Non-Parametric Smoothing for Relative Radiometric Correction on Remotely Sensed Data

MARIA LUIZA F. VELLOSO[1], FLÁVIO J. DE SOUZA[2]

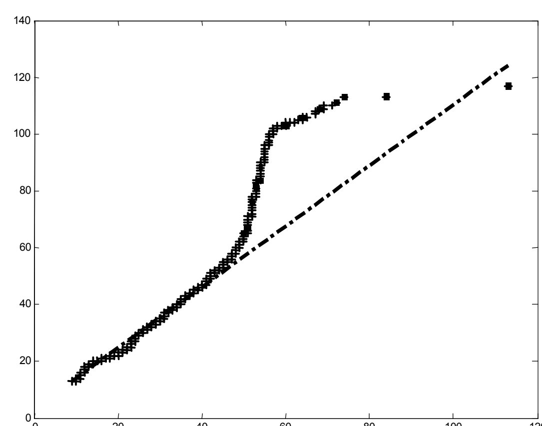
UERJ – Universidade do Estado do Rio de Janeiro – FEN/[1]DETEL, [2]DESC, Rua São Francisco Xavier, 524, 20550-013 Rio de Janeiro, RJ, Brasil
[1]mlfv@centroin.com.br, [2]fjsouza@eng.uerj.br

Abstract. Digital change detection methods have been broadly divided into either pre-classification spectral change detection or post-classification change detection. Since all spectral change detection methods are based on pixel-wise operations, or scene-wise plus pixel-wise operations, accuracy in image registration and scene-to-scene radiometric normalization is more critical for these methods than for other methods. A wide range of algorithms has been developed to adjust linear models. This paper proposes an automated radiometric normalization process that automatically extracts the training dataset, and uses a non-parametric smoother to adjust a non-linear mapping in order to minimize the effects of the influences of radiometric differences on image interpretation and classification. In order to investigate how the proposed normalization improves the performance classification, and assess the effectiveness of this technique, we carried out classification experiments on three image sets, and compare their results.

1. Introduction

The largest number of change detection techniques is found in the spectral change detection category [7]. These techniques utilize data from two images to generate new single-band, or multi-band unclassified image representing spectral change. The spectral change data must be further processed by other analytic methods such as a classifier to produce a labeled land cover change product. It has been shown that transformations designed to normalize or bypass radiometric differences between the raw images usually improves the results of change detection. However, removing or normalizing noises arising from atmospheric effect, changing view and illumination geometry, and instrument errors, are challenges of satellite-based land cover characterization. In an ideal remote sensing system all images in a long time series could be corrected for atmospheric and illumination effects and calibrated to surface reflectance factors for direct comparison. Numerous studies have shown this to be an elusive, if not unattainable, goal and image-based relative calibration procedures generally perform better. Relative radiometric correction procedures can be applied to scale all images to a single reference, and adjust the radiometric properties of subject images to match the reference. This kind of normalization does not require ancillary datasets on. For instance, atmospheric temperature, relative humidity, and/or aerosol backscatter that are normally very demanding when it comes to the logistical and personnel time required. Most relative calibration procedures assume that the reflectance-stable pixel values from the late image are a linear function of the early, including effects from sensor calibration and the atmosphere. A wide range of algorithms has been developed to adjust linear models. Many situations, however, present non-linear behavior. Figure 1 shows the quantile-quantile plot (qq-plot) of bands 5 of multispectral images, from different dates, acquired from the Thematic Mapper sensor of the Landsat 7 satellite taken from a region in Brazil. The purpose of the qq-plot is to determine if the samples in two datasets come from the same distribution type. If the samples come from the same distribution (same shape) even if one distribution is shifted and re-scaled from the other

Figure 1. QQ-plot of band 5 of two multispectral images



acquired, in different dates, from the Thematic Mapper sensor of the Landsat 7 satellite.

(different location and scale parameters) the plot will be linear. The plot in Figure 1 shows strong non-linear behavior in band 5. A non-linear approximation for regression functions would be more appropriate. As soon as we leave the relatively straightforward world of linearity, we are faced with an infinitude of possible choices of models.

The task of approximating the regression function can be done essentially in two ways. First, the quite often-used parametric approach is to assume that the approximation regression curve has some pre-specified functional form. Second, as an alternative one could try to estimate the regression curve non-parametrically without reference to a specific form. The non-parametric smoothing approach offers a flexible tool in analyzing unknown regression relationships.

