

Radiometric correction effects in Landsat multi-date/multi-sensor change detection studies

LEONARDO PAOLINI*†, FRANCISCO GRINGS‡, JOSÉ A. SOBRINOS,
JUAN C. JIMÉNEZ MUÑOZ§ and HAYDEE KARSZENBAUM‡

†Laboratorio de Investigaciones Ecológicas de las Yungas, Universidad Nacional de Tucumán, Casilla de Correo 34, 4107 Yerba Buena, Tucumán, Argentina

‡Instituto de Astronomía y Física del Espacio, Universidad Nacional de Buenos Aires-CONICET, Ciudad Universitaria, 1428 Buenos Aires, Argentina

§Unidad de Cambio Global, Departamento de Termodinámica, Facultad de Física, Universidad de Valencia, Dr. Moliner 50, 46100 Burjassot, Spain

Radiometric corrections serve to remove the effects that alter the spectral characteristics of land features, except for actual changes in ground target, becoming mandatory in multi-sensor, multi-date studies. In this paper, we evaluate the effects of two types of radiometric correction methods (absolute and relative) for the determination of land cover changes, using Landsat TM and Landsat ETM+ images. In addition, we present an improvement made to the relative correction method addressed. Absolute correction includes a cross-calibration between TM and ETM+ images, and the application of an atmospheric correction protocol. Relative correction normalizes the images using pseudo-invariant features (PIFs) selected through band-to-band PCA analysis. We present a new algorithm for PIFs selection in order to improve normalization results. A post-correction evaluation index (Quadratic Difference Index (QD)), and post-classification and change detection results were used to evaluate the performance of the methods. Only the absolute correction method and the new relative correction method presented in this paper show good post-correction and post-classification results (QD index ≈ 0 ; overall accuracy $>80\%$; kappa >0.65) for all the images used. Land cover change estimations based on uncorrected images present unrealistic change rates (two to three times those obtained with corrected images), which highlights the fact that radiometric corrections are necessary in multi-date multi-sensor land cover change analysis.

1. Introduction

Detection, evaluation, and prediction of changes in natural environments are among the most important tasks in landscape ecology studies. Land cover change is the single most important variable of global change affecting ecological systems (Vitousek 1994), with an impact on the environment that is at least as great as that associated with climate change (Skole 1994). For many purposes, remote sensing provides the only means to assess habitat structure and land cover changes across broad areas (Foody 2003, Kerr and Ostrovsky 2003, Turner *et al.* 2003). The spatial and temporally repeated observations have significantly improved the quantity and quality of these environmental data. The information derived from satellite images

*Corresponding author. Email: leo_paolini@fibertel.com.ar

is becoming increasingly important for mapping and monitoring land cover (Pax-Lenney *et al.* 2001, Watson and Wilcock 2001, Scanlon *et al.* 2002, Fuller *et al.* 2003) and for detecting natural and human-induced environmental changes (Abuelgasim *et al.* 1999, Helmer *et al.* 2000, Mukai and Hasegawa 2000, Rogan and Yool 2001, Song *et al.* 2001, Miller and Yool 2002, Rogan *et al.* 2002, Ress *et al.* 2003).

A wide variety of methods have been developed and used to study land cover change (e.g. Singh 1989, Muchoney and Haack 1994, Collins and Woodcock 1996, Coppin and Bauer 1996, Kerr and Ostrovsky 2003, Turner *et al.* 2003, Coppin *et al.* 2004). Simple image differencing, differencing of images of derived indices, such as Normalized Difference Vegetation Index (NDVI), and change vector analysis (CVA) are mostly linear methods, which estimate change on the basis of a linear combination of the input image spectral bands. Other methods, for example principal-component analysis (PCA), use the statistical properties of the image to extract the change component, assuming that the variability caused by real change is different from other sources of variability in the images. But probably, the most popular approach in change detection analysis is the post-classification comparison method (Foody 2002), where all the images used are classified separately (based only on the information contained in each image), and then the thematic maps generated are compared. In change-detection analysis, data generalization across time and space is necessary. When using multi-date multi-sensor images, the post-classification comparison method could lead to wrong results due to the differences in the radiometric characteristics of the images from which thematic maps were obtained.

The underlying assumption when using remotely sensed data for change detection is that land cover changes can be recorded as significant differences in the spectral characteristics of the affected area between two or more dates. In addition, these differences must be larger or somehow distinguishable from other changes in the images, such as those related to atmospheric conditions (Song *et al.* 2001, Lu *et al.* 2002), illumination geometry, phenological variability (Lunetta *et al.* 2002, Rogan *et al.* 2002, Fuller *et al.* 2003), and sensor calibration (Teillet *et al.* 2001, Vogelmann *et al.* 2001). When using images from two or more dates in change detection analysis, a given target can have a different sensor's radiometric responses over time, due to several factors, including:

- differences in relative radiometric response between sensors;
- changes in satellite sensor calibration over time (i.e. aging);
- differences in illumination and observation angles;
- variation in atmospheric effects;
- reflectance anisotropy (i.e. BRDF effects);
- topography (i.e. slope-aspect effects); and
- actual changes in target reflectance.

The goal of radiometric corrections is to remove or compensate for all the above effects except for actual changes in ground target. For change-detection studies, some form of image matching or radiometric calibration is recommended to eliminate exogenous differences (Coppin *et al.* 2004). In this way, radiometrically corrected images should appear as if they were acquired with the same sensor and under the same atmospheric and illumination conditions.

Change-detection analysis using Landsat Thematic Mapper (TM) and Landsat Enhanced Thematic Mapper plus (ETM+) images requires the assessment of

radiometric consistency to ensure inter-image comparability, regarding the differences in relative radiometric response between corresponding Landsat-4/5 TM and Landsat-7 ETM+ spectral bands (Teillet *et al.* 2001). In this case, radiometric correction should be used not only to remove the atmospheric, illumination/observation and radiometric sensor aging effects, but also to reduce the radiometric discrepancy between TM and ETM+ sensors. This is a key factor when training data from one image (i.e. spectral signature) will be used to classify another image (Song *et al.* 2001, Woodcock *et al.* 2001). Otherwise, in multi-date multi-sensor analysis, the use of images none corrected radiometrically may lead to misclassification and incorrect land cover change estimations.

In subtropical ecosystems of north-western Argentina, remotely sensed data can be a very valuable tool to estimate land cover change and evaluate its impact on the ecosystems dynamics, especially regarding its large extension and environmental complexity. Recent studies conducted in this region have successfully used Landsat TM image to detect natural disturbances (Blodgett 1998, Paolini *et al.* 2002, Grau 2001). Although these works have achieved good discrimination of land cover categories, none of them has used multi-date, multi-sensor images for classification and change detection analysis. In land cover change studies of large areas, data generalization could be very useful. For example, classification results of the present date image can be easily checked through field surveys and then used as training data to classify past-time images. This kind of generalization will be dependent on the ability to perform good radiometric calibrations of images.

