

# Unsupervised Land Cover Change Detection: Meaningful Sequential Time Series Analysis

Brian P. Salmon, Jan Corne Olivier, Konrad J. Wessels, Waldo Kleynhans, Frans van den Bergh, and Karen C. Steenkamp

**Abstract**—An automated land cover change detection method is proposed that uses coarse spatial resolution hyper-temporal earth observation satellite time series data. The study compared three different unsupervised clustering approaches that operate on short term Fourier transform coefficients computed over subsequences of 8-day composite MODerate-resolution Imaging Spectroradiometer (MODIS) surface reflectance data that were extracted with a temporal sliding window. The method uses a feature extraction process that creates meaningful sequential time series that can be analyzed and processed for change detection. The method was evaluated on real and simulated land cover change examples and obtained a change detection accuracy exceeding 76% on real land cover conversion and more than 70% on simulated land cover conversion.

**Index Terms**—Change detection, clustering, satellite, time series.

## I. INTRODUCTION

THE transformation of natural vegetation by practices such as deforestation, agricultural expansion and urbanization has significant impacts on hydrology, ecosystems and climate [1]–[3]. Coarse spatial resolution satellite data provide the only regional, spatial, long-term and high temporal measurements for monitoring the earth's surface. Automated land cover change detection at regional or global scales, using hyper-temporal, coarse resolution satellite data has been a highly desired but elusive goal of environmental remote sensing [4]–[6]. Most change detection studies rely on image differencing, post-classification comparison methods and change trajectory analysis [7]–[13], and the data is mostly treated as hyper-dimensional, but not necessarily as hyper-temporal. These methods therefore do not fully capitalize on the high temporal sampling rate which captures the dynamics of different land cover types, nor do they provide automated change detection capabilities.

A time series is a sequence of data points measured at successive (often uniformly spaced) time intervals. Time series

analysis comprises methods that attempt to understand the underlying forces structuring the data. Analyzing this structure enables the identification of patterns and trends, detection of change, clustering, modeling, and forecasting [14]. In the time series context, complete clustering is when the entire time series is taken as a discrete object and clustered with conventional methods. In contrast, subsequence clustering is performed on streaming time series that are extracted with a sliding window from an individual time series. Time series analysis is less concerned with the global properties of a time series and more interest in the subsequences of a time series [15]. A subsequence  $x_p(t)$  for a given time series  $x(t)$  of length  $N$ , is given as

$$x_p(t) = (x(t_p), x(t_{p+1}), \dots, x(t_{p+Q})) \quad (1)$$

for  $1 \leq p \leq N - Q + 1$ , where  $Q$  is the length of the subsequence. The sequential extraction of subsequences in (1) is achieved by using a temporal sliding window that has a length of  $Q$  and position  $p$  that is incremented with a natural number  $\mathbb{N}$  to extract sequential subsequences  $x_p(t)$  from  $x(t)$  (Fig. 1). The signal processing and data mining communities have made wide use of the clustering of subsequence time series,  $x_p(t)$ , that were extracted using a temporal sliding window [16]–[18]. To date, it has found very limited applications on satellite time series data.

Recently the data mining community's attention was brought to a fundamental limitation of the clustering of subsequences that were extracted with a sliding window from a time series [15]. The sliding window causes the clustering algorithms to create *meaningless* results as it forms sine wave cluster centers regardless of the data set, which clearly makes it impossible to distinguish one dataset's clusters from another. This is due to the fact that each data point within the sliding window contributes to the overall shape of the cluster center as the window moves through the time series. This limitation was illustrated by using data sets from various fields, i.e., stockmarket and random walk data sets.

Keogh and Lin [15] demonstrated a tentative solution, claiming that non-overlapping sliding windows, with their positions incremented by exactly the periodic length, would produce valid clusters when applied to a periodic time series. Since remote sensing time series data has a very strong periodic component due to seasonal vegetation dynamics, the extracted sequential time series could potentially be processed to yield usable features. A feature extraction method is presented in Section II-D that will extract features from a time series with a sliding window that expands on Keogh and Lin's approach.

Manuscript received November 06, 2009; revised January 01, 2010; accepted May 26, 2010. This work was supported by the CSIR Strategic Research Panel.

B. P. Salmon and W. Kleynhans are with the Department of Electrical, Electronic and Computer Engineering, University of Pretoria and also with the Remote Sensing Research Unit, Meraka Institute, CSIR, Pretoria, South Africa (e-mail: bsalmon@csir.co.za).

J. C. Olivier is with the Department of Electrical, Electronic and Computer Engineering, University of Pretoria and also with the Defense, Peace, Safety and Security Unit, CSIR, Pretoria, South Africa.

K. J. Wessels, F. van den Bergh, and K. C. Steenkamp are with the Remote Sensing Research Unit, Meraka Institute, CSIR, Pretoria, South Africa.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2010.2053918

