ILLUMINATION CORRECTION OF LANDSAT TM DATA IN SOUTH EAST NSW

Elizabeth Roslyn McDonald[1], Xiaoliang Wu[2], Peter Caccetta[2] and Norm Campbell[2]

[1]Environmental Resources Information Network (ERIN), Department of Environment and Heritage, GPO Box 787, Canberra 2601, Ph (02) 6274 2381, fax (02) 6274 1333

[2]CSIRO Mathematical and Information Sciences, Leeuwin Centre for Earth Sensing Technologies, 65 Brockway Road, Floreat, 6014

Abstract

The NSW Forests Taskforce within Environment Australia purchased multi-temporal Landsat Thematic Mapper (TM) data ranging in dates from 1987 through to 1999 for use in the NSW Regional Forest Agreement (RFA) process. The geometrically and radiometrically corrected multi-temporal mosaics were used to assist in mapping and updating of wilderness areas, old-growth forest and disturbance history layers. The data are also used as an information layer in a Geographical Information System (GIS) to assist in the design of reserve areas.

Much of eastern NSW contains high topographic relief. The effects of the topography, coupled with a low sun angle at the time of satellite overpass, creates significant shadowing effects in the data. These effects constrain the application of image classification techniques to further ‘value-add’ the data for land management purposes. A collaborative project between Environment Australia and CSIRO Mathematical and Information Sciences was set up to test the effectiveness of selected illumination corrections for reducing shadowing effects.

A number of published algorithms (Teillet *et al.*, 1982 and Meyer *et al.*, 1993) were tested. The application of these methods required the use of a high resolution (25 metre) Digital Elevation Model (DEM). A correction method first published by Teillet *et al.*, (1982) and referred to by Meyer *et al.*, (1993) as the C-correction, was found to give the best results. This correction is similar to a simple cosine illumination correction but introduces an adjusted offset derived from the regression of the digital number against the calculated sun incidence angle.

A canonical variate analysis (CVA) was used to compare results before and after application of the correction. Test areas were selected over a range of incidence angles within four major land cover classes, including eucalypt forest, exotic plantations, dry-land agriculture and irrigated agriculture. As expected, the CVA for both original and corrected data showed good separation between the major land cover classes. The CVA for the original data showed that, within the major land cover classes, particularly the high relief forest and plantation sites, variation along CV1 was related to the incidence angle. A CVA applied to the corrected data showed that CV1 predominantly separated cover classes. Further analysis of the corrected data showed that location of the sites along CV2 could be related to vegetation ‘greenness’.

It was concluded that the illumination correction significantly reduced the shadowing effects in the image and can be recommended for use before classification of multi-temporal image data in eastern NSW.

Introduction

The NSW Forests Taskforce within Environment Australia purchased multi-temporal Landsat TM data ranging in date from 1987 through to 1999 for use in the NSW Regional Forest Agreement Process. The geometrically and radiometrically corrected multi-temporal mosaics were used to assist in the mapping and updating of old-growth forest, wilderness and disturbance history layers collected as part of the NSW Regional Forest Agreement. The data is also used as an information layer in a Geographical Information System to assist in the design of reserve areas.

Much of eastern NSW contains high topographic relief. The effects of the topography, coupled with a low sun angle at the time of satellite overpass, creates significant shadowing effects in the data. These effects constrain the application of image classification techniques to further 'value-add' the data for land management purposes. A collaborative project between Environment Australia and CSIRO Division of Mathematical and Information Sciences was set up to test the application of published illumination corrections for their suitability in addressing this problem. The application of these methods required the use of a high resolution (25 metre) Digital Elevation Model (DEM). The results of this project are given in the following report.