For studies conducted in Argentina, the lack of specific methodologies for correcting Landsat images generated by local stations is yet an unsolved issue. In general, all the algorithms and methodologies used for radiometric corrections of Landsat images are developed specifically for images produced in Europe or USA, which has some differences with those generated in other stations (i.e. format, pre-processing level, etc.). For example, the cross-calibration method between Landsat TM and ETM+ sensors developed by Teillet *et al.* (2001) is for level 0 images, whereas the images generally accessible by users in South America are of level 1. Thus, it will be of great help to test and improve these methodologies in order to be used with images locally generated.

The main goal of this paper is to evaluate the effect of two different radiometric correction algorithms, one absolute and other relative, based on post-correction, post-classification, and change-detection analysis, in a multitemporal study involving Landsat TM and ETM+ images. Although, in past studies, Collins and Woodcock (1996) have already examined the effects of different radiometric corrections levels on change detection results, their research was based on only one sensor type (Landsat TM), and the pseudo-invariant features (PIFs) for the relative correction process were selected through a subjective procedure.

2. Radiometric correction methods

Radiometric correction methods of satellite images can be grouped in two major categories: absolute and relative (Thome *et al.* 1997, Tokola *et al.* 1999, Teillet *et al.* 2001, Vogelmann *et al.* 2001, Du *et al.* 2002). Absolute radiometric correction converts the digital counts (Q), in which satellite image data are distributed, to at-sensor radiance ($\text{W m}^{-2} \mu\text{m}^{-1} \text{sr}^{-1}$) and then, through atmospheric correction, to reflectance at the surface of the Earth. Relative radiometric correction is used to remove or normalize the variation between images and yields radiometrically

normalized data in a common (reference) scale, not necessarily in physical units. It is also used to normalize differences between individual detectors, usually done at the product generation state.

In this paper, the absolute radiometric correction method applied is a combination of the cross-calibration method developed by Vogelmann *et al.* (2001), with an atmospheric correction algorithm based on the Dark Object Subtraction (DOS) approach, presented by Song *et al.* (2001). The relative radiometric correction used was developed by Du *et al.* (2002), while a new approach for an objective selection of pseudo-invariants features (PIFs) is presented. Also, we propose a modification and extension of Du *et al.*'s method, to overcome the limitations found when applying it to a combination of Landsat TM and ETM+ images. The method presented in this paper relies on the same hypothesis presented by Du *et al.* (2002) but introduces a new algorithm to improve the selection of PIFs, a key factor in the normalization process.

2.1 Absolute correction

DOS is perhaps the simplest yet one of the most widely used image-based atmospheric correction approach for classification and change-detection studies. Atmospheric corrections of satellite images based on the DOS approach have proved to be as reliable as other atmospheric-correction algorithms (Song *et al.* 2001). This approach assumes the existence of dark objects (zero or low surface reflectance) throughout a scene and a horizontally homogeneous atmosphere. If a dark object could be found in the image, this would correspond to the minimum digital count value (Q_{\min}) in the histogram of the entire scene. From this minimum, it is possible to estimate the path radiance and correct the entire scene for atmospheric scattering effects (Teillet and Fedosejevs 1995, Chavez 1996, Song *et al.* 2001).

When using both Landsat TM and ETM+ images in studies that require radiometric consistency between images, special attention has to be paid to the differences in sensors response. As mentioned before, there are significant differences in the radiometric response between Landsat ETM+ and Landsat TM spectral bands (Teillet *et al.* 2001). To reduce these differences and ensure image inter-comparability, a cross-calibration is needed before performing the absolute radiometric correction. In this way, any radiometric discrepancy between sensors could be reduced, while taking advantage of the superior radiometric calibration of Landsat ETM+. The procedure followed in this paper was presented by Vogelmann *et al.* (2001), where Landsat TM image (L5_Q5), in digital counts, is first converted to Landsat ETM+ image (L5_Q7), in digital counts, as follows:

$$L5_{Q7} = L5_{Q5} \text{slope} + \text{intercept} \quad (1)$$

using the slope and intercept values provided by Vogelmann *et al.* (2001) (table 1).

The *slope* and *intercept* values for cross-calibration were derived using two tandem Landsat images (one TM and one ETM+) with about 30 min difference in data acquisition time. The cross-calibration scheme derived in this way can be applied to any set of images of any place. This is due to the nearly linear response of both sensors. In addition, there are no other tandem scenes available. Once the Landsat TM image is cross-calibrated according to the Landsat ETM+ sensor, all the images are treated as Landsat ETM+, and an absolute radiometric correction can be applied. This method, based on the DOS approach, implies, first, eliminating

Table 1. Slope and intercept values for radiometric conversion of Landsat 5 TM DN values to Landsat 7 ETM+ DN values.

Band	Slope	Intercept
1	0.9398	4.2934
2	1.7731	4.7289
3	1.5348	3.9796
4	1.4239	7.0320
5	0.9828	7.0185
7	1.3017	7.6568

or reducing the effects resulting from the satellite sensor system calibration, transforming the Q values of the original image in values of at-satellite radiances,

$$L_{\text{sat}} = Q_{\text{gain}} + \text{offset}, \quad (2)$$

where L_{sat} is the at-satellite radiance, Q is the image value in digital counts, and gain and offset are those of Landsat ETM+ (table 2), since all the TM images have now been rescaled to be comparable to ETM+ images.

Then, it is necessary to convert the apparent at-satellite radiance to reflectance at the surface of the Earth, which involves the correction of effects caused by illumination geometry and atmospheric conditions. We used the improved DOS method (Chavez 1996), denoted DOS3 by Song *et al.* (2001). This includes the standard DOS approach calculation, except for the atmospheric transmittance along the path from the Sun to the ground surface (T_z). The T_z value was computed as,

$$T_z = \exp\left(-\frac{\tau_r}{\cos \theta_z}\right), \quad (3)$$

where θ_z is the solar zenith angle, and τ_r is the optical thickness of Rayleigh dispersion, calculated according to Kaufman (1989),

$$\tau_r = 0.008569\lambda^{-4}(1 + 0.0113\lambda^{-2} + 0.00013\lambda^{-4}), \quad (4)$$

where λ is the central wavelength of each band, in microns.

Equation (5) shows the final calculation to retrieve images in values of reflectance at the surface of the Earth:

$$\rho = \frac{(L_{\text{sat}} - L_{\text{haze}})\pi d^2}{E_0 \cos \theta_z T_z}, \quad (5)$$

where ρ is the reflectance at the surface of the Earth, L_{haze} is the upwelling atmospheric radiance or path radiance (haze), d is the Earth–Sun distance in

Table 2. Gain and offset values of Landsat 7 ETM+, used for calibration of images.

Band	Gain	Offset
1	0.775	-6.2
2	0.795	-6.4
3	0.619	-5.0
4	0.637	-5.1
5	0.125	-1.0
7	0.043	-0.35

astronomical unit, E_0 is the exoatmospheric solar irradiance, q_z is the Sun zenith angle, and T_z is the atmospheric transmittance along the path from the Sun to the ground surface.

L_{haze} was computed as

$$L_{\text{haze}} = Q_{\text{min}} \text{gain} + \text{offset} - \frac{(0.01 \cos \theta_z T_z E_0)}{\pi d^2}, \quad (6)$$

following Song *et al.* (2001).

In this paper, we will refer to this combination of methods (cross-calibration+atmospheric correction) as Vogelmann–DOS3. The Vogelmann–DOS3 method produces images of reflectance at the surface of the Earth that are radiometrically consistent.

