EVALUATION OF VEGETATION INDICES FOR RANGELAND BIOMASS ESTIMATION IN THE KIMBERLEY AREA OF WESTERN AUSTRALIA

Charity Mundava\textsuperscript{a,b}, Petra Helmholz\textsuperscript{b}, Antonius G.T. Schut\textsuperscript{c}, Robert Corner\textsuperscript{a}, Brendan McAte\textsuperscript{e,d}, David W. Lamb\textsuperscript{b,e}  

\textsuperscript{a}Department of Spatial Sciences, Curtin University, Kent Street, Bentley, Western Australia - charity.mundava@postgrad.curtin.edu.au, Petra.Helmholz@curtin.edu.au  
\textsuperscript{b}Cooperative Research Centre for Spatial Information, P.O. Box 672 Carlton South, Victoria 3053  
\textsuperscript{c}Wageningen University, Plant Production Systems Group, 6700AK, Wageningen, The Netherlands - tom.schut@wur.nl  
\textsuperscript{d}Western Australian Land information Authority, Landgate, 65 Brockway Rd, Floreat, WA 6014 - Brendon.McAte@landgate.wa.gov.au  
\textsuperscript{e}Precision Agriculture Research Group, University of New England, Armidale New South Wales - dlamb@une.edu.au

KEY WORDS: Analysis, Modelling, Estimation, Prediction, Imagery, Landsat

ABSTRACT:

The objective of this paper is to test the relationships between Above Ground Biomass (AGB) and remotely sensed vegetation indices for AGB assessments in the Kimberley area in Western Australia. For 19 different sites, vegetation indices were derived from eight Landsat ETM+ scenes over a period of two years (2011-2013). The sites were divided into three groups (Open plains, Bunch grasses and Spindifix) based on similarities in dominant vegetation types. Dry and green biomass fractions were measured at these sites. Single and multiple regression relationships between vegetation indices and green and total AGB were calibrated and validated using a "leave site out" cross validation. Results indicate that relationships based on single vegetation indices are moderately accurate for green biomass in wide open plains covered with annual grasses. The cross-validation results for green AGB improved for a combination of indices for the Open plains and Bunch grasses sites, but not for Spindifix sites. When rainfall and elevation data are included, cross validation improved slightly with a $Q^2$ of 0.49-0.72 for Open plains and Bunch grasses sites respectively. Cross validation results for total AGB were moderately accurate ($Q^2$ of 0.41) for Open plains but weak or absent for other site groups despite good calibration results, indicating strong influence of site-specific factors.

1. INTRODUCTION

Australian rangelands are economically significant and pastoral enterprises cover large parts of the continent. In the state of Western Australia (WA), rangelands cover 87% of the total land area (2.5 million square kilometres) (DAFWA, 2014). Livestock productivity and ecological health are driven mainly by the cover and composition of the perennial grass component of pastures (DRD, 2011). Rangeland vegetation comprises living and dead plant material in proportions that vary strongly within relatively short distances (Huete and Jackson, 1987). Ground coverage is low with a mixed species composition resulting in high spatial and temporal variability of biomass and dead/green vegetation components (O'Neil, 1996). Pastoralists need to know forage availability to assess carrying capacity and grazing periods.

Above Ground Biomass (AGB) estimates provide an important indication of feed availability and monitoring growth dynamics may give clues to the vegetation condition (Reeves et al., 2001). AGB in this context therefore refers to the total amount of edible biomass (including dead and green components) available to grazing cattle while green biomass refers specifically to green components of AGB.

Remote sensing provides temporal and spatial information on feed resources and makes measurements at large extents possible by reducing laborious field sampling and sample processing procedure (Starks et al., 2006). Remotely sensed vegetation indices represent an integrative measure of both vegetation photosynthetic activity and canopy structural variation that is useful in monitoring, time series analysis, and change detection studies (Huete et al., 2002). They can be used as a proxy for plant photosynthesis and relate to the leaf area index (LAI), determining the capacity of a vegetation canopy to intercept solar radiation.

The objective of this paper is to evaluate the robustness of vegetation indices derived from Landsat ETM+ imagery for AGB assessment in the Kimberley region in Western Australia (WA); a region of significance to Australia’s livestock production (EPA, 2004a). Firstly, relationships between selected vegetation indices and total and green AGB are compared. Secondly, the best performing vegetation indices are combined with rainfall and elevation data in a multiple regression model. Rainfall and elevation data are included as they are key drivers of soil moisture and ultimately plant growth.

