FRACTIONAL GROUND COVER MONITORING OF PASTURES AND AGRICULTURAL AREAS IN QUEENSLAND

Michael Schmidt, Robert Denham and Peter Scarth
Department of Environment and Resource Management
Climate Building
80 Meiers Rd
Indooroopilly QLD 4068
Phone: 07 3896 9297 **Facsimile** 07 3896 9843
michael.schmidt@nrw.qld.gov.au

Abstract

Vegetation cover is a critical attribute of the landscape, affecting infiltration, runoff, water erosion and wind erosion. Ground cover relates to living and non-living materials on the soil surface of the lower strata (<2m). Besides the direct economical value for graziers, varying levels of ground cover have indirect economical and major ecological implications on productivity, land condition and biodiversity. Remote sensing offers one of the few cost effective approaches to monitoring long term ground cover change over large spatial extents. In Queensland's dryland ecosystems, changes in ground cover are generally related to management practices and climate influences. The spatial resolution of Landsat TM and ETM+ (Thematic Mapper and Enhanced TM Plus) imagery provides relevant information to monitor ground cover and to inform natural resource management decision makers.

Three ground cover prediction algorithms are evaluated based on 692 field sites: (i) a regression based bare ground prediction, (ii) multiple logistic regression algorithms, and (iii) a spectral unmixing approach. The estimates of (ii) and (iii) include fractions of bare ground, green vegetation and senescent vegetation. The three models estimate bare ground with a root mean square error of 16.9%, 14.6% and 17.4% and a median average error of 10.1, 10.6 and 11.5, respectively. Although the algorithms were trained in pasture field sites across the State of Queensland, the three approaches show correlations with ground cover field observations in an agricultural area of r^2=0.40, r^2=0.68 and r^2=0.56, respectively. The utility of ground cover production models across pastoral and agricultural land uses requires further consideration for accuracy and operational efficiencies.

Introduction

Dynamic ground cover information is important for soil erosion and nutrient flux estimates into the stream network. These are of particular interest in catchments adjacent to the Great Barrier Reef (Karfs et al. 2009, Queensland Government 2009). Episodic interactions of low ground cover with heavy rainfall, often arising after overgrazing and during periods of droughts or low rainfall, have led to eventual catchment degradation in many Australian

rangelands (Bastin et al. 2009). Ground cover levels may vary due to anthropogenic management of grazing enterprises and agricultural land management practices, or natural changes in seasonal rainfall. Many extensive grazing areas and rangelands in the reef catchments are subject to high climate variability on seasonal, annual, decadal and longer timescales, making management for economic and environmental sustainability difficult (Ludwig et al. 2007).

A recent report (Leys et al. 2009) identified ground cover as a key indicator of land management practices. At both the national and regional levels, there remains a lack of comprehensive and consistent ground cover data at a temporal and spatial scale adequate for monitoring and assessing environmental targets related to soil erosion and land management.

To date, the Department of Environment and Resource Management (DERM) ground cover monitoring program has reported annually on the percentage of ground cover in Queensland based on Landsat imagery using a model described in Scarth et al. (2006). Recent research has resulted in an additional two improved ground cover models. One is based on a linear spectral unmixing approach (Scarth et al. in prep.) and the other is based on a multinominal regression (Schmidt, Denham & Scarth in prep). Both models predict three fractions of surface cover for the ground stratum: bare ground (bare), green vegetation (GV) and dry (or non-green) vegetation (NGV).

A comparison between these three ground cover models is described here using historic field observations in rangelands and field data from a pilot study in an agricultural area.

Data

Remote sensing data

Landsat TM and ETM+ imagery were historically obtained once per annum for Queensland from ACRES (Australian Centre for Remote Sensing). Additional imagery was acquired to coincide with field observations. A large archive of more than 2500 Landsat images is held by DERM (Trevithick & Gillingham 2010).

All freely available (largely cloud free) Landsat TM and Landsat ETM+ imagery from the United States Geologic Survey (USGS) EarthExplorer website (http://edcsns17.cr.usgs.gov/EarthExplorer/) covering Queensland have recently been downloaded and added to the DERM image archive. This includes more than 15,000 Landsat images and means that on average about 20 image dates per year from 1999 to 2010 are available for each scene (see Figure 1). This includes Landsat ETM+ with scan line corrector “off” imagery (http://landsat.usgs.gov/products_slcoffbackground.php).

The Landsat TM (Thematic Mapper) and ETM+ (Enhanced TM Plus) images (of both sources) were stored in a raw stage, but the higher level products have been corrected using methods developed and implemented by the SLATS project (http://www.derm.qld.gov.au/slats/). All images were pre-processed and subjected to rigorous radiometric corrections as per standard DERM remote

sensing methods (de Vries et al. 2007) and were standardised for surface reflectance and atmospheric effects. Standard masks for cloud, cloud shadow and water were applied to the ACRES imagery, while preliminary cloud, cloud shadow and water body masking (Muir, Danaher 2008) processes were implemented for USGS data (Schmidt & Trevithick 2010).

Field data

Ground cover field data were collected in across a broad range of land types in Queensland over more than a decade. A transect point-intercept method was used in grasslands and improved pastures with the same field methodology. Sample sites were chosen to be homogeneous and to represent a 100m x 100m area. Descriptions of general site details are performed according to the method described by Tongway and Hindley (1995).

All field data were collected using a modified discrete point sampling method of 300 individual measurements at each metre along three 100m tapes. For this study, the site measurements were summarised in percentages of the ground cover components: bare ground (including rock and cryptogam), green vegetation and dry vegetation (including litter and dead plant material). In the midstorey (woody plants <2m) and overstorey (woody plants >2m) stratas, components of dry/dead leaf, branch and green leaf were recorded also. For a detailed description of the field methodology see Schmidt, Muir and Scarth (Draft). For ground cover studies we selected field data which was collected from sites with low Foliage Projective Cover (FPC) in the overstorey, typically less than 15%.

Figure 1 gives an overview of the location and spatial distribution of field data sites, and the USGS data coverage for Queensland according to the World Reference System II path/row for Landsat imagery (http://landsathandbook.gsfc.nasa.gov/handbook.html).