2.2 Relative correction

The relative radiometric correction method normalizes images of the same area and different dates by using landscape elements (pixels) whose reflectance values are nearly constant over time. This procedure assumes that the pixels sampled at Time 2 are linearly related to the pixels, of the same locations, sampled at Time 1, and that the spectral reflectance properties of the sampled pixels have not changed during the time interval (no actual change during this period). The sampled pixels are considered pseudo-invariant features (PIFs) and are the key to the image regression method used in the normalization process. The main characteristic of PIFs is that they are considered objects spatially well defined and spectrally stable though time. The limitation of this kind of approach is that the landscape elements are normally selected by visual inspection, which could result in a subjective radiometric normalization.

In order to perform a relative radiometric correction, Du *et al.* (2002) proposes an objective methodology for the selection of PIFs. The main assumption in this method is that the linear effects impacting the images are much greater than nonlinear effects, so

$$Q = La + b, \quad (7)$$

where Q is the image value in digital counts, L is the surface radiance of the imaged scene, and a and b are linear coefficients that take into account changes in satellite sensor calibration over time, differences in illumination and observation angles, atmospheric effects, etc.

From equation (7), it can be shown that the statistical properties of the PIFs are constants for all the images (see Du *et al.* 2002 for a full description). There is an attribute $A(i)$ of each PIF, independent of the image characteristics,

$$\frac{(Q(i) - \bar{Q})^2}{\frac{1}{n} \sum_{i=1}^{1=n} (Q(i) - \bar{Q})^2} = \frac{(L(i) - \bar{L})^2}{\frac{1}{n} \sum_{i=1}^{1=n} (L(i) - \bar{L})^2} = A(i). \quad (8)$$

It can be seen, from equation (8), that $A(i)$ is a dimensionless factor that is independent of the a and b coefficients. Therefore, $A(i)$ represents a property of the PIFs that is independent of all linear variations affecting the image.

Through the selection of PIFs, the linear effects impacting the images can be established and, therefore, determine the correction coefficients to be applied to

normalize the images. The PIFs selection procedure proposed by Du *et al.* (2002) is an objective process based on PCA calculations between pair of analogous bands of different images, and involves:

1. Applying thresholds values to each band of each image to reject cloudy and water pixels.
2. Using the remaining pixels to compute the PCA (in a multitemporal mode) between each pair of band of each image (i.e. b2_1998 vs b2_1999).
3. Then, the pixels located around the primary major axis will be selected as PIFs, using an arbitrary threshold U perpendicular to the PCA major axis (i.e. $U < \text{minor axis} < U$).

Once the PIFs are selected, the mean and standard deviation of each band in each image are calculated, and the gain and offset to normalize the images are computed as,

$$\text{gain}_{(j)} = \frac{\sigma_{Q_{\text{ref}}}}{\sigma_{Q_j}} \quad (9)$$

$$\text{offset}_{(j)} = \bar{Q}_{\text{ref}} - \frac{\sigma_{Q_{\text{ref}}}}{\sigma_{Q_j}} \cdot \bar{Q}_j, \quad (10)$$

where j represents the image date, Q_{ref} and $\sigma_{Q_{\text{ref}}}$ are the reference mean and standard deviation values, respectively, and Q_j and σ_{Q_j} are the mean and standard deviation of each set of PIFs.

In this way, the entire set of images are normalized to a reference scale common to all the images, by applying new gain and offsets values for each band of each image.

3. Study area and data

3.1 Vegetation and climate

Our analysis was conducted over remotely sensed data from a portion of the subtropical mountain forest of north-western Argentina (figure 1). This area corresponds to the ‘Yungas’ biogeographic province (Cabrera and Willink 1980) that extends along the tropical Andes, reaching its southern limit in northern Argentina. The climate of the region is subtropical monsoonal, with 70–90% of precipitation occurring during the summer followed by a marked dry winter season. Therefore, cloud-free satellite images are mainly available from June to August.

In a simplified scheme, two forest types can be differentiated along the elevation gradient. The lower montane forest extends from 400 to 1700 m of elevation, and is dominated by a relatively diverse (20–30 tree species per hectare) semi-evergreen forest. The upper montane forest (1700–2700 m) is a mosaic of relatively simple forests largely dominated by *Alnus acuminate*, grassland, and shrublands (Grau *et al.* 2003). Above treelines, vegetation is dominated by tall grasslands.

3.2 Imagery processing

Radiometric consistency is hard to maintain between separate images with repeated coverage. Among the various aspects of image preprocessing for land cover change

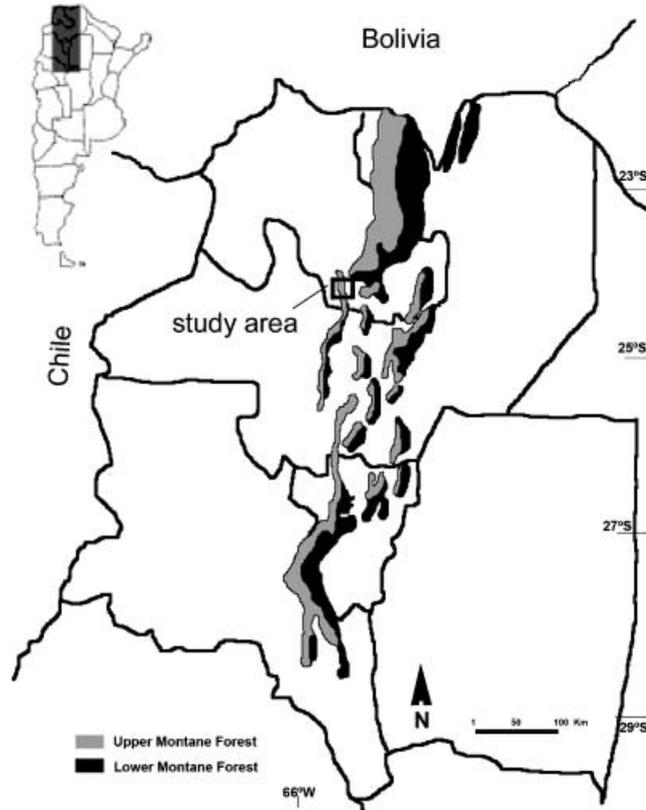


Figure 1. Location of subtropical mountain forests and study area in north-western Argentina.

detection, there are two outstanding requirements: multitemporal image registration and radiometric correction (Coppin and Bauer 1996). For this study, one Landsat TM and two Landsat ETM+ images of the same area and from different years were used (table 3). The three images are of near-anniversary dates to reduce the differences in illumination geometry and vegetation phenology. All images were first geometrically registered to one base image, using nearest-neighbour resampling. The resulting misregistration error (rms) is, in all cases, less than 0.3 pixel (table 3). An area common to all three images was selected for testing the different radiometric correction methods presented in this paper.

4. Application procedure

We first corrected our set of images with both the absolute (Vogelmann–DOS3) and the relative (Du *et al.*) radiometric correction methods presented above. When looking at the results after applying these methodologies, we found a key limitation in the relative method developed by Du *et al.* (2002). When working with a combination of TM and ETM+ images, it is not possible to find a common set of PIFs. It was this limitation that led us to look for a different procedure for PIFs selection.