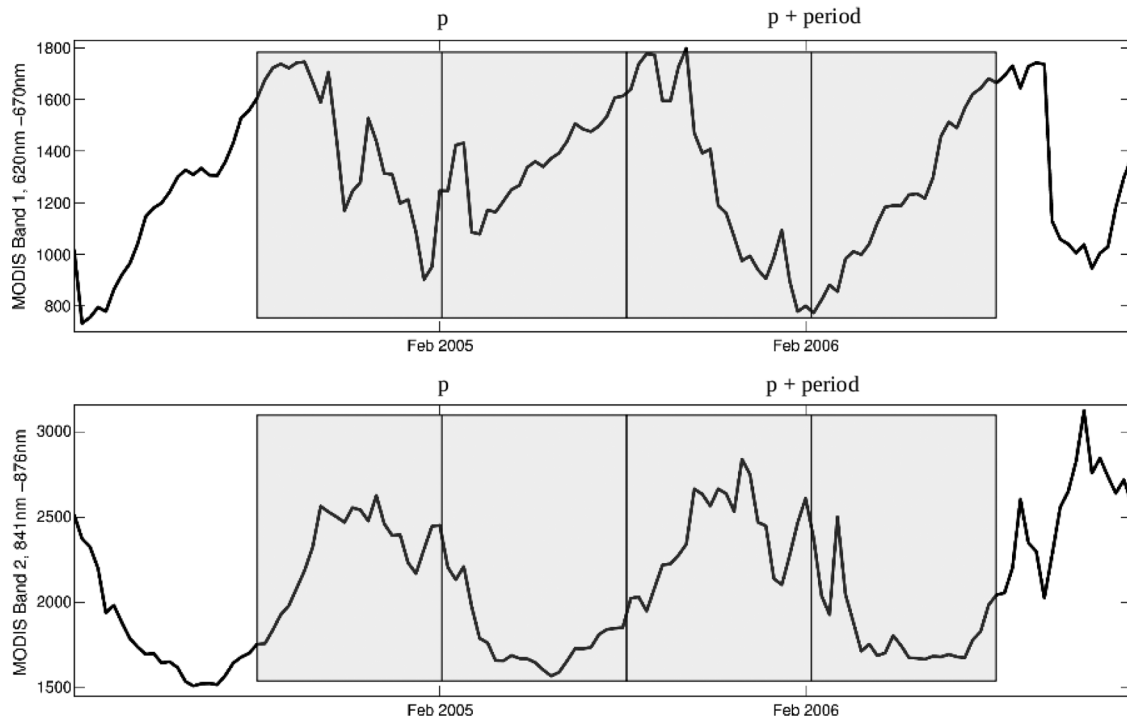


Fig. 1. Subsequence extraction through the use of a sliding window over the two spectral MODIS bands by incrementing by exactly a periodic cycle.

This feature extraction method will reduce the feature space's dimensionality and remove the restriction on the sliding window position  $p$  that will enable effective subsequence clustering that does not suffer from the afore-mentioned limitations, and potentially provide the basis for a land cover change detection method. Land cover change is defined here the transition of subsequences of a pixel's time series from one cluster to another cluster, after which it remains assigned to the second cluster for the remainder of the time series [13].

There are two general approaches to classification that can be applied to time series data, namely supervised and unsupervised [19], [20]. The supervised approach requires initial training on labelled pixels according to their land cover type. The disadvantage of using a supervised approach to perform change detection is the dependency on periodic high resolution imagery for updating the unchanged training sets over time. The supervised approach must also be robust to errors occurring within the training sets [21]. The unsupervised approach does not require any training and detects change in the inherent properties of the signal. The supervised approach can provide "from what, to what" information on land cover change [7], [13], while the unsupervised approach simply provides a "change alarm" [22] to highlight areas of change for further investigation using e.g., high resolution satellite data and field inspections. Generating training data at global and regional scales is a very labour-intensive and costly endeavour [23], which makes an unsupervised approach to automated land cover change detection a more attractive option.

The objective of this paper is to introduce the concept of unsupervised land cover change detection algorithm that operates on a temporal sliding window of MODIS time series data that uses a feature extraction method that does not suffer from the

limitation shown by Keogh and Lin [15]. Three well-known unsupervised clustering techniques were used within a land cover change detection algorithm and were evaluated specifically on new settlement development. The land cover change detection algorithm was tested on real and simulated land cover change using the 8-day composite MODIS land surface reflectance data product. The performance of the three unsupervised clustering techniques were measured against a supervised multilayer perceptron (MLP) that was used to provide an empirical upper limit to the performance [24]. The two-layer MLP network using sigmoidal activation functions was chosen as it can closely approximate any decision boundary in any feature space when enough hidden nodes are present [24].

The paper is organized as follows. Section II presents the methodology used, while Section II-D discusses the feature extraction approach. Section II-E gives a brief overview of the clustering algorithm used for the unsupervised change detection and Section III presents the results for the automated change detection on real and simulated land cover change. Section IV presents the conclusions.

## II. METHODOLOGY

### A. Study Areas

The area of interest was the Limpopo province which is situated in the northern part of South Africa. The province is still largely covered by natural vegetation used as grazing for cattle and wildlife. The development of settlements is one of the most pervasive forms of land cover change in South Africa. Areas within the province were selected where settlements and natural vegetation occur in close proximity to ensure that the rainfall, soil type and local climate were similar over both land cover