Materials and Methods

Illumination Corrections

To test the utility of illumination correction on Landsat TM data in NSW, a number of methods derived from the scientific literature (Teillet *et al.*, 1982, and Meyer *et al.*, 1993) were implemented and are listed in Table 1:

Table 1: Illumination corrections tested in the study

1. Cosine correction	$L_H = L_T \frac{\cos(sz)}{\cos(i)}$
2. Statistic-empirical method	$L_H = L_T - \cos(i)m - b + \overline{L_T}$
3. Minnaert correction	$L_H = L_T \left[\frac{\cos(sz)}{\cos(i)}\right]^k$

4. C-correction	$L_H = L_T\left[\dfrac{\cos(sz)+c}{\cos(i)+c}\right]$
5. Advanced Minnaert	$L_H = L_T\left[\dfrac{\cos(sz)}{\cos(i)}\right]^k\left[\dfrac{\cos(w)}{\cos(v)}\right]^{1-k}$
6. Advanced C-correction	$L_H = L_T\left[\dfrac{\cos(sz)+c}{\cos(i)+c}\right]\left[\dfrac{\cos(w)}{\cos(v)}\right]^{1-k}$
7. Semi-empirical	$L_H = L_T\left[\dfrac{\cos^k(sz)+c}{\cos^k(i)+c}\right]$

L_H = radiance observed for a horizontal surface;
L_T = radiance observed for sloped terrain;
$\cos(sz)$ = cosine of, sun's zenith angle;
$\cos(i)$ = cosine of, sun's incidence angle;
b = intercept of the regression line;
m = gradient of the regression line;
$c = \dfrac{b}{m}$
k = Minnaert constant (considered to be a measure of the extent to which a surface is Lambertian);
$\cos(w)$ = emergent angle;
$\cos(v)$ = rotation of mapping coordinates

The correction methods, listed above, were coded by Xiaoliang Wu, (1999) using C/C++ on a SUN . For the first four corrections: Cosine; Statistic-empirical; Minnaert and C-correction, the program requires four inputs:

- Geometrically and radiometrically corrected Landsat TM image;
- A high resolution DEM;
- location information in Australian Map Grid (AMG) coordinates;
- the sun elevation and azimuth for the time of satellite overpass;

The last three corrections: Advanced Minnaert; Advanced C-correction and Semi-empirical; require a further two inputs:

- Satellite height;
- map coordinates for the corners of the original image.

The program returns the corrected image and, if specified by the user, ASCII text files that contain both pre-corrected and post-corrected image values. This information can then be imported into a statistical package such as SPLUS for further analysis.

Study Sites

Each of the illumination corrections was tested at two study sites located in the Tallangatta Landsat TM Scene (091/085) taken on the 28 December 1998. This scene was chosen because it contained a good mix of both high-relief forested areas and low relief forest and agricultural areas. The Landsat TM scene was geometrically and radiometrically corrected using the methods of Furby (1999), before comparing the illumination corrections.

The first study site was a forested area located in mountainous terrain, refer to Table 4 for images of the test area. The second study site is an agricultural area. A Geographic Information System (GIS) coverage containing vegetation information derived from aerial photography (NPWS, 1999) was used to separate native and plantation forest areas. The location of the two study areas is given in Table 2 for AMG Zone 56.

Table 2: Forest and Agriculture field site locations used to test the illumination corrections listed above

Cover Type	Easting	Easting	Northing	Northing
Eucalypt Forest	658750	670775	6059625	6045900
Dryland Agriculture	512775	527979	6061166	6045534

For the Tallangatta scene (path/row: 091/85) taken on 28 December 1998 the position of the sun was as follows:

Elevation = 0.96587 radians;
Azimuth = 1.38213 radians.

The XY coordinates of the original Tallangatta image in AMG coordinates and the satellite height are as follows:

Top-Left corner: 509894 6109905
Top-right corner: 689852 6076023
Bottom-right corner: 648888 5908876
Bottom-left corner: 468841 5942775

Satellite Height = 702000 km

The information above is obtained from the GICS Radiometric Quality Assessment Report supplied with the original Landsat TM data from the Australian Centre for Remote Sensing (ACRES).

Results

Statistical Interpretation of the Various Correction Methods

The forest and agriculture areas were successively corrected using the algorithms listed in Table 1. The resulting trends for both the forested and agricultural areas were similar. Therefore, only the results from the forested test site are presented in this paper.