Table 3. Characteristics of the images used in the study and coregistration results.

Acquisition date	WRS path/ row	Source/ format	Satellite/ sensor	Number of GCPs ^a for registration	Registration RMSE ^b
26 July 1998	231/076	CONAE ^c /LPGS	Landsat 5/TM	34	0.26
06 August 1999	231/076	CONAE ^c /LPGS	Landsat 7/ ETM+	35	0.23
23 July 2000	231/076	CONAE ^c /LPGS	Landsat 7/ ETM+	–	Base image

^aGround control points.

^bRoot mean square error of x,y coordinate.

^cComisión Nacional de Actividades Espaciales, Argentina.

Although Du *et al.*'s method was developed based on a set of images composed exclusively by Landsat TM, the algorithm proposed should perform well, even with a combination of Landsat TM and ETM+ images (Du, personal communication). Even when the theoretical approach and the $A(i)$ calculations do not themselves present problems in a multisensor arrangement, the selection of PIFs through a paired (band-to-band) PCA analysis failed to consider all the spectral variation among the images.

Following Du *et al.*, we compute the PCA between the bands for all the possible combination of images (i.e. b2_1998 vs. b2_1999, b2_1999 vs. b2_2000 and b2_2000 vs. b2_1998), and found three sets of PIFs (one per each image pair) but with no common pixels to all images (figure 2). The results of this analysis for the other bands are similar (not shown). When working with images with differences in the radiometric response (as is the case of TM and ETM+ sensors), this PIFs selection technique based on paired bands PCA calculation fails to find a common set of PIFs to all the images from which to construct the common reference scale. From a theoretical viewpoint, the lack of common PIFs to all of the images involved in any study means that no portion of the study area remains unchanged, so the pixels sampled cannot be considered PIFs. Thus, having no real PIFs common to all the images, it is not possible to build a common reference scale. Then, the gains and offsets computed for the normalization procedure will not be the correct ones. Although the method presented by Du *et al.* (2002) has a solid theoretical basis, the PIFs selected does not represent the entire radiometric variability of all the images when the sensor gains have considerable differences.

To test our hypothesis that the radiometric difference between sensors is responsible for the difficulty of finding common PIFs, we first cross-calibrate the Landsat TM image according to Vogelmann *et al.* (2001) and then re-normalize the images following Du *et al.* (2002). In this paper, we will refer to this methodology as Vogelmann–Du. The results achieved following this procedure show a better classification accuracy than those obtained after applying the original Du *et al.* (2002) method. Although this supports our hypothesis, the radiometric consistency assessed is reduced in band-to-band PIFs radiometric response, compared with the radiometric consistency reached after applying another more sophisticated radiometric correction method (i.e. absolute correction method).

In order to overcome the aforementioned difficulties, we develop a new algorithm, which simplifies the relative calibration process and improves final results. The methodology we propose follows the same theoretical approach presented by Du

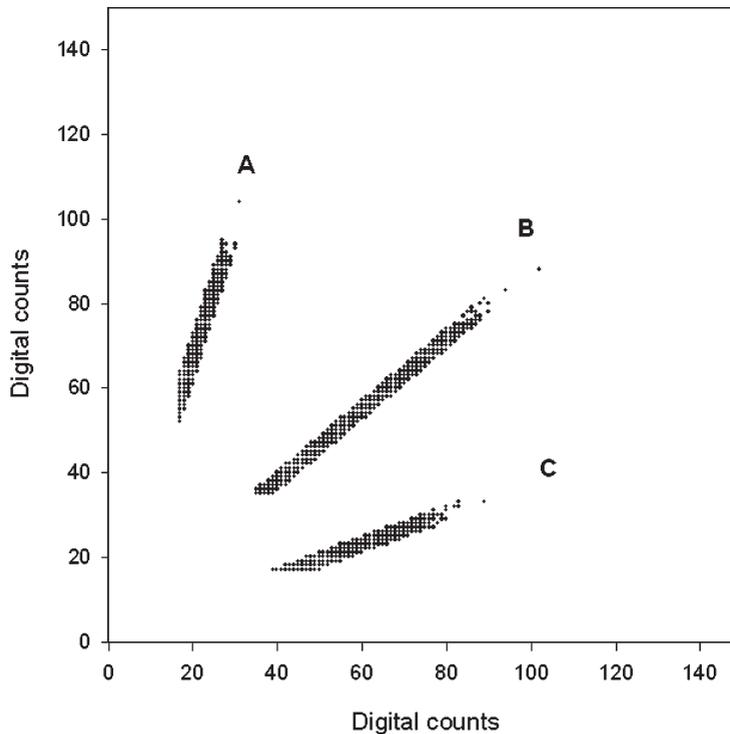


Figure 2. Scatter plots of the PIFs selected from band 2 of the three images, according to Du *et al.*'s (2002) methodology. A, B, and C correspond to the PIFs selected from paired analysis of 1998–1999, 1999–2000, and 2000–1998 images, respectively.

et al. (2002), but developing a simpler yet more effective procedure for selecting common PIFs to all the images, without cross-calibrating the Landsat TM and ETM+ images. We will refer to this new method developed by us as the multi-dimensional PIFs selection (MDPS) procedure.

4.1 MDPS method

The selection of PIFs following the MDPS methodology is based on a three-dimensional principal-component analysis (3D PCA). After 3D PCA calculations, a cylinder is constructed where its axis is defined by the PCA major axis, and its radius is an arbitrary threshold U , generated through an iteration process until there is a sufficient number of pixels included (i.e. 1% of the pixels of the images) (figure 3). All the pixels common to the three images that are contained in the cylinder are selected as PIFs. To decide if a pixel is contained in the cylinder, the minimum distance between that pixel and the 3D PCA major axis must be determined. If that distance is less than the cylinder radius, the pixel is contained in the cylinder and will be selected as a PIF; otherwise the pixel will be rejected.

The 3D PCA major axis could be defined by the vector (x', y', z') , and any pixel P common to the three images could be defined by the vector (x, y, z) , where x =band 1_1998, y =band 1_1999, and z =band 1_2000. The minimum distance d between any

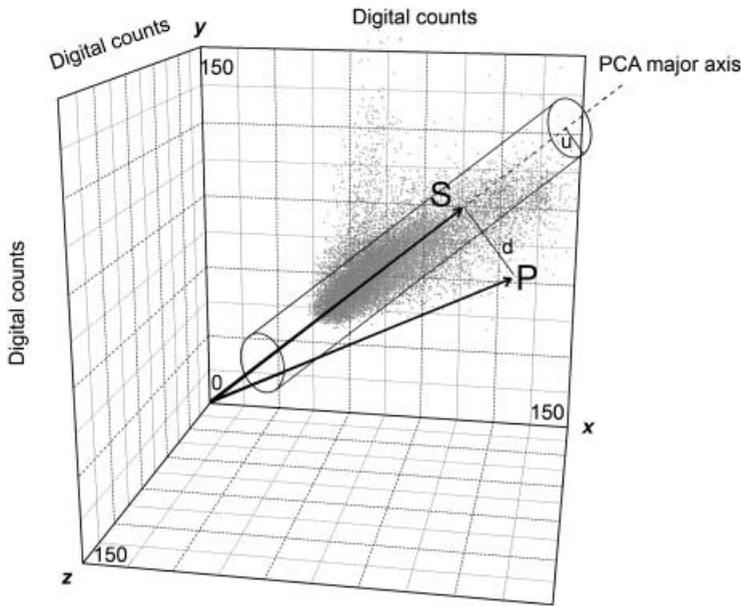


Figure 3. 3D Scatter plot for band 2 of 1998 (X), 1999 (Y) and 2000 (Z) images, in digital counts (Q). P is the pixel analysed. S is the closest point on the 3D PCA axis to the pixel P. d is the minimum distance between P and S. u is the cylinder’s radius (threshold).

