Multi-temporal landsat algorithms for the yield prediction of sugarcane crops in Australia

Muhammad Moshiur Rahman[*1], Jasmine Muir [1] and Andrew J Robson[2]

*1 Research Fellow, Precision Agriculture Research Group: Agricultural Remote Sensing Team, University of New England, Armidale, NSW 2351, Australia

2 Associate Professor and Principal Research Fellow, Precision Agriculture Research Group: Agricultural Remote Sensing Team University of New England, Armidale, NSW 2351, Australia

Abstract

Accurate with-in season yield prediction is important for the Australian sugarcane industry as it supports crop management and decision making processes, including those associated with harvest scheduling, storage, milling, and forward selling. In a recent study, a quadratic model was developed from multi-temporal Landsat imagery (30 m spatial resolution) acquired between 2001–2014 (15th November to 31st July) for the prediction of sugarcane yield grown in the Bundaberg region of Queensland, Australia. The resultant high accuracy of prediction achieved from the Bundaberg model for the 2015 and 2016 seasons inspired the development of similar models for the Tully and Mackay growing regions. As with the Bundaberg model, historical Landsat imagery was acquired over a 12 year (Tully) and 10 year (Mackay) period with the capture window again specified to be between 1st November to 30th June to coincide with the sugarcane growing season. All Landsat images were downloaded and processed using Python programing to automate image processing and data extraction. This allowed the model to be applied rapidly over large areas. For each region, the average green normalized difference vegetation index (GNDVI) for all sugarcane crops was extracted from each image and overlayed onto one time scale 1st November to 30th June. Using the quadratic model derived from each regional data set, the maximum GNDVI achieved for each season was calculated and regressed against the corresponding annual average regional sugarcane yield producing strong correlation for both Tully (R2 = 0.89 and RMSE = 5.5 t/ha) and Mackay (R2 = 0.63 and RMSE = 5.3 t/ha). Moreover, the establishment of an annual crop growth profile from each quadratic model has enabled a benchmark of historic crop development to be derived. Any deviation of future crops from this benchmark can be used as an indicator of widespread abiotic or biotic constraints. As well as regional forecasts, the yield algorithms can also be applied at the pixel level to allow individual yield maps to be derived and delivered near real time to all Australian growers and millers.

Introduction

Sugarcane (*Saccharum* spp. L.) is one of the most important agricultural crops to the Australian economy, producing around 4.5 to 5 million tonnes of raw sugar each year (Davis et al., 2007) at a value of around AUS$2 billion (Australian Sugar Milling Council, 2016). The Australian sugarcane industry is located over a 2100 kilometre region between Grafton in New South Wales and Mossman in far north Queensland. Considering the size and value of the industry, it is imperative that accurate and timely predictions of annual production are available not only for monitoring annual variability but also to support well informed harvesting, storage, milling, cane transport arrangements and forward selling decisions.

Currently, estimates of annual production in Australia are provided by visual or destructive sampling techniques, which are time consuming, labour intensive and can produce significant errors if the sample locations are not representative of the greater crop. Numerous studies have demonstrated satellite based remote sensing as an accurate and cost effective method for the yield prediction of sugarcane crops internationally and within Australia (Simões et al., 2005; Abdel-Rahman & Ahmed, 2008; Nascimento et al., 2009; Fernandes et al., 2011; Robson et al., 2012; Rahman & Robson, 2016a, 2016b). The high accuracies achieved are the result of the crop biomass, leaf area index and ultimately crop yield being strongly correlated to the spectral reflectance characteristics of the sugarcane canopy.

One of the benefits of satellite based remote sensing technologies is the large archive of historical imagery available and the high temporal frequency of data collection. This information when interrogated can be used to identify the influence of seasonal or environmental factors such as rainfall, drought, nutrient deficiency and other related factors on annual crop production. Although many studies have reported the use of time series analysis of satellite based remote sensing techniques for yield prediction, this research has been generally limited to crops other than sugarcane (Bastidas-Obando & Carbonell-Gonzalez, 2007; Nascimento et al., 2009; Duveiller et al., 2013; Mulianga et al., 2013; Morel et al., 2014b; Rahman & Robson, 2016a, 2016b). Several researchers have also integrated crop models with the multi temporal remote sensing approach to predict sugarcane yield (Gonçalves et al., 2009; Morel et al., 2012; Morel et al., 2014a).

