI. At a regional level, yield prediction algorithms derived from GNDVI calculated from single SPOT5 capture were less susceptible to varying cultivars and crop classes than from the influence of growing location, seasonal variation and timing of image capture. Additionally, high accuracies for the prediction of individual crop yield were only acceptable for the Bundaberg region.

Although DPI021 successfully delivered on its objectives, the preliminary results clearly indicated that additional analysis methodologies were required to address the influence of location and seasonal effect on the accuracies of regional forecasts and to improve accuracies at the individual block level. Additionally, since the completion of the DPI021 the satellite platforms SPOT 5, IKONOS and QuickBird have been retired and therefore additional research was required to extrapolate the algorithms to the new suite of platforms including SPOT 6 & 7 and Worldview 2 & 3.

Project DPI025, provided the opportunity to continue the momentum and application development of DPI021, but also include the evaluation of remote sensing technologies to support improved nitrogen management.

The industry issues:

Accurate with-in season yield forecasting is essential for Australian sugar mills as an indicator of annual production and to support harvesting, milling and forward selling decisions. At the block or farm scale, accurate yield mapping provides growers with a stronger understanding of in-crop variability, thus supporting the adoption of precision agricultural practices. Currently, predictions of in- season yield are made by experienced productivity officers whom visually estimate, or manually cut samples from representative crop locations across a growing region. Although this process generally produces accuracies of 95% (Pitt pers. comm. 2011), it is time consuming and errors can occur when the sample locations do not truly represent the greater region, especially when influenced by constraints such as flood, pest or disease. Harvester mounted yield monitors are also used in some growing regions (Jensen *et al*. 2012), but these too can be inaccurate with data only available post-harvest, therefore too late to assist with-in season decision making. Climate forecasting has also been identified as a viable tool for the prediction of regional sugarcane yield (Everingham *et al*. 2016), but it is unable to provide predictions at a small spatial scale such as the with-in farm or individual crop level. As detailed in Robson *et al.* (2012) and in the final project report submitted for SRA project DPI021, satellite imagery has been identified as an accurate method for predicting regional sugarcane yield both

internationally and domestically. However, prediction accuracies can be severely influenced by irregular seasonal weather conditions, timing of image capture and look angle of the satellite. In an attempt to compensate for these sources of inaccuracy, a preliminary study using time series Landsat imagery similar to that employed by Mulianga *et al*. (2013); Duveiller *et al*. (2013) and Morel *et al*. (2014), was undertaken in an attempt to develop a more robust methodology for predicting regional yield for the Australian sugarcane industry. Additionally, to improve yield prediction accuracies at the individual cane block level, a number of statistical models were evaluated over the Bundaberg and Condong growing regions using temporal productivity and attribute data for each individual crop (provided by the respective mills) as well as corresponding canopy reflectance data.

As well as yield forecasting, Nitrogen management has also been identified as a key industry objective for the Australian sugar industry. A recent environmental report identified nitrogen runoff from agricultural systems, particularly sugarcane, as a major cause of harm to the Great Barrier Reef (Australian government 2014; 3rd Reef Plan Report Card (State of Queensland 2013C). Apart for the environmental concern, improved Nitrogen management has the direct benefit to growers by reducing input costs, whilst retaining yield potential. In order to achieve improved N management at the industry level, growers need to understand the spatial and temporal fertiliser requirements of their crop, while at the industry level, some understanding of nitrogen usage at the region and subregional level is also required. Currently, there are few options for obtaining this information. Soil sampling and tissue testing can provide an accurate measure of plant available nutrition during a given crop cycle, but this is time consuming, expensive and history shows that few growers actually base their fertiliser applications on regular soil or leaf tissue sampling (Johnson, 1995; Wood *et al*. 2003).

Remote sensing has again been identified as one technology that can provide a non-invasive measure of foliar nitrogen concentration in a range of agricultural systems including corn (Bagheri *et al*. 2013; Tahir *et al*. 2013; Barker and Sawyer 2012); wheat (Rodriguez *et al*. 2006; Johnson and Raun *et al.* 2003; Raun *et al*. 2002; Zillmann *et al*. 2006; Yao *et al*. 2013; Erdle *et al*. 2011; Tilling *et al*. 2007); cotton (Rao *et al*. 2008; Raper *et al*. 2013): potatoes (Goffart et al. 2004; Nigon *et al*. 2014) and rice (Rao *et al*. 2008; Dunn 2012), and to a lesser extent sugarcane crops (Jackson *et al.* 1980; Abdel-Rahman *et al*. 2008; Portz *et al*. 2012 and Lofton *et al*. 2012). From this extensive research, higher concentrations of N have been identified to increase absorption in the visible spectral red region (630 - 690 nm) due to increased chlorophyll concentration, increase reflectance in the near infrared (770 - 895 nm) and mid infrared (1300- 2500 nm) spectral regions, and cause a shift in slope of the transitional red-edge region (700 - 780 nm). The integration of these discrete spectral regions, through the derivation of vegetation indices, has also been shown to be strongly correlated to changes in vegetation condition. The research conducted through this project compares three remote sensing technologies: very high resolution (VHR) satellite imagery, field spectroscopy and airborne hyperspectral imagery for accurately determining leaf nitrogen content in Australian sugarcane crops, thus determining its potential for supporting improved nitrogen management in the Australian sugar industry.

The Australian sugar industry is well suited for the broad scale adoption of remote sensing applications. For over a decade the Australian sugar industry have employed a fully integrated geographical information system (GIS) for recording and managing grower information at the individual crop level, making it one of the most progressive industries in Australia, if not globally. This system, predominantly developed by AGTrix (http://www.agtrix.com/), has greatly increased the integration of mill and productivity datasets, thus enabling greater efficiencies in data retrieval and analysis of client information, improved the coordination and planning of the cane harvest, and the identification of consignment errors (Markley *et al.* 2008). In terms of the research presented in this report, the annual updating of individual crop boundaries and associated attribute information (class, cultivar, plant date etc) provides a crucial data layer that supports the use of

remote technologies. Essentially, the digitised crop boundaries for each growing region enables the efficient and accurate extraction of canopy specific reflectance data for each individual crop. This spectral data can then be processed, manipulated into an array of vegetation indices, converted into specific parameters such as crop vigour, derived yield and foliar N concentration and re-distributed back to each mill in a complimentary GIS format. Additionally the compilation of current and historic productivity and associated attribute information for each individual crop, along with the extracted canopy reflectance information, creates the unique opportunity to undertake intensive statistical interrogation to better understand the spatial and temporal trends in productivity at the block, farm and regional level.

Project DPI025 aimed to deliver on three main objectives:

1. ***To refine the accuracies and delivery of crop, farm and regional scale yield forecasts derived from satellite imagery;***

2. ***To evaluate multispectral and hyper-spectral tools as a method for screening research and breeding trials and;***

3. ***Evaluate multispectral and hyper-spectral tools as a method for measuring canopy nitrogen status.***

Section 3: Outputs and Achievement of Project Objectives

Project objectives, methodology, results and discussion

Objective 1: To refine the accuracies and delivery of crop, farm and regional scale yield forecasts derived from satellite imagery.

Regional Yield Prediction from Single Date SPOT Satellite Imagery:

For the yield prediction component of this research, SPOT imagery was acquired over 6 Queensland growing regions: Gordonvale (Mulgrave), Herbert, Burdekin, Bundaberg, Isis and Maryborough; and 3 New South Wales (NSW) growing regions: Condong, Broadwater and Harwood during the 2013 and 2014 growing season. This was repeated again for the 2015 and 2016 growing seasons, with the addition of the South Johnstone and Tully regions, bringing the total number of individual crops included in the study to approximately 60,000 over the 11 growing regions (Figure 1).

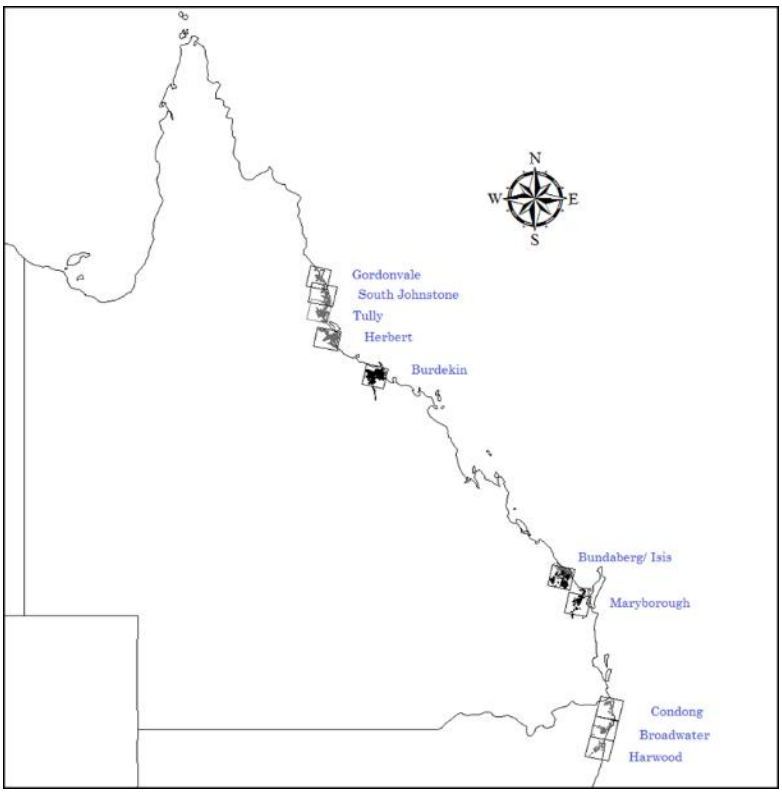


Figure 1: 11 Sugarcane growing regions covered during 2014/2015 for the yield forecasting component. Note that South Johnston and Tully were added to the project in 2015.

Following the outcomes of DPI021, all image captures were acquired between March - May, the period of stabilized growth between vegetative and senescence. As seen from Table 1, imagery capture generally occurred within this timeframe, except for those occasions where continual cloud cover delayed captures (i.e. until June).

Table 1: Satellite imagery acquired over the course of the project for yield prediction. All imagery was acquired from SPOT 5, 6, and 7 satellites.

	Year			
Region	**2013**	**2014**	**2015**	**2016**
North Bundaberg & Isis	25/04/2013 (SPOT 5)	19/04/2014 (SPOT 5)	18/03/2015 (SPOT 5) 22/04/2015 (SPOT 7)	04/05/2016 (SPOT 5) 29/05/2016 (SPOT 6)
Burdekin	05/05/2013 (SPOT 5) 29/12/2013 (SPOT 5)	08/05/2014 (SPOT 5)	15/03/2015 (SPOT 5)	11/04/2016 (SPOT 7)
Herbert	25/05/2013 (SPOT 5)	09/07/2014 (SPOT 5)	06/04/2015 (SPOT 6)	26/05/2016 (SPOT7) 15/06/2016 (SPOT 7)
Maryborough	30/04/2013 (SPOT 5)	29/04/2014 (SPOT 5)	18/03/2015 (SPOT 5)	22/04/2016 (SPOT7) 29/05/2016 (SPOT6)
Gordonvale (Mulgrave)	18/04/2013 (SPOT 5)	23/05/2014 (SPOT 5)	02/05/2015 (SPOT 6)	26/05/2016 (SPOT 6)
South Johnstone	N/A	N/A	15/03/2015 (SPOT 5)	08/06/2016 (SPOT 7)
Tully	N/A	N/A	02/05/2015 (SPOT 6)	08/06/2016 (SPOT 7)
South NSW (Harwood)	25/04/2013 (SPOT 5)	01/05/2014 (SPOT 5)	18/03/2015 (SPOT 5)	22/04/2016 (SPOT 7)
Mid NSW (Broadwater)	16/05/2013 (SPOT 5)	24/04/2014 (SPOT 5)	08/03/2015 (SPOT 5)	01/04/2016 (SPOT 7)
North NSW (Condong)	20/04/2013 (SPOT 5)	19/04/2014 (SPOT 4)	18/03/2015 (SPOT 5)	03/05/2016 (SPOT 6) 04/05/2016 (SPOT 7) 23/05/2016 (SPOT 7)

From Table 1, 3 satellite platforms SPOT 5, SPOT 6 and SPOT 7 are indicated. The SPOT 5 satellite was decommissioned in March 2015 and replaced with SPOT 6 & SPOT 7. These sensors offer higher spatial resolution (6 m), compared to the 10 m provided by SPOT 5, and a spectral resolution (Blue, Green, Red and NIR) that fortunately supported the continued development of the existing SPOT 5 yield forecasting algorithms. The cost of SPOT 6 & 7 imagery was higher than SPOT 5 i.e. \$2.25/km^2 versus \$1.05/km^2. However, with no longer the stipulation to purchase a full scene (60 * 60 km), an area of interest (AOI) could be refined to include only relevant targets i.e. sugarcane crops. Whilst for the larger growing regions such as Bundaberg/Isis this change resulted in a minimal cost increase, 5 of the growing regions benefited from a price reduction. As an example, Figure 2 compares the refined SPOT 6 & 7 capture areas for the three NSW growing regions (Broadwater, Harwood and Condong) to that previously required from SPOT 5, ultimately reducing the cost of imagery by \$4,200.

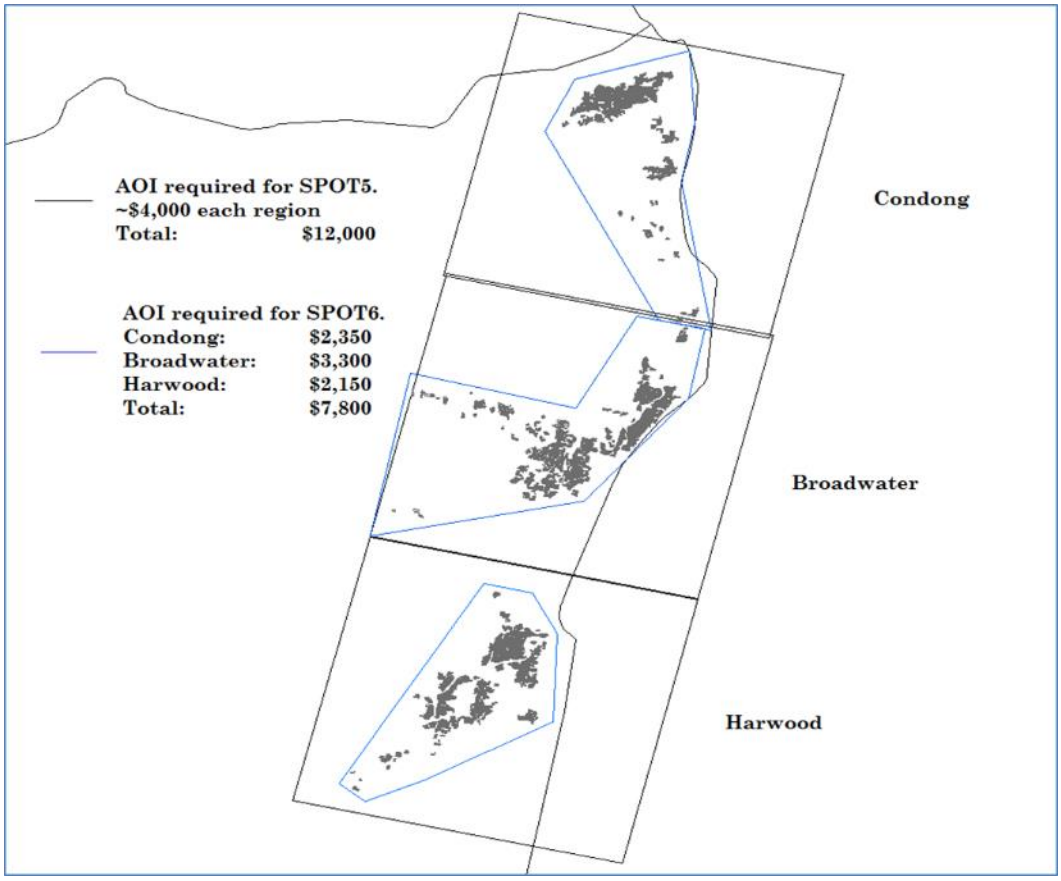


Figure 1: Comparison of the image coverage and associated costs of SPOT 5 (60 * 60 km) versus SPOT 6 & 7 for the three NSW growing regions (Harwood, Broadwater and Condong). * AOI refers to area of interest.

The annual acquisition of SPOT imagery has supported the development of the regional yield forecasting models that for some regions has extended up to 6 growing seasons. From Table 2 and Table 3, the accuracies of prediction for each region and for each growing season, including those obtained for the 2010-2012 seasons (from project DPI021), can be seen. A negative difference (%) value indicates an under prediction, whilst a positive value indicates an over prediction. A '0.00' difference indicates an exact prediction of yield. At the time this final project report was submitted, harvesting for the 2016 season had not been completed and therefore prediction accuracies for 2016 are not included. Also, as the development of the prediction algorithms for the Tully and South Johnstone growing regions only commenced in 2015, predictions have not been included.

Table 2: Comparison of actual yield (TCH) to predicted yield (TCH) for Queensland growing regions. Note * refers to results from the previous DPI021 project.

	Region	Number of Crops	Predicted Yield (TCH)	Actual Yield (TCH)	Difference (%)
2010*	**North Bundaberg**	3544	80	82	-2.44
	Isis	2772	84	84	0.00
2011*	**North Bundaberg**	3824	80	73	9.59
	Isis	4205	98	83	18.07
	Burdekin	4999	119	120	-0.83
	Herbert	8596	51	55	-7.27
2012*	**North Bundaberg**	3217	88	89	-1.12
	Isis	4000	93	96	-3.13
	Burdekin	6921	110	105	4.76
	Herbert	15463	75	72	4.17
2013	**North Bundaberg**	3348	77	72	6.94
	Isis	3227	82	76	7.89
	Burdekin	6876	121	105	15.24
	Herbert	11184	97	74	31.08
	Maryborough	2718	72	61	18.03
	Gordonvale (Mulgrave)	3127	102	93	9.68
2014	**North Bundaberg**	4016	77	72	6.94
	Isis	3502	72	76	-5.26
	Burdekin	6876	112.5	113.3	-0.71
	Herbert	12968	64.5	74.4	-13.31
	Maryborough	3495	62	55.5	11.71
	Gordonvale (Mulgrave)	3004	66.6	91	-26.81
2015	**North Bundaberg**	4953	92	87	5.75
	Isis	5401	98	87.5	12.00
	Burdekin	7441	134	118	13.56
	Herbert	14913	74.5	79	-5.70
	Maryborough	3592	59	76	-22.37
	Gordonvale (Mulgrave)	3254	74	87	-14.78

Due to the slightly different farming system of the NSW regions i.e. cane crops that are grown for either 1 year or 2 to 3 years, yield prediction algorithms were derived for each age class (Table 3).

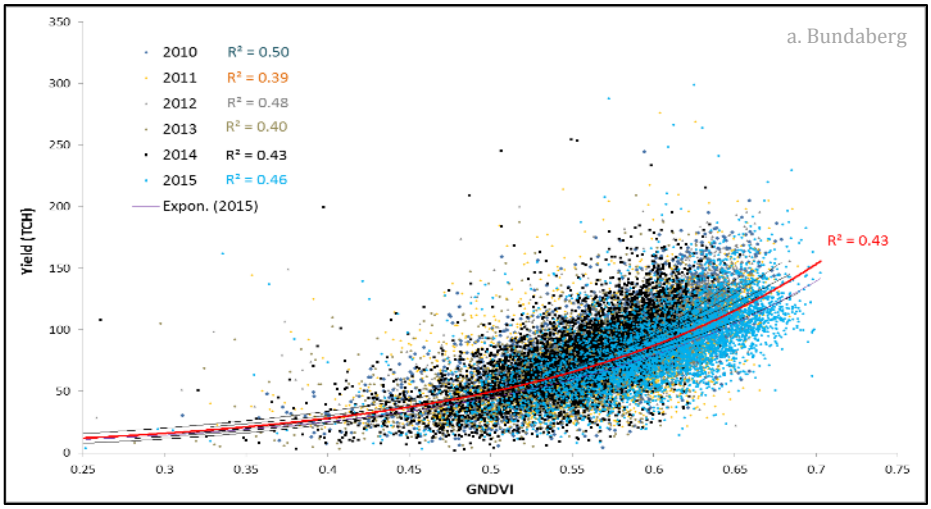
Table 3: Comparison of actual yield (TCH) to predicted yield (TCH) for New South Wales growing region.

		Predicted Yield (TCH)		Actual Yield (TCH)		Difference (%)	
		1 Year	2/3 Year	1 Year	2/3 Year	1 Year	2/3 Year
2013	**Condong**	59	80	70	95	-15.71	-15.79
2014	**South NSW (Harwood)**	79	87.7	120.4	115.1	-34.39	-23.81
	Mid NSW (Broadwater)	77.7	92.8	88.8	110.4	-12.50	-15.94
	Condong	71.4	94.5	116.3	119.8	-38.61	-21.12
2015	**South NSW (Harwood)**	94.2	142.6	120.4	157.6	-21.76	-9.52
	Mid NSW (Broadwater)	55.7	89	100.2	142.7	-44.41	-37.63
	Condong	76.1	121.5	105.5	175	-27.87	-30.57

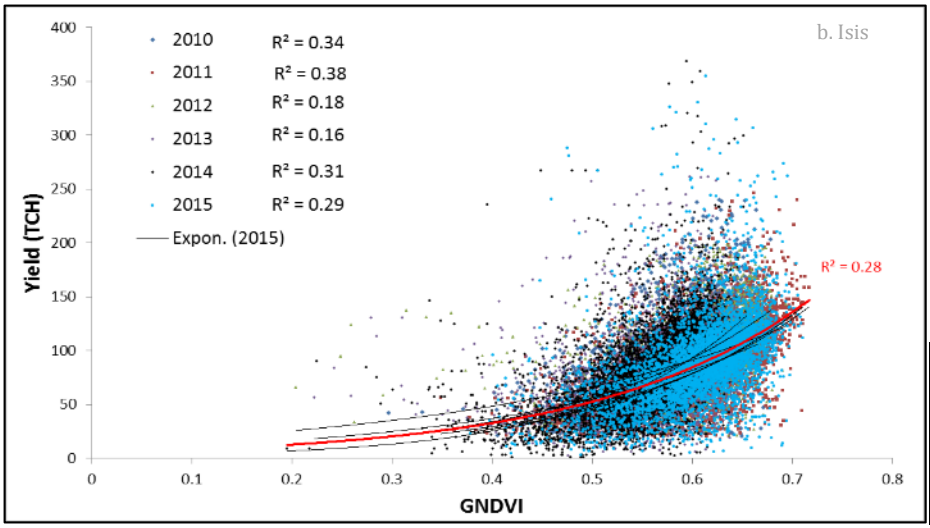
For the Bundaberg and Isis regions, the annual yield predictions have been highly comparable to actual harvested yield i.e. ~10%, a result likely attributed to the consistent seasonal relationship between GNDVI and yield (TCH) (Figure 3 a and b). During the 2015 growing season, the predicted yield for the Bundaberg region (SPOT 5 captured 18th March) was abnormally low (~ 27 %), with a second image acquisition on the 22nd April producing a better result (5.7 % over predicted). This result provided the first indication of the influence of seasonal variation on the rate of annual crop growth and the limitations of the single capture approach. This result is further discussed in the following section *'An improved 'time-series' approach for sugarcane yield forecasting'*.

The prediction accuracies for the Burdekin and Herbert regions were less consistent, with accuracies ranging from near exact for the 2011 (- 0.83 %) and 2014 (- 0.71 %) seasons, to greater than 10 % for 2013 (15.2 %) and 2015 (13.6 %). Similar results were identified for the Herbert with the 2013 season producing the highest inaccuracy (31.08 %). Again, these results indicated the influence of varied seasonal conditions and timing of image capture on the accuracy of regional yield predictions. Figure 3 c and d, clearly shows the variation in the linear relationship between GNDVI and Yield (TCH) achieved for each season. This variation is even more apparent for the Maryborough and Mulgrave regions (Figure 3 e and f), with the 2013 points for the later region being very much separated from the 2014 and 2015 data. The poor predictions made for these two regions over the three year period, each averaging ~ 17 % difference, was likely attributed to highly variable seasons.

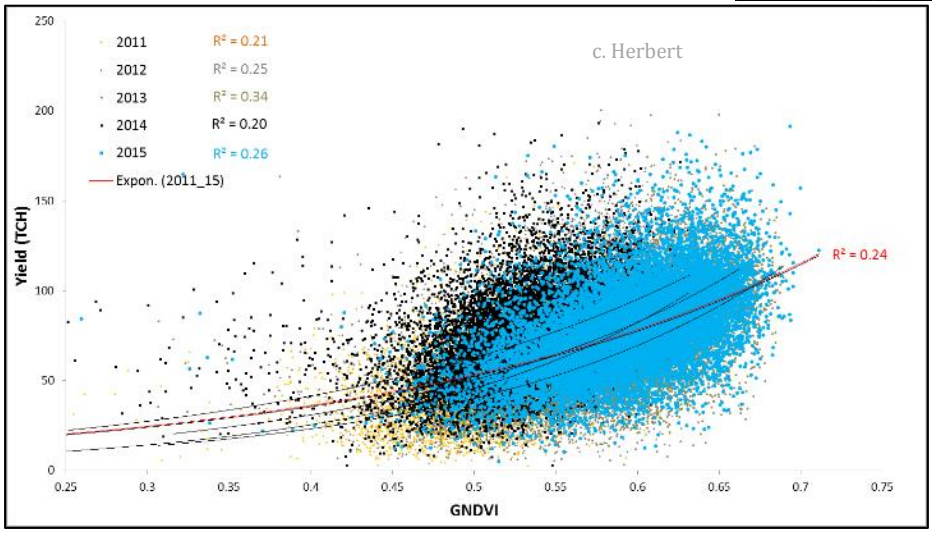
From Table 3, the prediction accuracies for the three NSW regions have been poor. This result is believed to be attributed to a number of influences including the development of the algorithm spanning years of highly variable production, the influence of the 2 year growing system and timing of image capture. The 2012 and 2013 growing seasons produced a very different point cloud distribution and exponential linear relationship between crop GNDVI and yield (TCH) compared to the 2014 and 2015 season (Figure 3 g to l). The initial yield forecasting algorithm was developed from 2012/ 2013 data, both poor growing years, and therefore it was unable to compensate for the improved growing conditions observed during 2014 and 2015. Additionally the image capture timing for the 2015 season (8th and 18th of March) was early compared to previous years.



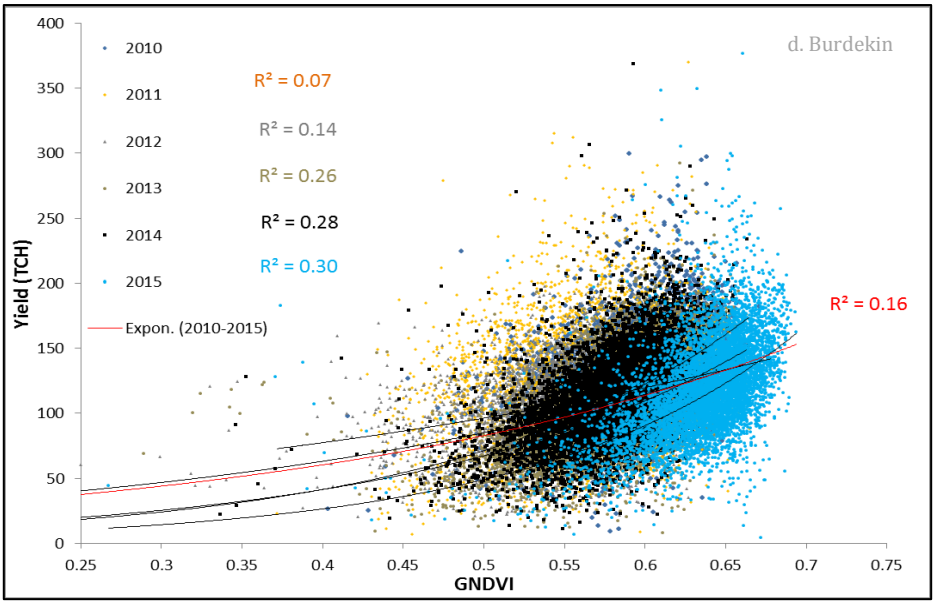
Year	Av GNDVI	image capture
2010	0.5669	10-May-10
2011	0.5697	27-Mar-11
2012	0.5854	1-Apr-12
2013	0.5602	25-Apr-13
2014	0.5513	19-Apr-14
2015	0.5928	22-Apr-15



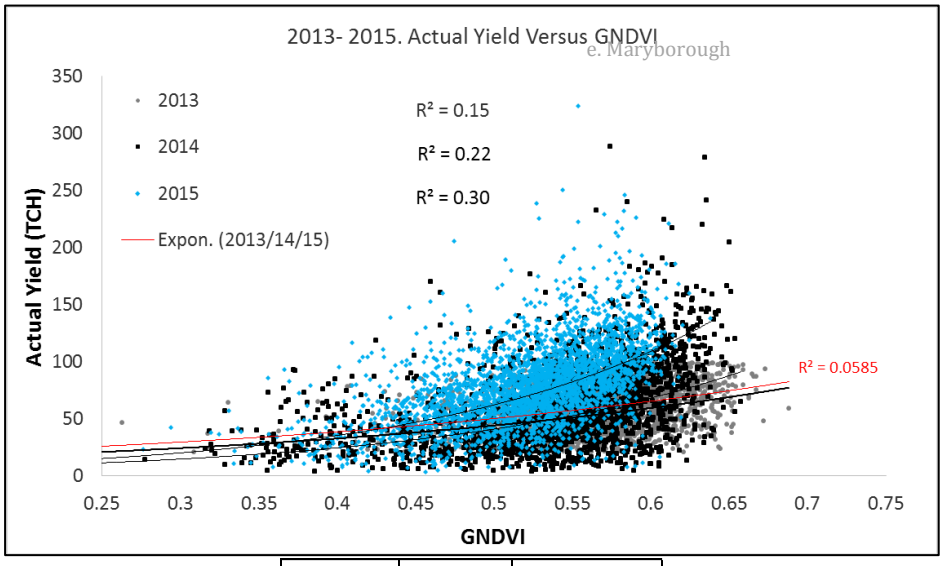
Year	Av GNDVI	image capture
2010	0.5762	10-May-10
2011	0.6039	27-Mar-11
2012	0.5968	1-Apr-12
2013	0.5715	25-Apr-13
2014	0.5550	19-Apr-14
2015	0.6068	22-Apr-15



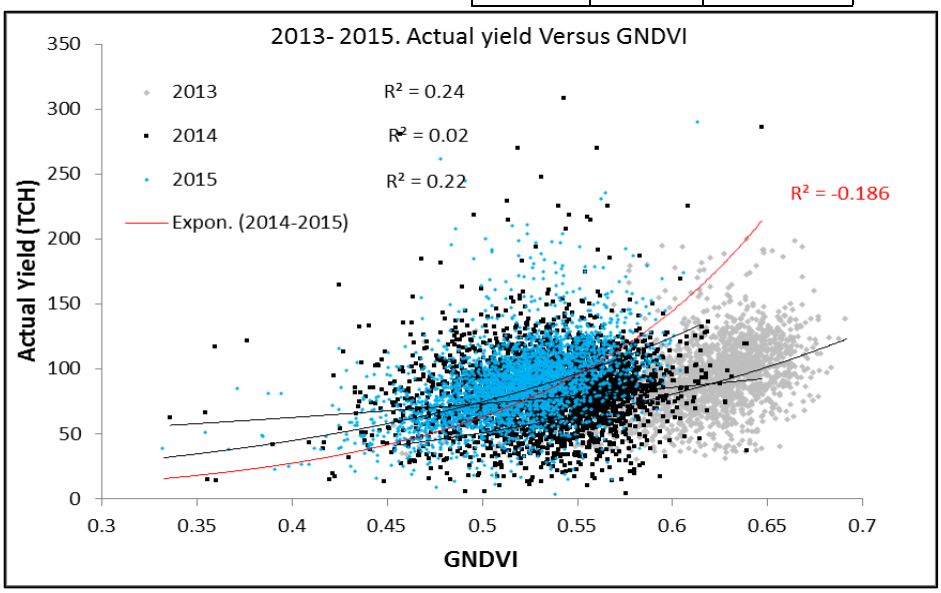
Year	Av GNDVI	image capture
2011	0.4898	5-May-11
2012	0.5593	4-Apr-12
2013	0.6011	25-May-13
2014	0.5294	9-Jul-14
2015	0.5858	6-Apr-15



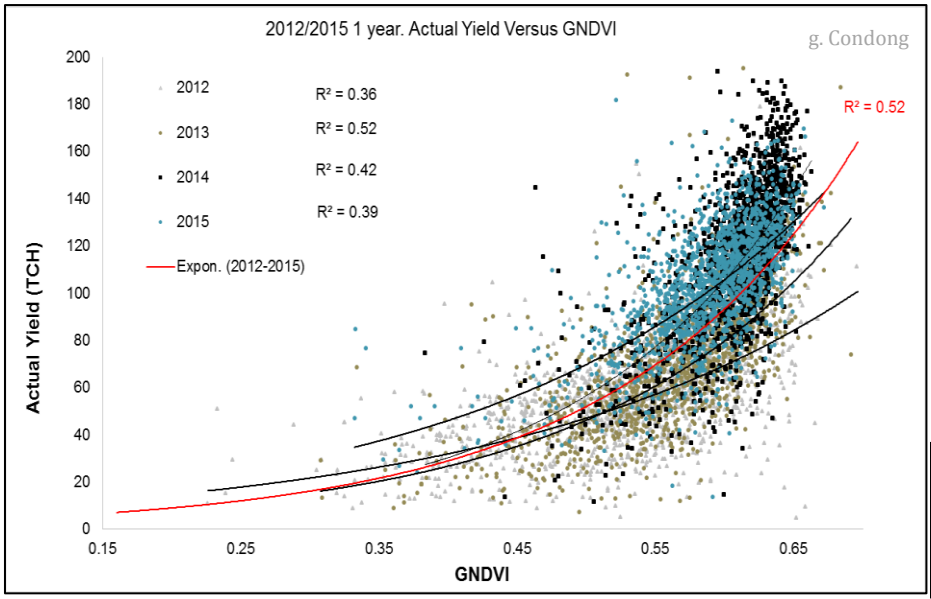
Year	av GNDVI	image capture time
2010	0.59186	14-May-10
2011	0.57444	22-Apr-11
2012	0.56002	16-May-12
2013	0.58017	5-May-13
2014	0.57797	29-Apr-14
2015	0.63642	15-Mar-15



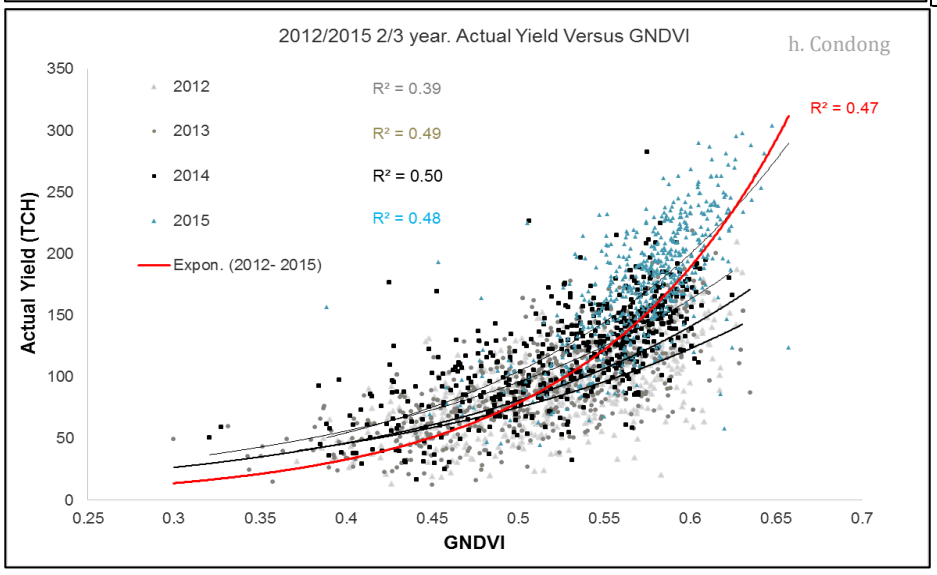
Year	Av GNDVI	Image capture
2013	0.5712	30-Apr-13
2014	0.5414	29-Apr-14
2015	0.5223	18-Mar-15



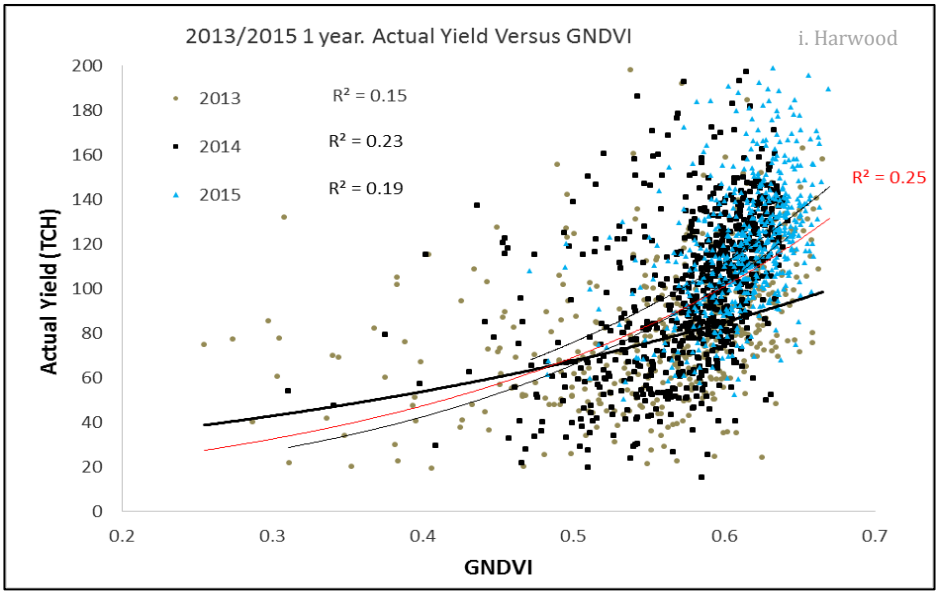
Year	Av GNDVI	image capture
2013	0.6152	18-Apr-13
2014	0.5354	23-May-14
2015	0.5195	2-May-15



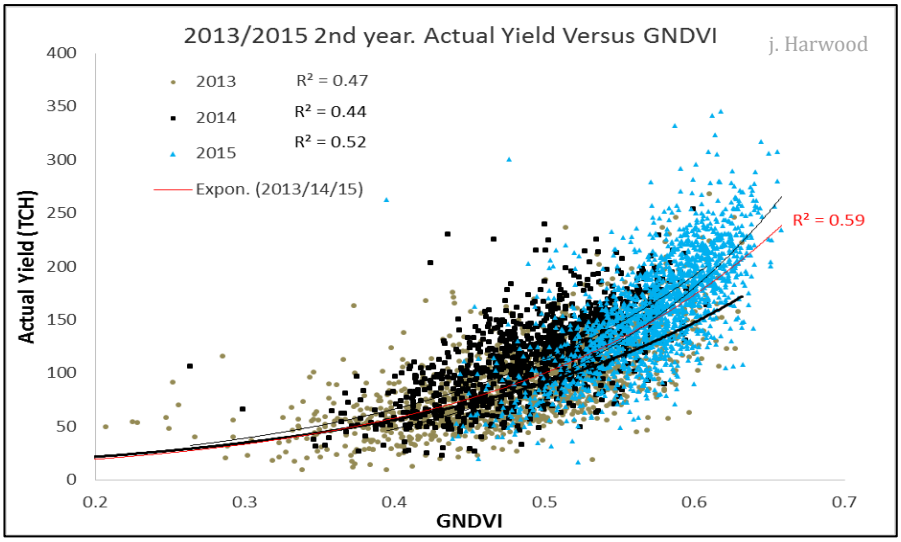
Year	Av GNDVI	image capture
2012	0.5374	29-Feb-12
2013	0.5597	13-Apr-13
2014	0.6081	19-Apr-14
2015	0.5883	18-Mar-15



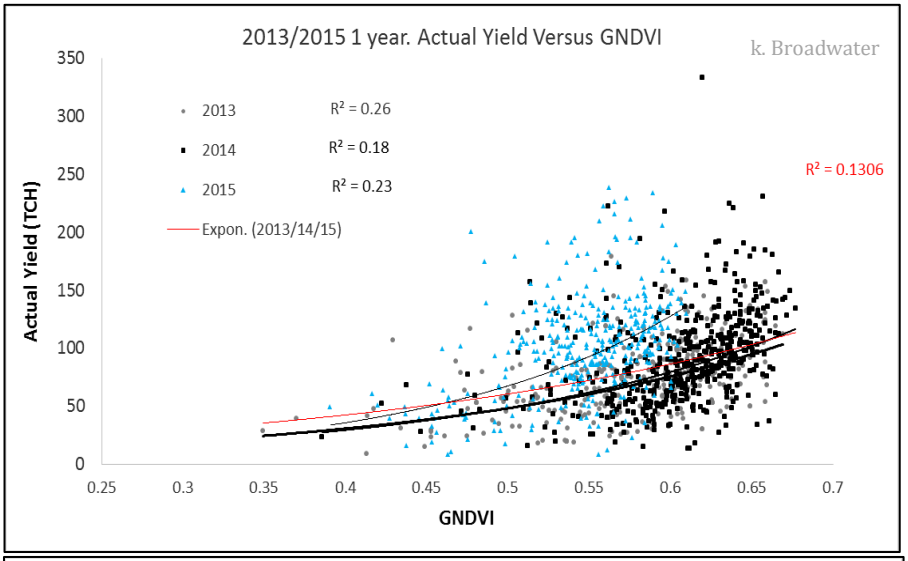
Year	Av GNDVI	image capture
2012	0.52125	29-Feb-12
2013	0.51263	13-Apr-13
2014	0.52934	19-Apr-14
2015	0.57237	18-Mar-15



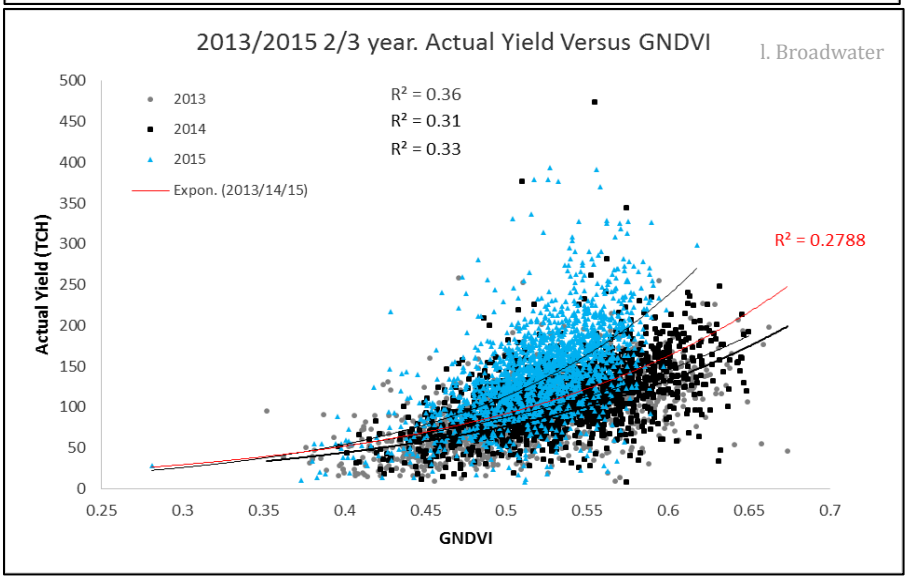
Year	Av GNDVI	image capture
2013	0.5619	25-Apr-13
2014	0.5786	10-May-14
2015	0.6138	18-Mar-15



Year	Av GNDVI	image capture
2013	0.4864	25-Apr-13
2014	0.4907	10-May-14
2015	0.5729	18-Mar-15



Year	Av GNDVI	image capture
2013	0.5764	25-Apr-13
2014	0.6064	24-Apr-14
2015	0.5472	8-Mar-15

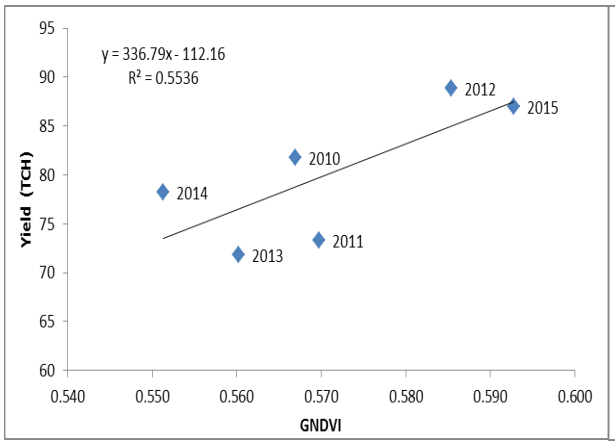


Year	Av GNDVI	image capture
2013	0.5186	25-Apr-13
2014	0.5349	24-Apr-14
2015	0.5205	8-Mar-15

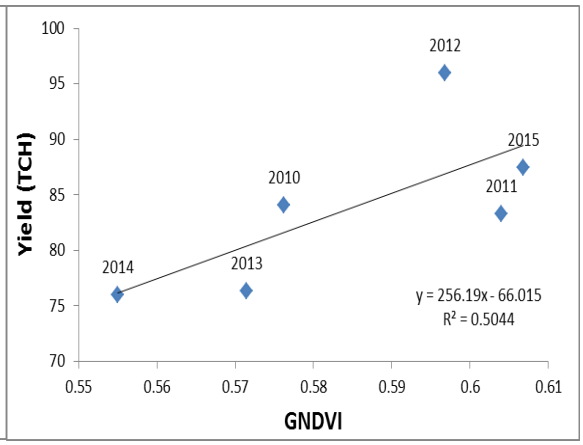
Figure 3. Relationship between actual harvested crop yield (TCH) and extracted average crop GNDVI for each individual crop grown across 9 growing regions. The timing of imagery capture for each season as well as the associated average GNDVI value for all crops within that region are also provided.