pixel P and the 3D PCA major axis is defined by:

$$d = \left[(S_x - x)^2 + (S_y - y)^2 + (S_z - z)^2 \right]^{\frac{1}{2}}, \tag{11}$$

where S is the closest point on the 3D PCA axis to the pixel P. S has the same 3D PCA major axis direction and a magnitude computed as the dot product of P and 3D PCA major axis,

$$S = (P \cdot PCA) PCA = (xx' + yy' + zz')(x', y', z') = R(x', y', z'). \tag{12}$$

where $R = (x \cdot x' + y \cdot y' + z \cdot z')$.

The distance between S and P is calculated as shown in equation (13),

$$d = \left[(Rx' - x)^2 + (Ry' - y)^2 + (Rz' - z)^2 \right]^{\frac{1}{2}} < U, \tag{13}$$

where U is the radius of the cylinder. Equation (13) represents the condition for the pixel P to be included in the cylinder.

This calculation could be extended to be used with any number of images, according to the general form,

$$d = \left[(Rx' - x)^2 + (Ry' - y)^2 + (RN' - N)^2 \right]^{\frac{1}{2}} < U, \tag{14}$$

where N is the total number of images and R is computed as

$$R = (xx' + yy' + + NN') \tag{15}$$

After applying this procedure, all the pixels contained in this cylinder are selected as PIFs common to the three images. Then, they are used to compute the new gain and offset for the set of images, according to equations (9) and (10).

5. Evaluation procedure

5.1 Post-correction analysis

To assess the quality of the radiometric corrections, after applying all the radiometric correction methods (Vogelmann–DOS3, Du *et al.*, Vogelmann–Du and MDPS), we developed a Quadratic Difference index (QD). Through the QD index, we analyse the slope of the PCA major axis (SL) between each pair of analogous bands of each image combination (e.g. b2_1998 vs. b2_1999, b2_1999 vs. b2_2000 and b2_2000 vs. b2_1998), after the radiometric correction process. If a pair of images do not show land cover changes and are not affected by radiometric linear effects (case A), the slope of the major axis of a multi-temporal band-to-band PCA would be equal to 1, and the scatter plot pattern for each paired bands λ of each image combination n would be a straight line. When there is no land cover change between the two images, but there are differences in its radiometric responses due to linear effects (i.e. different gain) (case B), the SL will be greater or less than 1. The purpose of relative radiometric corrections is to recover case A data from case B data.

To measure the differences between the actual slope of the PCA major axis and the ideal case of a PCA slope equal to unity, we compute the QD Index, for all the possible images combinations as:

$$QD_{\lambda} = \sum_{n=1}^{n=N} (1 - SL_{n,\lambda})^2, \quad (16)$$

where $SL_{n,\lambda}$ is the slope of PCA major axis for each paired bands λ of each image combination n . N is the total number of image combinations.

For the case of our set of images, the computation of the QD index for band 2 (QD_2) is as follows,

$$QD_2 = (1 - SL_{(A,2)})^2 + (1 - SL_{(B,2)})^2 + (1 - SL_{(C,2)})^2, \quad (17)$$

where A , B , and C are the 1998–1999, 1999–2000, and 2000–1998 images combination, respectively.

In this way, the closer the QD index to 0, the better the overall radiometric correction process.

5.2 Classification

To further evaluate the impact of the different radiometric correction methods, we classified both corrected and uncorrected images, following a hybrid classification technique (unsupervised + supervised). First, an unsupervised classification (ISODATA) was applied to the newest image (Landsat ETM+, year 2000) to retrieve 50 different spectral classes. After labelling classes based on a field survey, we grouped them into five informational classes to be used as training data for a supervised classification. The classification was designed to evaluate the radiometric correction methods proposed in this paper. Defining a few, broad and spectrally differentiated classes increases the probability of finding all the classes labelled in the

last year image in all the images composing the time series. To check if the five classes defined have an adequate radiometric separability among classes (Richards and Jia 1999), we compute the Transformed Divergence (DT) index. A TD value of 1.88 was considered the threshold above which the classes are different enough to be used in the supervised classification (table 4).

In order to train the supervised classifier, we used as training data the spectral signature of each of the five classes defined from the newest image (year 2000). Beside the data needed to train the supervised classification algorithm, site data also are necessary to validate the classification results (Muchoney and Strahler 2001). A stratified random (proportionate) sampling method was used to select 70% of the pixels of each class for training the classifier and the other 30% for testing results. These spectral signatures statistics were used as input data in a supervised maximum likelihood classification algorithm (probability threshold 0.05), to classify all the images. This generalization procedure allows us to obtain thematic maps of past years (years 1998 and 1999) from data acquired and field-checked at the present time (year 2000)

The value of a thematic map obtained as a product of the classification process is clearly a function of the accuracy of the classification (Foody 2002). The most widely promoted and used accuracy-assessment technique may be that derived from a confusion or error matrix. The confusion matrix is used to provide a site-specific assessment of the correspondence between the image classification and ground conditions. After classifying the images, we assessed the accuracy of the thematic maps by computing the kappa coefficient (k) and the overall accuracy.

5.3 Land cover change-detection analysis

Change-detection results depend on the quality of the thematic maps used. Land cover change assessment was computed comparing time series of thematic maps produced from uncorrected and corrected images. The ‘uncorrected’ set of images corresponds to the original images without any radiometric correction. We identify as ‘corrected’ the set of images normalized with the MDPS method, which were used for further analysis. To illustrate the effects of radiometric correction on change detection assessment, we compare the overall land cover change for the period 1998–2000, between uncorrected images and corrected images using MDPS method. The comparison of the overall change detected is used only as an estimation of the error extent that could be achieve when working with uncorrected images in multi-date multi-sensor studies.

Table 4. Transformed divergence class separability index^a.

Class	Shadow	Forest	Grassland	Bare soil	Rock
Shadow	2.00	1.88	1.99	1.99	2.00
Forest		2.00	1.89	1.92	1.99
Grassland			2.00	1.90	1.99
Bare soil				2.00	1.94
Rock					2.00

^aTransformed Divergence index range from 0 to 2.

6. Results

Table 5 shows the results of each paired band-to-band analysis for all images combinations. This analysis was applied to images not radiometrically corrected (original images in digital counts) and to images in which four different radiometric correction procedures were used (Vogelmann–DOS3, Du, Vogelmann–Du and MDPS). The SL and QD index values are presented as indicators of the level of radiometric correction achieved. The lower the QD index, the better the radiometric correction level.