This paper proposes an automated radiometric normalization process to adjust a non-linear mapping to minimize the effects of the influences of radiometric differences on image interpretation and classification. This process computes tables, for each band of image, generated with a non-parametric regression technique.

Most relative image-to-image calibration models require that the percentage of total changed pixels in the image is small relative to the entire image. One outstanding issue is how to identify the subset of non-changing pixels automatically. We propose a correlation-based method for automated extraction of set of pixels used to adjust the model, the training set. "Chips", or extracts of images process the images band to band in order to generate a correlation map for each band. These maps are used to guide the composition of the training set. With this selected set of pixels we adjust the non-parametric regression to perform the relative correction.

In order to investigate how the proposed normalization improves the performance classification, and assess the effectiveness of this technique, we carried out classification experiments on three image sets, and compare their results. We used multitemporal data sets composed of multispectral images acquired from the Thematic Mapper sensor of the Landsat 7 satellite taken from a region in Brazil under severe degradation process. These three sets have the same reference image, with the original radiometric characteristics, but the second image was radiometrically corrected by different methods. In the first set the second image was corrected by simple linear regression, the most used method. In the second set robust regression was used. Robust regression eliminates or attenuates outliers and is used in relative radiometric correction because it interprets the changing pixels as outliers. Finally, the second image in the third set was corrected by the proposed technique.

This paper is organized as follows. The next section presents the procedure used to select the training set of the used techniques, a subset of non-changing pixels. Section 3 reviews the techniques used for comparison purposes: linear regression and robust linear regression. Sub-section 3.2 describes the proposed technique radiometric normalization. Section 4 describes an evaluation experiment and discusses the results.

2. Automated Extraction of Training Set

Relative radiometric correction is a correction method that applies one image as a reference, and adjusts the radiometric properties of subject images to match the reference [3], [9], [11]. Rectified images appear to have been acquired with the reference image sensor under atmospheric and illumination conditions equal to those in the reference scene [3]. If atmospheric and calibration parameters are available for the reference date then atmospheric correction procedures based on radiance transfer models can be applied to the images [3].

One outstanding issue is how to identify the subset of non-changing pixels automatically. Hall's concept of radiometric rectification is based on stable reflectance control sets derived from the extremes of the image scattergram. Other methods have used stable reflectance targets, or the so-called pseudo-invariant features [10]. Pixel values of these stable targets are used to find linear coefficients to scale each image to the reference. Such stable targets, however, may not be available in some areas, or may be obscured by clouds or missing data in some dates. Some approaches interpret changed pixels as outliers, and adopt a strategy to eliminate, or attenuate them. This approach, however, can be used only if the percentage of total pixels, whose digital numbers (DNs) has changed in the image, is small relatively to the entire image.

To identify the subset of non-changing pixels automatically we used a correlation procedure, as the procedures used in location of ground points in geometric correction. "Chips" or extracts of image of size 21x21 pixels were used in a procedure that processes the images from different dates band to band. This procedure results in a correlation map.

A "chip" of the reference image is placed over the chip centered on the same coordinates in the other image. Then, bi-dimensional correlation between the "chips" pixels is calculated and recorded in a bi-dimensional map, with the same dimensions of the images, at the same position of the centers of the "chips" of two images. This operation is repeated for all coordinates. It was assigned a correlation coefficient for each pixel, for all bands, obtained on the “chips” centered in these pixels.

Figure 2 shows a contour plot of the correlation map of band 3 of two multispectral images acquired, in different dates by Thematic Mapper sensor of the Landsat 7 satellite taken from a region in Brazil, and the reference image of this image set. For better visualization, this plot was divided in three regions: coefficients equal or lower than zero in black, coefficients between zero and 0.5 in gray, and coefficients greater than 0.5 in white.

This procedure was repeated for all bands. After this, we apply a threshold criterion, for each band, based on overall correlation of the band, in order to select pixels that would belong to the training set used to adjust radiometric normalization. A pixel was assumed to belong to the training set if it had correlation coefficients in all bands greater than the respective thresholds.

