Multiscale remote sensing data segmentation and post-segmentation change detection based on logical modeling: results for cost-effective forestland cover change quantification with case of Mau Forest Complex in Kenya[+]

Yashon O. Ouma [1,*] and Ryutaro Tateishi [2]

[1] *Department of Civil & Structural Engineering, Moi University*
P. O. Box 1188-30100, ELDOERT, Kenya
[2] *Center for Environmental Remote Sensing, Chiba University*
1-33 Yayoi-cho, Inage-ku, Chiba, 263-8522, JAPAN

Abstract

Quantification of forestland cover extents, changes and causes thereof are currently of global research priority. Remote sensing data plays a significant role in this exercise. However, *in-situ* and supervised-based forest dynamics mapping from remotely sensed data are limited by lack of ground-truth data collection logistics and spectral-only based methods respectively. In this paper, first results of a methodology to detect change/no-change based on unsupervised multiresolution image transformation are presented, as a cost-effective strategy for forest inventory data acquisition in the context of sustainable development are presented. The technique combines directional wavelet transformation texture and multispectral imagery in an anisotropic diffusion aggregation algorithm, using unsupervised self-organizing feature map neural network implementation. Using Landsat TM (1986) and ETM+(2001), logical-operations based change detection results for part of the Complex Mau forests in Kenya are presented. An overall accuracy for change detection of 88.4% and kappa of 0.8265 was obtained. The methodology is able to predict the change information *a-posteriori* as opposed to the conventional methods that require land cover classes *a-priori*, for change detection. Most importantly, the approach can be used to predict the existence, location and extent of disturbances within natural environmental systems. This system is to be used in the National Forestry Inventory Systems Project for applications in modeling of: climate change; national water resources budget; desertification and agricultural yield systems for food security.

Keywords: forestland cover, multiscale-multispectral anisotropic diffusion, segmentation, wavelets transformation, logical modeling, unsupervised change detection

1. Introduction

Reliable and updated forest information is necessary for many scientific and land management applications. For example, monitoring and mapping of afforestation, reforestation and deforestation (ARD) processes plus related land cover changes is important for national carbon budget inventories (potential source and sink function of vegetation as determined in the Kyoto protocol) and for regional biosphere modeling. However, lack of reliable data and survey information on forest structural changes or ARD has made the estimation of areas of intact forest and/or under land use change and their relation to economic indicators surprisingly difficult to establish. Given the importance of land cover and land use change data in conservation planning, reproducible, accurate and efficient techniques to detect forest change from multitemporal satellite imagery are desired for implementation by local conservation organizations.

Locating and characterizing areas of significant forestland cover change using remotely sensed data is important for many applications (Fung and Ledrew, 1988). These include:

[*] Corresponding author: *E-mail*: yashon@graduate.chiba-u.jp (Dr. Yashon O. Ouma).
[+] Paper published in: Computers and Geosciences Journal, 2008.

resource management (Pilon et al., 1988; and Collins and Woodcock, 1994) crisis management and response, urban planning (Lo and Shipman, 1990) and impacts of human activities on the environment (Leemans and Zuidema, 1995; Adams et al., 1995). Forestland cover information is often derived from remotely sensed images using classification algorithms e.g. (Franklin et al., 1986; Mickelson et al., 1998), many of which require a substantial amount of reference data (Townshend, 1992; Hall et al., 1995). One of the challenges to the management of forestland cover is the lack of reliable and adequate reference data. In areas where some reference data exists, they may have been collected in different ways and may have different levels of accuracy. Such heterogeneity limits efficiency, consistency and accuracy in deriving forest information. Another problem with most of the traditional procedures is that they are time consuming and rely heavily on human recognition.

Traditionally used techniques for forest change detection are able, with different levels of accuracies, to timely and effectively answer the following questions with regards to change detection, which are necessary for efficient environmental change understanding and modeling: (a) whether there is change or no-change within a given timeframe, (b) where change has occurred and the spatial extent of the change, and (c) the kind of change, that is categorization of change. From remote sensing data, general land cover change detection and analysis faces the following problems: (a) each imagery source has its own geometric, spatial, temporal and spectral characteristics. (b) Noise is often visualized as change and is non-linear for every sensor, and (c) data amounts used in change detection are becoming much voluminous.

Current change detection systems use a variety of image processing tools to make changes visible, but typically rely on manual interpretation by expert analysts to delineate the change areas. Most systems look for changes between two images: one "before" and one "after". These approaches may only work well for research projects that are geographically limited or one-time efforts, but for a production environment, an automated process is required that locates only those areas of change that are significant to a specific application (Singh, 1989). In Sgrenzaroli et al. (2002), it is mentioned that the most widely used unsupervised change detection techniques involve the pixel-by-pixel comparison of the candidate multitemporal images pairs.

Objectives of this research

Objective of the current work is on the derivation of a general framework for the identification of changes in forested areas using optical images, with minimal *a-priori* class information. In particular, the technique can be applied to pairs of multitemporal images, and for each pixel, the objectives are to determine whether the corresponding area (i) was covered with forest in both dates, (ii) was not forested in both dates, (iii) was cleared between the two considered dates, or (iv) vice versa. The proposed technique requires some pre-processing steps, which are: geometric correction and radiometric normalization selection of the band(s) for texture extraction and spectral input. The entire approach is composed of three steps: (1) 2D-discrete wavelets transform (DWT) decomposition of the band selected for texture extraction, considering 5 different scales and three detail images (horizontal, vertical and diagonal) for each scale or level. (2) Integrated smoothing of texture and spectral data, by means of anisotropic diffusion. This step aims at reducing noise without blurring edges, and it is implemented using the self-organizing feature map (SOFM) neural network. The multitemporal images are processed separately. (3) Change detection - this is achieved by: (i) classifying each image obtained from the previous step in forest/not forest by comparing it to a threshold, (ii) identifying the areas of change by comparing the previously achieved classification. Finally the accuracy of the results is assessed through the computation of the overall accuracy and kappa index and class-producers' and users' accuracy.

We illustrate the workability of this approach using 512 x 512 pixels (15.36 km x 15.36 km) of Landsat TM and Landsat ETM+ datasets dated 1986 and 2001. Figure 1 gives a summary of the workflow for this research. In Figure 1, TDM refers to the transformed divergence measure for class separability. The terms smoothing, aggregation and segmentation are used interchangeably in this paper, to mean the same thing.

