Radiometric correction techniques and accuracy assessment for Landsat TM data in remote forested regions

Darren T. Janzen, Arthur L. Fredeen, and Roger D. Wheate

Abstract. Subtle change detection analysis in remote sensing relies on some form of radiometric consistency. Radiometric correction techniques developed in previous studies often require ancillary information such as climate data, illumination geometry, ground reference data of pseudo-invariant features (PIFs), and satellite calibration data. Most studies do not have the luxury of all these data. A relative radiometric correction technique of consistent quality applicable to study areas that lack urban development has not been generally accepted by the remote sensing community. A series of Landsat-5 thematic mapper (TM) and Landsat-7 enhanced thematic mapper plus (ETM+) images spanning 18 years was obtained for a primarily forested area in central British Columbia, Canada. Different techniques of radiometric correction that do not rely on ground reference data, climate data, or the subjective selection of PIFs were assessed for these images. They included an atmospheric transfer model that requires no ancillary climate data, a simple scaling function, and two scatterplot-based regression functions. Assessment of radiometric consistency was performed qualitatively by using edge detection and quantitatively using analysis of old-growth forests in equilibrium and measures of biomass accumulation in clearcuts. For these three methods of assessment, the two scatterplot-based regression functions yielded the best radiometric fidelity. These two techniques can be completely automated and are equally applicable in any Landsat TM- or ETM-based change detection studies.

Résumé. La détection de changements subtils en télédétection est basée sur une certaine forme de cohérence radiométrique. Les techniques de correction radiométrique développées dans des études précédentes requièrent souvent des informations auxiliaires telles que des données climatiques, la géométrie d'illumination, des données de référence au sol des attributs pseudo-invariants (PIF) ainsi que les données d'étalonnage du satellite. La plupart des études ne peuvent se permettre le luxe d'autant de données. Aucune technique de correction radiométrique relative, de qualité constante et applicable à des zones d'étude caractérisées par l'absence de développement urbain, n'a encore été acceptée par l'ensemble de la communauté de télédétection. Une série d'images TM (« thematic mapper ») de Landsat 5 et ETM+ (« enchanced thematic mapper plus ») de Landsat 7 couvrant 18 ans a été acquise pour une zone principalement forestière dans le centre de la Colombie-britannique, au Canada. Différentes techniques de correction radiométrique qui ne sont pas basées sur des données de référence au sol, des données climatiques ou la sélection subjective de PIF ont été évaluées pour ces images. Ces techniques comprenaient un modèle de transfert atmosphérique qui ne requiert pas de données climatiques auxiliaires, une fonction d'échelle simple et deux fonctions de régression. L'évaluation de la cohérence radiométrique a été réalisée qualitativement, à l'aide de la technique de détection de contours, et quantitativement, à l'aide d'une analyse aux forêts anciennes en équilibre et des mesures d'accumulation de la biomasse dans des coupes à blanc. Pour ces trois méthodes d'évaluation, les deux fonctions de régression ont donné la meilleure fidélité radiométrique. Ces deux techniques peuvent être entièrement automatisées et sont applicables à toute étude de détection du changement basée sur les données TM et ETM de Landsat.
[Traduit par la Rédaction]

Introduction

Satellite image analysis has played a key role in environmental monitoring and modelling over the past few decades. Repeat observation of a given area over time yields the potential for many forms of change detection analysis. These can be confounded by radiometric inconsistency due to changes in sensor calibration, differences in illumination and observation angles, and variation in atmospheric effects (Eckhardt et al., 1990).

Radiometric correction of satellite imagery falls into two broad categories, namely absolute and relative. Absolute radiometric correction converts the digital number of a pixel to a percent reflectance value using established transformation

equations or atmospheric models (Richter, 1990; Song et al., 2001). Relative radiometric correction normalizes multiple satellite scenes to each other. For both categories, the majority of techniques developed require ancillary data or the manual selection of pseudo-invariant features (PIFs) in the imagery. A more exhaustive comparison of such techniques can be found

in Yuan and Elvidge (1996), Yang and Lo (2000), and Song et al. (2001).

Most forms of absolute radiometric correction rely on any combination of sensor calibration coefficients, atmospheric correction algorithms, and illumination and observation geometry coefficients. These data are used in a radiative transfer model to correct the imagery to reflectance values. Although a considerable amount of research into radiative transfer models has been conducted, the application of these models to a satellite scene often requires both atmospheric and sensor properties for the acquisition date of that scene. For the majority of archived satellite data detailing remote areas, the atmospheric properties are not readily available.

Relative radiometric correction is usually simpler than absolute radiometric correction and requires less computer operating time and less theoretical understanding. However, relative correction may hinder comparisons between different sensor types and yields values that are unitless as opposed to direct measurements of reflectance. Thus image data that have been corrected relatively are only comparable to other imagery within that set.

Relative radiometric correction often involves the selection of ground targets whose reflectance values are considered constant over time, otherwise known as PIFs, and relating these targets to all imagery in the study. Selection of such ground targets results in radiometric normalization that is entirely dependent on the abilities and local knowledge of the analyst. There are five generally accepted criteria for a PIF or PIF set (Eckhardt et al., 1990): (*i*) the targets should be approximately the same elevation so that the thickness of the atmosphere over each target is approximately the same; (*ii*) the targets should contain only minimal amounts of vegetation because vegetation spectral reflectance is subject to change over time; (*iii*) the targets must be in relatively flat areas so that changes in sun angle between images will produce the same proportional increases or decreases in insolation to all normalization targets; (*iv*) the spatial pattern of the normalization target should not change over time; and (*v*) a set of targets must have a wide range of brightness values for the regression model to be reliable.

Features used as PIFs in previous studies have included lakes, beaches, new asphalt, old asphalt, concrete, and gravel (Caselles and Lopez Garcia, 1989; Coppin and Bauer, 1994; Elvidge et al., 1995; Pax Lenney et al., 1996; Yuan and Elvidge, 1996; Michener and Houhoulis, 1997; Yang and Lo, 2000). In many studies, the selection of appropriate PIF sets is not problematic, and high-quality radiometric correction is possible. In other areas, however, the presence of suitable PIFs can be confounded by any combination of variable cloud cover, variable climate leading up to the date of image capture, high topographic complexity in the imagery, and lack of urban development. Additionally, the longer the time interval between satellite images, the higher the probability that any given pixel will have experienced change. It has been noted that, in some areas, manually determinable PIFs with constant reflectance do not exist (Du et al., 2001; 2002; Yuan and Elvidge, 1996); this is particularly true for remote areas with little or no urban development (Olthof et al., 2005).

