SEASONAL LAND USE / LAND COVER MAPPING: ACCURACY COMPARISON OF VARIOUS BAND COMBINATIONS

Priyakant Sinha*, Lalit Kumar, Nick Reid

Ecosystem Management, School of Environmental and Rural Science, University of New England, Armidale 2351, NSW, Australia
(psinha, lkumar, nrei3)@une.edu.au

Abstract – The study determines if seasonal spectral variability significantly affects land-cover classification accuracy of Landsat TM in northern New South Wales, Australia, by comparing accuracies in different seasons. Image transformations such as TC, PCA and NDVI were performed and the resultant images were used alone and in combination with original bands for classification. Results from PC1–3, B1–5 and B1–4, 7 were found better than the others; while B1–4 combination was the best in all dates, producing an overall highest accuracy of 96.7% (Kappa 0.96) in January and overall lowest accuracy of 85.7% (Kappa 0.83) in September. Pair-wise tests showed the January accuracy to be significantly higher than in other months at the 0.001 levels and considered the best period for land-cover mapping in the region. The results were supported by analysis of additional January and September images under a range of environmental conditions in three different years.

Keywords: land-cover mapping, transformation, classification, band combinations, accuracy assessment.

1. Introduction

Land-cover mapping is one of the most successful and widely used applications of satellite-based remote sensing, (Laba *et al.,* 1997). Timely and accurate remote sensing-based information about vegetation and land-use and land-cover (LULC) have contributed immensely to the efforts of planners and policy makers in making resource management and land-use planning decisions.

For a landscape characterized by seasonal vegetation features, the number of images required and the date of image acquisition can be of prime importance, but difficult to determine, due to weather constraints (Coppin and Bauer, 1996) and variations in the location and application of the research (Pax-Lenney and Woodcock, 1997). Single-date imagery may be sufficient for some applications. However, multi-seasonal data are useful to better understand seasonal dynamism and improve classification accuracy. Maximum Likelihood (MLC) is the classic, most widely used technique applied in different studies (Laba *et al.,* 1997). In addition to the use of image classifiers with original band data, image transformation techniques have been developed to generate new images for improving classification. Two linear techniques, the Tasseled Cap (TC) transformation (Kauth and Thomas, 1973) and Principal Components Analysis (PCA) are commonly used for image transformation and enhancement. In a few studies, ratio images (e.g. NDVI) have also been used as one of the bands for vegetation or land-cover classification. For example, Miller *et al.* (1998) applied a supervised classification algorithm to selected Landsat MSS spectral bands generated by PCA and NDVI for LULC classification and identified changes from 1973–1991 in New England, USA. Selection of optimum band combinations in digital image classification is also important as using all bands may not provide the best result (Bruzzone and Serpico, 2000; Murakami, 2004). A few researchers have examined optimum band selection (Stenback and Congalton 1990) for image classification through comparison of accuracies attained with different band combinations from one or a few scenes.

The task of selecting appropriate seasonal data and methods to map seasonal land-cover in northern New South Wales (NSW), Australia, using Landsat TM data was attempted through investigation of the electromagnetic radiation response of each cover class in different seasons. The objectives were to:

(a) develop a derived dataset through image enhancement, including PCA, TC and NDVI;
(b) use raw and derived images for land-cover classification with a wide range of band combinations;
(c) undertake a comprehensive accuracy assessment of the land-cover map, and
(d) identify the most suitable season for land-cover mapping in the region.

2. Study region

The study region in northern NSW, Australia, between 29°30’S and 31°0’S latitude and 150°15’E and 152°15’E longitude, covered nearly 34 200 km^2. The area included major parts of the New England Tablelands and Nandewar bioregions and a small portion of the Brigalow Belt South bioregion. Rainfall decreases from east to west across the region. Table 1 describes the LULC identified in the study region.

Table 1. LULC identified in the study region and their descriptions (BRS, 2006).