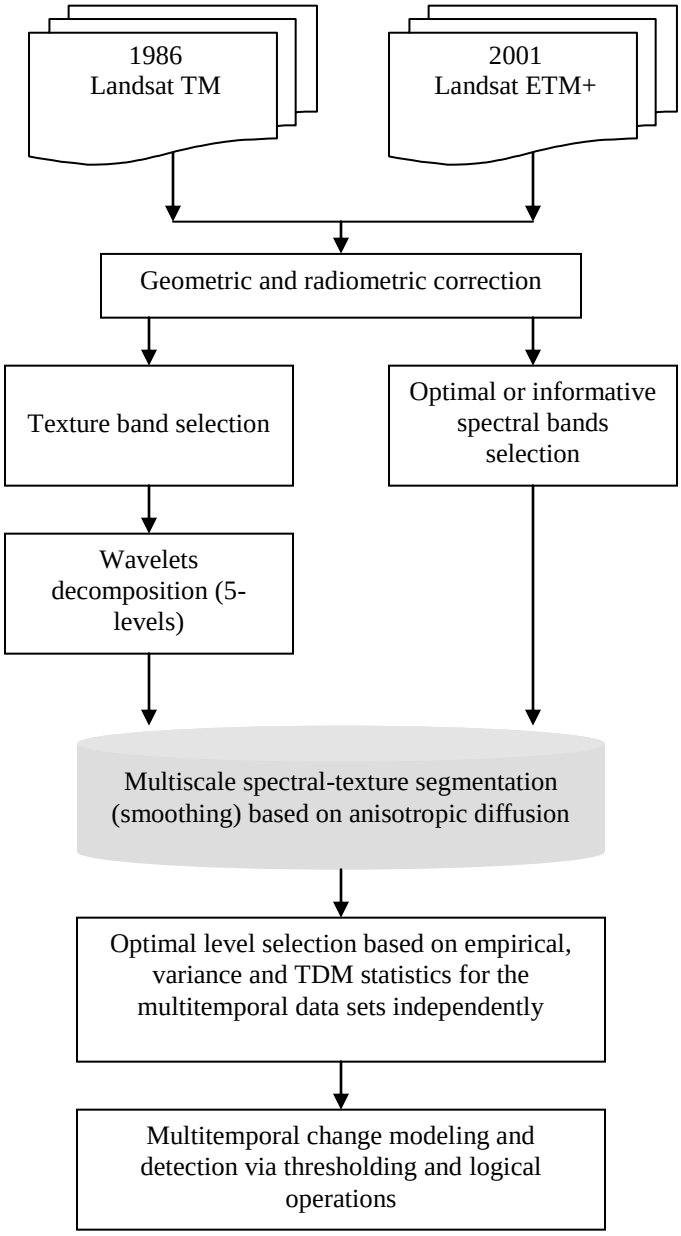


Figure 1. Proposed approach: block diagram.

2. Data and study area

2.1 Study area

The test area for this study is situated in the Mau complex forests, which lies in the Kenyan Rift Valley. In the last 25 years there has been a rapid increase in human settlement and urban related activities, great amounts of deforestation with little forestation or afforestation and agricultural activities mostly in the eastern part of the escarpment (Figure 2). This has caused immense environmental degradation within this region that has called for rapid monitoring (identification and quantification) of the changes. Conventional methods like supervised post-classification change detection that require ground-truth information *prior* to the change detection procedure, may not be suitable for the assessment of this area due to the economic logistics

associated with remoteness and data. Alternative techniques, that are less demanding and flexible in terms in input data and human interactions, are thus sought for.

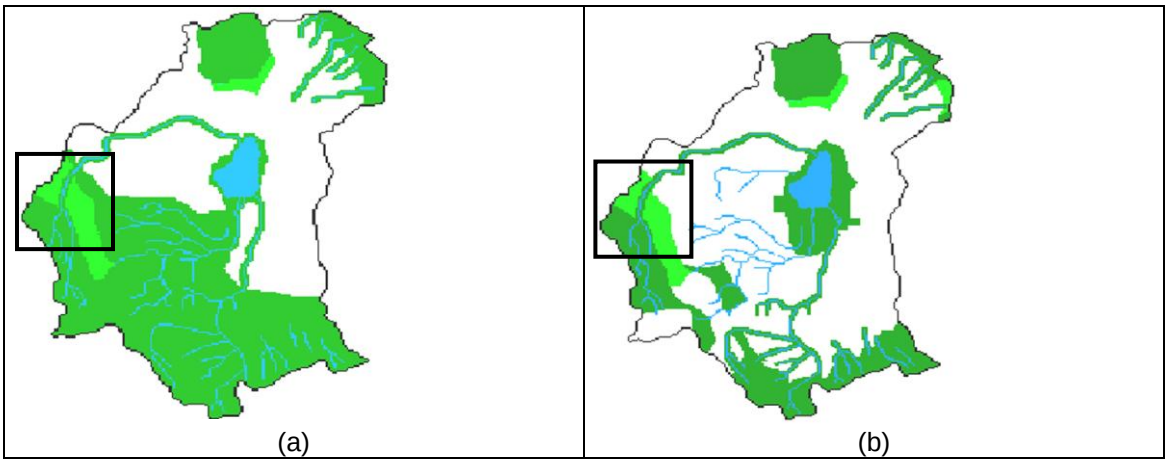


Figure 2. Lake Nakuru Basin and part of the Mau Complex. Forestland covers in 1970 (a) and in 1992(b). The rectangle shows part of the Mau Complex assessed in this study. Green color shows the forest areas, blue represents water (lake and rivers). (*Source:* Baharini Wildlife Trust.)

2.2 Experimental Data

The study site (Figure 3) chosen to test the proposed methodology approach represents part of the Mau complex forests. The size of the site chosen was 512 pixels x 512 pixels of the 30 m - Landsat TM and ETM+ dated 1986 and 2001 (Table 1). This region was subset from an entire (full) Landsat scene (path 169, row 60).

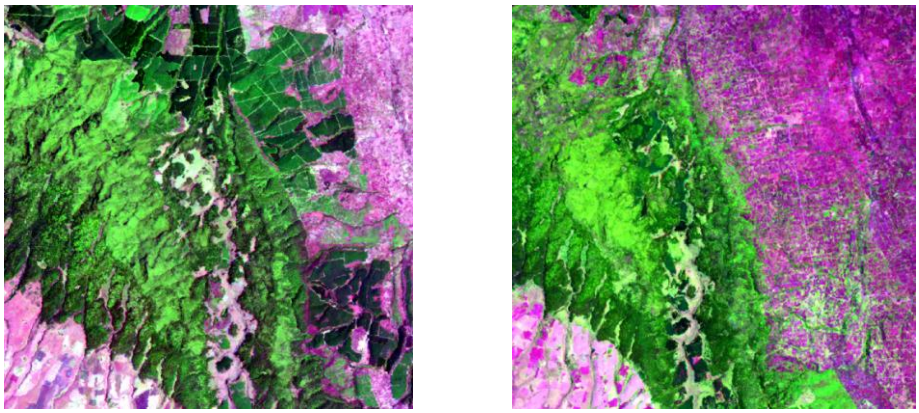
Table 1. Summary of experimental data.

Test Site	Data (Satellite & Reference)	Size	Resolution	Date of Acquisition
the Mau escarpments in Kenya	Landsat ETM+	512 x 512	30 m x 30 m	3Apr2001
	Landsat TM	512 x 512	30 m x 30 m	28Mar1986
	1: 50, 000 (TOPO)	1 Sheet	Scale: 1:50,000	1997
	1:250,000(VEGETATION)	1 Sheet	Scale: 1:250,000	1986

2.3 Data preprocessing – geometric correction and radiometric normalization

Geometric correction is the process for rectifying the common areas test data to the same projection system for producing spatially and geometrically correct land cover information. The study data was geometrically rectified to the Universal Transverse Mercator (UTM) map projection system, zone 37-east, using GCPs collected mostly around the accessible lower parts of the mountain. A 1997- 1: 50 000 topographic map and DGPS derived GCPs of the study area were used to geometrically rectify the Landsat 7 Enhanced Thematic Mapper Plus (ETM+). The RMSE report was 0.2 pixel. The Landsat 5 Thematic Mapper (TM) was then rectified to the ETM+ through image-to-image registration with RMSE of 0.25 pixel. Both the images were resampled to 30 m x 30m using the nearest-neighbor resampling method and for both images the transformations were done using the first-order polynomial. The nearest-neighbor resampling conserved the original values of the images much better, through visual inspection of the

resampling results, than the bilinear interpolation and cubic convolution because in the test area (Figure 3), there was very distinct (spectral) boundary between the forest and the no-forest extents making the closest pixel assignment more reliable than using 2 x 2 or 4 x 4 neighbors, which seemed to distort the edge pixels.



(a) (b)

Figure 3. Study site-section of the Mau escarpments (512 x 512 pixels). (a) 1986-Landsat TM bands 543 and (b) 2001-Landsat ETM+ bands 543.

2.4 Texture and spectral band selection

To determine the optimal bands for texture mapping and spectral separation of forest from non-forest, the average of the variance within each class (5-forest classes and 3-non-forest classes) was computed. The classes were chosen based in the land cover type and geographical location within the test image. Further, using ellipses, the separability of signatures was evaluated by displaying the ellipses or rectangles describing the signatures in a two-dimensional feature space. The results for the average variance comparison are presented in Table 2. In Table 2, the 5 forest representative classes were combined together and the 3 non-forest representative classes were also combined to one class so as to clearly visualize the overall separability between these two classes.