The paper is structured as follows: the background section reviews related work in AGB prediction using vegetation indices. The following section introduces the study area (a rangeland cattle station in the Kimberley

*Corresponding author

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.  
doi:10.5194/isprsannals-II-7-47-2014
region of Western Australia) and the methodology. The results section evaluates the potential for the prediction of AGB using Landsat ETM+ imagery. The paper closes with a conclusion and outlook for future research.

2. BACKGROUND

Remote sensing derived, ratio-based vegetation indices have been in use for a long time in AGB estimation and the number of publications is immense (Lu, 2006). A vegetation index’s main intention is to enhance the vegetation signal while minimising solar irradiance and soil background effects (Jackson, 1991). Reflectance data are affected by soil properties (Rondeaux et al., 1996), atmosphere (Huete et al., 2006), viewing angle effects (López Díaz and González-Rodríguez 2003) pixel heterogeneity caused by variation in dead/green ratios, ground cover and leaf angle distribution (Hilker et al., 2008).

For the Landsat ETM+, commonly used indices include various ratios of the NIR and RED wavebands, such as Normalised Difference Vegetation Index (NDVI), Normalised Difference Weighted Index (NDWI), and Soil Adjusted Vegetation Index (SAVI). These have been successfully used for AGB estimation at various scales (Lu, 2006). The most widely applied index in biomass estimation is the NDVI (Anderson et al., 1993; Todd et al., 1998). Although the NDVI has been useful in many studies, it has clear limitations. It is most sensitive to ground coverage with chlorophyll containing leaves (Purevđerj et al., 1998), as present in the vegetative growth phase (Donald et al., 2010), but the response saturates when the ground is completely covered, typically at a LAI of 2 to 3 (Baret and Guyot, 1991).

Previous studies have identified strong relationships between vegetation indices and AGB for grassland ecosystems (Wylie et al., 2002). For instance, Dancy et al. (1986) reported an $R^2$ of 0.72 when deriving ground cover in Botswana rangelands and Beeri et al. (2007) reported a prediction error of 18% for total biomass in moderately grazed pastures due to overestimation by NDVI. However, Beck et al. (1990) found no correlation between green or total biomass and ground derived NDVI. The most probable reason for this was the spatially-variable soil background and the contribution of senescent material.

In the Australian context, relationships between biomass and vegetation indices have been successfully developed for the Mediterranean and temperate zones (Donald et al., 2010). However, little information is available for the North Western tropical zone as many vegetation indices have not been extensively tested. Commonly used indices are not suitable as rangelands are dominated by perennial vegetation and of total ground cover at the landscape scale. This points out that the both near and middle infrared wavelengths are useful in monitoring vegetation in semi-arid environments.

However, relationships between vegetation indices and AGB are species and often location specific (Eisfielder et al., 2011), hence there is the need to evaluate which vegetation indices work in this environment. The main aim of this paper is to explore and verify the use of vegetation indices derived from remote sensing data (Landsat) as well as elevation and rainfall data for the estimation of AGB in the Kimberley region of WA. Indices evaluated in the work are based on findings of the “Pastoral Lease Assessment using Geospatial Analysis project” for rangelands in this region of WA (Robinson et al., 2012) (see table 2).

3. MATERIALS AND METHODS

The study area is the Milne Agri-group owned Liveringa cattle station located in the Kimberley area of WA (Figure 1).

Figure 1: The location of Liveringa station in the West Kimberley area of North Western Australia.

The Kimberley region has a tropical or monsoonal climate in the Köppen climate classification (Köppen,
1923), with a wet and a dry season. The station covers an area of approximate 263 000 ha and is located along the Fitzroy river. It includes several land systems and land types that vary in soil type, plant species composition and land condition. Selected sites covered the seven major land systems and included three sites per land system. These land systems were: Djada, Egan, Calwynyardah, Myrooda, Luluigu, Camelgooda and Wanganut (DAFWA, 2014).