The digital values for both the pre-corrected and post-corrected data were imported into the Splus Statistical Analysis software. Linear regression methods were then used to compare the results of the pre-corrected and post-corrected data and for each subsequent illumination corrections. An example of the relationship between TM Band 4 digital number (DN) and incidence angle for both uncorrected and corrected data is given in Figures 1 and 2 below. A full listing of the numerical results for each of the corrections is given in Table 3.

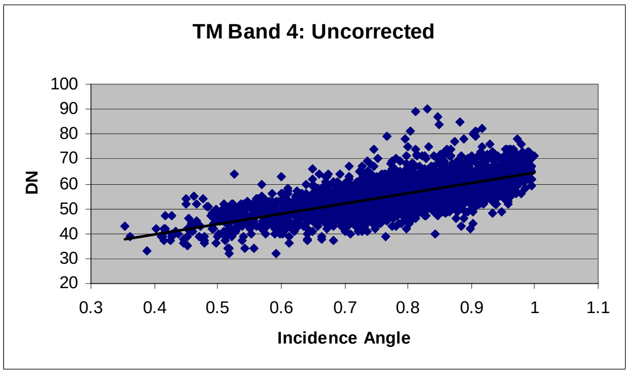


Figure 1 shows a positive linear relationship between the Digital Numbers (DN) of TM Band 4 and the incidence angle. The equation for the trend line gives a *y* intercept of 23 and a slope of 41. The multiple r-squared is r^2=0.47 with a residual standard error of 5.50.

Figure 1: Linear regression of incidence angle versus band 4 for uncorrected data in the forested test area.

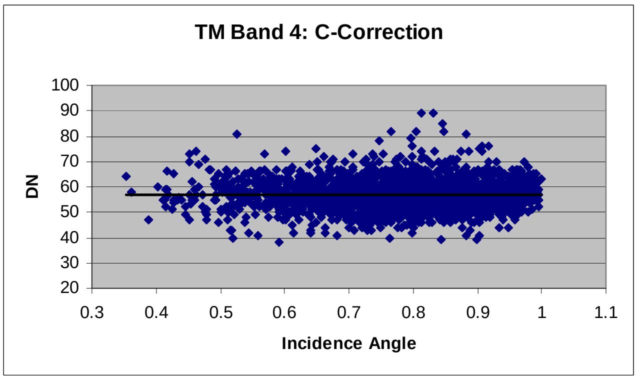


Figure 2 shows the effect of applying a C-correction on the same data set shown in Figure 1. The relationship between the DN and the incidence angle is removed from the data. The equation for the trend line gives a *y* intercept of 57 and a slope of -0.4. The multiple r squared is r^2=0 with a residual standard error of 5.69.

Figure 2: Linear regression of incidence angle versus band 4 for C-Corrected data in the forested test area.

As can be seen from Figures 1 and 2, the purpose of applying an illumination correction is to reduce the effect of the incidence angle on the image digital values. It should be noted that in Table 3 a low r^2 value following correction indicates the effectiveness of the correction. Consequently, a low r^2 value suggests that a classification of land cover types applied after correction should be less biased by changes in illumination

Table 3: Results from analyses of the forested test area. The results were obtained by regressing DN values against incidence angle for both uncorrected and corrected data. Ideally the slope for the corrected data is zero