Table 2. Average variance for the forest and non-forest classes in the visible, NIR and MIR bands.

Bands and wavelengths	Blue (0.45-0.52)	Green (0.52-0.60)	Red (0.63-0.69)	NIR (0.76-0.90)	MIR1 (1.55-1.75)	MIR2 (2.08-2.35)
Forest	12.10	9.06	66.42	262.31	227.58	20.61
Non-forest	33.18	18.75	43.17	46.92	204.12	106.92

3. Multispectral-multitemporal imagery smoothing

3.1 Multiresolution wavelet decomposition (MWD)

A detailed theoretical analysis on wavelets transformation and its application to remote sensing data analysis can be found in (Ouma et al., 2005). Deviating from the monoscale-based occurrence and co-occurrence textures (Mallat, 1989) developed a theory for multiresolution signal decomposition using the orthonormal wavelet basis. The multiresolution wavelet transform decomposes a signal into low-frequency approximation and high-frequency detail information at successively coarser resolutions (multiscale texture concept). The resultant approximation is then decomposed into second level of approximation and detail, iteratively. MWD in discrete-time

corresponds to successive band filters decomposing the signal into details and overall pattern. It separates high from low frequencies recursively using the same transform at the new scale (Carvalho et al., 2001). The result is multiscale image texture represented in the vertical, diagonal and horizontal directions. The MWD results are to be used in the next phase of the anisotropic diffusion.

3.2 Multispectral anisotropic diffusion (MAD)

In this paper, we propose the anisotropic diffusion scheme, first proposed by Perona and Malik (1990), to tackle the problem of noise reduction and segmentation via edge preservation and homogenous land cover derivation in multitemporal forest land cover data. This approach is basically a modification of the linear diffusion (or heat equation), and the continuous anisotropic diffusion given by:

$$\frac{\partial I_t(x,y)}{\partial t} = div[c_t(x,y)\nabla I_t(x,y)] \tag{1}$$

where $I_t(x,y)$ is the is the image at time *t*, *div* the divergence operator, $\nabla I_t(x,y)$ the gradient of the image, and $c_t(x,y)$ the diffusion coefficient. If $c_t(x,y)$ a constant, then Eq. 1 is reduced to the *isotropic diffusion* equation, and is equivalent to convolving with a Gaussian. The idea of *anisotropic diffusion* is to adaptively choose c_t such that intra-regions become smooth while edges of inter-regions are preserved. The diffusion coefficient c_t is generally selected to be a non-negative function of gradient magnitude so that small variations in intensity such as noise or shading can be well smoothed, and edges with large intensity transition are distinctly retained.

3.2.1 MAD parametization

If we let $I_t(x,y)$ be a gray level image at coordinates (x,y) of a digital image at theoretical time t, and $I_0(x,y)$ be the original input image, the continuous anisotropic diffusion in Eq. 1 can be discretely implemented suing 4-nearest neighbors (Fig. 4a) and the Laplacian operator (Torkamani-Azar and Tait, 1996) as:

$$I_{t+1}(x,y) = I_t(x,y) + \frac{1}{4}\sum_{i=1}^{4}[c_t^i(x,y)\nabla I_t^i(x,y)] \tag{2}$$

where $\nabla I_t^i(x,y)$, *i=1, 2, 3* and *4*, represent the gradients of the 4-neighbors in the north, south, east and west directions, respectively, as shown in Fig. 4a:

$$\begin{aligned}
\nabla I_t^1(x,y) &= I_t(x,y-1) - I_t(x,y) \\
\nabla I_t^2(x,y) &= I_t(x,y+1) - I_t(x,y) \\
\nabla I_t^3(x,y) &= I_t(x+1,y) - I_t(x,y) \\
\nabla I_t^4(x,y) &= I_t(x-1,y) - I_t(x,y)
\end{aligned} \tag{3}$$

and $c_t^i(x,y)$ is the diffusion coefficient associated with $\nabla I_t^i(x,y)$, and can be considered as a function of the magnitude of gradient $\nabla I_t^i(x,y)$, i.e.,

$$c_t^i(x,y) = g(\nabla I_t^i(x,y). \tag{4}$$

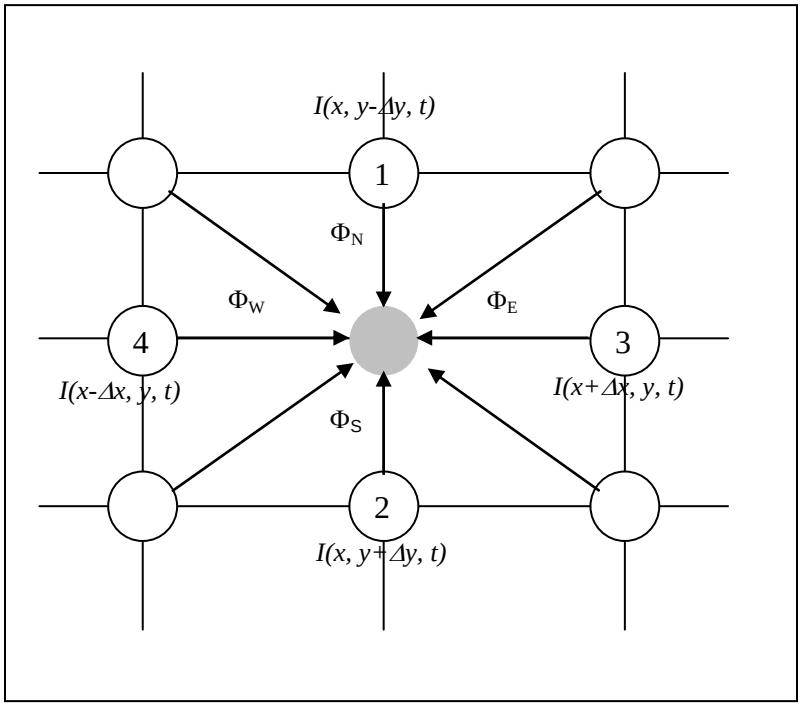


Figure 4. (a) Visualization of 2D diffusion among pixels in an image – the basic cell of 4-nearest neighbors lattice.

For the sake of simplicity, $\nabla I_t^i(x,y)$ is subsequently denoted by ∇I. $g(\nabla I)$ has to be a non-negative monotonically decreasing function with $g(\nabla I) = 1$ and $\lim_{|\nabla I \to \infty|} g(\nabla I) = 0$. The selection of $g(\nabla I)$ is to have a low coefficient value within image regions so that unwanted noise are thoroughly smoothened and inter-region edges of defects are preserved. Two possible diffusion coefficient functions are

$$g(\nabla I) = c_1(x,y,t) = \exp\left(-\left(\frac{|\nabla \mathrm{I}(\mathrm{x},\mathrm{y},\mathrm{t})|}{\kappa} \right)^2 \right) \tag{5}$$

and

$$g(\nabla I) = c_2(x,y,t) = \frac{1}{1 + \left(\frac{|\nabla \mathrm{I}(\mathrm{x},\mathrm{y},\mathrm{t})|}{\kappa} \right)^2} \tag{6}$$

In the anisotropic diffusion model of Perona and Malik (1990), the parameter κ is a constant, and must be fine-tuned for a particular application. κ is a threshold parameter, which influences the anisotropic smoothing process. It is also called the diffusion constant or the flow constant. κ in the diffusion coefficient function acts as an edge strength threshold. If the κ value is an overly small constant in all diffusion iterations, the diffusion will stop in early iterations and the background or truly homogenous areas cannot be sufficiently smoothed. This may cause false rejection of faultless homogenous land cover surfaces. Reversely, if the κ value is a large constant, the diffusion process will over-smooth in early iterations and both the background texture and edges will be removed, and this may cause false acceptance. Figures 4b depict the diffusion coefficient functions of Eqs. (5) and (6), respectively.