From Figure 3, the distribution of points around the exponential relationship of annual yield versus GNDVI, reported as the coefficient of determination or R^2 value, is also highly variable. For some regions such as

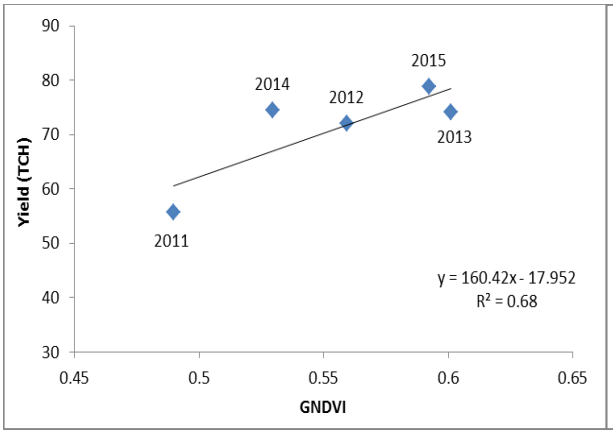
Bundaberg, the distribution of these points is relatively small considering the large sample number (n = 3,217 to 4,953), and relatively consistent across growing season (R^2 value 0.39 to 0.50). This result may indicate why prediction accuracies at the block level were found to be more accurate for the Bundaberg region compared to those achieved for the Burdekin region which produced poor coefficients of determination (R^2 value 0.07 – 0.30). In an attempt to improve the accuracies of yield prediction at the block level a number of statistical models were developed for the Bundaberg and Condong growing regions. These are discussed in the section *'Statistical analysis of Mill and imagery data for improved yield prediction accuracies at the block level'*. The coefficient of determination displayed in 'red text' within each region figure, is derived from the relationship of all seasonal points for each region. It was hypothesized that the corresponding algorithm developed from the exponential relationship would be less sensitive to seasonal variation than those algorithms developed from one season only. Interestingly, the linear relationship between average crop GNDVI for each season, versus average crop harvested yield produced a strong coefficient of determination, particularly for the Condong (1 year and 2 year cane), Herbert, Bundaberg and Isis regions (Figure 4). If these relationships persist over future seasons, then this simple linear algorithm may serve as a useful prediction algorithm in its own right.



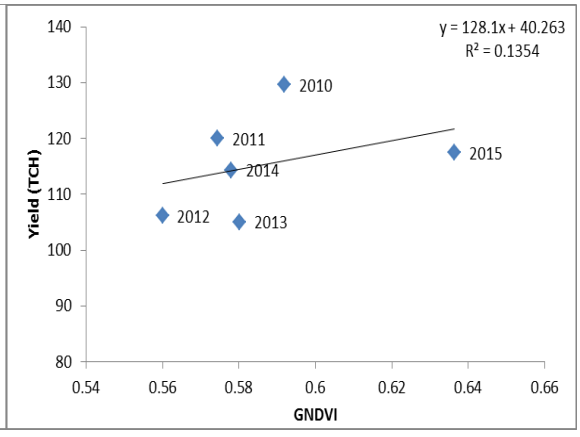
Bundaberg



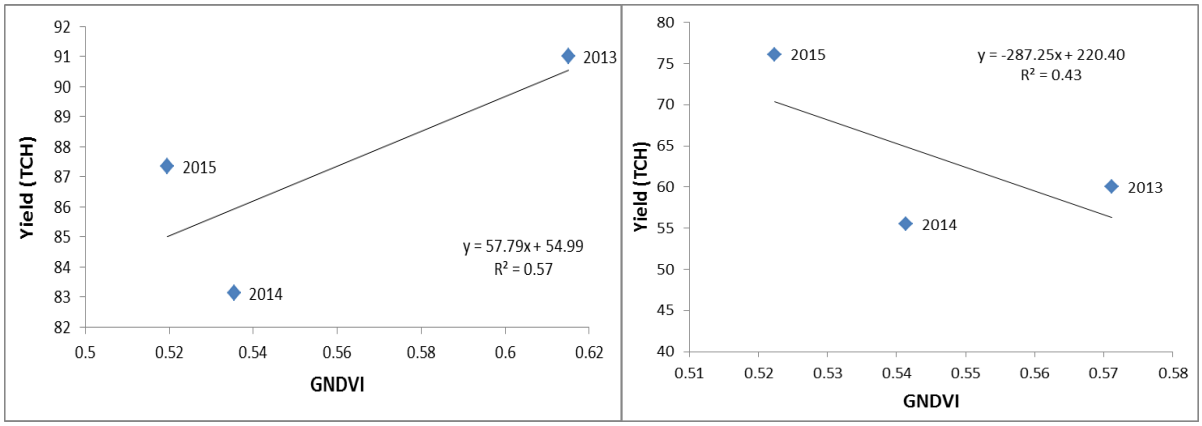
Isis



Herbert

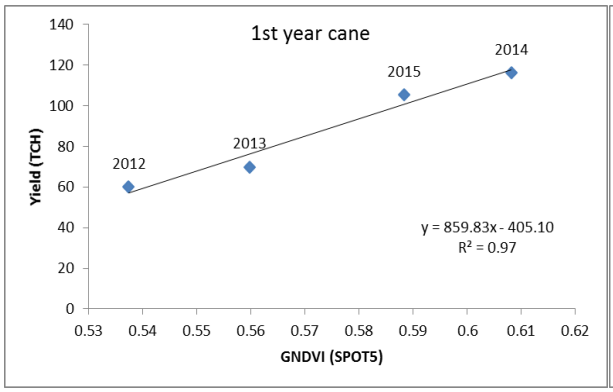


Burdekin

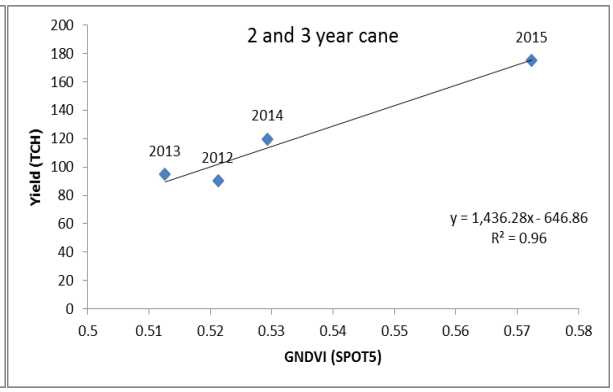


Gordonvale

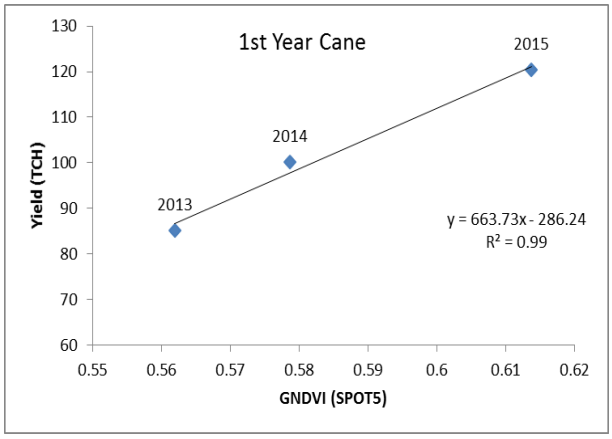
Maryborough



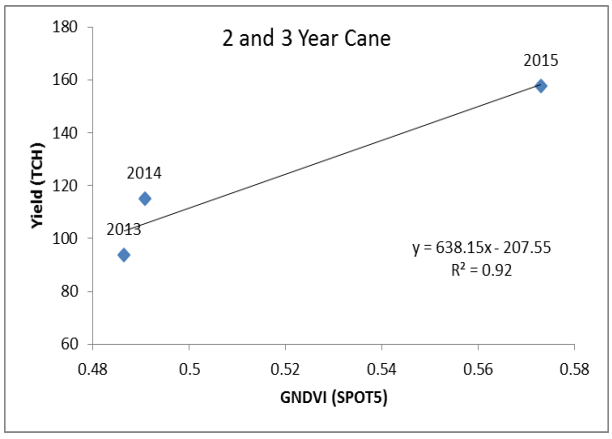
Condong 1 year cane



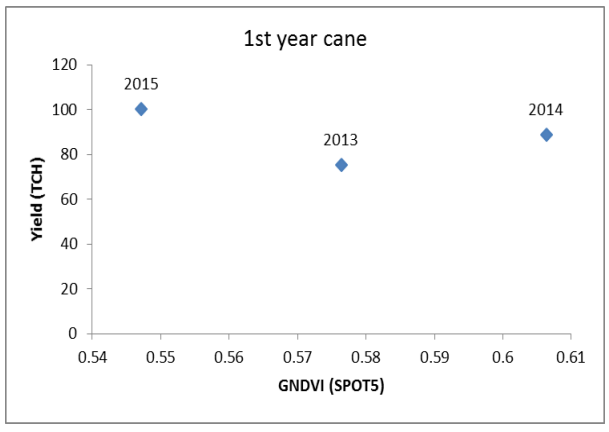
Condong 2 year cane



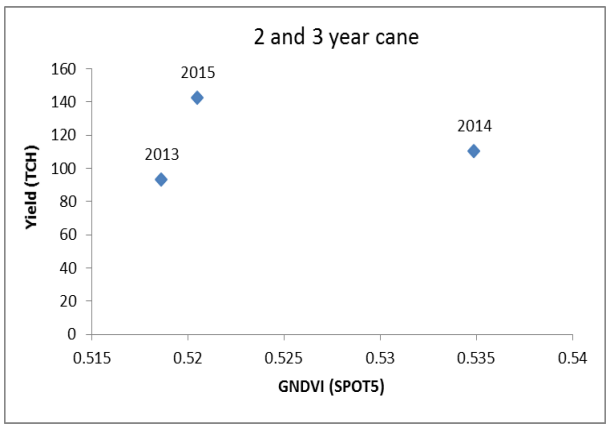
Harwood 1 year cane



Harwood 2 year cane



Broadwater 1 year cane



Broadwater 2 year cane

Figure 4. Distribution of average annual GNDVI vs yield for all crops, across each growing region and the respective number of season investigated.

As demonstrated, the development of a yield algorithm from the GNDVI versus yield relationship for only one or two seasons is unlikely to produce accurate predictions for an upcoming season, particularly if that season experiences significant climatic variation. In the following example (Table 4), yield forecasts for the 2016 season were made for the three NSW growing regions (1 year and 2 year cane) using variants of the GNDVI versus yield algorithm. These included:

1. *Exponential algorithm developed from the 4 years of collected data (red trend lines in Figure 3);*
2. *Exponential algorithm produced from the 2015 season only, as the 2016 was suggested be very similar;*
3. *Linear algorithms produced from average crop yield versus average GNDVI for each season (Figure 4). The Broadwater region produced a very poor relationship and as such an algorithm wasn't used*.

Table 4. Comparison of yield predictions made for the three NSW growing regions 2016 harvest, derived from 3 differing GNDVI versus yield algorithms. Only 'grower's estimate' is provided as validation of the predictions as the final harvest had not been completed at the time this report was submitted.

Mill	Grower estimate (t/ha)	4 years (t/ha)	2015 only (t/ha)	Av. Linear Alg. (t/ha)
Condong 1YO*	99	95.9	**107.8**	114.7
Condong 2YO*	152	122.3	**148.6**	149.8
Broadwater 1YO	93.6	**92**	144	
Broadwater 2YO	134.6	**130**	178	
Harwood 1YO	92	82.2	**95.3**	76
Harwood 2YO	153	97.2	104.2	115.8

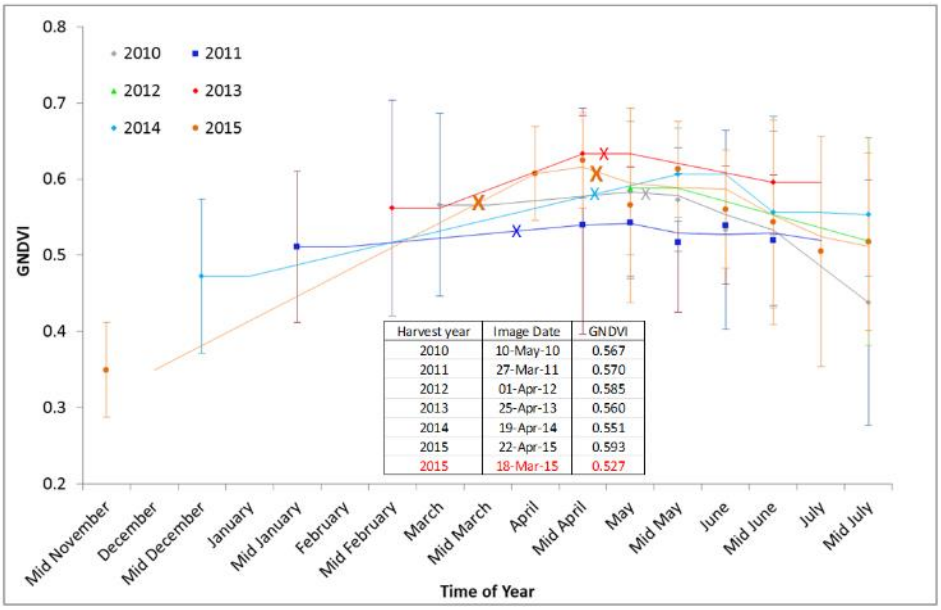
From Table 4, the regional forecast (1 and 2 year cane) for the Condong region derived from the 2015 season algorithm only were close to those estimated by growers. For the Broadwater region, predictions made from the 4 year algorithm were closer to grower estimate, whilst for Harwood all models fell short of the estimated average yield of 2 year cane. The poor result for the Harwood region is believed to be the result of the highly oblique look- angle of the SPOT 6 satellite capture (> 30 degrees), which would have reduced the amount of solar reflectance measured from the crop canopies (GNDVI) and therefore under estimated average crop yield.

This example again indicates the limitations of the single capture algorithm approach. If a season experiences environmental conditions that have not been accounted for in the development of the prediction algorithm; or if the image is either captured too early or too late in the growing season; or at an oblique look angle, then the yield predictions accuracies are likely to be poor. A reduced reliance on the single capture approach, coupled with a stronger understanding of the historic relationship between annual crop development and canopy spectral response, through a time series model approach, addresses these limitations. The following section examines the development of such a model for the Bundaberg growing region.

An improved 'time-series' approach for sugarcane yield forecasting:

In an attempt to understand the poor accuracy of prediction obtained for the Bundaberg growing region during the 2015 growing season (18th of March SPOT 5 capture), a scoping study was undertaken using time series Landsat imagery. By obtaining a measure of average canopy GNDVI multiple times between 2010 and 2015 a large variation in the annual development rate was identified, with peak growth for all seasons achieved in mid-April (Figure 5). The 18th of March image was acquired well before peak crop vigour was achieved and as a result the initial prediction was well under reported actual yield (27% under predicted). A

second image capture on the 22nd April, closer to peak growth, produced a more accurate prediction (5.7 % over predicted).



Harvest year	Image Date	GNDVI
2010	10-May-10	0.567
2011	27-Mar-11	0.570
2012	01-Apr-12	0.585
2013	25-Apr-13	0.560
2014	19-Apr-14	0.551
2015	22-Apr-15	0.593
2015	18-Mar-15	0.527

Figure 5: Annual growth trends of Bundaberg sugarcane crops identified from GNDVI values extracted from Landsat images between 2010 and 2015. The error bars indicate the GNDVI range for all crops. 'X' indicates timing of SPOT image capture for each season.

As well as identifying the historic growth trends of Bundaberg crops, this preliminary study produced an improved method for predicting sugarcane yield from satellite imagery. Using the modified algorithm produced from the linear relationship between average regional yield to average crop GNDVI value (extracted from Landsat images acquired between February and April 2010 -2015), a stronger regression coefficient was identified (R^2=0.91) than from that achieved from the single SPOT5 capture (R^2=0.52) (Figure 6).

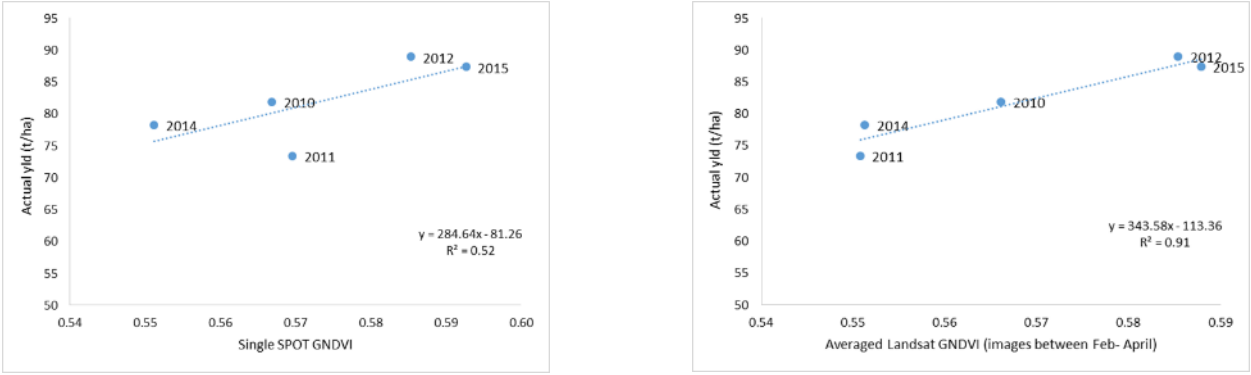


Figure 6: Correlation between GNDVI (SPOT5) with actual correlation between average GNDVI (Landsat) harvested yield (t/ha) between Feb- April, with actual harvested yield (t/ha).

Additionally, a further extraction of average crop GNDVI from 98 Landsat images (2001- 2015) was undertaken, where the annual growth cycle from vegetative, stabilisation to senescence is clearly visible (Figure 7).

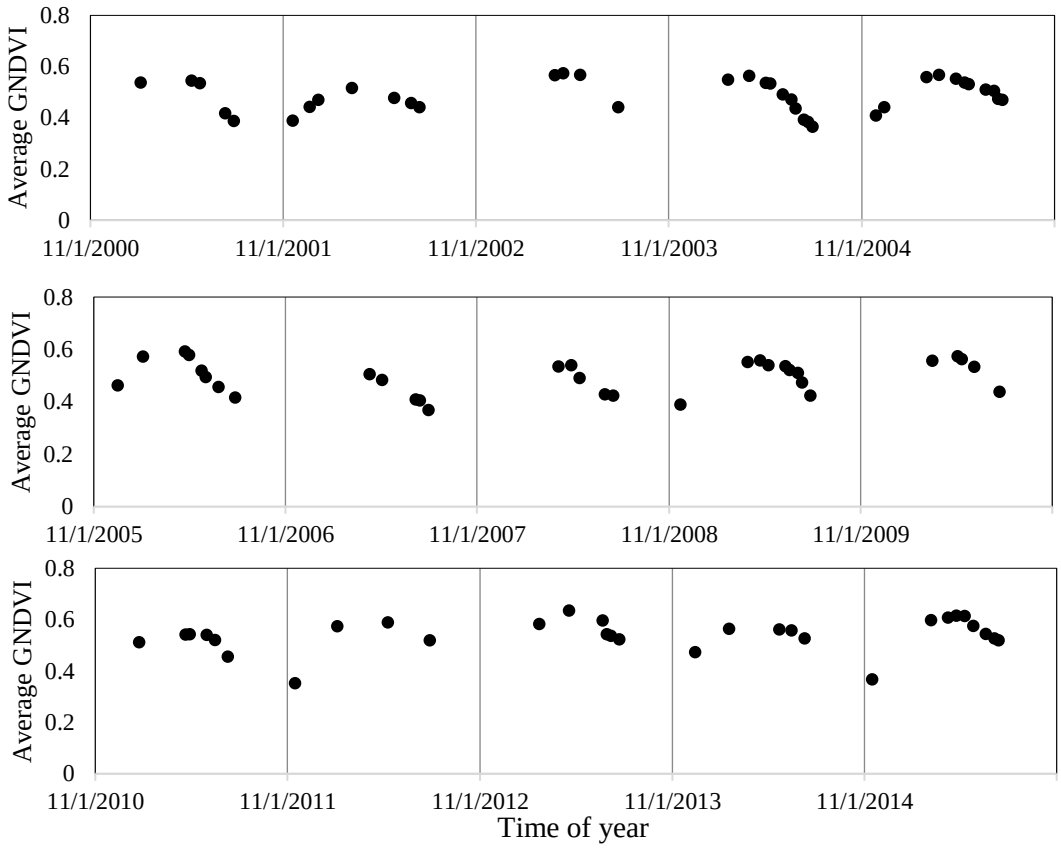


Figure 7. Time series of Landsat GNDVI data over 15 year's period (2001 to 2015) illustrating the sugarcane crop cycles in Bundaberg region. The images were captured only during the growing period from mid-November to July each year.

A quadratic model developed from this time series accurately explained annual crop growth between November and July (R^2=0.72) (Figure 8). This result is very encouraging considering the range of extreme weather events (drought and flood) experienced by the Bundaberg region within the period examined. The development of this model allows for predictions of average regional yield to be made earlier in the growing season i.e. from November, and in years where persistent cloud prevents a satellite image capture around the crucial April period, the maximum GNDVI value can still be calculated from earlier image acquisitions.

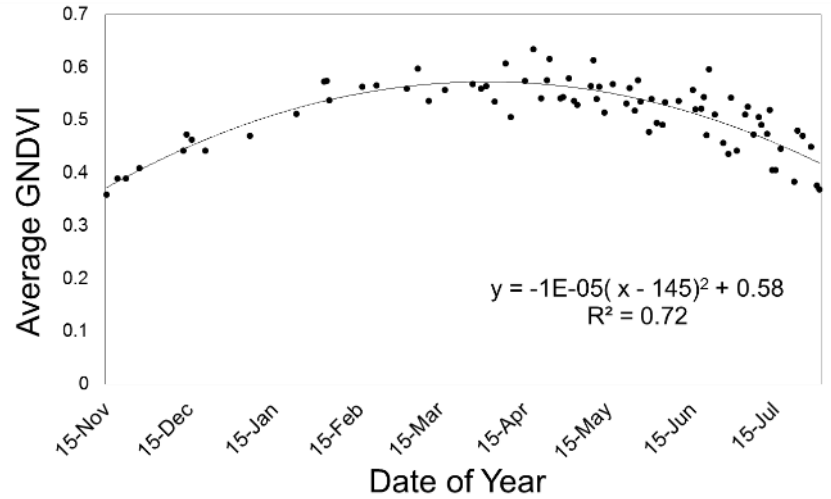


Figure 8: The average GNDVI data from 2001 to 2015 in the growing period of sugarcane (Mid-November to July) Regional crop GNDVI extracted from 98 images (2001 and 2015: mid Nov- July)

A scatter plot comparing the maximum GDNVI calculated from the quadratic model for each year (2001 to 2015) to the annual harvested yield (t/ha) produced a strong linear relation (R^2 = 0.69 and RMSE = 4.2 t/ha) (Figure 9). Note, 2013 is excluded due to extensive flooding that prevented the harvest of around 40% of final yield.

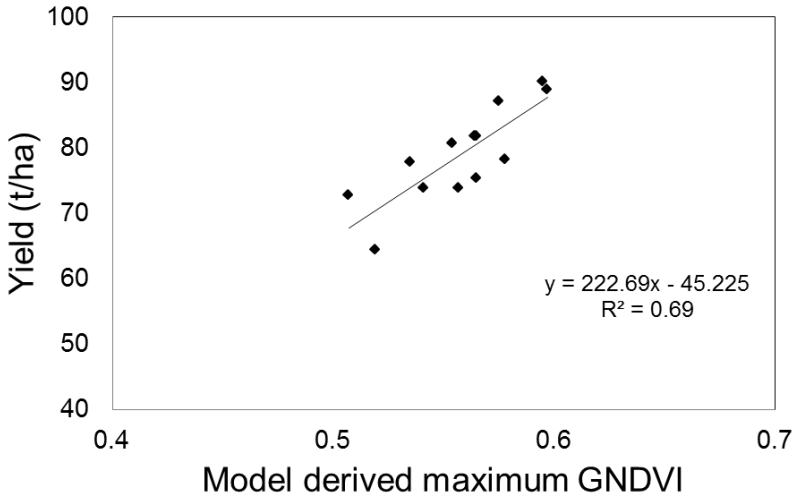


Figure 9: Scatter plot of model derived maximum GNDVI vs annual harvested yield (t/ha) from 2001 to 2015.

The time series analysis methodology offers significant improvement to remote sensing based yield forecasting than that derived from a single capture. However, similar models need to be developed for each growing region.

Statistical analysis of Mill and imagery data for Improved yield prediction accuracies at the block level.

Throughout Project DPI025 a number of statistical approaches using historic mill data were investigated to improve yield forecasting accuracies at the regional, farm and block level. These included a preliminary evaluation of the influence of variety and class on the regional yield/ GNDVI prediction algorithms (Bundaberg/ Isis, Burdekin and Herbert) undertaken by DAF Qld statistician Kerry Bell (Appendix 1); a continuation of this analysis undertaken by Greg Falzon (UNE) looking at the influence of 31 varieties on regional yield and the relationship between yield/ GNDVI for the Bundaberg region (data from 2010- 2013) (Appendix 2) as well as a further evaluation of generalised linear models (GLM) to determine the influence of additional covariates on the Bundaberg yield/ GNDVI model (Appendix 3); and an investigation of seasonal weather variability completed by Yvette Everingham (JCU), as a potential driver of the annual shift of the exponential linear relationship between cane yield/ GNDVI (Herbert, Bundaberg and Condong) (Appendix 4).

For the Bundaberg/Isis, Burdekin and Herbert growing regions, Kerry Bell identified that the inclusion of variety and ratoon (class) information as co-variates did improve the accuracies of the yield to imagery (GNDVI) forecasting models (yield/GNDVI). However, due to the lack of consistency of this interaction from one harvest year to the next, the benefit of this approach was questionable. It is worth noting that this analysis only included those seasons where satellite imagery was available i.e. 2010-2012.

The analysis of historic mill data undertaken by Greg Falzon, identified significant differences between the median yield achieved by varieties grown in the Bundaberg region (2010- 2013). The varieties Q240 (100.71 Tonnes/Ha), Q238 (94.41 Tonnes/Ha), Q183 (94.20 Tonnes/Ha), Q242 (94.04 Tonnes/Ha) and Q135 (89.05 Tonnes/Ha) were identified as the top five highest median yielding (Appendix 2). This information alone can

assist growers with future varietal selection. A positive relationship between GNDVI and Yield/Ha was identified for each variety. However, the intercept and slope did vary. This result was supported by generalised linear models (GLM) which identified the effect of variety on the yield/GNDVI model to be highly significant although surprisingly, the inclusion of variety information provided little improvement on the prediction errors associated with yield. An additional preliminary evaluation of univariate (GNDVI) and covariate predictor (GNDVI, MIDIR, NIR, Red and Green, Block Area, Variety) models for improved yield predictions at the block level, undertaken for the Bundaberg region, are provided in Appendix 3. The prediction accuracies for the Bundaberg and Condong growing regions, 2014 harvest season, are provided in the following section.

For Bundaberg:

This analysis compared the standard GNDVI algorithm prediction linear model (m1) verses a generalised linear model (m2). These covariates were selected using a stepwise regression routine on a much larger dataset containing variables such as cane class, growing year and irrigation method. This model was developed from Bundaberg data collected through SRA funded projects DPI021 and DPI025. Predictions of block yield were made for the 2014 harvest season using both m1 and m2 and then compared to actual yield provided from Bundaberg Sugar Ltd following harvest.

Table 5: Numerical summary of m1 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH).

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-108.2	-27.0	-0.7	28.3	259.7

Table 6: Numerical summary of m2 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH)

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-143.9	-26.2	1.4	30.3	272.2

From Tables 5 and 6, the univariate model which incorporates only GNDVI information is outperforming the more complicated model which contains additional covariate information such as which variety of cane is planted. For both models approximately 75% of all predictions were within ±30 TCH of the true yield. Model m1 was slightly less biased than model m2 with model m1 typically under-predicting by 0.7 TCH as compared to model 2 typically over-predicting by 1.4 TCH. Model m2 also had more extreme errors as compared to model m1, for instance the greatest under-prediction by model m1 was out by 108.2 TCH as compared to greatest under-prediction by model m2 which was out by 143.9 TCH. For these reasons model m1 appears to be a more reliable predictive model for the Bundaberg system. The take home message from this analysis is that the additional covariate information is not improving the prediction accuracies. Three potential explanations are: (i) a fundamental limit to prediction accuracy has been reached using remote sensing methodology, (ii) for the Bundaberg system these additional covariates are not important but in other growing areas they may be and (iii) the linear models proposed are too simplistic to capture the subtle non-linearity between predictor covariates and therefore more sophisticated models need to be developed. The latter two points are highlighted as key avenues of further investigation.

For Condong:

For the Condong region the classes of cane were stratified into two groups: one year old (Y1) cane and cane two years of age or greater (Y2+). A total of four different models were assessed for each of these stratified groups. The first model (m1) utilised only GNDVI information, whilst the second model (m2) incorporated further multi-spectral indices (GNDVI, SWIR, NIR, Red and Green sensor readings). The third model also included cane variety with these multi-spectral measurements (GNDVI, SWIR, NIR, Red, Green and Variety) and the fourth model (m4) added Class information (GNDVI, SWIR, NIR, Red, Green, Variety and Class). All models were created and predictions made prior to the 2014 harvest.

Condong Y1 Cane

Table 7: Numerical summary of model 1 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH).

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-140.40	-49.58	-34.34	-17.86	63.68

Table 8: Numerical summary of model 2 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH)

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-134.40	-33.68	-20.01	-3.29	69.77

Table 9: Numerical summary of model 3 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH).

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-135.20	-35.49	-21.17	-5.14	69.32

Table 10: Numerical summary of model 4 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH)

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-135.40	-38.50	-24.69	-8.50	69.49

From Tables 7- 10, model 2 (m2) produced the lowest median error, under-predicting yield by approximately 20 TCH. In fact all models 'under-predicted' yield in the 2014 season for Condong Y1 cane. It is important to observe that all GLM's substantially out-performed model one (m1) univariate model in terms of median prediction error. This is significant as model one (m1) is the standard single capture SPOT GNDVI model used for forecasting yield in the Condong region. Comparison of the errors reveals that the use of the GLM improved the median prediction error by 41.76% compared to a univariate model (m1 compared to m2). Incorporation of both variety and class information in models three (m3) and four (m4) respectively led to slightly poorer median prediction errors suggesting that the utility of incorporating this information for the Condong Y1 cane is questionable.

Extreme prediction errors (maximum and minimum) were identified in all models and of similar magnitude. Such errors are highly undesirable and should be avoided if possible. Identifying whether the causes of these

errors are a result of poor consignment, the modelling process or physical on-ground processes such as disease is a high priority task.

The following Tables 11- 14, indicate the accuracies of prediction from models developed for Condong Y2 cane. Model 5 (m5) was the univariate model using GNDVI, model 6 (m6) was the GLM using spectral indices (GNDVI, SWIR, NIR, Red and Green sensor readings), model 7 (m7) utilising variety (GNDVI, SWIR, NIR, Red, Green and Variety) and model 8 utilising class (GNDVI, SWIR, NIR, Red, Green, Variety and Class).

Condong Year 2 Cane

Table 11: Numerical summary of model 5 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH).

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-138.90	-38.09	-20.52	-2.09	50.99

Table 12: Numerical summary of model 6 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH)

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-134.90	-29.91	-12.34	-13.56	55.73

Table 13: Numerical summary of model 7 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH).

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-122.30	-28.00	-11.04	-12.80	50.84

Table 14: Numerical summary of model 8 prediction error for actual compared to predicted Tonnes Cane per Hectare (TCH).

Minimum	1st Quantile	Median (50%)	3rd Quantile	Maximum
-121.30	-29.78	-11.23	-13.25	55.41

From Table 11- 14, model 7 (which included cane variety) produced the best performing median prediction error (-11.04 TCH). This was a 46.20% improvement in prediction accuracy over the standard prediction model (model 5). Models 6 and 8 were also competitive models in terms of median prediction error. Inclusion of variety information along with the spectral information did result in a 10.53% improvement in median prediction accuracy for Condong Y2 cane indicating that cane variety information should be included in this case. Whilst the median prediction error for Condong Y2 cane was substantially better than the Condong Y1 cane, large magnitude prediction errors were still identified in all models evaluated.

The yield prediction models developed for the Condong Y1 2014 and Condong Y2 2014 datasets significantly outperformed the standard univariate model which was previously used for yield forecasting. Several areas of improvement remain including focusing greater effort on producing a statistical model which produces 0 TCH median prediction errors and if possible removing those observations that are likely to result in very large magnitude prediction errors, pre-forecast.

For the 2015 season the accuracies of prediction from the univariate and GLM's were again assessed. For the Bundaberg region the median prediction accuracy for all crops was slightly higher at 1.6 TCH, but 77.5 % of all the crop predictions were accurate to within 30 TCH using a univariate (GNDVI * yield) model. The low median error identified for the Bundaberg region was attributed to the relatively consistent annual (2010-2015) relationship between individual crop GNDVI and yield (Figure 3 a). For the Condong region, 1 year cane produced a median error of 27 TCH with 53.8% of crops to within 30 TCH (Yield*GNDVI*4 SPOT spectral bands); whilst for 2 year cane a median error of 70.6 TCH was identified with 25% of crops to within 30 TCH (Yield* GNDVI*4 SPOT spectral bands* Variety). For the Condong region the relationship between individual crop GNDVI and yield for 2012 and 2013 year was very different to that produced during 2014 and 2015. This variation between seasons and the limited years of available data were the likely drivers of poorer accuracies compared to the Bundaberg region.

The study undertaken by Yvette Everingham (JCU) differed to the previous approaches as it aimed to explain if weather was a potential driver of the seasonal shift observed between the yield/GNDVI relationship for each growing region (Figure 3). For this analysis Mill productivity information was obtained for the Herbert (2011-2014), Bundaberg (2010-2014) and Condong (2012-2014) growing regions, as well as annual corresponding crop GNDVI measures from SPOT5 captures and SILO weather data. Unfortunately, as seen from Appendix 4, the results were inconclusive, most likely the result of a limited time series of data i.e. maximum 5 seasons. As a positive result, the inclusion of an in season measure of crop performance was suggested to be highly beneficial to crop modelling. Also having an independent prediction strategy i.e. through the GNDVI approach would provide a useful validation for forecasts derived through crop modelling.

The results from the varying statistical studies all support the potential of historic mill data as a resource for better understanding past spatial and temporal patterns of crop production and as such help to predict similar trends into the future. The addition of complimentary information such as that from climate or remote sensing, offer improved knowledge of crop behaviour and therefore supports an improved ability to predict and therefore compensate for future anomalies such as those resulting from weather events, pest or disease.

Deriving Yield maps from imagery.

Using the algorithms developed from the relationship between GNDVI and yield, classified yield maps were derived for every block in each region (~60,000 individual crops) for each year of the project. Examples of these maps are provided in the *'Output and Outcomes'* section of this report.

Objective 2: To evaluate multispectral and hyper-spectral tools as a method for screening research and breeding trials

The use of replicated field trials is a well-established strategy for screening agricultural commodities for a range of purposes including pest and disease resistance, Nitrogen Use Efficiency (NUE), Water Use Efficiency (WUE) and for cultivar screening in breeding trials. However, the analysis of trials rarely takes into account the influences of inherent variability expressed by the trial location. This may include the influences of soil type, drainage or water holding capacity, topography or other nutritional issues. Additionally, the incidence of pest or disease infestations may also influence crop performance during the life of the trial. Remote sensing can improve the monitoring and analysis of replicated field trials through a number of ways, these include:

1. *Improved selection of trial location via an understanding of inherent spatial trends.*
2. *Improved monitoring of treatment response during the life of a trial.*
3. *Identification of representative sampling locations under a strip trial scenario.*
4. *The non- invasive screening of breeding trials in response to applied treatments.*

1. Improved selection of trial location via an understanding of inherent spatial trends.

The selection of a homogenous trial sites is of vital importance particularly when conducting breeding selections. The influence of a non- treatment related constraints can lead to the unfair elimination of a potentially beneficial genotype. The following example of a Nitrogen Use Efficiency (NUE) breeding trial grown in the Burdekin clearly demonstrates the influence of trial location on plot performance. As seen in Figure 10 b, a large high yielding region running diagonally through the crop, possibly a 'prior stream' (indicated by a purple polygon), results in high vigour crop response (expressed as GNDVI) (Figure 11) and potentially yield (Figure 12). Conversely, a small region to the south of this feature (yellow polygon) produces a region of low vigour (Figure 11) that translated into reduced yield and nitrogen concentration (Figure 12).

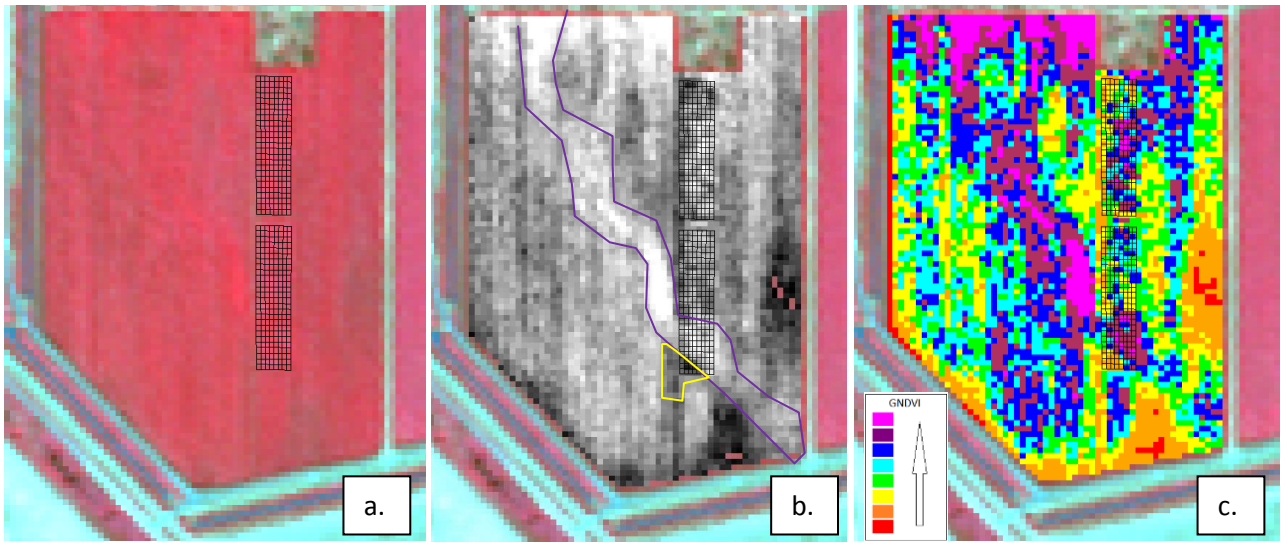


Figure 10. a. False colour SPOT5 image (captured 8 May 2014) of a Burdekin cane crop, with the extent of replicated trial overlaid. b. GNDVI and c. classified GNDVI layer.

Although the SPOT 5 image presented in Figure 10 was acquired after the completion of the trial i.e. not historical, it does clearly show the inherent spatial variability of the trial block which transcends through the life of the trial (Figure 11). Note, the specifics of this trial are detailed in Objective 3.

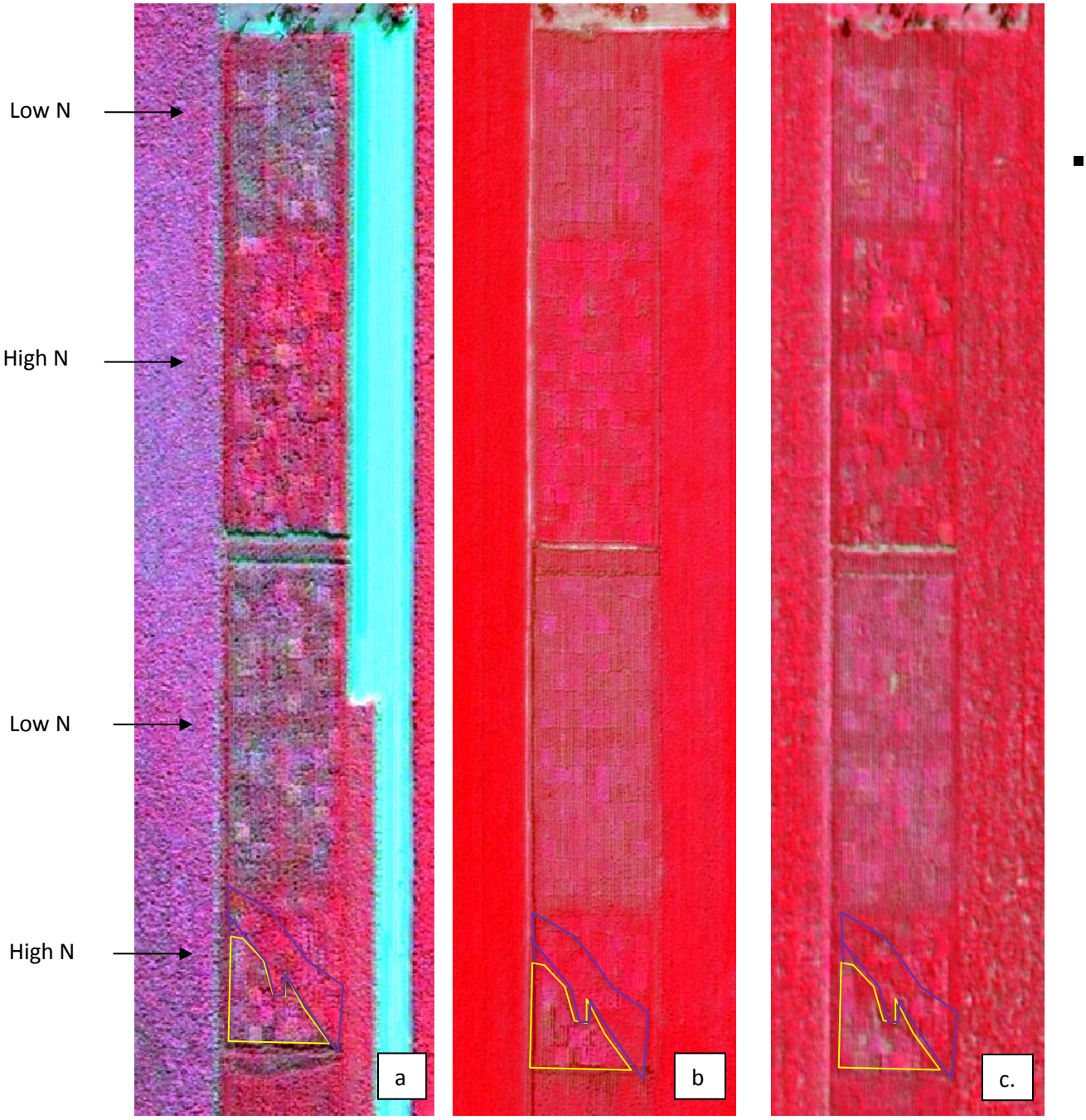


Figure 11. a. False colour IKONOS image (captured 28 May 2011). b. False colour GeoEYE satellite image (captured 14 Jan 2013). c. False colour Worldview-2 satellite image (captured 24 May 2013). False colour includes near infrared, with the brighter red colour indicating more vigorous crop growth. Yellow polygon indicates poor vigour area, purple polygon indicates high vigour area.