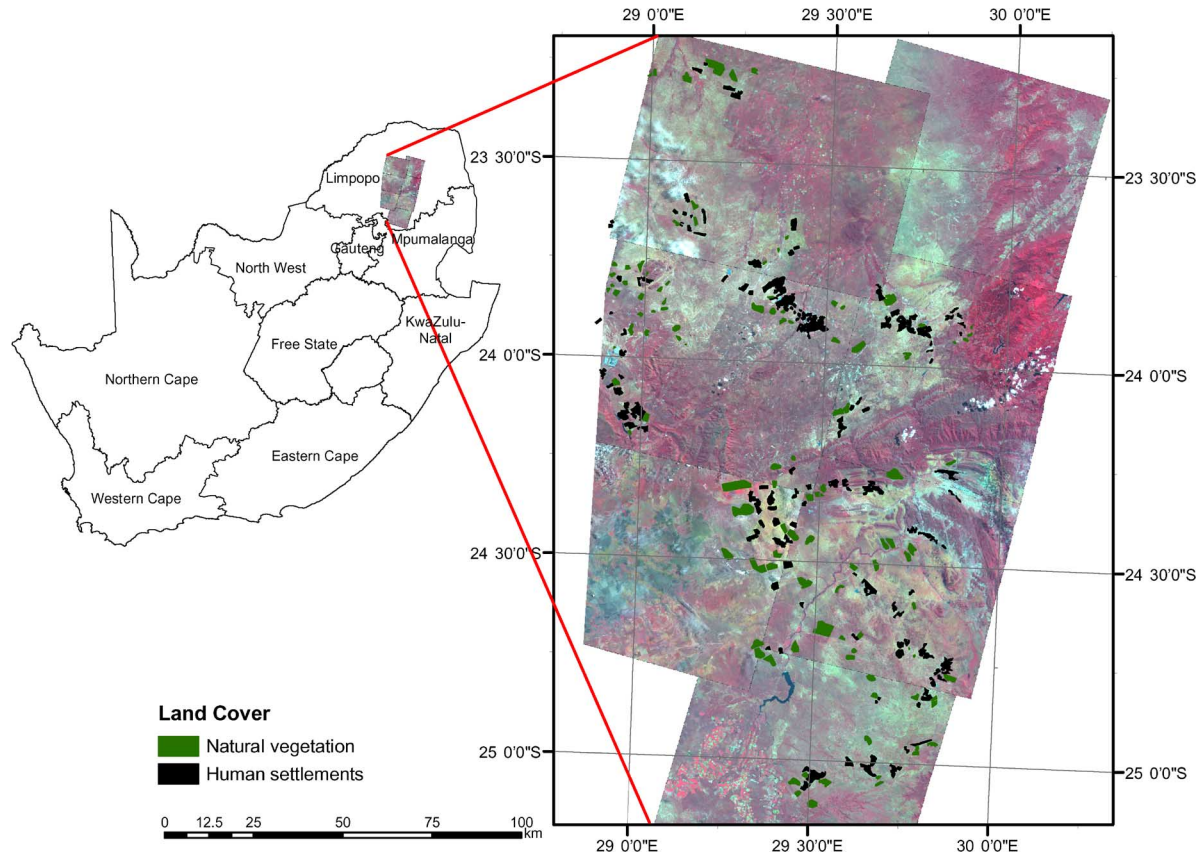


Fig. 2. Location of the study area in the Limpopo province, South Africa with land cover types polygons overlaid using Albers projection on SPOT5 RGB 321 imagery that was acquired between March 2006 and May 2006. The SPOT2 images were taken of the same area in May 2000.

TABLE I  
NUMBER OF PIXELS PER LAND COVER TYPE USED  
FOR VALIDATION AND TESTING DATA SETS

Study Area	Time Series Data Set	Simulated Change Time Series
Vegetation	1235	500
Settlement	997	500

types. The selected areas of interest are composed of 433.75 km<sup>2</sup> of natural vegetation and 374.25 km<sup>2</sup> of human settlements that are distributed throughout the study area (Fig. 2). The total number of time series (pixels) available for each class is given in Table I and were evaluated over the time period of February 2000 to February 2008.

### B. MODIS Time Series Data

The 500-meter MODIS MCD43A4 land surface reflectance product was used because it offers nadir and bidirectional reflectance distribution function (BRDF) adjusted spectral reflectance bands. This significantly reduces the anisotropic scattering effects of surfaces under different illumination and observation conditions [25], [26]. Initial tests on the uncorrected 250 meter surface reflectance data (MOD09) were not successful due to the afore-mentioned BRDF effects. The MCD43A4 product is produced by acquiring 16 samples from each of the two MODIS sensors (one on the Aqua satellite, one on the Terra satellite), that are processed to yield one output

value every 8 days for each spectral band. For each pixel a time series was extracted from only the first two spectral bands of the 8-day composite MODIS MCD43A4 data set (tile H20V11) (year 2000–2008) as these were shown to have considerable class separation when the features are analyzed [27]. The quality flags in the MODIS data were used to identify cases where quality was low due to persistent cloud cover (or other atmospheric factors) over the 8 day period of data collection [25], [26]; these samples were replaced with interpolants obtained using a cubic spline fitted through temporal neighbours.

### C. Data Sets: Validation, Simulated and Real Land Cover Change

The unsupervised clustering methods' generalization accuracy was assessed on a validation set. This validation set is composed of time series that were extracted from the MCD43A4 product. The time series were selected using visual interpretation of SPOT2 images in the year 2000 and SPOT5 images in the year 2006 to map areas of change and no change in land cover type during the study period. Through the manual interpretation of these SPOT images, new settlement developments were discovered within the Limpopo province around the known settlements. These new settlements had to be build over the course of the 6 years after the natural vegetation has been removed. The total area of newly formed settlements amounted to 5.25 km<sup>2</sup>. The total number of time series (pixels) available for each class in the validation set is given in Table I. Information on known

TABLE II  
MATCHING MATRIX USED FOR LAND COVER CHANGE DETECTION

	Change simulated Subsets 1 and 2	No change simulated Subsets 3 and 4
Detected change	True Positive	False Positive
Detected no change	False Negative	True Negative

land cover change is generally very limited [28], thus the land cover change was also simulated.

Land cover change was simulated by concatenating a set of time series from the natural vegetation class to another set of time series from the settlement class and vice versa. This made it possible to control both the type, rate and timing of change in order to quantitatively evaluate the change detection methods. As a control, testing sets containing no land cover change were also created by concatenating the same land cover type time series to each other. Hence, there were four testing data subsets based on concatenating time series of different combination of time series:

- subset 1: natural vegetation time series (class 1) spliced to settlement time series (class 2);
- subset 2: settlement time series (class 2) spliced to natural vegetation time series (class 1);
- subset 3: settlement time series (class 2) spliced to another settlement time series (class 2);
- subset 4: natural vegetation time series (class 1) spliced to another natural vegetation time series (class 1).

These four subsets were used to produce a matching matrix (Table II) to test if the unsupervised methods can detect change reliably in an automated fashion on subsets 1 and 2, while not falsely detecting change for subsets 3 and 4. The number of simulated land cover change time series available for the analysis process is also given in Table I. The concatenation of two time series produced an abrupt change in the time series, which does not necessarily represent the reality of human-induced change such as a new settlement, which may take several months to develop. Initial experiments were conducted where the signals of the two different classes were linearly blended over a time period of 12 to 24 months. These experiments revealed that the blending period merely translated into an extended period of classification uncertainty without adding any more depth to the analysis [29]. Therefore, only abrupt change was considered here.