Landsat TM Band Number and Correction Number From Table 1	**Illumination Correction Name**	**Slope** ***m***	**Intercept** ***b***	**Coefficient of det.** r^2
Band1	**Uncorrected data**	**8.0**	**73.4**	**0.1409**
1	Cosine	-116.0	176.5	0.9171
2	Statistic-empirical	0.14	72.7	0
3	Minnaert	0.28	79.3	0.0002
4	C	0.03	79.5	0
5	Advanced Minnaert	14.13	66.6	0.2379
6	Advanced C	14.03	66.6	0.2315
7	Semi-empirical	21.26	60.8	0.4084
Band 2	**Uncorrected data**	**7.0**	**22.7**	**0.2155**
1	Cosine	-35.9	58.0	0.8219
2	Statistic-empirical	0.16	22.1	0.0001
3	Minnaert	0.25	27.8	0.0003
4	C	0.15	27.9	0.0001
5	Advanced Minnaert	4.42	23.9	0.0848
6	Advanced C	4.36	23.9	0.0817
7	Semi-empirical	9.53	19.7	0.2967
Band 3	**Uncorrected data**	**12.9**	**18.4**	**0.2931**
1	Cosine	-29.3	53.0	0.6236
2	Statistic-empirical	-0.05	17.9	0
3	Minnaert	0.22	28.4	0.0001
4	C	0.03	28.5	0
5	Advanced Minnaert	4.06	24.9	0.0372
6	Advanced C	3.85	25.0	0.0334
7	Semi-empirical	11.34	18.8	0.2322
Band 4	**Uncorrected data**	**41.2**	**23.4**	**0.4734**
1	Cosine	-37.9	88.5	0.383
2	Statistic-empirical	0.03	22.9	0
3	Minnaert	0.27	56.7	0
4	C	-0.41	57.1	0
5	Advanced Minnaert	5.53	51.9	0.0151
6	Advanced C	4.73	52.4	0.0110
7	Semi-empirical	22.0	38.2	0.2008
Band 5	**Uncorrected data**	**57.9**	**9.9**	**0.3301**
1	Cosine	-17.1	71.2	0.0357
2	Statistic-empirical	1.3	8.8	0.0002
3	Minnaert	-0.24	57.2	0
4	C	-0.82	57.6	0
5	Advanced Minnaert	2.11	55.1	0.0006
6	Advanced C	1.56	55.5	0.0003
7	Semi-empirical	13.56	45.57	0.0249
Band 7	**Uncorrected data**	**23.8**	**2.9**	**0.1976**
1	Cosine	-5.6	26.7	0.0115
2	Statistic-empirical	1.12	2.5	0.0007
3	Minnaert	-0.72	22.6	0.0002
4	C	-0.85	22.7	0.0003
5	Advanced Minnaert	-0.02	21.9	0
6	Advanced C	-0.11	22.0	0
7	Semi-empirical	3.66	18.9	0.0053

Table 3 shows that the performance across corrections varies. Overall, the worst result was given by the Cosine correction. In areas where the incidence angle approaches 90 degrees (that is, where *cos(i)* tends towards zero) the fraction becomes very large and when multiplied by the pixel DN it creates a disproportionate brightening effect. For pixels in complete self shadow *(cos (i)=0))*, a division by 0 occurs leading to the creation of artifacts in the data (Meyer, 1993). When applied to Bands 1, 2 and 3, the Cosine corrected data is highly correlated with the topography, with highest correlation occurring in Band 1 (r^2 = 0.917) in comparison to the uncorrected data (r^2= 0.141). This correlation decreases as wavelength increases from the visible to the short wave infrared and becomes almost negligible at Band 7 (r^2 = 0.012).

The Advanced Minnaert and Advanced C-corrections show an over-correction in Band 1, but improve with increasing wavelength and show reasonable corrections for Bands 4, 5 and 7. The Semi-empirical correction shows an over-correction in Bands 1 and 2 and poor corrections in the other wavelengths.

The remaining corrections, the Statistic-empirical, the Minnaert and the C-corrections, give the best results over all bands. The statistic-empirical correction works best in the visible bands but is not as effective in the short-wave infrared bands. A visual inspection of the Statistic-empiric corrected image shows that the histograms for both bands 5 and 7 are being truncated for low DN values, therefore decreasing their dynamic range in comparison with the uncorrected data.

There is very little difference between the results obtained from the Minnaert and the C-corrections, though the C-correction operates best on the visible bands (1,2,3) and the Minnaert correction operates best in the infrared and short wave infra red bands (4,5,7). Therefore, for forested areas, it is recommended that either the Minnaert or C-corrections be used to reduce the effects of the illumination.

Visual Interpretation of Corrections

For the forested region, each of the corrections was applied and output images were displayed in ERMapper. The same histogram stretch was applied to all of the data (except for the statistic-empirical correction) so that a visual comparison of the pre-corrected versus post-corrected data could be made. The images show Bands 2, 4 and 5 in blue, green and red respectively. All of the images show a decrease in illumination effects when the correction is applied, except in the case of the Cosine and the Semi-empirical corrections.