MDPS procedure (presented in this paper) (table 5(e)) yields SL values very close to 1 and QD values that are almost 0 for all the bands. This means that the

Table 5. Slopes of the PCA 1ST component and QD index of uncorrected and corrected images.

Bands	SL ^a	SL ^a	SL ^a	QD index ^b
	L5 1998–L7 1999	L7 1999–L7 2000	L7 2000–L5 1998	
<i>(a) Uncorrected images</i>				
1	1.5124	0.8158	0.7936	0.3391
2	3.1053	0.8414	0.3707	4.8535
3	2.6529	0.9678	0.3938	3.1006
4	1.4004	0.9801	0.6822	0.2617
5	1.2874	1.0189	0.7243	0.1590
7	1.9380	0.9690	0.5015	1.1293
<i>(b) Images in reflectance at surface of the Earth (Vogelman–DOS3 correction method)</i>				
1	1.2348	1.0004	0.8041	0.0935
2	1.1569	1.0072	0.8565	0.0453
3	1.3032	1.0066	0.7609	0.1491
4	0.8235	1.0695	1.1396	0.0555
5	1.0220	1.0648	0.9197	0.0111
7	1.2614	1.0041	0.7897	0.1126
<i>(c) Images in digital counts Q (Du et al.'s correction method)</i>				
1	1.3083	0.8727	0.8683	0.1286
2	1.0407	0.8785	1.0935	0.0252
3	0.8048	1.1294	1.1027	0.0654
4	0.8089	1.2302	1.0048	0.0895
5	0.6540	1.1727	1.2984	0.2386
7	0.7436	0.9911	1.3562	0.1927
<i>(d) Images in digital counts Q (Vogelman–Du correction method)</i>				
1	1.0925	0.9526	0.9584	0.0125
2	1.3652	0.9660	0.7550	0.1946
3	0.9306	0.9518	1.1292	0.0238
4	0.8876	1.2874	0.8712	0.1118
5	0.8252	1.0509	1.1498	0.0556
7	0.7997	0.9259	1.3487	0.1672
<i>(e) Images in digital counts Q (MDPS correction method)</i>				
1	0.9723	1.0203	1.0052	0.0012
2	0.9960	1.0075	0.9964	0.0001
3	0.9816	1.0052	1.0135	0.0005
4	0.9562	1.0184	1.0286	0.0031
5	0.9937	0.9942	1.0118	0.0002
7	0.9659	1.0112	1.0183	0.0016

^aSlope of the PCA major axis.

^bQuadratic Difference Index.

radiometric differences among PIFs (after correction) are minimal, so the images are correctly normalized.

The SL values for the Vogelmann–DOS3 absolute radiometric correction method (table 5(b)) differ from 1, except for the L7_1999–L7_2000 image pair. The QD index values show values close to 0 but have some deviation from ideal values in bands 3 and 7. Looking at SL values, a better radiometric correction level is reached when images of the same sensor (in this case ETM+) are normalized.

The results obtained after applying Du *et al.*'s method, with and without cross-calibration (table 5(d) and table 5(c)), present several inter-band variations in SL and QD values, although they show an improvement when cross-calibration is applied as a first step. Some SL values differ significantly from 1, reaching values greater than 1.3, while the QD index presents values greater than 0.1 for bands 1 and 7 and greater than 0.2 for band 5. These values do not fit the theoretical assumption of SL values close to 1 (for normalized images) and can lead to poor inter-image comparability.

As expected, uncorrected images (table 5(a)) show SL values very different to 1 and QD index values much larger than 0 for most of the bands.

Figure 4 shows all the classification and post-classification results of corrected and uncorrected images. In the same way as in the post-correction analysis, as discussed previously in table 5, post-classification results show that the MDPS method and the Vogelmann–DOS3 method reach the best overall accuracy (OA) and kappa coefficient (k) values, with minimal differences between them.

The results of MDPS method (figures 4(M), (N), and (O)) show similar values of both, OA and the k coefficient, of the order of 80% and 0.7, respectively, for L7_1999 and L5_1998. Also, these values do not differ by more than 15% from the L7_2000 image taking as the base image for training samples selection.

The Vogelmann–DOS3 absolute correction method (figures 4(d), (e), and (f)), shows results very close to those obtained through the MDPS approach. The small differences between these two methods arise in L7_1999 image classification, where Vogelmann–DOS3 presents a 5–10% improvement with respect to MDPS in OA and k coefficient.

In the case of the Vogelmann–Du method (figures 4(j), (k), and (l)), the results present differences in OA of the order of 15% between L7_2000 and L7_1999, and of 30% between L7_2000 and L5_1998, even considering that a cross-calibration was applied first.

When using Du *et al.*'s method (figures 4(g), (h), and (i)), the OA and k values are very low for L7_1999, showing a significant difference in OA with respect to L7_2000 image. This could be due to the fact that the PIFs selected do not take into account the radiometric variations among all images, so it is not possible to build a common reference scale to use for normalization.

For completeness, the classification of the uncorrected images included in figures 4(a), (b), and (c) shows, once more, the importance of performing a radiometric correction to ensure inter-comparability of the images.

Table 6 shows the results of change detection analysis for uncorrected images and corrected images using MDPS method. Table 6(a) shows that change-detection estimations of uncorrected images present unrealistic change rates. This happens when training samples obtained from an uncorrected ETM+ image are used as a training set to classify an uncorrected TM image.