In a recent study, Rahman and Robson (2016b) developed a new model to predict sugarcane yield for the Bundaberg region using multi temporal Landsat images acquired over a 14 year period (2001 to 2014). A linear relationship was established between model derived maximum GNDVI and actual sugarcane yield (t/ha) with a coefficient of determination (R^2) of 0.69 and RMSE = 4.2 (t/ha). The model was validated over the 2015 growing season and showed an overestimation of only 3.54 t/ha sugarcane. In 2016, model predicted yield was 7.2 t/ha less than the actual yield, an error believed to be the result of late rainfall that occurred after the image capture and subsequent prediction. In this study, time series models for the Tully and Mackay regions are presented.

Methods

The study was undertaken over two North Queensland cane growing regions, Tully and Mackay. The Tully sugarcane growing region is located between longitudes 145.71°E and 146.12°E, and latitudes 17.72°S and 18.27°S, while Mackay is located between longitudes 148.50°E and 149.31°E, and latitudes 20.79°S and 21.58°S. The climate of both Tully and Mackay is tropical with most months of the year experiencing high rainfall with the exception of a short dry season. The average annual temperature in Tully is 28.7°C and average annual rainfall is 4095.1 mm, whereas the average annual temperature of Mackay is 27.3°C and average annual rainfall is 1585 mm (BOM, 2017).

All available Landsat scenes for Tully (Path 95, Row 72 and 73) and Mackay (Path 93, Row 74 and 75) were downloaded via the National Computation Infrastructure Thredds server: http://dapds00.nci.org.au/thredds/catalogs/rs0/catalog.html. Data downloaded included Landsat 5, 7 and 8 surface reflectance imagery and the cloud and cloud shadow mask for each image.

Historic Geographical Information System (GIS) crop boundary data and associated yields were sourced from Mackay Sugar Ltd. and Tully Sugar Ltd. 140 scenes (70 dates) of Landsat imagery was downloaded for the growing period (1st November to 30th June) during the years 2007-2016 for the Tully region, and 300 scenes (150 dates) for Mackay for the years 2005 – 2016. Cloud and cloud shadow within each image was masked, and the two scenes required to cover each respective region were mosaicked together to form one image for each date. Spectral information specific to each cane crop was extracted from each image using the corresponding years GIS Mill boundary layer, followed by the calculation of the average Green Normalized Difference Vegetation Index (GNDVI) value for each region at each image date. The following equation was used to derive GNDVI (Gitelson et al., 1996).

$$GNDVI = \frac{R_{NIR}-R_{Green}}{R_{NIR}+R_{Green}} \qquad (1)$$

where, R_{NIR} and R_{Green} are the reflectance measured in the near infrared and green spectral bands.

The number of pixels affected by cloud/cloud shadow with in each image date was calculated, with those images that presented a value over 30% removed from further analysis. From the remaining images, the extracted average GNDVI value for all crops in each region were plotted against image acquisition date (1st November to 30th June). The subsequent fitting of a quadratic model to this data allowed the annual growth trend of sugarcane to be identified as well as the time of year maximum growth was achieved. The identification of maximum growth is important out come as it suggests the optimal time that a single image capture should be acquired if it is to be used to predict future yield or derive yield maps. Using the vertex form produced by the quadratic model shown in equation (2) the maximum vigour period of a specific year could be calculated by shifting the vertical axis of the curve according to the acquired GNDVI value.

$$Y = -a(x-h)^2 + k \qquad (2)$$

where, "a" is a value in the curve that indicates the curvature of the line, h is the horizontal shift of the curve from x = 0, for any standard quadratic equation $y = ax^2 + bx + c$, $h = \left(-\frac{b}{2a}\right)$ and k is the vertical shift of the curve from x = 0, for any standard quadratic equation $y = ax^2 + bx + c$, $k = \left(\frac{4ac-b^2}{4a}\right)$.