#### D. Feature Extraction—Subsequence Time Series

In this section a method is shown that will create usable features from time series  $x_p(t)$  extracted from MODIS data. The fixed acquisition rate of the MODIS product and the seasonality of the vegetation in the study area makes for an annual periodic signal  $x(t)$  that has a phase offset that is correlated with rainfall seasonality and vegetation phenology. The Fast Fourier Transform (FFT) [30] of  $x_p(t)$  was computed, which decomposes the time sequence's values into components of different frequencies with phase offsets. This is often referred to as the frequency (Fourier) spectrum of the time series. Because the time series  $x_p(t)$  is annually periodic, this would translate into frequency

components in the frequency spectrum that have fixed positions. This can be viewed as a fixed location for each of the classes for the clustering algorithm in the feature space regardless of the sliding window position in time, which overcomes the main disadvantage to a sliding window [15]. Because of the seasonal attribute typically associated with MODIS time series and the slow temporal variation relative to the acquisition interval [31], the first few FFT components dominate the frequency spectrum. This reduces the number of features needed to represent the feature space and thus reduced the dimensionality, making clustering a feasible option [32].

Another limitation was that the sliding window position  $p$  had to be shifted by exactly a periodic cycle [15]. This limitation was addressed by computing the magnitude of all the FFT components, which removed all the phase offsets. This made it possible to compensate for both the restrictive position  $p$  of the sliding window and the rainfall seasonality. This means that  $p$ , which is the position of the sliding window, does not only have to be incremented by a fixed annual period, but can be incremented by any natural number  $\mathbb{N}$ . The features  $X_p(f)$  for the clustering method were extracted from the sliding window  $x_p(t)$  by the methodology discussed above as

$$X_p(f) = |\mathcal{F}(x_p(t))| \quad (2)$$

where  $\mathcal{F}(\cdot)$  represents the Fourier transform. From the discussion above, a sliding window of any length can be applied to the MODIS time series and moved along the time axis at any rate as long as the feature extraction rule in (2) is applied. Fig. 3 illustrates how the features that are extracted using two different sliding window positions in time maintain their position in the feature space, even though the two sliding windows are arbitrarily positioned in time. The mean and annual FFT components from (2) were considered as it was shown by Lhermitte [27] that considerable class separation can be achieved from these components. Many FFT based classification and segmentation methods consequently only consider a few FFT components [27], [33], [34]. The sliding window length was fixed at  $Q = 46$  samples to correspond to the length of the annual cycle, thereby minimizing the spectral smear and increasing the power in each feature.

#### E. Unsupervised Change Detection

The clustering method was required to process subsequences of time series data and detect land cover change as a function of time. Land cover change is declared when consecutive subsequences that are extracted from one MODIS time series transitions from one cluster to another cluster and remains in the newly assigned cluster for the rest of the time series. The temporal sliding window was designed to operate on a subsequence of the time series to extract information from two spectral bands from the MODIS product (Fig. 4). These extracted features were analyzed with three different clustering techniques: Ward,  $K$ -means and Expectation-Maximization (EM) [24].

Clustering techniques are broadly divided into hierarchical and partitional clustering approaches [15]. The Ward clustering algorithm is an agglomerative hierarchical clustering method

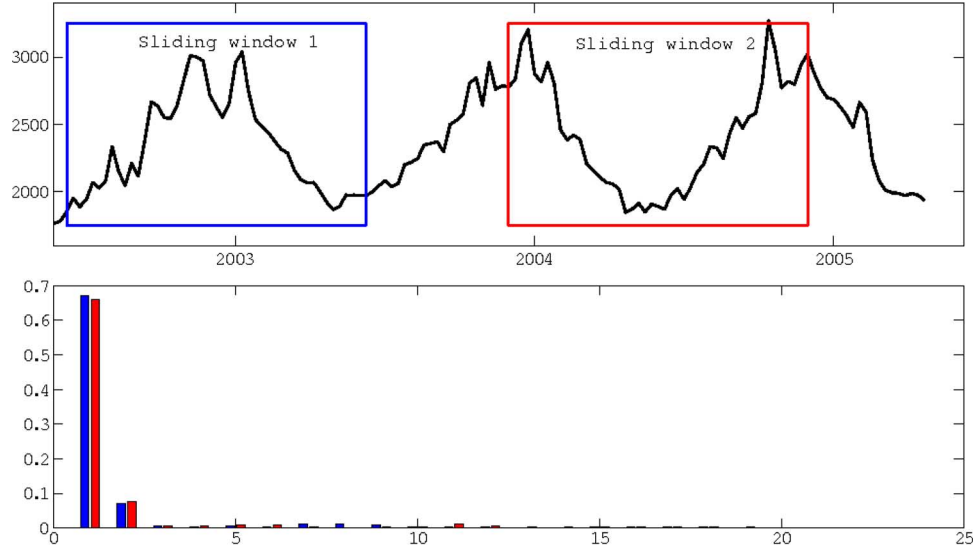


Fig. 3. Feature components  $X_p(f)$  extracted from two sliding windows at random positions using (2).

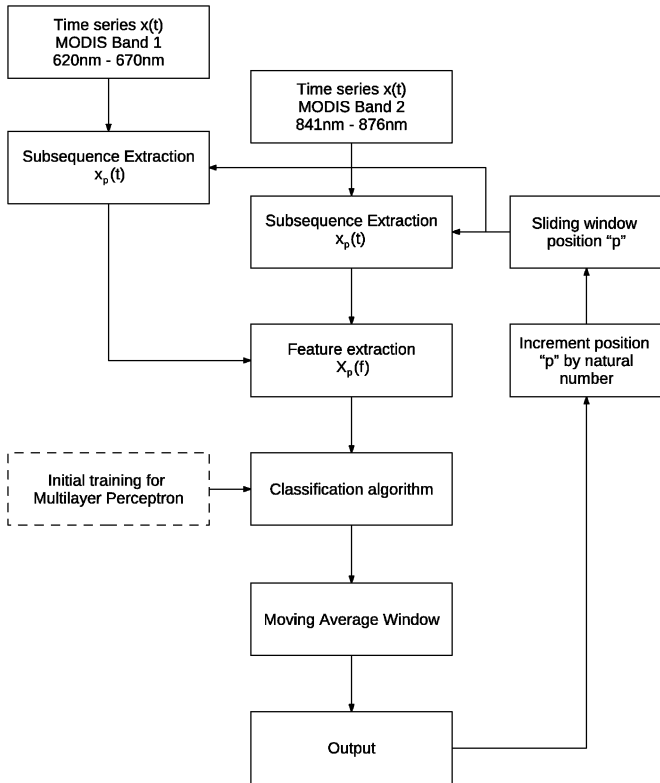


Fig. 4. Subsequences of the time series extracted from the two spectral MODIS bands that are processed for clustering and change detection.