Although cultivars do respond differently to applied treatments, the concentration of low vigour plots within the south- western corner of the trial, producing both reduced yield (Figure 12 a) and nitrogen concentration (circled) (Figure 12 b), is clearly non-treatment related. As well as having a visual understanding of this influence, an actual metric of crop response, such as a GNDVI value, can serve as a useful co-variate for the statistical analysis of trial results, thus helping to reduce the wrongful elimination of cultivars.

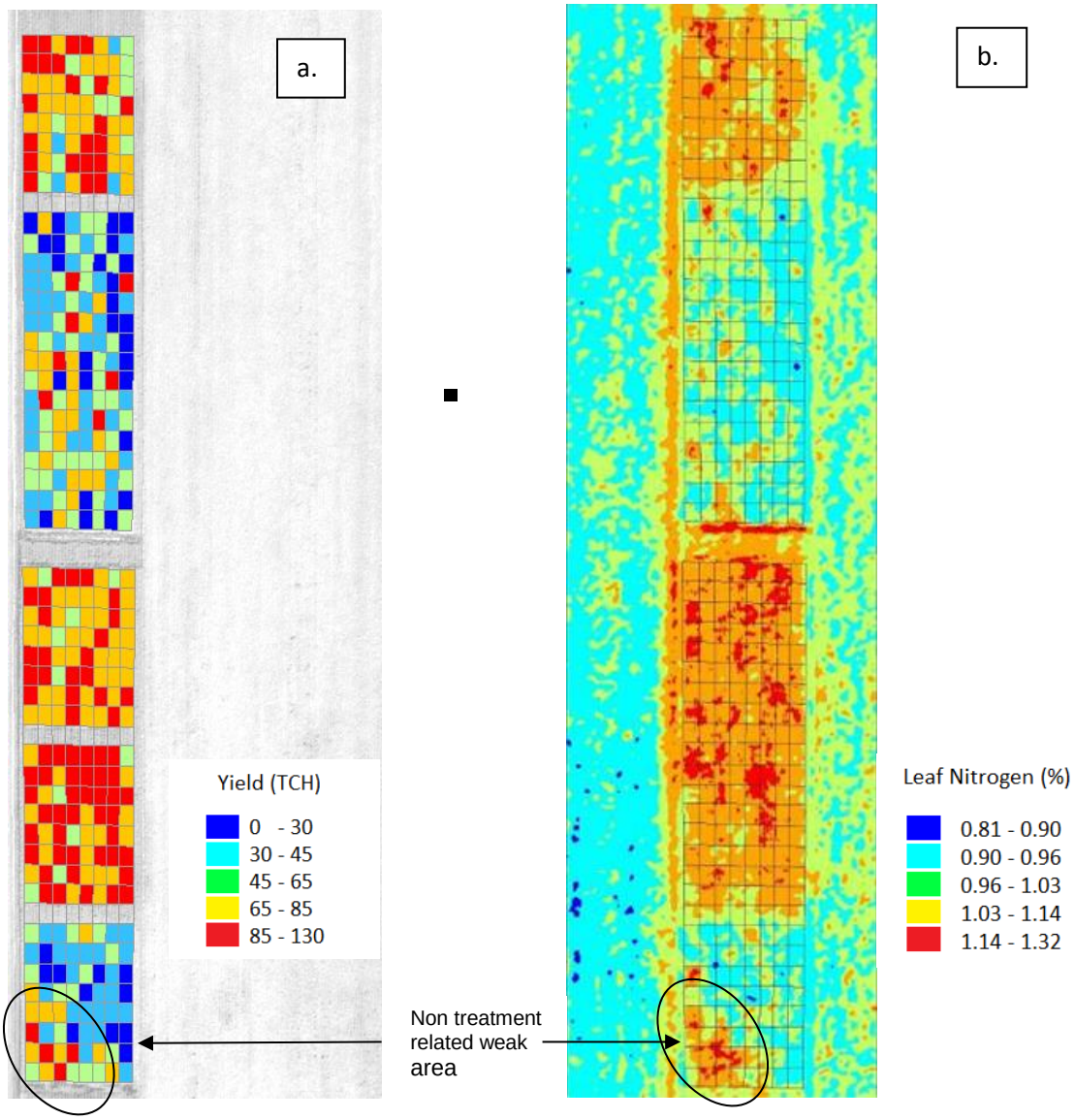


Figure 12. a. Distribution of measured yield (TCH) from plot samples taken July 2013. b. map of Nitrogen Concentration (%) derived from field sampling 23 February 2013 and Worldview-2 image (24 May 2013).

This example strongly demonstrates the benefit of remote sensing for the placement of replicated trials. Ideally historic imagery should be sourced during the planning phase of every trial to ensure it is placed in a homogenous area.

2. *Improved monitoring of treatment response during the life of a trial.*

As well as identifying the optimal location for trial placement, obtaining a measure of crop performance during the life of trial is also highly advantageous. Figure 13, provides a comparison of a replicated field trial grown at the Mackay Sugar Research Australia (SRA) station, with imagery acquired on the 19 April 2013 (Figure 13 a) and the 3 June 2015 (Figure 13 b). Although these dates are sometime apart, the example clearly identifies the later ratoon of some plots are significantly influenced by a constraint (yellow polygon). The advent of drones or UAV's may provide a useful platform for the multi-temporal image acquisition of replicated trials during a growing season due to size of the area of interest and the cost of high resolution satellite imagery for multiple repeat captures. However, significant research is required to ensure the appropriate protocols for acquiring imagery (choice of UAV and sensor, flying direction, flying height, time of day etc) as well as image pre-processing are developed.

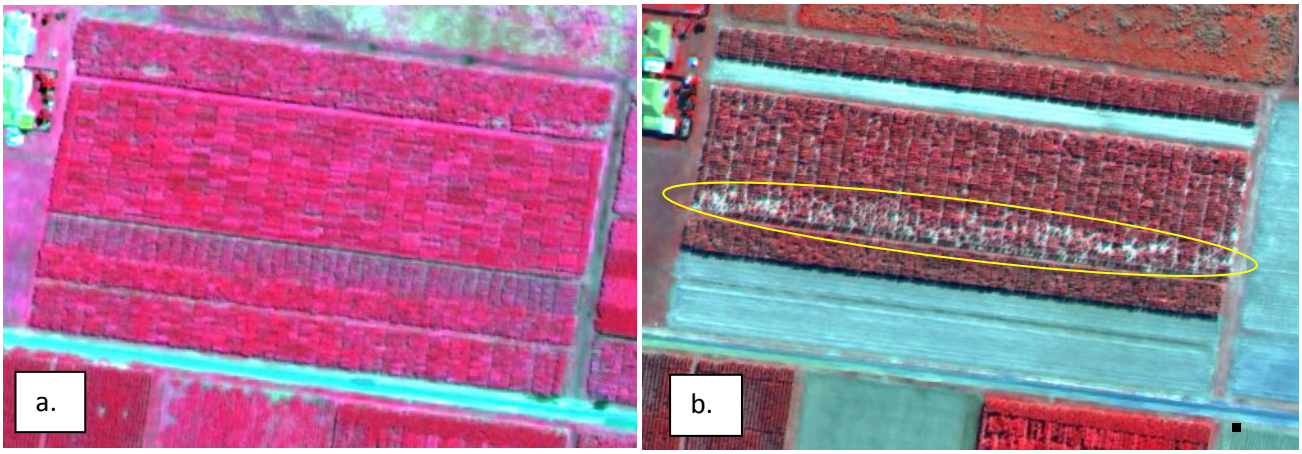


Figure 13. a. False colour image of a replicated sugar cane trial (Worldview-2 image acquired trial 19 April 2013). b. Repeat false colour image acquisition of the same trial (Worldview-3 image acquired trial 3 June 2015), with a large area of reduced plant growth indicated by a yellow polygon.

3. Identification of representative sampling locations under a strip trial scenario.

As demonstrated, an understanding of the spatial variability of crop vigour across a replicated trial can greatly improve the interpretation and analysis of results. Additionally, a similar understanding for 'strip plot' design trials can assist in determining appropriate sampling locations that accurately represent the applied treatment. Figure 14, provides two examples of 'strip plot' trials, where a large degree of variation in terms of crop vigour can be observed. Again specifics of these trials are provided in Objective 3. For Figure 14 a and b, there is little indication in terms of crop response to the varying treatments of applied N. However, a large region of poor performing plants can be seen (Figure 14 a: yellow polygon). At least three sample points have been collected within this poor performing region, the result of a compaction layer at a depth of between 30-50 cm, and as such likely produced results not representative of the treatment. For Figure 14 c and d, the Nitrogen treatments are clearly identifiable, but in some locations the treatment response is over shadowed by other constraints (Figure 14 c: yellow polygon).

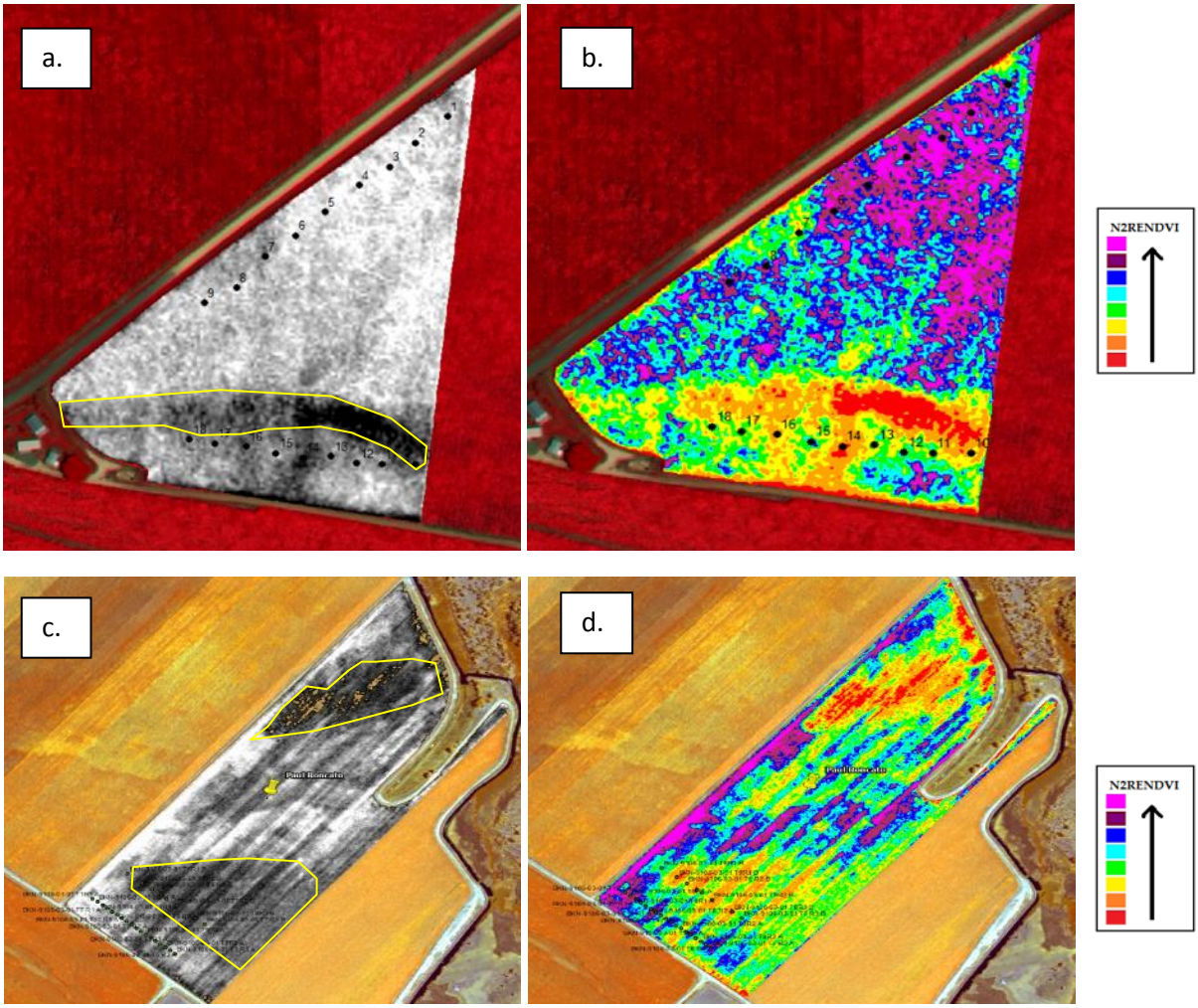


Figure 14. a. GNDVI image and b. classified N2RENDVI image of EHP (grower 14) Nitrogen 'strip plot'. c. GNDVI image and b. classified N2RENDVI image of site Farmacist managed 'Roncato' Nitrogen 'strip plot' trial. Both trials grown within the Burdekin with field sampling locations overlaid.

4. The non- invasive screening of breeding trials in response to applied treatments.

It has been suggested that remote sensing, or more specifically the extraction of canopy reflectance data from replicated trial plots, can be used as a non- invasive screening tool for varietal discrimination and for evaluating varietal response to an imposed treatment. In the following example, a spectral comparison of the replicated trial Mackay UQ/SRA is presented (Figure 15 a). This trial consisted of 384 plots of which only 86 consisting of 15 genotypes with 6 replicates of Q186, SP80-1816, N14, Q138, QN01-1075, QN01-1735, KQ228, QBYC04-10559, SP79-2313, R570, QC91-580, Q183, Q208, 5 replicates of Q240 and 3 replicates of Q218, were intensively sampled. The trial was planted on the 21st November 2014, with 40 kg/ha of Nitrogen applied to the low nitrogen treatments and 160 kg/ha to the high. The spectral data used for this analysis was acquired from a number of sensors, these included:

- Worldview2 satellite image (acquired 19 April 2013 and 5 January 2014) were selected for the imaging of the replicated field sites as they offer very high spatial resolution (0.5 m panchromatic and 2 m for the 8 bands) and 8 spectral bands. The spatial resolutions of the 8 spectral bands are as follows:

Band 1 (Coastal):	400 - 450 nm	Band 5 (Red):	630 - 690 nm
Band 2 (Blue):	450 - 510 nm	Band 6 (RedEdge):	705 - 745 nm
Band 3 (Green):	510 - 580 nm	Band 7 (Near- IR1):	770 - 895 nm

Band 4 (Yellow): 585 - 625 nm Band 8 (Near- IR2): 860 -1040 nm

All imagery used for this study was orthorectified and top of atmospheric (TOA) corrected (website. DigitalGlobe).

- Field based spectrometer measures were undertaken on the 3-4th December 2013 using an Analytical Spectral Devices™ (Fieldspec, 1997) (ASD). The ASD provides a spectral resolution of 350-2500 nm with 1 nm increments. The sampling regime consisted of three reflectance measurements taken 50 cm above the leaf and three measurements taken 20 cm above. The 25° field of view (fov) spectrometer foreoptics orientated at a downward look angle of 60 degrees elevation in the plane of the sun. This produced a target area of 22 * 22 cm and 8.8 * 8.8 cm. The initial radiance measurements from each plant were converted to reflectance values through the use of coincident measurement from a Spectralon panel, and re-acquired a number of times throughout the sampling session.

- Hyperspectral Airborne imagery was captured over the Mackay UQ/SRA 2014 trial on the 12th December 2013 by Airborne Research Australia (ARA) and included two sensors: SPECIM EAGLE (400-100nm) and SPECIM HAWK. The Eagle instrument has 252 bands ranging from 400.7 nm – 999.2 nm with a swath width of 965 pixels. The Hawk instrument has 241 bands ranging from 993.1 nm – 2497.4 nm with a swath width of 296 pixels. Imagery swaths from each sensor were mosaicked and resampled to 1 m spatial resolution. Radiance measures were converted to reflectance via the use of the ASD spectrometer and ground control targets.

As with commonly reported practice, the water absorption bands (1350- 1420, 1760- 1965 and 2450- 2500 nm) from both hyperspectral data sets were removed as they are considered major sources of spectral noise (Carter, 1994, Zhao et al. 2005).

From Figure 15b, the Worldview2 image (19 April 2013) of the Mackay trial clearly identified the six sub- block treatments of high (brighter red colour in the false colour image) and low N applications as well as some variation in the spectral response of the different cultivars. Although the in- season visual assessment of varietal response is beneficial, a qualitative measure that can be employed for the 'non- destructive' phenotyping of breeding trials would be of more use, particularly if the measure could be employed as a field –based screening tool.

Figure 16 identifies the spectral reflectance signatures from each of the 15 varieties (averaged from the 3 replicates of high applied N and then the 3 replicates of low) measured by the field spectrometer (Figure 16 a and b), ARA airborne hyperspectral sensor (Figure 16 c and d) and from a Worldview2 satellite image (Figure 16 e and f).

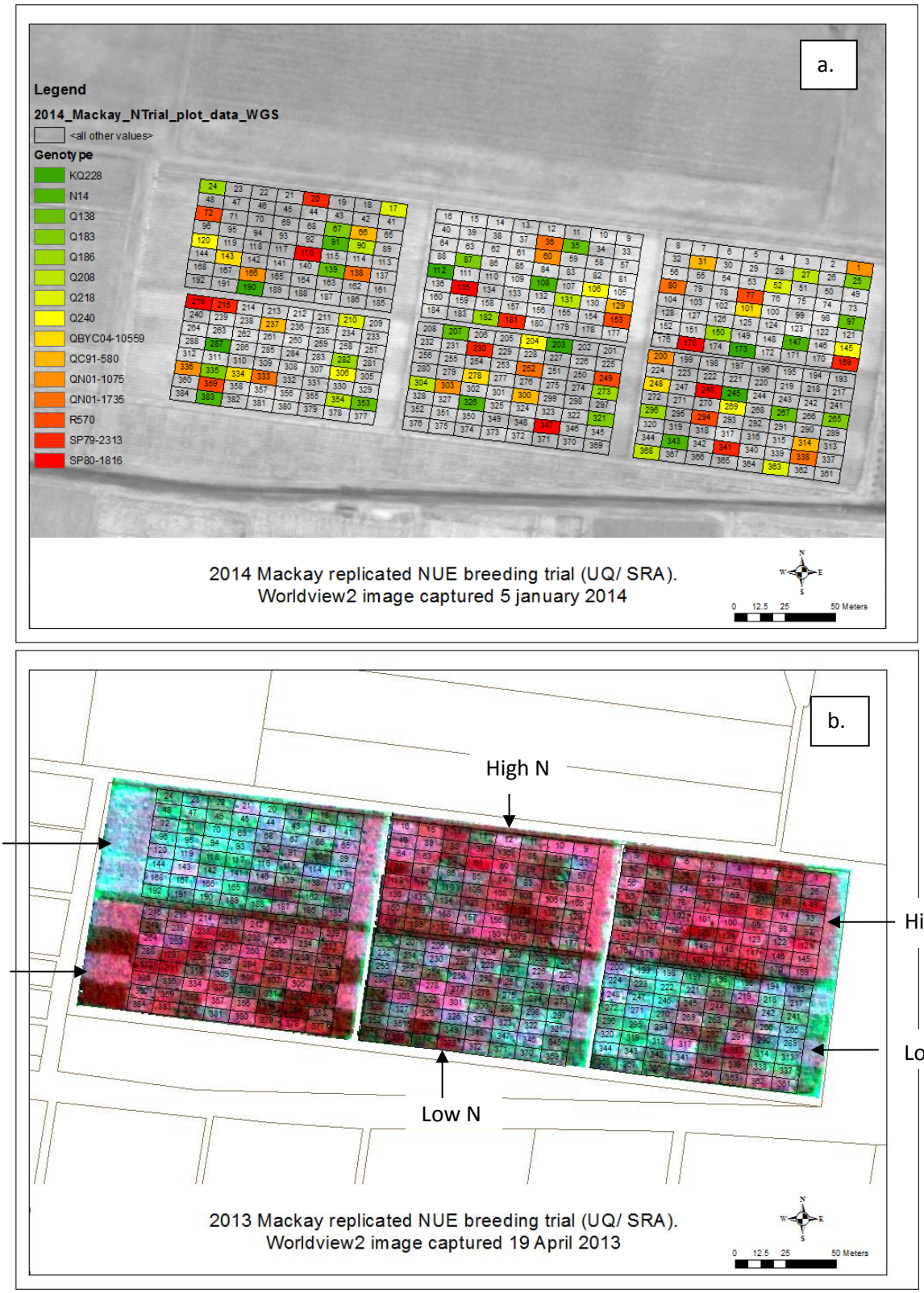
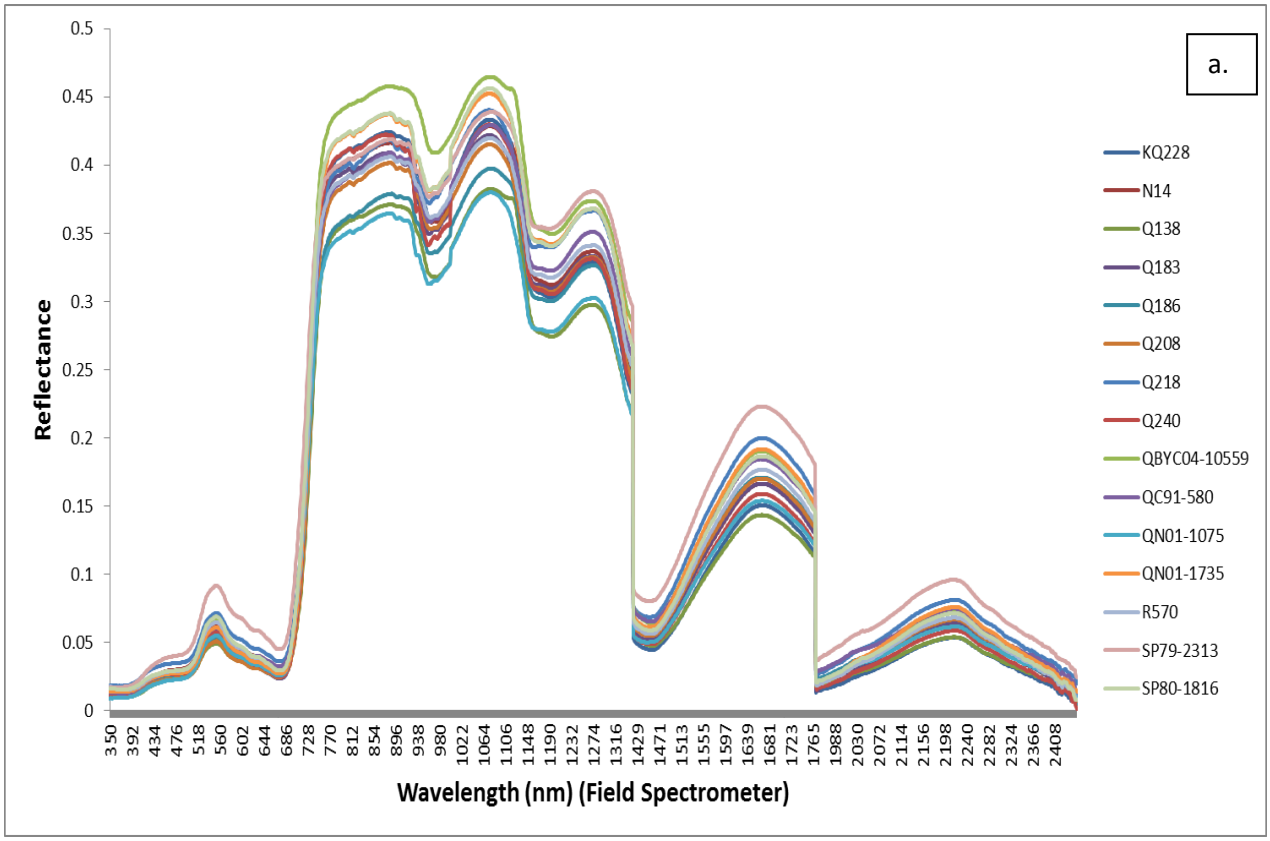
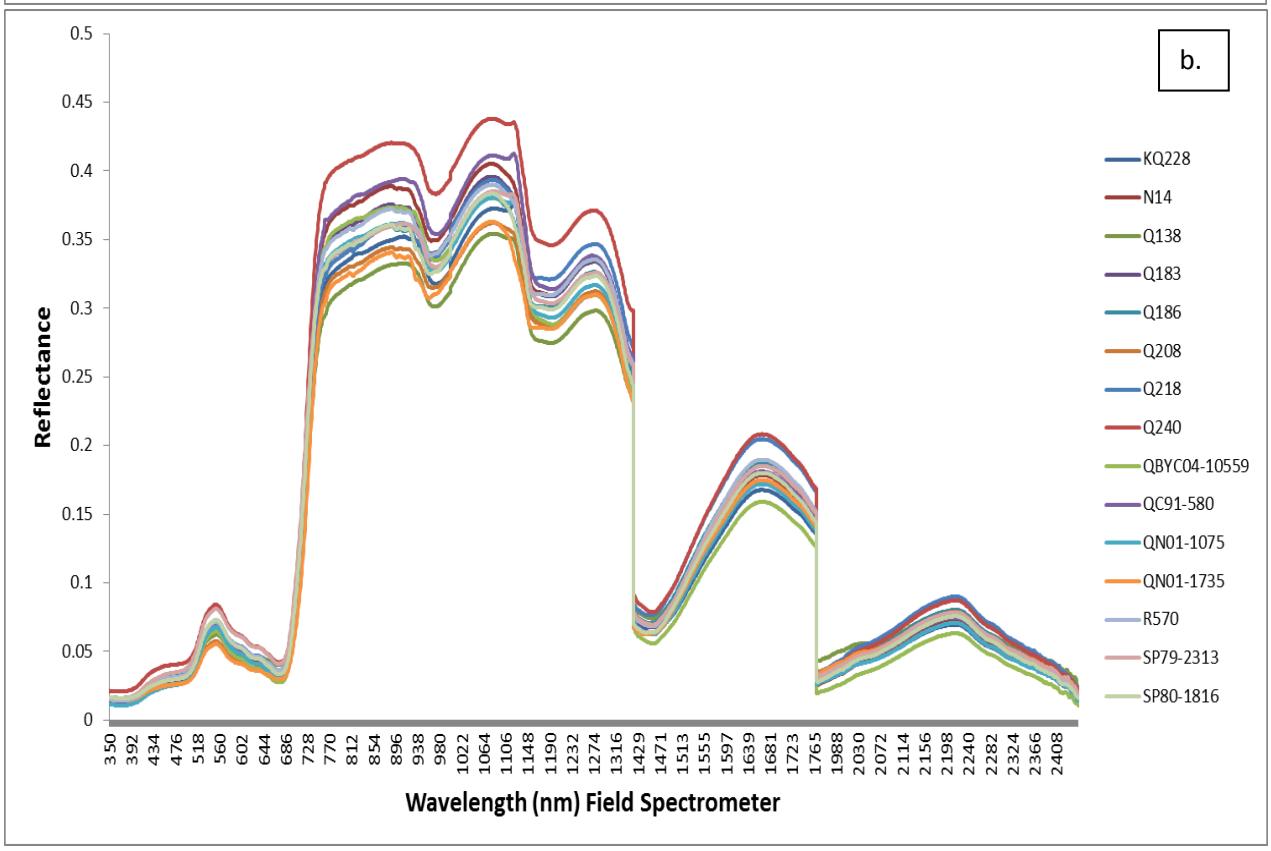
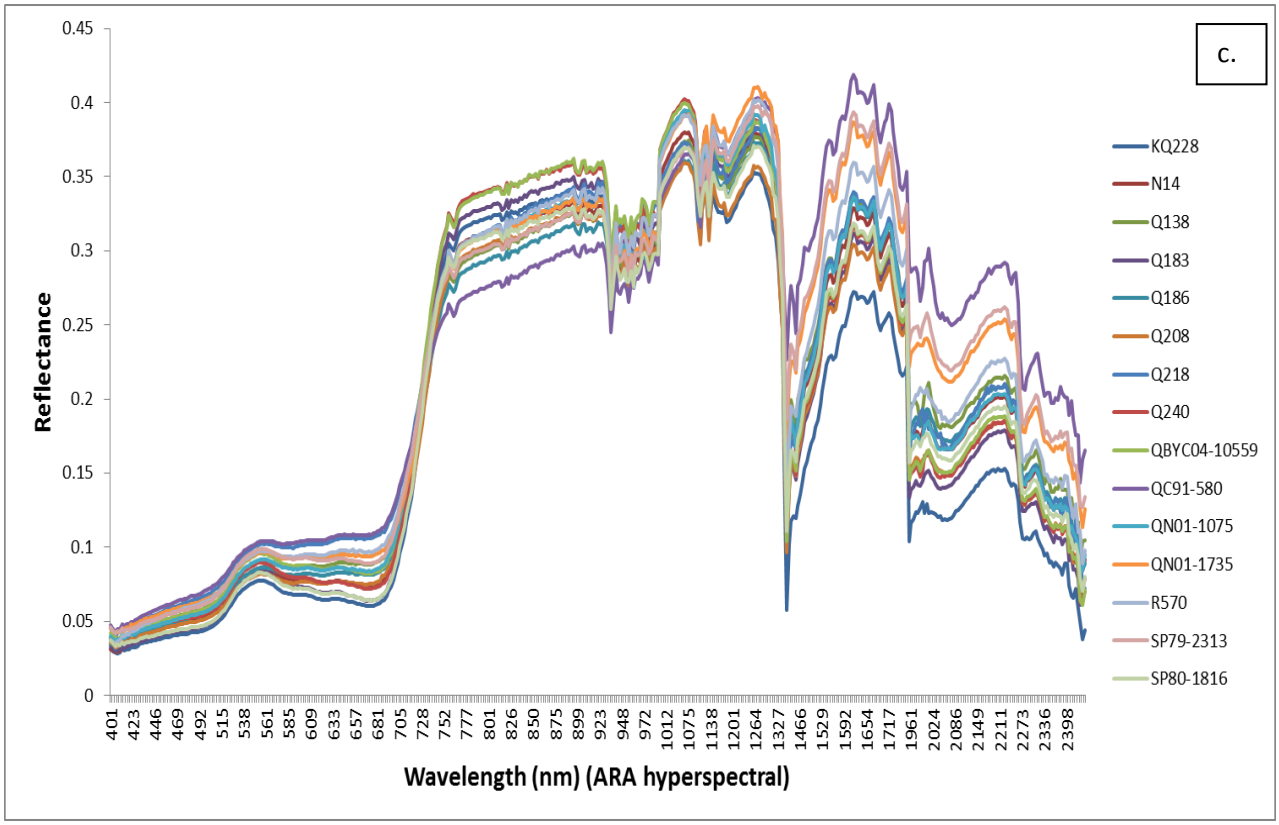
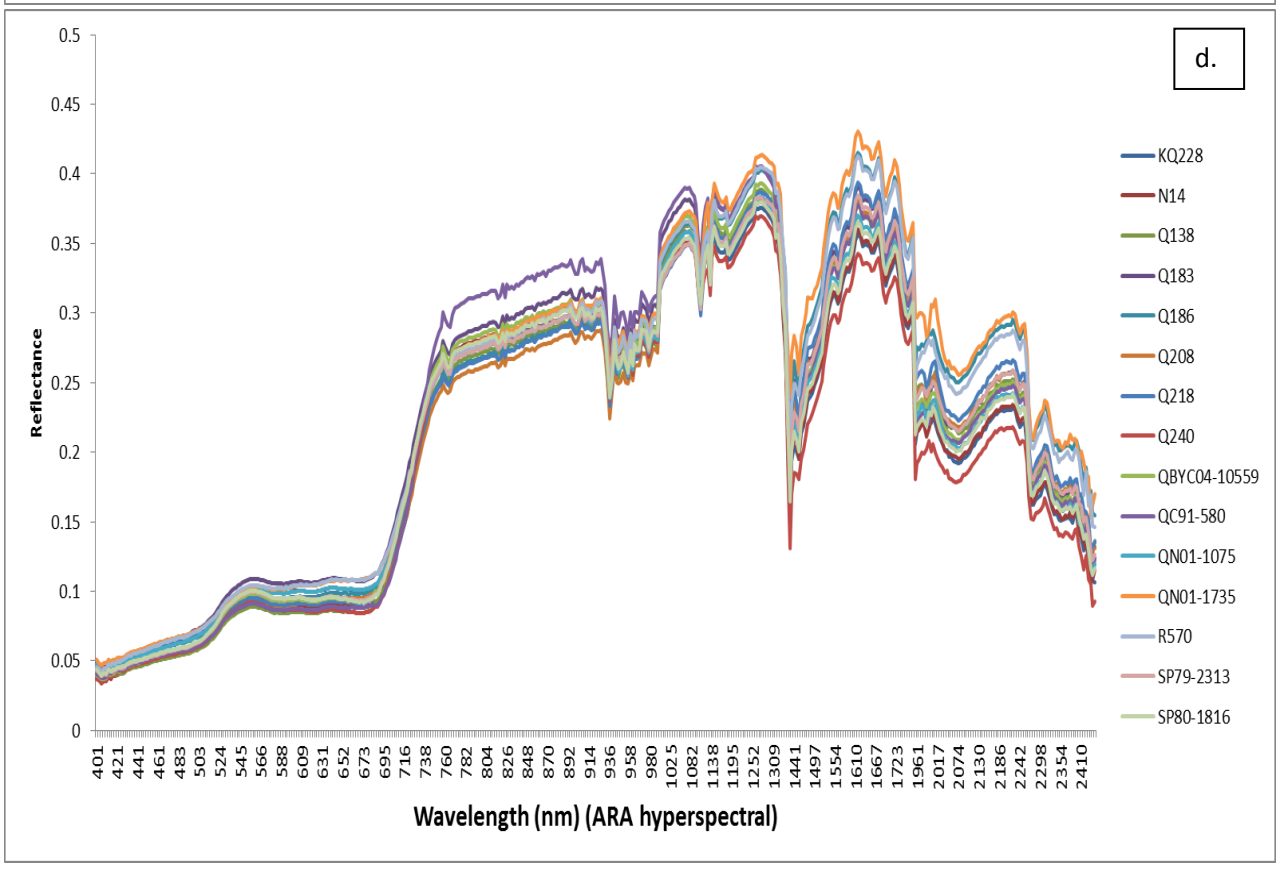


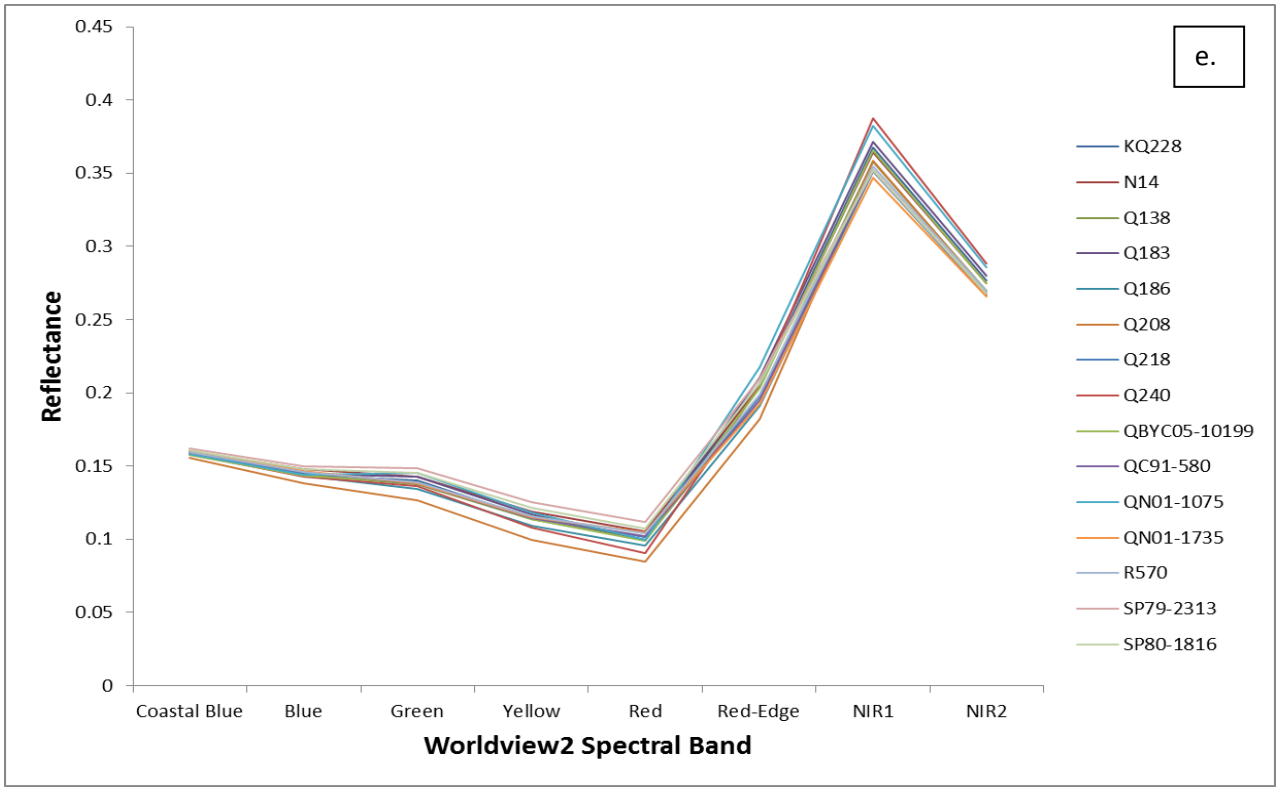
Figure 15. a. Design of the UQ/ SRA Mackay NUE trial with those cultivars/ plots intensively harvest for yield and foliar N concentration colour coded. b. False colour Worldview2 image of the trial (19 April 2013).


a.
Reflectance
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
350 392 434 476 518 560 602 644 686 728 770 812 854 896 938 980 1022 1064 1106 1148 1190 1232 1274 1316 1429 1471 1513 1555 1597 1639 1681 1723 1765 1988 2030 2072 2114 2156 2198 2240 2282 2324 2366 2408
Wavelength (nm) (Field Spectrometer)
KQ228
N14
Q138
Q183
Q186
Q208
Q218
Q240
QBYC04-10559
QC91-580
QN01-1075
QN01-1735
R570
SP79-2313
SP80-1816


b.
Reflectance
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
350 392 434 476 518 560 602 644 686 728 770 812 854 896 938 980 1022 1064 1106 1148 1190 1232 1274 1316 1429 1471 1513 1555 1597 1639 1681 1723 1765 1988 2030 2072 2114 2156 2198 2240 2282 2324 2366 2408
Wavelength (nm) Field Spectrometer
KQ228
N14
Q138
Q183
Q186
Q208
Q218
Q240
QBYC04-10559
QC91-580
QN01-1075
QN01-1735
R570
SP79-2313
SP80-1816


c.
Reflectance
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
401 423 446 469 492 515 538 561 585 609 633 657 681 705 728 752 777 801 826 850 875 899 923 948 972 1012 1075 1138 1201 1264 1327 1466 1529 1592 1654 1717 1961 2024 2086 2149 2211 2273 2336 2398
Wavelength (nm) (ARA hyperspectral)
KQ228
N14
Q138
Q183
Q186
Q208
Q218
Q240
QBYC04-10559
QC91-580
QN01-1075
QN01-1735
R570
SP79-2313
SP80-1816


d.
Reflectance
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
401 421 441 461 483 503 524 545 566 588 609 631 652 673 695 716 738 760 782 804 826 848 870 892 914 936 958 980 1025 1082 1138 1195 1252 1309 1441 1497 1554 1610 1667 1723 1961 2017 2074 2130 2186 2242 2298 2354 2410
Wavelength (nm) (ARA hyperspectral)
KQ228
N14
Q138
Q183
Q186
Q208
Q218
Q240
QBYC04-10559
QC91-580
QN01-1075
QN01-1735
R570
SP79-2313
SP80-1816



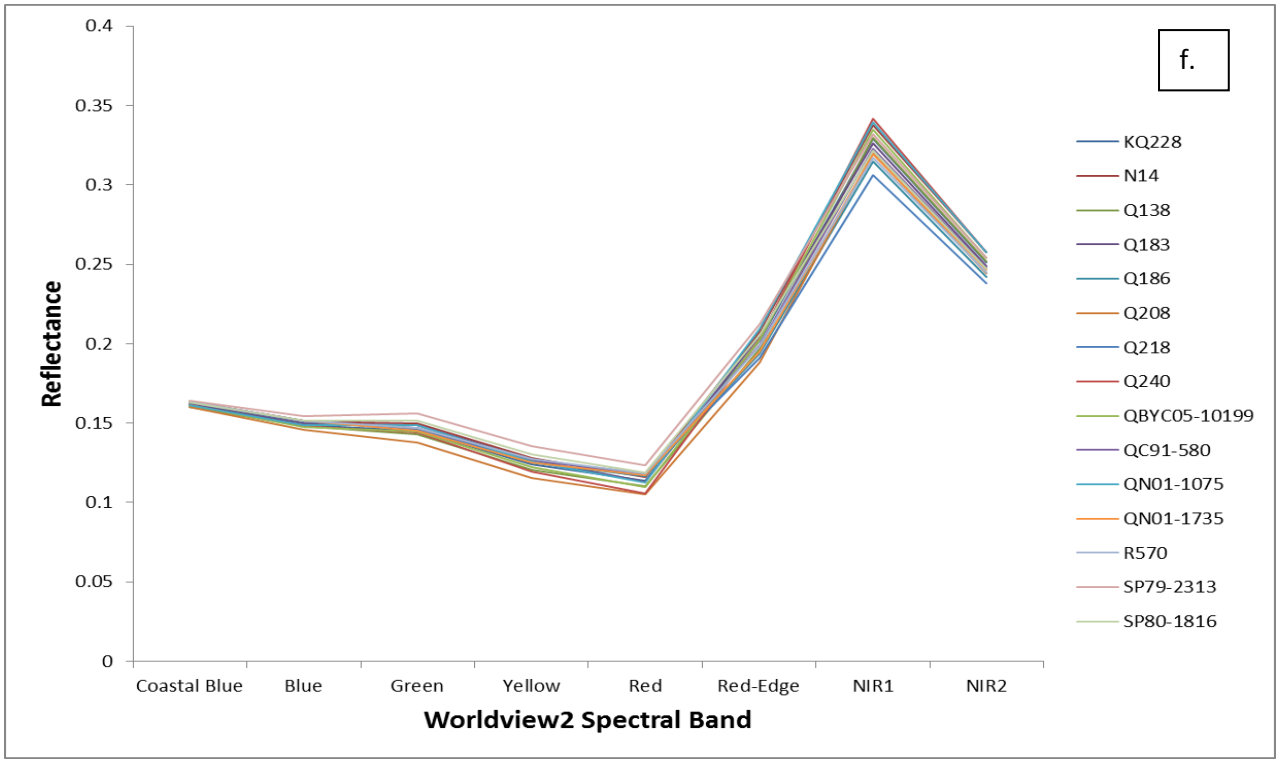


Figure 16. The separation in spectral reflectance from each of the 15 varieties, averaged from the 3 replicates of high applied N (a, c and e) and then the 3 replicates of low (b, d and g), measured by the field spectrometer (Figure 16 a and b), ARA airborne hyperspectral sensor (Figure 16 c and d) and from a Worldview2 satellite image (Figure 16 e and f).

The differentiation of spectral reflectance signatures from each of the varieties can be clearly seen in Figure 16, irrespective of the sensor used. Differences in the order in which the varieties have spectrally separated are evident both in response to the two N treatments and the different sensors used. In terms of the varying N treatments, this may be attributed to individual varietal response, i.e. increased NUE. However, in terms of the

differing sensors, this is may be the result of the varying spectral resolutions as well as spatial resolutions or the size of the 'foot print' measured.

Although the full spectral response (350 -2450 nm) is provided it is likely that there are a number of specific wavelengths that are more sensitive to varietal discrimination. Although additional statistical analysis was conducted over these data sets to identify such wavelengths, the limited number of replicates i.e. 3 High N and 3 Low N, used in the trial, produced poor results.