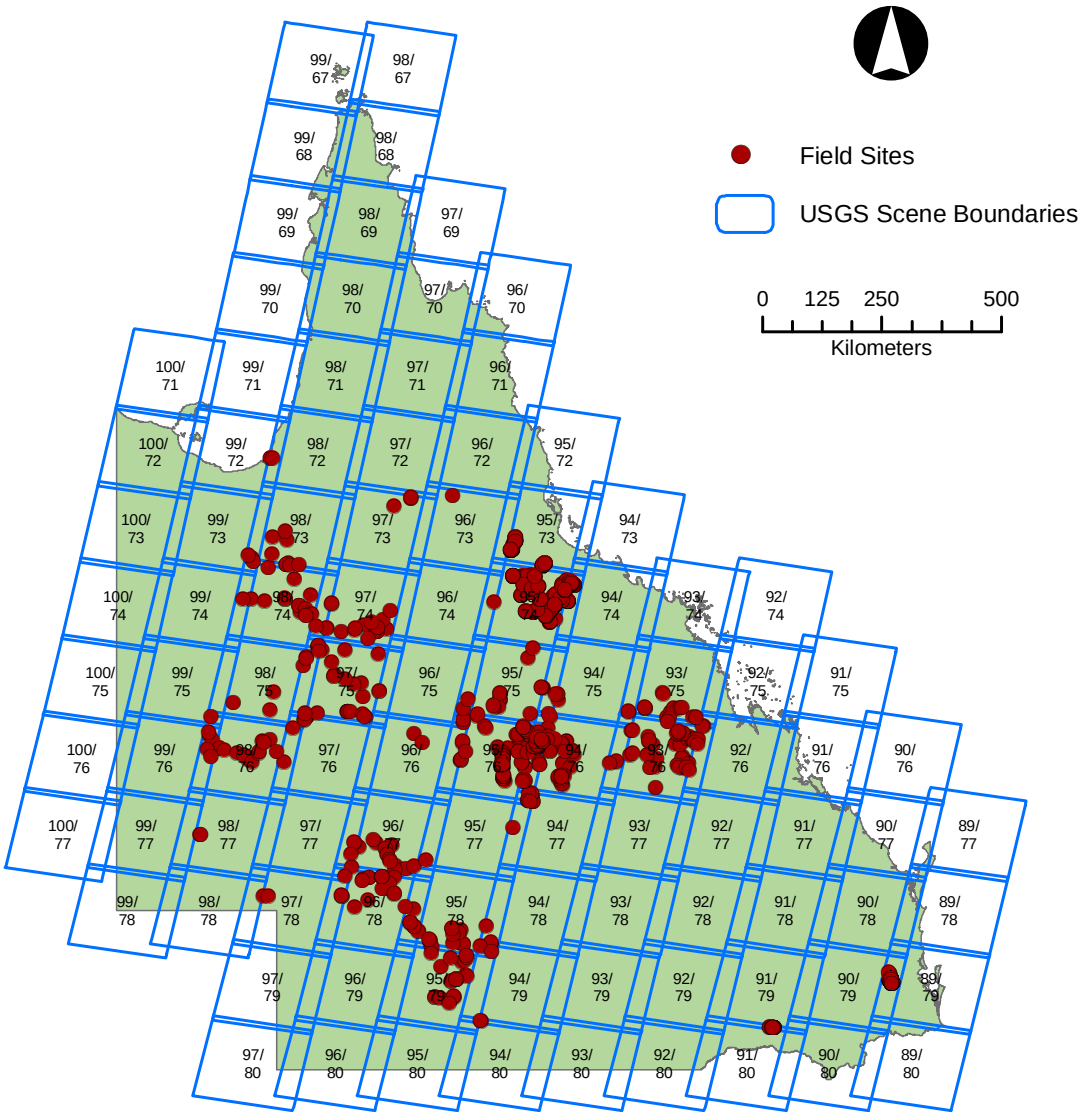

Figure 1: Location of ground cover field observations in Queensland and USGS data coverage in path/row; some sites were repeatedly visited.

Three field work campaigns were undertaken at different crop growth stages in an agricultural area near Goondiwindi (data point in scene 91/80) in 2009 with a modified sampling strategy of two 100m tapes in a 45-degree angled tape layout across linearly sown crops, as suggested by Schmidt et al. (2010).

Methods

Ground cover predictions of three algorithms were compared independently against field observations. For each of the field observations the Landsat image observation with the closest date was used, with a time difference of no more

than 60 days. A brief description of the three methods and previously derived error statements follows.

(i) Model 1 is based on a linear regression approach as described by Scarth et al. (2006). The approach estimates the amount of vegetative ground cover as Ground Cover Index (GCI) as a percentage, or its inverse bare ground as Bare Ground Index (BGI) using data from 431 field sites in Queensland with a reported RMSE of 13.8%;

(ii) Model 2 uses a multinominal logistic regression as described in Schmidt, Denham & Scarth (in prep). Three fractions are predicted as percentages: bare ground, GV and NGV – which sum up to 100%. 512 field sites in Queensland were used with a reported RMSE of 14.9 % in this approach;

(iii) Model 3 estimated ground cover fractions via a (linear) spectral mixture analysis (SMA) as described by Scarth et al. (in prep.). Three fractions of ground cover are predicted as: bare ground, GV and NGV plus an error term of the SMA, summing up to 100%. This approach used 577 field sites in Queensland and additional 83 sites in New South Wales with a reported RMSE of 11.8%.

For this analysis the root mean squared error (RMSE) of the observed and predicted values was reported as well as the median average error (MAE) as a more robust (less sensitive to outliers) description. This was performed on 692 field sites which are now available for grasslands and improved pastures in Queensland. Data from more than 110 additional site data, which were not part of any model calibration, were used for validation. All available field data were used without the removal of outliers or bad data.

A second analysis step was performed in an agricultural area based on 25 field observations during three field campaigns.

All three ground cover models were developed on the basis of Landsat data from the ACRES which have a slightly different data calibration compared to the USGS data. A data inter-comparison between the two data sources for five different land-use types was performed for Landsat bands 3, 5 and 7 which were used in Model 1.

Results

Rangeland sites

The first row in Figure 2 shows scatter-plots of field observations and bare ground predictions of the three ground cover models. The plots show regression lines as well as the RMSE and MAE per model. Only data from the ACRES archive were used here for all three model comparisons.