3. Parametric and Non-parametric Regression

A regression curve describes a general relationship between an explanatory variable X and a response variable Y. If n data points have collected, the regression relationship can be modeled as

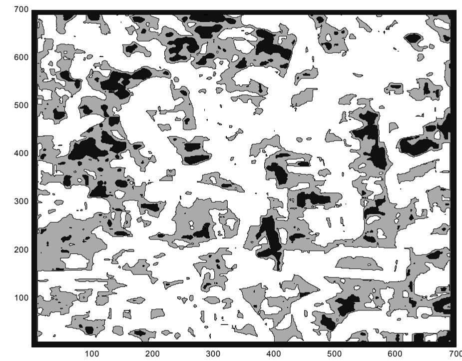
$$Y_i = m\left(X_i\right) + \varepsilon_i, \quad i = 1,..,n \tag{1}$$

with the unknown regression function *m* and observation errors ε_i. A look at a scatter plot of X_i versus Y_i does not always suffice to establish an interpretable regression relationship. An example is given in Figure 3, a scatter plot where X_i and Y_i are DN’s of band 5 of the same image set used in Figure 1.

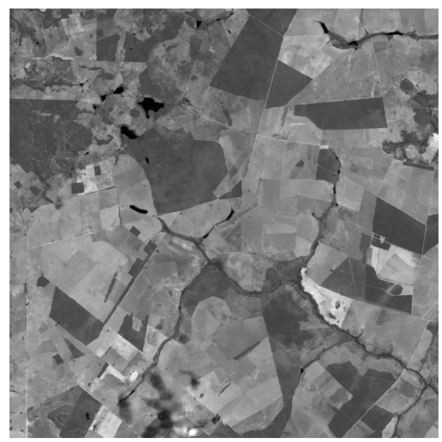
The aim of a regression analysis is to produce a reasonable approximation to the unknown response function *m*. By reducing the observational errors it allows interpretation to concentrate on important details of the mean dependence of Y on X. This curve approximation procedure is commonly called “smoothing”. This task of approximating the mean function can be done essentially in two ways.

First, the parametric approach that is often used, and assumes that the mean curve *m* has some pre-specified functional form, for example, a line with unknown slope and intercept – the linear regression model and its robust variant, used in this work for comparative purposes.

As an alternative one could try to estimate m non-parametrically without reference to a specific form. The non-parametric smoothing approach offers a flexible tool in analyzing unknown regression relationships.



(a) Contour plot of correlation coefficients.



(b) Reference image.

Figure 2. A contour plot (a) of the correlation map of band 3 of two multispectral images acquired, in different dates by the Thematic Mapper sensor of the Landsat 7 satellite from a region in Middle West Brazil and the same band of the reference image (b).

3.1 Linear Robust Regression

A wide range of algorithms has been developed for radiometric normalization based on linear assumptions [1], [2], [10],[11]. A common form of linear radiometric rectification is

$$Y_k = a_k.X_k + b_k, \tag{2}$$

where X_k is the original value of DN of band k in image X and Y_k is the normalized DN of the same band on other date. The derivation of the normalization coefficients, a_k and b_k, varies according to the selected modeling. Simple regression normalization [11], for instance, applies the least-squares regression equation in order to derive normalization coefficients.

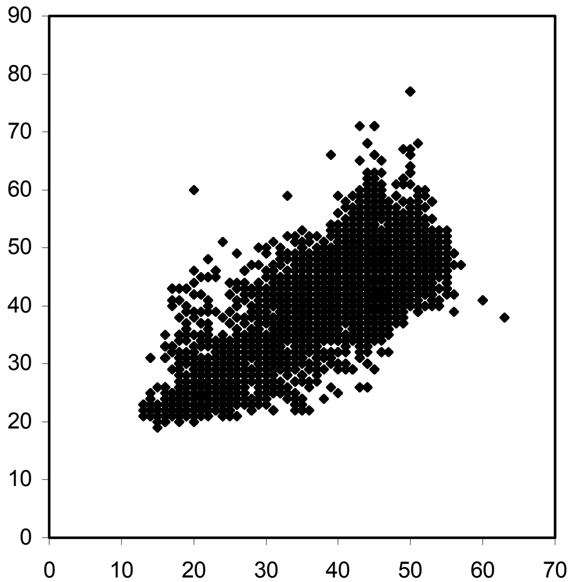


Figure 3. Scatter plot of band 5 of two multispectral images acquired, in different dates by the Landsat 7 satellite Thematic Mapper sensor from a region in Brazil.