Linear regression analysis was performed to evaluate the relationship between the model derived maximum GNDVI value and sugarcane yield. The root means square error (RMSE) was also calculated using the following equation.

$$\text{RMSE} = \sqrt{\sum(\text{actual-predicted})^2/(n-1)} \qquad (3)$$

where, n is the number of observations.

Results and discussion

For the Tully region, the average crop GNDVI values extracted from all available Landsat images over the 10 year period (2007 – 2016) were plotted on to one time series (1st November to 30th June), (Figure 1). A quadratic model was fitted to the data producing a coefficient of determination (R^2) of 0.80, indicating the Tully region exhibits a relatively consistent annual growth trend irrespective of seasonal variation. From the quadratic curve, the GNDVI value reaches its peak after 145 days (from 1st November), which is consistent with the previous findings of Rahman and Robson (2016b), who reported that sugarcane grown within the Bundaberg region also achieves maximum GNDVI value after 145 days of planting.

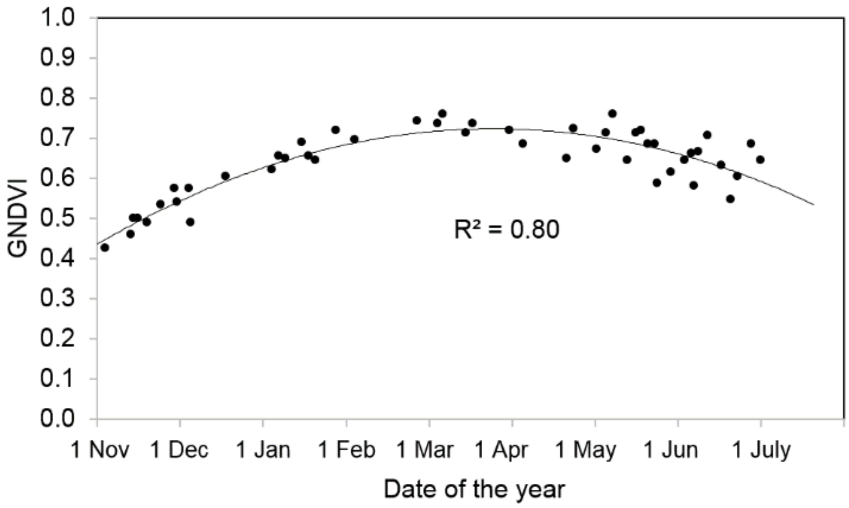

Figure 1. The average GNDVI data for Tully region from 2007 to 2016 (1 November to 30th June) were plotted against image acquisition date over the sugarcane growing season.

The model derived maximum GNDVI values for each year were regressed against the final actual yield (t/ha) achieved for each corresponding year, producing a very strong coefficient of determination ($R^2 = 0.89$) and RMSE = 5.5 (t/ha) (Figure 2). For validation purpose of the model a cloud free image from 12th February 2017 was acquired and used to predict the total regional yield for Tully in 2017. Although the image capture is well before the peak growth (i.e. mid-February) the prediction was highly comparable to that produced by the 2017 mill estimate (pers comm Dale Thomas, Tully Sugar ltd). The accuracy of this prediction will be confirmed after the 2017 harvest.

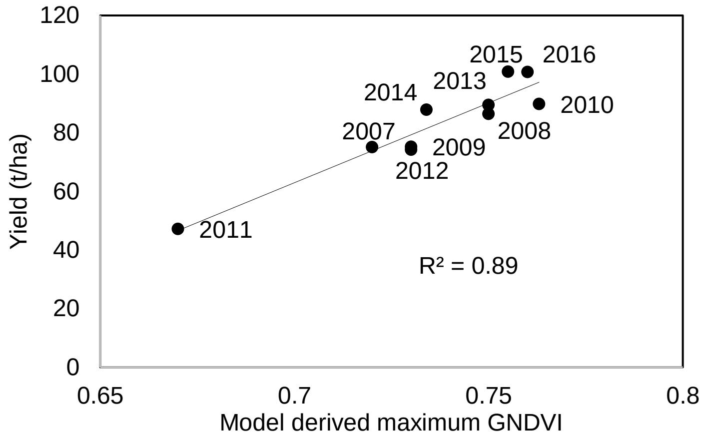

Figure 2. The scatter plot of model derived maximum GNDVI vs annual actual yield (t/ha) from 2007 to 2016 for Tully region.