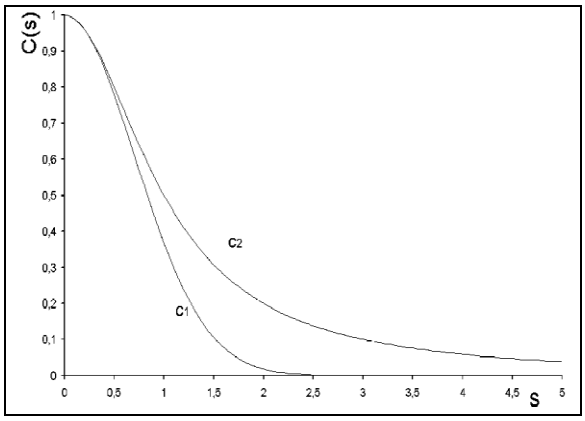


Figure 4. (b) Diffusion functions plotted as a function of image gradient, where $C(s) = g(\nabla I)$, and $s = \kappa$. (c) Flow functions plotted as a function of image gradient, where $\Phi(s) = \Phi(\nabla I)$, and $s = \kappa$.

If we define $\phi(\nabla I)$ to be a flux defined as $\phi(\nabla I) = g(\nabla I)\nabla I$, which can be expressed as $\Phi_1(\kappa)$ and $\Phi_2(\kappa)$, for c_1 and c_2 respectively, then the following plots (Figure 4c) are obtained. Simplifying further the equations above for graphical representation, we denote $h = |\nabla I(x, y, t)|$, we can denote $s = h/\kappa$, we obtain the following expressions for the conductivity functions: $c_1(x, y, t) = \exp(-s^2)$ and $c_2(x, y, t) = 1/(1 + s^2)$.

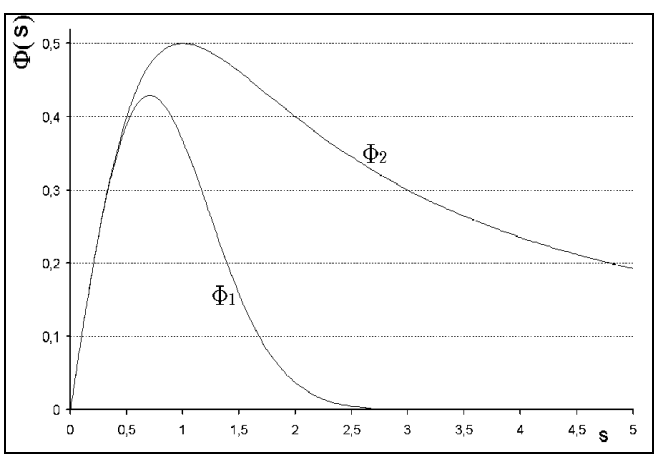


Figure 4. (c) Flow functions plotted as a function of image gradient, where $\Phi(s) = \Phi(\nabla I)$, and $s = \kappa$.

A large flux value indicates a strong effect on smoothness. Fig. 4c gives the graphs of the flux functions of the respective diffusion coefficient functions in Eqs. (5) and (6). For a given κ value, it can be seen from Fig. 5 that the diffusion coefficient function of Eq. (5) drops dramatically and approximates to zero when the gradient magnitude $|\nabla I|$ is larger than 2κ, i.e. the diffusion stops as soon as $|\nabla I| > 2\kappa$. The maximum smoothness occurs at $|\nabla I| = 0.75\kappa$ as shown in the corresponding flux function. The diffusion coefficient function of Eq. (6), instead, decreases more gradually even when $|\nabla I| > 2\kappa$. Its corresponding flux function shows that the maximum smoothness is at $|\nabla I| = 1\kappa$. Compared to Eq. (6), Eq. (5) privileges high-contrast edges over low-contrast ones. In the application of forestland cover segmentation, diffusion coefficient function of Eq. (6) is more desirable than that of Eq. (5), as it will favor even low-contrast edges and adequately considers the large homogenous patches.

The forestland cover surfaces involve inhomogeneous textures in nature and some homogenous regions may contain noise. The diffusion coefficient function of Eq. (5) may cause the diffusion process to stop in the early iterations, and the background spectral-textural will not be sufficiently removed. Given that the gradient threshold κ is a constant, the selection of a best κ value becomes extremely crucial. A large κ value will over-smooth both background information and edges. An overly small κ value disables the diffusion process and the unwanted background texture will be preserved.

In order to alleviate the limitations of the use of a constant κ, we used the annealing *n*th root function for the gradient threshold κ, to determine the suitable thresholds. Its value will be reduced as the diffusion iteration increases. In each diffusion iteration, the gradient magnitude, that is intensity-contrast-texture, will be reduced in the filtered image. A constant κ will eventually smooth out the irrelevant information like noise. However, as the gradient threshold adaptively decreases with the increment of iterations, the diffusion process has no effect on the inhomogeneous regions while it can gradually remove the background textures as long as the decrement of gradient magnitude in homogeneous regions is competitive with the decrement of the κ value. The annealing *n*th root function used in this study is defined by:

$$\kappa(t) = \kappa(0)t^{-(1/n)} \tag{7}$$

where $\kappa(t)$ is the gradient threshold at iteration *t*, $\kappa(0)$ is the initial value, and *n* is a positive integer. Figures 5(a)-(d) presents the plots of four root functions with *n*=1, 2, 3 and 4 respectively.

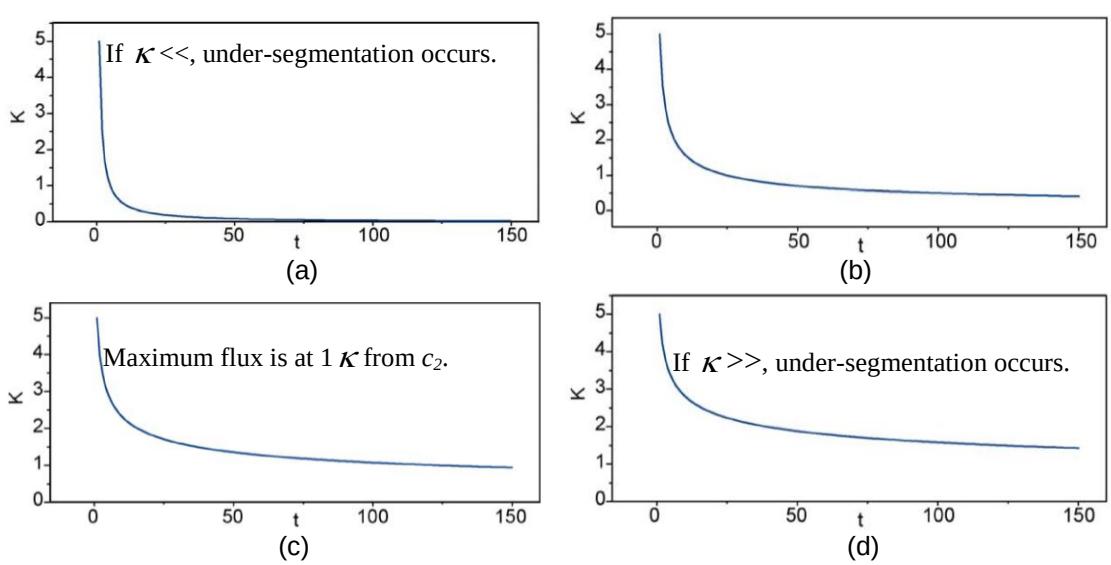


Figure 5(a)-(d). Plots for the first four root functions from (a)-(d) for *n* = 1, 2, 3 and 4 respectively. *t* is the number of iterations; *n* is the level or scale of diffusion and κ is the gradient threshold.