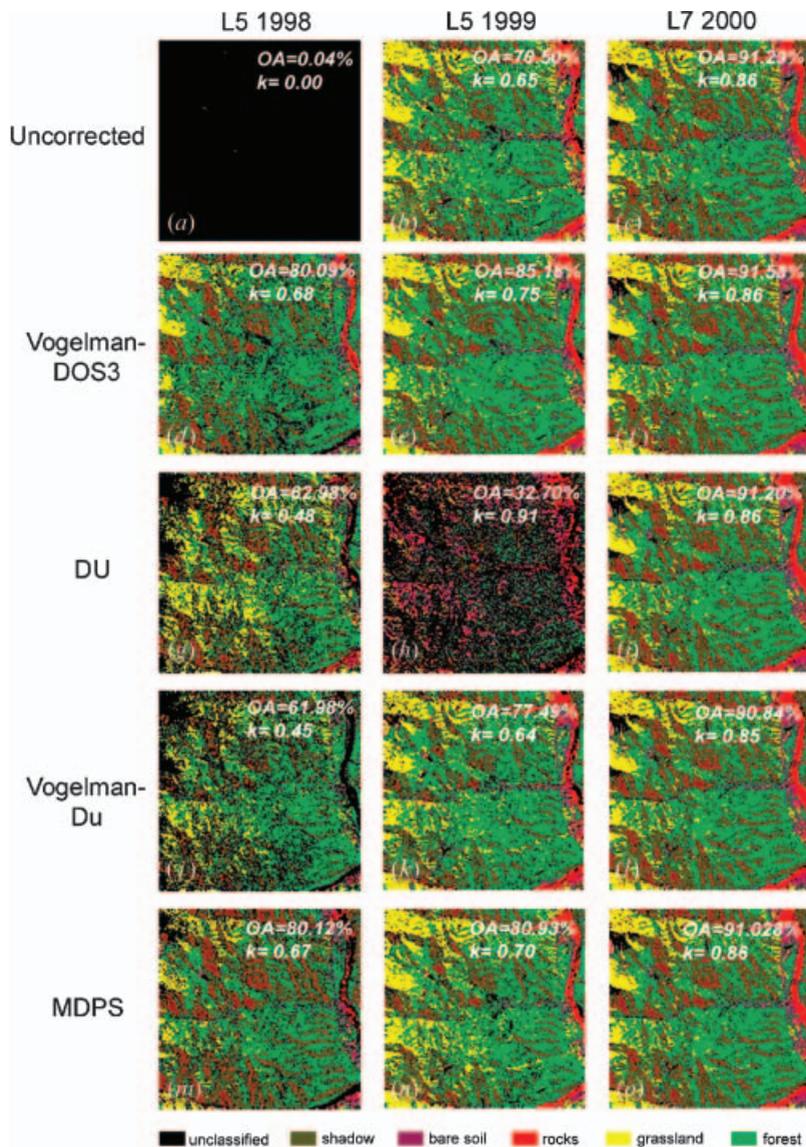


Figure 4. Thematic maps obtained through hybrid classification (unsupervised + supervised) of uncorrected and corrected Landsat TM (L51998) and Landsat ETM+ (L71999 and L72000) images. The base image is L72000. Radiometric corrections include all the methodologies addressed in this paper (Vogelman–DOS3, Du *et al.*, Vogelman–Du, and MDPS methods). OA=overall accuracy (%); k =kappa coefficient.

Change-detection estimations using images corrected with MDPS method (table 6(b)), show the expected change rate between 1998 and 2000. As an example, the overall change in the forest class is about 5% for the corrected images and about 55% for the uncorrected images (unrealistic value for this area). These results highlight the importance of radiometric correction in change-detection analysis, especially when different sensors are used.

Table 6. Area covered (km²) and change detection analysis for the period 1998–2000.

<i>(a) Images in digital counts Q (uncorrected)</i>							
Class year	1998	1999	2000	1998–1999	1999–2000	1998–2000	Overall change (%) 1998–2000 ^a
Shadow	0.04	21.85	33.18	21.81	11.33	33.14	24.35
Forest	0	82.04	75.89	82.04	–6.15	75.89	55.76
Grassland	0	17.56	13.15	17.56	–4.41	13.15	9.66
Bare soil	0	9.35	9.85	9.35	0.5	9.85	7.24
Rock	0	4.01	4.02	4.01	0.01	4.02	2.95

<i>(b) Images in digital counts Q (corrected with MDPS method)</i>							
Class year	1998	1999	2000	1998–1999	1999–2000	1998–2000	Overall change (%) 1998–2000 ^a
Shadow	38.56	22.77	33.21	–15.79	10.44	–5.35	–3.938
Forest	69.24	80.96	76.13	11.72	–4.83	6.89	5.071
Grassland	7.9	20.28	13.15	12.38	–7.13	5.25	3.864
Bare soil	6.61	5.92	9.37	–0.69	3.45	2.76	2.031
Rock	1.51	3.39	4.01	1.88	0.62	2.5	1.840

^aOverall change 1998–2000=((area2000–area1998)/total area) × 100.

7. Conclusions

Radiometric corrections of satellite images, either absolute or relative, are necessary in multi-date multi-sensor change-detection studies, especially if generalization of remotely sensed data is needed. Unless some processing is performed to ensure the radiometric comparability between images, any additional procedure, such as classification and/or change detection, will not be reliable. Furthermore, land cover change estimates should be carefully evaluated depending on the level of radiometric consistency among the data used.

It is clear from this paper that classification and change-detection estimates from radiometrically uncorrected images yield incorrect results. Four different radiometric correction methods with different levels of performance and complexity were tested, one absolute (Vogelmann–DOS3) and three relative (Du, Vogelmann–Du and MDPS). The results obtained from the absolute radiometric correction method show a good level of radiometric comparability among images but require data about sensor characteristics, illumination and observation geometry, estimations of atmospheric components (i.e. path radiance and molecular absorption), and the relationship between gain and offsets of the sensors involved. If these data are available, this is a recommended method because it leads to consistent results and has an intermediate level of complexity (it requires a certain level of knowledge on atmospheric corrections).

From the three relative radiometric correction methods tested, only one (MDPS) provides a good level of radiometric image comparability, similar to the quality reached with the absolute correction method. The advantage of this method is the fact that it does not require other data than the images itself. One of the disadvantages is related to the selection of the Pseudo Invariant Features (PIFs), which requires the application of a set of procedures previous to normalization of images. Another point to consider is that as the total number of images increases, this also increases the difficulty in finding common PIFs. Therefore, the selection of

one method (absolute correction) or the other (relative correction) involves a performance–complexity trade-off.

The aim of this paper was to gain a better understanding of the radiometric corrections effects in land cover change analysis, in order to improve data generalization across time, space, and sensors. In regions where past field data are not available, the possibility of using data from a present-date image to interpret past-date images is a key factor to assess ecosystems dynamics. With the added benefit of sufficient temporal coverage, remote sensing can now be used to make predictions of the Earth's vegetation dynamics with respect to future climate scenarios based upon an analysis of these past satellite observations.

Acknowledgements

This paper was funded by grants from CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), Argentina and MAE-AECI (Ministerio de Asuntos Exteriores-Agencia Española de Cooperación Internacional), Spain. Satellite images were provided by CONAE (Comisión Nacional de Actividades Espaciales), Argentina as part of a L. Paolini's A.O. project. The work presented in this paper was carried out while L. Paolini was a visiting researcher at the Global Change Unit (Universitat de Valencia, Spain) and at the Instituto de Astronomía y Física del Espacio (CONICET, Argentina), whose support is gratefully acknowledged. The authors would like to thank Dr. Héctor R. Grau for his helpful suggestions and comments.

References

- ABUELGASIM, A.A., ROSS, W.D., GOPAL, S. and WOODCOCK, C.E., 1999, Change detection using adaptive fuzzy neural networks: environmental damage assessment after the Gulf War. *Remote Sensing of Environment*, **70**, pp. 208–223.
- BLODGETT, T.A., 1998, Erosion rate on the NE escarpment of the Eastern Cordillera, Bolivia derived from aerial photographs and thematic mapper images. PhD thesis. Cornell University.
- CABRERA, A.L. and WILLINK, A., 1980, *Biogeografía de América Latina* (Washington, DC: Organización de Estados Americanos).
- CHAVEZ, P.S., 1996, Image-based atmospheric correction—revisited and improved. *Photogrammetric Engineering and Remote Sensing*, **62**, pp. 1025–1036.
- COLLINS, J.B. and WOODCOCK, C.E., 1996, An assessment of several linear change detection techniques for mapping forest mortality using multitemporal Landsat TM data. *Remote Sensing of Environment*, **56**, pp. 66–77.
- COPPIN, P.R. and BAUER, M.E., 1996, Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sensing of Environment*, **13**, pp. 207–234.
- COPPIN, P., JONCKHEERE, I., NACKAERTS, K., MUYS, B. and LAMBIN, E., 2004, Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, **25**, pp. 1565–1696.
- DU, Y., TEILLET, 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*, **82**, pp. 123–134.
- FOODY, G.M., 2002, Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, **80**, pp. 185–201.
- FOODY, G.M., 2003, Remote sensing of tropical forest environments: towards the monitoring of environmental resources for sustainable development. *International Journal of Remote Sensing*, **24**, pp. 4035–4046.