In the case of the Cosine correction, the over-exaggeration is dramatic and in the Semi-empirical less so. In the case of the statistic-empirical correction, investigation of the histograms showed that the correction is truncating the low DN values in bands 5 and 7. The reason for this is unknown, though it meant an equivalent histogram stretch could not be applied to this image.

Table 4: forested test area showing the application of successive illumination corrections as listed in Table 1.

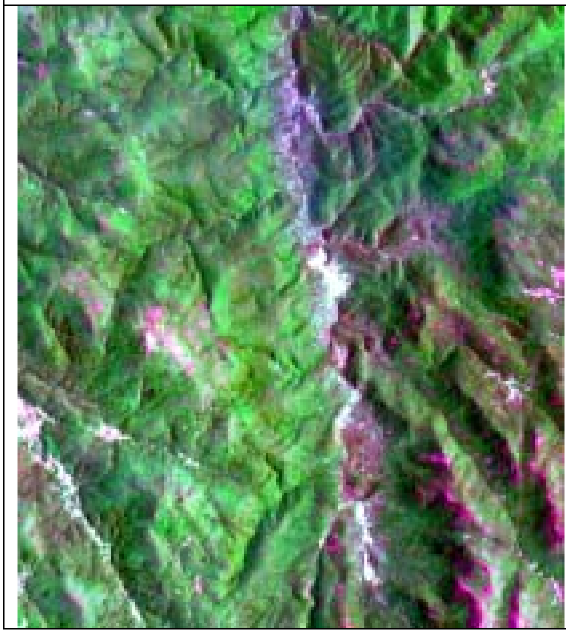
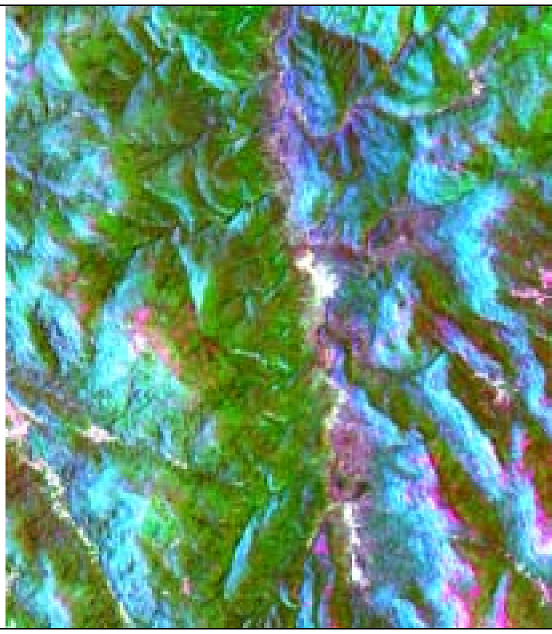
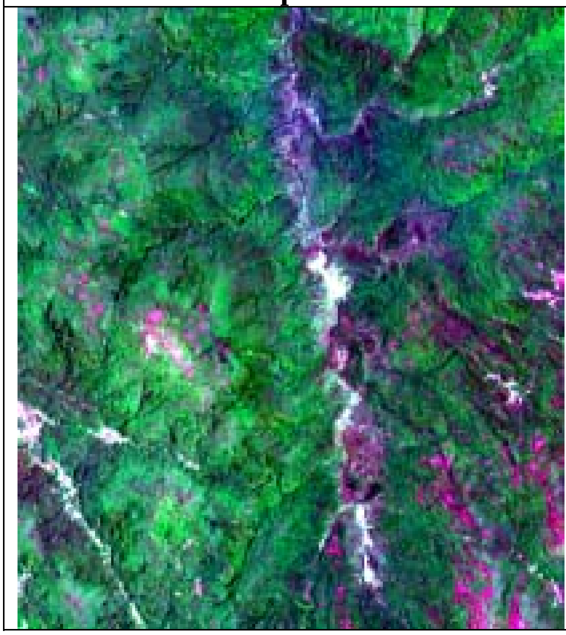
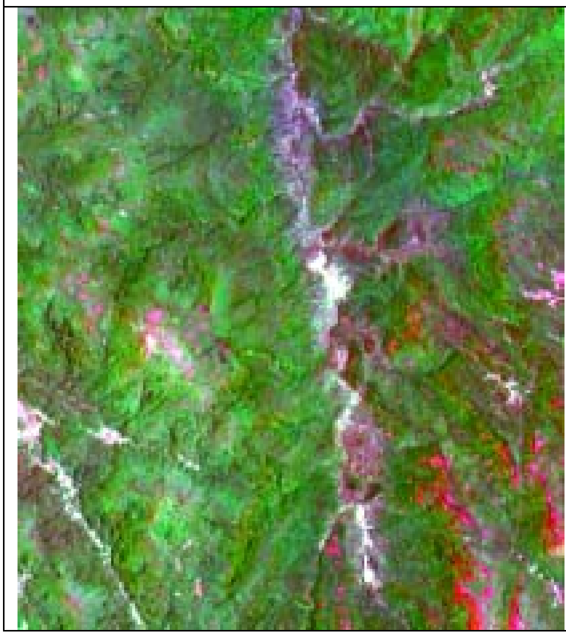
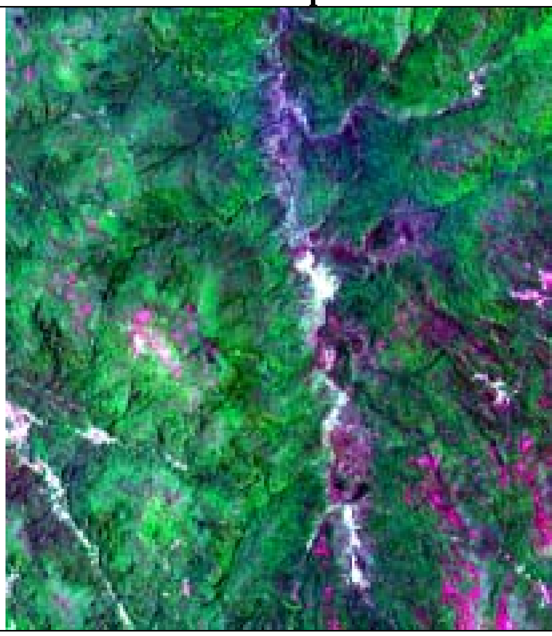
Uncorrected Data	1. Cosine Correction
	