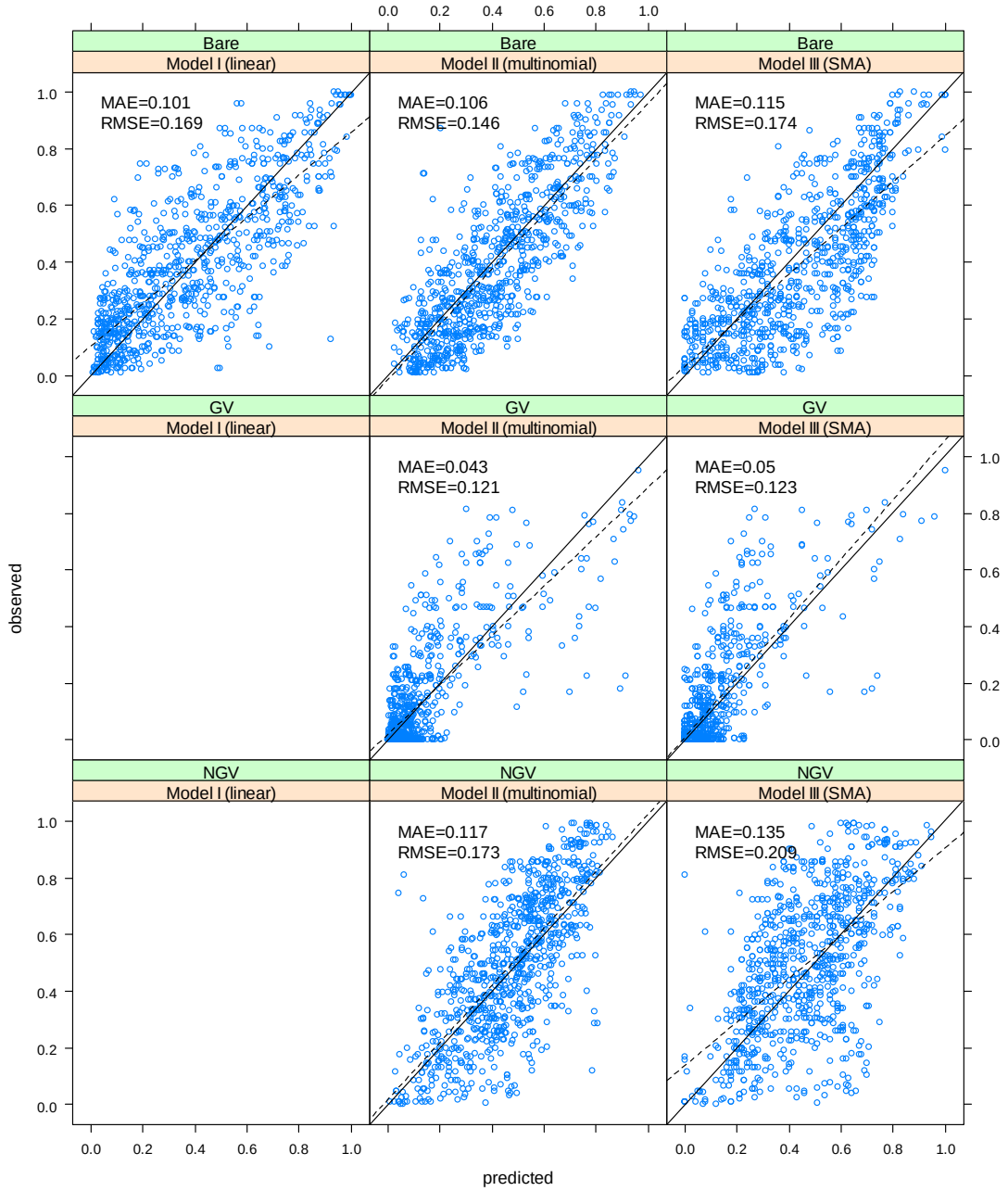

Figure 2: Comparison for three different ground cover models with field observations. Line 1 compares the ground cover estimate of all three models, while in lines 2 and 3 the GV and NGV predictions of Model 2 and 3 are shown. The solid line is a 1:1 line, the dotted line a regression.

The error statistics as well as the data scattering around the plotted regression line show that all three models give useful estimates of bare ground with MAE of 10.1% in Model 1, 10.6% in Model 2 and 11.5% in Model 3. The RMSE of 14.6% is lowest in Model 2, with less data scatter occurring. The regression line in Model 2 is also very close to the 1:1 of field observations and predicted bare ground, while the regression lines in Model 1 and Model 3 have a slightly different slope and bias.

Row 2 shows the observations and predictions for GV, both Model 1 and 2 have similar errors with an RMSE of 12.1% and 12.3% RMSE respectively and 4.3% and 5% for the MAE. Model 2 seems to have a slightly greater deviation from the 1:1 line. The predictions for NGV seem to work better in Model 2 than in Model 3 with an RMSE of 17.3% compared to 20.9 % in Model 3, also the MAE (11.7%) is lower in Model 2 than in Model 3 (13.5%). The point cloud seems to represent the data in Model 2 well with a regression very close to the 1:1 line, whereas there is a different slope and bias for the predictions of Model 3. Figure 3 shows the residuals of Figure 2 summarised in deciles.

