AN AUTOCORRELATION ANALYSIS APPROACH TO DETECTING LAND COVER CHANGE USING HYPER-TEMPORAL TIME-SERIES DATA

†*W. Kleynhans,, †*B.P. Salmon, †J.C. Olivier, *K.J. Wessels, *F. van den Bergh

† Electrical, Electronic and Computer Engineering University of Pretoria, South Africa

*Remote Sensing Research Unit Meraka Institute, CSIR, Pretoria, South Africa

ABSTRACT

Human settlement expansion is one of the most pervasive forms of land cover change in the Gauteng province of South Africa. A method for detecting new settlement developments in areas that are typically covered by natural vegetation using 500 m MODIS time-series satellite data is proposed. The method is a per pixel change alarm that uses the temporal autocorrelation to infer a change metric which yields a change or no-change decision after thresholding. Simulated change data was generated and used to determine a threshold during a preliminary off-line optimization phase. After optimization the method was evaluated on examples of known land cover change in the study area and experimental results indicate a 92% change detection accuracy with a 15% false alarm rate.

1. INTRODUCTION

Remote sensing satellite data provide researchers with an effective way to monitor and evaluate land cover changes. In most cases two spatially registered high resolution images acquired at two different instances are compared and based on a change metric and threshold selection method, each pixel is classified as either belonging to the change or no-change class. Such a comparison of only two images is not always reliable, as similar land cover types can appear significantly different at various stages of the natural growth seasonal cycle [1]. To mitigate this problem, the temporal frequency of medium resolution remote sensing data acquisitions should be high enough to distinguish change events from natural phenological cycles. The Moderate-resolution Imaging Spectroradiometer (MODIS) data product used in this study utilizes daily Terra and Aqua satellite overpasses to produce a 500 m resolution composite image every 8 days, and as such offer a high enough temporal frequency of the remote sensing data for change detection through time-series analysis [2].

In this paper, a semi-supervised approach is proposed. The semi-supervised nature of the method is attributed to the fact the training database requirement is limited to no-change examples which are numerous and can be obtained in large numbers as opposed to change examples that are rare at a regional scale [1]. A land cover change was then simulated using no-change examples of typical natural vegetation and settlement time-series data. Both the no-change and simulated change datasets are then used to determine a set of parameters in an off-line optimization phase after which the algorithm is run in an operational and unsupervised manner for the entire study area.

The autocorrelation function (ACF), in the temporal context, have been used selectively in remote sensing [3], but is mostly applied in the spatial context [4]. In this study the temporal ACF of a pixel's time-series was considered. An ACF of a time-series that is stationary behaves differently from an ACF of a time-series that is non-stationary due to land cover change. By determining suitable detection parameters using only a no-change database (as explained above), it will be shown that real land cover change can be detected reliably in a semi-supervised fashion.

The goal of this study was to detect new human settlement formation in the Gauteng province of South Africa using MODIS time-series data with minimal operator assistance. The new proposed method exploits the non-stationary property that is typically associated with a time-series that undergoes land cover change by using a time-series ACF.

2. DATA DESCRIPTION

2.1. Study Area

The Gauteng province is located in northern South Africa and because of a high level of urbanization it has seen significant human settlement expansion during the 2001 and 2008 period. A total area of approximately 17000 km^2 was considered being centred around $26°07'29.62''$S, $28°05'40.40''$E.

2.2. MODIS Data

The time-series for all seven MODIS land bands, as well as NDVI derived from 8 daily composite, 500 m, MCD43 BRDF-corrected, MODIS data [5] was used for the period 2001/01 to 2008/01.