The time period covered in this study is 18 years, represented by a series of eight satellite images. Because of this long time interval, the study area has experienced a substantial degree of change. The terrain surfaces include forests, lakes–wetlands, and dirt or two-lane roads. Although these conform to two or three of the criteria listed previously, none conform to all five criteria and only one (deep lakes) has four of the five criteria. They fail to meet all five criteria because they are the only suitable terrestrial surface, and therefore the resulting PIF set would have a very low range in brightness values. The manual selection of ground targets in this scene would therefore introduce a high degree of subjectivity. Since the concept of PIFs was first outlined (Schott et al., 1988), many quality correction techniques have been developed that rely on PIFs. For this study area, however, PIF features were not available.

Four techniques of radiometric correction that require little or no ancillary data or subjectivity were performed on the data. The first technique was a form of absolute radiometric correction (ATCOR) available from PCI Geomatics Enterprises Inc. (Richmond Hill, Ont.) and based on the Richter model (Richter, 1990; Franklin, 2001). Approaches similar to this have been used in previous studies, primarily the dense dark vegetation approach (Liang et al., 1997) and the modified dense dark vegetation approach (Song et al., 2001). The second technique tested adjusted the origin and scale for all images to a common origin and common scale (Yuan and Elvidge, 1996). The third technique used the major axis of scatterplots from image pairs to determine PIFs and adjusts each image based on the mean and standard deviation of the PIFs (Du et al., 2001; 2002). Similar techniques have been developed such as the ridge technique (Song et al., 2001) and the no-change set (Yang and Lo, 2000). The fourth technique performed median-based directional scaling on the transformation from the third technique.

Quality control of radiometric correction is essential to obtaining a meaningful result. The best method for assessment of the fidelity of radiometric correction is through field measurements of reflectance, but such data are rarely available. Comparing the visual appearance of multitemporal imagery is the most common method of testing the fidelity of radiometric correction techniques. Although the visual distinction between images is useful for large differences between images, it is highly prone to subjectivity when differences are more subtle. Another common method is to compare the outputs of a simple classification performed on each radiometrically corrected image in the time series. As an example of this method, training statistics are derived from one image and used to classify all imagery. Similar classification accuracies signify quality radiometric correction (Heo and Fitzhugh, 2000; Song et al., 2001). This has proven successful for coarse class structures that would not be expected to change over time, e.g., deciduous, coniferous, wetlands, lakes. Change would occur for more discrete class structures, however, such as classification of forests by levels of biomass. More simply, subtle change

detection, such as forest succession, places a higher demand on noise reduction than does land-cover change detection (Song and Woodcock, 2003).

A third method is to compare the root mean square error (RMSE) between images (Yang and Lo, 2000). This method can determine discrete differences if care is taken to ensure that the data used in the RMSE calculation have not experienced change, otherwise that change is incorporated as error. The assessment methods used in this study are similar because they measure error, but specifically for areas that are expected to experience either no change or relatively linear change.

The first goal of this study was to test techniques of radiometric correction that require no ancillary data while maintaining high radiometric consistency for multitemporal satellite images. The second goal was to develop quantitative methods for accuracy assessment that provide more detail of radiometric consistency than that provided by classification similarity across a satellite image time series.

Radiometric correction techniques

Landsat-5 thematic mapper (TM) and Landsat-7 enhanced thematic mapper plus (ETM+) imagery was obtained for the Aleza Lake Research Forest (~121 km^2) in central British Columbia. The satellite scenes were captured on 20 July 1985, 24 August 1992, 29 July 1994, 23 September 1997, 4 August 1999, 13 September 1999, 23 September 2000, and 22 July 2003. The only Landsat-7 ETM+ scenes were captured on 4 August 1999 and 23 September 2000. Each scene was image-to-image georectified to the median temporal image, 23 September 1997. The RMSEs for all georectifications were less than half a pixel.

Two of the radiometric correction techniques required a working definition for what would constitute a significant digital number for this study. A significant digital number occurs in such proportion within an image that it is unlikely to be an outlier or an extreme deviation from the mean. In previous studies, a significant digital number was defined as any having a pixel count of at least 1000 pixels (Teillet and Fedosejevs, 1995; McDonald et al., 1998; Song et al., 2001). This does not account for the spread of the dataset, however. In this study a significant digital number was mathematically defined by the following equation, making the pixel count inversely related to the spread in the dataset:

$$PC_{ij} = P_j \sigma_{ij} / \sigma_{\max j} \qquad (1)$$

where i is for a given image, j is for a given band, PC_{ij} is the minimum pixel count that will define a significant digital number, P_j is a constant arbitrary percentage of the image area, σ_{ij} is the standard deviation for old-growth spruce stands, and $\sigma_{\max j}$ is the maximum σ_{ij}.

P_j was defined such that the resulting mean PC_{ij} for all imagery of a given band was equal to 0.1% of the entire image. Significant pixel counts were defined in this way to ensure that data with a higher range of digital numbers had a PC_{ij} inversely proportional to that range. For the remainder of this paper, a

significant digital number is one that meets a minimum pixel count of PC_{ij}.

The value for σ_{ij} used old-growth spruce stands because these stands are at a life stage where growth is equal to mortality, and therefore most stand structural attributes are in equilibrium, provided the system is not changed by any kind of disturbance (Kneeshaw and Burton, 1998; Wells et al., 1998). Using a forest cover dataset, the spruce stands in the study area that were greater than 210 years old and 30 m high for all imagery were identified as old-growth spruce stands (Kneeshaw and Burton, 1998; Wells et al., 1998). The underlying assumption for the use of old-growth spruce stands is that these stands do not change in reflectance over time. A forested stand will experience phenological changes throughout the growing season, even if that forested stand is in equilibrium on an annual basis. Phenological changes will alter forest reflectance, and thus removing phenological effects is one of the primary challenges in using multitemporal imagery to monitor subtle changes in forests (Song et al., 2002). Therefore, the analysis of the output of the radiometric correction techniques was assessed for possible errors resulting from phenological differences between images.

Technique 1: Richter model atmospheric correction (ATCOR; Richter, 1990)

The majority of absolute radiometric correction techniques rely on data that were not available for this time series. The algorithms available from PCI Geomatics Enterprises Inc. for ATCOR were used because the only ancillary data required are the solar zenith angle of each image and the location of old-growth spruce stands.