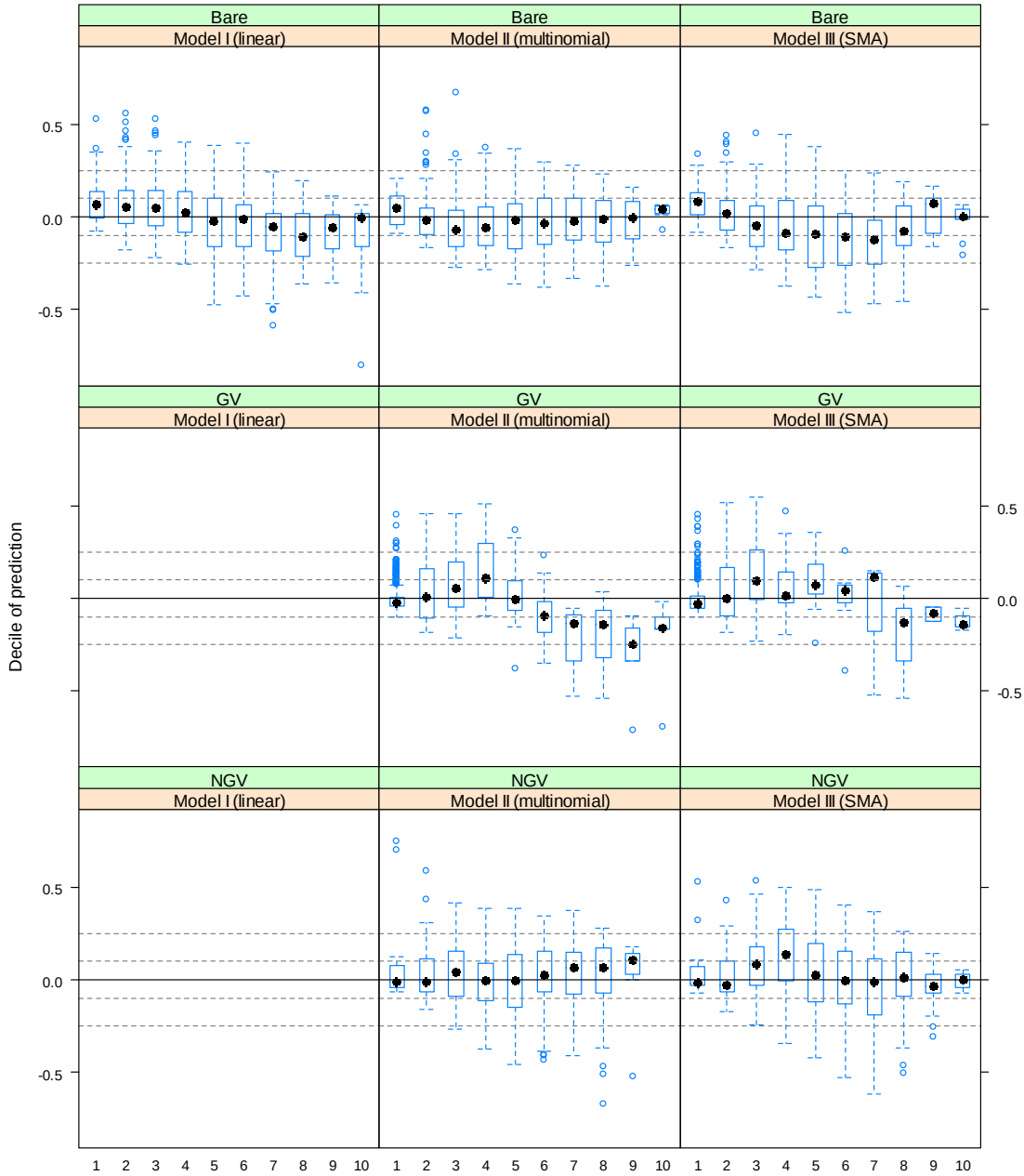

Figure 3: Comparison of three different ground cover prediction models and field observations as box and whisker plots of deciles. In row 1 the ground cover estimates of all three models are compared, while rows 2 and 3 show the GV and NGV

predictions of Model 2 and 3 are shown (the doted lines are plotted at the 0.1 and 0.25 values)

Figure 3 shows box and whisker plots of the three model predictions. The residuals of the bare ground predictions in Model 1 and 2 are very evenly distributed and close to zero, while Model 3 shows some systematic deviation from zero with increasing standard deviation in the deciles 4 to 7 with the highest deviations. The range of the residuals appears lowest in Model 2. Some data outliers are plotted, which might be due to artefacts of cloud and cloud shadow masking.

The GV estimation seems to show a systematic behaviour of the residuals per decile in Model 2 and to a lesser extent in Model 3. The residuals per decile for NGV in both Model 2 and 3 do not appear to have a systematic behaviour, with a lower range and standard deviation in Model 2.

Agricultural areas

An ACRES and USGS Landsat data comparison for different land cover types for the Landsat spectral band used for the BGI in model 1 are shown in Table 1.

Table 1: Differences in BGI for ACRES and USGS data in Landsat bands 3, 5 and 7 top of the atmosphere reflectance (used in Model 1), as well as the BGI.

	ACRES				USGS				BGI
Description	**b3**	**b5**	**b7**	**BGI [%]**	**b3**	**b5**	**b7**	**BGI [%]**	**Difference**
Pasture	69	149	86	79	66	145	87	.83	**4**
Pasture	31	117	39	01	28	113	37	.01	**0**
Pasture	69	148	83	73	67	143	82	.75	**2**
Open woodland	42	104	55	22	39	100	56	.23	**1**
Open woodland	76	143	56	11	72	137	56	.16	**5**
Riverbed (grassy)	73	142	84	86	74	146	88	.89	**3**
Shrubs	33	118	52	03	31	114	51	.02	**1**
Agricultural field	30	56	34	86	29	56	33	.79	**7**
Claypan	103	193	108	94	97	187	108	.96	**2**

The differences between the two Landsat data sources used for BGI calculation in Model 1 were 5% or less in natural environments and had a maximum difference of 7% in an agricultural site. The differences in BGI were below the model accuracy (Scarth et al. 2006) and the data was thus deemed adequate for further model comparison. For this analysis only cloud free USGS data were utilised.

Field work data from three campaigns near Goondiwindi (QLD) were collated at three different growth stages of a wheat field and a fallow field of Sorghum and a fallow wheat field. The utilised wheat field (Site 2 in Figures 4 to 6) was visited firstly after ploughing (May 2009), secondly during greening up (August 2009) and thirdly after harvesting (October 2009).

Figure 4 shows a spatial representation of the utilised fields of the property ‘Monte Christo’ in a true colour Landsat TM 5 image, the three ground cover

models and field photos. Some of the field sites from May 2009 (near photo 4) where 'short' transects (10m) for test purposes and excluded from this analysis (Schmidt et al. 2010). So that 13 full transect field sites were used in the following, including two data points from the site near photo point 1 in Figures 4 and 5 from an improved pasture site.