Note how the shape of the function is affected as the value of *n* is changed. The graphs show that a small *n*, such as *n*=1, will make the κ value drop rapidly and cause the diffusion process to stop at a small number of iterations *t*. As *n* increases, the κ value will decrease gradually and result in fast smoothness in a small number of iterations. An overly large value of *n* may over-smooth both background textures and subtle defects in early iterations. From the results in Figure 5, *n*=3 gave the optimal solution whereby the maximum diffusion occurs at approximately 1κ. Thus the used diffusion function was defined by:

$$g(\nabla I) = c_2(x,y,t) = \frac{1}{1+\left(\frac{|\nabla \mathrm{I}(\mathrm{x},\mathrm{y},\mathrm{t})|}{\kappa}\right)^2} = \frac{1}{1+\left(\frac{|\nabla \mathrm{I}(\mathrm{x},\mathrm{y},\mathrm{t})|}{\kappa(0)t^{-(1/3)}}\right)^2} \tag{8}$$

Since the gradient threshold value κ is adaptively decreased as the iteration number increases, the selection of the initial value $\kappa(0)$ is not as crucial as that of a constant κ in the Perona and Malik's model. As seen in 4c, the flux function of the diffusion coefficient in Eq. (6) shows that the maximum smoothness is given by $|\nabla I| = 1\kappa$. We therefore set the initial value $\kappa(0)$ of the annealing cubic-root function to the average gradient magnitude, Eq. 9, of the whole image used to control the diffusion the spectral diffusion, and compared this with other possible image spatial parameters like spectral and textural gradient information.

$$\kappa(0) = \frac{1}{4MN}\sum_{i}^{4}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\nabla I_t^i(x,y) \tag{9}$$

For example, the implementation, for an image $\{I(x,y)\}, 0 \le x, y < N$, the wavelet transform produces four components {*LL, LH, HL,* and *HH*}, *LL* corresponds to the smoothed image, *HL* contains the horizontal edges, *LH* contains the vertical edges, and *HH* contains edges not parallel to the axes. The estimation of conduction parameter $\kappa(x,y)$ based on wavelet transform upon experimentation is given by:

$$\kappa(x,y) = \frac{\kappa_0}{1+\alpha|HL(x,y)|^t + \beta|LH(x,y)|^t + \gamma|HH(x,y)|^t} \tag{10}$$

where κ_0 is the 90% value of average value of $|\nabla I_t^i(x,y)|$, the weights satisfy the conditions: $0 \le \alpha, \beta, \gamma \le 1; \alpha+\beta+\gamma = 1$ and $|HL(x,y)^t|$, $|LH(x,y)^t|$, $|HH(x,y)^t|$ are the wavelet transform texture gradients in the horizontal, vertical and diagonal directions respectively at *t* optimal iterations. A thorough experimentation showed that the selected $\kappa(0)$ works successfully for this application. This was determined purely based on empirical testing ad may vary from one application to another. Thus Eq. 9 may be considered as a more universal approach.

For example in related utility of anisotropic diffusion, Barash (2002) showed that an iteration of adaptive smoothing: $I_{t+1}(x,y) = \frac{\sum_i\sum_j I_t(x+i,y+j)w_t(x+i,y+j)}{\sum_i\sum_j w_t(x+i,y+j)}$, is an implementation

of the discrete version of the anisotropic diffusion equation if the weight (w_t) is taken as the same as the diffusion coefficient c. This weight factor can be modeled in the case remote sensing data as the input texture features at the corresponding levels of diffusion or scale, which is represented by the conduction parameter κ. At each level, each band is smoothed based on the homogeneity of the input directional textures.

3.3 Self-organizing feature map (SOFM)

The SOFM is trained iteratively. In each training step, one sample vector x from the input data set is chosen randomly and the similarity between it and all prototypes of the map are calculated using a certain distance measure, typically Euclidian distance. The unit whose incoming connection weights have the greater similarity with the input pattern $\boldsymbol{x}$ is called the Best-Matching Unit (BMU), denoted as c:

$$\|x - m_c\| = \min_i \{\|x - m_i\|\} \tag{11}$$

where $\|\cdot\|$ is the distance measure.

After finding the BMU, the prototype vectors (or connection weights) of the SOFM are adjusted. SOFM creates a topological mapping by updating not only the BMU's weights, which are adjusted (i.e. moved in the direction of the input pattern by a factor determined by the learning rate), but also adjusting the weights of the adjacent output units in close proximity to the neighborhood of the winner. So not only does the BMU get updated, but the whole neighborhood of output neurons gets moved closer to the input pattern.

After presentation of the data vector **x**, the weight vectors of the map neurons are updated according to the learning rule. The SOFM update rule for the weight vector of neuron i is:

$$m_i(t+1) = m_i(t) + \eta(t) h_{ci}(r(t))[x(t) - m_i(t)] \tag{12}$$

where t denotes time, $\eta(t)$ is the learning rate, that is a monotonically decreasing function of time between (0,1) and regulates the amount of the weight update at time t, and $h_{ci}(r(t))$ is the neighborhood kernel centered on the winner unit (the BMU) c, with neighborhood radius $r(t)$, which typically decreases with time. There are two alternatives for performing the prototype vectors' update: sequential and batch training algorithms (Kohonen, 2001). The algorithm utilized in this work is the sequential one.

Generalized SOFM Algorithm

Step 1: Initialize the synaptic weights of the network, $V_j(0)$, to small, different, random numbers at iteration $k = 0$.
Step 2: Draw a sample y from the input set.
Step 3: Find the best-matching (winning) neuron $r(y)$ at iteration k, using the minimum distance Euclidean-criterion

$$r(y) = \min\{\|y - V_j\| \, j = 1,2,.....,L\}.$$

Step 4: Update the synaptic weight vectors using the update formula

$$V_{r(y)}^{k+1} = V_{r(y)}^{k} + \eta^k (y - V_{r(y)}^{k}),$$

$$V_j^{k+1} = V_j^k + (\eta^k)^2 (y - V_j^k)$$

$$\forall j \in \Omega_{r(y)}(k),$$

where $\Omega_{r(y)}(k)$ is the neighborhood of $r(y)$.

Step 5: Increment k by 1, goto *step 2*, and continue until the synaptic weights V_j reach their steady state values.

4. Methodology implementation

4.1 MAD-SOFM implementation

Based on the SOFM concept, the aggregation process using the multiscale anisotropic diffusion is accomplished by constructing multiscale or pyramid structure with successive anisotropic diffusion of the original image and then linking nodes at each level to the final node at root level. The node value at (*i, j*) of level *l* is initialized by anisotropic diffusion result in the 4-directions or nearest-neighbors level of level *l*-1. The root level can be selected for segmentation and thus, the root level determines the number of final segments. Generally, linking process where node values are updated and each node is inked to one of the root node is iteratively performed; each node is linked to the father that has the root value closest to its value, the node value is updated based on the node values of its sons, and finally, root values are assigned to the descendent nodes successively, as in Figure 6.

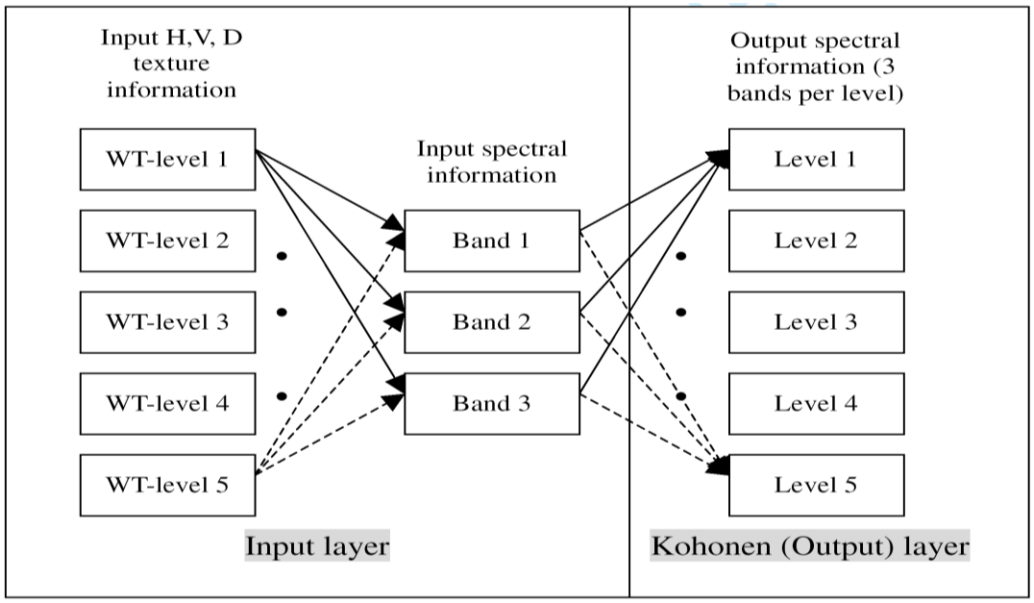


Figure 6. Structure of the SOM-NN consisting, in principal of: input layer - image wavelets textural and spectral information and output (Kohonen) layer – smoothened multiscale visible, NIR and MIR bands.