LULC	Descriptions
Evergreen Forest (EF))	Mostly strict nature reserves, national parks and other protected landscapes consisting of evergreen forest mainly dominated by *Eucalyptus* species.
Evergreen Woodland (EW)	Characterized by low to medium tree density, mostly on rugged steep rocky hills and peaks.
Grazed natural vegetation (GNV)	Mainly native grassland on poor soil, supporting livestock with low productivity throughout year.
Grazed modified pastures (high density) (GMP-HD)	Well maintained grazed pastures used for livestock production based on active modification or replacement of original vegetation.
Grazed modified pasture (low to medium density) (GMP-LMD)	Similar to GMP-HD except that density is less due to more heavily grazed pastures.
Irrigated cropping land (ICL)	Land under crop at time of mapping. This land may be in rotation as described earlier, but classification was based on

* Corresponding author

	land-use at time of mapping.
Fallow or ploughed land (wet soil conditions) (FLWS)	Agricultural lands where irrigated production is carried out but did not contain crops or pastures (in case of rotations) at time of image generation.
Fallow or ploughed land (dry soil conditions) (FLDS)	Agricultural land with no standing crop and where soil was dry.
Waterbody (WB)	All forms of waterbodies.

3. Materials and methods

A multitemporal, multispectral digital image data set of Landsat TM data for six dates (January 2007, March 2008, May 2007, August 2007, September 2007 and November 2009); (path 90, row 81) were used for LULC classification. Images were selected in all four seasons (summer, autumn, winter and spring) from 2007 to 2009 All six TM image Digital Numbers (DNs) were converted to top-of-atmosphere (TOA) reflectance values. A relative radiometric normalization technique was applied to each band from each season's imagery to normalize the variation in solar illumination conditions, atmospheric conditions and other properties of target images with respect to the November base image (Mas, 1999). NDVI was computed for assessing the type, extent, and condition of vegetation in the study region as a part of land-use investigations and to separate vegetated and non-vegetated areas. For all six dates TM, NIR (B4) and Red (B3) were used to calculate NDVI using:

NDVI = (B4–B3) / (B4 + B3)

TC transformation (Crist and Cicone 1984), a method for enhancing spectral information content of TM data was calculated for each scene using six TM bands (1–5 and 7) (Lillesand and Kiefer, 1994). The index fit a linear transformation to six TM bands using a set of empirically derived coefficients (Crist and Cicone, 1984) and the information present in the six original bands was compressed into three TC transformed bands: TC1 (brightness, measure of soil), TC2 (greenness, measure of vegetation), and TC3 (wetness, relationship between soil and canopy moisture). Standardized PCA was used to enhance image contrast by reducing data redundancy and dimensionality in a dataset while retaining most of the information to identify water, vegetation and soil features (e.g. Haack and Jampoler, 1995).

Land-cover classification was carried out on both original and derived images through several band combinations such as (a) three original bands (B24 and B35), (b) four original bands (B1–4, B1,3–5, B2–5 and B3–7), (c) five original bands (B1–5 and B1–4,7) (d) three derived bands TC1–3 and PC1–3, and (e) mixed original and derived bands (B4–5,PC1 and NDVI; B3–5,TC1; B4–5,PC1 and TC2). For convenience in image handling and computer processing, these bands were stacked together to generate a final image containing 13 bands. A supervised approach to image classification was used for LULC mapping and comparison of classification accuracy. The classification was performed through (1) identification of feature and selection of training areas, (2) evaluation and analysis of training signature statistics and spectral pattern, and (3) classification of the images. For each land-cover class, two to three sample plots containing a minimum of 120 pure pixels (i.e. area larger than 1 ha) were chosen as training areas. From the signature polygons (ROI) of each class, a few polygons (nearly 30%) were selected randomly to define the spectral signatures for classification and the remaining 70% were left to evaluate the classification results. Supervised classification was performed using MLC.

The classification accuracy of each image was expressed in the form of an error matrix in terms of producer's error (error of omission), user's error (error of inclusion or commission) and overall accuracy (Congalton *et al.,* 1983). Overall accuracy was calculated by adding the number of pixels classified correctly (diagonal pixels) and dividing by the total number of pixels (sum of all pixels in all ground truthed classes). Kappa Coefficient ($\hat{K}$) (Congalton *et al.,* 1983) was also used as a measure of classification accuracy, subtracting the effect of random accuracy. Pair-wise tests of significance (Gong and Howarth, 1990) were used to compare classifications. The Z-statistic computed in this test compared the $\hat{K}$ statistics obtained from the error matrices of two classifications, to determine if they were significantly different. Two sets of approaches were evaluated namely (i) pair-wise performance of different months and various band combinations taking January as a reference, and (ii) pair-wise performance of different band combinations taking B1–4 as the reference. The classification accuracies obtained from different band combinations were grouped into four broad sets (Table 2).