2.2.1. No-Change Data

A dataset of no-change pixel time-series (n=964) consisting of natural vegetation (n=592) and settlement (n=372) pixels,

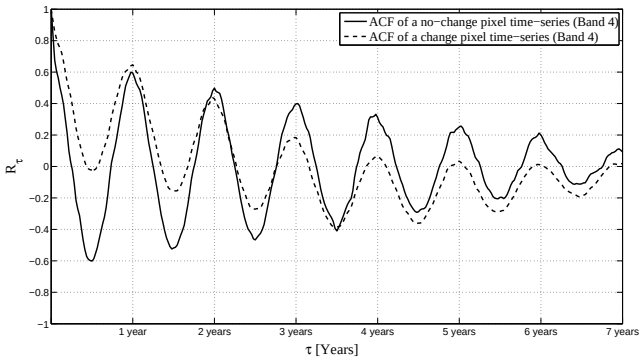

Fig. 1. Autocorrelation of a change and no-change pixel's MODIS band 4 time-series.

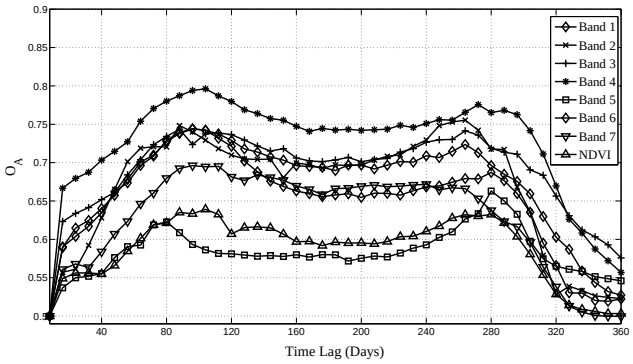

Fig. 2. Overall change detection accuracy for each band and time-lag combination for a maximum time-lag of up to 360 days.

were identified by means of visual interpretation of high resolution Landsat and SPOT images in 2000 and 2008 respectively.

2.2.2. Simulated Change Data

A simulated change dataset (n=592) was generated by linearly blending a time-series of a pixel covered by natural vegetation with that of a settlement pixel time-series. The resulting simulated change database had a uniformly spread change date between 2001/01 and 2008/01. The blending period was found not to influence the method's performance, and a representative blending period of 6 months was chosen. The simulated change data was used together with a subset of the no-change dataset (n=482) in an off-line optimization phase to determine the detection parameters (section 3).

2.2.3. Real Change Data

Examples of confirmed settlement developments during the study period were also obtained by means of visual interpretation of high resolution Landsat and SPOT images in 2000 and 2008 respectively. All settlements identified in 2008 were referenced back to 2000 and all the new settlement polygons were mapped and the corresponding MODIS pixels (n=181) were so identified. The real change pixels and remaining pixels of the no-change dataset (n=482) were used in an unsupervised operational mode to test the change detection capability of the method.

3. METHODOLOGY

3.1. Temporal ACF method

The temporal ACF method uses a two stage approach. Firstly, the simulated change dataset together with the no-change dataset (Section 2.2.2) is used in an off-line optimization phase to determine the appropriate parameters (band, lag and threshold selection). Second, the method is run in an unsupervised manner using the parameter set that was determined during the aforementioned off-line optimization phase. These two stages will be discussed in further detail in the following sections.

3.1.1. Off-line optimization phase

Assume that the time-series for any given band of MODIS is expressed as:

$$X_n^b\ ,\ n \in \{1, 2, ..., N\}\ b \in \{1, 2, ..., 8\}, \tag{1}$$

where X_n^b is the observation from spectral band b at time n and N is the number of time-series observations available. It should be noted that band 8 in (1) refers to computed NDVI. It is assumed that N is equal for all seven bands.