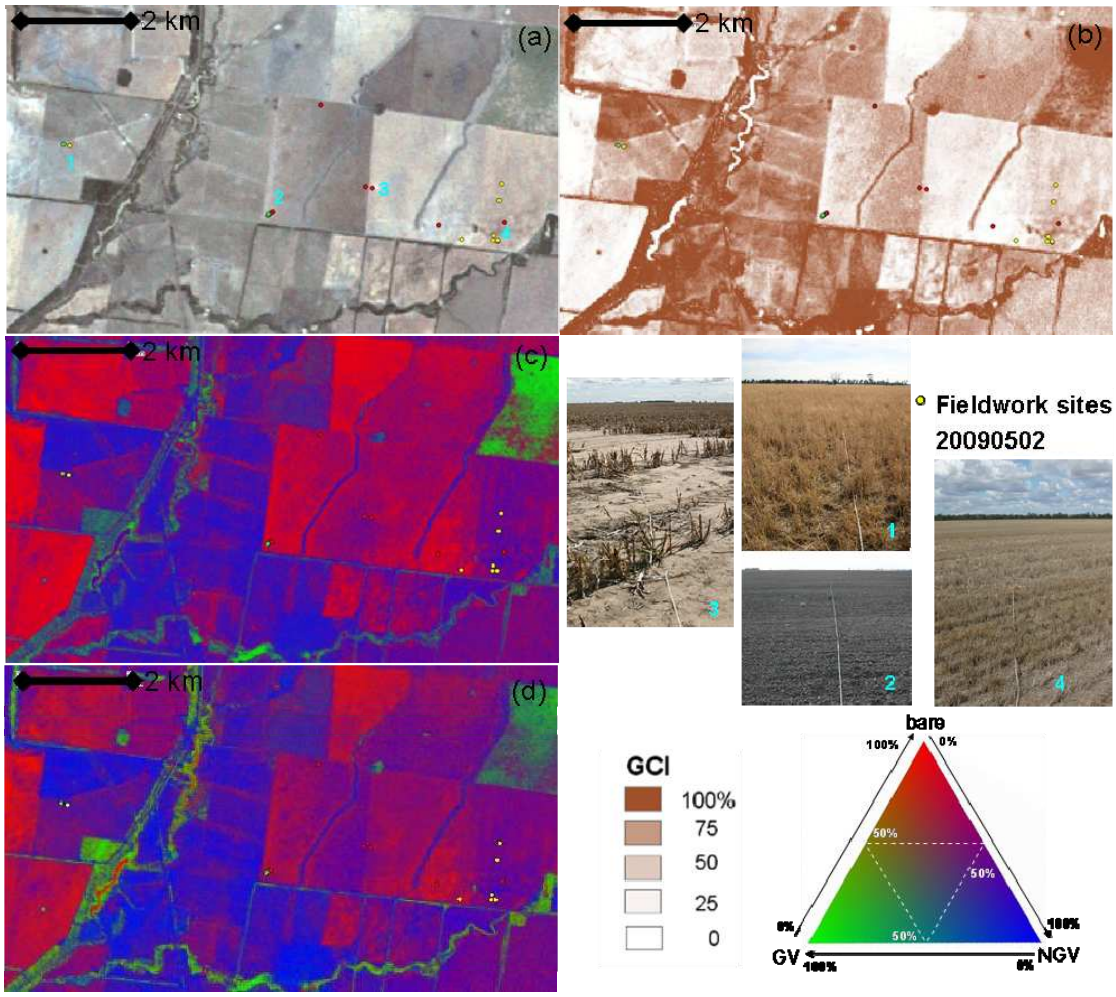

Figure 4: Spatial representation of fieldwork in May 2009 on an agricultural site near Goondiwindi: a) Landsat true colour composite showing field site and field photo locations; b) Model 1 - Ground Cover Index (GCI) as described in Scarth et al. (2006); c) Model 2 - fractional cover descriptions of bare, GV and NGV as described in Schmidt, Denham & Scarth, (in prep); and d) Model 3 - fractional cover descriptions of bare, GV and NGV as described in Scarth et al. (in prep.). Landsat TM image observation date: 2009/04/30.

Figure 4 shows a better differentiation within and across field in Models 2 and 3 than in Model 1. Models 2 and 3 estimate noticeably different greenness intensities in one of the fields and also in the riparian vegetation. The freshly ploughed field in photo 2, with 94% observed bareness is predicted as being 62% bare by Model 3, 70% by Model 1 and 76% by Model 2.

Figures 5 and 6 show the same area at a different crop growth stage and time of year.