The atmospheric models require the selection of atmospheric properties. These are predefined and are tropical, mid-latitude, or the US standard atmosphere and are also rural, urban, desert, or maritime. For this area the US standard rural atmosphere best described the study area. The atmospheric model also uses sensor calibration defaults and solar zenith angles to calculate reflectance values. A variable, optical visibility has to be calculated for each image to perform the final algorithms in the atmospheric correction package. This requires the selection of a target with known reflectance values, which are compared with the reflectance values calculated within each image.

Technique 2: origin fix with scaling (OFS; Yuan and Elvidge, 1996)

The origin and range for each set of TM bands in the imagery were found to be variable. Minimum and maximum significant digital numbers were determined with Equation (1). The significant origin for each band was the minimum significant digital number in the band. The significant range was defined as the maximum significant digital number minus the minimum significant digital number. The true origin and range of digital numbers were not used because extremely bright and dark outlier pixels were present in some images but not in others. **Table 1** shows the true and significant origin and range for each

TM band 4 in the image series. The image data were transformed using the following equation based on Yuan and Elvidge (1996):

$$Q_{ij}^* = (Q_{ij} - O_{ij})\,(R_{\max\,j}/R_{ij}) + A_{\max\,j} \qquad (2)$$

where i is for a given image; j is for a given band; Q_{ij}^* is the adjusted pixel value of the transformation; Q_{ij} is the original pixel value; O_{ij} is the significant origin (the minimum significant digital number); R_{ij} is the significant range (the range between the maximum and minimum significant digital numbers); $R_{\max\,j}$ is the maximum R_{ij}; and $A_{\max\,j}$ is the maximum value of A_{ij}, where $A_{ij} = (Q_{ij} - O_{ij})(R_{\max\,j}/R_{ij})$ when Q_{ij} is the true origin. The terms $R_{\max\,j}$ and $A_{\max\,j}$ are used for quality control to ensure that the resulting imagery does not get compressed or contain values less than zero.

Technique 3: pseudo-invariant feature regression (PIFR; Du et al., 2001; 2002)

In previous studies, a technique for nonsubjective PIF selection was developed and showed a significant decrease in radiometric inconsistency (Du et al., 2001; 2002; Song et al., 2001). Consider the scatterplot of TM band 4 for two images (**Figure 1**). The major axis is the solid black line and the two broken parallel lines are thresholds defined by a deviation l from the major axis. For this technique, the variations in pixel values during the period represented by the scatterplot are assumed to be linear, spatially homogeneous, and normally distributed. All pixels that fall within the threshold are considered PIFs for that image pair. A scatterplot was created for each TM band using the median temporal image (23 September 1997) as the y axis and every other image as an x axis in a separate scatterplot. As the time series has eight images, seven comparisons with the 23 September 1997 image were possible. Any pixel that fell within the thresholds for every scatterplot (i.e., for all image bands) is considered a PIF. Thresholds were determined by ensuring that the scatterplot of each image pair under the resultant PIFs had a correlation coefficient (r^2) greater that 0.9 and selected a minimum of 300 pixels (~0.07% of the image

area) to ensure a quality transformation and a maximum of 900 pixels (~0.20% of the image area) to minimize the presence of outliers due to land cover changes in the transformation. If any one of these three criteria was not met, the deviation l from the major axis was adjusted to determine new thresholds. The minimum and maximum pixel counts were chosen because lower values tended to produce erratic transformation coefficients and higher values tended to include image elements that had undergone change in one or more images.

Each TM band has a single dataset describing the PIFs for every image in that band. For each TM band, the appropriate PIF dataset is used to calculate the standard deviation and mean for the pixel values for every image. **Table 2** shows the standard deviations and means for TM band 4 and the gains and offsets used in the image transformations.

Gain for any given image is defined as the maximum standard deviation for all imagery for a given TM band divided by the standard deviation of the TM band for that given image. This ensures that the gain is greater than or equal to one so that the data are stretched (not compressed). S_{ref} is defined as the gain multiplied by the mean. The offset for any given image is defined as the maximum S_{ref} for all imagery of a given TM band minus the S_{ref} of the TM band for that given image. This calculation for offset ensures that it is greater than or equal to zero, so that the output data cannot be negative. Because of the large variation in the standard deviation and mean of the imagery under the PIFs in this study, the resulting gain and offset also have large variation. When transformed using the gain and offset, the imagery falls outside of the eight-bit range (0–255), and therefore 32-bit storage was used for the

Table 1. True and significant origin and range for Landsat thematic mapper band 4.

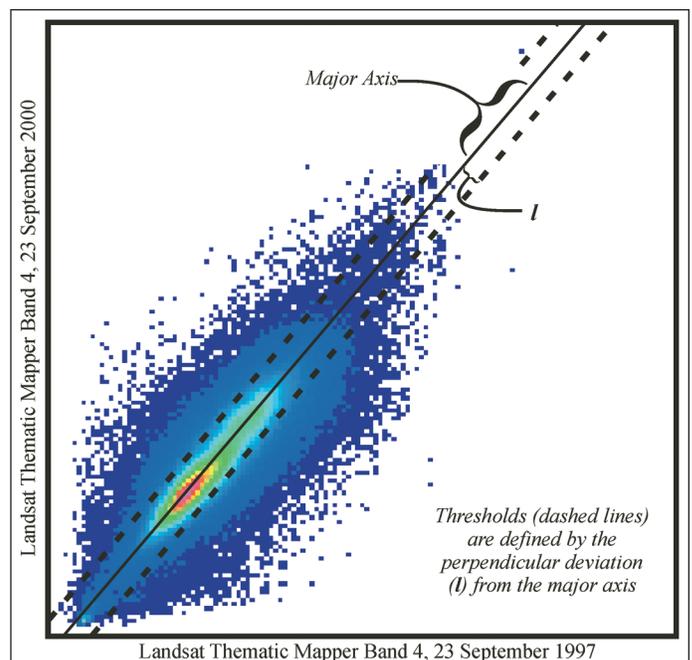| Image date | True | | Significant | | R_{true}/R_{ij} |
	Origin	Range (R_{true})	Origin (O_{ij})	Range (R_{ij})	
20 July 1985	6	131	9	105	1.25
24 August 1992	6	99	13	73	1.36
29 July 1994	5	129	8	103	1.25
23 September 1997	2	87	7	57	1.53
4 August 1999	12	193	18	147	1.31
13 September 1999	0	119	5	78	1.53
23 September 2000	13	115	14	70	1.64
29 July 2003	4	149	7	114	1.31

Figure 1. Scatterplot of Landsat thematic mapper band 4 for two images, showing the major axis and the deviation l from the major axis to determine thresholds.