2. Statistic Empiric Correction	**3. Minnaert Correction**
	

Table 4: Continued

4. C-correction	**5. Advanced Minnaert**
	

6. Advanced C-correction	**7. Semi-empirical**
	

Accuracy Assessment

The results of the statistical and visual analysis of the data indicate that the most effective illumination corrections are the Minnaert and C–corrections.

To further test the robustness of the C-correction, a canonical variate analysis (CVA) was used to examine the difference between the same cover types both before and after application. Ideally, areas with the 'same' cover type, but occurring at different incidence angles, should have a similar spectral response after illumination correction.

To test this premise, four areas with homogenous cover types were selected from within the image. The location of these areas is given in Table 5. Within each of the four cover types, DN values at successively increasing incidence angles were sampled.

Table 5: Location of the four major cover classes used in the Canonical Variate Analysis

Cover Type	Easting	Easting	Northing	Northing
Eucalypt Forest	658750	670775	6059625	6045900
Exotic Plantation	587788	591075	6055257	6051134
Dryland Agriculture	512775	527979	6061166	6045534
Irrigated Agriculture	571397	592742	6055500	6037555

Canonical Variate Analysis of Uncorrected Data

The canonical variate analysis for the four cover types given in Table 5 has the following canonical roots: CV1 = 9.076; CV2 = 1.387. The greatest proportion of variation in the test data is exhibited by CV1 - the brightness gradient. As expected, the results showed good separation between the forested sites (i.e. eucalypt and plantation) and non-forested sites, (i.e. agriculture and horticulture). There was, however, mixing of the mean values within the forested sites, showing an overlap for the eucalypt and plantation sites. For all classes, a large proportion of the variation in CV1 is being driven by incidence angle, with lower CV1 values indicating shadowed terrain and successively higher CV1 values indicating a solar brightening in the terrain.