that produces a nested hierarchy of clusters of discrete objects according to some kind of proximity matrix (similar or dissimilar distance matrices) [35]. A summary of the Ward clustering algorithm is given in Table III. The Ward clustering method [36] was used because it provided the highest cophenetic correlation coefficient (minimum loss from original information) when compared to minimum, maximum and average link clustering [37].

TABLE III

AN OUTLINE OF THE WARD HIERARCHICAL CLUSTERING ALGORITHM

Algorithm - Ward hierarchical clustering
1. Compute a complete proximity matrix of all the discrete objects.
2. Find two most similar clusters / objects that minimizes the increase in the error sum of squares.
3. Join clusters / objects to produce a new cluster.
4. Recompute proximity matrix of new cluster and all other clusters.
5. Repeat steps 2 to 4 until only one cluster exist.

TABLE IV

AN OUTLINE OF THE  $K$ -MEANS PARTITIONAL CLUSTERING ALGORITHM

Algorithm - $K$ -means partitional clustering
1. Choose the number of clusters $K$ .
2. Initialize the $K$ cluster centroids (randomly)
3. Assign cluster membership of each discrete object to nearest cluster centroid.
4. Recompute the $K$ centroids with new cluster assignments.
5. Repeat steps 3 to 4 until none of the discrete object's memberships changes cluster.

The second approach to clustering is partitional clustering, a family which includes the  $K$ -means and the EM algorithm [38]. These partitional clustering techniques are usually used as a benchmark for other algorithms, and have been used in many other fields [37].

The partitional clustering method creates an unnested partitioning of the data points with  $K$  clusters. A silhouette graph [39] was used to determine the optimal number of clusters for partitional clustering and resulted in two clusters being the best choice.  $K$ -means is a heuristic, hill climbing algorithm and can be viewed as a gradient descent approach which minimizes the sum of squared error of each feature point from the nearest cluster centroid in the feature space (Table IV) [40].



TABLE VI  
AVERAGE CLASSIFICATION ACCURACY OF THE VALIDATION SET FOR EACH OF THE CLUSTERING METHODS PRESENTED, WITH STANDARD DEVIATION IN PARENTHESIS

Clustering method	Hierarchical clustering Ward algorithm	Partitional clustering $K$ -means algorithm	Partitional clustering Expectation Maximization	Supervised clustering Multilayer Perceptron
Vegetation validation	80.70% (14.29)	81.30% (3.65)	81.03% (4.06)	84.28% (3.37)
Settlement validation	77.47% (10.19)	81.17% (2.76)	81.43% (2.55)	86.08% (2.47)
Overall validation	79.20%	81.24%	81.22%	85.11%

TABLE VII  
MATCHING MATRIX REPRESENTING THE LAND COVER CHANGE DETECTION ACCURACY ON THE SIMULATED LAND COVER CHANGE DATA SET FOR ALL THE CLUSTERING METHODS ON THE STUDY AREA

Clustering method	Hierarchical clustering Ward algorithm	Partitional clustering $K$ -means algorithm	Partitional clustering Expectation Maximization	Supervised clustering Multilayer Perceptron
True positive	71.49%	70.52%	70.99%	82.70%
True negative	75.81%	75.20%	75.76%	83.65%

TABLE V  
AN OUTLINE OF THE EM PARTITIONAL CLUSTERING ALGORITHM

Algorithm - EM partitional clustering	
1.	Choose the number of clusters $K$ .
2.	Initialize the $K$ models parameters
3.	E-step: Construct a local lower bound to the posterior distribution
4.	M-step: Optimizes the bound found in step 3.
5.	Repeat steps 3 to 4 until algorithm converges to local maximum.

The EM algorithm attempts to fit  $K$  Gaussian Mixture models to the features that would produce the highest a-posterior probability to all features (Table V). It was observed that all three clustering techniques produced minor oscillations in the cluster assignments of consecutive subsequences in areas of high cluster membership uncertainty. To smooth out all transitory oscillations in the clustering labels, a moving average window of length 3 was applied to the three clustering method's output.

An overfitted MLP was used to provide an upper bound on the performance that can be expected from the given features. The MLP comprises an input layer, one hidden layer and an output layer. All hidden and output layers used a tangent sigmoid activation function in each node. The weights in the training phase of the MLP were determined using a steepest descent gradient optimization method, with gradients estimated using backpropagation [24]. The MLP architecture was optimized at each time increment in the sliding window and a moving average window length of 3 was applied to the MLP outputs to smooth out all transitory oscillations in all classifications.

### III. EXPERIMENTAL RESULTS

#### A. Clustering Accuracy—No Change Validation Set

The clustering algorithms were tested on all the *no change* time series in the validation set; the experimental accuracies are reported in Table VI. Each entry in Table VI lists the average clustering accuracy calculated over 48 independent experiments (standard deviation in parentheses) using cross validation [41].

The  $K$ -means outperformed the Ward clustering in overall clustering accuracy by 2.04% (Table VI). The more significant result is the low standard deviation obtained by the  $K$ -means algorithm to cluster the time series data. An EM algorithm was used to fit two Gaussian mixture models over all the features in the feature space and produced results comparable to the  $K$ -means algorithm. The MLP had a average classification accuracy of 85.11% and thus performed better than the unsupervised techniques by 3.87%, when using the same features.