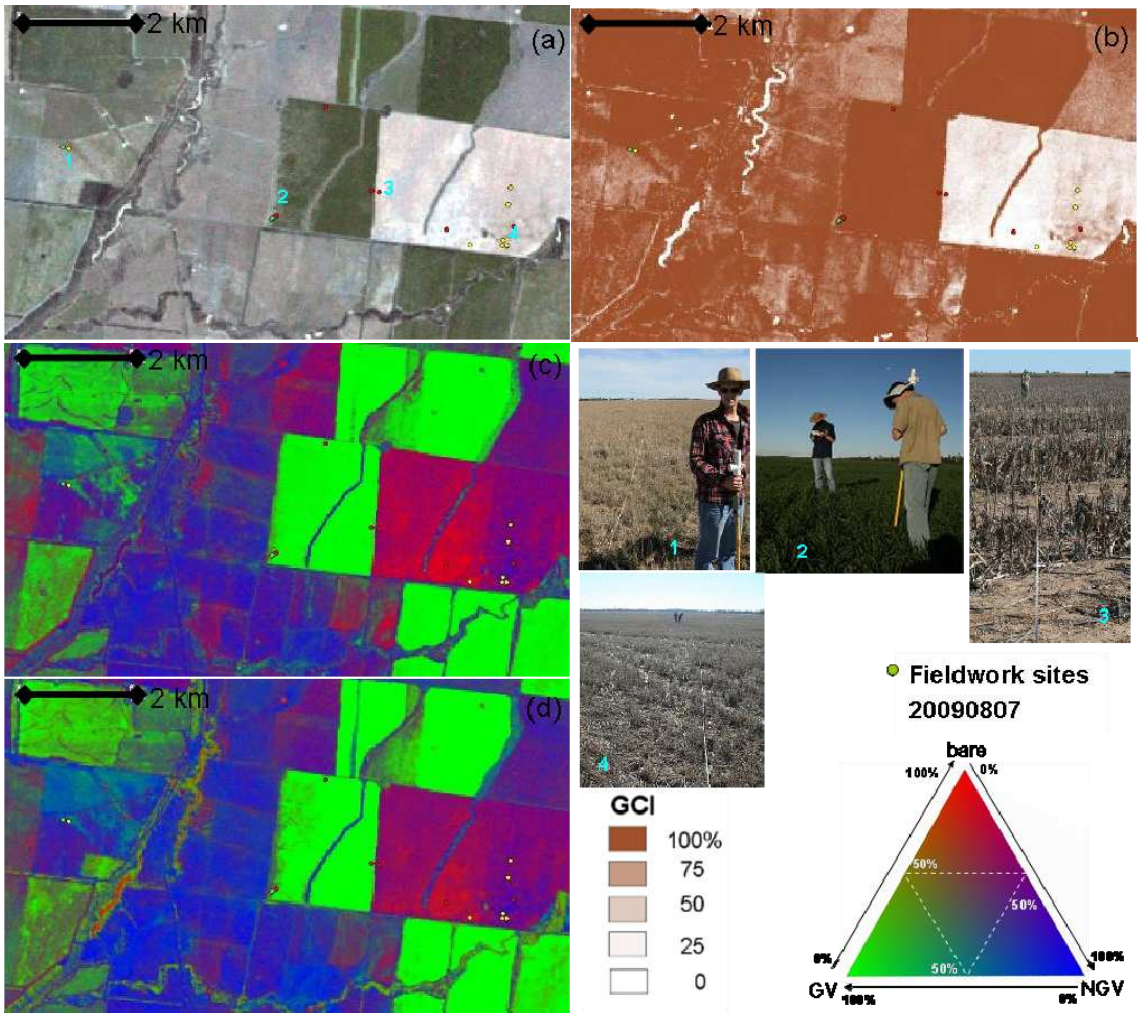

Figure 5: Spatial representation of fieldwork in August 2009, on an agricultural site near Goondiwindi: a) Landsat true colour composite, indicating field site and field photo locations; b) Model 1 - Ground Cover Index (GCI) as described in Scarth et al. (2006); c) Model 2 - fractional cover descriptions of bare, GV and NGV as described in Schmidt, Denham & Scarth (in prep); and d) Model 3 - fractional cover descriptions of bare, GV and NGV as described in Scarth et al. (in prep.). Landsat TM image observation date: 2009/08/04.

Some of the fields in Figure 5 show active cropping (near Photo 2), while other fields (near Photos 3 and 4) remained unutilised with very little change in ground cover.

Figure 6 shows the same area after harvesting. It is noticeable that the harvested wheat stubble (near Photo 2) shows a different cover than the older wheat stubble (near Photo 4). Observations: 65.5% cover (65% NGV plus 0.5% GV), – Model 1: 60% cover, Model 2: 63% cover (56% NGV plus 7% GV), Model 3: 72% cover (65% NGV plus 8% GV).

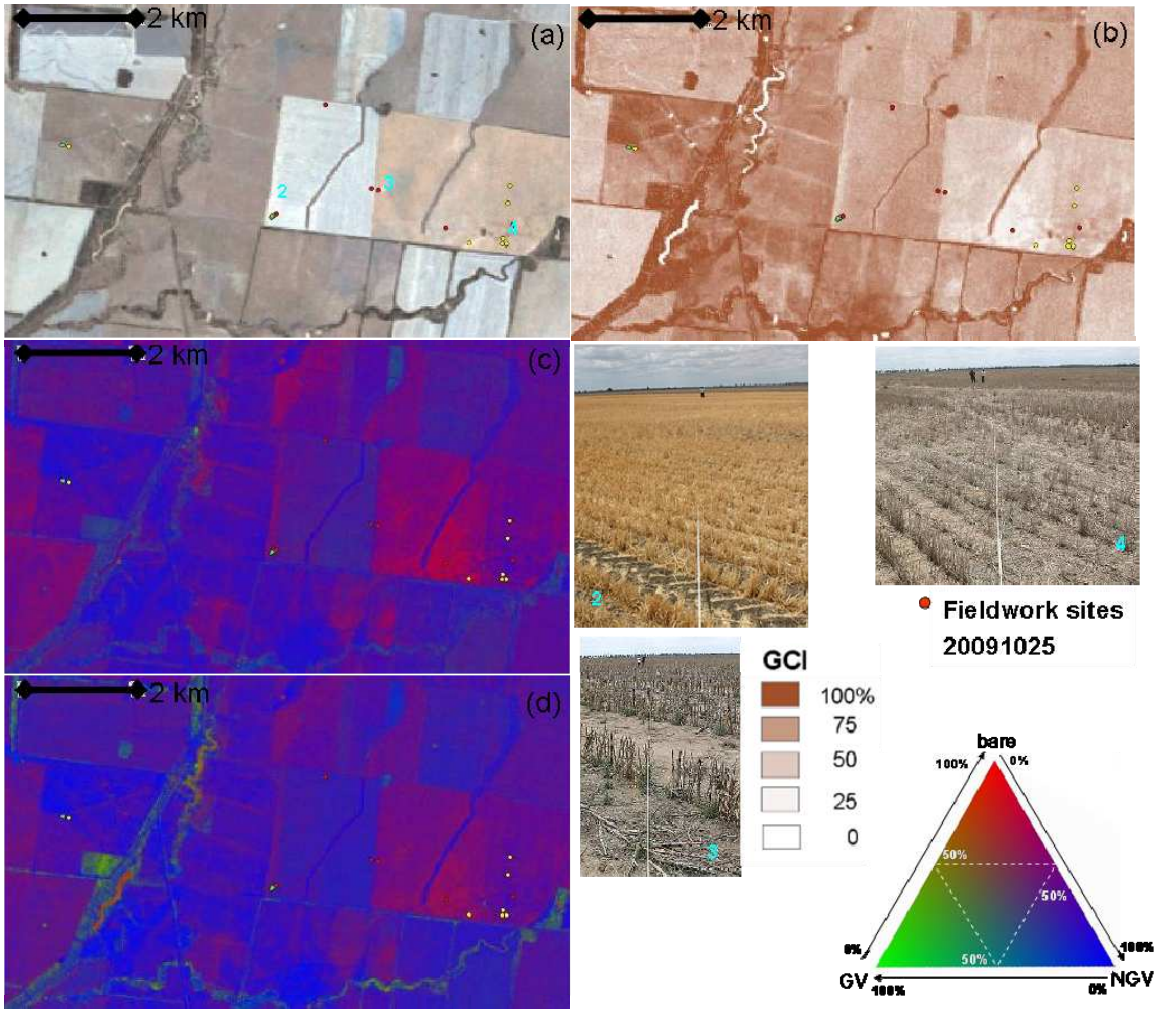

Figure 6: Spatial representation of fieldwork in October 2009 on an agricultural site near Goondiwindi: a) Landsat true colour composite, indicating field site and field photo locations; b) Model 1 - Ground Cover Index (GCI) as described in Scarth et al. (2006); c) Model 2 - fractional cover descriptions of bare, GV and NGV as described in Schmidt, Denham & Scarth (in prep); and d) Model 2 - fractional cover descriptions of bare, GV and NGV as described in Scarth et al. (in prep.). Landsat TM image observation date: 2009/10/23.

A regression plot of all 13 observed field sites for the agricultural area is shown in Figure 7. The three cover predictions of bare, GV and NGV are plotted separately for Models 1, 2 and 3.