In terms of usefulness, the non- invasive discrimination of parental lineage and inherited traits of each progeny would be highly beneficial for breeding selections as well as for the monitoring and enforcement of plant breeding rights. Following these results, it is evident that more research is required to evaluate the non-invasive phenotyping of sugar cane breeding trials.

The ability to spectrally measure the foliar nitrogen concentration of sugar cane and the influence variety has on that relationship is discussed in the following Objective.

Objective 3: Evaluate multispectral and hyper-spectral tools as a method for measuring canopy nitrogen status

A number of remote sensing (multispectral and hyperspectral) platforms and sensors were evaluated to determine their accuracies in measuring the nitrogen concentration of sugarcane leaves. To determine the robustness of this relationship a number of replicated trial sites, both plot and strip configurations, including a range of genotypes, growing locations and growing seasons were spectrally measured and sampled for nitrogen concentration (Table 15). These trials were conducted by a range of research groups (SRA, UQ and Farmacist), whom actively collaborated with project DPI025.

Table 15: Nitrogen trial sites spectrally assessed over the course of the project

location	Season	sample date	image date	platform
Tully				
Euramo (SRA)	2014/15	05-Dec-14	28-Jan-15	Worldview2
	2015/16	05-Dec-15	5-Jan-16	Worldview2
Tully Station (SRA)	2014/15	05-Dec-14	28-Jan-15	Worldview2
	2015/16	05-Dec-15	5-Jan-16	Worldview2
Mackay				
Mackay (UQ/SRA)	2012/13	01-Aug-12	26-Dec-12	GeoEYE
		10-Dec-12	19-Apr-13	Worldview2
		10-Mar-13		
	2013/14	03-Dec-13	2-Dec-13	Field spectrometer
		05-Feb-14	13-Dec-13	ARA hyperspectral
			5-Jan-14	Worldview2
SRA (B Salter)	2013/14	21-Nov-13	19-Apr-13	Worldview2
	2014/15	01-Dec-14	5-Jan-14	Worldview2
Mackay (Strip Trials) Farmacist	2014/15	22-Dec-14	18-Dec-14	Worldview2
Burdekin				
(Farmacist) Ryan Jones	2014/15	9-Jan-15	18-Dec-14	Worldview2
(Farmacist) Paul Roncato	2014/15	9-Jan-15	23-Dec-14	Worldview2
			23-Dec-14	Worldview2
EHP Site (Growers 14)	2013/14	25-Nov-13	29-Jan-14	Worldview2
(UQ/SRA Trial)	2012/13	1-Oct-12	14-Jan-13	GeoEye
		1-Feb-13	24-May-13	Worldview2

Trial details:

- **Site 1 (Tully SRA)**

48 plots (6 row x 30m long plots) of Q208 with 4 replicates of 12 applied nitrogen rates (0, 30, 60, 75, 90, 105, 120, 135, 150, 180, 210 and 240 kg/ha). (Figure 17 a)

- **Site 2 (Euramo SRA)**

36 plots (6 row x 27m long plots) of Q208 with 3 replicates of 12 applied nitrogen rates (0, 30, 60, 75, 90, 105, 120, 135, 150, 180, 210 and 240 kg/ha). (Figure 17 b)

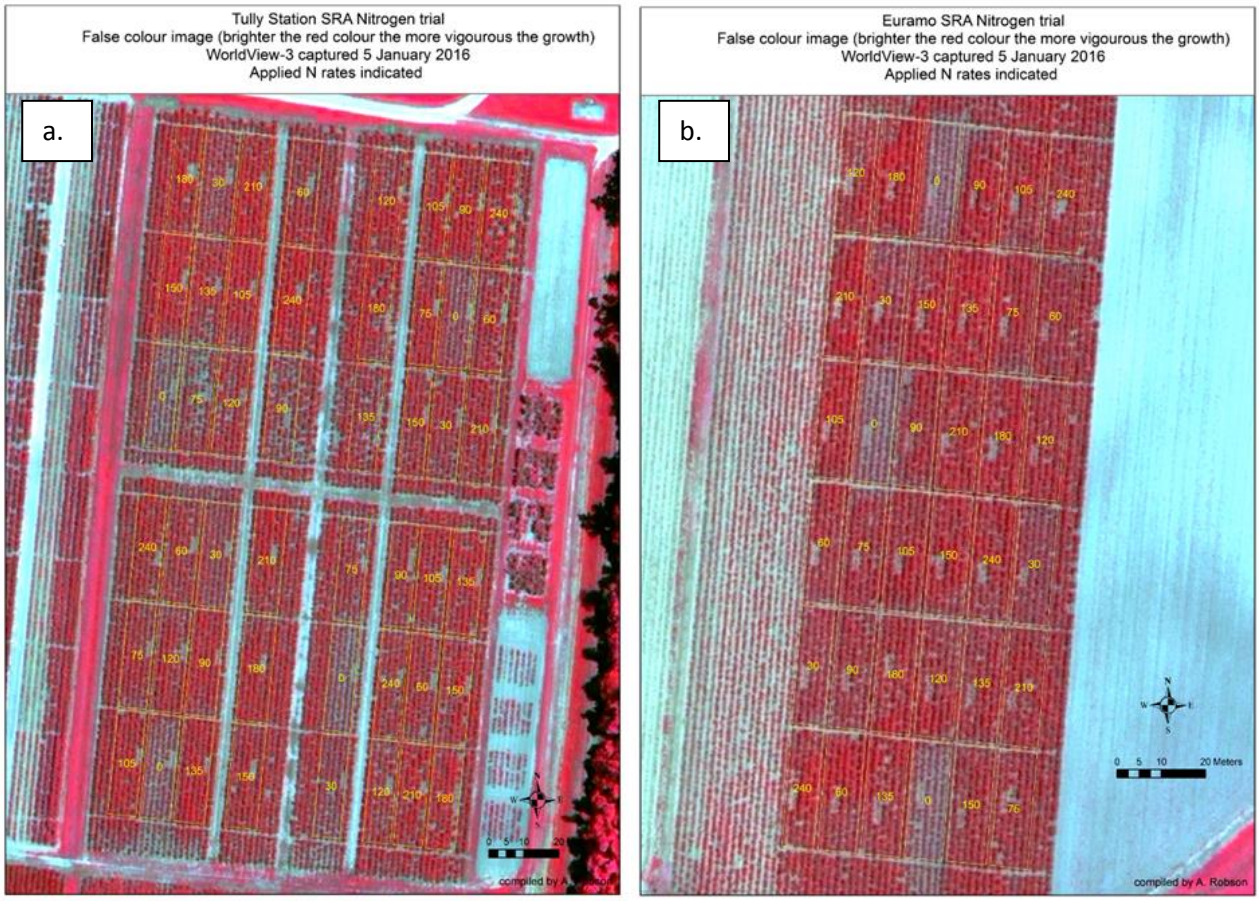


Figure 17. False colour Worldview2 image (acquired 15 Jan 2016) of SRA nitrogen trials: a. Tully and b. Euramo.

- **Mackay UQ/SRA Trial:**

  384 plots of which only 86 were intensively sampled for foliar N concentration at the out-of-hand stage. These plots consisted of 15 genotypes with 6 replicates of Q186, SP80-1816, N14, Q138, QN01-1075, QN01-1735, KQ228, QBYC04-10559, SP79-2313, R570, QC91-580, Q183, Q208, 5 replicates of Q240 and 3 replicates of Q218, The trial was planted on the 21st November 2014, with 40 kg/ha of Nitrogen applied to the low nitrogen treatments and 160 kg/ha to the high. (Figure 18 a).

- **Burdekin UQ/SRA Trial:**

  This trial is a replicate of the Mackay UQ/SRA trial and as such 86 of the 384 plots planted were again intensively sampled for N concentration at the out-of-hand stage. The 14 genotypes slightly varied to include 6 replicates of Q186, SP80-1816, N14, Q138, QN01-1075, QN01-1735, KQ228, QBYC04-10559, SP79-2313, QC91-580, Q232, QS95-6004, QC91-580, Q208 and Q218. The trial was planted on the 2011, with 40 kg/ha of Nitrogen applied to the low nitrogen treatments and 160 kg/ha to the high. (Figure 18 b).

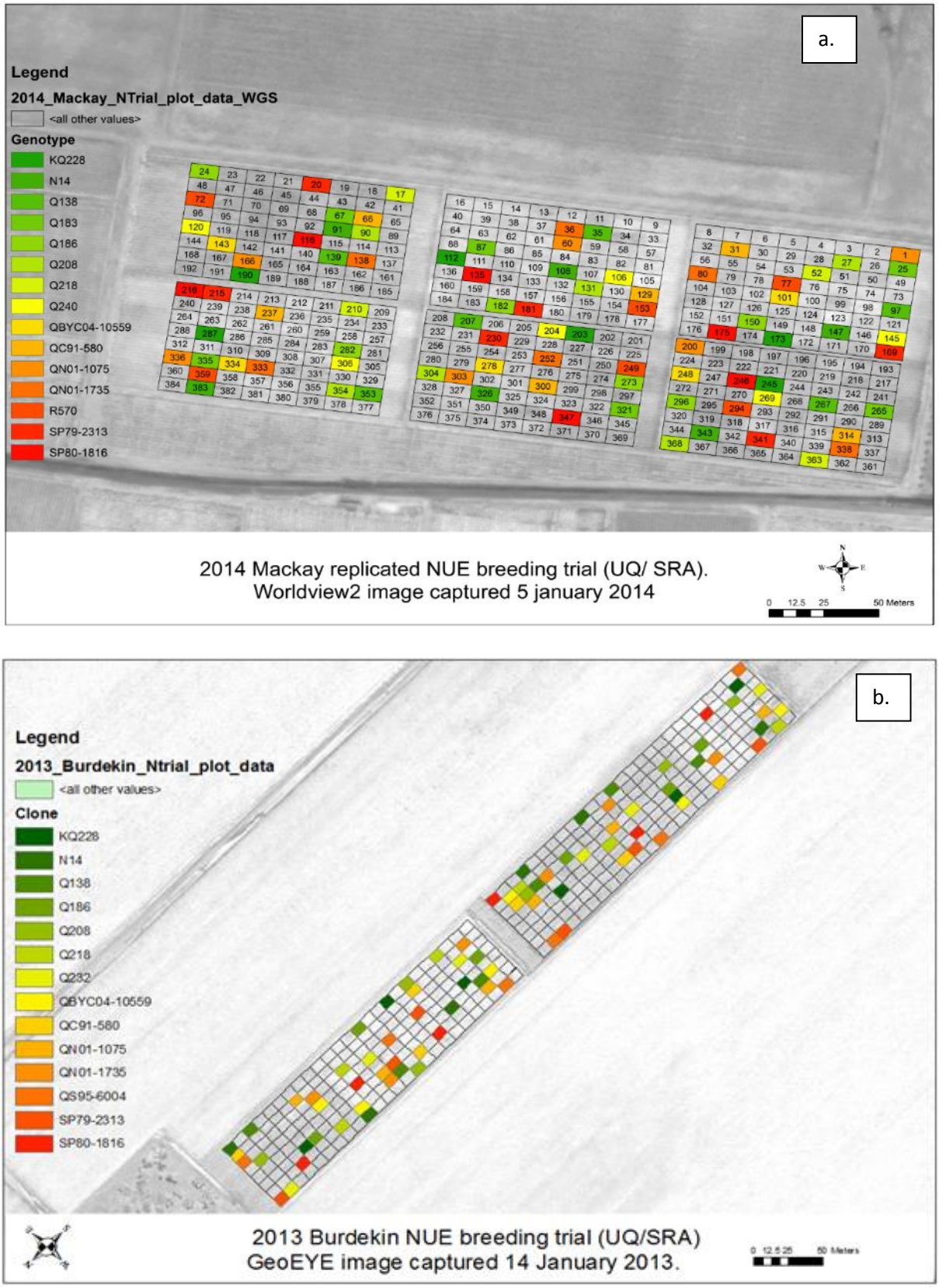


Figure 18. a. Trial design of the Mackay UQ/SRA trial and b. Burdekin UQ/SRA trial, the 86 plots sampled for N concentration are highlighted.

- **Mackay SRA (B Salter) Trial (2015):**

24 plots of Q208, 2nd ratoon, consisting of 2 separate trials:

Trial 1: 4 applied nitrogen rates (0, 95, 150 or 200 kg/ha) plus two trash treatments (green cane trash blanket (GCTB) or burnt). 2 of the four 4 replicates sampled 21st Nov 2014 (plots highlighted in the top left and bottom right quadrant of the trial in Figure 19).

Trial 2: 2 Nitrogen rates (0 or 150 kg/ha) plus two fallow treatments (soy beans or bare fallow). Again 2 of the four 4 replicates sampled 21st Nov 2014 (plots highlighted in the top right and bottom left quadrant of the trial in Figure 19).

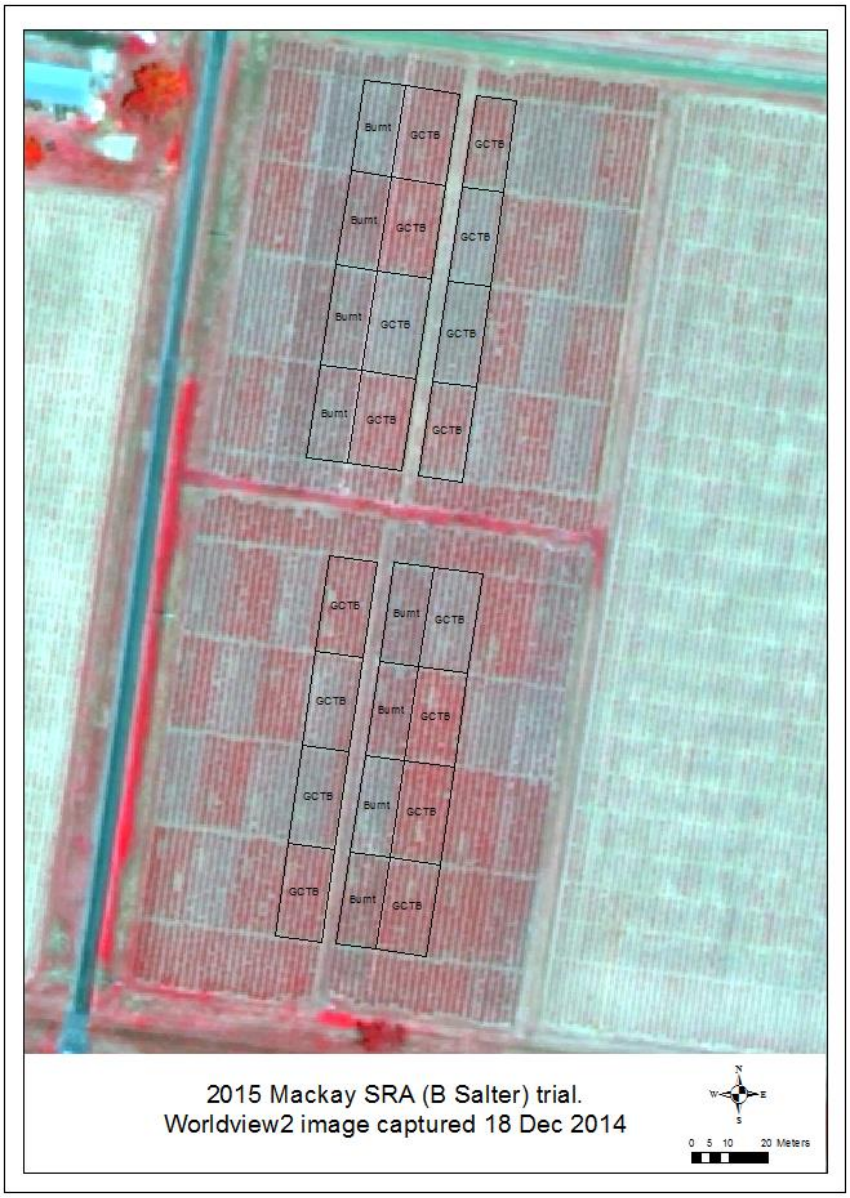


Figure 19. Trial design of the Mackay SRA (B Salter) trial, the 24 plots sampled for N concentration are highlighted.

- **Burdekin EHP trials**:

Grower 14: 1st ratoon crop. 3 nitrogen rates 210, 250 and 290 Kg/ha applied in a strip trial format. Sampled in 18 locations (Point source samples only: 3rd leaf). (Figure 20).

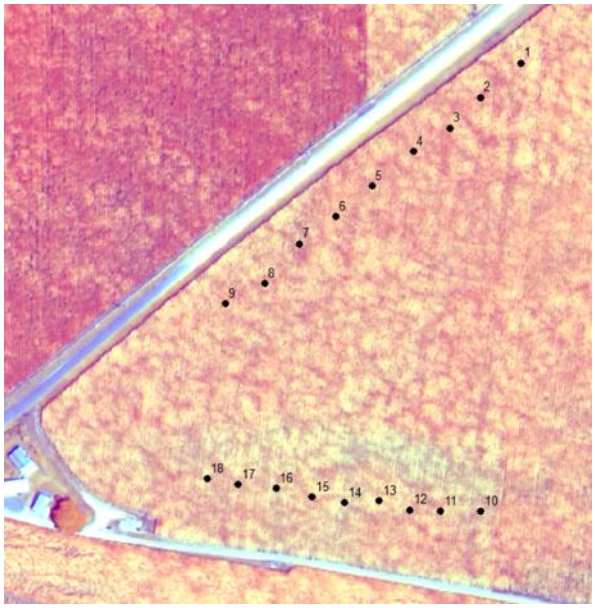


Figure 20. False colour Worldview2 image (29 Jan. 2014) of the Burdekin EHP (grower 14) site with field sampling locations overlaid.

- **Burdekin (Farmacist) Strip trials:**

The trial design of the Burdekin 'Roncato' (cv. KQ228) and 'Jones' (cv. Q208) sites are provided in Figure 21 a, with the false colour Worview2 imagery of each trial provided in Figures 21 b and c. These trials were evaluating crop response to a varying forms and rates of applied Nitrogen.

a.

		Rates versus soil types	
		10	11
		Sand	Loam
	GIS Zone	1	1
	Nitrate (mg/L)	0.3	0.4
	CEC	9.98	11.1
		Ryan Jones	***Roncato***
Treatment 1	Grower	Urea @ 220N	Urea @ 220N
Treatment 6	IMPACT	CR25% @ 220N	CR25% @ 220N
Treatment 7	IPL	Entec @ 220N	Entec @ 220N
Treatment 8	Grower	Urea @ 160N	Urea @ 160N
Treatment 9	IMPACT	CR25% @ 160N	CR25% @ 160N
Treatment 10	IPL	Entec @ 160N	Entec @ 160N
Application Date		25-Sep	29-Aug
Workspace			

Figure 21. a. trial design of the Burdekin Roncato and Jones sites. b. false colour of the Roncato and c. jones sites (acquired 23 Dec 2014) with the sample locations overlaid.

- **Mackay (Farmacist) Strip trials:**

Simpson trial: Q232 2nd ratoon. 16 plots (8 sampled). 4 treatments (Granular Urea 120 kg/ha, Entec Urea 120 kg/ha, Agrocote (controlled release) Urea 120 kg/ha and 0 application (Figure 22 a).

Young trial: Q232 2nd ratoon. 15 strip plots (10 sampled). 5 treatments (UreaS 160 kg/ha, Urea75: Agricote25 160 kg/ha, Urea75: Agricote25 128 kg/ha, Entec Urea 160 kg/ha and Entec Urea 128 kg/ha) (Figure 22 b)

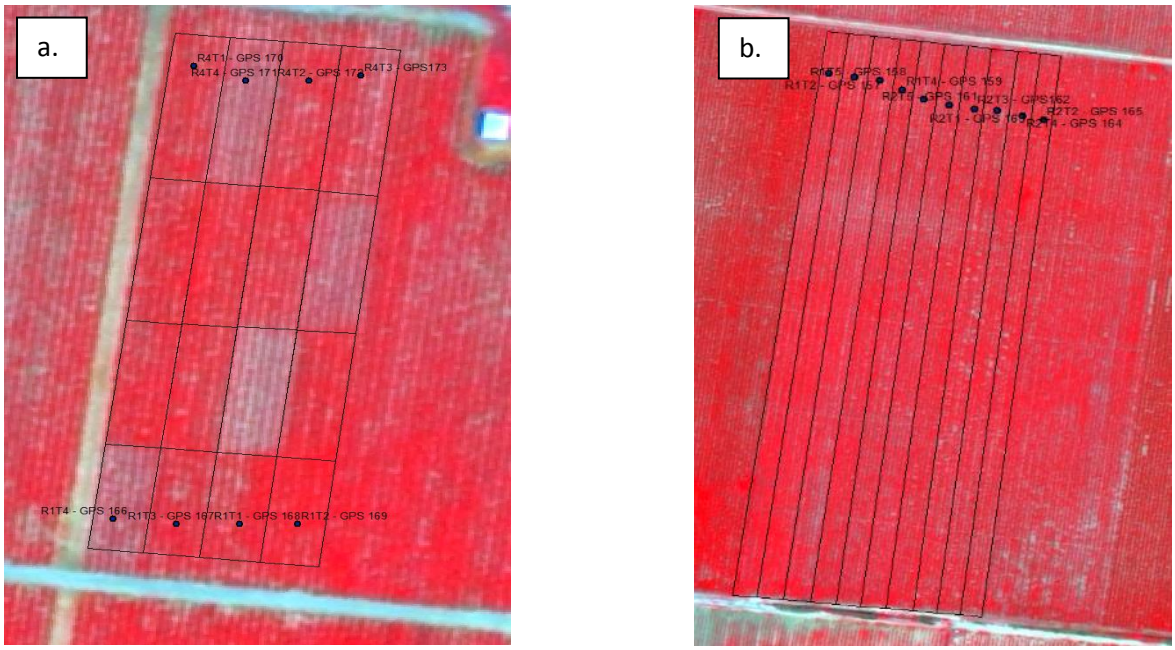


Figure 22. a. trial design of the Mackay Simpson and b. Young trial sites. False colour worldview2 image acquired 18 Dec 2014, with the sample locations overlaid.

Field sampling methodology:

As seen in Table 15, the collection of field data occurred at various time throughout the growing season. For some trials sample timing coincided with crop age i.e. 3, 6 and 9 months as well as at harvest, whilst for others only the 3rd month sample was undertaken. The latter is commonly referred to the out- of – hand stage as this is the last occasion when crop inputs such as N can be applied without the risk of cane damage from machinery.

Field sampling of the replicated field trials consisted of two methods: 'total' and '3rd leaf'. For the Mackay UQ/SRA Trial (2014) the 'total' sample consisted of 8 stalks cut from each plot with all leaves removed. Dead leaves were disposed of, whilst 16 green leaves were retained, weighed and then dried at 60°C. Stalk material was shredded and mixed with a 200 g subsample weighed and then dried at 60°C for at least 3 days. Dried leaf and stalk subsamples were combined, weighed, ground in Retsch ZM 200, and then ground again in a ball mill Retsch MM400. The nitrogen concentration of each sample was measured by LECO CN analysis.

For the Mackay SRA Trial (2015) and Tully SRA Euramo Trial (2015) trials only 'total' samples were collected. 5 whole shoots were shredded including all leaves with few dead leaves or stalk material present due to the young age of the crop.

As the name suggests, the '3rd leaf' samples consisted of only the third leaf from the top of the stalk being removed from 10 plants within the centre of each plot. Samples were processed as per 'total'. This sampling protocol was conducted for the EHP (grower 14) and Mackay UQ/SRA (2014) site.

Foliar nitrogen concentration was calculated as both a percentage (%) and as kg/ ha, depending on whether biomass weights were available.

Ground based hyperspectral measures of the Mackay UQ/SRA 2014 trial were obtained on the 3- 4th December 2013 using an Analytical Spectral Devices TM (Fieldspec, 1997) (ASD). The protocols of data collection are provided in the previous section *'The non- invasive screening of breeding trials'*.

Spectral analysis:

The collection of spectral data from the various sensors generally coincided with the out-of-hand field sampling i.e. December (Table 15). However, continued cloud cover did delay the acquisition of some satellite imagery acquisitions until January. Additional imagery was also captured in April and May to coincide with pre-harvest.

The specific details of the multispectral Worldview2 platform and the hyperspectral spectrometer and ARA airborne sensor are provided in the previous section *'The non- invasive screening of breeding trials in response to applied* treatments'. Additionally, The GeoEYE satellite (http://www.satimagingcorp.com/satellite-sensors/geoeye-1/) provides a spatial resolution of 1.81 m (multispectral) and 0.46 m (panchromatic band) and 4 band spectral resolution: Blue: 450 - 510 nm Green: 510 - 580 nm Red: 655 - 690 nm and Near Infra Red: 780 - 920 nm. As GeoEYE only offers 4 spectral bands, compared to the 8 provided by Worldview2, there is a limited ability to derive vegetation indices other than visible/ NIR ratios, such as NDVI.

To allow the extraction of spectral data from each plot, the boundary of each trial was manually digitized using ArcGIS ver. 10. An internal buffer of 50 cm was then applied to ensure non-treatment related information, such as those associated with edge rows, was excluded. The extraction of both the satellite and airborne ARA hyperspectral data was undertaken using the software SAGA GIS GUI. For point source samples, such as those collected for the within strip trials, a 3 m buffer was applied around each point.

From the extracted satellite imagery data, a number of structural and pigment based vegetation indices (VI) were derived (Table 16). Each VI was then correlated against measured N concentration (% and Kg/ha) to identify that which consistently produced the highest coefficient of determination (R^2). For the extracted hyperspectral data, the water absorption bands (1350- 1420, 1760- 1965 and 2450- 2500 nm) were initially removed from both the ground based and airborne data. Simple linear regressions between reflectance value and N concentration (total and 3rd leaf) were then undertaken to determine the specific wavelengths and band ratios best correlated to N concentration.

Table 16: Vegetation Indices investigated, where R represents the reflectance measured in the corresponding spectral band. Note, a number of other MCARI, OSAVI, MSAVI, CVI ratios were investigated.

Normalised Difference Vegetation Index (NDVI)	$R_{NIR} – R_{Red} / R_{NIR} + R_{Red}$
GreenNDVI	$R_{NIR} - R_{Green} / R_{NIR} + R_{Green}$
MidIRNDVI	$R_{MIR} – R_{Red} / R_{MIR} – R_{Red}$
Plant Cell Density (PCD)	R_{NIR} / R_{Red}
MidIRPCD	R_{MIR} / R_{Red}
MidIRNDVIPCD	MidIRNDVI / R_{Red}
Transformed chlorophyll absorption reflectance index (TCARI)	$-3*(R_{Red} – R_{Green}) – 0.2*(R_{Red} – R_{Green}) *(R_{Red} / R_{NIR} + Red))$
Two-band Enhance Vegetation Index (EVI_2)	$2.5*((R_{NIR} - R_{Red})/(R_{NIR} +(2.4* R_{Red})+1))$
Structure insensitive pigment index (SIPI)	$R_{NIR} – R_{Blue} / R_{NIR} – R_{Red}$
Modified Simple Ratio (MSR)	$(R_{NIR}/ R_{Red})- 1)/(SQRT(R_{NIR}/ R_{Red})+ 1)$
REGNDVI	$R_{Red-edge} - R_{Green} / R_{Red-edge} + R_{Green}$
RENDVI	$R_{NIR} – R_{Red-edge} / R_{NIR} + R_{Red-edge}$
N2RENDVI	$R_{MIR} – R_{Red-edge} / R_{MIR} + R_{Red-edge}$
N1/Red/RENDVI	$R_{NIR} – R_{Red} / R_{NIR} + R_{Red-edge}$

Results:

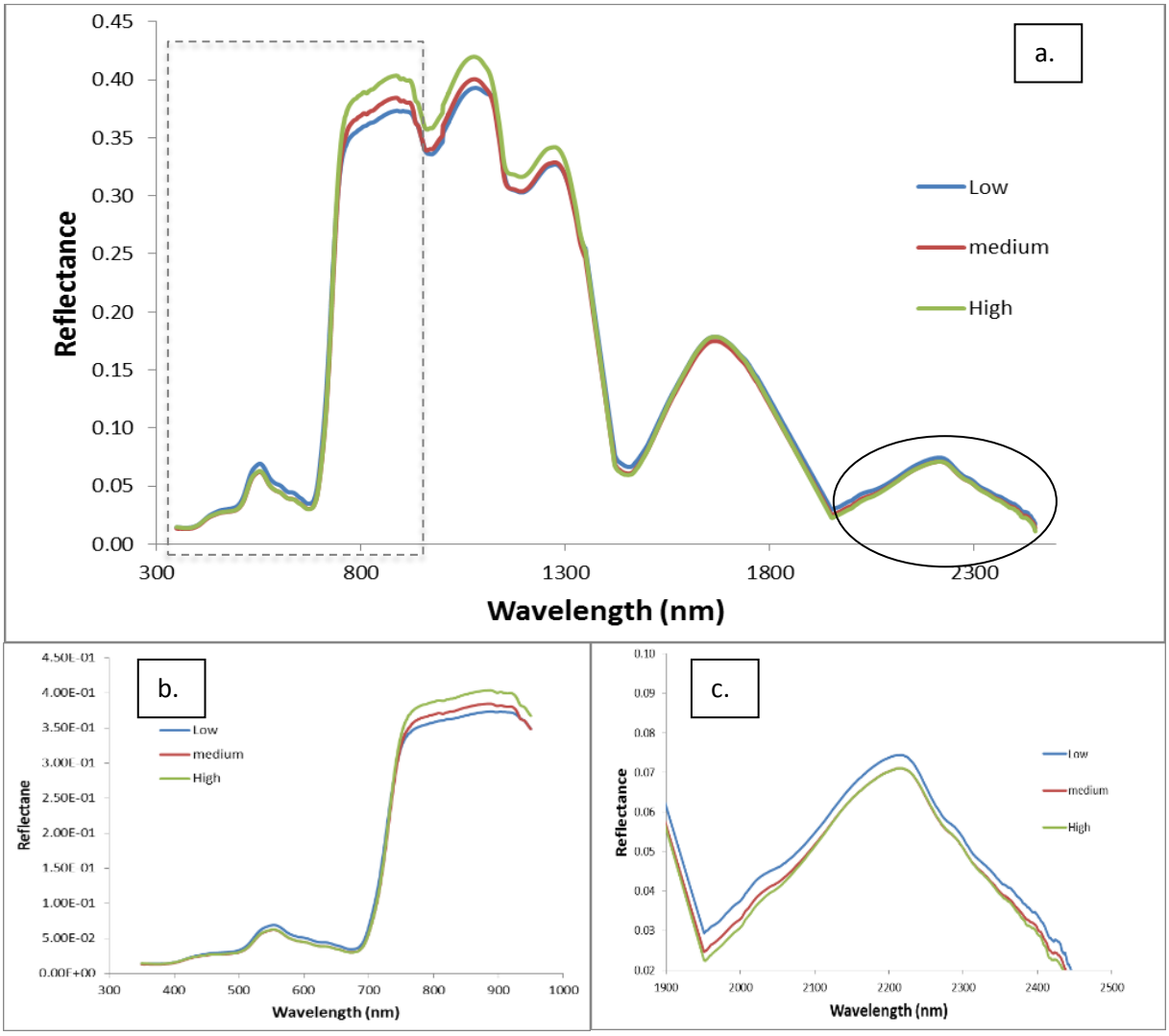
Identifying specific wavelengths and vegetation Indices best correlated to foliar Nitrogen concentration:

The following results are from the Mackay UQ/SRA trial where all sensors were used to measure the canopy reflectance of 86 plots (Figure 18 a) at the 'out–of-hand' stage. The spectral measures coincided with the in-field sampling of 'total' and '3rd leaf' N (%). From Figure 23, the spectral response curves are characterised by strong absorption within the visible part of the electromagnetic spectrum (400 – 680 nm) followed by a transition region from low to high reflectance, termed the red- edge (RE) (680- 780 nm); strong leaf reflectance within the near infrared (NIR) region (700-1300 nm); and a gradual decline through the mid near infrared region (MNIR) (1300-2500 nm).

The relatively low reflectance in the visible region (400- 680 nm) is the result of light absorbing compounds, namely pigments with chlorophyll absorption 'troughs' at 450 nm (blue) and 650 nm (red). The small reflectance peak at (510 – 580 nm) corresponds to the visible green wavelength and is attributed to chlorophyll pigment and is the reason leaves appear green to the human eye. The red- edge region (RE) indicates a transition between the strongly absorbing red wavelength region of the leaf through to the highly scattering near infrared region. The near infrared spectral region often provides a strong insight to the 'health' or 'vigour' of a plant. When exposed to radiation within the NIR spectral region, the atomic bonds of certain molecules display harmonic vibrations at specific wavelengths within the MNIR. This bond vibration is considered anharmonic at higher vibration states, resulting in overtone bands. These overtone and

combination bands typically correspond to the presence of the lightest atom hydrogen, and combinations of this atom with nitrogen as well as carbon, oxygen, or sulphur.

The spectral signatures derived from each sensor exhibit a clear separation in response to the varying canopy Nitrogen concentrations: low (0.95 – 1.3 %), medium (1.3 – 1.6 %) and high (1.6 - 2.1 %). The greatest separation occurs between 800 nm and 950 nm (midIR), followed by the visible region 400 – 680 nm, and to a lesser extent the far Mid IR regions 1950- 2000 nm and 2350 -2400 nm. These results are consistent with previous research that has identified increased nitrogen concentration lowers reflectance within the visible spectral region, increases reflectance in the near infrared region, and a shift in the transitional red edge slope (Rodriguez et al. 2006, Hansen and Schjoerring 2003). More specifically the wavelengths 870 – 874 nm, 908 – 910 nm, 1976- 1996 nm and 2058- 2076 nm have all been previously correlated to N- H bonds.



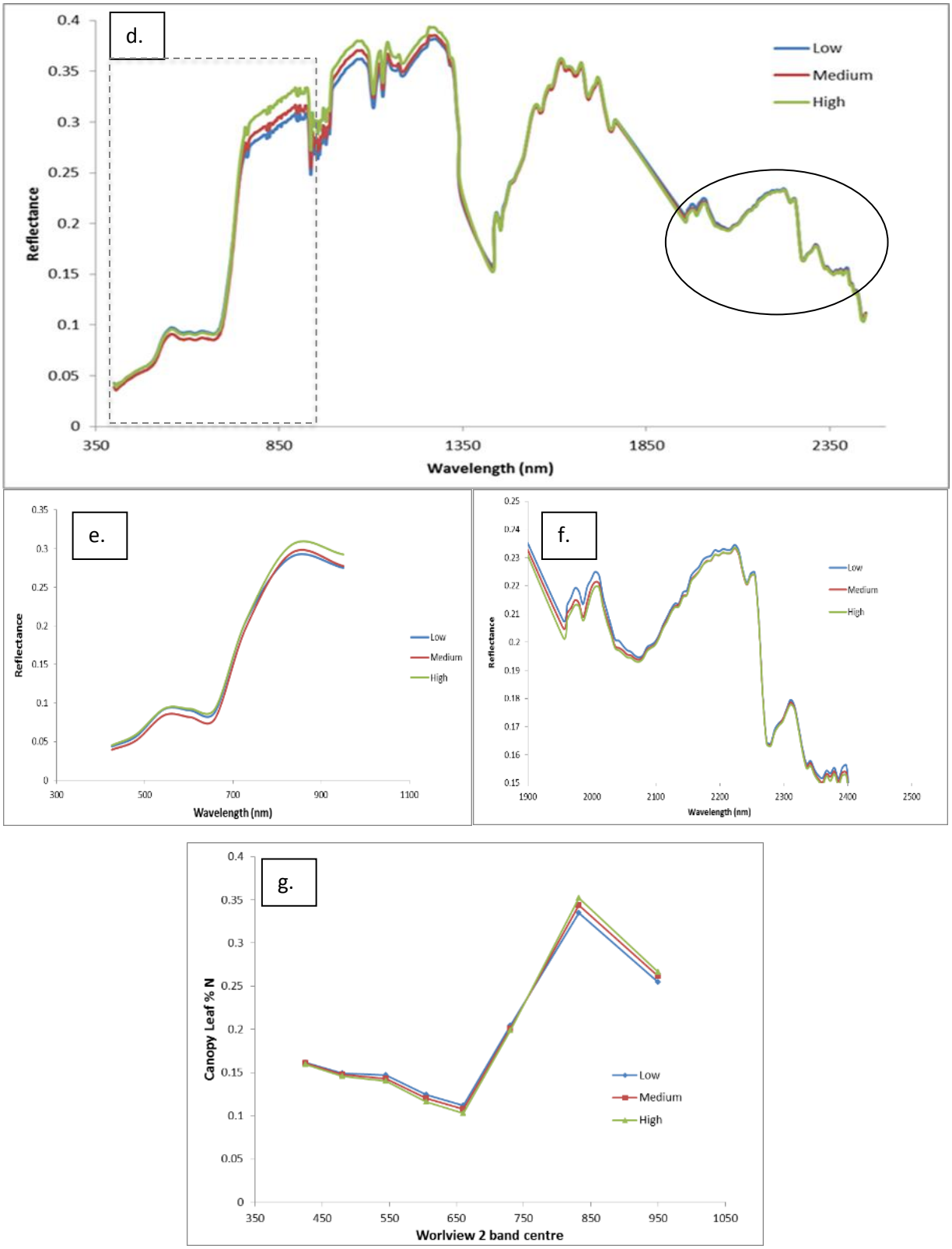


Figure 23. Spectral response curves of sugarcane leaves with varying levels of measured N low (0.95 – 1.3 %), medium (1.3 – 1.6 %) and high (1.6 - 2.1 %), measured with an ASD field spectrometer (a, b and c); Airborne hyperspectral sensor (ARA) (d, e and f) and Worldview2 satellite (g). The 'dashed line' rectangle in a and d, and enlarged subsection (b and e) indicate the spectral range encompassed by the Worldview2 satellite (g).

The analysis of hyperspectral airborne data (ARA) for the Mackay UQ/SRA Trial (2014) identified the spectral bands 621 nm (Red), 742 nm (RedEdge) and 913 nm (NIR) as best correlated to N % (3rd leaf). The optimal band ratio was identified as $REN2NDVI_{ARA}$ producing a coefficient of determination of R^2=0.52 to N % (3rd leaf).

$$REN2NDVI_{ARA} = \frac{(913 - 742)}{(913 + 742)}$$

For the field hyperspectral radiometer (ASD) measure of the Mackay UQ/SRA Trial (2014), the wavelengths 634 nm, 750 nm and 880 nm were identified to produce the highest correlations to N % (3rd leaf). The optimal band ratio was identified as $REN2NDVI_{ASD}$ producing a coefficient of determination of R^2=0.41 to N % (3rd leaf).

$$REN2DVI_{ASD} = \frac{(880 - 750)}{(880 + 750)}$$

For the Worldview2 sensor, the VI $REN2NDVI_{WV}$ derived from the RedEdge (705 - 745 nm) and Near- IR2 (860 - 1040 nm) spectral bands produced the strongest coefficient of determination of R^2=0.31 to N % (3rd leaf).

$$REN2NDVI_{WV} = \frac{(\text{WV Band 8} - \text{WV Band 6})}{(\text{WV Band 8} + \text{WV Band 6})}$$

The identification of similar spectral bands from all three sensors is highly encouraging as it demonstrates the findings are not sensor specific, and more importantly they are comparable to those reported by Patil and Nadagouda (2008) (R620 -R680 nm and R770 – R860 nm), Jackson et al., (1980) (R760- R900)/(R630- R690) and Portz *et al*., (2012) (In R760 – In R730) * 100) as well as those identified in Rice (Dunn *et al*. 2016).

Determining the Robustness of $REN2NDVI_{WV}$.

Following the results of the Mackay UQ/SRA trial, further evaluation of the relationship between $REN2NDVI_{WV}$ and N concentration was conducted across multiple trials, as detailed in Table 15. As seen in Figure 24, the Euramo SRA trial (2014/15) produced the highest coefficient of determination (R^2=0.59) followed by with Mackay SRA trial (2014/15) (R^2=0.59), Euramo SRA trial (2015/16) (R^2=0.48), Tully SRA trial (2015/16) (R^2=0.43), Mackay UQ/ SRA Trial (2013/14) (R^2=0.31), Tully SRA trial (2014/15) (R^2=0.30), Mackay (Simpson) Farmacist 2014/15 (R^2=0.27), Mackay (Young) Farmacist 2014/15 (R^2=0.11), Burdekin EHP (Grower 14) 2013/14 (R^2=0.04), Burdekin (Jones) Farmacist 2014/15 (R^2=0.01), and Burdekin (Roncato) Farmacist 2014/15 producing a negative relationship (R^2=0.02). The obvious separation of data points from each trial is considered to be predominantly the result of the trials being grown in different locations, seasons and with different cultivars, a hypothesis supported by Rodriguez et al. (2006).

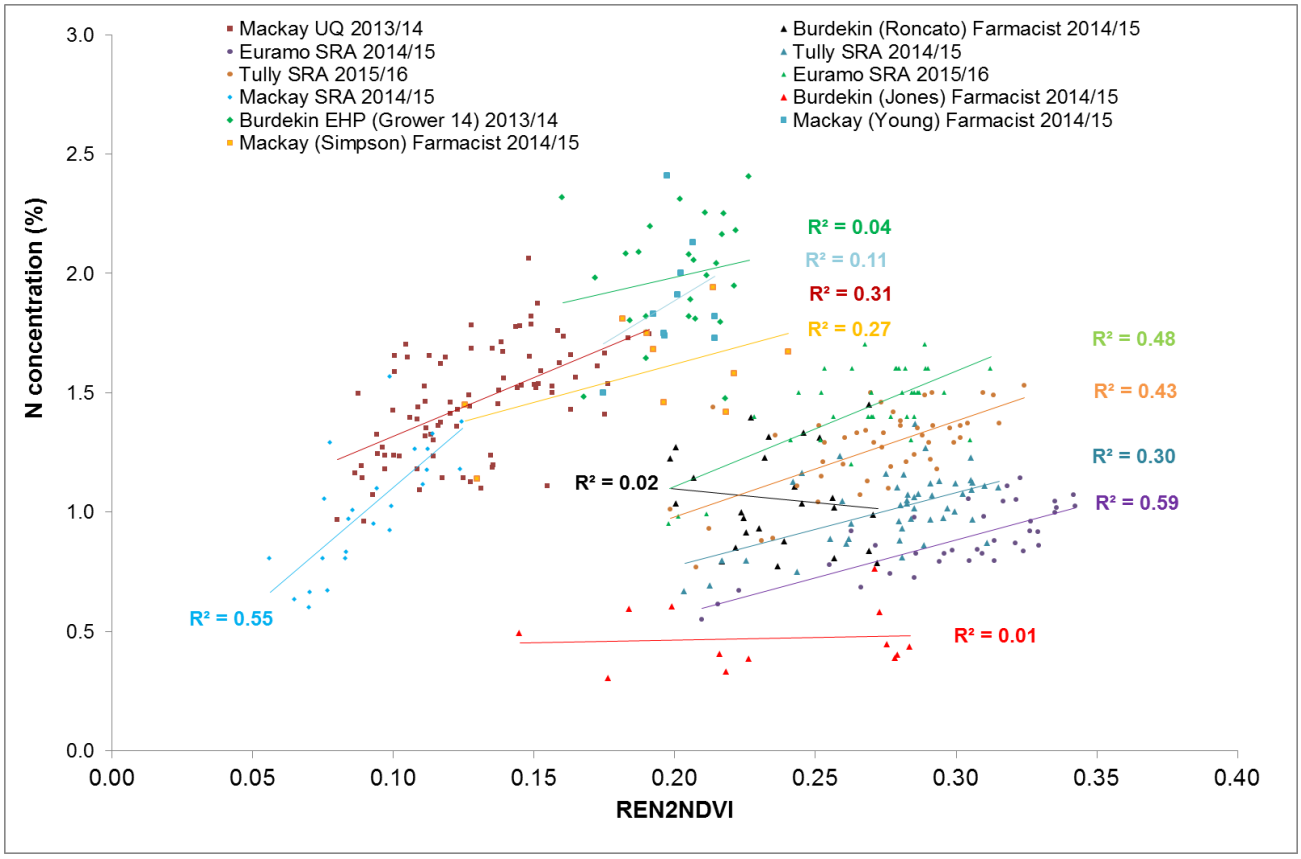


Figure 24. Scatter plots indicating the relationship between N 3 rd leaf concentration (%) and the derived vegetation index $REN2NDVI_{WV}$.