#### B. Change Detection—Simulated Land Cover Change

In Section II-C four testing data subsets were introduced which correspond to four possible outcomes of the land cover change detection analysis (Table II). Only the *true positive* and *true negative* cases are reported, as the other two cases are simply the complement. The outcome of the change detection simulations is summarised in the matching matrix shown in Table VII. The land cover change detection accuracy differs by less than 1% between the different clustering algorithms (Table VII). The  $K$ -means was considered the better option, due to the lower standard deviation reported in the average clustering accuracy in the *no change* time series (Table VI). The MLP had a better performance on the *true positive* by 11.21% and 7.84% on the *true negatives* than the three unsupervised methods.

#### C. Change Detection—Real Land Cover Change

Fig. 5 illustrates SPOT images of real land cover change from natural vegetation (2 May 2000) to a new human settlement (10 May 2006) in the Limpopo province. This shows an example of a new settlement that had been established in the last six years. All the clustering algorithms were tested on all the known new settlements developed on previously natural vegetated areas in the Limpopo province (Table VIII). Even though the accuracy of 76.12% reported in Table VIII were exactly the same for all the unsupervised clustering techniques, the EM algorithm detected land cover change in different areas to the  $K$ -means algorithm. The Ward clustering detected change on the same time series as the EM algorithm, but the Ward clustering provided transitory oscillations in all its *false negative* reports. This could be due to

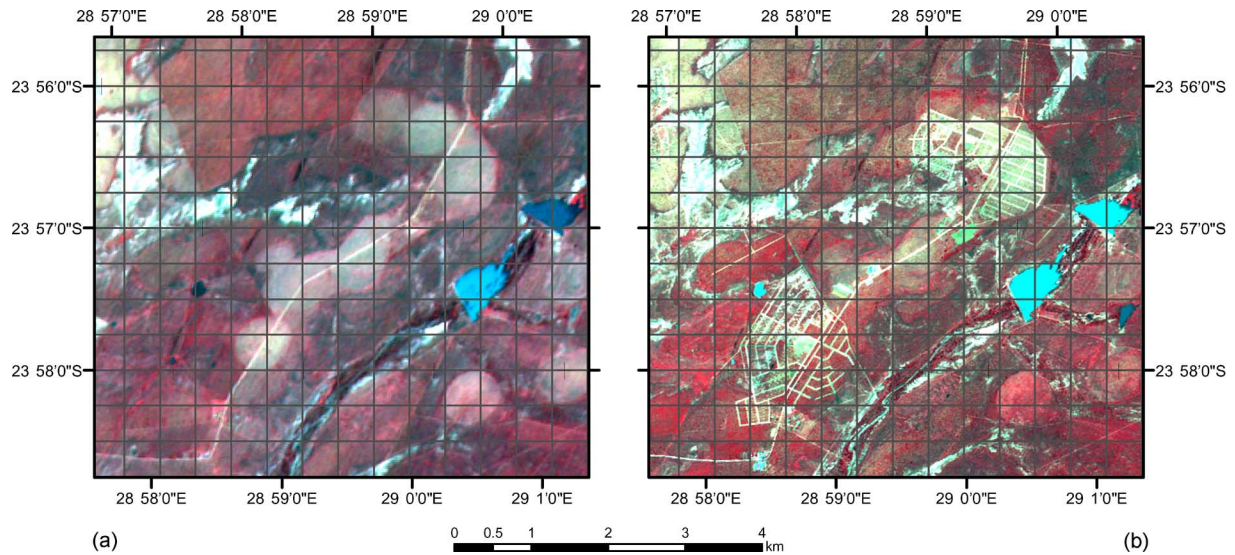


Fig. 5. SPOT2 image taken on 2 May 2000 of natural vegetation area in the Limpopo province (a) and a SPOT5 image taken on 10 May 2006 of a new human settlement called Sekuruwe (28.94 E, 23.94 S) in (b).

TABLE VIII  
LAND COVER CHANGE DETECTION ACCURACY FOR EACH OF THE CLUSTERING METHODS ON THE 5.25 km<sup>2</sup> AREA OF REAL NEW HUMAN SETTLEMENT DEVELOPMENT

Clustering method	Hierarchical clustering Ward algorithm	Partitional clustering <i>K</i> -means algorithm	Partitional clustering Expectation Maximization	Supervised clustering Multilayer Perceptron
Change detection accuracy	76.12%	76.12%	76.12%	80.95%

the high standard deviation observed in the average clustering accuracy (Table VI).

#### IV. CONCLUSIONS

In this paper, a method for unsupervised land cover change detection incorporating a temporal sliding window, operating on MODIS time series data was demonstrated. The unsupervised approaches reported *true positive* measurements of higher than 70.5% on all simulated land cover change using cross validation [41]. The results for the detection of simulated land cover change was compared to real mapped settlement development and a *true positive* accuracy of 76.12% was achieved.

The difference in change detection accuracy between the real and simulated land cover change was still acceptably small in these experiments, even though only a limited number of real land conversion examples were available. The average classification accuracies of the unsupervised approaches were similar to that of a supervised MLP (Table VI). The supervised training of the MLP however did ensure a strict boundary within the feature space, which allowed better change detection accuracy (Table VII).

Since the MODIS time series has a very strong periodic component due to seasonal vegetation growth, it provides the remote sensing community with a special type of data which, if processed correctly, is immune to the limitation pointed out by Keogh and Lin [15]. This is mainly due to the extraction process which produced a short term FFT that fixed the feature positions, which allows the features to be analyzed and permits the temporal sliding window to be moved in any time increment.

This feature extraction method will enable the analysis of meaningful sequential subsequence extraction that will incorporate the hyper-temporal properties that is provided by the MODIS product for the use of land cover change detection, which aids in providing another dimension for most change detection studies [7]–[13]. This should rekindle the remote sensing community's quest for automated change detection using time series as it allows for the application of different types of algorithms and methodologies to the sequential subsequences that were extracted from satellite data time series. This is especially relevant as emissions and other impacts of land cover change is expected to have a very large impact on climate change [42].