For the Mackay region, the corresponding extraction of annual crop GNDVI from all available Landsat data over a 12 year period (from 2005 to 2016) (Figure 3) produced a comparably high coefficient of determination (R^2 = 0.89) to that produced for Tully. From this model the peak growth period was identified to be at 155 days after planting.

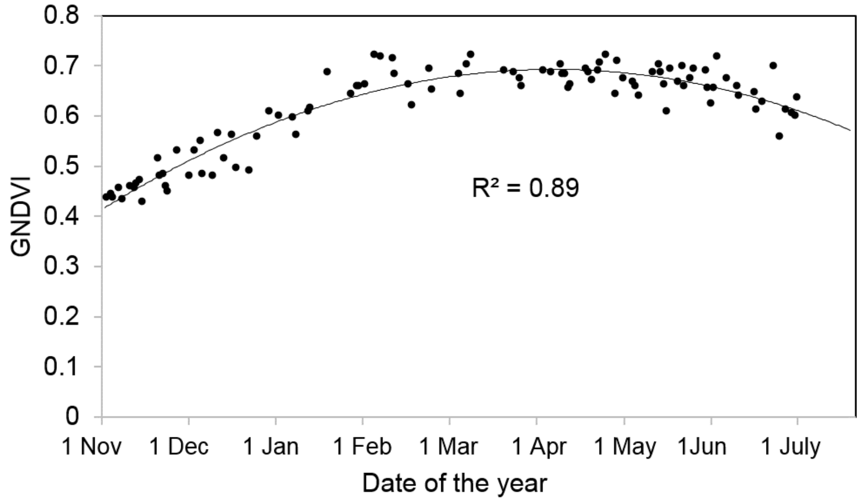

Figure 3. The average GNDVI data for sugarcane crops grown within the Mackay region between 2007 and 2016 (1 November to 30th June) were plotted against image acquisition date over the sugarcane growing season.

A scatter plot of model derived maximum GNDVI for each year against the corresponding average regional yield achieved (t/ha) is shown in Figure 4. A linear relationship produced a lower coefficient of determination (R^2 = 0.63) than that achieved by the Tully model although the RMSE of 5.3 (t/ha) was still encouraging. Using an image from 14th February 2017, the predicted yield for the Mackay 2017 crop was again highly comparable to the growers estimate provided by Mackay Sugar Ltd (pers comm. John Tait, Mackay Sugar ltd), although this estimate was provided before the onset of Cyclone Debbie which passed over the Mackay region on 23 March 2017, bringing widespread flooding and destructive winds. Again the accuracy of prediction will be confirmed following the 2017 harvest.

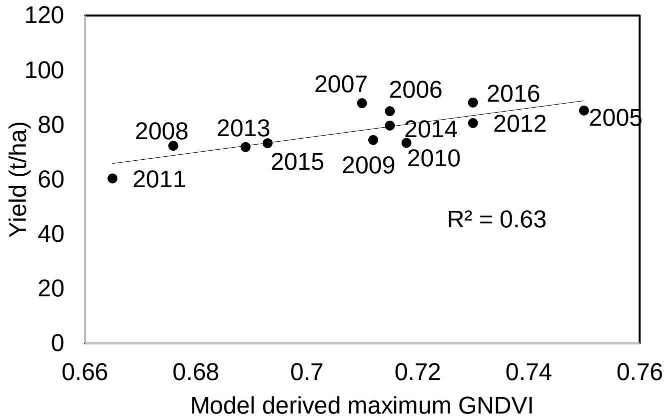

Figure 4. The scatter plot of model derived maximum GNDVI Vs annual actual yield (t/ha) from 2007 to 2016 for Mackay region.