Regression and analysis of variance procedures depend on certain assumptions, such as, a normal distribution for the error term. Sometimes such an assumption is not warranted. For example, if the distribution of errors is asymmetric or prone to extreme outliers, it is a violation of the assumption of normal errors. Robust regression function is useful when there may be outliers. Robust methods are designed to be relatively insensitive to large changes in a small part of the data.

The robust regression used in this work for comparison purposes in Section 4 uses an iteratively re-weighted least squares algorithm, with the weights at each iteration calculated by applying the bi-square function to the residuals from the previous iteration. This algorithm gives lower weight to points that do not fit well. The results are less sensitive to outliers in the data as compared with ordinary least squares regression.

3.2 Non-parametric Regression

The non-parametric approach to estimating provides a versatile method of exploring a general relationship between two variables. It gives predictions of observations yet to be made without reference to a fixed parametric model, and constitutes a flexible method of substituting for missing values or interpolating between adjacent X values.

The basic idea of non-parametric smoothing is a local averaging procedure to estimate *m* in (1). More formally this procedure can be defined as

$$\hat{m}(x) = n^{-1}\sum_{i=1}^{n} W_{ni}(x) Y_i \qquad (3)$$

where $\{W_{ni}(x)\}_{i=1}^{n}$ denotes a sequence of weights which depend on the whole vector $\{X_i\}_{i=1}^{n}$.

Most smoothing methods are, at least asymptotically, of the form (3).

A conceptually simple approach for a representation of the weight function $\{W_{ni}(x)\}_{i=1}^{n}$ is to describe its shape by a density function with a scale parameter that adjusts the size and the form of the weights near x. It is quite common to refer to this shape function as a kernel K. The kernel is a continuous, bounded and symmetric real function K that integrates to one,

$$\int K(u)du = 1.$$

The weight sequence for kernel smoothers (for one-dimensional x) is defined by

$$W_{ni}(x) = \frac{K_{hn}(x - X_i)}{\hat{f}_{hn}(x)} \qquad (4)$$

where

$$\hat{f}_{hn}(x) = n^{-1}\sum K_{hn}(x - X_i) \qquad (5)$$

and where

$$K_{hn}(u) = hn^{-1}K\left(\frac{u}{hn}\right) \qquad (6)$$

is the kernel with scale factor hn. Suppressing the dependence of h=hn on the sample size n, the kernel sequence (4) is conveniently abbreviate as $\{W_{hi}(x)\}_{i=1}^{n}$.

The function $\hat{f}_h(.)$ is the Rosenblatt-Parzen kernel density estimator of the (marginal) density of X. The form (4) of kernel weights $\{W_{hi}(x)\}_{i=1}^{n}$ has been proposed by Nadaraya and Watson [4] and

$$\hat{m}_h(x) = \frac{n^{-1}\sum_{i=1}^{n} K_h(x - X_i) Y_i}{n^{-1}\sum_{i=1}^{n} K_h(x - X_i)} \qquad (7)$$

is often called the Nadaraya-Watson estimator. K determines the shape of the kernel weights, whereas the size of the weights is parameterized by h, which is called the bandwidth.

The kernel estimate was defined as a weighted average of the response variables in a fixed neighborhood around x. However, in some applications is more appropriate to work with a varying neighbor.

The k-nearest neighbors (k-NN) estimate is a weighted average in a varying neighborhood. This neighborhood is defined through those X-variables, which are among the k-nearest neighbors of x in Euclidian distance. The simplest k-NN smoother is defined as

$$\hat{m}_k(x) = n^{-1}\sum_{i=1}^{n} W_{ki}(x) Y_i \qquad (8)$$

where

$$W_{ki}(x) = \begin{cases} n/k, & if\ i \in J_x \\ 0 & \text{otherwise} \end{cases},$$

and $J_x = \{i : X_i \text{ is one of the k nearest observations to x}\}$.

Instead of using uniform weights we used triangular k-NN weights [4] and the neighbors more distant from x were weighted more weakly. Generally speaking, the weights can be thought of as being generated by a kernel function K,

$$W_{Ri}(x) = \frac{K_R(x - Xi)}{\hat{f}_R(x)} \qquad (9)$$

where

$$\hat{f}_R(x) = n^{-1}\sum K_R(x - X_i) \qquad (10)$$

is a kernel density estimate of f(x) with kernel sequence

$$K_R(u) = R^{-1}K(u/R) \qquad (11)$$

and $R=R_n$ is the distance between x and its k-th nearest neighbor. K is

$$K(u) = (1-|u|)I(|u| \le 1) \qquad (12)$$

where I(.) is the indicator function, or Dirac delta,

$$I(u) = \begin{cases} 1 & u = 0 \\ 0 & otherwise \end{cases}.$$

We applied the triangular k-NN smoother for each band and generated tables, based on training sets, in order to correct radiometrically the other image.