#### ACKNOWLEDGMENT

The authors would like to thank Willem Marais of the Remote Sensing Research Unit (RSRU) at the CSIR, for his many comments and inputs. Alex Fortesque and Naledzani Mudau of CSIR, Satellite Application Centre (SAC) for providing the data on settlements.

#### REFERENCES

- [1] R. S. DeFries, L. Bounoua, and G. J. Collatz, "Human modification of the landscape and surface climate in the next fifty years," *Global Change Biology*, vol. 8, no. 5, pp. 438–458, May 2002.
- [2] J. A. Foley *et al.*, "Global consequences of land use," *Science*, vol. 309, no. 5734, pp. 570–574, Jul. 2005.
- [3] G. Gutman, A. Janetos, C. Justice, E. Moran, J. Mustard, R. Rindfuss, D. Skole, and B. Turner, *Land Change Science: Observing, Monitoring, and Understanding Trajectories of Change on the Earth's Surface*. New York: Springer, 2004.

- [4] J. R. G. Townshend and C. O. Justice, "Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations," *Int. J. Remote Sens.*, vol. 9, no. 2, pp. 187–236, Feb. 1988.
- [5] T. Loveland *et al.*, "A strategy for estimating the rates of recent united states land-cover changes," *Photogramm. Eng. Remote Sens.*, vol. 68, no. 10, pp. 1091–1099, Oct. 2002.
- [6] M. C. Hansen and R. S. DeFries, "Detecting long-term global forest change using continuous fields of tree-cover maps from 8-km Advanced Very High Resolution Radiometer (AVHRR) data for the years 1982–99," *Ecosystems*, vol. 7, no. 7, pp. 695–716, Nov. 2004.
- [7] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2407, Jun. 2004.
- [8] S. Gopal, C. E. Woodcock, and A. H. Strahler, "Fuzzy neural network classification of global land cover from a 1 degree AVHRR data set," *Remote Sens. Environ.*, vol. 67, no. 2, pp. 230–243, 1999.
- [9] G. A. Carpenter *et al.*, "A neural network method for efficient vegetation mapping," *Remote Sens. Environ.*, vol. 70, no. 3, pp. 326–338, Dec. 1999.
- [10] H. Nemmour and Y. Chibani, "Neural network combination by fuzzy integral for robust change detection in remotely sensed imagery," *EURASIP J. Appl. Signal Process.*, vol. 2005, pp. 2187–2195, Jan. 2005.
- [11] T. Westra and R. R. D. Wulf, "Monitoring Sahelian floodplains using Fourier analysis of MODIS time-series data and artificial neural networks," *Int. J. Remote Sens.*, vol. 28, no. 7, pp. 1595–1610, Apr. 2007.
- [12] B. H. Braswell, S. C. Hagen, S. E. Frolking, and W. A. Salas, "A multivariable approach for mapping sub-pixel land cover distributions using MISR and MODIS: Application in the Brazilian Amazon region," *Remote Sens. Environ.*, vol. 87, no. 2–3, pp. 243–256, Oct. 2003.
- [13] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: A review," *Int. J. Remote Sens.*, vol. 25, no. 9, pp. 1565–1596, May 2004.
- [14] M. Natrella, *Engineering Statistics Handbook. NIST/SEMATECH e-Handbook of Statistical Methods*. Washington DC: NIST, 1963 [Online]. Available: <http://www.itl.nist.gov/div898/handbook/>, [2009/06/25]
- [15] E. Keogh and J. Lin, "Clustering of time-series subsequences is meaningless: Implications for previous and future research," *Knowledge Inf. Syst.*, vol. 8, no. 2, pp. 154–177, Aug. 2005.
- [16] J. Vermaak and E. C. Botha, "Recurrent neural networks for short-term load forecasting," *IEEE Trans. Power Syst.*, vol. 13, no. 1, pp. 126–132, Feb. 1998.
- [17] X. Wang, L. Xiu-Xia, and J. Sun, "A new approach of neural networks to time-varying database classification," in *Proc. Machine Learning and Cybernetics 2005*, Guangzhou, China, Aug. 18–21, 2005, vol. 4, pp. 2050–2054.
- [18] J. Roddick and M. Spiliopoulou, "A survey of temporal knowledge discovery paradigms and methods," *IEEE Trans. Knowledge Data Eng.*, vol. 14, no. 4, pp. 750–767, Aug. 2002.
- [19] L. Bruzzone and S. Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 858–867, Jul. 1997.
- [20] A. Singh, "Digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, Jun. 1989.
- [21] R. S. DeFries and J. C. W. Chan, "Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data," *Remote Sens. Environ.*, vol. 74, no. 3, pp. 503–515, Dec. 2000.
- [22] X. Zhan, R. Sohlberg, J. Townshend, C. DiMiceli, M. Carroll, J. Eastman, M. Hansen, and R. DeFries, "Detection of land cover changes using MODIS 250 m data," *Remote Sens. Environ.*, vol. 83, no. 1, pp. 336–350, Nov. 2002.
- [23] M. Hansen, R. DeFries, J. Townshend, and R. Sohlberg, "Global land cover classification at 1km resolution using a decision tree classifier," *Int. J. Remote Sens.*, vol. 21, no. 6, pp. 1331–1365, Apr. 2000.
- [24] C. Bishop, *Neural Networks for Pattern Recognition*, 1st ed. New York: Oxford Univ. Press, 1996.
- [25] W. Wanner *et al.*, "Global retrieval of bidirectional reflectance and Albedo over land from EOS MODIS and MISR data: Theory and algorithm," *J. Geophys. Res.*, vol. 102, no. D14, pp. 17 143–17 162, Jul. 1997.
- [26] C. Schaaf *et al.*, "First operational BRDF, albedo nadir reflectance products from MODIS," *J. Remote Sens. Environ.*, vol. 83, no. 1–2, pp. 135–148, Nov. 2002.
- [27] S. Lhermitte *et al.*, "Hierarchical image segmentation based on similarity of NDVI time series," *Remote Sens. Environ.*, vol. 112, no. 2, pp. 506–521, Feb. 2008.
- [28] R. S. Lunetta *et al.*, "Land-cover change detection using multi-temporal MODIS NDVI data," *Remote Sens. Environ.*, vol. 105, no. 2, pp. 142–154, Nov. 2006.
- [29] B. Salmon, J. Olivier, W. Kleynhans, K. Wessels, and F. van den Bergh, "The quest for automated land cover change detection using satellite time series data," in *IEEE Int. Geoscience and Remote Sensing Symp.*, Cape Town, South Africa, Jul. 12–17, 2009.
- [30] A. Oppenheim, R. Schaffer, and J. Buck, *Discrete-Time Signal Processing*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, Signal Processing series, 1999.
- [31] R. S. Lunetta, D. Johnson, J. Lyon, and J. Crotwell, "Impacts of imagery temporal frequency on land-cover change detection monitoring," *Remote Sens. Environ.*, vol. 89, no. 4, pp. 444–454, Feb. 2004.
- [32] R. Bellman, *Adaptive Control Processes: A Guided Tour*. Princeton, NJ: Princeton Univ. Press, 1961.
- [33] M. Jakubauskas, D. Legates, and J. Kastens, "Crop identification using harmonic analysis of the time-series AVHRR NDVI data," *Computers and Electronics in Agriculture*, vol. 37, no. 1–3, pp. 127–139, Nov. 2002.
- [34] R. Juarez and W. Liu, "FFT analysis on NDVI annual cycle and climatic regionality in northeast Brazil," *Int. J. Climatol.*, vol. 21, no. 14, pp. 1803–1820, Dec. 2001.
- [35] A. Jain and R. Dubes, *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice Hall, 1988.
- [36] J. Ward, "Hierarchical grouping to optimize an objective function," *J. Amer. Statist. Assoc.*, vol. 58, no. 301, pp. 236–244, Mar. 1963.
- [37] A. Jain, M. Murty, and P. Flynn, "Data clustering: A review," *ACM Comput. Surveys*, vol. 31, no. 3, pp. 264–323, Sep. 1999.
- [38] T. Mitchell, *Machine Learning*. New York: McGraw Hill, 1997.
- [39] L. Kaufman and P. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, 9th ed. Hoboken, NJ: Wiley-Interscience, 1990.
- [40] M. R. Anderberg, *Cluster Analysis for Applications*, ser. Monographs and Textbooks on Probability and Mathematical Statistics. New York: Academic Press, 1973.
- [41] S. Salzberg, "On comparing classifiers: Pitfalls to avoid and a recommended approach," *Data Mining and Knowledge Discovery*, vol. 1, no. 3, pp. 317–327, Sep. 1997.
- [42] F. Achard, H. Eva, P. Mayaux, H. Stibig, and A. Belward, "Improved estimates of net carbon emissions from land cover change in the tropics for the 1990s," *Global Biogeochemical Cycles*, vol. 18, p. GB2008+, May 2004.