In summary, MAD-SOFM algorithm modifies the pixel values of the input image by consecutive weighted averaging with neighboring pixels. The rate of smoothing is a function of the sum of the local spectral and textural similarity. The similarity values are scaled by the spectral rate and textural rate, respectively. The following are the main characteristics of the applied MAD-SOFM algorithm:

(i) Simultaneous modification of all input spectral bands (multispectral anisotropic diffusion).

(ii) Operation at five levels of scale/resolution of the multispectral image. Each network level is a two dimensional array of nodes containing the current modified image at a given scale and location. The bottom network level consists of one node for each pixel in the full scale (initial image). This level is initialized using the input multi-spectral image. Successive levels undergo a 2:1 downsampling.

(iii) Both spectral and texture information of images are used to determine the amount of modification at a given scale and location. Two kinds of data are used in this step, (1) multispectral images of two dates, three bands (3,4,5) for each date, and (2) five successive levels of texture maps (consisting of three detail images for each level), resulting from wavelet transform. Considering the largest detail image, which

contains half of the pixels and lines of the initial image in wavelet transform, the window size of the initial image for segmentation, was 256 pixels and 256 lines (equal to the largest detail image resulting from wavelet transform). The algorithm was run for the multitemporal images separately at five levels using the same parameters.

5. Results

5.1 Wavelets transformation results

The *db8* (Daubechies, 1992) wavelet filter coefficients as in MATLAB WaveLab were chosen based on earlier experiments (Ouma et al., 2005) to create multiscale texture maps resulting in a set of detail images at different scales. Detail images occur in groups of three (horizontal, vertical and diagonal) at each scale. The first group of detail images will be of size P/2 pixels and L/2 lines from the initial image of P pixels and L lines. The *n*th group will be of size P/(2^n) pixels and L/(2^n) lines. The transformation was carried out at 5 scales for both data sets. In multilevel segmentation using anisotropic diffusion technique, the first 5 groups of detail images are used as input texture maps.

5.2 MAD-SOFM results

5.2.1 Testing of the MAD-SOFM performance

For the MAD-SOFM performance evaluation, a subset of band 4 (NIR-band) of the latest Landsat data, that is 2001 ETM+, was selected as test area. The results for the selection of optimal parameters were evaluated via boundary overlay, of the subset image, of manually delineated boundary and diffusion or segmentation output boundaries, for the 5 levels. First, the ground truth feature of a geometrical region is extracted manually from a test image, and then compared with the results extracted through segmentation results. Next, the deviation of the segmented feature from the original feature was numerically evaluated. For this, two features were overlapped and their boundaries were compared. Then, the numbers of pixels in the interior and exterior of the original feature that is not matched in the segmented feature were counted as the interior and exterior error, respectively. That is, the results were compared in terms of interior-exterior boundary geometric errors (outliers). The results are presented in Table 3 for the original imagery against the five-levels. It was found that level 3 gave minimal boundary errors, as compared to the manually derived portions. Level 1 and level 5 gave the poorest results, of course owing to the observations already discussed above. The results in Table 3 are comparable to those of Figure 7. Implying that for *n*=1, there is an obvious pre-mature segmentation effects, irrespective of the number of iterations.

Table 3. Evaluation of segmentation performance by the multiscale anisotropic diffusion.

Segmentation level	Interior	Exterior	Total error
Level-1	911	843	1754
Level-2	209	120	329
Level-3	99	107	206
Level-4	112	143	255
Level-5	459	302	761

5.2.2 MAD-SOFM output results

The output results for this phase are 5-levels or scale with level 1 being of size 256 x 256; level2 – 128 x 128; level3 – 64 x 64; level4 – 32 x 32; and level5 – 16 x 16. Each level contains

three bands corresponding to the selected bands 3, 4 and 5. Not all the levels are relevant in mapping forest and no-forest information. Thus by computing the TDM separability between the two classes (forest and non-forest) at each level for each band, we statistically determine the optimal or most informative level amongst the five. The average variances for the test classes were also computed in the respective output bands. The TDM average was computed as follows: first, for each forest class, its separability (TDM) with the three non-forest classes was computed. The TDM and class average-variances were computed and compared as illustrated in Table 4.

Table 4. Illustration of TDM computation procedure and class average-variance comparisons.

On the use of TDM and class average variance. Example illustration with band x and level y, whereby y=1,2,3,4,5 and each y has three bands (x=1,2,3 for each level).

Forest (F_i)	Non-Forest (nF_j)	TDM between (F_i and nF_j)				Variance for F_i
F_1	nF_1	TDM_{11}	TDM_{12}	TDM_{13}	Average TDM_1	$\sigma^2_{F1:level1-band1}$
F_2	nF_2	TDM_{21}	TDM_{22}	TDM_{23}	Average TDM_2	.
F_3	nF_3	TDM_{31}	TDM_{32}	TDM_{33}	Average TDM_3	.
F_4		TDM_{41}	TDM_{42}	TDM_{43}	Average TDM_4	.
F_5		TDM_{51}	TDM_{52}	TDM_{53}	Average TDM_5	$\sigma^2_{F5:level1-band1}$
Example: For level 1, band 1 (L1, B1)					$TDM_{L1,B1}$_AVERAGE	$VAR_{L1,B1}$_AVERAGE

The TDM_AVERAGE are plotted as in Fig. 11. To avoid multiple variances, we only state the VAR_AVERAGE performances. (The same concept of average variances is used in Table 2).

TDM_{11} is derived from F_1 and nF_1
TDM_{12} is derived from F_1 and nF_2 and
TDM_{13} is derived from F_1 and nF_3
.
.
and so on for the rest of the F_i-classes.

This resulted in three TDM measures for each forest class, for every band in a given level. This was then averaged to give a single TDM measure per forest class such that for band *i* in level *j*, five TDM measures (corresponding to the five-forest classes) were obtained and the average taken. These results were plotted as shown in Figure 7. The average variances were also computed in a similar manner for the forest classes only. Figure 7 shows the TDM results for TM and ETM+ datasets respectively.