The normalized ACF for time-series $\mathbf{X}^b = [X_1^b, X_2^b, ..., X_N^b]$ can then be expressed as:

$$R^b(\tau) = \frac{E[(X_n^b - \mu^b)(X_{n+\tau}^b - \mu^b)]}{\mathrm{var}(\mathbf{X}^b)}, \tag{2}$$

where τ is the time-lag and E denotes the expectation. The mean of $\mathbf{X}^b$ is given as μ^b and the variance, which is used for normalization, is given as $\mathrm{var}(\mathbf{X}^b)$. Figure 1 shows the typical ACF of an actual change and no-change pixel's time-series. It is clear that the no-change pixel has a symmetrical form relative to the $R^b(\tau) = 0$ axis, whereas the change pixel shows a strong non-symmetrical property. The reason for this is the stationarity requirement of the ACF in (2). The mean and variance of the time-series of X_n^b in (2) is required to remain constant through time to determine the true ACF of the time-series. The inconsistency of the mean and variance typically associated with a change pixel's non-stationary time-series thus becomes apparent when analyzing the ACF of the

Table 1. Confusion Matrix, overall accuracy (O_A) and optimal threshold (δ^*) showing the best land cover change detection performance during the off-line optimization phase using MODIS band 4 550 nm with a lag of 96 days

	Simulated change (n=592)	No Change (n=482)	δ^*	O_A
Change Detected	75.17%	14.73%	0.16	80.22%
No Change Detected	24.83%	85.27%		

Table 2. Confusion Matrix, overall accuracy (O_A) and threshold (δ) for the case of real change detection using the MODIS band 4 (550 nm) with a lag of 96 days as determined during the off-line optimization phase

	Real change (n=181)	No Change (n=482)	δ	O_A
Change Detected	92.27%	15.35%	0.16	88.46%
No Change Detected	7.73%	84.65%		

time-series. The change metric is thus simply equivalent to the temporal correlation of a specific band (b) at a specific lag (τ)

$$R^b(\tau) = \delta^b_\tau. \quad (3)$$

It is clear, however, that the distribution of δ^b_τ in the case of change and no-change,will vary for different values of τ and b. The aim is thus to determine the value of τ and b in δ^b_τ that will result in the most separable distributions between δ^b_τ for the change and no-change case respectively. The value of the optimal threshold (δ^{b*}_τ) also needs to be determined. The aim is to determine the time-lag (τ), band/s (b) and threshold (δ) which provide the best separation between the ACF of change and no-change pixel time-series taken from the simulated change and no-change datasets respectively.

3.1.2. Operational phase

After the off-line optimization phase is complete, the resulting parameters are used to run the algorithm in an unsupervised manner for the entire area of interest. A pixel is labeled as having changed by evaluating the following,

$$\text{Change} = \begin{cases} \text{true} & \text{if } R^b(\tau) > \delta \\ \text{false} & \text{if } R^b(\tau) < \delta \end{cases}$$

where $R^b(\tau)$ is the ACF of band b evaluated at lag τ and δ is the decision threshold. The value of τ, b and δ, was provided in the the aforementioned off-line optimization phase. The results obtained for both the off-line optimization phase and operational phase are presented in section 4.

4. RESULTS

4.1. Optimal band and lag selection using a simulated change dataset

The right sided normalized ACF for band b can be expressed as $R^b(\tau) = [R^b(0), R^b(1), ..., R^b(N)]$. The task at hand is to determine the separation between the ACF of the change and no-change dataset for each band at each lag. The Bayesian decision error was calculated based on the distribution of the inferred change metric $\delta^b_\tau = [R^b_0(\tau), R^b_1(\tau), ..., R^b_\alpha(\tau)]$ for the change and no-change dataset and the overall accuracy was computed. The overall accuracy of the ACF change detection method, for each band and lag is presented in figure 2. It is evident that Band 4 (550 nm) shows the best separation between the no-change and simulated change datasets for the study area. The lag that shows the highest separability is 96 days. Table 1 shows the confusion matrix when running the algorithm on the simulated change dataset and using the optimal parameters obtained in the off-line optimization stage.

4.2. Real change detection

After the band, lag and optimal threshold selection was completed, the performance of the proposed method was validated using the test dataset described in section 2.2.3. Table 2 summarizes the performance of the method using the parameters obtained during the off-line optimization phase. For comparison, the performance of the NDVI differencing method [1] using an optimal threshold (z value) for the same dataset is shown in Table 3.