To further investigate the relationship between the eucalypt and plantation forest sites, the agricultural and horticultural sites were removed and the analysis rerun, the results of which are shown in Figure 3. As expected, the canonical variate analysis showed an increase in the separation of the forest and plantation sites, with canonical roots: CV1 = 5.167 and CV2 = 2.023. The plot shows that the variation in CV1 is being driven by the incidence angle, with low CV1 values indicating shadowed areas and successively higher CV1 values indicating a solar brightening in the terrain.

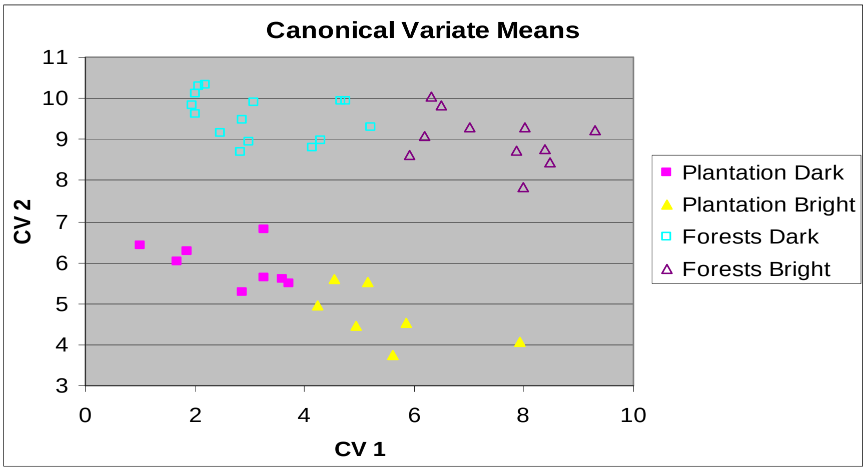


Figure 3: Scatter plot from a canonical variate analysis of forests and plantation sites for uncorrected data

Canonical Variate Analysis of C-corrected Data

The C-correction was applied to the data and the canonical variate analysis rerun, the results of which are shown in Figure 4. The canonical roots were CV 1 = 2.24; CV 2 = 0.69. For both the forest and plantation classes there is a significant decrease in the dependence on incidence angle in CV1, with classes of various incidence angles showing a mixing along CV1. This indicates that the differences in the classes are no longer attributable to brightness value but by some other factor such as 'greenness' in CV2.

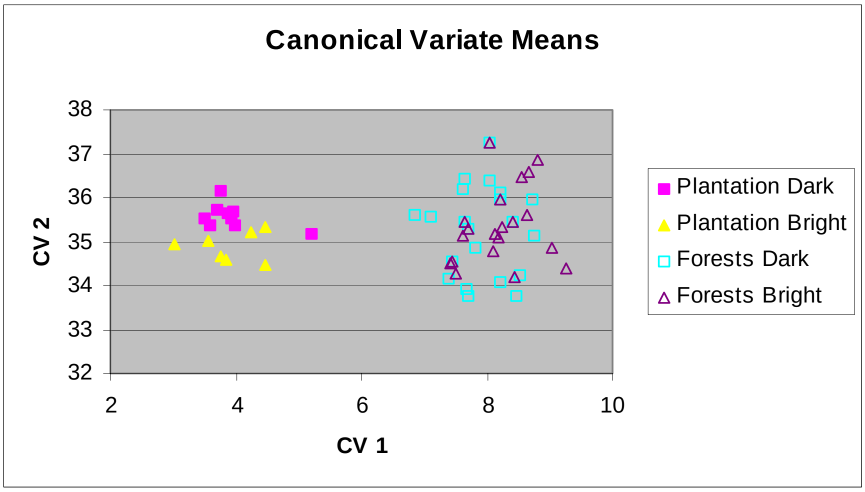


Figure 4: Scatter plot from a canonical variate analysis of forests and plantation sites for C-corrected data

Discussion

The results above show that the most effective illumination correction is the C-correction. It is worth noting that while the Minnaert correction gave similar results to those of the C-correction, the coefficients for the C-correction are easier to obtain from the data, making its application simpler.