Table 2 classification accuracy performed on different sets of band combinations

Group	Combinations
Three bands- original or derived bands	(a) B2–4 (b) B3–5 (c) TC1–3 and (1(d) PC1–3
Four bands - only original bands	(a) B1–4, (b) B1,3–5 (c) B2–5 and (d) B3–5,7
Four mixed bands - original and derived bands	(a) B4–7,PC1 (b) B4–5,PC1,NDVI (c) B3–5,TC1 and (d) B4–5,PC1,TC2
Five bands containing only original bands	(a) B1–5 and (b) B1–4,7

Multi-year LULC classification

The purpose of the study was to pick the best season for land-use classification in the region when spectral mixings between land use classes are least. However, the analyses done so far are solely based on one image per date. Monthly rainfall data from 20 different stations for the period of 1990–2010 were analyzed for summer and winter seasons. Summers of 2004, 2007 and 2010 were identified as wet, drought and normal, respectively, while 1998, 2002 and 2007 were identified as wet, drought and normal winters, respectively. The classification results obtained with different band combinations for January and September and accuracies were compared with previous results.

4. Results and discussion

NDVI

Maximum contrast in mean NDVI values was observed between land-use classes of GMP-HD and ICL on the one hand and WB on the other, the two extremes in the NDVI output across all dates. For ICL and GMP-HD and the high NDVI values reflected the greenness of sown, fertilized crops and sown pastures. Low to moderate mean NDVI values were found in wooded areas on all dates, but values in March and August were slightly higher than at other times. The low NDVI values were attributed to the high moisture content and maturity of tree stands. With age, the spaces between tree crowns and branches become larger, resulting in less incident light falling on trees and low IR (TM B4) reflectance and NDVI value (Fiorella and Ripple, 1993). The low mean NDVI of land-use class GMP-LMD was mainly due to low biomass. Land-use category GNV, mostly consisting of native pastures resulted in a very low mean NDVI value.

PCA

The first three PCs contained more than 99% of the total variance at all dates, explaining the entire information content of the image. For comparison, for example, factor loadings for the different bands for January and August for standardized PCA (StdPCA) are shown in Table 3. Little difference in factor loadings on the three PCs was observed except for positive loadings on PC2 in B5 in the January image, and negative loadings of this band on PC2 in August image. StdPC1s were considered to be brightness measures as there was a contribution from all bands. StdPC2 had highest factor loadings in the near infrared (B4) at two dates and were related to dense green vegetation. PC3s were found to be the differences between positively loaded infrared bands and negatively loaded visible bands. Therefore, these reflected dry soil (low visible reflectance and high infrared reflectance).

Table 3. Standardized eigen vectors of January and March

Data	Comp	B1	B2	B3	B4	B5	B7
Jan	PC1	0.41	0.42	0.41	0.39	0.41	0.41
	PC2	0.19	0.06	0.27	0.76	**0.20**	0.51
	PC3	-0.60	-0.32	-0.23	0.42	0.47	0.30
Aug	PC1	0.41	0.42	0.42	0.39	0.41	0.40
	PC2	0.33	0.15	-0.18	0.63	**-0.34**	-0.57
	PC3	-0.59	-0.29	-0.19	0.63	0.33	0.15

LULC classification: comparison of accuracy

A wide set of band combinations were selected for supervised classification and their performance evaluated. Table 4 shows LULC class accuracies in terms of overall accuracy for all months from B1–4 combination by comparing the location and class of each ground-truthed pixel with the corresponding location and class in the classification images. In the case of the January image, almost all of the land-cover classes were spectrally separable producing high classification accuracies except for intermixing of the DEF and MEF categories. Although confusion between the ICL, GMP-LMD, GMP-HD and GNV categories and also between DEF and MEF was observed in all months, the former set was particularly difficult to distinguish in August and September, resulting in high intermixing between these classes, while DEF and MEF were most inter-mixed in January image. Table 5 shows the results of Kappa analysis in terms of the Z-statistics computed from B1–4 classification error matrices between January and the other months. The January classification was significantly better than the May classification at the 92% level, while it was significantly better than other months at the 99.9% level.