Table 2. Pseudo-invariant feature statistics for Landsat thematic mapper band 4.

Image date	SD	Mean	Gain	Offset	S_{ref}
20 July 1985	13.86	60.08	1.15	71.79	68.84
24 August 1992	7.04	50.03	2.25	27.85	112.78
29 July 1994	12.28	56.42	1.29	67.68	72.95
23 September 1997	3.92	33.27	4.05	6.03	134.60
4 August 1999	15.88	88.44	1.00	52.19	88.44
13 September 1999	6.99	39.90	2.27	49.97	90.66
23 September 2000	5.54	49.05	2.87	0.00	140.63
29 July 2003	12.53	60.75	1.27	63.67	76.96

Note: SD, standard deviation.

transformation output. The output was then scaled as 32-bit data to match the 0–255 range of the other radiometric correction techniques.

Technique 4: median-based directional scaling (MBDS)

The large variation in gain and offset from the PIFR technique resulted in the minimum and maximum significant digital numbers being highly variable across the imagery for a given TM band using PIFR radiometric correction. To account for this, median-based directional scaling was performed on the resulting imagery from the PIFR technique. The median was chosen over the mean because the median is less sensitive to these extreme outliers. For each channel, two range calculations were performed using the following equation, the median minus the minimum significant digital number and the maximum significant digital number minus the median:

$$Q_{ij}^* = (Q_{ij} - m_{ij})\ (R_{\max\ j}^{\mathrm{a}}/R_{ij}^{\mathrm{a}}) + m_{ij} \qquad (3)$$

where i is for a given image, j is for a given band, Q_{ij}^* is the adjusted pixel value of the transformation, Q_{ij} is the original pixel value, m_{ij} is the median value, R_{ij}^{a} is the range above or below the median, and $R_{\max\ j}^{\mathrm{a}}$ is the maximum range above or below the median

Radiometric correction accuracy assessment

In addition to change detection studies, radiometric correction is also utilized in satellite imagery mosaics where adjacent images are aligned. In most of these studies the only method of quality control is edge detection along the image transition seams. Although this method gives a strong indication of the quality of image matching, the assessment is qualitative and subjective. The radiometric correction technique is deemed accurate when the results match what is desired and not necessarily what is correct (Du et al., 2001). A mosaic of the imagery from this study was used to artificially create a seam for every image pair. The mosaic is composed of alternating rows, in which the first row contains eight cells with each cell depicting a different image and the second row is a
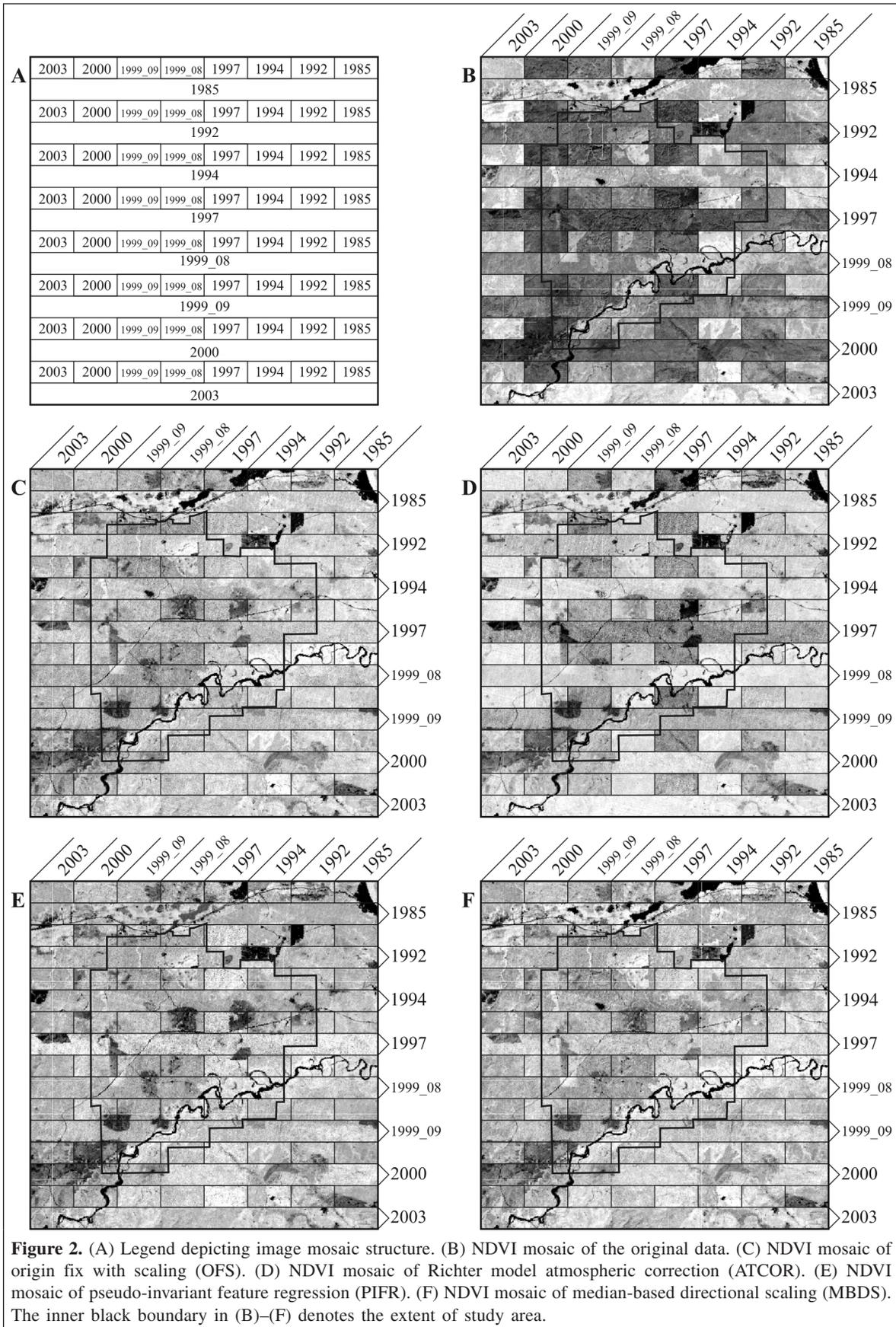
single cell depicting a single image. A mosaic of the normalized difference vegetation index (NDVI) from each image was used because it enabled easier edge detection (**Figure 2**). OFS performed better than ATCOR but not quite as well as PIFR and MBDS. The PIFR technique had results very similar to those of the MBDS technique for the majority of the scene. Extreme dark and bright objects, such as the lake in the top row in the 1997 and 1994 cells, usually had a more significant edge with the PIFR technique than with the MBDS technique.