Field data was collected in rangeland pastures in 8 field visits within the 2011/2012 and 2012/2013 seasons, according to the protocol introduced in Mundava et al (2013). At each site location, radials extended 50 m from the centre in North, South, West and East directions. The effective sampling size was 0.785 ha. Direct (destructive) and indirect (non-destructive) sampling was carried out. Direct sampling consisted of the clipping and harvesting of quadrats. Indirect sampling involved calibration against these harvested samples.

Satellite images from the Enhanced Thematic Mapper Plus (ETM+) on-board Landsat 7 were used. Atmospheric correction with an estimate of transmissivity was performed using the Landsat calibration module in ENVI (version 5.0, EXELIS, 2013). The Landsat 7 ETM+ has a Scan Line Corrector (SLC) failure which causes some areas to be imaged twice while no data is recorded for others (Figure 2).

An approximated one arc second (30 m) Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) was produced by Geoscience Australia in 2009. The daily gridded total rainfall (mm) data was obtained from the Australian Bureau of Meteorology (BOM) corresponding to the sampling dates. The spatial resolution of the grids is 5 km x 5 km. The positional accuracy of the observational stations is approximately 1 km. Extraction and processing of rainfall grids was done with ArcMap (version 10.2).

Vegetation indices used in this study are presented in Table 2 and were calculated using the ESRI ArcMap software model builder tool (version 10.2). In order to match the image pixels to the size of the ground represented in the sample sites, index values were bilinearly interpolated from adjacent cells with valid values using a weighted distance method. This step was required because the study sites covered a circular area with a 50 m radius, equivalent to four Landsat pixels.

For the data analysis the assumption of linearity was made. In a first test, separate linear regression models were developed for each vegetation index and total or green AGB measured at all sites. The independent variables are the derived vegetation indices while the dependent variables are the total or green biomass. Sampling sites were further merged into three main groups according to the vegetation type: Open plains (6

### Table 1: Sampling Dates and Landsat passes used in the study

<table>
<thead>
<tr>
<th>Sampling Date</th>
<th>% Cloud</th>
<th>Landsat pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>11-18 December 2011</td>
<td>0</td>
<td>23 December</td>
</tr>
<tr>
<td>18-25 February 2012</td>
<td>7</td>
<td>25 February</td>
</tr>
<tr>
<td>19-26 March 2012</td>
<td>0</td>
<td>29 March</td>
</tr>
<tr>
<td>16-23 June 2012</td>
<td>8</td>
<td>16 June</td>
</tr>
<tr>
<td>30 November - 7</td>
<td>1</td>
<td>10 December</td>
</tr>
<tr>
<td>December 2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-22 February 2013</td>
<td>3</td>
<td>9 February</td>
</tr>
<tr>
<td>12-19 April 2013</td>
<td>0</td>
<td>(13 April)</td>
</tr>
<tr>
<td>15-22 June 2012</td>
<td>0</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table 1: Sampling Dates and Landsat passes used in the study. *72/73 refers to the Landsat rows.

### Table 2: Vegetation indices used in this study (Adapted from Robinson et al (2012)).

<table>
<thead>
<tr>
<th>Vegetation Index</th>
<th>Example studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalised Difference</td>
<td>Tucker et al (1983)</td>
</tr>
<tr>
<td>Vegetation Index (NDVI)</td>
<td></td>
</tr>
<tr>
<td>Land Monitor</td>
<td>Curry et al (2008)</td>
</tr>
<tr>
<td>Normalised Difference Senescent Vegetation Index (NDSVI)</td>
<td>Dean (2005)</td>
</tr>
<tr>
<td>Stress-related Vegetation Index (STVI-1)</td>
<td>O’Neill (1996)</td>
</tr>
<tr>
<td>Stress-related Vegetation Index (STVI-3)</td>
<td>Jafari et al (2007)</td>
</tr>
<tr>
<td>Green +Red band</td>
<td>Wallace and Thomas (1998)</td>
</tr>
</tbody>
</table>

Figure 2: Landsat ETM+ false colour image including a RGB combination of bands 6, 5 and 4 with site locations at Liveringa station from 23 December 2011.
sites), Bunch grasses (9 sites) and Spinifex (5 sites) (Mundava et al., 2013). Open plains refer to sites in areas with periodic flooding with mostly annual grasses, Bunch grasses include mostly perennial grasses and cover sandy Pindan country and other higher terrain while Spinifex sites were dominated by hard or soft spinifex (Triodia spp). In a second test, relationships between the site groups and the vegetation indices were established.