Different band combinations produced different classification accuracies. For all three-band, four-band and five-band combinations, the January image produced the highest overall accuracy while accuracies obtained with the August and September images were lowest. Among three-band combinations, B2–4 was slightly better than B3–5. For derived bands, the PC bands had slightly higher accuracies than the TC bands and higher again than original three band combinations. Combination PC1–3 resulted in a maximum overall accuracy of 95.9% in January, followed by 92.7% in each of March and May. The same combination produced overall accuracies of 85.3%, 80.5% and 87.7% with the August, September and November images, respectively. The overall accuracies obtained with the TC1–3 were 95.1%, 90.7%, 91%, 82.8%, 79.3% and 86.1%, respectively, from January to November.

Table 4 Overall LULC classification accuracy under different band combinations.

3 Band Combination						
	Jan	Mar	May	Aug	Sept	Nov
B 2-4	94.7	89.5	91.6	86	82.7	90.6
B 3-5	90.7	88.7	87.1	81.8	77.9	83.4
TC1-3	95.1	90.7	91	82.8	79.3	86.1
PC1-3	95.9	92.7	92.7	85.3	80.5	87.7
4 Band Combination						
B 1- 4	96.7	91.1	96	86.9	85.7	93.9
B 1, 3-5	96.7	89.7	94.1	86	86.3	88.1
B 2-5	94.7	91.5	91.6	85.5	83.7	87
B 3-5, 7	91.9	88.7	87.8	83.5	78.2	82.1
B4-5, 7, PC1	96.4	88.1	91.2	84.8	83.3	87.2
B4-5,PC1,NDVI	96.4	92	93.1	83.3	81	86.5
B3-5,TC1	96.6	91	92.6	83.9	82.3	85
B4-5,PC1,TC2	95.9	90	90.3	82.9	78.1	83.6
5 Band Combination						
B 1-5	96.7	90.1	94.1	86.2	85.8	89.7
B 1-4, 7	96.7	91.6	91.8	86.5	81.1	92.3

Table 5. Comparison of Jan vs. other months using B1–4

Comparison	Z-statistics	Confidence level
Jan vs Mar	13.08	99.9%
Jan vs May	1.94	94%
Jan vs Aug	21.64	99.9%
Jan vs Sept	23.83	99.9%
Jan vs Nov	7.24	99.9%

Accuracy obtained from four band combinations was highest in January for all the combination used. Both combinations B1–4 and B1, 3–5, resulted in overall accuracies of 96.7% in January, respectively while the same combinations were least accurate (85.7% and 86.3%) in September. The two other combinations, B2–5 and B3–7 showed similar trends, with accuracies ranging from 94.7% and 91.9%, respectively, in January to 83.7% and 78.2%, respectively, in September. The accuracies for remaining dates lay between these two extremes. Because of the capability of B1 and B2 to distinguish the soil surface and vegetation, band combinations B1–4, B1, 3–5 and B2–5 produced higher accuracies than B3–7, which contained mid-infrared bands mostly effective in detecting soil or plant moisture content. B1–4 was the best among this group as it consistently produced higher accuracies across all dates. Similar results were obtained with mixed four-band combinations with overall accuracy greatest in January and least in August and September. Combination B4–5, PC1, NDVI was slightly more accurate than B4–5, PC1, TC2, demonstrating the superiority of NDVI over TC2 bands in measuring greenness.

The results obtained with five-band combinations were very similar to PC1–3, B1–4 and B1, 3–5 with overall accuracies of 96.7% with both B1–5 and B1–4, 7 combinations in January. The performance of band combination B1–4, 7 was in slightly better than B1–5 in March, August and November, while in September B1–5 was more accurate.

In summary, maximum accuracy was achieved in January with all band combinations. Accuracy from original three band combination was substantially lower than that with B1–4 combinations. However, four and five- band combinations produced similar results, suggesting that addition of the fifth band in land-cover classification did not increase classification accuracy significantly (Z statistics-NS) but did increase data volume and computing time, similar to conclusion made by Schriever and Congalton (1995) .