Since edge detection is highly qualitative, it was coupled with quantitative analysis. Two forms of quantitative assessment that do not rely on field measurements or classification similarity were performed on the radiometric correction techniques. The first was a comparison of the old-growth spruce stands, which were examined for the similarity between means and standard deviations for each TM band. Although phenological differences would have occurred between image dates, these changes are thought to be relatively small. The best technique will be that which produces the most identical reflectance patterns for these old-growth stands (**Figures 3, 4**)

The mean pixel value for each image in a TM band group should be equal with perfect radiometric correction under the assumption of constant reflectance over time in old-growth spruce stands, and therefore a line with zero slope in **Figures 3 and 4** would illustrate perfect radiometric correction.

Under perfect radiometric correction, the means for a spectrally constant spatial unit should be identical across all imagery for a given band and a given technique. Likewise, the range of digital numbers for a given spectrally constant spatial unit should also be identical. For this assessment of accuracy, the range of digital numbers is indicated by standard deviations. For a given band and technique, eight measurements of the mean and standard deviation were calculated for old-growth spruce stands, one measurement for each image. The correction techniques were evaluated according to the consistency of these values (**Tables 3, 4**). The MBDS performed slightly better than the PIFR and significantly better than the others. **Table 3** shows that the MBDS performed the best with two exceptions. PIFR was slightly better for TM band 1, and ATCOR was better for TM band 5. In **Table 4**, MBDS is significantly better overall than every other technique.

The other form of quantitative assessment performed on the data was a measurement of growth in clearcuts that were planted prior to the first image in the time series and have undergone no silvicultural treatment that would affect biomass, stand volume, or leaf area index (LAI), such as fertilization, thinning, manual brushing, and herbicide spraying. The expected growth pattern of a forest stand after a stand-initiating event is highly dependent on site factors, species, and climate. Within a short time frame, the annual increases in LAI and crown cover of a forest have been found to be relatively constant (Coppin and Bauer, 1994). There is a strong correlation between either of LAI or crown cover with NDVI in developing coniferous stands (Ripple et al., 1991; Gong et al., 1995; Chen, 1996; McDonald et al., 1998; Franklin, 2001).

Figure 2. (A) Legend depicting image mosaic structure. (B) NDVI mosaic of the original data. (C) NDVI mosaic of origin fix with scaling (OFS). (D) NDVI mosaic of Richter model atmospheric correction (ATCOR). (E) NDVI mosaic of pseudo-invariant feature regression (PIFR). (F) NDVI mosaic of median-based directional scaling (MBDS). The inner black boundary in (B)–(F) denotes the extent of study area.
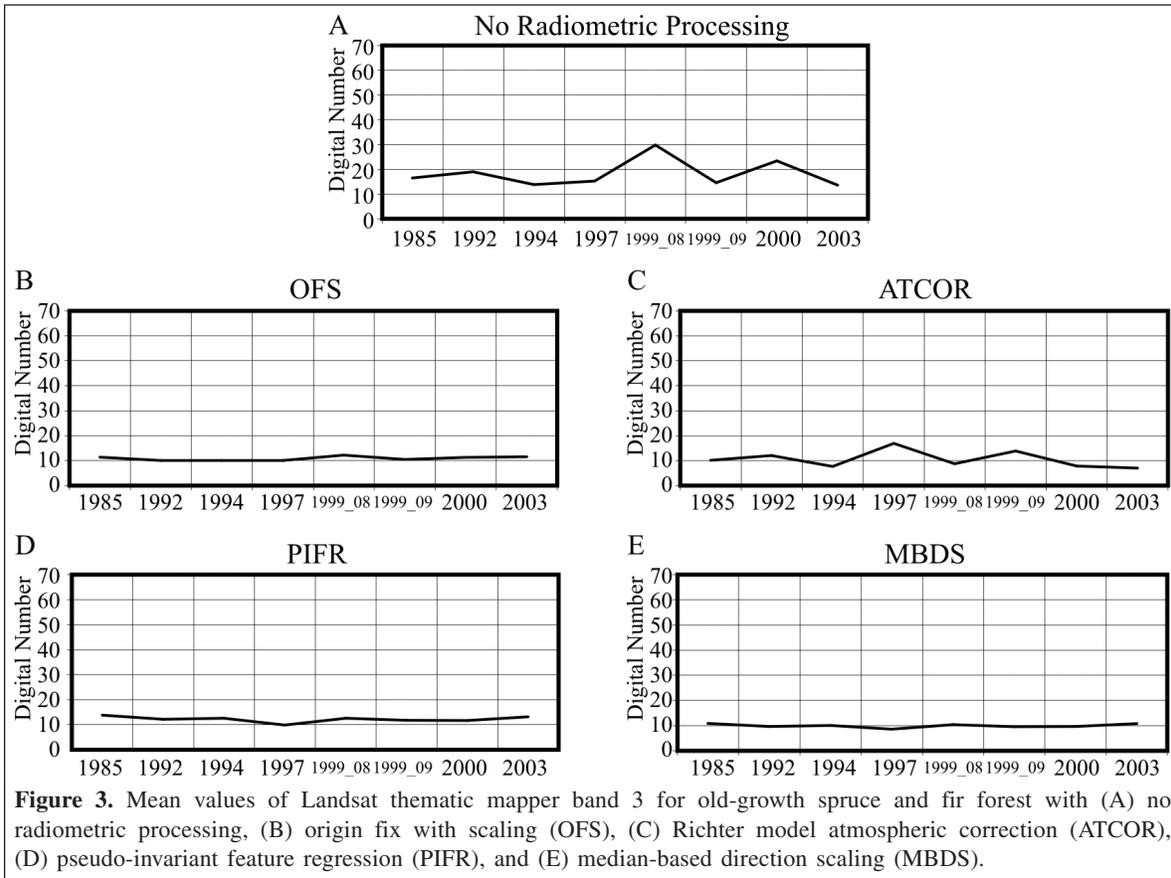
Figure 3. Mean values of Landsat thematic mapper band 3 for old-growth spruce and fir forest with (A) no radiometric processing, (B) origin fix with scaling (OFS), (C) Richter model atmospheric correction (ATCOR), (D) pseudo-invariant feature regression (PIFR), and (E) median-based direction scaling (MBDS).

Table 3. Standard deviations of mean old-growth radiance across all years by radiometric correction technique for thematic mapper bands 1–5 and 7 (TM1–TM5 and TM7).