It was suggested by Teillet *et al*., (1982) in a similar study that the additive parameter *c,* given in the formula for the C-correction, and the power constant *k* in the Minnaert correction may mimic the effect of the diffuse light (path radiance) component. Mathematically the effect of the *c* parameter is similar to that to the Minnaert constant, in that it increases the denominator and weakens the over-correction of faintly illuminated data (Meyer, 1993). In contrast, the Cosine correction only models the direct component of the incoming solar radiation. The Cosine correction is generally used to correct variations in sun angle for multi-temporal data. However, when applied in areas of steep terrain with faint illumination, the denominator tends to zero and the fraction becomes very large. This has an exaggerated multiplier effect on the DN value that leads to an over- brightening of the data, as shown by the bright blue areas in Table 4 for the Cosine correction.

One of the issues concerning illumination correction is that its application, at least in a broad sense, is land cover dependent. This fact was illustrated when deriving the coefficients for the C-correction. The values of *c* (i.e. $c=b/m$ derived from the regression analysis) for the forested area were significantly different to those obtained for the agricultural area. This suggests that scattering of the solar radiation is dependent on the vegetation cover. Another factor is that agricultural areas are generally found on flatter terrain, thus the coefficients in the C-correction become very small and the correction tends towards a Cosine correction.

As a result of these findings, it is suggested that major land cover classes such as bush/non-bush should be separated and appropriate correction coefficients obtained before correcting the data. Further testing in this area is required, but preliminary results suggest that coefficients derived from these two major land cover classes are robust enough to be applied to a whole Landsat TM scene.

It has been shown that application of the C-correction provides a significant reduction in the illumination-driven variation over similar land cover types. This leads to an increase in the separation of spectrally similar classes such as eucalypt forest and plantation areas. Without the aid of an illumination correction, the results of a visual or digital classification may become confused. That is, areas of similar land cover type, but different illumination, will be classified as different classes, thus masking the variation of interest. Because it is the variation in land cover type that will most often be of interest, it follows that removal of illumination effects before classification will improve the results. Exhaustive testing of this premise is the subject of ongoing work.

Conclusion

The main findings from this study are:

- That application of an illumination correction is landcover dependent, thus indicating that major land cover classes such as bush/non-bush should be separated before correction, and appropriate correction coefficients obtained for each;
- The overall best performing illumination correction is the C-correction. Note: though the statistical and visual results for the Minnaert correction were similar to those of the C-correction, the coefficients for the C-correction are easier to obtain;
- Application of the C-correction provides a significant reduction in the illumination-driven variation observed in areas of similar cover type;
- Application of the C-correction provides an increase in the separation of spectrally similar classes such as eucalypt forest and plantation areas.

References

Furby, S. (1999). *Improving the Scene-to-Scene Registration of Overlapping Images Using Terrain Correction. A Case Study Using NSW RFA (South) Data*. CSIRO Mathematical and Information Sciences, Perth, WA.

Meyer, P., Itten, K., Kellenberger, T., Sandmeier, S., Sandmeier, R. (1993). *Radiometric corrections of topographically induced effects on Landsat TM data in an alpine environment. ISPRS Journal of Photogrammetry and Remote Sensing.* 48(4): 17-28.

NPWS (1999). NSW, Southern, Comprehensive Regional Assessment, API Derived Vegetation Mapping.

Teillet, P.M., Guindon, B., and Goodenough, D.G. (1982). *On the Slope-Aspect Correction of Multispectral Scanner Data*. Can. J.Remote Sensing, 8(2): 84 – 106.

Xiaoliang, Wu. (1999). *A Program for Illumination Corrections for Landsat TM Data.* CSIRO Mathematical and Information Sciences, Perth WA.