The results obtained under different environmental conditions in wet and dry years in summer and winter once again showed the higher land-cover classification accuracy of images obtained in January–February compared to September. In the case of January and February images, comparable results were obtained with all band combinations under all conditions (wet, drought and normal). However, in the case of September data, the highest accuracy was achieved in a drought year (2002) followed by the normal year (2007) and finally the wet year (1998). The consistency in results confirmed that mid-summer was the best season for LULC classification in the region and therefore a January image was used for final LULC mapping.

5. Conclusion

The study shows the potential of satellite data for land-cover mapping through application of various digital image processing techniques. The aim of this study was to identify the season in which landcover classes showed maximum separability. A thorough investigation of seasonal variability in land-cover classes over four seasons using Landsat TM data was performed to explain how seasonal spectral variability affects classification accuracy. Since the region is characterized by high tablelands, slopes and plains, supporting a variety of open forests, woodlands and other vegetation types such as grazing pastures and cropland, different LULC classes responded differently to seasonally varying meteorological conditions and human activities. Among different band combinations used the B1–4 classification was the best across all dates and was used in final LULC mapping. Most inter-mixing of land-cover classes was between ICL, GMP-LMD, GMP-HD and GNV categories in most but these classes were particularly difficult to distinguish in August and September images. Because of the substantially higher accuracy achieved in the January image and lower accuracy in September, the former was considered to be best and the latter the worst period for LULC map generation in the region. This result was further supported by analysis of additional images for January and September months under different environmental (wet, drought and normal) conditions. These results confirmed the superiority of land-cover classification accuracy in January compared to September.

6. References

Bruzzone, L., and S.B. Serpico, 2000. A technique for feature selection in multiclass problems, International Journal of Remote Sensing, 21:549-563.

Bureau of Rural Sciences (BRS), 2006. Guidelines for land use mapping in Australia: principles, procedures and definition. Canberra, ACT 2601.

Coppin, P.R., and M.E. Bauer, 1996. Digital change detection in forest ecosystem with remotely sensing imagery, Remote Sensing of Environment, 13:207-234.

Congalton, R. G., R. G. Oderwald, and R. A. Mead, 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques, *Photogrammetric Engineering and Remote Sensing*, 49:1671–1678.

Crist, E. P., and R. C. Cicone, 1984. Application of the tasseled cap concept to simulated Thematic Mapper data, *Photogrammetric Engineering and Remote Sensing*, 50:343–352.

Gong, P. And Howarth, P. J., 1990. An assessment of some factors influencing multispectral land-cover classification, *Photogrammetric Engineering and Remote Sensing*, 56:597-603.

Haack, B., and S. Jampoler, 1995. Colour composite comparisons for agricultural assessments, *International Journal of Remote Sensing*,16(9): 1589-1598.

Kauth, R.J., and G.S. Thomos, 1976. The tasseled cap- a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proceedings, Symposium on Machine Processing of Remotely Sensed Data,* West Lafayette, IN: LARS, 41-51.

Laba, M., S. D. Smith, and S. D. Degloria, 1997. Landsat based land cover mapping in the lower Yuna river watershed in the Dominican Republic, *International Journal of Remote Sensing*, 18:3011-3025.

Lillesand, T.M., and R.W. Kiefer, 1994. *Remote Sensing and Image Interpretation*, Newyork:Wiley.

Mas, J. F., 1999. Monitoring land-cover changes: a comparison of change detection techniques, *International Journal of Remote Sensing*, 20:139–152.

Murakami, T., 2004. Seasonal variation in classification accuracy of forest-cover types examined by a single band or band combinations, *Journal of Forest Research*, 9:211-215.

Pax-Lenney, M., and C.E. Woodcock, 1997. Monitoring agricultural lands in Egypt with multitemporal landsat tm imagery: how many images are needed?, *Remote Sensing of Environment*, 59:522-529.

Schriever J.R., and R.G. Congalton, 1995. Evaluating seasonal variability as an aid to covertype mapping from Landsat Thematic mapper data in the northeast, *Photogrammetric Engineering and Remote Sensing*, 61:321-327.

Stenback, J.M., and R.G. Congalton, 1990. Using thematic mapper imagery to examine forest understory, *Photogrammetric Engineering and Remote Sensing*. 56:1285-1290.