Then, the proposed method consists of three major steps:

1) Selection of training set based on the correlation map of all bands;
2) For each band generate a table x versus m(x), where m(x) is obtained by applying triangular k-NN smoother, only on training set pixels; and
3) For each band of the image to be corrected apply the respective table.

Figure 4 shows the curve obtained by non-parametric regression on two multispectral images acquired, in different dates by the Thematic Mapper sensor of the Landsat 7 satellite taken from a region in Brazil.

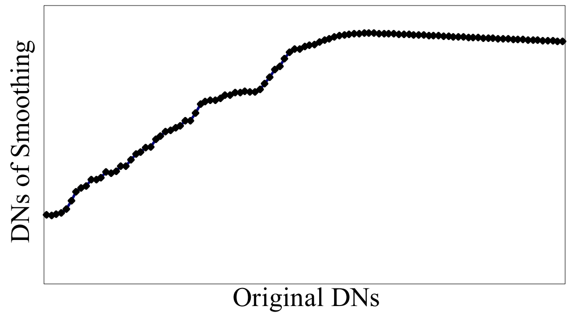


Figura 4. Non-parametric regression on band 5 of two multispectral images acquired, in different dates by the Thematic Mapper sensor of the Landsat 7 satellite.

4. Experiments and Results

In order to assess the effectiveness of the proposed technique, and to investigate how the proposed normalization improves the performance classification, we carried out classification experiments on three multitemporal data sets, and compared their results. These sets had the same reference image, acquired in 1996 by the Thematic Mapper sensor of the Landsat 7 satellite from Middle West of Brazil, with the original radiometric characteristics. In particular, a section (700x700 pixels) of a scene was selected.

The second image was radiometrically corrected by three different methods, one for each dataset: simple linear regression, robust regression and non-parametric regression.

We used the difference image, in the Change Vector Analysis context [8], [12], [5], in order to perform change-detection. Thus, the classification was binary and had two classes: changed and not changed pixels.

The subtraction operation is carried out on the pair of co-registered images. Image differencing is performed on a pixel-by-pixel basis and a difference image tends to have a histogram that is normal in shape. The pixels near the peak in zero represents pixels that have not changed very much while the pixels in the histogram tails have changed substantially.

The classification method was the same for the three datasets, a simple thresholding classifier. The same band thresholds were used for all cases. Then we used one logic rule

if ((not changed in band i) **and** …(pixel not changed in band n)) **then** pixel not changed.

In order to interpret classification accuracies we used two descriptive measures: the overall accuracy, and the Kappa coefficient. The overall accuracy is computed by dividing the total number of correctly classified pixels by the total number of reference pixels. Kappa coefficient of agreement is a measure for overall thematic classification accuracy and ranges from 0 to 1. It is a measure of the difference between the actual agreement between reference data and an automated classifier and the chance agreement between the reference data and a random classifier. A true agreement (observed) approaches 1, and chance agreement approaches 0. A Kappa coefficient of 0 suggests that a given classification is no better than a random assignment of pixels [6].

Results in Tables 1, 2, 3, and 4, refer to subsets in the images, where we synthetically made the changes to decrease the dependence on appropriate ground truth.

Tables 1, 2 and 3 show the confusion matrix for the datasets with image radiometrically corrected by linear regression, robust regression, and non-parametric regression, respectively.

Table 4 resumes the results and presents the overall accuracy and the Kappa coefficient for the three datasets. The worst performance was obtained by the dataset with an image radiometrically corrected by simple linear regression, Kappa coefficient 0.867, and the best performance was obtained through the use of the non-linear modeling of the non-parametric regression, with a Kappa coefficient equal to 0.922.

The user's accuracies are computing by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified in that category (the row total). This figure is a measure of commission error and indicates the probability that a pixel classified into a given category actually represents that category on the ground.

What are often termed producer's accuracies result from dividing the number of a correctly classified pixels in each category (on the major diagonal), by the number of training set pixels used for that category (the column total). This figure indicates how well training set pixels of the given cover type are classified.