The vertical shift in the linear relationships between the temporal (2014/15 and 2015/16) Tully station and Euramo SRA trials, does support the hypothesis of a seasonal influence. However, it is possible that ratoon age may also be having an effect. In terms of growing location, the data does indicate some clustering of points from the 4 Mackay trials as well as from the 4 Tully trials, suggesting that a 'locational specific' algorithm may be possible. However, the data points from the 3 Burdekin trials are well dispersed which may be again attributed to an additional influence such the point source sampling nature of the trials. Unlike the plot trials that provide a clear boundary for spectral data extraction, the point source sampling technique is exposed to locational errors associated with the non- differential GPS geotagging of sites and the orthorectification accuracy of the image (~ 1m). The compounding of these two sources of error can result in the extracted spectral data not aligning with the location that the sample was collected from and as seen in Figure 14, this can make a significant difference in reflectance values when the trial crop expresses considerable inherent spatial variability.

Cultivar or genotype is also suspected to influence the relationship between $REN2NDVI_{WV}$ and Nitrogen concentration (%). As displayed by Figure 25, the 15 genotypes grown in the Mackay UQ/SRA trial clearly exhibit different linear relationships between $REN2NDVI_{WV}$ and Nitrogen concentration (%), ranging from strongly positive (R^2=0.94) for Q208, to negative QC91-580. This result may be attributed to differing crop geometries, potentially the way N is partitioned or even from increased susceptibility of some genotypes to external stresses i.e. water, disease not related to the applied N treatments.

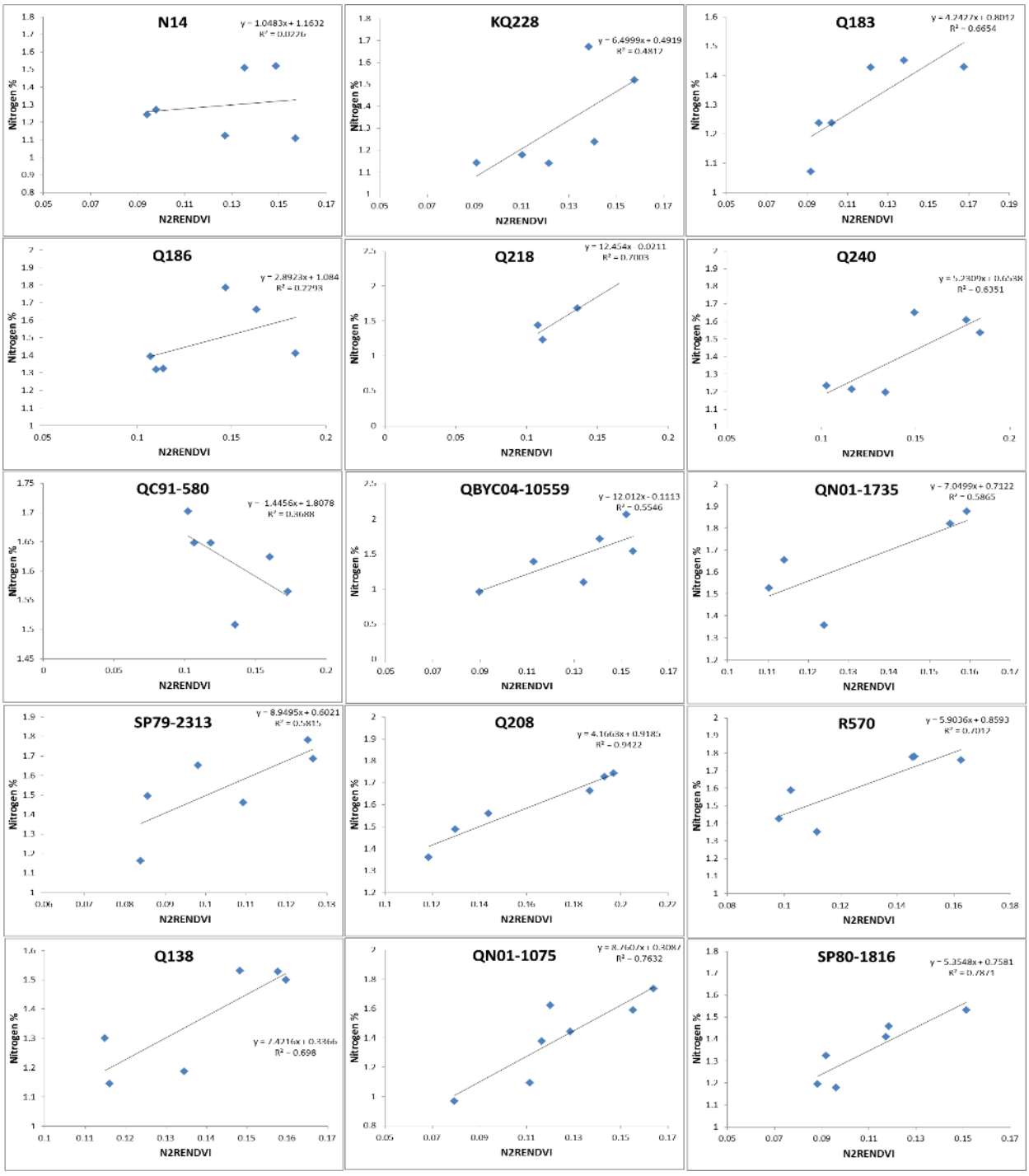


Figure 25: Linear relationships between N2RENDVI (Worldview2 acquired 05/01/2014) and canopy N % for the 15 genotypes grown at the UQ/ SRA Mackay site (2014/ 2015).

To further evaluate the influence of genotype, Table 17 identifies the correlation coefficients (R) calculated between N% (3rd leaf) and REN2NDVI measured by each sensing platform. On average, genotypes Q183 and QC91-580 produced the higher correlations to N %, whilst Q186 and SP80-1816 produced the lowest. These results indicate that the development of a 'non- genotype' specific REN2NDVI:N% predictive algorithm is unlikely.

Table 17. Correlation coefficients between REN2NDVI and 3rd leaf foliar N (%) derived from each of the 15 genotypes grown in the Mackay UQ/SRA Trial (2014).

Varieties	WV2	Field Spec	ARA
Q186	0.09	0.17	0.26
Q218	0.45	0.99	0.95
SP80-1816	0.28	0.42	0.19
N14	0.31	0.41	0.4
Q138	0.27	0.98	0.73
QN01-1075	0.43	0.82	0.75
QN01-1735	0.48	0.61	0.4
KQ228	0.17	0.75	0.5
QBYC04-10559	0.59	0.82	0.37
SP79-2313	0.67	0.62	0.53
R570	0.70	0.73	0.71
Q240	0.65	0.84	0.58
QC91-580	0.69	0.87	0.7
Q183	0.85	0.87	0.8
Q208	0.85	0.75	0.27

The variations between trials observed in Figure 24, may also be the result of discrepancies between the sampling methodologies employed by the varying research groups. As a simple demonstration, Figure 26 provides an interesting comparison of foliar N% measured from '3rd leaf' samples compared to those sampled from the whole canopy technique. Both samples were collected from the same plot trials (Mackay UQ Trial (2013/14)) and on the same day (December 'out-of-hand sampling event). This result clearly demonstrates a disparity between results obtained from each sampling technique. Although the coefficient of determination between a VI and the measures would be similar, the slope of the relationship and therefore linear algorithm would differ. If this methodology was to be extended across the Australian sugar industry then a consistent sampling protocol would need to be employed.

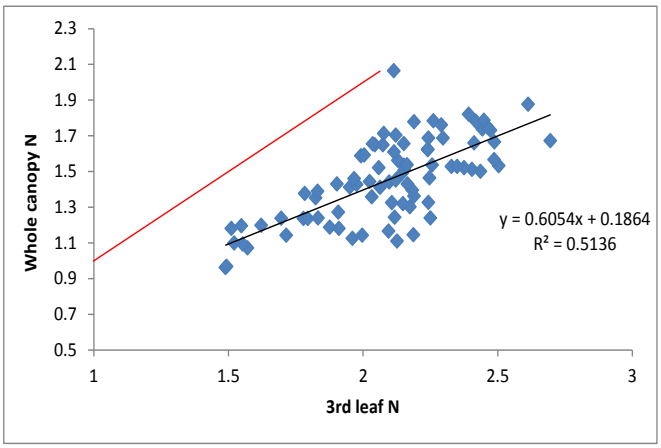
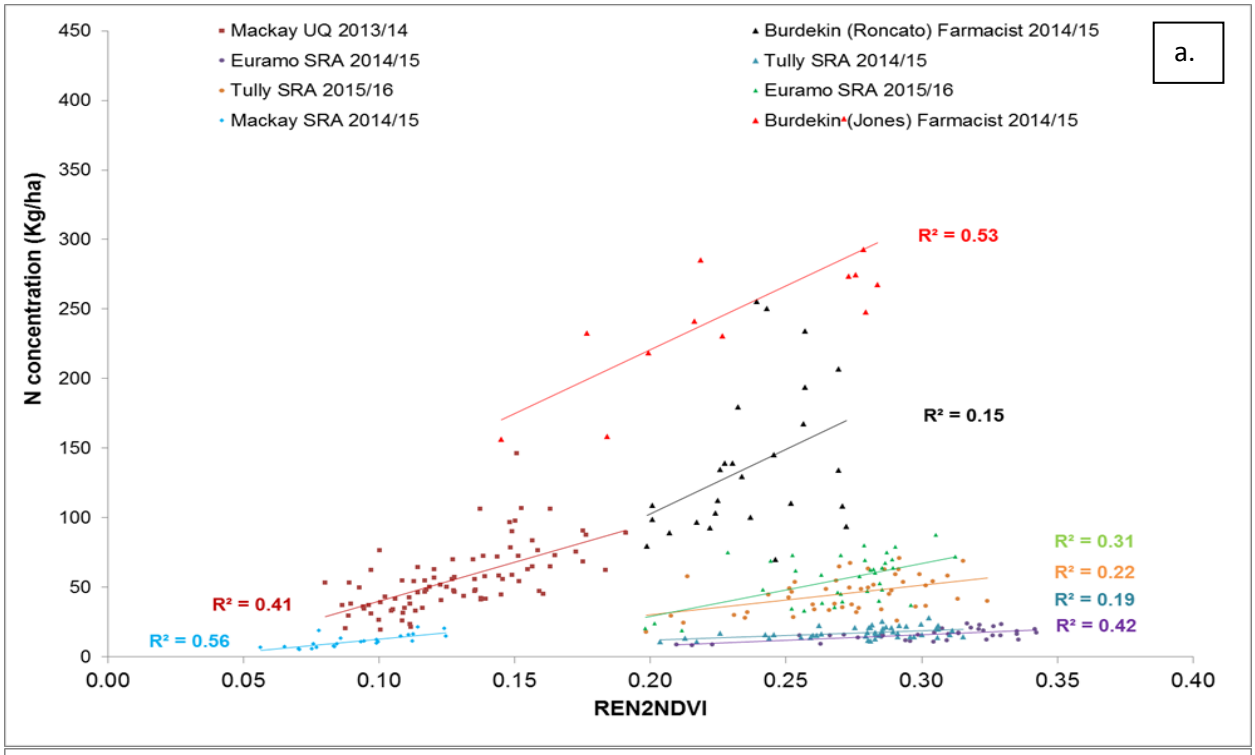


Figure 26: Red line indicates the 1:1 relationship between the whole canopy sampling technique for Foliar N%, whilst the blue markers/ black trend line indicate the relationship between the 3rd leaf sampling technique and the whole canopy technique from the same plots (n= 88).

As well as correlating $RENDVI_{WV}$ to foliar Nitrogen concentration as a percentage, an analysis of all trials was also conducted looking at N as kg/ ha. As seen in Figure 27 a, a wide separation of trials was identified. However, a random forest analysis of the same data set, identified the coastal blue band as having the highest explanation of variance (%) in explaining N (kg/ha). Figure 27 b, indicates why this may have occurred.



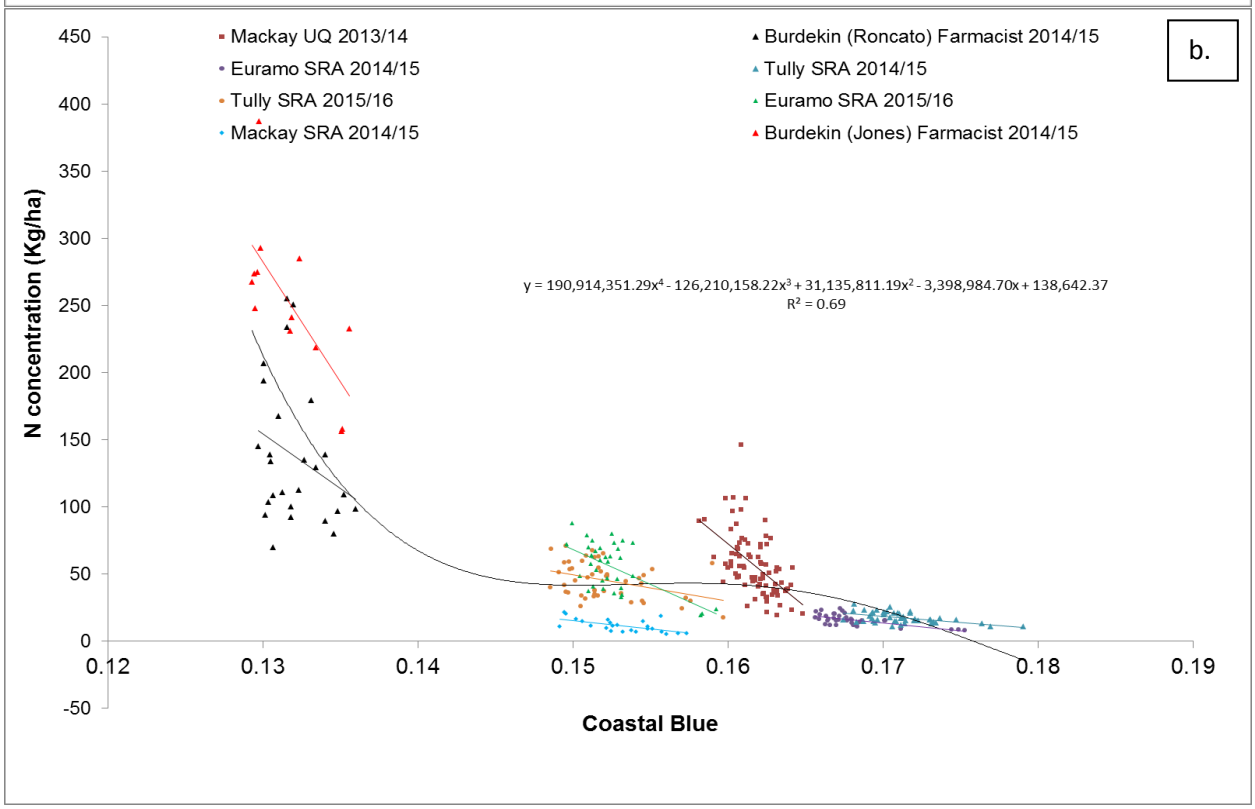


Figure 27.a. Scatter plots indicating the relationship between 3 rd leaf N concentration (Kg/ ha) and the derived vegetation index REN2NDVI$_{WV}$, and b. reflectance measured from the Worldview2 Coastal Blue band only.

Although the Coastal Blue band does appear to be less influenced by location and season etc., a major limitation of using only one spectral band compared to a ratio, is it is more susceptible to errors from atmospheric attenuation, distance from sensor to target and look angle. Therefore, unless those sources of error can be mitigated the ration RENDVI$_{WV}$ would be more appropriate.

Finally, as presented in Objective 1, timing of imagery can have a major influence on the prediction accuracy of yield. In terms of nitrogen concentration, imagery captured at times other than the out-of-hand stage when samples were collected, still produced positive correlations, especially for kg/ ha N later in the growing season (Table 18). Considering the strong relationships identified between the spectral reflectance of the sugarcane canopy and yield, it is likely that biomass is the main driver of the strong correlations. Although predictions made later in the season may be considered more appropriate considering sugarcane hasn't completed uptake until 6 months of growth, the ability to apply any remediation to the current crop at the late stage is impossible without causing major mechanical damage to the crop.

Table 18. Coefficients of determination (R^2) achieved between leaf N concentration (% and Kg/ha) from Worldview2 imagery captured later in the growing season.

trial	timing of sampling event	timing of WV2 image capture	R2 (REN2NDVI to %N total)	R2 (REN2NDVI to kg/ ha N total)
Burdekin UQ/SRA	Oct-12	24-May-13	0.37	0.04
Burdekin UQ/SRA	Feb-13	24-May-13	0.28	0.52
Mackay UQ/SRA	Dec-12	19-Apr-13	0.08	0.18
Mackay UQ/SRA	Mar-13	19-Apr-13	0.13	0.48

Deriving Nitrogen maps from imagery.

Using the algorithm developed from the relationship between $REN2NDVI_{WV}$ and nitrogen concentration, foliar nitrogen maps were derived for each region using the extent of the WV2 image. Figure 28, provides an example of a derived classified N % map with actual leaf % measures collected in Dec 2013 overlaid ('0' indicates no data was collected in those plots). A further example of this relationship extrapolated to the extent of the Worldview2 image capture is provided in Figure 28. Further examples are provided in the Outcome section of this report.

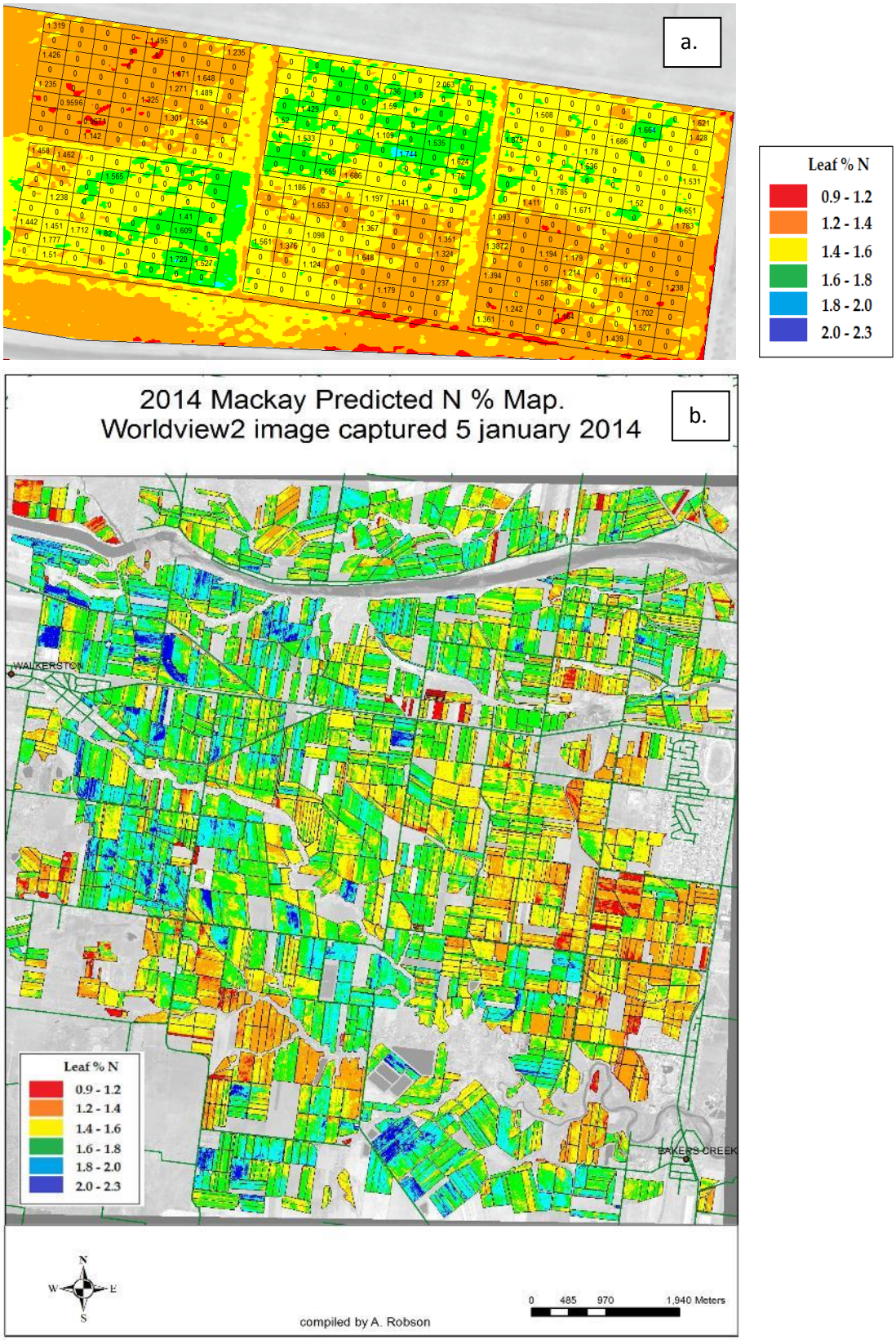


Figure 28. Classified foliar N % maps of the Mackay UQ/SRA trial (a) and the extent of the WV2 capture (b) derived from the linear relationship between N2RENDVI (Worldview 2 image: 5 Jan 2014) and field samples of N % (full canopy) collected in the first week of December 2013.

Summary.

The vegetation index (REN2NDVI) consistently produced the strongest correlation to foliar N concentration (% and kg/ha) than other VI's tested, although the relationship was influenced by trial location, growing season, cultivar grown and sampling method. As such, the use of this index should be considered purely as a qualitative measure of N concentration to either support targeted in- field sampling to produce a quantitative map or to provide a general understanding of the spatial and temporal variability of N concentration at the block, farm and regional level. With this information, growers can form a stronger understanding of within crop NUE and as such refine their management to employ variable rate applications of N, offering both financial and environmental benefit.

Section 4: Outputs and Outcomes

Outcomes:

DPI021 and now DPI025 have successfully introduced remote sensing technologies to the wider Australian sugarcane industry, including sugar mills, productivity services, growers and researchers. Through direct collaboration with these industry groups the potential of RS technology is being realised with new applications regularly being suggested and then actively pursued. Such applications include the monitoring of pest (cane grub and soldier fly) outbreaks, flood and disease, irrigation efficiencies, nutrition and soil health as well as the current outputs that support harvest scheduling and yield forecasting.

As outlined in "*Section 3: Output and Achievement of Project Objectives*" SPOT satellite imagery was identified as having the most appropriate spatial resolution, spectral resolution and price for yield mapping of sugarcane. The recent availability of SPOT 6 and 7 offers increased spatial resolution (6 m resolution for SPOT 6 and 7, as opposed to 10m resolution for SPOT 5), 1 day revisit time, and an ability to explicitly define the image order area, thus reducing costs for most regions.

A strong correlation between GNDVI and sugarcane yield was achieved for most areas. However, variation in climatic conditions between growing seasons and timing of imagery capture did influence the regional yield prediction accuracy. The development of the time series Landsat yield prediction model is a significant advancement to achieving higher accuracy yield predictions at the regional level. In addition, the statistical analysis of historic mill data, coupled with in season spectral measures of crop performance, offers improved accuracy at the farm and block level.

In terms of Nitrogen, results from this project have confirmed remote sensing as an effective tool for monitoring the spatial and temporal variation in foliar N concentration, and as such offers a highly beneficial tool to support management strategies such as 6 easy steps. The identification of REN2NDVI (mid near infrared red-edge normalized difference vegetation index) is a significant step in identifying spectral bands more sensitive to N concentration and therefore helps to refine what satellite and ground based sensors would be more appropriate for N measurement. For the 'on-the-go' tractor based application of N, sensors such as the Crop Circle, OPT-RX and Yara N- Sensor that offer spectral bands in the red- edge, green or higher NIR region are likely to be more appropriate than an NDVI based sensor e.g. Greenseeker, that is likely to suffer reflectance saturation.

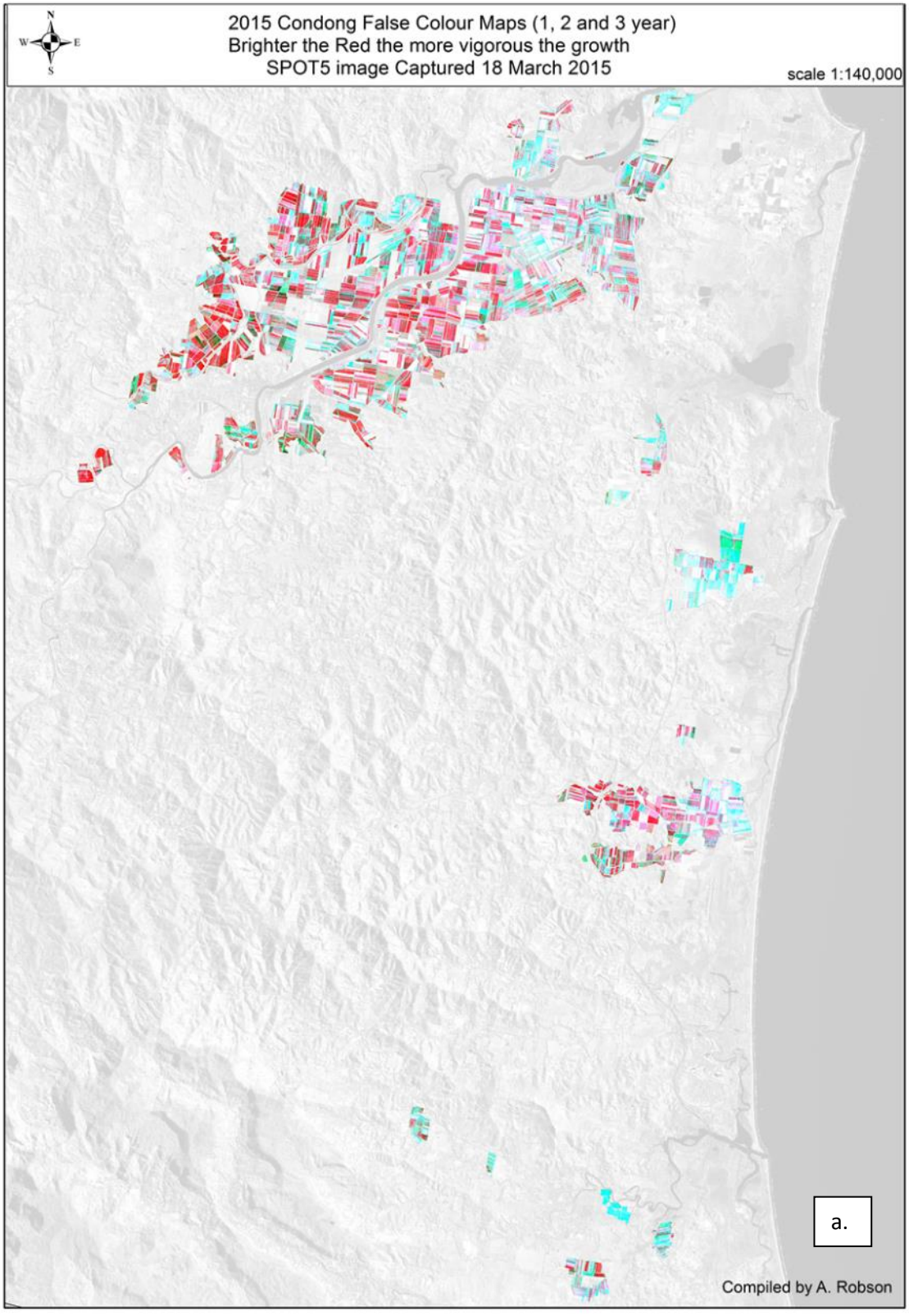
Finally, remote sensing technologies were demonstrated as an essential tool for supporting replicated field trials by presenting as an improved method for trial site selection, identifying the in-season incidences on non-treatment related constraints and as a non- destructive discriminator of genotype or phenotypic response to an applied treatment.

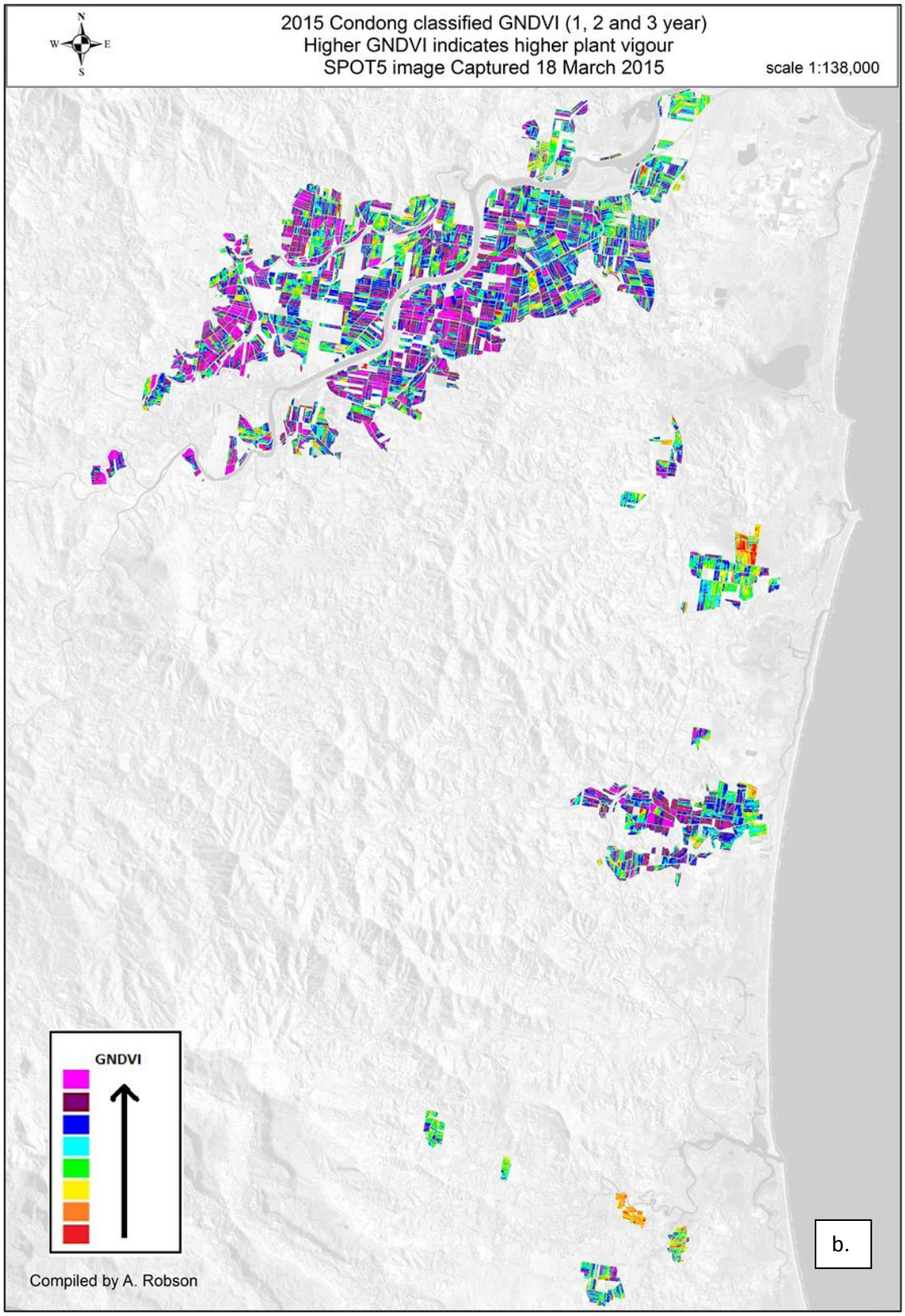
In order to extend DPI025 outcomes, the project team engaged with a range of industry groups, participated in field sampling with collaborative agencies, presented results at a range of forums, including ASSCT, field days and grower meetings, participated in SRA led reviews of precision agriculture and Nitrogen management and published a number of papers. A full list of these communications is provided in *'section 6: Industry Communication and Adoption of Outputs'*.

Outputs:

Project DPI025, produced many outputs including:

- Predictions of annual regional yield for up to 11 growing regions over the 3 years (described in the Objective 1 section). Predictions provided as a numerical value for each region (Table 2).
- Annual derivation and distribution of classified crop vigour and yield maps for each collaborating mill, for every block in each growing region (~60,000). Condong example provided in Figure 29.
- Remote sensing derived products (derived yield, GNDVI vigour) for individual farms. These outputs have been generated in response to direct requests from industry (mills, researchers, consultants and growers) and used for a range of applications including soil health, disease and pest monitoring. Some examples are provided in Figure 31.
- Qualitative nitrogen concentration maps for regions where Worldview2 imagery was acquired and where trials and associated ground truth data was available (Figure 33).
- In-season measure of plot response for replicated trials conducted by a number of entities including SRA, UQ, Farmacist, Incitec and CSIRO (Figures 10- 15 and 17- 22).


N
W
E
S
2015 Condong False Colour Maps (1, 2 and 3 year)
Brighter the Red the more vigorous the growth
SPOT5 image Captured 18 March 2015
scale 1:140,000
a.
Compiled by A. Robson


2015 Condong classified GNDVI (1, 2 and 3 year)
Higher GNDVI indicates higher plant vigour
SPOT5 image Captured 18 March 2015
scale 1:138,000
N
W
E
S
GNDVI
Compiled by A. Robson
b.

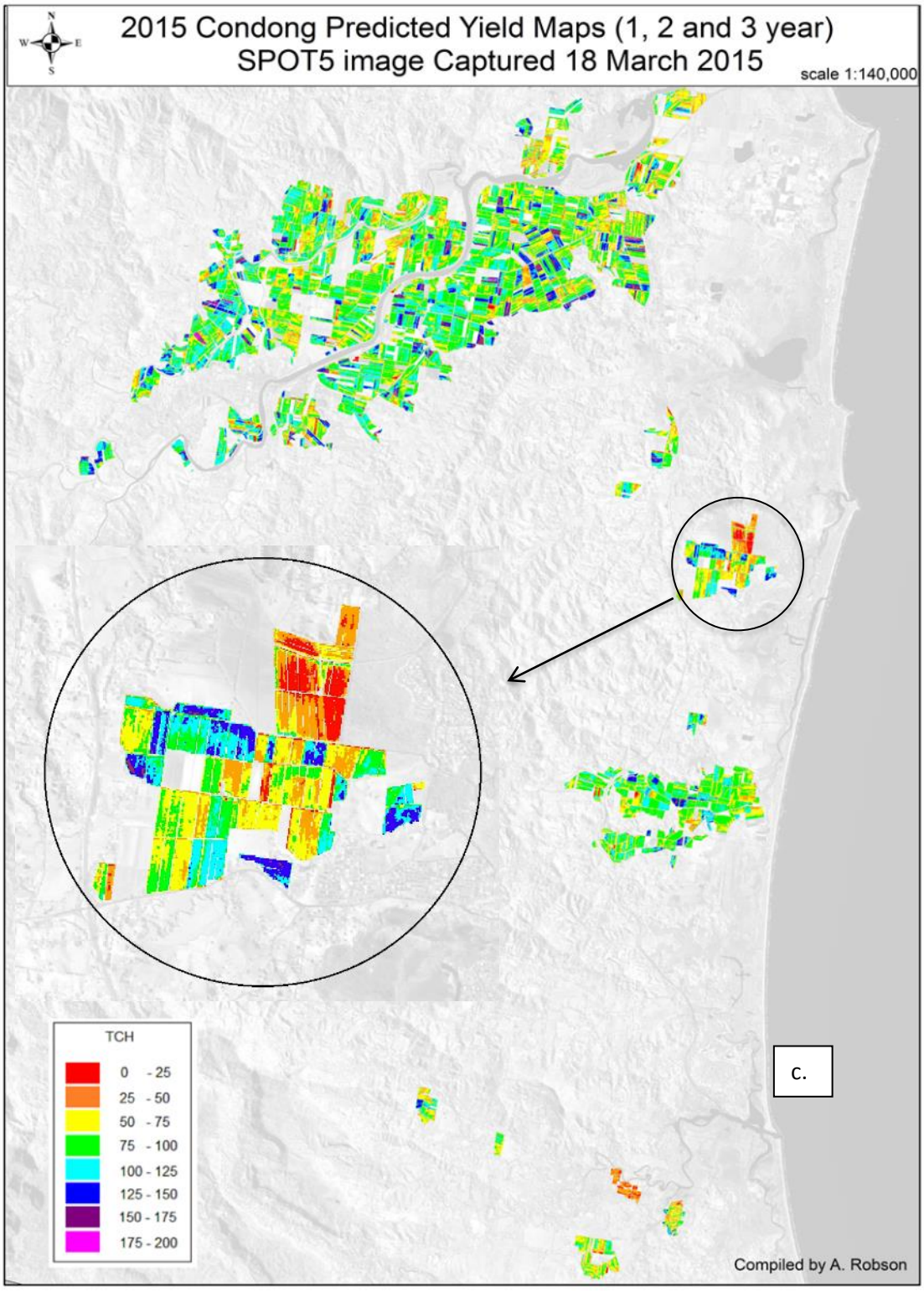


Figure 29 a. Stretched false colour image of all sugarcane crops grown in the Condong growing region in 2015. The brighter the red colour the more vigorous the growth. b. Classified GNDVI map of the same cane crops, with a higher GNDVI value indicating more vigorous growth. c. Classified derived yield from the Condong year 1 and year 2/3 algorithms. The magnified section clearly shows the within crop resolution, in this instance a region of poor performing cane, provided by 10 m SPOT5 imagery.

Regional Vs block level mapping

As seen in Figure 29 a, the derivation of regional yield maps from satellite imagery is extremely beneficial for identifying the spatial variation of yield within the growing season. Millers have used the information to assist with post-harvest decision making i.e. indication of tonnage, harvest scheduling, handling, forward selling etc. Additionally, the identification of sub- regions that are expressing poor growth, such as that identified in the magnified circle, has enabled productivity officers to conduct onsite inspections to identify the constraint. Without this derived map layer, the only indication of a low performing area would come post-harvest, therefore too late to conduct an in-field inspection. In the event that the constraints are the result of a new disease or pest out break this method of early detection is essential.

A limitation of the classification methodology currently employed by project DPI025 is that the entire region is separated into 8 classes (of crop vigour based on GNDVI) by the classification. Whilst the result is still extremely useful for identifying regional and sub- regional trends (Figure 29 b), there is a loss of information at the farm and block level. The following example (Figure 30) demonstrates this limitation. From Figure 30 a, most blocks are characterized by only 1 or 2 GNDVI colour classes following a classification using the current practice. In comparison, Figure 30 b, identifies greater within block variation when each block is individually classified into the 8 classes. It is an understanding of this within block variability which is important to growers when considering improved management strategies such as the adoption of precision agriculture and variable rate technologies.

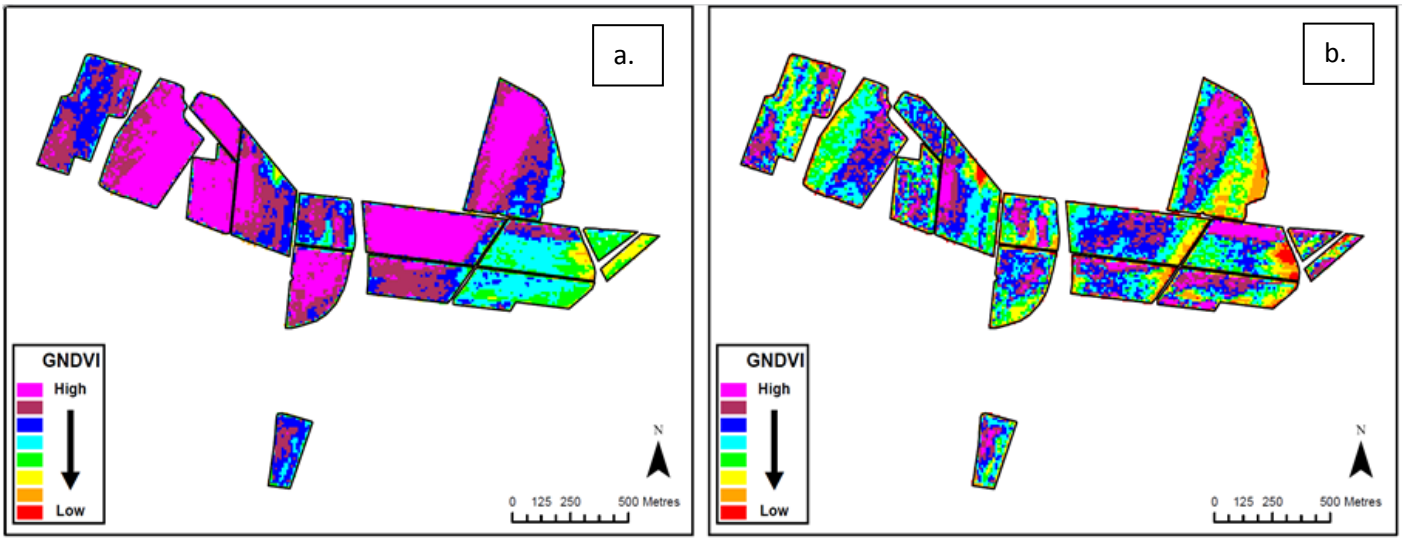
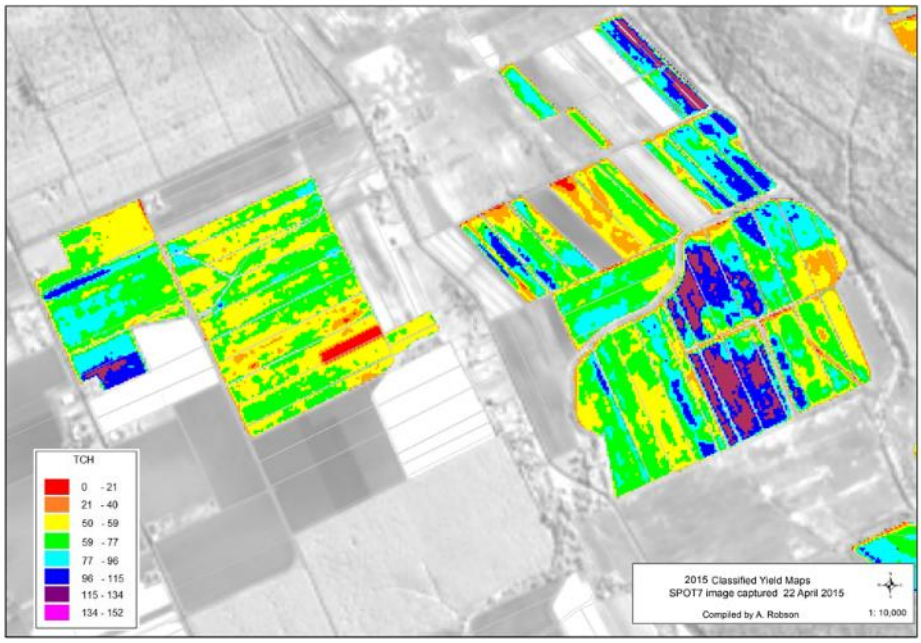
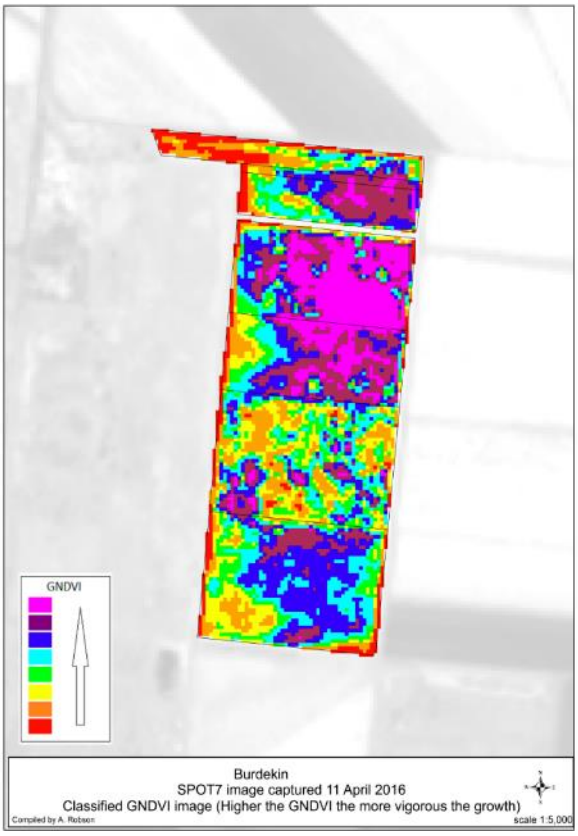
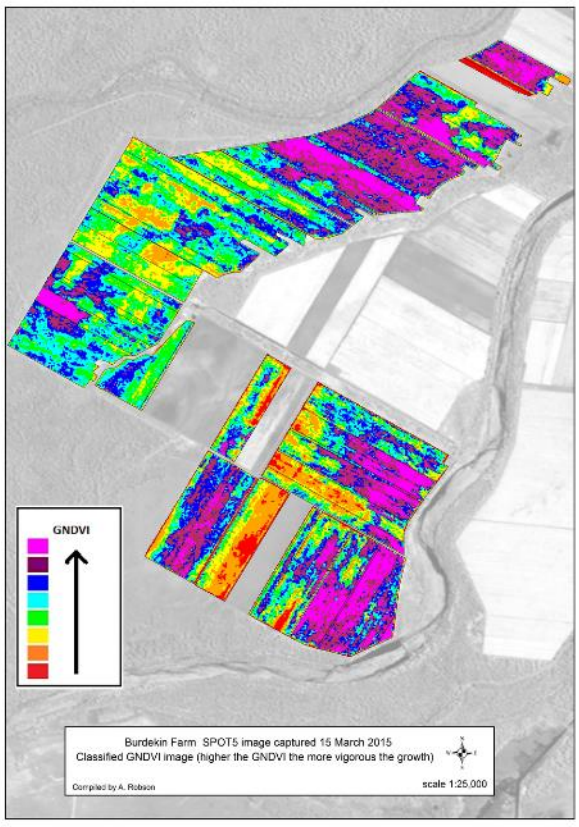


Figure 30: a). GNDVI classification of a sugarcane farm using the current approach i.e. 8 classes applied at the regional level b). GNDVI classification of each individual block. SPOT 5 image (15th March 2015).