Technique	TM1	TM2	TM3	TM4	TM5	TM7
Original	5.75	8.45	5.69	14.58	6.32	4.45
OFS	1.52	1.20	0.85	4.48	5.53	2.39
ATCOR	5.93	2.70	3.48	7.98	1.80	2.99
PIFR	0.78	1.52	1.16	3.23	2.72	0.80
MBDS	0.89	0.76	0.73	1.89	2.42	0.81

Note: Lower values indicate higher homogeneity of the means.

Table 4. Standard deviations of standard deviations of old-growth radiance across all years by radiometric correction technique for thematic mapper bands 1–5 and 7 (TM1–TM5 and TM7).

Technique	TM1	TM2	TM3	TM4	TM5	TM7
Original	0.32	0.35	0.17	2.44	1.08	0.46
OFS	0.47	0.34	0.30	0.56	0.89	0.44
ATCOR	2.11	0.98	1.45	1.88	1.53	1.47
PIFR	0.52	0.86	0.43	2.21	1.38	0.61
MBDS	0.22	0.30	0.09	0.38	0.38	0.29

Note: Lower values indicate higher homogeneity of the standard deviations.

Three cutblocks, clearcut harvested and replanted with white spruce, were used for linear regression analysis with mean NDVI as the dependent variable and number of months since planting as the independent variable. Each cutblock was analyzed separately, and mean NDVI was calculated for eight approximately equal areas in each stand to increase the number of samples. The average annual NDVI increase and r^2 value for each clearcut and each radiometric correction technique are shown in **Table 5**. Although the calculation of absolute values of the NDVI depend on the radiometric correction technique, the average annual increase for NDVI on these cutblocks was approximately 0.0031–0.0047 for all radiometric correction techniques.

All forms of radiometric correction improved the correlation between NDVI and time since planting as compared with the original data. In all cases, PIFR and MBDS performed better than the other techniques. In terms of average r^2 value, the PIFR technique yielded slightly better results than the MBDS technique. In all cases, the greatest sources of error in regression were the images from 23 September 2000, 13 September 1999, and 23 September 1997, indicating a strong phenological effect in these young stands. The calculation of r^2 values using all non-September imagery increased the average r^2 value to 0.67 for the original data and 0.87 for PIFR; however, the number of samples was reduced to 40 from 64 with the omission of the September imagery. The order of r^2 values from non-September imagery by radiometric correction technique was the same as the order of r^2 values by radiometric correction technique from **Table 3**. **Table 6** shows the average annual NDVI increase and r^2 values by radiometric correction technique with the omission of September imagery.
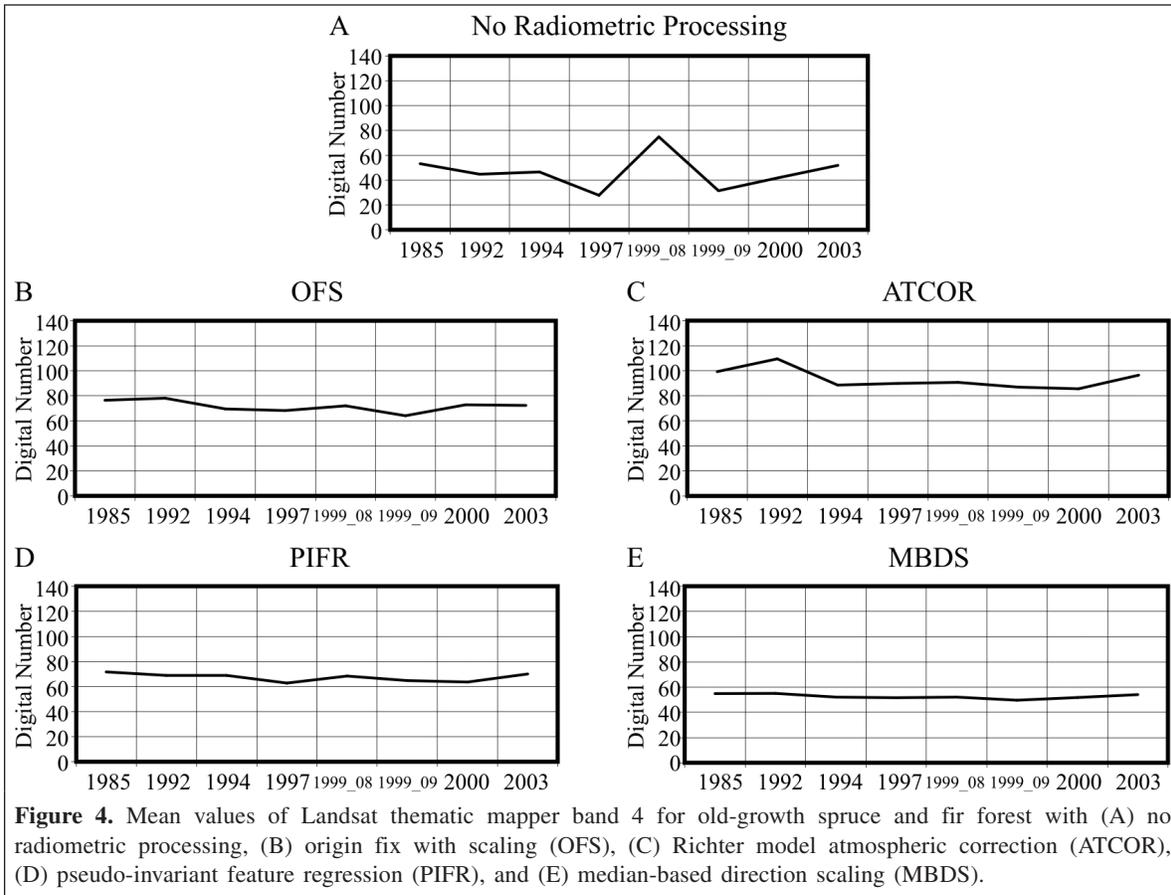
Figure 4. Mean values of Landsat thematic mapper band 4 for old-growth spruce and fir forest with (A) no radiometric processing, (B) origin fix with scaling (OFS), (C) Richter model atmospheric correction (ATCOR), (D) pseudo-invariant feature regression (PIFR), and (E) median-based direction scaling (MBDS).

Table 5. Linear regression statistics for normalized difference vegetation index (NDVI) values versus time for three cutblocks for each radiometric correction technique.