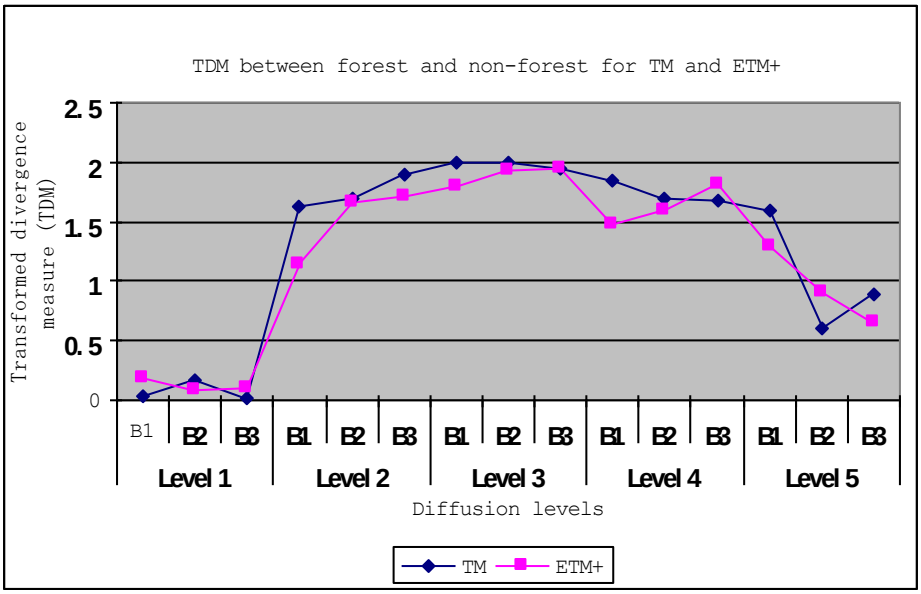
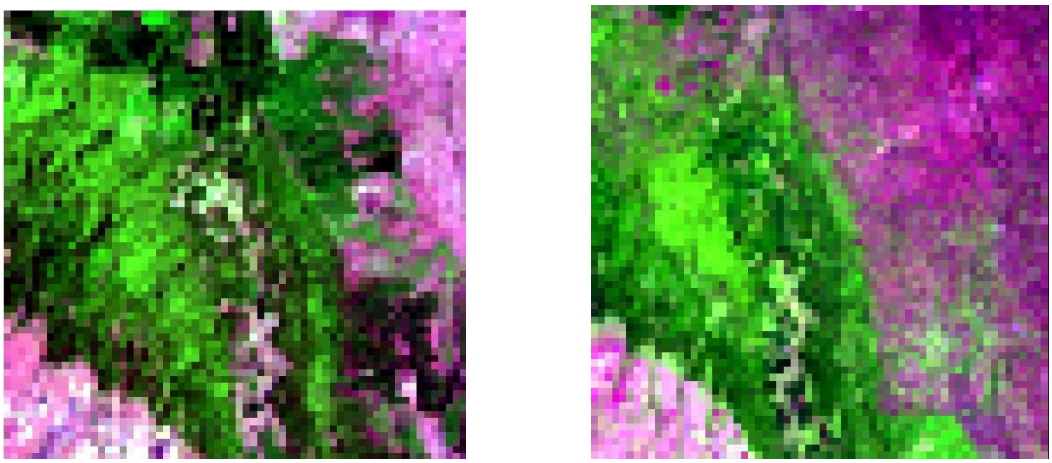


Figure 7. TDM between forest and non-forest classes for TM and ETM+ for the five levels of diffusion. *B1, B2* and *B3* refer to TM and ETM+ bands 3, 4 and 5 respectively.

Overally, level 3 gave the highest TDM meaning highest separability between forest and non-forest (Figure 8 for TM and ETM+ results). Level 1 had the least separability with TDM values lower than 0.500, followed by level 5. Levels 2 and 4 were the only levels that comparable with level 3. From the above results, the third scale/level gave the highest separability between the test classes. This observation was also confirmed by comparing the average variance for the levels 3 and 4 diffusion results for the forest classes. Level 3 gave a higher average variance than level 4.



(a) (b)

Figure 8. Anisotropic diffusion results form (a) TM image and (b) ETM+ image presented as RGB for levels 1, 2 and 3 respectively.

As the levels increases from 1 through 5, the nature of the aggregation of neighboring pixels also changes. The results of the different levels are influenced by the scale or resolution of the texture information. Level 1 result corresponds to the initial network output. Obviously as already discussed, as *n* increases from 1 through 5, there is observed pre-mature termination of the

segmentation process, resulting into level 1 results. Stability is gained at level 3 and over-segmentation is observed past this level 3. Thus from level 4, there is evident on-set of over-aggregation such that the well aggregated features at level 3 begin to be disintegrated.

6. Multiscale change detection based on logical modeling

6.1 Logical modeling – concept and formulation

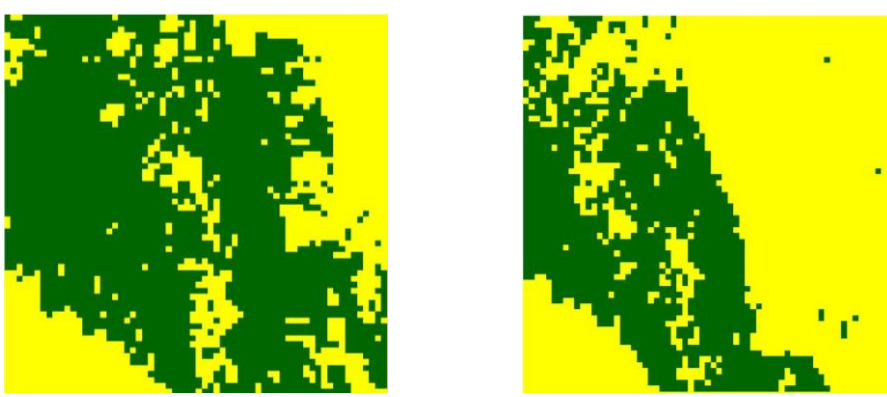
Change detection is derived from the segmented imagery. We introduce the use of logical modeling (LM) concept in change detection. LM is considered as formalism for reasoning under uncertainty, with special advantages in its treatment of ambiguous data and the ignorance arising from or about it, as is demonstrated below. In our case, we have two hypothesis or subsets {forest (*F*), non-forest (*nF*)}, the plausible combinations upon change/ no-change are {[*F, F*], [*nF, nF*], [*F, nF*], [*nF, F*]}. These sets imply that between two temporal intervals, it is possible to have no-changes [*F, F*] and [*nF, nF*], or changes of the order [*F* to *nF*] and [*nF* to *F*]. Polygons or theme maps, are generated for reasoning from the {*F*} and {*nF*}. Logical operations are performed on the database {[*F, F*], [*nF, nF*], [*F, nF*], [*nF, F*]} bitmap segments with the following "boundary" condition:

$$L\{[F,F],[nF,nF],[F,nF],[nF,F]\}=1 \tag{13}$$

This boundary condition (Eq. 13) states that within the two time periods, the expected probability for change and no-change irrespective of any other conditions is constant and equal to 1.

6.2 Theme map generation

In order to quantify the changes, allocation of a unique value to each smoothed area (representing one class) is required. Because the forest/non-forest areas have been isolated via the diffusion smoothing procedure, it is direct and easy to determine the threshold DN values that define these forest and non-forest regions. This was achieved by comparing: density slicing, *k*-means clustering and histogram for the forest/non-forest regions. The former two methods gave similar results, which were then used in this study. Note that level 3 results consisting of VIS, NIR and MIR were used in this stage. The thresholding were then used to allocate a unique value to the range of values representing forest areas. Then the logical operation *NOT* was applied to separate the forest class from non-forest areas. Figure 9a and 9b shows the results of this procedure for the TM and ETM+ data sets respectively derived from level 3 results.



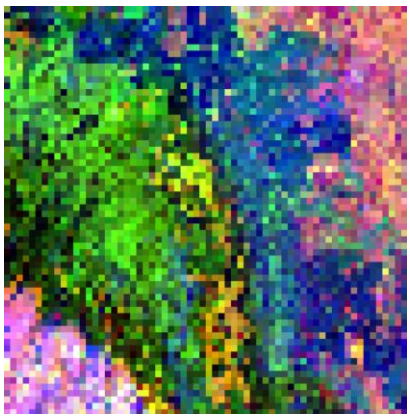
(a) (b)

Figure 9. Theme map of (a) 1986-TM data (b) 2001-ETM+ images. Green color is forest and yellow color represents non-forest information.