Due to the large number of individual crops grown within each region, it is currently unfeasible to manually subset and classify each block separately. As such, a recommendation of DPI025 is to develop an automated process, i.e. through Python scripting, that will allow the regional wide classification to be derived as well as separate classifications for every individual block. To achieve this would be a significant advancement in the development and then adoption of remote sensing outputs for the Australian sugar industry. The following Figure 31, provides some examples of Remote Sensing provided by DPI025, where the classification of individual blocks was critical. These examples were specific requests from growers and consultants. The outputs have been used to identify both poor and high performing areas within a block, and can be used in combination with on ground assessment to determine if yield is constrained by abiotic or biotic factors. For example, higher yield may be associated with a particular sugarcane variety, while lower yield/vigour may indicate a soil or water constraint is affecting growth within a particular block.


TCH
0 - 21
21 - 40
50 - 59
59 - 77
77 - 96
96 - 115
115 - 134
134 - 152
2015 Classified Yield Maps
SPOT7 image captured 22 April 2015
Compiled by A. Robson
1: 10,000


GNDVI
Burdekin
SPOT7 image captured 11 April 2016
Classified GNDVI image (Higher the GNDVI the more vigorous the growth)
Compiled by A. Robson
scale 1:5,000


GNDVI
Burdekin Farm SPOT5 image captured 15 March 2015
Classified GNDVI image (higher the GNDVI the more vigorous the growth)
Compiled by A. Robson
scale 1:25,000

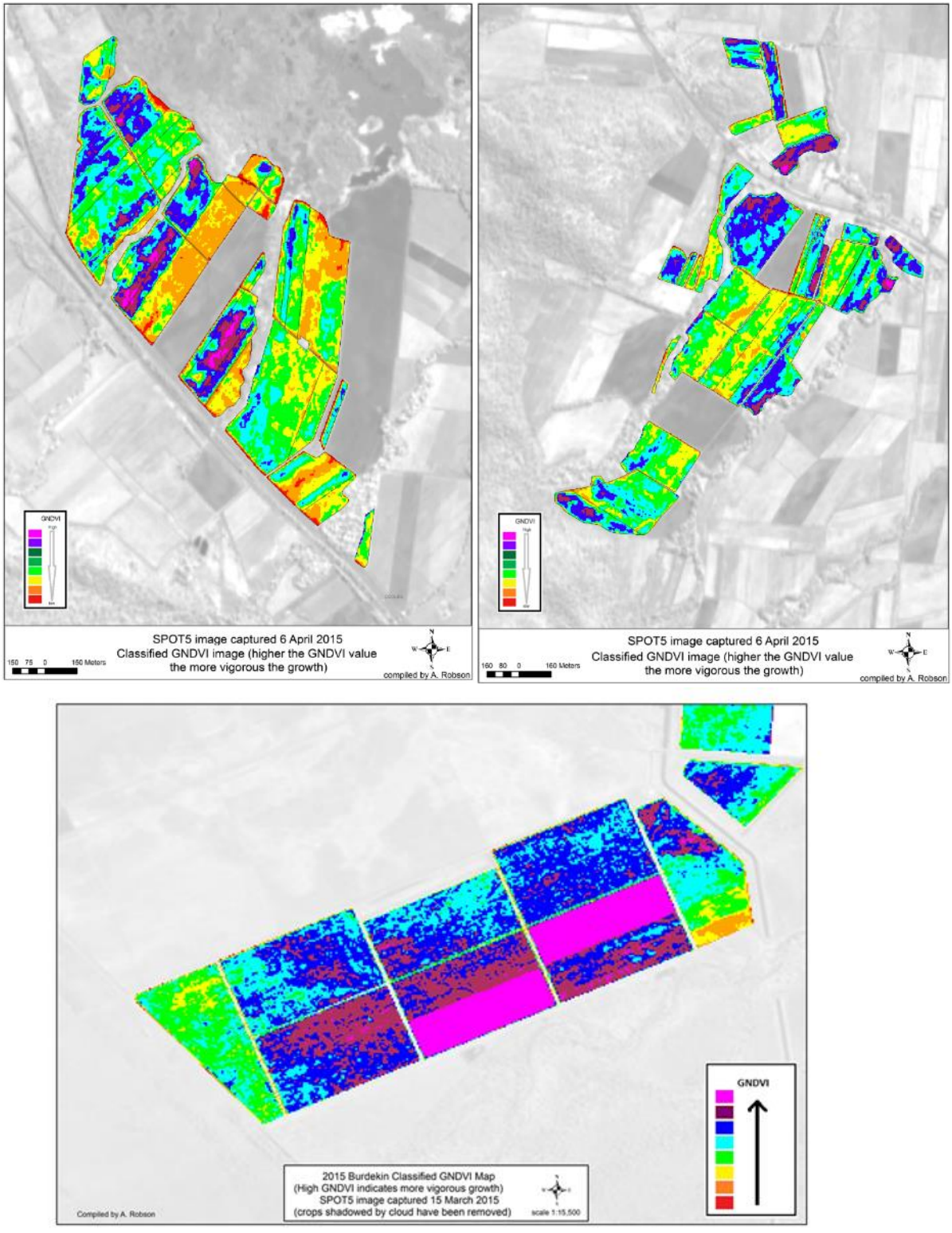


Figure 31. Examples of imagery products i.e. GNDVI (vigour) and derived yield provided to individual growers, consultants and researchers on request.

Corresponding with the range of applications investigated, was the file formats in which the imagery products were requested. Overall, the GeoTiff file format was identified to be the most compatible to the end user system, whether it be for Mill GIS software (ArcGIS, Mapinfo and AGdat), precision agriculture software platforms (i.e. John Deere and SST), or drafting software (Figure 32). However, ASCII grid files to support tractor guided systems, static pdfs and Google Earth KML files were also generated.

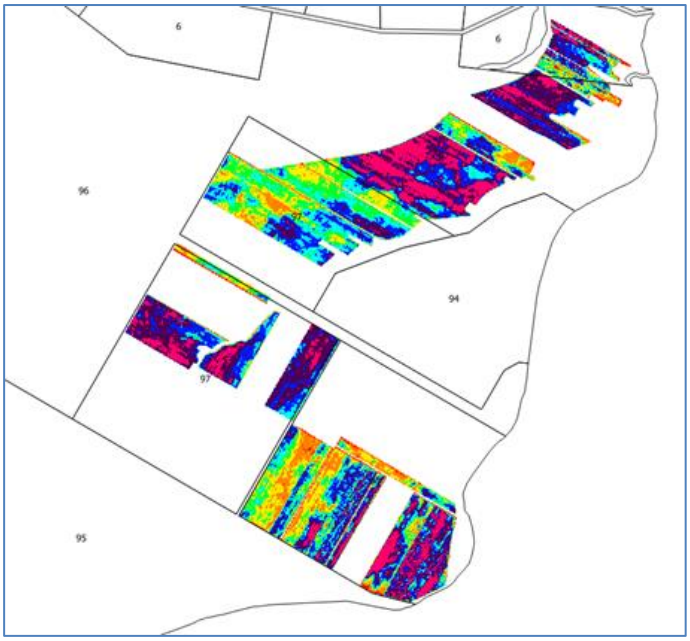
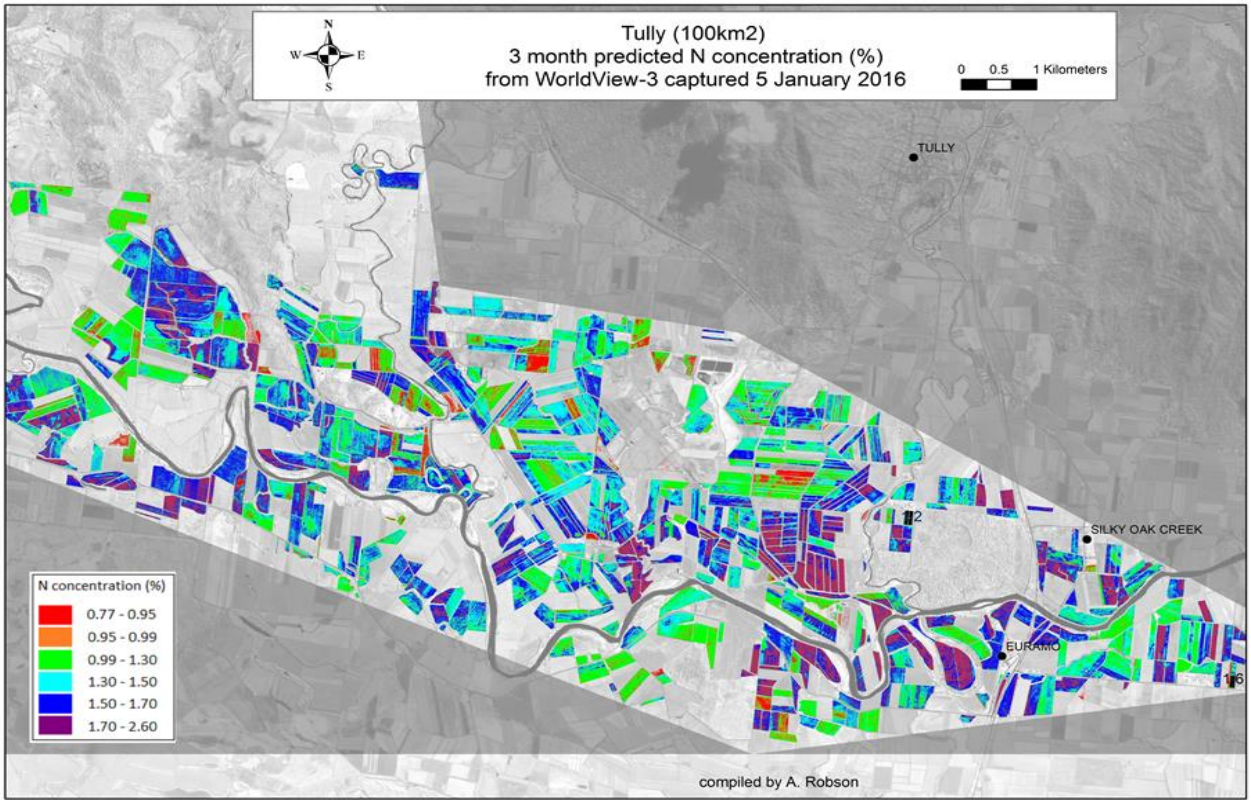


Figure 32: Geotiff classified GNDVI layer (8 colour classes) displayed in drafting software.

Qualitative foliar nitrogen maps.

As discussed in Objective 3, outputs of project DPI025 included regional 'qualitative' foliar Nitrogen maps. The term 'qualitative' is used as the relationship between measured N concentration and the vegetation index $REN2NDVI_{WV}$, was shown to be influenced by a range of parameters (Figure 24) and therefore the derivation of an absolute value is unlikely without the collection of in-field calibration measures. As well as the Mackay map presented in Figure 28, similar maps were derived for the Burdekin and Tully regions (Figure 33 a and b). As with the derived yield maps, a qualitative split into low – high classes applied at the individual crop level, as opposed to across all blocks within the region, would provide additional information on foliar N variation at the individual block level.



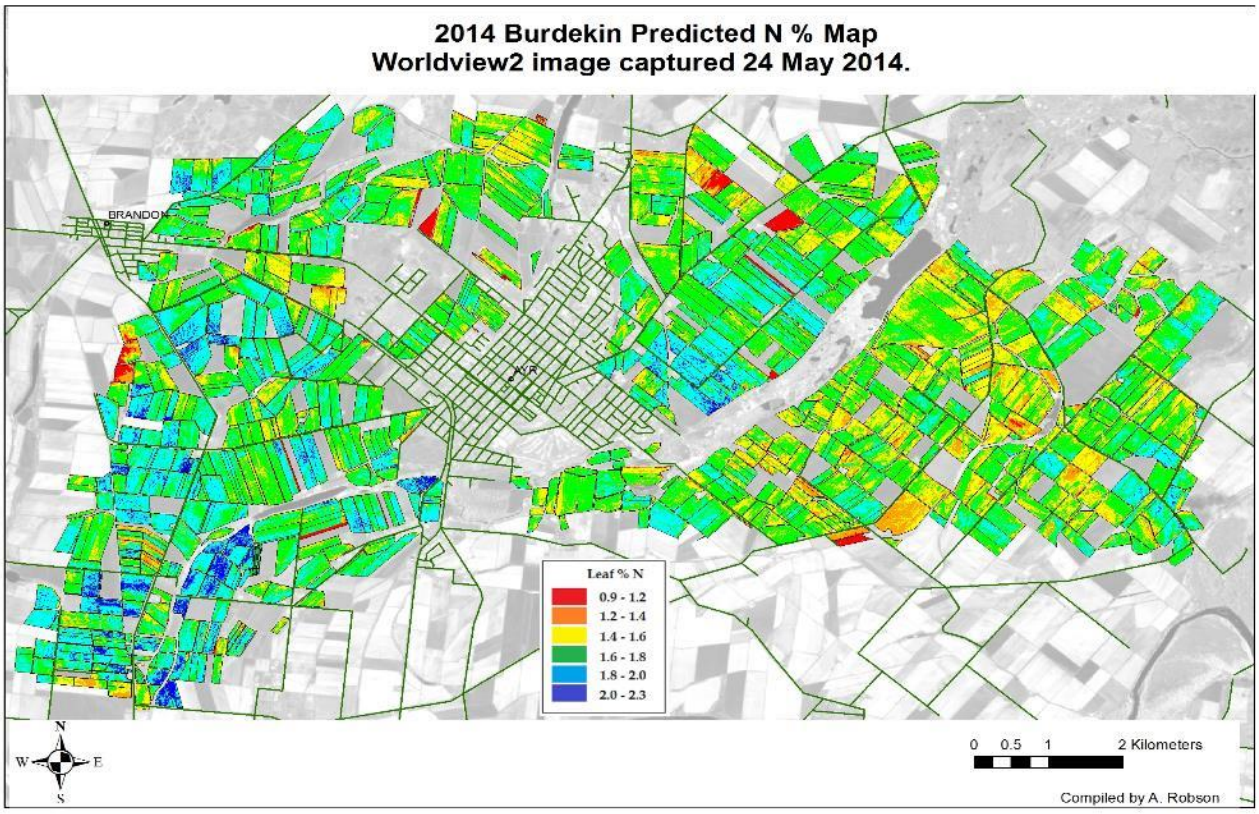


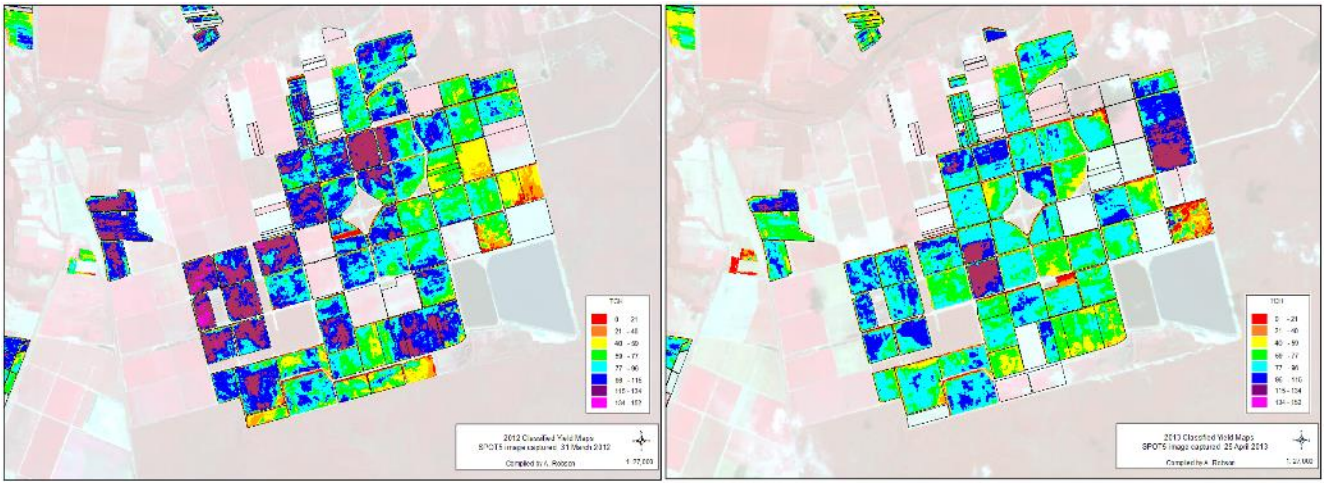
Figure 33. a. Classified % N foliar concentration map derived for all crops within the 100km^2 image capture area for the a. Tully and b. Burdekin growing regions.

The generation of these maps, even in its 'qualitative' form, provides useful information regarding the spatial trends of sub-region foliar N concentration. For example, the Mackay map indicates higher N concentrations near the township of Walkerston (Figure 28 b), the Tully map (Figure 33 a) indicates N concentrations are higher in crops along the varying water courses, whilst for the Burdekin map concentrations are higher to the

west of the township of Ayr. When compared with additional spatial layers such as soil type, topography and derived yield maps, this information can be extremely useful for improving N application efficiencies in the future. Similarly an understanding of the spatial variability can assist growers with targeted soil and tissue N sampling that again ultimately supports the in-crop application of N management strategies such as 6 easy steps. A similar process has been adopted by the rice industry for the application of N at panicle initiation (Dunn 2012).

Establishment of an image library to assist in the understanding of temporal change.

Through DPI021 and now DPI025, a database of raw imagery as well as derived imagery products has been accumulated, that in some regions extends up to 7 years of data. This resource is extremely useful for understanding both the temporal and spatial variability of cop production at the regional, farm and within crop level. Figure 32, provides an example of derived yield maps generated for a Bundaberg farm over a four year period (2012- 2015), with the full region supplied in Appendix 5. There is a clear indication of what regions of the farm are consistently underperforming, and what variation may be occurring as result of changes in seasonal weather conditions. When overlaid with Mill data, rapid assertions can be made on what cultivars are best suited for each sub-region or conversely what crop rotations are producing the most vigorous response in the following plant cane. In terms of nutrition management, understanding inherently high and low performing crop regions can indicate likely response to inputs, therefore guiding targeted applications such as Nitrogen.



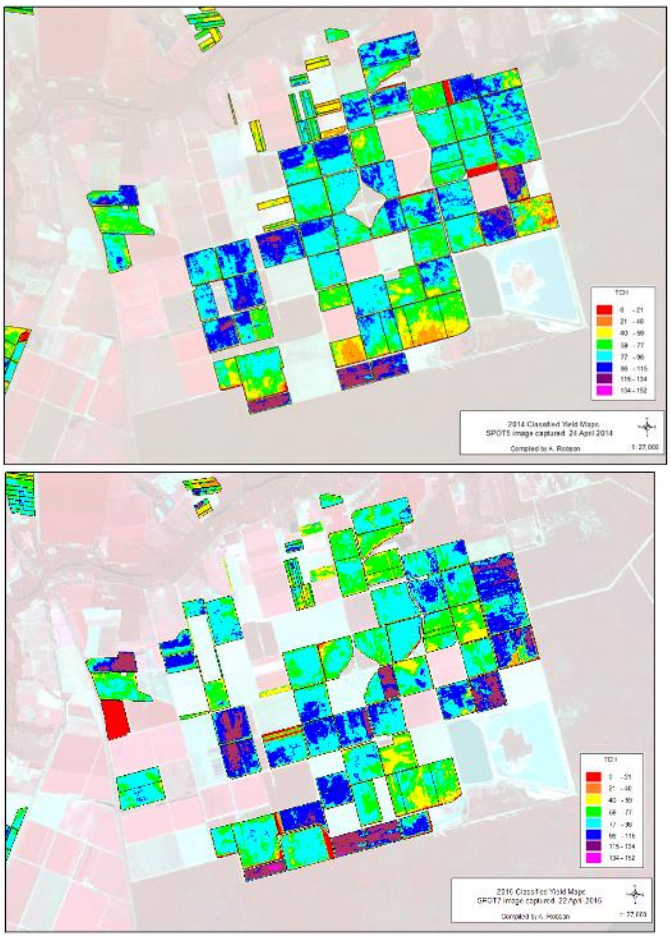


Figure 34. Temporal variability in derived yield for a Bundaberg farm between 2012- 2015.

Due to existing confidentiality agreements with the mills, the output and use of the maps and data generated by project DPI025 has been somewhat limited. It is envisaged that through automation and improved data access, the temporal image databases will become more accessible to the broader industry through established Mill portals. In order to facilitate such uptake, it is necessary to engage with all levels of the industry, to ensure they are receiving this information in a timely manner to facilitate decision making – while ensuring confidentiality is maintained.

Section 5: Intellectual Property (IP) and Confidentiality

N/A

Section 6: Industry Communication and Adoption of Outputs

DPI021 successfully introduced remote sensing technologies to the wider Australian sugarcane industry and research community. Findings from the project have continued to develop and evolve not only within the current project but also in other funded projects including projects '2015/038: using remote sensing to improve cane grub management in North Queensland cane fields', 'PL006: Licence to Farm: Nitrogen use efficient varieties to meet the future environmental targets' and 'Preliminary investigating into the effectiveness of remote sensing and GIS for identifying Yellow Canopy Syndrome (YCS) in Sugarcane'. Additionally, remote sensing outputs have been included in projects BPS001, CSE022, and as a chapter in a SRA funded Nitrogen review (http://elibrary.sugarresearch.com.au/handle/11079/14735)

Publications, newsletters, fact sheets and other media coverage.

The results of project DPI025 have been communicated widely through field days, ASSCT conferences, relevant industry and technical forums and publications:

Presentations:

- 2013 (27 -30 August): Presented at Asia- Pacific Economic Cooperation (APEC): Training Course on the application of Remote Sensing and GIS Technology in Crop Production. Beijing, China.
- 2013 (26- 28 June): Presentation at Digital Rural Futures Conference. University of New England. Armidale. NSW. http://www.une.edu.au/about-une/academic-schools/school-of-science-and-technology/news-and-events/events/digital-rural-futures-conference/presentations#! Theme 3, talk 4.
- 2013 (13 June): Guest speaker at Precision Farming Field Day. Bundaberg, Queensland.
- 2014 (1 April): Invited speaker at 2014 Herbert Walk and Talk field Day.
- 2014 (5 May): Presented results and provided an update of current project direction to millers at the annual AGDAT conference.
- 2015 (2-6 Feb). Presentation at 2015 IEEE RAS summer school on Agricultural Robotics (http://www.acfr.usyd.edu.au/education/ssar2015.shtml).
- 2015 (11- 12 May) AGDAT annual millers meeting, Coolangatta.
- 2015 (14 May). SRA Precision Agriculture workshop.

Media:

- 2014 (4 April): SRA CaneClip “Connecting with Growers”.
- 2014 (11 April): SRA CaneClip “The benefits of Remote Sensing for Growers and Millers.
- 2014 (5 August): Article in The Australian Newspaper ‘Precision system finds the sweet spot on sugar yields’.
- 2015 (January): Farmacist post on Facebook. (Appendix 6).
- 2015 (May): Geoimage Newsletter http://www.geoimage.com.au/Newsletter/May2015Newsletter#farewell
- 2015 (15 Oct): SSSi webinar. (https://attendee.gotowebinar.com/recording/8329034046758099713).
- 2016 (May): Sunshine sugar newsletter: Around the Paddocks (Appendix 6).

Publications (further detail provided in Section 9):

- Robson, A., Johansen, K., Schmidt, S., Robinson, N., Lakshmanan, P., Connellan, J., Everingham, Y. (2014). Assessing remote sensing and GIS for the prediction of crop nitrogen status in sugarcane as well as the influences of seasonal weather variability on image based yield forecasting. In: *ASSCT 2014: Proceedings of*

the 36th Annual Conference of the Australian Society of Sugar Cane Technologists, 28 April – 1 May, 2014, Broadbeach, Queensland, Australia.

- Chapter written for SRA/ EHP review: *'Investigating remote sensing technologies and their potential for improving nitrogen management in Sugarcane'* Andrew Robson, John Stanley and David Lamb.

- Andrew Robson (2014). Remote Sensing. Precision Ag News. SPAA Precision Agriculture Australia. Vol 11(1), Spring 2014.

- Everingham, Y., Sexton, J., and Robson, A. (2015). A statistical approach for identifying important climatic influences on sugarcane yields. In: *ASSCT 2015: Proceedings of the 37th Annual Conference of the Australian Society of Sugar Cane Technologists*, pp. 8-15, 28-30 April 2015, Bundaberg, QLD, Australia.

- Andrew Robson, Muhammad Moshiur Rahman, Gregory Falzon, Niva Kiran Verma, Kasper Johansen, Nicole Robinson, Prakash Lakshmanan, Barry Salter and Danielle Skocaj (2016). Evaluating Remote Sensing technologies for improved Yield Forecasting and for the Measurement of foliar nitrogen concentration in Sugarcane. Proceedings of the 38th Annual ASSCT Conference, Mackay Entertainment & Convention Centre, 2016.
  (note: this paper was awarded the Presidents medal for best paper).

- Rahman, M. M., Robson, A. J. (2016). A Novel Approach for Sugarcane Yield Prediction Using Landsat Time Series Imagery: A Case Study on Bundaberg Region. *Advances in Remote Sensing*, *5*, 93–102.

- Rahman, M. M. and Robson, A. J. (2016). Multi-temporal remote sensing for yield prediction in sugarcane crops. In: *19th Society of Precision Agriculture Australia Symposium Proceedings*, 12-13th September 2016, Toowoomba, QLD, Australia.

- Andrew Robson, Muhammad Moshiur Rahman, Gregory Falzon, Niva Kiran Verma, Kasper Johansen, Nicole Robinson, Prakash Lakshmanan, Barry Salter and Danielle Skocaj (2016). Evaluating Remote Sensing technologies for improved Yield Forecasting and for the Measurement of foliar nitrogen concentration in Sugarcane. December issue ISSCT. 2016.

Other:

- Letter from local member of the Northern Tablelands Adam Marshall, congratulating the UNE project team for their president's medal at the 2016 ASSCT conference (Appendix 6).

b) Identify any further opportunities to disseminate and promote project outputs at seminars, field days etc.

The results from multi temporal analysis of Landsat imagery for yield prediction of sugarcane crops on block will be presented at the 11th European Conference on Precision Agriculture (ECPA) in Edinburgh on 16 – 20th July, 2017.

Section 7: Environmental Impact

N/A

Section 8: Recommendations and Future Industry Needs

As discussed in the outcomes and output section of this report *'Regional Vs block level mapping'*, a major limitation to the current processing methodology is the ability to generate classified maps for individual crops on mass, as opposed to a classification applied to all crops in a region. The loss of spatial information (refer to Figure 30), greatly reduces a growers' ability to understand spatial variability within individual crops and therefore implement improved management strategies i.e. variable rate technologies. As such, there is a great need to develop an automated processing methodology i.e. python scripting, that will allow for each crop to be classified separately. By using grower identifiers i.e. Linkcodes, already established within existing mill data, maps for individual growers could be generated and disseminated on mass, through the respective mills, with privacy protection ensured. It would follow that the automation of this process would also include an improved method of data dissemination to all end users including growers, researchers, productivity services etc, a point that was identified by a 2015 review of precision agricultural conducted by SRA (Iain Yule and Allan Garside).

The development of the multi-temporal Landsat yield prediction algorithm for the Bundaberg region was a significant advancement for not only improved regional yield forecasting, but also for better understanding the annual variation in regional crop growth. In order for the entire Australian sugar industry to take advantage of this new methodology, it is necessary that models similar to the one developed for the Bundaberg region, be developed for each growing region. Additionally, the derivation of the time series algorithm created the opportunity to make regional yield predictions earlier in the growing season i.e. from December. The recent launch of the Sentinel 2 satellite and the subsequent availability of free 10 m resolution (currently 20 day repeat time – although this will decrease to 5 days with the addition of improved data downloading capabilities, and the launch of the next satellite in the constellation), provides the opportunity to extrapolate the Landsat time series models to Sentinel. However, this would have to be validated.

The identification of REN2NDVI as being consistently well correlated to foliar N concentration was also a major outcome of this research. To further build on this result, there is an essential need to establish practical adoption pathways that will assist growers to implement this knowledge within their farming system, either through proximal sensors or from derived imagery maps. Following the methods presented by Johnson and Raun *et al.* 2003 and Lofton *et al.* 2012, limited and non- limited nitrogen reference strips could be used to calibrate commercially available active or passive sensors such as the Crop Circle or possibly the Yara N- Sensor. If validated, this approach would enable growers with tractor based sensors to develop a rapid calibration using the reference strips and deliver variable rate 'on-the-go' nitrogen rates. This approach is currently adopted within cereal, wheat and corn crops (Reiter *et al.* 2014; Raun *et al.* 2010). Additionally, the implementation of the satellite derived 'qualitative' N concentration maps into a precision farming system would greatly compliment additional spatial layers including harvester yield mapping (where available) soil and topography maps. When combined, these spatial layers could define productivity zones that through targeted ground truthing can be used to direct applications of crop inputs, particularly nitrogen.

Section 9: Publications

A total of 8 peer reviewed publications (1 Technical report, 1 journal article and 6 conference papers) were produced in this project and one conference proceedings will be delivered in the 11th European Conference of Precision Agriculture (ECPA) next year in Edinburgh.

1) Rahman, M. M., Robson, A. J. (2016). A Novel Approach for Sugarcane Yield Prediction Using Landsat Time Series Imagery: A Case Study on Bundaberg Region. *Advances in Remote Sensing*, *5*, 93–102. **(Journal Paper)**

Abstract
Quantifying sugarcane production is critical for a wide range of applications, including crop management and decision making processes such as harvesting, storage, and forward selling. This study explored a novel model for predicting sugarcane yield in Bundaberg region from time series Landsat data. From the freely available Landsat archive, 98 cloud free (<40%) Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) images, acquired between November 15th to July 31st (2001-2015) were sourced for this study. The images were masked using the field boundary layer vector files of each year and the GNDVI was calculated. An analysis of average green normalized difference vegetation index (GNDVI) values from all sugarcane crops grown within the Bundaberg region over the 15-year period identified the beginning of April as the peak growth stage and, therefore, the optimum time for satellite image based yield forecasting. As the GNDVI is an indicator of crop vigour, the model derived maximum GNDVI was regressed against historical sugarcane yield data, which showed a significant correlation with R^2 = 0.69 and RMSE = 4.2 t/ha. Results showed that the model derived maximum GNDVI from Landsat imagery would be a feasible and a modest technique to predict sugarcane yield in Bundaberg region.

2) Rahman, M. M. and Robson, A. J. (2016). Multi-temporal remote sensing for yield prediction in sugarcane crops. In: *19th Society of Precision Agriculture Australia Symposium Proceedings*, 12-13th September 2016, Toowoomba, QLD, Australia. **(Conference Proceedings)**

Abstract
Sugarcane yield prediction is critical for in season crop management and decision making processes such as harvest scheduling, storage and milling, and forward selling. This presentation reports on a recently published method of predicting sugarcane yield in the Bundaberg region (Qld) using time series Landsat data. From the freely available Landsat archive, 98 cloud free (<40%) Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) images, acquired between November 15th to July 31st (2001-2015), were sourced for this study. The images were masked using the field boundary layer vector files of each year and the green normalized difference vegetation index (GNDVI), an indicator of crop vigour was calculated. An analysis of average GNDVI values from all sugarcane crops grown within the Bundaberg region over the 15-year period using a quadratic model identified the beginning of April as the peak growth stage and, therefore, the decisive time of image capture for a single satellite image based yield forecasting. The model derived maximum GNDVI was regressed against historical sugarcane yield data obtained from the mill. The coefficient of determination showed a significant relation between the predicted and actual sugarcane yield (t/ha) with R^2 = 0.69 and RMSE 4.2 t/ha. Results showed that the model derived maximum GNDVI from Landsat imagery would be a feasible technique to predict sugarcane yield in Bundaberg region. This research, however, warrants further investigation to improve and develop accurate operational sugarcane yield prediction model across other domestic and global growing regions, as the influence of environmental conditions and cropping practices will likely vary the relationship between GNDVI and yield (t/ha).

3) Robson, A., Rahman, M.M., Falzon, G., Verma, N.K., Ohansen, K.J., Robinson, N., Lakshmanan, P., Salter, B. and Skocaj, D. (2016). Evaluating Remote Sensing Technologies for Improved Yield Forecasting and for the Measurement of Foliar Nitrogen Concentration in Sugarcane. In: *ASSCT 2016: Proceedings of the 38th*

Annual Conference of the Australian Society of Sugar Cane Technologists, 27-29th April, 2016, Mackay, Qld Australia. **(Conference Proceedings) (Also published in December Issue of ISSCT)**

Abstract

An analysis of time series Landsat imagery acquired over the Bundaberg region between 2010 and 2015 identified variations in annual crop vigour trends, as determined by greenness normalised difference vegetation index (GNDVI). On average, early to mid-April was identified as the crucial period where crops achieved their maximum vigour and as such indicated when single image captures should be acquired for future regional yield forecasting. Additionally, the regional crop GNDVI averaged from Landsat images between February to April, produced a higher coefficient of determination to final yield (R^2 = 0.91) than the average crop GNDVI value from a single mid-season SPOT5 image capture (R^2 = 0.52). This result indicates that the time series method may be more appropriate for future regional yield forecasting. For improved prediction accuracies at the individual crop level, a univariate model using only crop GNDVI values (SPOT5) and corresponding yield (t/ha) produced a higher prediction accuracy for the 2014 Bundaberg harvest than a multivariate model that included additional historic spectral and crop attribute data. For Condong, a multivariate model improved the prediction accuracy of individual crops harvested in 2014 by 41.8% for one-year-old cane (Y1), and 46.2% for two-year-old cane (Y2). For the non-invasive measure of foliar nitrogen (N%), the specific wavelengths 615 nm, 737 nm and 933 nm (Airborne hyperspectral), and 634 nm, 750 nm and 880 nm (ground based field spectroscopy) were found to be the most significant. These results were supported by satellite imagery (Worldview-2 and Worldview-3) acquired over three replicated field trials in Mackay (2014 and 2015) and Tully (2015), where the vegetation index (VI) REN2NDVIWV, a ratio of the red-edge band (705–745 nm) and the Near–IR2 band (860–1040 nm), produced a higher correlation to nitrogen concentration (%) than NDVI.

4) Robson, A., Stanley, J., and Lamb, D. (2015). Investigating remote sensing technologies and their potential for improving nitrogen management in Sugarcane. In: *A Review of Nitrogen Use Efficiency in Sugarcane*. Sugar Research Australia Research Report, pp. 154 -179. **(Report)**

Abstract

This paper investigates remote sensing as a technology for accurately determining leaf nitrogen content in agricultural crops, thus determining its potential for supporting improved nitrogen management in the Australian sugar industry. A review of commercially available active and passive, proximal and airborne, multispectral and hyperspectral sensors as well as relevant research on various cropping systems identified a number of benefits and limitations to the varying technologies.

On the basis of these out comes four strategies for the possible application of remote sensing to measuring foliar Nitrogen concentration in sugar are proposed. These include:

- Identifying specific spectral wavelengths measured from a growing sugar plant canopy that directly correlates to leaf Nitrogen concentration. These wavelengths could form the basis of specific vegetation indices to be derived from imagery collected from airborne sensors i.e. Worldview 3 satellite or alternatively used to develop specific active sensors to be mounted on ground based vehicles or manually deployed;
- Implementing a 'fertiliser response index' approach where limited and non- limited nitrogen reference strips are used to calibrate vegetation indices such as NDVI, measured by a range of commercially available sensors (active and passive); These calibrations when applied across the remaining crop can indicate the spatial variability of leaf N concentration, thus supporting targeted applications of N.

- Integrating imagery derived crop vigour maps with additional spatial layers such as harvester yield maps, soil surveys, topography etc. to identify inherent crop 'productivity' zones that exist both temporally and spatially. This information can direct strategic plant sampling to establish if a relationship exists between the productivity zones and leaf nitrogen content. If such a relationship exists, then again an opportunity exists for the strategic application of N.
- Identifying if the nitrogen content of harvested material measured by NIR at the mills can be used to calibrate the extracted spectral information from imagery captured over the corresponding crops. If such a correlation can be identified, then the opportunity exists to extrapolate that relationship across all crops within the growing region, therefore deriving crop, farm and regional N maps.

Additionally, Remote sensing offers a number of indirect benefits for improving the monitoring and management of foliar N such as 'value adding' crop modelling by offering a within season measure of crop performance; as a non-destructive screening tool for assessing the nitrogen use efficiency (NUE) of cultivars in breeding trials; identifying sub regional locations better suited to high NUE cane varieties; as a non- invasive tool for measuring the response and effectiveness of various nitrogen application strategies i.e. mill mud, six easy steps etc.

5) Everingham, Y., Sexton, J., and Robson, A. (2015). A statistical approach for identifying important climatic influences on sugarcane yields. In: *ASSCT 2015: Proceedings of the 37th Annual Conference of the Australian Society of Sugar Cane Technologists*, pp. 8-15, 28-30 April 2015, Bundaberg, QLD, Australia. **(Conference Proceedings)**

Abstract

Inter-annual climate variability impacts sugarcane yields. Local climate data such as daily rainfall, temperature and radiation were used to describe yields collected from three locations–Victoria sugar mill (1951–1999), Bundaberg averaged across all mills (1951–2010) and Condong sugar mill (1965–2013). Three regression methods, which have their own inbuilt variable selection process were investigated. These methods were (i) stepwise regression, (ii) regression trees and (iii) random forests. Although there was evidence of overlap, the variables that were considered most important for explaining yields by the stepwise regressions were not always consistent with the variables considered most important by the regression trees. The stepwise regression models for Bundaberg and Condong delivered a model that was difficult to explain biophysically, whereas the regression trees offered a much more intuitive and simpler model that explained similar levels of variation in yields to the stepwise regression method. The random forest approach, which extends on the regression tree algorithm generated a variable importance list which overcomes model sensitivities caused by sampling variability, thereby making it easier to identify important variables that explain yield. The variable importance list for Victoria indicated that maximum temperature (February–April), radiation (January–March) and rainfall (July–October) were important predictors for explaining yields. For Bundaberg, emphasis clearly centred on rainfall, particularly for the period January to April. Interestingly, the random forest model did not rate rainfall highly as a predictor for Condong. Here the model favoured radiation (February to April), minimum temperature (March–April) and maximum temperature (January to April). Improved understanding of influential climate variables will help improve regional yield forecasts and decisions that rely on accurate and timely yield forecasts.

6) K. Johansen, A. Robson, P. Samson, N. Sallam, K. Chandler, A. Eaton, L. Derby and J. Jennings (2014). Mapping Canegrub Damage from High Spatial Resolution Satellite Imagery. In: *ASSCT 2014: Proceedings of*

the 36th Australian Society of Sugar Cane Technologists, 29 April – 1 May 2014, Gold Coast, QLD, Australia. **(Conference Proceedings)**

Abstract

Canegrubs feed on the roots of sugarcane plants, reducing plant vigour and yield, and if left untreated they have the potential to rapidly increase the impacted area in the following year. For the targeted control of the canegrub, it is essential that the location of the affected areas is identified. However, identifying canegrub damage in the field is difficult due to the often impenetrable nature of sugarcane. The objective of this research was to use geographic object-based image analysis (GEOBIA) and high spatial resolution satellite imagery to map canegrub damage. The GEOBIA mapping approach used in this research was based on the following key steps for three selected study sites in Queensland, Australia: (1) initial segmentation of sugarcane block boundaries and further segmentation of each block into smaller homogenous objects; (2) classification and subsequent omission of fallow/harvested fields, tracks and other non-sugarcane features; (3) identification of 'potentially' grub-damaged areas within each block based on low NDVI and high image texture values; and (4) identification of 'likely' grub affected areas based on the absolute difference in NDVI and texture values between the 'potentially' grub damaged areas and the remaining parts of each block. Overall accuracies were between 53-79%. Further research will focus on improving these mapping accuracies. The results of this research will help cane growers to manage and reduce damage caused by canegrubs and increase future yields.

7) Zellner, P., Lelong, C., Soti, V., Tran, A., Sallam, N., Robson, A., Goebel, F. (2014). A remote sensing and GIS approach to the relationship between canegrub infestations and natural vegetation in the sugarcane landscape of Queensland, Australia, Meeting on Landscape Management for Functional Biodiversity, 6, Poznan, Pologne, 21-23rd May 2014. *IOBC/WPRS Bulletin*, 100: pp. 153-158. **(Conference Proceedings)**

Abstract

The greyback canegrub, Dermolepida albohirtum, remains a major problem in sugarcane areas in the north Queensland, Australia, despite years of control efforts. The adult beetle spends most of its life span swarming and feeding on trees along river banks, rainforest edges and scrubs or even in cane farms, where they aggregate, mate and then return to sugarcane fields to lay eggs. Remote sensing methodologies and Geographic Information System (GIS) analyses were developed to better understand the relationship between natural vegetation bordering the sugarcane fields and damage distribution. Very high spatial resolution images were processed through an object-based classification to map the sugarcane fields and different types of surrounding vegetation, which harbor the cane beetles feeding and mating trees. Preliminary results indicate that the mean distance to the closest vegetation patch is 35 m and does not exceed 154 m and therefore suggest that the trees bordering the sugarcane fields play a major role in damage occurrence.

8) A. J. Robson and G. C. Wright (2013). Accurate Regional to Field Scale Yield forecasting of Australian Sugarcane and Peanut Crops using Remote Sensing and GIS. *Asia- Pacific Economic Cooperation (APEC): Training Course on the application of Remote Sensing and GIS Technology in Crop Production*. Beijing, China. 27- 30 August 2013. **(Conference Proceedings)**

Abstract

The following paper demonstrates the accuracy of satellite imagery for the prediction of average regional yield for both Sugar cane and Peanut. The development of non- cultivar and non- class specific algorithms that are relatively insensitive to seasonal and location variability offer a useful tool for validating current forecasting

methods. The further development of analysis protocols for the rapid derivation and distribution of surrogate yield maps, on mass, offer many applications to their respective industries. These include improved monitoring capabilities to quantify lost productivity resulting from within season incidences of pest and disease outbreaks, as well as unseasonable weather events.

References

Abdel-Rahman, E. M., & Ahmed, F. B. (2008). The application of remote sensing techniques to sugarcane (Saccharum spp. hybrid) production: a review of the literature. *International Journal of Remote Sensing, 29*(13), 3753-3767. doi:10.1080/01431160701874603.

Bagheri N, Ahmadi H, Alavipanah S K, Omid M (2013) Multiplspectral remote sensing for site-specific nitrogen fertilizer management. Pesquisa Agropecuária Brasileira. 48 (10): 1394-1401. DOI: 10.1590/S0100‑204X2013001000011.

Barker D W, Sawyer J E (2012) Using active canopy sensors to quantify corn nitrogen stress and nitrogen application rate. Agronomy Journal. 102 (3): 964-971.