5. DISCUSSION

The performance of the false alarm rate for both the off-line optimization (14.73%) and operational phase (15.35%) is very similar with a difference of less than one percent. The change detection accuracy on the other hand for the off-line optimization (75.17%) and operational phase (92.27%) is considerably different (Tables 1 and 2). It might seem counter intuitive that the simulated change is more difficult to detect than real change examples, but this does make sense when considering the timing of the change. The mean start of change date of the real change dataset is 2004 with a standard deviation of two years. The simulated change date on the

Table 3. Confusion Matrix, overall accuracy (O_A) and threshold (z) for the case of real change detection using the NDVI differencing method [1].

	Real change (n=181)	No Change (n=964)	z	O_A
Change Detected	75.14%	23.96%	1.7	75.59%
No Change Detected	24.86%	76.04%		

Table 4. O_A performance for different start of change dates

Mean start of change	O_A
2001/06	70.67%
2002/06	83.57%
2003/06	85.33%
2004/06	85.43%
2005/06	84.92%
2006/06	81.74%
2007/06	76.66%

other hand, was distributed uniformly over the entire date range of the time-series. Therefore, when the change occurs in the center of the time-series, the non-stationarity of the time-series will be at a maximum and will decrease as the change date moves towards the beginning or end of the time-series. The performance of the simulated change detection is shown for different start years (Table 4). It is clear that the ACF change detection method is slightly compromised when change occurs in the first or last year with no significant decrease in the performance for the others years.

Combining multiple bands in this study did not significantly improve on the separability achieved using only band 4. This does not suggest that band 4 is the best for all types of land cover change. However, for our study area and land cover change case, the ACF of the band 4 time-series showed the highest separability between the no-change and simulated change datasets. Multiple band combinations could also be used to improve the separability at the cost of increased computational complexity in cases where no single band gives adequate separability.

The proposed Temporal ACF method was also compared to the NDVI differencing method [1]. The optimal threshold (z value) for this method was used for a fair comparison of the two methods and was determined iteratively by evaluating a range of possible realizations of z. The NDVI differencing method was found not to be very successful, having a change detection accuracy of 75.14% and false alarm of 23.19% for the study area.

6. CONCLUSION

In this paper, a simple but effective method was proposed as a land cover change detection alarm. The simplicity of the algorithm is achieved by using a two step approach. Firstly, in an off-line optimization phase, the time-series ACF of all seven MODIS land bands of a no-change and simulated change dataset is used to determine the band (b), lag (τ) and threshold (δ) value that shows the highest separability between the two datasets. Second, in the operational phase, the time-series ACF of band b at lag τ is computed per pixel and compared to the threshold (δ) to yield a change or no-change decision. This approach requires no significant pre-filtering, iterative annual differencing or spatial analysis. The method was effectively used to determine the location of new settlement developments in the Gauteng province of South Africa. A change detection accuracy of 92% with a 15% false alarm rate was achieved.

7. REFERENCES

[1] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," *Remote Sensing of Environment*, vol. 105, no. 2, pp. 142–154, Nov. 2006.

[2] W. Kleynhans, J. C. Olivier, K. J. Wessels, B. P. Salmon, F. van den Bergh, and K. Steenkamp, "Improving land cover class separation using an extended Kalman filter on MODIS NDVI time series data," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 2, pp. 381–385, Apr. 2010.

[3] H. You, J. Garrison, G. Heckler, and D. Smajlovic, "The autocorrelation of waveforms generated from ocean-scattered gps signals," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 1, pp. 78–82, Jan. 2006.

[4] D. Jupp, A. Strahler, and C. Woodcock, "Autocorrelation and regularization in digital images i. basit theory," *IEEE Geoscience and Remote Sensing*, vol. 26, no. 4, pp. 463–473, Jul. 1988.

[5] C. Schaaf *et al.*, "First operational BRDF, albedo nadir reflectance products from MODIS," *Remote Sensing of Environment*, vol. 83, no. 1/2, pp. 135–148, Nov. 2002.