In a third test, relationships including a combination of indices for each site group were evaluated. This time a multiple regression model was used; the dependent variables are green and total AGB and the independent variables considered are all the indices derived from the Landsat data. The three best independent variables were selected in a step wise regression approach. Robustness of the relationships was tested with a “leave site out” cross validation approach. Cross validation is a statistical procedure that tests the predictive ability of a model. The result from the cross validation is a $Q^2$ value, a measure of the predictive ability of the model (Loague and Green, 1991). A correlation is assumed to be satisfying when $Q^2$ values exceed 0.5. Note that $Q^2$ can become negative when predictions are poor. The final test is similar to the third test but elevation and rainfall are included as additional independent variables in the multiple regression model. For the evaluation of all the tests, we consider that a successful correlation is detected when $R^2$ values of between 0.4 and 0.6, occasionally achieving the required threshold value of 0.5. For Open plains, green biomass has a moderate correlation with the vegetation indices STVI_1, SAVI and NDVI with a $R^2$ of 0.6. However Bunch grasses have a low correlation with green biomass for all indices with $R^2$ values ranging from 0.4-0.4. Similar results are obtained for Spinifex ($R^2$ values between 0.2 for green AGB. The correlation of the indices with total AGB was low for all sites with $R^2$ values of between 0 and 0.3

When all indices were combined together in a multiple regression model (test 3), there was no correlation found for total biomass for each of the site groups (Open plains, Bunch grasses and Spinifex). The adjusted coefficient of determination ($R^2$ adj) of 0.26 for Open plains, 0.33 for Bunch grasses and 0.39 for Spinifex.

The tests are summarised in Table 3. The left column represents the test number, the middle column the test configuration and the right column, the number of data samples available for each test.

<table>
<thead>
<tr>
<th>Test no.</th>
<th>Name</th>
<th>Number of test samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Single vegetation indices for 19 sites in a simple linear regression model</td>
<td>152 (19 sites sampled 8 times)</td>
</tr>
<tr>
<td>2</td>
<td>Single vegetation indices for three site groups in a simple linear regression model</td>
<td>O=48, B=64, S=40 (3 groups at 8 time points)</td>
</tr>
<tr>
<td>3</td>
<td>Combined vegetation indices for three site groups in a multiple regression model</td>
<td>O=48, B=64, S=40 (3 groups at 8 time points)</td>
</tr>
<tr>
<td>4</td>
<td>Combined vegetation indices and additional data for three site groups in multiple regression model</td>
<td>O=48, B=64, S=40 (3 groups at 8 time points)</td>
</tr>
</tbody>
</table>

Table 3: Overview of the processed tests where O=Open plains, B=Bunch grasses and S=Spinifex.

4. RESULTS

When all sites are included (test 1), relationships between total or green AGB and selected vegetation indices resulted in consistently weak relationships with $R^2$ values below 0.2. Relationships remained far below the defined threshold of 0.5 and were therefore not useful to assess AGB. When sites were grouped according to vegetation types (test 2), relationships did not improve for total biomass as indicated by a maximum $R^2$ value of 0.3 (Table 4).

<table>
<thead>
<tr>
<th>Index</th>
<th>Green biomass $R^2$</th>
<th>Total biomass $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green + Red</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>STVI_3</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>STVI_1</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>NDSVI</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>SAVI</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>NDVI</td>
<td>0.6</td>
<td>0.2</td>
</tr>
<tr>
<td>Land Monitor</td>
<td>0.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 4: Coefficients of determination for vegetation indices with green and total biomass where O=Open plains, B=Bunch grasses and S=Spinifex. Values equal or above 0.5 are highlighted.

However, there is an improvement in correlation with green biomass for Open plains, with $R^2$ values of between 0.4 and 0.6, occasionally achieving the required threshold value of 0.5. For Open plains, green biomass has a moderate correlation with the vegetation indices STVI_1, SAVI and NDVI with a $R^2$ of 0.6. However Bunch grasses have a low correlation with green biomass for all indices with $R^2$ values ranging from 0.4-0.4. Similar results are obtained for Spinifex ($R^2$ values between 0-0.2) for green AGB. The correlation of the indices with total AGB was low for all sites with $R^2$ values of between 0 and 0.3.
Bunch grasses. The most significant term for the Open plains was STVI3 while for Bunch grasses it was STVI1. However, the validation for total biomass was poor with $Q^2$ values of -0.05 for Open plains, -0.09 for Bunch grasses and -1.74 for Spinifex. For green biomass, Open plains and Bunch grasses had LSO-$Q^2$ values of 0.53 and 0.42. Spinifex had poor $Q^2$ values at -0.22.