Dunn, B., Dunn, T., Hume, I., Orchard, B., Dehaan., and Robson, A. (2016). Remote sensing PI nitrogen uptake in Rice. IREC Farmers' Newsletter No. 195- Rice R&D 2016.

Dunn B (2012) Nitrogen tissue test for rice panicle initiation. RIRDC publication No. 12/047.

Duveiller G, López-Lozano R, Baruth B (2013) Enhanced Processing of 1-km Spatial Resolution fAPAR Time Series for Sugarcane Yield Forecasting and Monitoring. Remote Sensing, 5, 1091-1116.

Erdle K, Mistele B, Schmidhalter (2011) Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. Field Crops Research. 124: 74–84.

Everingham, Y., Sexton, J., Skojac, D., Inman-Barber, G. (2016). Accurate prediction of sugarcane yield using a random forest algorithm. Agronomy for Sustainable Development, 36:27. doi:10.1007/s13593-016-0364-z.

Goffart J P, Olivier M (2004) Management of N fertilization of the potato crop using total N-advice software and in-season chlorophyll-meter measurements. In D. K. L. MacKerron & A. J. Haverkort (Eds.), Decision support systems in potato production: bringing models to practice (pp. 69-83). Wageningen, the Netherlands.

Great Barrier Reef (Australian government 2014; 3rd Reef Plan Report Card (State of Queensland 2013C).

Hansen P M and Schjoerring J K (2003) Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sensing of Environment. 86 (4):542-553.

Jackson R D, Jones C A, Uehara G, Santo L T (1980) Remote detection of nutrient and water deficiencies in sugarcane under variable cloudiness.

Jensen TA, Baillie C, Bramley RGV, Panitz JH (2012) An assessment of sugarcane yield monitoring concepts and techniques from commercial yield monitoring systems. Proceedings of the Australian Society of Sugar Cane Technologists 34.

Johnson G V, Raun W R (2003) Nitrogen response index as a guide to fertilizer management. Journal Plant Nutrition. 26:249–262. doi:10.1081/PLN-120017134.

Johnson A.K.L. (1995) Risk perceptions and nutrient management responses in the Australian sugar industry: preliminary results from the Herbert River District. Proceedings of the Australian Society of Sugar Cane Technologists, 17, 172–178.

Keating, B. A., Verburg, K., Huth, N. I., & Robertson, M. J. (1997). *Nitrogen Management in Intensive Agriculture: Sugarcane in Australia.* Paper presented at the Intensive Sugarcane Production: Meeting the Challenges Beyond 2000, CAB International, Wallingford, UK.

Koochekzadeh, A., Fathi, G., Gharineh, M. H., Siadat, S. A., Jafari, S., & Alami-Saeid, K. (2009). Impacts of Rate and Split Application of N Fertilizer on Sugarcane Quality. . *International Journal of Agricultural Research, 4*, 116-123.

Lofton J (2012) Improving nitrogen management in sugarcane production of the mid-south using remote sensing technology. PhD Thesis. School of Plant Environmental and Soil Sciences. Oklahoma State University.

Lofton J, Tubana B S, Kanke Y, Teboh J, Viator H (2012a) Predicting sugarcane response to nitrogen using a canopy reflectance-based response index value . Agronomy Journal. 104: 106-113.

Lofton J, Tubana B S, Kanke Y, Teboh J, Viator H, Dalen M (2012b) Estimating sugarcane yield potential using an in-season determination of normalized difference vegetative index. Sensors. 12: 7529-7547; doi:10.3390/s120607529.

Markley, J., Ashburner, B., and Beech, M. (2008). The development of a spatial recording system for productivity service providers. Proceedings of the Australian Society of Sugar Cane Technologists. 30: 10-16.

Morel J, Todoroff P, Bégué A, Bury A, Martiné J-F, Petit M (2014) Toward a Satellite-Based System of Sugarcane Yield Estimation and Forecasting in Smallholder Farming Conditions: A Case Study on Reunion Island. Remote Sensing, 6, 6620.

Mulianga B, Bégué A, Simoes M, Todoroff P (2013) Forecasting Regional Sugarcane Yield Based on Time Integral and Spatial Aggregation of MODIS NDVI. Remote Sensing, 5, 2184-2199.

Nigon T J, Mulla D J, Rosen CJ, Cohen Y, Alchanatis V, Rud R (2014) Evaluation of the nitrogen sufficiency index for use with high resolution, broadband aerial imagery in a commercial potato field. Precision Agriculture. 15: 202-226.

Patil V C, Nadagouda B T (2008) Precision nutrient management in sugarcane. Proceedings of the 9th International Conference on Precision Agriculture. July 20-23, Denver, Colorado USA.

Portz G, Molin J P Jasper J (2012) Active crop sensor to detect variability if nitrogen supply and biomass on sugarcane fields. Precision Agriculture. 13: 33-44.

Rahman, M. M., & Robson, A. J. (2016a). *Multi-temporal remote sensing for yield prediction in sugarcane crops.* Paper presented at the 19th Symposium of Society of Precision Agriculture Australia Towoomba, QLD, Australia.

Rahman, M. M., & Robson, A. J. (2016b). A Novel Approach for Sugarcane Yield Prediction Using Landsat Time Series Imagery: A Case Study on Bundaberg Region. *Advances in Remote Sensing, 5*, 93-102. doi:10.4236/ars.2016.52008.

Rao N R, Garg P K, Ghosh S K, Dadhwal V K (2008) Estimation of leaf total chlorophyll and nitrogen concentrations using hyperspectral satellite imagery. Journal of Agricultural Science. 146: 65–75. DOI:10.1017/S0021859607007514.

Raper T B, Varco J J, Hubbard K J (2013) Canopy- based normalized difference vegetation index sensors for monitoring cotton nitrogen status. Agronomy Journal. 105 (5): 1345-1354.

Raun WR, Solie J B, Johnson G V, Stone M L, Mullen R W, Freeman K W, Thomason W E, Lukina, E V (2002) Improving nitrogen use efficiency in cereal grain production with sensing and variable rate application. Agronomy Journal. 94:815–820. DOI:10.2134/agronj2002.0815.

Raun B, Solie J, May J, Zhang H, Kelly J, Taylor R, Arnall B, Ortiz-Monasterio I (2010) Nitrogen rich strips for wheat, corn and other crops. Oklahoma State University extension brochure E-1022.

Reiter M S, Mason J E, Thomason W E (2014) Sensor- Based, variable-rate nitrogen applications in Virginia. Virginia Cooperative extension. Virginia Tech. Virginia State University. Publication CSES-90P.

Report, R. (2013). *Reef Water Quality Protection Plan - Securing the health and resilience of the Great Barrier Reef World Heritage Area and adjacent catchments*. Retrieved from http://www.reefplan.qld.gov.au/resources/assets/reef-plan-2013.pdf.

Robson, A., Abbott, C., Bramley, R., & Lamb, D. (2013). *Remote sensing- based precision agriculture tools for the sugar industry*. Retrieved from http://elibrary.sugarresearch.com.au/handle/11079/12504

Robson, A., Stanley, J., & Lamb, D. (2014). *Investigating remote sensing technologies and their potential for improving nitrogen management in Sugarcane*. Retrieved from http://elibrary.sugarresearch.com.au/handle/11079/14735

Robson A, Abbott C, Lamb D, Bramley R (2012) Developing sugar cane yield algorithms from satellite imagery. Proceedings of the Australian Society of Sugar Cane Technologists. 34.

Rodriguez D, Fitzgerald G J, Belford R, Christensen L K (2006) Detection of nitrogen deficiency in wheat from spectral reflectance indices and basic crop eco-physiological concepts. Australian Journal of Agricultural Research. 57: 781-789.

Tahir M N, Li J, Liu B, Zhao G, Fuqi Y, Chengfeng C (2013) Hyperspectral estimation model for nitrogen contents of summer corn leaves under rainfed conditions. Pakistan Journal of Botany. 45 (5): 1623-1630.

Tilling A K, O'Leary G J, Ferwerda J G, JonesS D, Fitzgerald G J Rodriguez D, Belford R (2007) Remote sensing of nitrogen and water stress in wheat. Field Crop Research. 104: 77-85.

Wood A W, Schroeder B L, Stewart R L (2003) The development of site-specific nutrient management guidelines for sustainable sugarcane production. Proceedings of the Australian Society of Sugar Cane Technologists. 25, no-pagination.

Yao X, Yao X, Jia W, Tian Y, Ni J, Cao W, Zhu Y (2013) Comparison and Intercalibration of Vegetation Indices from Different Sensors for Monitoring Above-Ground Plant Nitrogen Uptake in Winter Wheat. Sensors (Basel). Mar 2013; 13(3): 3109–3130.

Zillmann E, Graeff S, Link J, Batchelor W D, Claupein W (2006) Assessment of cereal nitrogen requirement derived by optical on-the-go sensors on heterogeneous soils. Agronomy Journal. 98: 682-690.

Acknowledgements:

This research would not have been possible without the ongoing funding and support from Sugar Research Australia. As mentioned through the report, Project DPI025 received continual engagement and essential collaboration from a range of groups, including:

- Researchers:
  - Nicole Robinson (UQ),
  - Prakash Lakshmanan, Barry Salter, Danielle Skocaj and Julian Connellan (Sugar research Australia)
  - John Hughes (QDAF)

- Millers:
  - Ric Beattie, Johan Lambrechts, Ben Mayo, Simon Hollis (Sunshine Sugar),
  - Gavin Lerch and Allan Pitt (Bundaberg Sugar Ltd),
  - Sharon Dickson (Isis Central Sugar Mill Co Ltd),
  - Greg Wieden, Michael Sefton, Raymond De Lai, John Jardine and Dave Langham (Wilmar),
  - Peter Downs, Chris Coutts- Smith, Matt Hession, Yolande Kliese, Ann Stephensen, Michael Porta, Anna Sacchi, Glyn Peatey (Maryborough Sugar Factory).
  - Greg Shannon and Dale Thomas (Tully Sugar)

- Consultants:
  - Robert Crossley and Monica Poel (AGTrix)
  - Peter McDonnell, Jayson Dowie, Rob Sluggett (Farmacist)
  - Lawrence Di Bella (HCPSL)
  - Rob Milla (BPS)
  - Rob Dwyer (Incitec Pivot Fertilisers)
  - Steve Attard (AgriTech Solutions)
  - Bryan Granshaw (BMS Lasersat)
  - Dale Armbrust (Northern Consulting Engineers)
  - Don Pollock and Ross Coventry

DPI025 also had a number of 'cameo' appearances by a number of researchers whom provided vital contributions; these include David Lamb, Greg Falzon, Derek Schneider and Niva Verma (UNE), Yvette Everingham (JCU), Kasper Johansen (UQ), Rob Bramley (CSIRO), Kerry Bell (QDAF), Phil Moody (DSITI) and Mike Bell (UQ). DPI025 also received extensive grower engagement namely with Dennis Pozzebon, Jason Loeskow, Lyndel Owens and Mark Savina.

Appendices:

Appendix 1: Predicting cane yield from satellite imagery for Andrew Robson

Compiled by Kerry Bell (DAFFQ) August 2013

Background

There were 8 harvest (region/year) that cane yield data and satellite imagery (top-of-atmosphere corrected SPOT 5) were collected for:

- Bundaberg2010
- Bundaberg2011
- Bundaberg2012
- Isis2010
- Burdekin2011
- Burdekin2012
- Herbert2011
- Herbert2012

Isis can be considered part of the Bundaberg region. The three major regions are Bundaberg, Burdekin and Herbert.

Previously an exponential equation was used to fit predicted yield based on the GNDVI (green normalised difference vegetation index). The aim of this report was to see if there was a way to improve the predictions of cane yield. Information available was variety, ratoon status, actual yield and GNDVI, and for Bundaberg1012, Burdekin2012 and Herbert2012 the four band widths used to calculate GNDVI.

The GNDVI was calculated by (NIR-Green)/(NIR+Green), where in this data set NIR is labelled Band 3, and Green is labelled Band 1.

Methodology

Varieties were excluded if the number of cases was less than 30 blocks. This left 42 varieties out of the original 83 varieties.

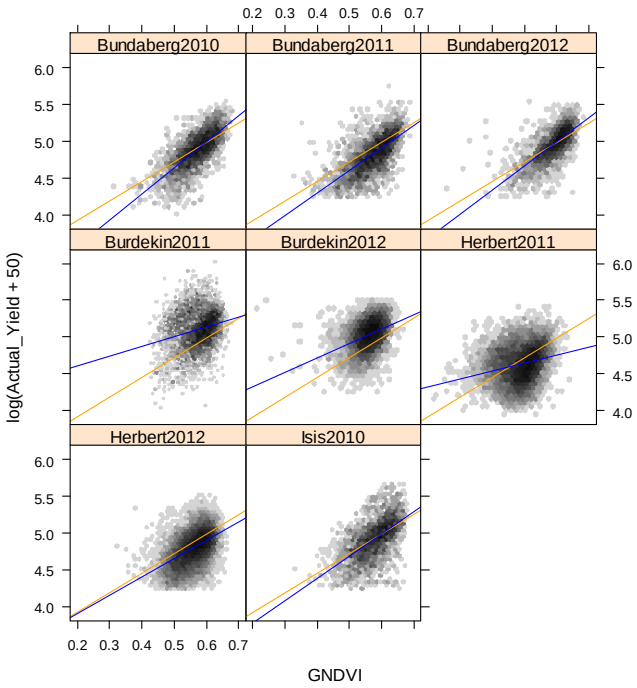
Plotting the data using symbols for each data point obscured the plot as there were too many points. To get an idea of where the majority of the points were located a graphing methods called hexagonal binning was used where the depth of colour of each hexagon represents the number of values in that area of the plot. This gave an indication of where the majority of the points positioned themselves on the graph.

In this study we wanted to determine if levels of harvest (site/year), populations and ratoon status were having different trends in the fit of yield versus gndvi. Using an exponential curve to model the fit proved to be problematic as the parameters often vary dramatically especially when curvature or lower asymptote couldn't be fitted easily with the data provided. To overcome this issue the logarithm (base e) of yield was used and this tended to make the trend linear and the data spaced more evenly (removing some issues with bias from extreme points). The effect of the log transformation was improved by using log(yield + 50).

Once the trend could be fitted with a straight line the different levels of the factors could be explored based on the intercept and slope.

Results

The following plot shows the transformed yield against GNDVI for each harvest.



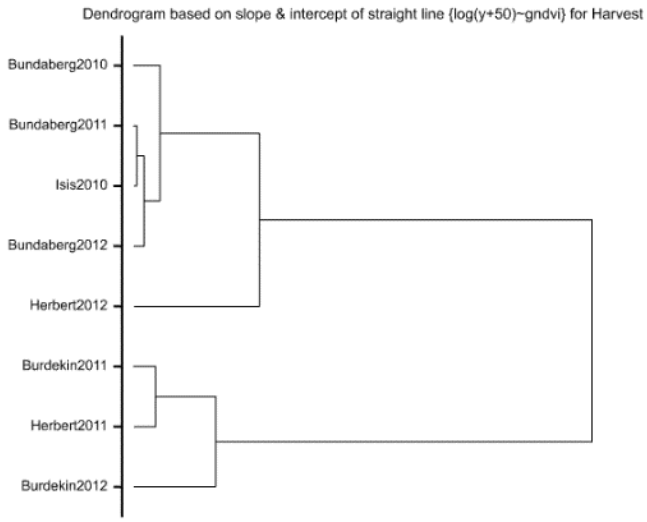
The orange lines show a common line fit across all harvests, and the blue lines are lines fitted individually to each harvest.

	A	B	seA	seB	AdjR2	
Bundaberg2010	2.879	3.520	0.035	0.061	53.0	this group has steeper
Bundaberg2011	3.075	3.052	0.035	0.061	44.3	slope than the next group
Bundaberg2012	3.020	3.283	0.038	0.064	49.9	and generally a better fit
Isis2010	3.179	3.006	0.046	0.080	34.8	
Herbert2012	3.396	2.504	0.021	0.037	25.3	
Herbert2011	4.106	1.081	0.024	0.049	5.7	poor fit
Burdekin2011	4.339	1.341	0.043	0.075	6.2	poor fit
Burdekin2012	3.941	1.951	0.034	0.061	13.6	
Overall	3.378	2.679	0.011	0.019	30.1	

A grouped regression is oversensitive to finding differences between harvests due to the large number of values available for analysis. However through visual assessment it can be seen that Bundaberg 2010, Bundaberg 2011, Bundaberg 2012 and Isis 2010 have similarly behaving lines (i. e. Bundaberg region). A dendrogram based on the slope and intercepts for each harvest supports this grouping, and also loosely groups Herbert 2012 with this group.

The harvests for Burdekin 2011, Burdekin 2012 and Herbert 2011 form a group according to the dendrogram, and it can be seen that the regressions have a lesser slope.

In general the harvests with the Bundaberg region and the Burdekin region seem to be behaving similarly, but the harvests from the two years in the Herbert region are acting differently.



Varieties and Ratoon classes

The slope and intercepts of varieties and ratoon classes within each harvest (site/year) was explored using scatter plots of slope and intercept, and dendrograms similar to above, however there appeared to be no consistency in the groupings from one harvest to the next. As the aim of the analysis was to produce a simple way of getting better predictions of yield, the use of picking out major groups of variety or ratoon class was abandoned.

Exploring band widths

Data for the four band widths were only provided for Bundaberg 2012, Burdekin 2012 and Herbert 2012. A multiple regression was used to explore the fit of the band widths in predicting yield, using the transformation of log(yield+50).

Adjusted R^2		Y=log(yield+50)	
	Bundaberg2012	Burdekin2012	Herbert2012
Band3 (NIR)	41.4%	26.2%	17.1%
add Band1 (Green)	51.0%	26.3%	25.4%
add Band2	51.2%	28.3%	26.0%
add Band4	51.2%	29.4%	28.1%
GNDVI only	49.9%	13.6%	25.3%

The NIR band3 was fitted first as it tended to explain most of the variation within the regressions. The Green band1 was then fitted next as it was used in calculating GNDVI.

Scanning the adjusted R^2 showed that adding the Green band1 provided an improved model than using NIR band3 alone for Bundaberg2012 and Herbert2012, but little benefit for Burdekin2012. The bands 2 and 4 had a minimal impact on improving the fit of the regression.

The fit of GNDVI (=(band3-band1)/(band3+band1)) was approximately showing the same percentage of variation accounted for (=adjusted R^2) as using an additive model of band1 and band 3 for Bundaberg 2012 and Herbert 2012, but was a worse fit for Burdekin 2012.

The following table shows the estimates of the regression coefficients:

Y=log(yield+50)	Constant	se	Band1	se	Band3	se	n	Adj R^2
Bundaberg2012	4.8801	0.0530	-12.246	0.538	3.3397	0.0689	2658	51.0%
Burdekin2012	3.8778	0.0351	0.776	0.361	3.6213	0.0769	6520	26.3%
Herbert2012	5.0432	0.0274	-14.464	0.372	3.1339	0.0483	13716	25.4%
Combined	4.1684	0.0144	-2.289	0.216	3.1078	0.0369	22894	28.1%

Fitting exponential curves using GNDVI for each region

The exponential regressions initially used to predict yield did not have a constant in the model, i.e. the A was missing from:

Yield = A + B * R^{GNDVI}

The form of the exponential curve was also slightly different using Yield = b exp(-k*GNDVI), though this should be comparable when consider k=-log(r). This may be an artefact of using Excel.

With most harvests the same regression equation was used in the data provided.

Given that the regions seem to be behaving differently and that there is a possibility that the constant in the model may add some extra information, the GNDVI was fitted for each region using an exponential curve as described above. Fitting curves to individual harvests may have been more accurate but does not allow for application to future predictions.

	A	Se	B	SE	R	SE	Adj R^2
Bundaberg	15.74	3.54	1.376	0.356	853	302	44.8%
Burdekin	69.18	6.81	1.023	0.782	696	735	9.5%
Herbert	35.17	1.36	0.3175	0.0782	4665	1718	26.3%

With this form of the equation the y-intercept is A + B (in the form fitted in excel the y-intercept is b). The B parameter is like a stretching factor along the x-axis, the R value reflects how quickly the curve bends.

Predicting Yield

Based on the information so far in this report, we can suggest that fitting separate lines to the different regions is worthwhile investigating for improving the predictions of yield.

There are 3 methods to explore:

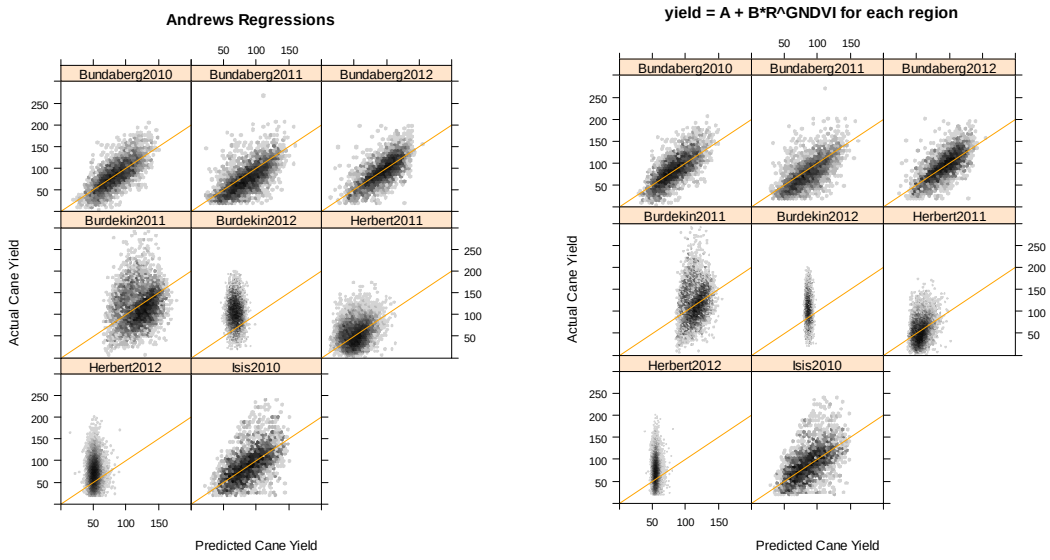
i) Use the existing regressions provided by Andrew Robson

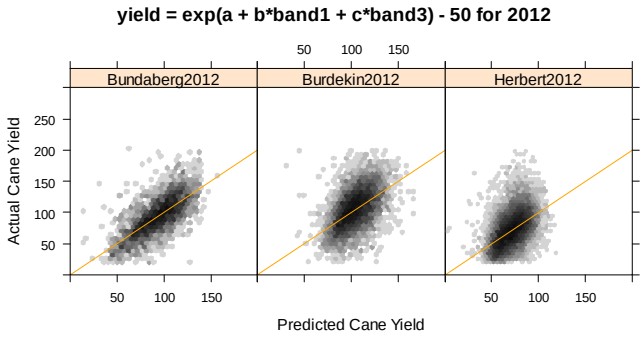
ii) Develop new exponential regressions based on GNDVI for each region

iii) Develop new linear models using NIR (band3) and Green (band1) to predict the log(yield+50)

- however iii) can only be used for the harvests Bundaberg 2012, Burdekin 2012 and Herbert 2012 as the band widths are only provided for these harvests.

The following graphs show the actual versus predicted yields based on these three different methods





By fitting the multiple linear regression of the band widths 1 and 3 for Green and NIR on the log(yield + 50), the predicted yield for Burdekin 2012 and Herbert 2012 has taken on a broader range of values.

	Intercept	se	t-prob	slope	se	Adj R2
Bundaberg2010						
i) Andrew	-1.05	1.54	0.496	1.0339	0.0177	53.3%
ii) GNDVI exp	-5.28	1.61	0.001	1.0914	0.0187	53.2%
Bundaberg2011						
i) Andrew	2.17	1.56	0.166	0.8816	0.0175	45.1%
ii) GNDVI exp	-1.69	1.63	0.300	0.9327	0.0184	45.4%
Bundaberg2012						
i) Andrew	3.52	1.80	0.051	0.9792	0.0188	50.5%
ii) GNDVI exp	0	1.86	0.999	1.0279	0.0197	50.7%
iii) Bands 1&3	1.20	1.86	0.519	1.0041	0.0194	50.2%
Burdekin2011						
i) Andrew	69.40	3.48	<0.001	0.4227	0.0285	4.4%
ii) GNDVI exp	42.37	5.32	<0.001	0.6785	0.0460	4.3%
Burdekin2012						
i) Andrew	24.58	2.53	<0.001	0.7201	0.0220	14.1%
ii) GNDVI exp	-28.44	4.10	<0.001	1.2217	0.0370	14.3%
iii) Bands 1&3	3.82	2.15	0.075	0.9827	0.0203	26.4%
Herbert2011						
i) Andrew	31.04	1.01	<0.001	0.4638	0.0181	7.4%
ii) GNDVI exp	9.12	1.76	<0.001	0.8217	0.0306	8.1%
Herbert2012						
i) Andrew	18.672	0.872	<0.001	0.7025	0.0108	23.6%
ii) GNDVI exp	1.65	1.14	0.148	0.9859	0.0153	23.1%
iii) Bands 1&3	2.57	1.11	0.021	0.9925	0.0152	23.7%

The t-prob for the intercept gives a guide to bias in the fit. Ideally the intercept should not be significantly different from zero and the slope not significantly different from one (roughly the difference from one shouldn't be greater than twice the standard error) if the relationship follows the 1:1 line.

The actual versus predicted yield give a similar adjusted R^2 for Andrew's predictions and refitting the GDNVI with the intercept also fitted.

Where the bandwidth data was available in the 2012 data, there was a major improvement in the adjusted R2 for Burdekin2012 data set, but not so for the Bundaberg and Herbert data sets. The regressions based on the bandwidths also gave a broader range of predicted values for the Burdekin and Herbert data sets.

Discussion

In a lot of the harvests there wasn't any major improvement in the adjusted R^2 with trying the different ways of predicting yield, except for the Burdekin 2012 where using a multiple regression of bands 1 and 3 (Green and NIR). It may be worth investigating using the band widths in the other five harvests where this data was not available.

Appendix 2: Effect of Variety on Yield - Bundaberg Region. Falzon G. & Robson A. 23/5/14.

Overview: The objective of this work is to assess the impact of cane variety on yield in the Bundaberg region. The yield data from 2010 to 2013 is assessed to determine if different cane varieties produce different yields. In all 31 different varieties of sugar cane were assessed. These varieties are specified in table 1.

Analysis: A box plot of yield with respect to all varieties immediately indicates apparent differences between varieties. For instance variety CP51 has a small and narrow range of yield which barely overlaps with the majority of measurements of variety Q135. A large number of outliers are also apparent (empty circles) which indicate that certain varieties have a subset of measurements that are of either extremely low or extremely high magnitudes. Such outliers are a likely cause of error in the yield prediction models. Confidence intervals about the median yield were also estimated for the different varieties. The median statistic was used (rather than the mean) due to the presence of many outliers.

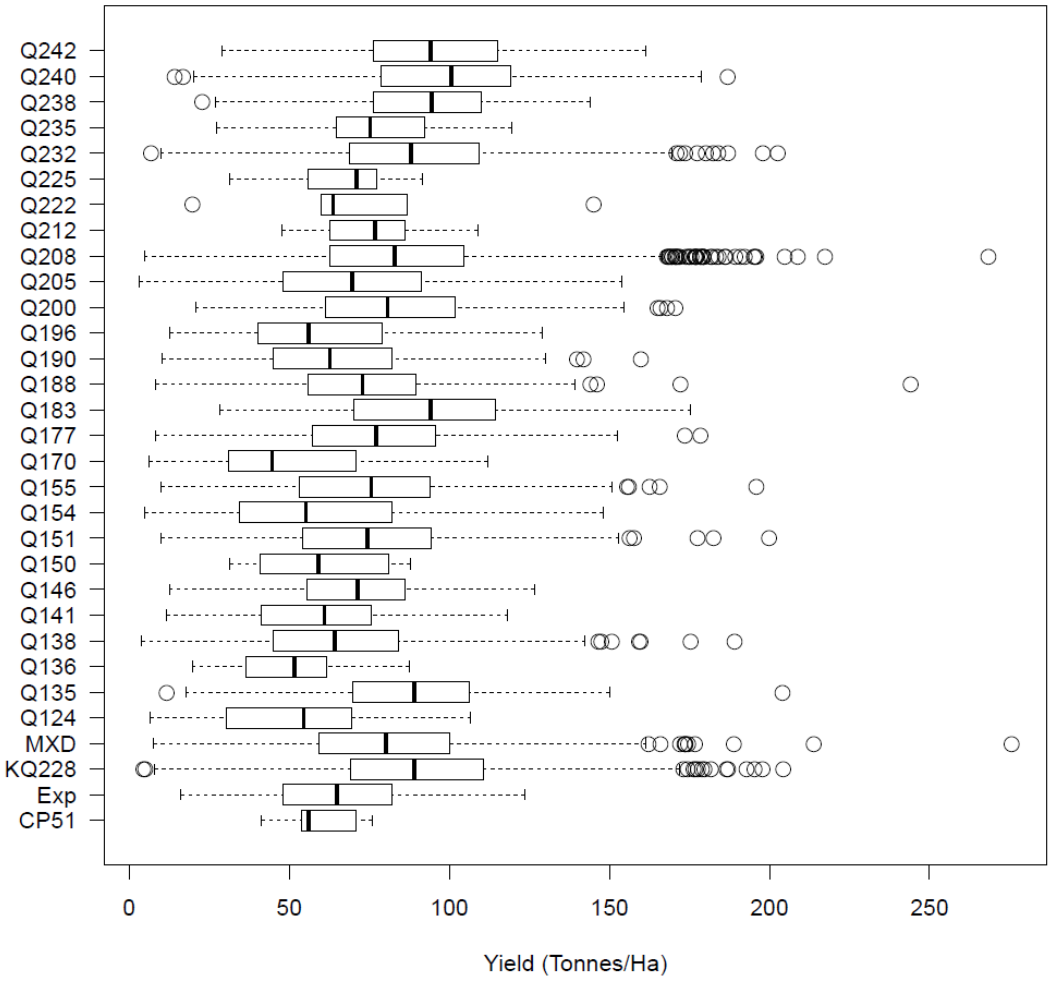


Figure 1: Boxplot of Yield (Tonnes/Ha) for all varieties.

Variety	Lower	Upper	n
CP51	53.27	75.10	7
Exp	54.75	74.38	44
KQ228	87.38	90.97	1598
MXD	78.16	82.44	1299
Q124	44.94	63.97	40
Q135	83.24	94.50	167
Q136	47.21	63.54	15
Q138	61.36	66.82	736
Q141	47.27	65.50	72
Q146	67.22	77.17	84
Q150	33.17	87.48	8
Q151	72.19	76.63	1029
Q154	40.73	69.46	43
Q155	70.66	77.84	438
Q170	39.14	55.36	91
Q177	72.55	80.61	251
Q183	84.93	98.76	114
Q188	70.62	75.60	520
Q190	58.96	68.77	128
Q196	49.29	71.93	49
Q200	78.05	83.40	585
Q205	66.57	73.83	358
Q208	82.01	84.13	3959
Q212	47.71	108.80	5
Q222	19.69	145.06	5
Q225	51.73	90.03	9
Q232	86.27	90.29	1198
Q235	71.85	82.53	79
Q238	78.02	105.98	20
Q240	98.21	105.11	295
Q242	88.58	99.59	132

Table 1: Estimated 95% confidence intervals of the median Yield/Ha for the different varieties in the Bundaberg data set and the associated numbers of data points (n).

The top five highest median yielding varieties were Q240 (100.71 Tonnes/Ha), Q238 (94.41 Tonnes/Ha), Q183 (94.20 Tonnes/Ha), Q242 (94.04 Tonnes/Ha) and Q135 (89.05 Tonnes/Ha). Based on the results in table 1 there is clearly significant differences in median yield which depend on variety. Statistical bootstrap methods were also used to estimate the 95 % confidence intervals of the inter-quartile range (IQR) for each variety. This information which is reported in table 2 might be useful to a producer to assess which variety has the most consistent yield across blocks. Observe that these intervals were quite large when sample size was relatively small (e.g. Q222 n = 5) and in such cases the estimates should be interpreted with caution. For those varieties with larger sample sizes e.g. Q188 (n = 520) and Q232 (n = 1198) greater confidence can be placed in the accuracy of the estimated intervals. Differences in the IQR with respect to variety are apparent e.g. Q188 [29.46, 36.78] and Q232 [37.59, 43.54] as the confidence intervals do not overlap.

The relationship between GNDVI and Yield/Ha was also examined for each variety. Figure 2(a)-(d) displays this relationship for select varieties. Inspection of these figures suggests that prediction models should account for the different varieties. The intercept and the slope might vary between varieties. A multi-variate generalised linear model (GLM) was fit to assess the statistical significance of variety, as displayed in table 3 the effect of variety is highly statistically significant. Comparison of the prediction errors between a GLM with and without variety information is presented in table 4. Inclusion of variety information provides little improvement on the prediction errors associated with yield despite the fact that variety information has been demonstrated to be a statistically significant effect.

One possibility is that the GLM is fitting poorly for particular varieties and hence artificially inflating error estimates. Table 5 displays the relative frequencies of observation above 13 Tonnes/Ha for each variety. The

13 Tonnes/Ha corresponds to approximately the 3rd quartile of the error magnitudes in table 4 and errors above 13 Tonnes/Ha can be considered as large. The median relative frequency across all varieties is 0.53 but for some varieties this proportion is relatively low (CP51, 0.29) or relatively high (Q222, 0.80). This suggests the GLM is fitting well for particular varieties but not others. Table 6 displays the GLM predictions for varieties CP51, MXD and Q222, separate models specific to each variety has the capability to greatly reduced the errors. The maximum error on the MXD variety is still quite large indicating that the issue associated with observations of extremely high yielding blocks has not been fully resolved.

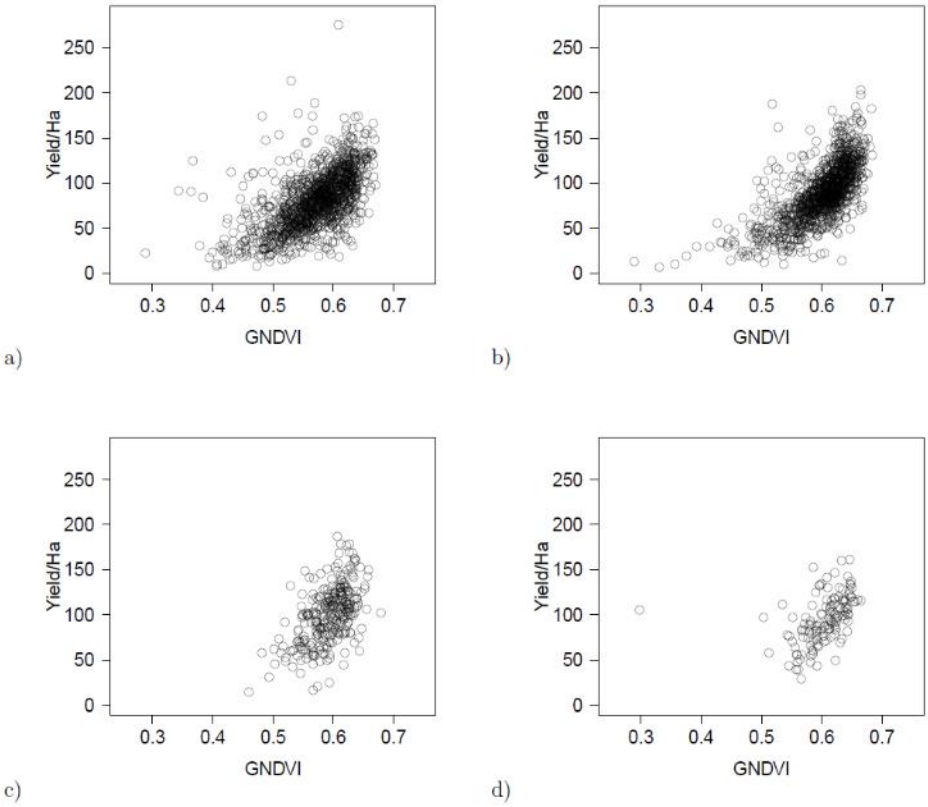


Figure 2: Relationships between GNDVI and Yield/Ha for the (a) MXD, (b) Q232 , (c) Q240 and (d) Q242.

Variety	Lower	Upper
CP51	1.95	28.15
Exp	23.78	44.48
KQ228	38.83	43.22
MXD	38.37	43.91
Q124	21.61	52.89
Q135	30.82	43.67
Q136	7.41	48.21
Q138	35.65	41.63
Q141	26.58	43.66
Q146	24.38	38.02
Q150	8.79	54.80
Q151	37.16	42.10
Q154	30.44	56.27
Q155	34.06	45.52
Q170	25.79	47.36
Q177	32.88	45.56
Q183	34.44	53.46
Q188	29.46	36.78
Q190	30.01	43.62
Q196	28.84	48.02
Q200	36.21	43.26
Q205	36.2	47.47
Q208	40.35	43.65
Q212	0.00	61.09
Q222	0.00	125.38
Q225	3.84	45.82
Q232	37.59	43.54
Q235	18.86	37.69
Q238	14.78	54.52
Q240	36.11	47.28
Q242	29.69	42.85

Table 2: Estimated 95% confidence intervals of the Interquartile Range Yield/Ha for the different varieties in the Bundaberg data set.

Term	$\Pr(>\chi^2)$
exp(GNDVI)	< 2.2e-16
Variety	< 2.2e-16

Table 3: Statistical significance of GLM using analysis of variance .

	Min	Q1	Med	Mean	Q3	Max
With	-94.08	-14.61	-1.09	0.00	12.88	180.20
Without	-92.98	-15.00	-1.20	0.00	12.93	179.60

Table 4: Error comparisons between models including (with) or omitting (without) the Variety information. The notation 'Min' denotes minimum, 'Q1' first quartile, 'Med' median, 'Mean' the arithmetic mean, 'Q3' third quartile and 'Max' the maximum error.

Variety	Frequency
CP51	0.29
Exp	0.57
KQ228	0.53
MXD	0.54
Q124	0.65
Q135	0.51
Q136	0.67
Q138	0.52
Q141	0.50
Q146	0.45
Q150	0.50
Q151	0.50
Q154	0.56
Q155	0.57
Q170	0.58
Q177	0.56
Q183	0.61
Q188	0.47
Q190	0.57
Q196	0.57
Q200	0.48
Q205	0.55
Q208	0.52
Q212	0.40
Q222	0.80
Q225	0.33
Q232	0.49
Q235	0.29
Q238	0.55
Q240	0.62
Q242	0.49

Table 5: Relative frequencies of observations in each variety above 13 Tonnes/Ha in error.

	Min	Q1	Med	Mean	Q3	Max
CP51	-11.32	-7.63	-3.40	0.00	8.34	13.28
MXD	-76.21	-15.84	-1.63	0.00	13.56	181.80
Q222	-40.58	-32.06	-6.23	0.00	36.66	42.21

Table 6: Error comparisons for models specific to the CP51, MXD and Q222 varieties. The notation 'Min' denotes minimum, 'Q1' first quartile, 'Med' median, 'Mean' the arithmetic mean, 'Q3' third quartile and 'Max' the maximum error.

Conclusion

Variety influences yield and is likely to also influence the relationship between GNDVI and yield. Models fit specifically for each variety revealed that the prediction errors were within an acceptable range for particular varieties and it was the presence of extremely high yields for some observations and varieties which case the greatest prediction errors. Variety information might be of some use for increasing prediction accuracy but priority should be given to identifying the factors which lead to extremely high yields as it is these observations which negate the potential improvements of including variety information.

Appendix 3: Sugar Cane Yield Prediction – Bundaberg

This document reports on the current progress to improve models utilised in the sugar cane industry to estimate yield using remote sensing imagery. Current models are based on an assumed exponential relationship between GNDVI and Yield but as will be demonstrated this model leads to a very poor model fit and extremely large residuals under some circumstances. The ultimate objective of this work is that of forecasting (or predicting) sugar cane yield from remote sensing data but before doing so sufficient evidence must be produced that predictions of acceptable accuracy can be obtained using pre-existing data. To investigate whether or not such accuracy is possible a range of alternative modelling approaches are explored and proposed. In particular the relationship between Yield/Ha and GNDVI is comprehensively investigated.

Data Set: The data supplied consists of 13379 records from 16 variables obtained via remote sensing of the Bundaberg sugar cane growing region over 2010, 2011, 2012 and 2013. These variables are listed in Table 1.

Current Model: The current model in use examines the univariate relationship between Yield/Ha and GNDVI. The specific mathematical form of this model is,

$$Y = \beta_0 + \beta_1 \exp(X) + \epsilon \qquad (1)$$

where Y is the Yield/Ha, X is the GNDVI magnitude β_0, β_1 are the intercept and slope terms respectively and $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is a normally distributed error term.

The model of equation 1 can be estimated using the linear regression techniques available in the statistical software R. Table 2 reports on the coefficient estimates and the corresponding analysis fit using model 1 and all of the available Yield/Ha and GNDVI data irrespective of year. Both the intercept and slopes are statistically significantly different from zero and are estimated to a relatively small level of uncertainty. Further assessment of the models residuals is reported in Table 3. It is apparent that the residuals are not normally distributed as they have positive skew thereby breaking a key modelling assumption. Furthermore although the average residual is 0 tonnes/Ha in the worst cases the model can either under estimate yield by 93 tonnes/Ha or over-estimate yield by 180 tonnes/Ha.

Figure 1(a) displays the plot of all GNDVI values against Yield/Ha values 2 as well as the fit (dotted line) of model 1. An increasing trend between GNDVI and Yield/Ha is apparent but there is a considerable amount of scatter (variation) as well as apparent outliers (large values) which suggests that a regression model might not be able to accurately predict Yield/Ha for all observations. The fit obtained using model 1 reinforces this viewpoint, there is wide variation about fitted line as indicated by a model R^2_{adj} = 0.4383. Importantly the parametric form of model 1 seems to be mis- specified that is the true relationship between Yield/Ha and GNDVI is not in fact based on an exponential function. Model 1 seems to be inadequate in capturing the trend at high GNDVI magnitudes. Figure 1(b) indicates that the response (Yield/Ha) is not in fact normally distributed and in fact that a Gamma distribution would be more appropriate. The fact that the assumptions of model 1 are not adequate for the Yield/Ha response provides immediate grounds for model improvement.

Variable	Description	Type
YEAR	Growing Year	Numeric
Image Date	Image Acquisition Date	Date
LINKCODE	Computer Processing Code	Categorical
Mill	Harvest Hill	Categorical
Irrigate	Irrigation Method	Categorical
Variety	Cane Variety	Categorical
Class	Cane Class	Categorical
AREA Ha	Area of field	Numeric
Act TCH	Yield/Ha	Numeric
Green	Remote Sensing Data	Numeric
Red	Remote Sensing Data	Numeric
NIR	Remote Sensing Data	Numeric
MIDIR	Remote Sensing Data	Numeric
GNDVI	Remote Sensing Data	Numeric
X	UTM location	Numeric
Y	UTM location	Numeric

Table 1: List of variables in the Bundaberg data set.

Estimate	SE	t-value	Pr(> |t|)
-321.752	3.953	-81.39	<2e-16
227.435	2.226	102.16	<2e-16

Table 2: Coefficient estimates and analysis of model 1, $R^2_{adj} = 0.4383$.

Min	Q1	Median	Mean	Q3	Max
-93	-15	-1	0	13	180

Table 3: Residual analysis of model 1 (nearest tonnes/Ha). The entries denote: 'Min' - minimum, 'Q1' - first quartile, 'Median' - median, 'Mean' - mean, 'Q3' - third quartile, 'Max' - maximum residual magnitudes.

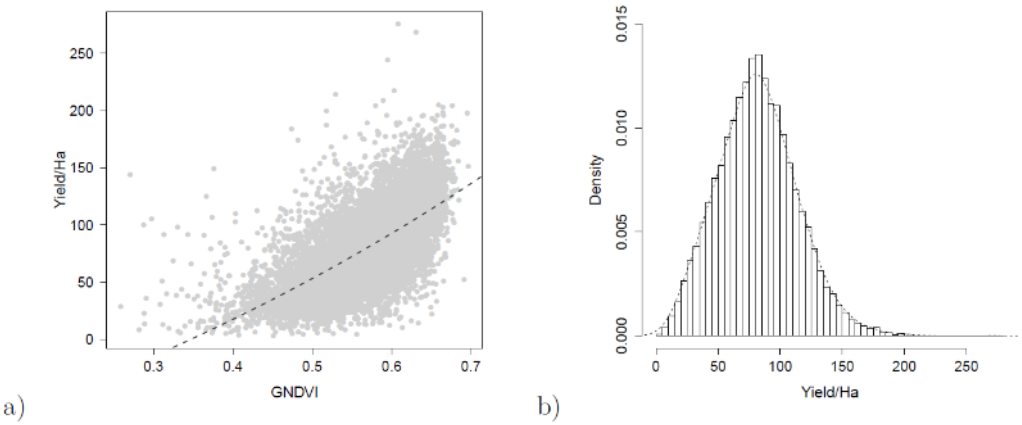


Figure 1: Current model assessment: (a) model fit (dotted line) of the univariate relationship between Yield/Ha and exp(GNDVI), (b) histogram and kernel density estimate of the Yield/Ha magnitudes.