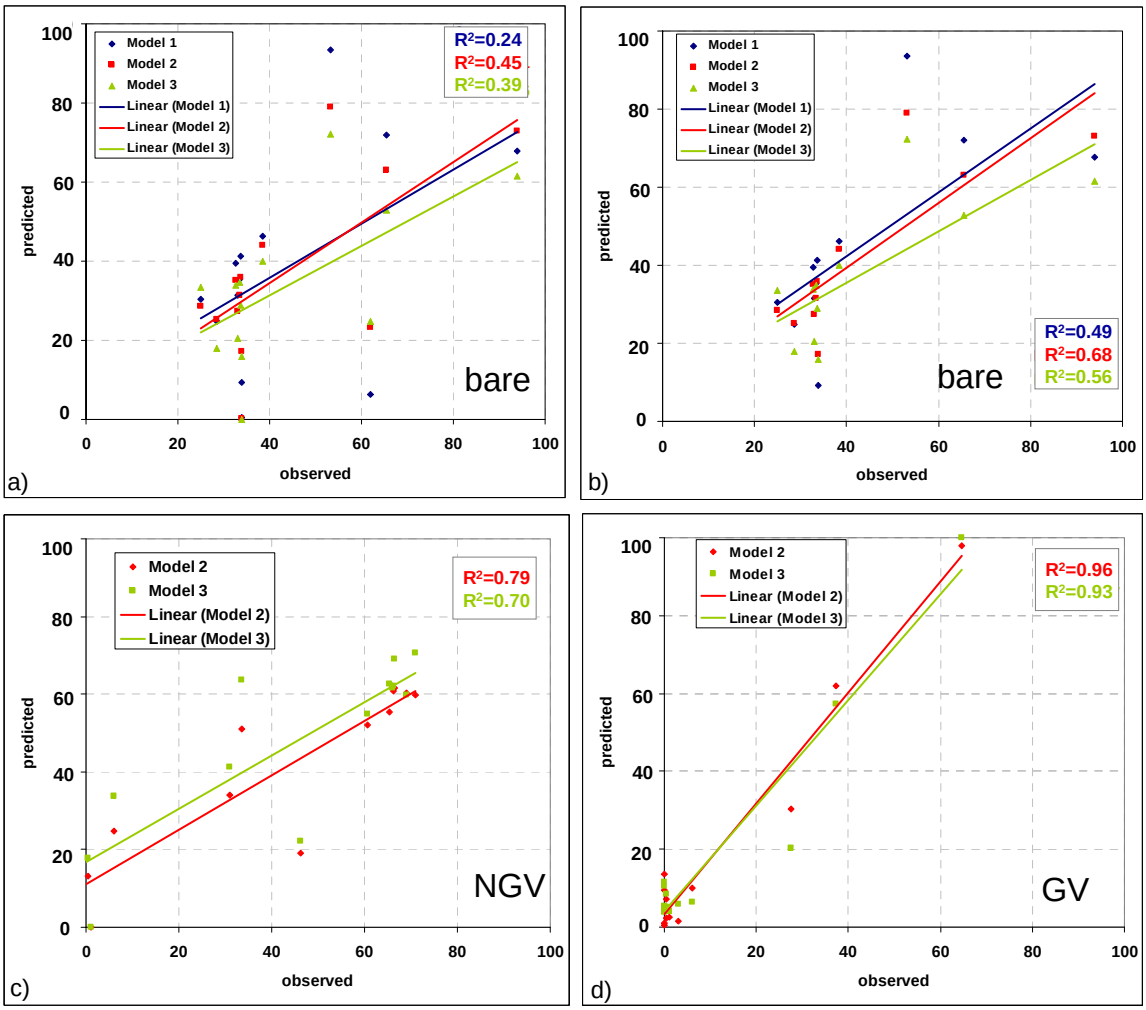

Figure 7: Model comparison for the agricultural field sites near Goondiwindi for three different ground cover prediction models for the fractions bare (a), NGV (c) and GV (d); (b) shows the bare fractions of all models without the two sites with entirely green vegetation coverage.

The bare component shows relatively low R^2 values in all three models: 0.24 in Model 1, 0.45 in Model 2 and 0.39 in Model 3. This might be because the training sites for all three models were chosen in pastoral environments. In fact, two field sites were measured during the greening up phase in August 2009 with no NGV components, which is very different from natural or pastoral environments. The observed cover components were entirely composed of a lush green wheat crop, growing in a black, ploughed soil. Not only has it been proven difficult in estimating cover in black soils, but also the greenness is outside the range of our training data. When these two sites were taken out of the field observations (Figure 7 b) the R^2 values for observed and predicted data increased to 0.49, 0.68 and 0.56 for Models 1, 2 and 3 respectively.

The plot with the NGV data (Figure 7 c) shows a much better fit with R^2 values of 0.79 and 0.70 for Model 2 and 3, respectively, which both seem to appear with an offset value in the linear regression.

The R^2 values of the GV components in Models 2 and 3 (Figure 7 d) are with 0.96 and 0.93 high although a clustering of low data points limits the interpretability of this plot.

Discussion and conclusions

In grasslands and improved pastures all three models seem to predict bare ground (or its inverse ground cover) well, with a MAE of 10.1 in Model 1, 10.6 in Model 2 and 11.5 in Model 3. The scatter in the data in Model 2 is slightly lower than in Models 1 and 3. The bare ground predictions in Model 2 are less biased. Model 2 represents GV slightly better than Model 3 while Model 3 seems to have lower bias in the predictions. NGV is better represented in Model 2 than in Model 3 with almost no bias.

The RMS errors in all three models are higher than the reported values in the model calibration; this can be seen as a result of several factors:

- a large time difference of up to 60 days between field and image observations was allowed;
- no weighting function for the time difference was applied;
- no outlier removal was applied;
- additional field sites were used as compared to the model calibration;
- some of the additional field data may have (depending on the field operator) differences in the separation between green and non green vegetation in the field observations which can lead to false estimated in these components.

The error estimates presented here should be seen as a (very) conservative error estimate. Figure 8 shows an improved plot with a maximum time difference of 15 days between field and image observations (351 data points). The scatter is much reduced with improved error statistics particularly in Model 3 (14.6%). Further analysis of the Model sensitivities is underway.