Technique	Stand 1 Annual NDVI increase	r^2	Stand 2 Annual NDVI increase	r^2	Stand 3 Annual NDVI increase	r^2
Original	0.0081	0.08	0.0070	0.06	0.0211	0.30
ATCOR	0.0112	0.25	0.0110	0.19	0.0250	0.37
OFS	0.0088	0.53	0.0085	0.38	0.0064	0.45
PIFR	0.0090	0.64	0.0083	0.49	0.0127	0.80
MBDS	0.0100	0.64	0.0094	0.45	0.0083	0.71

Note: All p values significant at $\alpha = 0.05$.

Table 6. Linear regression statistics for normalized difference vegetation index (NDVI) versus time for three cutblocks for each radiometric correction technique with the omission of all September imagery.

Technique	Stand 1 Annual NDVI increase	r^2	Stand 2 Annual NDVI increase	r^2	Stand 3 Annual NDVI increase	r^2
Original	0.0152	0.58	0.0151	0.53	0.0278	0.81
ATCOR	0.0154	0.73	0.0162	0.65	0.0282	0.81
OFS	0.0109	0.83	0.0116	0.75	0.0070	0.67
PIFR	0.0099	0.87	0.0103	0.85	0.0122	0.87
MBDS	0.0116	0.83	0.0123	0.77	0.0088	0.80

Note: All p values significant at $\alpha = 0.05$.

Discussion

All four radiometric correction techniques vastly improved the radiometric consistency from the original images. The poorest results were produced by the ATCOR technique, probably due to a weak correspondence between the aerosol model used and the study area conditions. With the inclusion of atmospheric data and an aerosol model specifically developed for this study area, the ATCOR technique may have yielded a better result.

The performance of the OFS technique was of medium quality. It may be suitable for low-detail classifications or change detection analysis, but not for studies analyzing subtle differences within the imagery. The output produced by PIFR and MBDS was significantly better than that by the other techniques. PIFR produced results that appear to be slightly better for monitoring vegetative growth in disturbed areas (**Tables 5, 6**), and the MBDS appears to be slightly better for monitoring overall vegetative presence (**Tables 3, 4**). The MBDS yielded the best visual result, similar to PIFR, except for extremely dark and extremely bright objects where MBDS performed better than PIFR.

Table 7. Pixel count and percentage of pseudo-invariant features by thematic mapper band and type.

Band	Deep dark lakes Count	Percent	Late seral stage deciduous Count	Percent	Late seral stage coniferous Count	Percent	Late seral stage mixedwood Count	Percent	Total count
TM1	9	2.8	26	8.2	269	84.6	14	4.4	318
TM2	0	0.0	9	3.5	225	86.5	26	10.0	260
TM3	0	0.0	92	17.3	330	62.0	110	20.7	532
TM4	0	0.0	10	1.5	636	93.9	31	4.6	677
TM5	0	0.0	216	27.1	377	47.3	204	25.6	797
TM7	0	0.0	180	29.1	315	50.9	124	20.0	619

Table 8. Ranking of radiometric correction technique performance from best (row 1) to worst (row 5, i.e., original data).

Edge detection	Old-growth assessment	NDVI versus time regression
MBDS	MBDS	PIFR
PIFR	PIFR	MBDS
OFS	OFS	OFS
ATCOR	ATCOR	ATCOR
Original data	Original data	Original data

The PIF set selected for each TM band was of particular interest, as the automatically selected targets were unrelated to the features that have been selected manually in other studies (Caselles and Lopez Garcia, 1989; Coppin and Bauer, 1994; Elvidge et al., 1995; Pax Lenney et al., 1996; Yuan and Elvidge, 1996; Michener and Houhoulis, 1997; Yang and Lo, 2000). Many of these studies used urban areas, which are not present in this study area. The study area is roughly 90% forest, 4% lakes, 4% fields, and 2% other features. Of the forested area, roughly 75% is coniferous, and mixedwoods and deciduous equally share the remaining forested area. **Table 7** shows the count of PIFs for each band for the features selected by the PIFR technique. Although deep, dark lakes are probably the most commonly selected ground target, this study indicates that the reflectance of the lakes in this study area was not constant. This discrepancy could be a mere anomaly or could be caused by a lack of lake depth, a high-wind event, or variable sediment loads, among other possibilities.

The use of unnecessary radiometric precision leads to increased computational complexity, cost, and time (Duggin and Robinove, 1990). The computational complexity is highest for the ATCOR techniques. OFS, MBDS, and PIFR all use very simple image transformation functions, which are easily automated and are suitable for large datasets. The outputs from MBDS and PIFR for this study required 16- or 32-bit data to prevent loss of radiometric range, however. Overall, MBDS produces the most widely applicable results, although PIFR may be more specifically suited to studies interested solely in low-density vegetation because it produced more consistent results for clearcuts.

OFS, PIFR, and MBDS are all radiometric correction techniques that could be automated in most commercial image processing software packages. For PIFR and MBDS, the only inputs beyond image channels would be a minimum acceptable correlation coefficient and, for MBDS, a mask describing a no-change area to calculate significant pixel counts. A minimum correlation coefficient of 0.90 was possible, even with the small area extent of this study ($\sim$121 km^2).

MBDS and OFS require that the highest and lowest significant reflectance be determined across all imagery. Significant reflectance is the reflectance over an arbitrarily defined area-based proportion of the time series of imagery, i.e., 0.1% of the image area. An example of a feature in this study that violates that requirement would be white clouds. Any such features should be masked out for scaling operations, bearing in mind that this will reduce the effective image area. The equation for calculating PC$_{ij}$ incorporates the standard deviation of old-growth spruce stands which may not be present in other studies. A quick comparison was performed on the output range of significant digital numbers (DNs) using PC$_{ij}$ equal to an arbitrary constant, and the results were similar to those using PC$_{ij}$ as defined in Equation (1). The correlation between the range of significant DNs and the true range of DNs was higher using Equation (1), but not significantly.

The assumption of constant reflectance in old-growth spruce forests appears to be true for this study area. For every TM band, none of the radiometric correction techniques or the original data indicated a consistent decline or increase in reflectance for these forests over time. The radiometric correction technique that performed best was MBDS, which in all cases was only slightly better than PIFR.

The assumption of constant or near-constant rates of increase for the NDVI over regenerating clearcuts over time also appears valid. An increase in the correlation between NDVI and time since planting for these regenerating clearcuts was observed from the original data to those from the PIFR technique. For this method of assessment, PIFR was deemed to be slightly more accurate than MBDS. Although inclusion of the September imagery was not statistically detrimental to the old-growth analysis, it was found to confuse the relationship between mean NDVI and time.