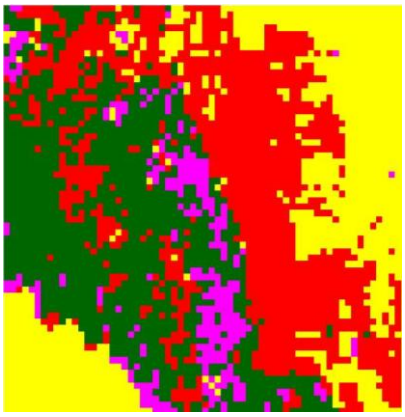
6.3 Multiscale change detection

Change/no-change theme maps resulting from segmentation are then extracted using the LM logical operation on the already created theme maps. The logical *AND* binary operator separated no-change areas of forest and non-forest areas, and logical *SUB* operation was used to separate change areas (forest (*F*) to non-forest (*nF*), and non-forest (*nF*) to forest (*F*)). With this approach, areas with different properties (unknown) are mapped out automatically. A composite of the oldest image, which is: TM bands 4 and 5, and channel 1-level3 of ETM+ (Figure 10a), is used to visualize the change/no-change. This visualization approach is synonymous to the multitemporal hyper-clustering whereby a post classification comparison of is carried out between the time 1 data and time 1 plus time 2 data sets. This approach gives a quick and rough indication of the change/no-change indication, but cannot be wholly relied upon. The false color composite in Figure 10a can be cross-correlated to the color information in Figure 10b. It thus acts as complementary information to the change/no-change information depicted in Figure 10b.

The result of this step was a change/no-change map. This map was generalized into the following four classes: (1) no-change for forest areas, (2) no-change for non-forest areas, (3) change of forest to non-forest areas, and (4) change of non-forest to forest areas. A summary on the proportion statistics is illustrated in Figure 10c.



(a)



(b)

Class	Class description	Proportion, %
	Non-forest-unchanged	27
	Deforestation	28
	Forest-Unchanged	38
	Forestation (young and different species)	7

(c)

Figure 10. (a) False color composite using channel 1 of 3rd scale of ETM+, and bands 4 and 5 from TM images. (b) LM based change/no-change (transition) map, and (c) legend and land cover proportions.

The legend of the classes given in Figure 10c is explained as follows: (i) non-forest - soil, pasture and crops; (ii) forest-unchanged – primary forest (preserved tropical rainforest); (iii) forestation – needleleaf old and young woody secondary forest cover; and (iv) deforestation – same as non-forest. Reforestation took place with different species of trees i.e. from natural broadleaves forest to needleleaves pine forest. Thus the observed difference is as a result of leaf/height texture of the young and different species. It is also possible to conclude that part of the forested land cover had been cleared in between 1986 and 2001 and then reforestation took place. This process may be called forest transition. While the excised forest cover is

predominantly used for subsistence farming, the other parts of the forest are illegally logged for timber and charcoal burning (fuel). This detailed description is not within the scope of the current study, however we present the accuracy of the results in the next subsection for accuracy assessment.

6.4 Change results accuracy assessment

The results for the four classes (refer to Figure 10c) are presented in Table 5.

Table 5. Contingency table for accuracy assessment.

		Change map class (pixel counts)					
	Cover type	Non-forest (unchanged)	Deforestation	Forest-unchanged (Broadleaf)	Afforestation (Pine)	Total	Producers' Accuracy (%)
Ground reference	Non-forest (unchanged)	910	20	1	0	931	97.7
	Deforestation	40	343	72	2	457	75.1
	Forest-unchanged (broadleaf)	0	62	376	14	452	83.2
	Afforestation (Pine)	0	0	19	127	146	87.0
	Total	950	425	468	143	1,986	
	Users' Accuracy (%)	95.8	80.7	80.3	88.8		

Overall accuracy = 88.4% and kappa = 0.8265

The overall accuracy achieved was 88.4% (kappa=0.8265). This can be said to be highly accurate results especially given the fact that the respective class accuracies were generally high as depicted by the users' and producers' accuracies and the efforts made to collect the ground-truth data.

7. Conclusions and some future considerations

The change detection and recognition system developed in this study could reliably detect and identify deforested, forested and no-changes as per the accuracy assessment results. However, a number of choices has to be made, which also influence the reliability of the results: (a) selection of the appropriate bands for texture and spectral information extraction; (b) selection of the suitable or optimal detail level to be used for change detection; and (c) threshold selection, from the smoothened optimal level bands, to discriminate between forested and non-forested areas. The focus in this study is to examine the forested areas. Therefore the selection of the optimal scale is tailored to maximization of forest information. Thus at this optimal scale, the methodology is expected to work well for any other similar type of forest (species), irrespective of the non-forest land cover types.

The results of the study tried to answer the questions posted in this research as: (a) change/no-change detection, (b) where the change occurred and spatial extent of the change, and (c) categorization of change/no-change. Using average variance and separability ellipse, it is also concluded that the red (band 3), near infrared (band 4) and the mid-infrared (band 5) were most significant spectral bands in isolating forest from non-forest regions.

This system is to be used in the National Forestry Inventory Systems Project for applications in modeling of: climate change; national water resources budget; desertification and agricultural yield systems for food security.

Acknowledgments

The authors wish to thank the Japanese Ministry of Education, for sponsoring this study.

References

Adams, J.B., Kapos, D.E.V., Filho, R.A., Roberts, D.A., Smith, M.O., Gillespie, A., 1995. Classification of multispectral images based on fractions of endmembers: application to land-cover change in the Brazilian Amazon. Remote Sensing of the Environment 52, 137-152.

Barash, D., 2002. A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 844–847.

Carvalho, L.M.T., Fonseca, L.M.G., Murtagh, F., Clevers, J.G.P.W., 2001. Digital change detection with the aid of multi-resolution wavelet analysis. International Journal of Remote Sensing 22, 3871-3876.

Collins, J.B., Woodcock C.E., 1994. Change detection using the Gramm-Schmidt transformation applied to mapping forest mortality. Remote Sensing of Environment 50, 2647-2679.

Daubechies, I., 1992. Ten Lectures on Wavelets. Philadelphia: SIAM, 1992.

Franklin, J., Logan, T.L., Woodcock, C.E., Strahler, A.H., 1986. Coniferous forest classification and inventory using Landsat and digital terrain data. IEEE Transactions on Geoscience and Remote Sensing 24, 139-149.

Fung, T., LeDrew, E., 1988. The determination of optimal threshold levels for change detection using various accuracy indices. Photogrammetric Engineering and Remote Sensing 54, 1449-1454.

Hall, F.G., Townshend, J.R., Engman, E.T., 1995. Status of remote sensing algorithms for estimation of land surface state parameters. Remote Sensing of Environment 51, 138-156.

Leemans, R., Zuidema, G., 1995. Evaluating changes in land cover and their importance for global change. Trends in Ecology and Evolution 10, 76-81.

Lo, C.P., Shipman, R.L., 1990. A GIS approach to land-use change dynamic detection. Photogrammetric Engineering and Remote Sensing 56, 1483-1491.

Mallat, S.G., 1989. A theory for multi-resolution signal decomposition: The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2, 674-693.

Mickelson, J.G., Civco, D.L., Silander, J.A., 1998. Delineating forest canopy species in the northeastern United States using multi-temporal TM imagery. Photogrammetric Engineering and Remote Sensing 64, 891-904.

Ouma, Y.O., Ngigi, T.G., Tateishi, R., 2005. On the optimization and selection of wavelet texture for feature extraction from high-resolution satellite imagery with application towards urban-trees delineation. International Journal of Remote Sensing, accepted paper due for publication.

Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 629-639.

Pilon, P.G., Howarth, P.J., Bullock, R.A., Adeniyi, P.O., 1988. An enhanced classification approach to change detection in semi-arid environments. Photogrammetric Engineering and Remote Sensing 54, 1709-1716.

Sgrenzaroli, M., Baraldi, A., Eva, H., Grandi, G.D., Archard, F., 2002. Contextual clustering for image labeling: an application on degraded forest assessment in Landsat TM images of the Brazilian Amazon. IEEE Transactions on Geoscience and Remote Sensing 40, 1833-1848.

Singh, A., 1989. Review article: digital change detection techniques using remotely sensed data. International Journal of Remote Sensing 10, 989-1003.

T. Kohonen, *Self-Organizing Maps*. New York: Springer (SIAM), 2001, pp. 501.

Townshend, J.R.G., 1992. Land cover. International Journal of Remote Sensing 13, 1319-1328.

Torkamani-Azar F., Tait, K.E., 1996. Image recovery using the anisotropic diffusion equation, IEEE Transactions on Image Processing 5, 1573–1578.