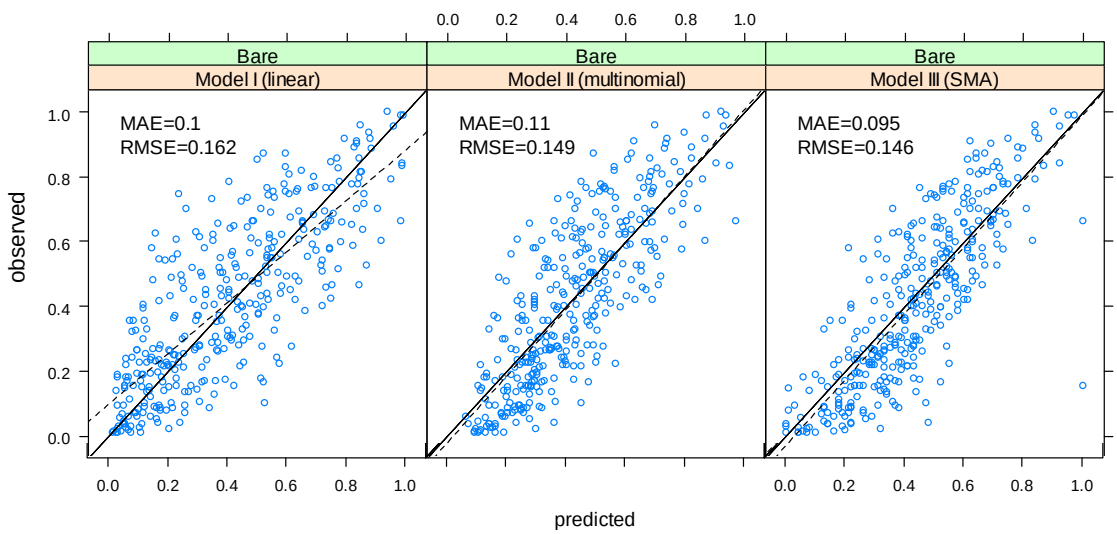

Figure 8: Comparison for the bare ground component with field observations with a maximum of 15 days time difference between field and image observation (see Figure 2).

The residuals per decile of all three models (Figure 3) show slight systematic errors in some fractions, which could potentially be improved with better model fits. The errors in the estimates can potentially be reduced by using USGS imagery in the analysis, as the time difference between image observation and field observation for a fast changing surface component, such as ground cover, is important. With the United States Geologic Survey (USGS) opening their Landsat image archive freely to the public, all archived images (after 2000) are now available. This coverage of up to 20 images per year creates new opportunities for Landsat based monitoring approaches and will potentially reduce the errors in the ground cover model.

The differences between the two Landsat data sources (ACRES and USGS) in Model 1 were 5% or less in natural environments and in an agricultural site, with the maximum difference of 7%, found to be acceptable for a 'fair' model comparison, as differences are all below the ground cover model accuracy.

The under-estimation of the bare ground in the freshly ploughed field in the agricultural example might be due to the fact that the surface is disturbed by the ploughing, but also dark soils have historically been the most difficult soil backgrounds for estimating ground cover. The field observations with high values of greenness in the wheat field also posed some difficulties for all three model predictions. The cases of absolutely bare and entirely lush green cover are extreme cases that were historically not captured in the rangelands and thus are outside of the data range used in the calibration of the data models. This leads to the conclusion that estimates at these extremes need to be interpreted carefully and that more data points in agricultural areas might be needed to account for these situations. However, the three models perform well in agricultural areas that hold a ground cover similar to rangeland conditions, for example with covers different from nearly 100% bare or 100% green. The predictions for GV and NGV appear useful in the agricultural area shown here, despite a small bias in the predictions.

Acknowledgements

Thanks to everyone involved in the Goondiwindi fieldwork and to the property owner Peter Russell who allowed the fieldwork. The authors would like to thank the reviewers who helped to improve this manuscript.

References

Bastin, G.N., Smith, D.M.S., Watson, I.W. & Fisher, A. 2009, "The Australian Collaborative Rangelands Information System: preparing for a climate of change", *The Rangeland Journal,* vol. 31, no. 1, pp. 111-125.

de Vries, C., Danaher, T., Denham, R., Scarth, P. & Phinn, S. 2007, "An operational radiometric calibration procedure for the Landsat sensors based

on pseudo-invariant target sites", *Remote Sensing of Environment,* vol. 107, no. 3, pp. 414-429.

Karfs, R.A., Abbott, B.N., Scarth, P.F. & Wallace, J.F. 2009, "Land condition monitoring information for Reef catchments: A new era", *Rangeland Journal,* vol. 31, no. 1, pp. 69-86.

Leys, J., Smith, J., MacRae, C., Xihua, J., Randall, L., Hairsine, P., Dixon, J. & McTanish, G. 2009, *Improving the capacity to monitor wind and water erosion: A review.*, Brureau of Rural Sciences, Canberra.

Ludwig, J.A., Bastin, G.N., Wallace, J.F. & McVicar, T.R. 2007, "Assessing landscape health by scaling with remote sensing: When is it not enough?", *Landscape Ecology,* vol. 22, no. 2, pp. 163-169.

Muir, J.S. & Danaher, T. 2008, "Mapping water body extent in Queensland through time series analysis of Landsat imagery", *Proceedings of the 14th ARSPC conference, Darwin.* ARSPC.

Queensland Government 2009, *Reef Water Quality Protection Plan 2009 for the Great Barrier Reef World Heritage Area and Adjacent Catchments*, Australian Government; Queensland Government.

Scarth, P., Byrne, M., Danaher, T., Henry, B., Hassett, R., Carter, J. & Timmers, P. 2006, "State of the paddock: monitoring condition and trend in groundcover across Queensland", *Proceedings of the 13th ARSPC conference, Canberra.* ARSPC, .

Scarth, P., Roder, A., Schmidt, M. & Denham, R. in prep., "Fractional groundcover using constrained spectral mixture analysis".

Schmidt, M., Muir, J.S. & Scarth, P. Draft, *Handbook for Surface Cover Field Observation,* Bureau of Rural Sciences, Canberra.

Schmidt, M., Tindall, D., Speller, K., Scarth, P. & Dougall, C. 2010, *Ground cover management practices in cropping and improved pasture grazing systems: ground cover monitoring using remote sensing*, Bureau of Rural Sciences, Canberra.

Schmidt, M. & Trevithick, R. 2010, " Seasonal ground cover monitoring in the grazing lands of the Great Barrier Reef catchments with Landsat time series data", *Proceedings of the 15th ARSPC conference, Alice Springs.* ARSPC.

Schmidt, M., Denham, R.J. & Scarth, P. in prep, "Large scale fractional vegetative ground cover monitoring based on long term field observations and LANDSAT satellite imagery".

Tongway, D.J. & Hindley, N. 1995, *Manual for soil condition assessment of tropical grasslands,* CSIRO Division of Wildlife and Ecology, Lyneham, A.C.T.

Trevithick, R. & Gillingham, S. 2010, "The use of open source geospatial software within the remote sensing centre, QLD", *Proceedings of the 15th ARSPC conference, Alice Springs.* ARSPC.