The ranking of each radiometric technique by assessment method, from best to worst, is shown in **Table 8**. All three

methods of assessment show similar results in terms of these rankings, although determining a ranking from edge detection is highly subjective, especially for the comparison among OFS, PIFR, and MBDS.

Conclusions

Many change detection study areas have few ground targets of constant reflectance that can be used for radiometric correction, especially in remote areas. This study utilizes radiometric correction techniques that require either no ground targets or only one target of constant reflectance. Although the literature has stated that areas with vegetation should be avoided as a ground target, they were found in this study area to be the most accurate in terms of having constant reflectance. Assessment methods showed that the phenological shift in reflectance of old-growth spruce stands from July to September was minimal compared with the overall reflectance of these stands.

The two best techniques of this study could be easily automated and incorporated into available remote sensing software. The MBDS and PIFR techniques demonstrated relative strengths and weaknesses, which should be a consideration in their selection. Overall, MBDS performed better for very dark and bright objects and for reducing edge effects, and PIFR performed better for discrete measurements of change in objects relatively close to the median digital number.

Acknowledgments

References

Caselles, V., and Lopez Garcia, M.J. 1989. An alternative simple approach to estimate atmospheric correction in multitemporal studies. *International Journal of Remote Sensing*, Vol. 10, pp. 1127–1134.

Chen, J.M. 1996. Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing*, Vol. 22, pp. 836–844.

Coppin, P.R., and Bauer, M.E. 1994. Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, pp. 918–927.

Du, Y., Cihlar, J., Beaubien, J., and Latifovic, R. 2001. Radiometric normalization, compositing, and quality control for satellite high resolution image mosaics over large areas. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, pp. 623–634.

Du, Y., Tiellet, P.M., and Cihlar, J. 2002. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sensing of Environment*, Vol. 82, pp. 123–134.

Duggin, M.J., and Robinove, C.J. 1990. Assumptions implicit in remote sensing data acquisition and analysis. *International Journal of Remote Sensing*, Vol. 11, pp. 1669–1694.

Eckhardt, D.W., Verdin, J.P., and Lyford, G.R. 1990. Automated update of an irrigated lands GIS using SPOT HRV imagery. *Photogrammetric Engineering and Remote Sensing*, Vol. 56, pp. 1515–1522.

Elvidge, C.D., Yan, D., Ridgeway, D.W., and Lunetta, R.S. 1995. Relative radiometric normalization of Landsat Multispectral Scanner (MSS) data using an automatic scattergram-controlled regression. *Photogrammetric Engineering and Remote Sensing*, Vol. 61, pp. 1255–1260.

Franklin, S.E. 2001. *Remote sensing for sustainable forest management*. Lewis Publishers, Washington, D.C.

Gong, P., Pu, R., and Miller, J.R. 1995. Coniferous forest leaf area index estimation along the Oregon transect using compact airborne spectrographic imager data. *Photogrammetric Engineering and Remote Sensing*, Vol. 61, pp. 1107–1117.

Heo, J., and Fitzhugh, T.W. 2000. A standardized radiometric normalization method for change detection using remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing*, Vol. 66, pp. 173–181.

Kneeshaw, D.D., and Burton, P.G. 1998. Assessment of functional old-growth status: a case study in the sub-boreal spruce zone of British Columbia, Canada. *Natural Areas Journal*, Vol. 18, pp. 293–308.

Liang, S., Fallah-Adl, H., Kalluri, S., Jaja, J., Kaufman, Y.G., and Townshend, J.R.G. 1997. An operational atmospheric correction algorithm for Landsat Thematic Mapper imagery over the land. *Journal of Geophysical Research*, Vol. 102, pp. 173–186.

McDonald, A.J., Gemmell, F.M., and Lewis, P.E. 1998. Investigation of the utility of spectral vegetation indices for determining information on coniferous forests. *Remote Sensing of Environment*, Vol. 66, pp. 250–272.

Michener, W.K., and Houhoulis, P.F. 1997. Detection of vegetation changes associated with extensive flooding in a forested ecosystem. *Photogrammetric Engineering and Remote Sensing*, Vol. 63, pp. 1363–1374.

Olthof, I., Pouliot, D., Fernandes, R., and Latifovic, R. 2005. Landsat-7 ETM+ radiometric normalization comparison for northern mapping applications. *Remote Sensing of Environment*, Vol. 95, pp. 388–398.

Pax Lenney, M., Woodcock, C.E., Collins, J.B., and Hamdi, H. 1996. The status of agricultural lands in Egypt: the use of multitemporal NDVI features derived from Landsat TM. *Remote Sensing of Environment*, Vol. 47, pp. 390–402.

Richter, R. 1990. A fast atmospheric correction algorithm applied to Landsat TM images. *International Journal of Remote Sensing*, Vol. 11, pp. 159–166.

Ripple, W.J., Wang, S., Isaacson, D.L., and Paine, D.P. 1991. A preliminary comparison of Landsat Thematic Mapper and SPOT-1 HRV multispectral data for estimating coniferous forest volume. *International Journal of Remote Sensing*, Vol. 12, pp. 1971–1977.

Schott, J.R., Salvaggio, C., and Volchok, W.J. 1988. Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment*, Vol. 26, pp. 1–16.

Song, C., and Woodcock, C.E. 2003. Monitoring forest succession with multitemporal Landsat images: factors of uncertainty. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, pp. 2557–2567.

Song, C., Woodcock, C.E., Seto, K.C., Pax Lenney, M., and Macomber, S.A. 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment*, Vol. 75, pp. 230–244.

Song, C., Woodcock, C.W., and Li, X. 2002. The spectral/temporal manifestation of forest succession in optical imagery: the potential for multitemporal imagery. *Remote Sensing of Environment*, Vol. 82, pp. 285–302.

Teillet, P.M., and Fedosejevs, G. 1995. On the dark target approach to atmospheric correction of remotely sensed data. *Canadian Journal of Remote Sensing*, Vol. 21, pp. 374–387.

Wells, R.W., Lertzman, K.P., and Saunders, S.C. 1998. Old-growth definitions for the forests of British Columbia, Canada. *Natural Areas Journal*, Vol. 18, pp. 279–292.

Yang, X., and Lo, C.P. 2000. Relative radiometric normalization performance for change detection from multi-date satellite images. *Photogrammetric Engineering and Remote Sensing*, Vol. 66, pp. 967–980.

Yuan, D., and Elvidge, C.D. 1996. Comparison of relative radiometric normalization techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 51, pp. 117–126.