Conclusion

This study has shown that Landsat time series models can be used to plot the annual growth trends of sugarcane development in both the Tully and Mackay growing regions. The composite model developed from these trends provides a strong measure of historic production that can be used as a benchmark for future sugarcane crop growth, potentially serving as an early warning of widespread abiotic or biotic constraints such as diseases, nutrient deficiencies, drought or flooding. For Tully the highest crop vigour (GNDVI) was achieved at 145 days after the first date of planting whilst for the Mackay region this was identified to be at 155 days. The identification of these peak growth times suggested that to obtain the highest yield forecasting accuracies for region from a single satellite acquisition, images should be obtained from late March to Mid-April. The verbal confirmation that yield predictions for both regions were close to 2017 mill estimates indicates the accuracies of the models may be comparable to those achieved by the previously developed Bundaberg model. More importantly these predictions were made early in the growing season i.e. Mid-February which is highly encouraging as it provides industry with a more time to obtain estimates that better support forward selling decisions as well as plan harvest scheduling.

Acknowledgments

The authors acknowledge Sugar Research Australia for funding this research. The authors gratefully acknowledge the assistance of Dale Thomas and Greg Shannon (Tully Sugar Limited) and John Tait (Mackay Sugar Limited) for providing the historical boundary layer map and the yield data.

References

Abdel-Rahman EM, Ahmed FB 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.

Australian Sugar Milling Council. (2016). Sugarcane Statistics. Retrieved from http://asmc.com.au/industry-overview/statistics/ on 10th May 2017.

Bastidas-Obando E, Carbonell-Gonzalez J 2007. Evaluating the applicability of MODIS data for forecasting sugarcane yields in Colombia. Paper presented at the International Society of Sugar Cane Technologists (ISSCT), Durban, South Africa.

BOM 2017. Bureau of Meteorology. Retrieved from http://www.bom.gov.au on 10th May, 2017.

Davis R, Bartels R, Schmidt E 2007. Precision agriculture technologies - Relevance and application to sugarcane production. Paper presented at the Precision Agriculture options for the Australian Sugarcane Industry, Sugar Research and Development Corporation: Brisbane.

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(3): 1091–1116. DOI: 10.3390/rs5031091.

Fernandes JL, Rocha JV, Lamparelli RAC 2011. Sugarcane yield estimates using time series analysis of spot vegetation images. Scientia Agricola 68: 139–146.

Gitelson AA, Kaufman YJ, Merzlyak MN 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment 58(3): 289–298. DOI: http://dx.doi.org/10.1016/S0034-4257(96)00072-7.

Gonçalves RRV, Nascimento CR, Zullo Jr J, Romani LAS 2009. Relationship between the spectral response of sugar cane, based on AVHRR/NOAA satellite images, and the climate condition, in the state of Sao Paulo (Brazil), from 2001 to 2008. Paper presented at the The Fifth International Workshop on the Analysis of Multi-temporal Remote Sensing Images, Groton, Connecticut.

Morel J, Bégué A, Todoroff P, Martiné J-F, Lebourgeois V, Petit M 2014a. Coupling a sugarcane crop model with the remotely sensed time series of fIPAR to optimise the yield estimation. European Journal of Agronomy 61: 60–68. DOI: http://dx.doi.org/10.1016/j.eja.2014.08.004.

Morel J, Martiné J-F, Bégué A, Todoroff P, Petit M 2012. A comparison of two coupling methods for improving a sugarcane model yield estimation with a NDVI-derived variable.

Morel J, Todoroff P, Bégué A, Bury A, Martiné J-F, Petit M 2014b. Toward a satellite-based system of sugarcane yield estimation and forecasting in smallholder farming conditions: A case study on Reunion Island. Remote Sensing 6(7): 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(5): 2184–2199. DOI: 10.3390/rs5052184.

Nascimento CR, Gonçalves RRV, Zullo Jr J, Romani LAS 2009. Estimation of sugar cane productivity using a time series of AVHRR/NOAA-17 images and a phenology-spectral model. Paper presented at the The Fifth International Workshop on the Analysis of Multi-temporal Remote Sensing images, Groton, Connecticut.

Rahman MM, Robson AJ 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 MM, Robson AJ 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.

Robson A, Abbott C, Lamb D, Bramley R 2012. Developing sugar cane yield prediction algorithms from satellite imagery. Paper presented at the 34th Annual Conference of the Australian Society of Sugar Cane Technologists, Palm Cove, Australia.

Simões MdS, Rocha JV, Lamparelli RAC 2005. Spectral variables, growth analysis and yield of sugarcane. Scientia Agricola 62: 199–207.