When elevation and rainfall information are added to the analysis (test 4), results slightly improve for total biomass for Open plains 0.58 (lower part of Table 5). However, the Bunch grasses and Spinifex values still fall below the defined threshold at 0.37 and 0.40 respectively. For green biomass, the Open plains improved from 0.60 to 0.78 (Figure 3) while for Bunch grasses and Spinifex there was only a slight improvement. Open plains are located in areas of high groundwater tables hence would have higher rates of greening. For all the site groups the most significant terms are the plant water sensitive indices in combination with rainfall and elevation. This shows that rainfall is a strong driver of plant growth in this environment.

The validation for total biomass had $Q^2$ values of 0.41 for Open plains, 0.17 for Bunch grasses and -1.14 for Spinifex. For green biomass, Open plains and Bunch grasses had LSO-$Q^2$ values of 0.72 and 0.49. Spinifex had very low $Q^2$ values at -0.21. The cross validation results for this test were comparable to test 3 with significant improvement for Open plains. This indicates that this more complex model did not significantly improve results for Bunch grasses and Spinifex.

![Figure 4: Clump of Spinifex grass in the field](image)

5. CONCLUSION AND OUTLOOK

In comparing vegetation indices derived from Landsat ETM+ imagery and AGB, in heterogeneous rangeland environments in the Kimberley region of WA, there was no correlation between vegetation indices and green or total AGB when combining all sites and samples in one single relationship. However, results improved for green AGB when sites were grouped in terms of similar botanical composition, i.e. Open Plains, Bunch Grasses and Spinifex in a multiple regression model, but not for total AGB. Open Plains had a correlation of 0.60 while Bunch grasses had 0.56. Spinifex which is predominantly perennial had a very low correlation of 0.19 for green biomass. Calibration results slightly improved for Open plains green biomass to 0.78 and validation results to 0.72 when elevation and rainfall information is added to the model. However validation results did not improve for Bunch grasses and Spinifex. Total biomass models did not provide any robust estimates when validated using the leave site out validation. $Q^2$ values ranged from -1.14 to 0.41. This indicates a strong variability within the sampled sites. A combination of vegetation indices provided robust estimates for green AGB Open plains and Bunch grasses groups and only for total AGB Open plains group. The introduction of environmental information did not further improve validation results.

These results indicate that Landsat vegetation indices are helpful to map green AGB across the most important grazing areas including the vegetation types Bunch grasses and Open plains. However, relationships were not robust for Spinifex and perform poorly for total AGB despite moderately strong calibrations, indicating a very strong influence of site-specific factors. The next step is to test if calibrated relationships of previous seasons enable accurate predictions of future seasons, or if further in-season calibrations are required. The prediction of AGB is important to rangeland farmers for setting stocking limits throughout the grazing periods.

ACKNOWLEDGMENT

The work has been supported by the Cooperative Research Centre for Spatial Information, whose activities are funded by the Australian Commonwealth's...
Cooperative Research Centres Programme. The authors also acknowledge the assistance of Mr Peter O’Brien (Liveringa Station, Milne Agrigroup) and his team. We also want to acknowledge and thank Landgate and CSIRO along with Richard Stovold, Graham Donald, Dr. Norman Santich, Derek Schneider, Vanessa Lonergan, Michael Schaefer, Mark Trotter, Andrew Toovey, Dr. Zheng-Shu Zhou, Elizabeth Hulm, Dr. Waqar Ahmed, Bob McCartney and Andrew Buchanan for their valuable fieldwork assistance and technical support.

REFERENCES


Environment Protection Authority (EPA) 2004a. Environmental protection and ecological sustainability of the rangelands in Western Australia. Position Statement No. 5, EPA, Perth


Köppen, W., (1923) Die Klimate der Erde. Walter de Gruyter, Berlin