**Brian P. Salmon** received the B.Eng. degree in computer engineering and the M.Eng. degree in electronic engineering (signal processing) from the University of Pretoria, Pretoria, South Africa, in 2004 and 2008, respectively.

He is currently with the Remote Sensing Research Unit at the Council for Scientific and Industrial Research. He is working towards the Ph.D. in electronic engineering and his research interests are machine learning and graph theory.



**Jan Corne Olivier** received the Ph.D. degree in electrical engineering from the University of Pretoria, South Africa, in 1990.

He is currently a Chief Researcher at the CSIR in Pretoria, South Africa, and an exceptional Professor at the University of Pretoria's Department of Electrical, Electronic and Computer Engineering. He was with Bell Northern Research (BNR) in Canada, and with Nokia Research Center in the United States prior to returning to South Africa in 2003. His research interests are in Estimation and Detection theory, as well as applications of Machine Learning.

Dr. Olivier serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.





**Konrad J. Wessels** received the M.Sc. degree in landscape ecology and conservation planning from the University of Pretoria, South Africa, in 1997 and the Ph.D. in geography from the University of Maryland, Baltimore, in 2005.

He was a research associate at NASA Goddard Space Flight Center, Hydrospheric and Biospheric Laboratory, in 2006. He is presently a principal researcher and leads the Remote Sensing Research Unit within the CSIR Meraka Institute in Pretoria, South Africa. His research interests include time-series analysis of satellite data for monitoring environmental change and the estimation ecosystem state variables and services with remote sensing.

ries analysis of satellite data for monitoring environmental change and the estimation ecosystem state variables and services with remote sensing.



**Waldo Kleynhans** received the B.Eng. degree in computer engineering and the M.Eng. degree in electronic engineering (OFDM channel estimation) from the University of Pretoria, South Africa, in 2004 and 2008, respectively.

He is currently with the Remote Sensing Research Unit at the Council for Scientific and Industrial Research in Pretoria, South Africa, where he is working towards the Ph.D. degree in electronic engineering. His research interests include wireless communications, statistical detection, estimation theory, and machine learning.

theory, and machine learning.



**Frans van den Bergh** received the M.Sc. degree in computer science (machine vision) and the Ph.D. degree in computer science (particle swarm optimization) from the University of Pretoria, Pretoria, South Africa, in 2000 and 2002, respectively.

He is currently a principal researcher at the Council for Scientific and Industrial Research. His research interests include automated feature extraction from high-resolution satellite images, as well as automated change detection. He maintains an active interest in particle swarm optimization and machine learning.



**Karen C. Steenkamp** received the M.Sc. degree in geography and environmental management from the University of Johannesburg, South Africa, in 1998.

She is presently a senior researcher in the Remote Sensing Research Unit within the CSIR Meraka Institute, Pretoria, South Africa. Her research interests include time-series analysis of satellite data and phenological studies of southern African vegetation. She is also active in fuel characterization for fire danger indices.