The user's accuracy for the category changing in Tables 1,2 and 3 were 88%, 90% and 93%, and the producer's accuracies for the category non-changing were 86%, 88% and 92%, respectively.

Performances of linear regression and robust regression did not differ substantially.

Classification Data	Reference Classification	
	Non-changing	Changing
Non-changing	11106	6
Changing	1768	14030

Table 1. Confusion Matrix for dataset with image radiometrically corrected by linear regression.

Classification Data	Reference Classification	
	Non-changing	Changing
Non-changing	11103	9
Changing	1597	14201

Table 2. Confusion Matrix for dataset with image radiometrically corrected by robust regression.

Classification Data	Reference Classification	
	Non-changing	Changing
Non-changing	11106	6
Changing	1029	14769

Table 3. Confusion Matrix for dataset with image radiometrically corrected by non-parametric regression.

	Kappa-coefficient	Overall accuracy
Linear Regression	0.8672	93.4%
Robust Regression	0.8798	94.0%
Non-parametric Regression	0.9217	96.2%

Table 4. Results for the three dataset.

5. Conclusions

In this paper, a technique for relative radiometric correction has been proposed. Such technique improved the performance of the simplest change detection method based on difference image by exploiting non-parametric regression characteristics. The technique also allows automatic selection of the training set.

This improvement in radiometric correction allows use of more simple classification methods.

Further research should be conducted to test the potential improvements associated with such approach. A strategy appropriate for neighborhood choice in k-NN smoother could be used, and others adaptation algorithms could be experimented. Criterions for specify the appropriate threshold for the correlation map could be tested. In spite of the simplicity adopted, experimental results confirm the effectiveness of the presented technique.

References

[1] Chavez, P.S., Mackinnon, D.J. "Automatic Detection of Vegetation Changes in the Southwestern United States Using Remotely Sensed Images", *Photogrammetric Engineering and Remote Sensing*, Vol 60, pp. 571-583, 1994.

[2] Elvidge, C.D., Yuan, D., Werackoon, R.D. and Lunneta, R.S. "Relative Radiometric Normalization of Landsat Multispectral Scanner (MSS) Data Using an Automated Scattergram Controlled Regression",. *Photogrammetric Engineering and Remote Sensing* Vol. 61, pp. 1255-1260, 1995.

[3] Hall, F.G., Strebel, D.E., Nickeson, J.E. and Goetz, S.J. "Radiometric Rectification: Toward a common Radiometric Response Among Multidate, Multisensor Images". *Remote Sensing of Environment* Vol. 35,pp. 11-27, 1991.

[4] Hardle, W., *Applied nonparametric regression*, Cambridge University Press, 1989.

[5] Johnson, R. D.; Kasischke, E. S. "Change Vector Analysis: a Technique for the Multispectral Monitoring of Land Cover and Condition", *International Journal of Remote Sensing*, No.19, pp. 411-426, 1998.

[6] Lillesand and Kieffer "Remote Sensing and Image Interpretation", 4th Edition, John Wiley & Sons, Inc., 2000.

[7] Lunetta, R. S.; Elvidge, C. D., *Remote Sensing Change Detection*, Ann Arbor Press, 1998.

[8] Malila, W. A., "Change Vector Analysis: An Approach for Detecting Forest Changes with Landsat", *Proceedings of the Machine Processing of Remotely Sensed Data Symposium*, Ann Arbor, ERIM,1980, pp. 326-335.

[9] Mathers, P. M., *Computer Processing of Remotely-Sensed Images*, John Wiley & Sons, 1999.

[10] Schott, John R. and Salvaggio, Carl and Volchok, William J., "Radiometric Scene Normalization Using Pseudoinvariant Features", *Remote Sensing of Environment*, Vol. 26., pp 1-16, 1998.

[11] Yuan, D. and Elvidge, C.D., "Comparison of relative radiometric normalization techniques". *ISPRS Journal of Photogrammetry and Remote Sensing* Vol. 51, No. 3, pp. 117-126, 1996.

[12] Zhan, X.; Huang, C.; Townshend, R. de F.; Hansen, M.; Dimiceli, C.; Sholberg, R.; Hewson-Sacardeletti, J.; Tompkins, A. "Land Cover Detection with Change Vector in the Red and Near-Infared Reflectance Space", Geoscience and Remote Sensing Symposium, 2000 IEEE International.