- FULLER, R.M., SMITH, G.M. and DEVEREUX, B.J., 2003, The characterization and measurement of land cover through remote sensing: problems in operational applications? *International Journal of Applied Earth Observation and Geoinformation*, **4**, pp. 243–253.
- GRAU, H.R., 2001, Regional-scale spatial patterns of fire in relation to rainfall gradients in sub-tropical mountains, NW Argentina. *Global Ecology and Biogeography*, **10**, pp. 133–146.
- GRAU, H.R., EASDALE, T.A. and PAOLINI, L., 2003, Subtropical dendroecology: dating disturbances and forest dynamics in northwestern Argentina mountain ecosystems. *Forest Ecology and Management*, **177**, pp. 131–143.
- HELMER, E.H., BROWN, S. and COHEN, W.B., 2000, Mapping mountain tropical forest successional stage and land use with multi-date Landsat imagery. *International Journal of Remote Sensing*, **21**, pp. 2163–2183.
- KAUFMAN, Y.J., 1989, The atmospheric effect on remote sensing and its correction. In *Theory and Application of Optical Remote Sensing*, G. Asrar (Ed.), pp. 336–424 (New York: Wiley).
- KERR, J.T. and OSTROVSKY, M., 2003, From space to species: ecological applications for remote sensing. *Trends in Ecology and Evolution*, **18**, pp. 299–305.
- LU, D., MAUSEL, P., BRONDIZIO, E. and MORAN, E., 2002, Assessment of atmospheric correction methods for Landsat TM data applicable to Amazon basin LBA research. *International Journal of Remote Sensing*, **23**, pp. 2651–2671.
- LUNETTA, R.S., EDIRIWICKREMA, J., JOHNSON, D.M., LYON, J.G. and MCKERROW, A., 2002, Impacts of vegetation dynamics on the identification of land-cover change in a biologically complex community in North Carolina, USA. *Remote Sensing of Environment*, **82**, pp. 258–270.
- MILLER, J.D. and YOOL, S.R., 2002, Mapping forest post-fire canopy consumption in several overstory types using multi-temporal Landsat TM and ETM data. *Remote Sensing of Environment*, **82**, pp. 481–496.
- MUCHONEY, D.M. and HAACK, B.N., 1994, Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing*, **60**, pp. 1243–1251.
- MUCHONEY, D.M. and STRAHLER, A.H., 2001, Pixel- and site-based calibration and validation methods for evaluating supervised classification of remotely sensed data. *Remote Sensing of Environment*, **81**, pp. 290–299.
- MUKAI, Y. and HASEGAWA, I., 2000, Extraction of damaged areas of windfall trees by typhoons using Landsat TM data. *International Journal of Remote Sensing*, **21**, pp. 647–654.
- PAOLINI, L., SOBRINO, J.A. and JIMENEZ, J.C., 2002, Detección de deslizamientos de ladera mediante imágenes Landsat TM: el impacto de estos disturbios sobre los bosques subtropicales del Noroeste de Argentina. *Revista de Teledetección*, **18**, pp. 21–27.
- PAX-LENNEY, M., WOODCOCK, C.E., MACOMBER, S.A., GOPAL, S. and SONG, C., 2001, Forest mapping with a generalized classifier and Landsat TM data. *Remote Sensing of Environment*, **77**, pp. 241–250.
- RESS, W.G., WILLIAMS, M. and VITEBSKY, P., 2003, Mapping land cover change in a reindeer herding area of the Russian Arctic using Landsat TM and ETM+ imagery and indigenous knowledge. *Remote Sensing of Environment*, **85**, pp. 441–452.
- RICHARDS, J.A. and JIA, X., 1999, *Remote Sensing Digital Image Analysis—An Introduction*, 3rd ed. (Berlin: Springer).
- ROGAN, J., FRANKLIN, J. and ROBERTS, D.A., 2002, A comparison of methods for monitoring multitemporal change using Thematic Mapper Imagery. *Remote Sensing of Environment*, **80**, pp. 143–156.
- ROGAN, J. and YOOL, S.R., 2001, Mapping fire-induced vegetation depletion in the Peloncillo Mountains, Arizona and New Mexico. *International Journal of Remote Sensing*, **22**, pp. 3101–3121.

- SCANLON, T.M., ALBERTSON, J.D., CAYLOR, K.K. and WILLIAMS, C.A., 2002, Determining land surface fractional cover from NDVI and rainfall time series for a savanna ecosystem. *Remote Sensing of Environment*, **82**, pp. 376–388.
- SINGH, A., 1989, Digital change detection techniques using remotely sensed data. *International Journal of Remote Sensing*, **10**, pp. 989–1003.
- SKOLE, D.L., 1994, Data on global land-cover change: acquisition, assessment and analysis. In *Changes in Land Use and Land Cover: A Global Perspective*, W.B. Meyer and B.L. Turner II (Eds), pp. 437–471 (Cambridge: Cambridge University Press).
- 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*, **75**, pp. 230–244.
- TEILLET, P.M., BARKER, L.L., MARKHAM, B.L., IRISH, R.R., FEDOSEJEVS, G. and STOREY, J.C., 2001, Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets. *Remote Sensing of Environment*, **78**, pp. 39–54.
- TEILLET, P.M. and FEDOSEJEVS, G., 1995, On the dark target approach to atmospheric correction of remotely sensed data. *Canadian Journal of Remote Sensing*, **21**, pp. 374–387.
- THOME, K., MARKHAM, B., BARKER, J., SLATER, P. and BIGGAR, S., 1997, Radiometric calibration of Landsat. *Photogrammetric Engineering and Remote Sensing*, **63**, pp. 853–858.
- TOKOLA, T., LÖFMAN, S. and ERKKILÄ, A., 1999, Relative calibration of multitemporal Landsat data for forest cover change detection. *Remote Sensing of Environment*, **68**, pp. 1–11.
- TURNER, W., SPECTOR, S., GARDINER, N., FLADELAND, M., STERLING, E. and STEININGER, M., 2003, Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution*, **18**, pp. 306–314.
- VITOUSEK, P.M., 1994, Beyond global warming: ecology and global change. *Ecology*, **75**, pp. 1861–1876.
- VOGELMANN, J.E., HELDER, D., MORFITT, R., CHOATE, M.J., MERCHANT, J.W. and BULLEY, H., 2001, Effects of Landsat 5 Thematic Mapper and Landsat 7 Enhanced Thematic Mapper Plus radiometric and geometric calibrations and corrections on landscape characterization. *Remote Sensing of Environment*, **78**, pp. 55–70.
- WATSON, N. and WILCOCK, D., 2001, Preclassification as an aid to the improvement of the thematic and spatial accuracy in land cover maps derived from satellite imagery. *Remote Sensing of Environment*, **75**, pp. 267–278.
- WOODCOCK, C.E., MACOMBER, S.A., PAX-LENNEY, M. and COHEN, W.B., 2001, Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment*, **78**, pp. 194–203.