Gamma Generalised Linear Model:
The natural first step in model improvement is to ensure the correct parametric assumptions are modelled for the response variable Yield/Ha. Visual inspection suggests that a distribution from the Gamma family of distributions would be adequate and this notion is supported by the fact that Yield/Ha is strictly greater than or equal to zero (and in this data set always greater than zero). The model has the form of equation 2 when an inverse link function generalised linear model is assumed.

$$Y = \frac{1}{\beta_0 + \beta_1 \exp(X)} + \epsilon \qquad (2)$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$.

The corresponding model _t is displayed as the dashed line of Figure 2. Observe that the _t produced by this model more accurately (at least by visual inspection) captures the overall trend between the Yield/Ha and GDNVI variables. Table 4 reports on the model 2 coefficient estimates and the associated p-values. Both the intercept and slope terms are significantly different from zero. Note that a pseudo R^2_{adj} term was also reported for this model, it should only be considered an approximation to the coefficient of determination and it is difficult to directly compare to other models. It does suggest that model 2 fits better overall than model 1. In particular the residuals of model 2 were found to be approximately normally distributed (analysis not displayed) indicating that another key modelling assumption was satisfied. Model 2 also has the desirable quality of not tending to over- or under-estimate Yield/Ha for a particular range of GNDVI values. This is in contrast to model 1 which is particularly ill suited for the lower and higher range GNDVI magnitudes. Despite significant improvements in the statistical assumptions and the predictive behaviour (similar behaviour across all GNDVI), model 2 has not improved the accuracy of the yield predictions and could in fact be slightly worse. Table 6 provides the summary statistics for the residuals of model 2. Observe that the first and third quartile residuals are of similar magnitude to those of model 1 but that the maximum error is now 2 tonnes/Ha greater. The fact that the minimum and maximum values in the residuals are of similar magnitude should be considered as further evidence of normality. The Gamma generalised linear model has improved both the rigour and the predictive behaviour of the methodology but it has not improved on the practical objective of more accurately estimating Yield/Ha. Therefore more advanced modelling needs to be considered.

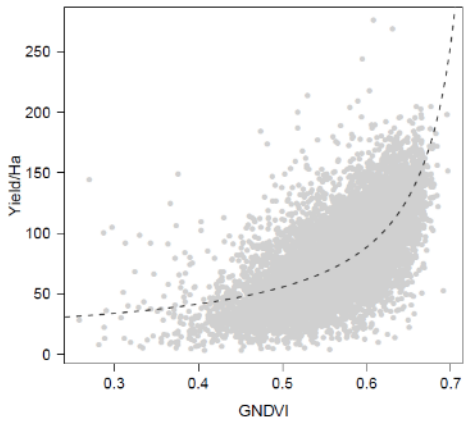


Figure 2: Gamma Generalised Linear Model fit (dotted line) of the univariate relationship between Yield/Ha and GNDVI.

Term	Estimate	SE	t-value	Pr(> |t|)
β_0	0.0815848	0.0007298	111.80	<2e-16
β_1	-0.0385778	0.0004007	-96.27	<2e-16

Table 4: Coefficient estimates and analysis of model 2, $R^2_{adj} = 0.6620255$.

Min	Q1	Median	Mean	Q3	Max
-166	-15	-1	0	14	182

Table 5: Residual analysis of model 2 (nearest tonnes/Ha). The entries denote: 'Min' - minimum, 'Q1' - first quartile, 'Median' - median, 'Mean' - mean, 'Q3' - third quartile, 'Max' - maximum residual magnitudes.

Generalised Linear Quantile Regression Model.

With the fundamental parametric modelling assumptions satisfied there is limited opportunity to improve on the linear regression model proposed (model 2) without running the substantial risk of over-fitting the dataset (and compromising forecast accuracy). There is however the opportunity to improve on the likely mis-specified exponential trend and also create an envelope of most likely predictions so that yield can be estimated to be between a certain range of values for each value of GNDVI. The exponential trend might be improved on by using a natural spline basis thereby allowing the data suggest the trend of the model. Whilst quantile regression can be used to estimate the most likely range of Yield/Ha values for a given GNDVI value. Use of a link function allows the data to still obey a Gamma distribution conditional on the GNDVI magnitude. Combining these three characteristics produces a Generalised linear quantile regression whose mathematical form is displayed in model 3.

$$Y = \eta\left(\beta_0 + \sum_i^n \alpha_i B_{i,k,t}(X)\right) + \epsilon \qquad (3)$$

where $\eta(\cdot)$ is a link function found via a Box-Cox transformation of the response Y to produce a conditional Gamma distribution, β_0 is an intercept term, $B_{i,k,t}(X)$ are terms of a natural spline basis with α_i weights at n knots and $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is the error term.

Model 3 was estimated using the VGAM package in R and the predictions for the 5%, 25%, 50%, 75% and 95% quantiles obtained. The prediction errors are reported in table 6. As expected model 3 has similar magnitude errors as those of model 2 of particular interest are the errors for the 50% quantile which corresponds to regression through the median Yield/Ha. The 50% quantile regression is the one most directly comparable to the results of models 1 and 2. Observe that the maximum (absolute) error in this case is 182 tonnes/Ha which is the same magnitude as that of model 2 and 2 tonnes/Ha greater than that of model 1. Note however there is now a slight skew in the residual distribution and a tendency for the model to underestimate the yield by 2 tonnes/Ha. The first (Q1) and third (Q3) quartiles of the residual error for the 50% quantile regression are similar in magnitude to that of models 1 and 2. That is most yield prediction errors will be up to 15 tonnes/Ha in magnitude. The pseudo R^2_{adj} = 0.6765617 which is a slight improvement on model 2. The advantage of the quantile regression approach is displayed in Figure 3, instead of one prediction there are now five (corresponding to the 5%, 25%, 50%, 75% and 95% quantile respectively). In particular the lower lines (5% quantile) and upper lines (95% quantile) display the region over which 90% of all Yield/Ha magnitudes for a given GDNVI. That is if a particular GNDVI value is supplied then the model can provide not only the most likely Yield/Ha estimate but also a prediction of the range of Yield/Ha values for the majority of cases. Note that the data set contains several high magnitude Yield/Ha for low GNDVI, the spline function used in model 3 indicates that Yield/Ha actually increases for very low (< 0.35) GNDVI values. This might be an artefact of this data set which can be resolved by either collecting more data at such low GNDVI (if available) or a priori constraining the behaviour of the spline at such GNDVI. Such improvements are a possible future direction of research.

	5%	25%	50%	75%	95%
Min	-226	-202	-182	-161	-127
Q1	-55	-33	-15	3	30
Median	-38	-18	-2	16	45
Mean	-39	-19	-2	16	45
Q3	-22	-3	13	31	60
Max	43	70	93	120	167

Table 6: Residual analysis of model 3 (nearest tonnes/Ha). The columns correspond to the 5%, 25%, 50%, 75% and 95% quantiles and the rows report the corresponding 'Min' - minimum, 'Q1' - first quartile, 'Median' - median, 'Mean' - mean, 'Q3' - third quartile, 'Max' - maximum residual magnitudes.

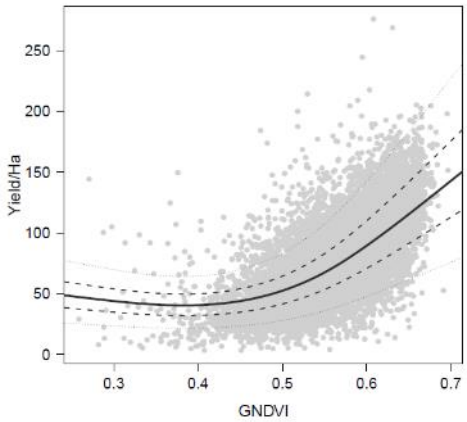


Figure 3: Gamma Quantile Generalised Linear Regression model fit. The lines (starting from the lowest Yield/Ha) correspond to the 5%, 25%, 50%, 75% and 95% quartile prediction curves. Note $R^2_{adj} = 0.6765617$.

Summary & Conclusions:

This report has examined three models designed to predict Yield/Ha from GNDVI as measured via remote sensing of the Bundaberg region. All models had a maximum error of approximately 182-183 tonnes/Ha but models 2 and 3 offered important benefits over model 1. In particular model 2 better met the underlying statistical distribution assumptions than that of model 1, this resulted in uniform behaviour (the model tended to equally over- and under-estimate to the same magnitude) across all GNDVI. The use of quantile regression in model 3 offered further capabilities in that a range of yield predictions could be provided with the true yield captured within this envelope for the majority of cases. Model 3 could be further improved if a greater number of Yield/Ha values were available for low GNDVI magnitudes, alternatively modelling assumptions could be made to specify the behaviour of the curve in this region. Irrespective of these developments there is still considerable scatter in the apparent relationship between Yield/Ha and GNDVI (for this data set). This imposes a fundamental limitation on the accuracy of any linear statistical model proposed. Future research should be directed towards reducing prediction error by the investigation of multi-variate regression models. That is models which incorporate readily available covariate information such as cane variety and spatial location. Different cane varieties are likely to influence Yield/Ha and adjacent fields are likely to have a level of correlation in their Yield/Ha values. Therefore the next line of investigation with this data set will be multi-variate spatial models.

Appendix 4: Report of statistical analysis of regional data from Yvette Everingham

Report Summary:

Question 1: Do the regression lines for modelling yield versus GNDVI change with major climate events?

Answer: We found that for all three locations the regression lines differ significantly from year to year, across most years. Given the instability in the regression model from year to year, it is difficult to determine particular climate events that may contribute to these instabilities. See section 1.2 for more details.

Question 2: Despite the year to year variability in the regression lines, were there any unusual climate effects for the three regions?

Answer: Based on the random forest models generated in Everingham et al. (2015), we closely examined the most important maximum temperature, radiation, rainfall and minimum temperature variables that were found to be important for explaining yields in the three regions. For Ingham the effects of the 2010 La Nina were evident. Here we found that yields harvested in 2011 experienced high Jul-Oct rainfall in 2010 and high minimum temperatures in Aug-Nov in 2010. There was nothing unusual for Condong or Bundaberg for the predictor variables found important for explaining yields. See section 1.2 for more details.

Question 3: Are there links between NDVI and climate signals?

Answer: Section 1.3 highlights some literature that indicates that climate signals like ENSO are embedded in the NDVI signal.

Question 4: Is NDVI useful for modelling sugarcane yields?

Answer: Yes! NDVI can be used as a surrogate for parameters in crop models (See section 1.4). NDVI models can also be used as a separate predictor and combined with predictions from a crop model as part of an ensemble modelling approach.

Future Research Suggestions:

1. Investigate if it is beneficial to sacrifice the higher resolution, short temporal history imagery for lower resolution imagery with a long temporal history.

2. Integrate NDVI signals into the Crop Model

3. Build ensemble models that incorporate crop model estimates and estimates from NDVI.

Important Note:

NDVI models are important especially in years when the crop is invaded by disease as crop models will underperform in these years. It is important the Australian sugarcane industry has a range of yield estimation methods so the strengths and benefits of each approach can be capitalised on. Moreover, it is important that the industry has a range of yield estimation methods that are not subject to human biases.

1. Updated analysis of yield predictions using GNDVI

1.1. Data

Satellite images of Ingham (2011 – 2014), Bundaberg (2010 – 2014) and Condong (2012 – 2014) were obtained from a satellite. From this image GNDVI values were calculated for individual sugarcane paddocks in each region. Table 1 lists the image date and number of individual farms identified for each season and year. GNDVI data for each paddock were paired with farm data including crop variety, age, class, area harvested, fibre and cane yields. Paddock data were sourced from local mills.

Table 4 **-** Image data collection information from three regions

Region	Year	Image Date	No. Paddocks
Ingham	2011	2nd June	8596
	2012	4th April	14736
	2013	25th May	11184
	2014	9th July	12968
Bundaberg	2010	10th May	3544
	2011	27th March	3824
	2012	1st April	3143
	2013	25th April	3348
	2014	19th April	4016
Condong	2012	29th February	1391
	2013	20th April	1354
	2014	19th April	1601

Climate data for Ingham, Bundaberg and Condong were sourced from the SILO patched point database (https://www.longpaddock.qld.gov.au/silo/). Following Everingham et al. (2015), climate data for each region was sourced from a single representative weather station (Table 2). Climate data were used to generate influential climate parameters identified in Everingham et al. (2015). For each region the growing season was defined as starting in June the year before harvest and ending in May the year of harvest. For example, the total rainfall for the period July to October (rain_JASO) was identified as an influential climate variable for harvest yields in Ingham (Victoria Mill). This means that harvest yields in Ingham in 2015 may be influenced by the rainfall received in July – October in 2014. Table 2 lists the weather stations used for each region and the most influential rainfall, maximum temperature, minimum temperature and radiation parameters identified by Everingham et al. 2015 using random forests.

Table 5 - Image data collection information from three regions

Region	Station Name	Latitude	Longitude	Influential Parameters
Ingham	Ingham Composite	-18.65°N	146.18°E	Max Temp: maxt_FM (most influential) Radiation: radn_JFMAM Rainfall: rain_JASO Min Temp: mint_ASON
Bundaberg	Bundaberg Airport	-24.91°N	152.32°E	Rainfall: rain_JFMA (most influential) Max Temp: maxt_JUN Min Temp: mint_JA Radiation: radn_JF
Condong	Murwillumbah (Bray Park)	-28.34°N	153.38°E	Radiation: radn_MA (most influential) Min Temp: mint_MA Max Temp: maxt_JFMA Rainfall: rain_FMA

1.2. Modelling yield on GNDVI

Harvested yields have an exponential relationship to GNDVI of the form

$yields = a.e^{b.GNDVI}$

To investigate if the parameters change for different years, a stepwise linear regression was performed between log transformed yields and GNDVI values in the following steps:

1. Yields measured in tonnes of cane per hectare were log transformed
2. GNDVI were left untransformed.
3. Data for every year in a single region was stored in a single array and binary indicator variables were created to identify paddock data from each year. For example in Condong 2012, 2013 and 2014 data were stored together. The variable Z1 had a value of 1 for data from 2012 and 0 for 2013 and 2014 while the variable Z2 was given a value of 0 in 2012 and 2014 and 1 for 2013. Finally Z3 was given a value of 1 for data from 2014 and a value of zero for data from 2012 and 2013.
4. First level interaction variables between binary variables and GNDVI values were computed.
5. A stepwise linear regression was used to build a model of log transformed yields (independent variable Y). The stepwise regression was allowed to select from the independent GNDVI (X), binary variables [Z1, …, ZN] (where N was the number of seasons available e.g. Z1= 2012, Z2 = 2013 and Z3 = 2014 for Condong) and first order interaction variables [Z1*X, …, ZN*X].

Crops in 2010 and 2011 in the Ingham and Bundaberg regions had larger than normal levels of standover cane which was shown to have significantly different GNDVI levels than other classes of cane. To remove the effect of standover cane, paddocks that could be identified as standover were removed from the analysis.

1.2.1. Ingham

The stepwise linear regression for Ingham (Table 3) suggested that year and year by GNDVI interactions were influential ($p < 0.001$) and resulted in statistically different profile curves for 2011, 2012, 2013 and 2014 (Figure 1). To avoid biases in predictions, separate prediction models are needed for 2011, 2012, 2013 and 2014. The separate regression lines are shown in Figure 1. The global model explains 31.7% of the variability in log transformed yields.

Table 3 - Stepwise regression model for Ingham. The overall model variables were recorded and deconstructed[1] to form linear equations for each season, R^2 represents the amount of variance explained by the model. P values represent the significance of the model. P values of less than 0.05 were considered significant.

Year	Parameters	Coefficient	R^2	P
All years	C (constant)	0.914	0.317	<0.001
	X (GNDVI)	5.545		
	Z4 (2014)	1.143		
	Z2*X (GNDVI in 2012)	-0.623		
	Z1*X (GNDVI in 2011)	0.272		
	Z4*X (GNDVI in 2014)	-1.380		
	Z2 (2012)	0.545		
2011	C (constant)	0.914		
	X (GNDVI)	5.817		
2012	C (constant)	1.459		
	X (GNDVI)	4.922		
2013	C (constant)	0.914		
	X (GNDVI)	5.545		
2014	C (constant)	2.057		
	X (GNDVI)	4.165		

The four climate parameters investigated for the Ingham region were maxt_FM, radn_JFMAM, rain_JASO and mint_ASON. Total rainfall from January to October (rain_JASO) and mean minimum temperature from August to November were unusually high for the 2011 season (Figure 2). As the 2011 season runs from June 2010 to May 2011 these high values coincide with the high rainfall of 2010. No other season was considered an outlier (unusually high or low) for any parameter considered.

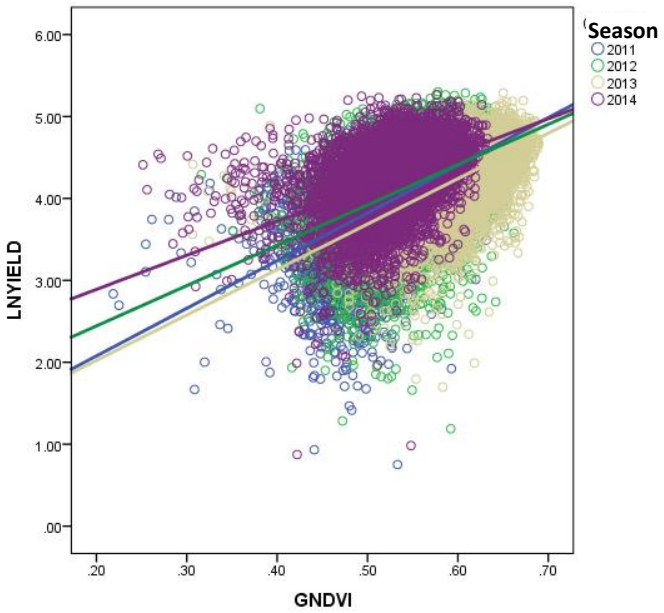


Figure 3 – Log transformed TCPH (LNYIELD) versus GNDVI for Ingham in 2011 (blue), 2012 (green), 2013 (gold) and 2014 (purple) seasons. Lines represent the regression equation for each season (Table 3).

[1] Deconstruction occurs by substituting 1 or 0 for the Z variables with either 1 or 0. Consider for example the year 2011, here, Z1=1, Z2=0, Z3=0 and Z4=0 which gives the equation log(yields)=0.914+5.817.GNDVI

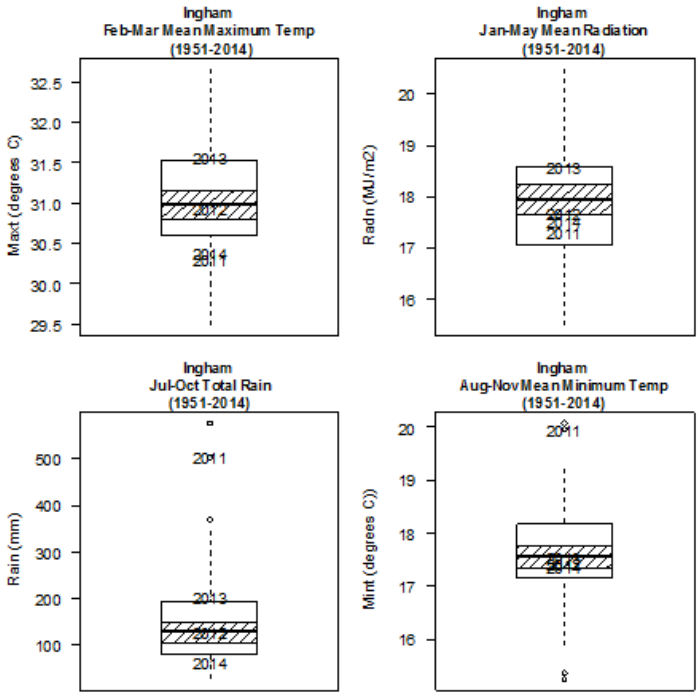


Figure 4 - Key climate parameters for Ingham for the period 1951 to 2014 based on random forest variable importance. Boxes represent 25th to 75th percentile, solid black line is median, shaded area represents an approximate 95% confidence interval on the long term median.

1.2.2 Bundaberg

The global stepwise regression explained 47.5% of the variance in observed log transformed TCPH (Table 4). The TCPH-GNDVI relationship was the same for 2010 and 2012. All other models were different.

Table 4 - Stepwise regression model for Bundaberg. The overall model variables were recorded and deconstructed to form linear equations for each season, R^2 represents the amount of variance explained by the model. P values represent the significance of the model. P values of less than 0.05 were considered significant.

Year	Parameters	Coefficient	R^2	P
All years	C (constant)	0.988	0.475	<0.001
	X (GNDVI)	5.934		
	Z2*X (2011*GNDVI)	-0.237		
	Z4*X (2013* GNDVI)	0.259		
	Z5 (2014)	-0.798		
	Z5*X (2014* GNDVI)	1.397		
	Z4 (2013)	-0.257		
2010/2012	C (constant)	0.988		
	X (GNDVI)	5.934		
2011	C (constant)	0.988		
	X (GNDVI)	5.697		
2013	C (constant)	0.731		
	X (GNDVI)	5.675		
2014	C (constant)	0.190		
	X (GNDVI)	7.331		

The four climate parameters investigated for the Bundaberg region were rain_JFMA, maxt_JUN, mint_JA and radn_JF (Figure 4). The five seasons analysed (2010 to 2014) were not considered outliers for any of the climate parameters investigated. The 2013 season had the highest total rainfall from January to April (rain_JFMA) and the lowest mean maximum temperature in June (maxt_JUN) of the five seasons.

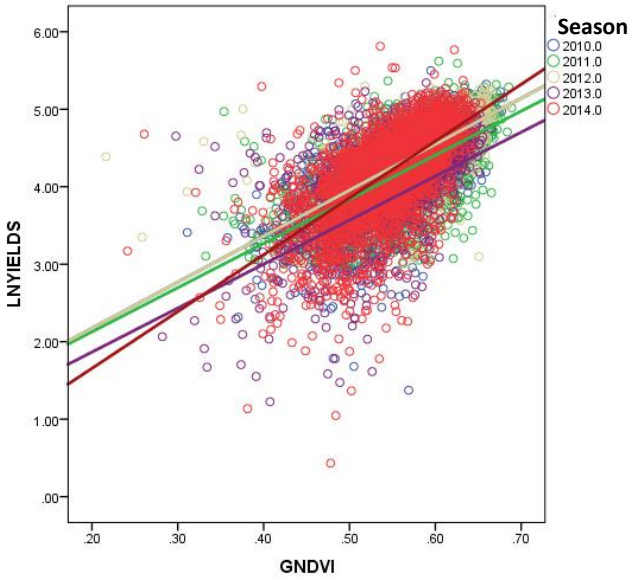


Figure 3 - Log transformed TCPH (LNYIELDS) versus GNDVI for Bundaberg in 2010 (blue), 2011 (green), 2012 (gold), 2013 (purple) and 2014 (red) seasons. Lines represent the regression equation for each season (Table 3). Gold line represents 2010 and 2012 seasons as these had the same equation.

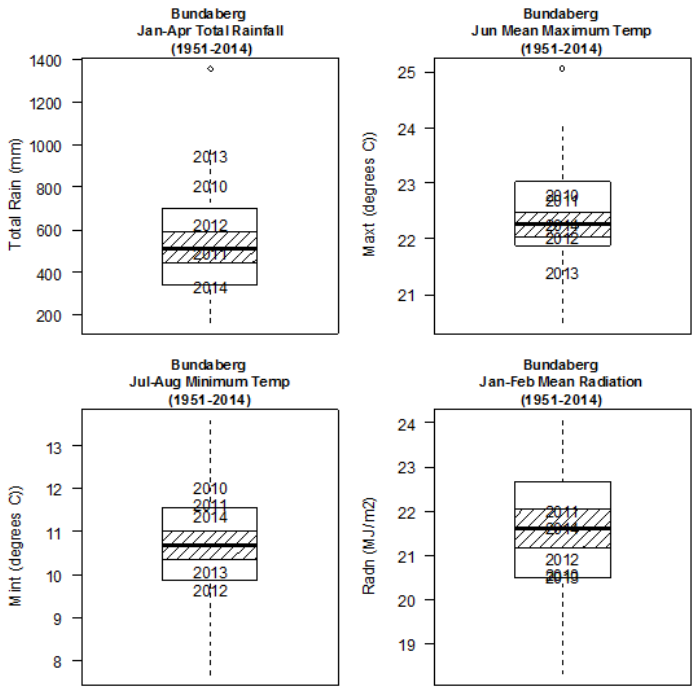


Figure 4 - Key climate parameters for Bundaberg for the period 1951 to 2014 based on random forest variable importance. Boxes represent 25th to 75th percentile, solid black line is median, shaded area represents an approximate 95% confidence interval on the long term median.

1.2.3 Condong

The global stepwise regression explained 63.6% of the variance in observed log transformed TCPH using GNDVI and year (Table 5). As for the Ingham region, the stepwise linear regression resulted in statistically different profile curves for each season (2012, 2013 and 2014).

Table 5 - Stepwise regression model for Condong. The overall model variables were recorded and deconstructed to form linear equations for each season, R^2 represents the amount of variance explained by the model. P values represent the significance of the model. P values of less than 0.05 were considered significant.

Year	Parameters	Coefficient	R^2	P
All years	C (constant)	1.054	0.636	<0.001
	X (GNDVI)	3.886		
	X*Z3(GNDVI*2014)	2.126		
	X*Z2(GNDVI*2013)	1.661		
	Z1 (2012)	0.854		
2012	C (constant)	1.908		
	X (GNDVI)	3.886		
2013	C (constant)	1.054		
	X (GNDVI)	5.547		
2014	C (constant)	1.054		
	X (GNDVI)	6.012		

The four climate parameters investigated for the Condong region were radn_MA, maxt_MA, mint_JA and rain_JF (Figure 6). Similarly to the Bundaberg region, the three seasons analysed (2012 to 2014) were not considered outliers for any of the climate parameters investigated. The 2012 and 2013 seasons had a lower mean radiation (March to April) and lower mean minimum temperature (July to August) compared to the 2014 season. Mean maximum temperature (March to April) was similar for all three seasons.

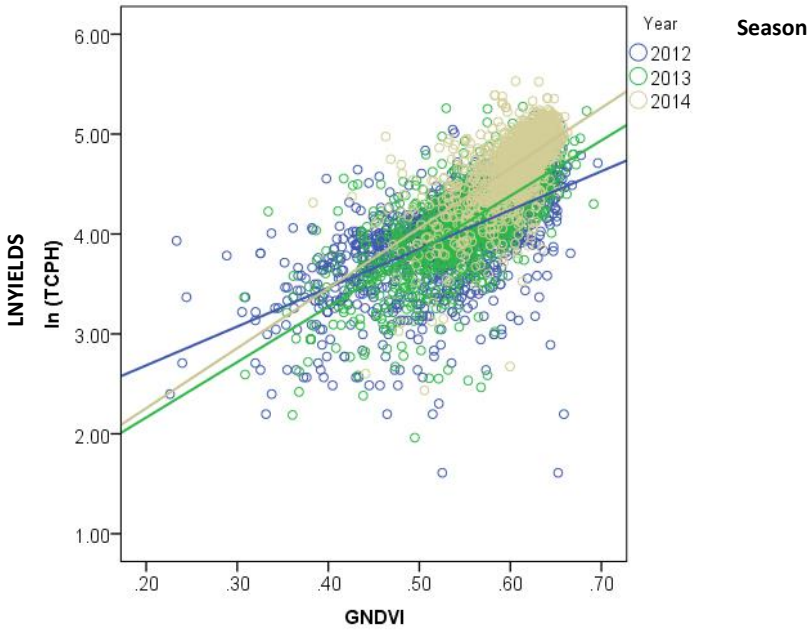


Figure 5 – LN TCPH versus GNDVI for Condong in 2012 (blue), 2013 (green) and 2014 (gold) seasons. Lines represent the regression equation for each season

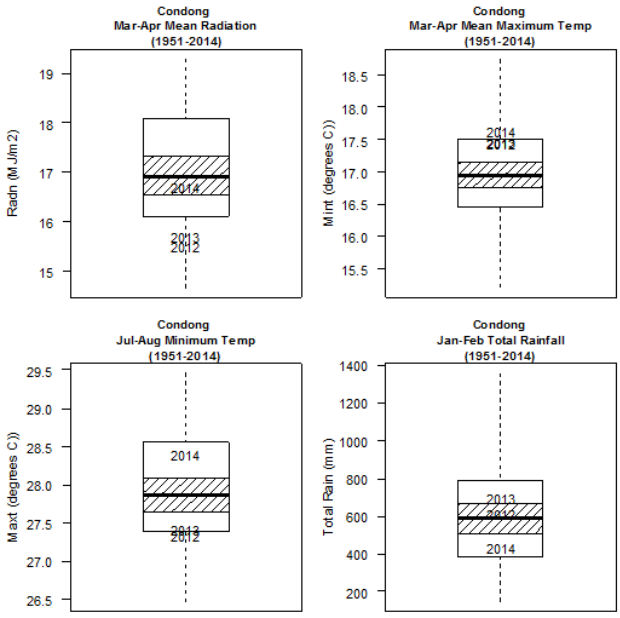


Figure 6 - Key climate parameters for Condong for the period 1951 to 2014 based on random forest variable importance. Boxes represent 25th to 75th percentile, solid black line is median, shaded area is an approximate 95% confidence interval on the long term median

1.3 NDVI and Climate Signals

Everingham et al. (2015) stated "the NDVI can be correlated to recent rainfall (Richard and Poccard, 1998) and even long range climate predictors such as ENSO phase (Anyamba and Eastman, 1996)." Specifically Richard and Poccard (1998) found strong correlations between NDVI measures and "bimonthly rainfall" that is the total rainfall of the two months before the image was taken. However, in their review of the literature Richard and Poccard (1998) point out that the sensitivity of NDVI to rainfall can be reduced in regions that consistently have high or very low rainfall.

Anyamba and Eastman (1996) performed a Principle Component Analysis of time-series NDVI data and found a link between one component and Interannual ENSO Index derived from sea surface temperature anomalies for NDVI in South Africa. Li and Kafatos (2000) performed a similar analysis to Anyamba and Eastman (1996) for the United States of America. Using a Principle Component Analysis (PCA) of the time domain Li and Kafatos (2000) found a link between one principle component and ENSO events based on Southern Oscillation Index.

1.4 NDVI in Agrometeorological Models

These studies suggest that satellite based vegetative indices such as NDVI and GNDVI may reflect the effects of local and long range climate variables. While previous studies have found links between the NDVI-Yield relationship and rainfall variables spatially (Mulianga et al. 2013), our study identified no significant relationship between the GNDVI-Yield relationship (b in the equation yield = $a.e^{b.GNDVI}$) and any of the climate variables found to influence yield. Mulianga et al. (2013) used a weighted time integral of NDVI values to forecast sugarcane yields in Kenya over nine years (2002-2010) for 6 climatic zones. These forecasts were based on MODIS 250 m pixel images. Due to the course/medium resolution of the MODIS data set and highly heterogeneous climate and land use of the regions studied, Mulianga et al. (2013) employed a novel weighting scheme to reduce the noise introduced by pixels that were not growing sugarcane.

High resolution remote images have been used to forecast yields through integration with process-based crop models. Morel et al. (2014) compared four methods of forecasting yields using high resolution images (SPOT-5 and SPOT-4; 10 m resolution), including empirical regression (Yield~NDVI) and forced-coupling of the fraction of absorbed photosynthetically active radiation (*f*APAR) in the MOSAICAS sugarcane crop model. In that study, forcing the MOSAICAS sugarcane model with observed fAPAR data derived from satellite images; improved modelled yields compared to the basic MOSAICAS results (RMSE 12.6 t/ha compared to 15.3 t/ha). Satellite image data can alternatively be used to calibrate parameters in a crop model, rather than forcing parameter values. For example Claverie et al. (2012) calibrated a crop model for sunflower and maize based on Green Area Index (GAI) time series, derived from Formosat-2 NDVI images.

References

Anyamba, A., Eastman, JR. (1996) Interannual variability of NDVI over Africa and its relation to El Niño/Southern Oscillation. *International Journal of Remote Sensing* **17**, 2533-2548.

Richard, Y., Poccard, I. (1998) A statistical study of NDVI sensitivity to seasonal and interannual rainfall variations in Southern Africa. *International Journal of Remote Sensing* **19**, 2907-2920.

Li, Z., Kafatos, M. (2000) Interannual Variability of Vegetation in the United States and Its Relation to El Niño/Southern Oscillation. *Remote Sensing of Environment*, **71**(3), 239-247.

Claverie, M., Demarez, V., Duchemin, B., Hagolle, O., Ducrot, D., Marais-Sicre, C., Dejoux, J-F., Huc, M., Keravec, P., Béziat, P., Fieuzal, R., Ceschia, E., Dedieu, G. (2012) Maize and sunflower biomass estimation in southwest France using high spatial and temporal resolution remote sensing data. *Remote Sensing of Environment* **124**, 844–857.

Everingham, Y., Robson, A., Sexton, J. (2015) A statistical approach for identifying important climatic influences on sugarcane yields. *Proceedings of the Australian Society of Sugar Cane Technologists* 37 (USB).

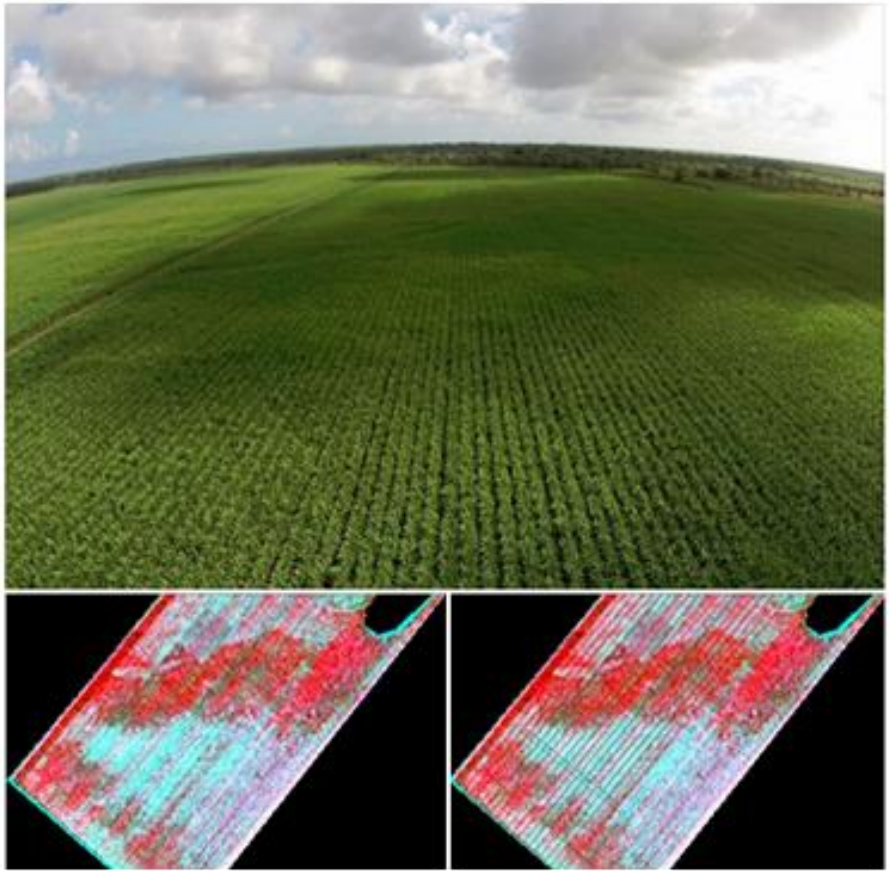
Mulianga, B., Bégué, A., Simoes, M., Todoroff, P. (2013) Forecasting regional sugarcane yield based on time integral and spatial aggregation of MODIS NDVI. *Remote Sensing* **5**, 2184-2199.

Morel, J., Todoroff, P., Bégué, A., Bury, A., Martiné, J-F., Petit, M. (2014) Toward a satellite-based system of sugarcane yields estimation and forecasting in smallholder farming conditions: A case study on Reunion Island. *Remote Sensing* **6**, 6620-6635.

Appendix 6. Extension articles.

Farmacist Burdekin added 3 new photos.
January 20 at 9:44am · Edited ·

Here is some interesting imagery of a trial we are currently working on in the Burdekin with NQDT. The imagery was supplied by Dr Andrew James Robson who is working on an SRA funded project with Digital Globe where they are looking at remotely sensing Nitrogen in crops using WorldView 2 imagery. You cant really see much with the naked eye but the imagery really highlights treatment effects. The third image has plot polygons laid over the top to show treatments. Treatments 1, 6 and 7 which are 220N (Urea, Entec and CR25%) and 8, 9 and 10 are 160N/ha (Urea, Entec and CR25%). Image shown is a false colour image that includes infrared, the brighter the red the more vigorous the growth. We have just finished sampling plots in both zones to ground truth N in plant material. Seriously humid in there this time of year!



May 2016

AROUND THE PADDOCKS

Ag newsletter for Broadwater farmers

RSD

Tom is about to get in to RSD sampling. We got some testing requests back but not too many.

Maybe we asked the question too early.

Now that the season is about to start and you're thinking about plans for planting give Tom a call on 0439 283404 and let him know if there are particular blocks you want sampled.

Yellow Canopy (YCS)

YCS has been positively identified in Maryborough and Bundaberg. It appears it's getting closer to us but so far we have seen no sight of it. Let's hope it stays where it is. Ag staff from Condong met up with Frikkie Botha from SRA a while back to get clued up on the latest developments with YCS. Interestingly SRA say that most varieties will show YCS but it seems that symptoms are very visible in Q240. This doesn't mean that it's any more affected just symptoms are clearer to see.

Breakfast information meetings

The season start date is officially set for Tuesday 31st May. This means it's time for our usual round of (almost pre-season) brekky meetings.

The first meeting will be at:

Pimlico Hall, Pimlico Road

Thursday, 2nd June, breakfast at 7am, meeting start at 8am.

The second meeting will be at:

Dungarubba Hall, Broadwater-Coraki road

Friday 3rd June, breakfast at 7am, meeting start at 8am.

Chris Connors will address the meeting plus mill staff on season plans, harvest logistics plus a range of ag topics.

Come along and get all the latest on the season ahead.

Satellite yield mapping

For the last 5 years we've been working with Andrew Robson who has been leading a sugar industry project on using satellite technology to determine crop yield and crop health. Andrew is based at UNE at Armidale. Andrew's work started in

Soil testing - a good time to start

Soybeans blocks are now clear so now is a good time to slip in

and soil sample prior to cane. Each year presents new challenges but there's a number of what we could call timeless issues in cane farming. Good crop nutrition is one of these issues.

To be honest, some farms out there haven't had a soil test for years and are running a fertiliser program without any idea of soil fertility or crop needs.

Fertiliser is our second biggest input to harvesting costs and a streamlined program based on good information can have a significant benefit to production and profitability.

Qld and so there has been some challenges adapting to 2 year crops. This technology works by measuring the vegetation cover of the crop and then using some fancy maths to determine yield.

Satellite yield estimation came up with these predictions for 2016:

- 92t/ha for 1-year cane (farmers estimate - 94t/ha)
- 130 t/ha for 2-year cane (farmers estimate - 135t/ha).

The prediction is pretty close.

With a few more years info a future plan might be to use the satellite estimate as the first indicator and rather than sending you a blank estimate sheet it will have this as as starting point.

Apart from yield estimating the technology is also being tested for determining crop nitrogen status.

It's possible to use the data to give you a farm map with the yield shown across paddocks. This might assist in identifying poor yielding areas and working out a plan to remedy this.

Yield maps for each mill can be viewed through the website. Go to Agricultural Information and then click on Satellite Estimation.

All the best for the season ahead.

Ag Services || Rick Beattie 0418 162964

Adam **Marshall** MP
MEMBER FOR NORTHERN TABLELANDS

27 June 2016

Assoc. Professor Andrew Robson
20 Aldred Avenue
ARMIDALE NSW 2350

Dear Professor ~~Robson~~ Andrew,

Please find enclosed a copy of a Notice of Motion I delivered in the NSW Legislative Assembly on Thursday 23 June 2016 congratulating you and the other members of the UNE research team on being awarded the President's Medal at the Australia Society of Sugar Cane Technologists conference.

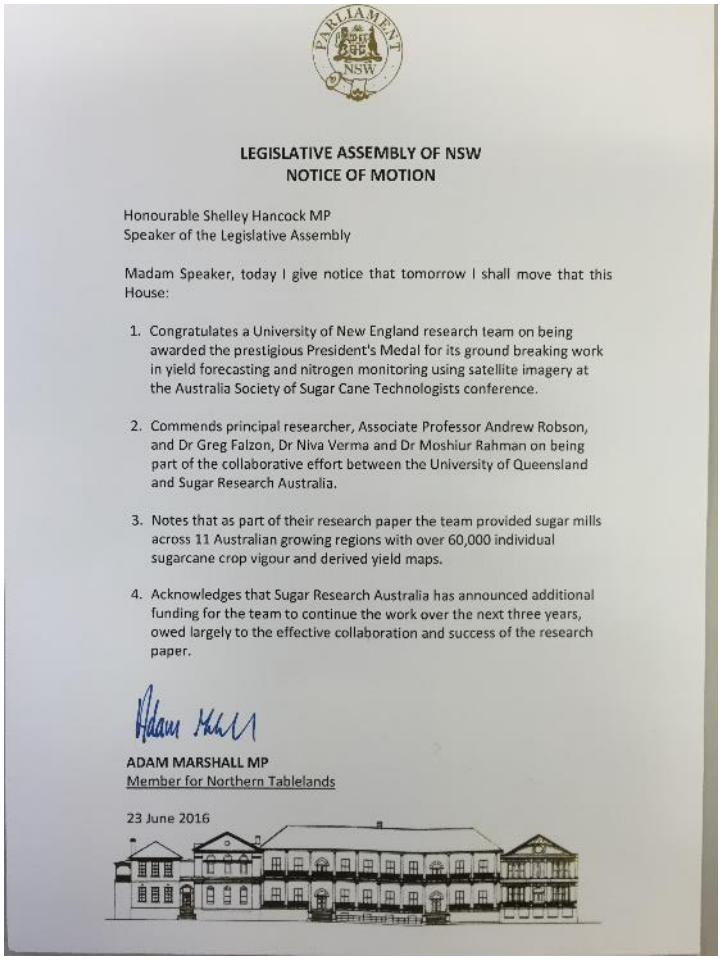
It was an honour to be able to share information about the details of your research paper with my parliamentary colleagues. Well done to all involved!

Congratulations, and I wish you all the best for the future.

Yours sincerely

Great work Andrew!

ADAM MARSHALL MP
Member for the Northern Tablelands
Enc:



LEGISLATIVE ASSEMBLY OF NSW
NOTICE OF MOTION

Honourable Shelley Hancock MP
Speaker of the Legislative Assembly

Madam Speaker, today I give notice that tomorrow I shall move that this House:

1. Congratulates a University of New England research team on being awarded the prestigious President's Medal for its ground breaking work in yield forecasting and nitrogen monitoring using satellite imagery at the Australia Society of Sugar Cane Technologists conference.

2. Commends principal researcher, Associate Professor Andrew Robson, and Dr Greg Falzon, Dr Niva Verma and Dr Moshiur Rahman on being part of the collaborative effort between the University of Queensland and Sugar Research Australia.

3. Notes that as part of their research paper the team provided sugar mills across 11 Australian growing regions with over 60,000 individual sugarcane crop vigour and derived yield maps.

4. Acknowledges that Sugar Research Australia has announced additional funding for the team to continue the work over the next three years, owed largely to the effective collaboration and success of the research paper.

ADAM MARSHALL MP
Member for Northern Tablelands

23 June 2016