Improving Land Cover Class Separation Using an Extended Kalman Filter on MODIS NDVI Time-Series Data

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

Abstract—It is proposed that the normalized difference vegetation index time series derived from Moderate Resolution Imaging Spectroradiometer satellite data can be modeled as a triply (mean, phase, and amplitude) modulated cosine function. Second, a nonlinear extended Kalman filter is developed to estimate the parameters of the modulated cosine function as a function of time. It is shown that the maximum separability of the parameters for natural vegetation and settlement land cover types is better than that of methods based on the fast Fourier transform using data from two study areas in South Africa.

Index Terms—Discrete Fourier transforms, Kalman filtering.

I. INTRODUCTION

LAND cover classification based on multitemporal satellite data can capitalize on seasonal variation in land surface reflectance due to vegetation phenology to provide better classification than single-date imagery [1], [2]. Multitemporal coarse-resolution satellite imagery such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Advanced Very High Resolution Radiometer has been widely used to map land cover at regional to global scales [3]–[5]. Land cover classification methods are often based on a series of secondary metrics derived from the normalized difference vegetation index (NDVI) time series and include principal component analysis [2], [6], [7], phenological metrics [8], or Fourier (spectral) analysis [9], [10]. Fourier (spectral) analysis expresses a time series as the sum of a series of cosine waves with varying frequency, amplitude, and phase [11]. The frequency of each cosine component is related to the number of completed cycles over the defined interval. The fast Fourier transform (FFT) is an effective and computationally efficient algorithm to compute the discrete Fourier transform (DFT) [11] and is often used when evaluating NDVI time-series data [10], [12]–[14]. In many applications where the FFT of NDVI time-series data is used for classification and segmentation, only the first few FFT components are considered as they tend to dominate the spectrum [10], [12], [13]. The reason for this is because of the strong seasonal component and slow variation relative to the sampling interval of the time series (eight days for MODIS). It has been found that, even when considering only the mean and seasonal FFT components [10], reliable class separation can be achieved. A drawback of using FFT-based methods is that the underlying process is assumed to be stationary. This assumption is often invalid in the case of NDVI time-series data, particularly if a land cover change is present. The extended Kalman filter (EKF) is a nonlinear estimation method that can potentially be employed to estimate unobserved parameters (process model) using noisy observations of a related measurement model. EKF techniques in remote sensing have been used for parameter estimation of values related to physical, biogeochemical processes, or vegetation dynamics models [15], [16]. In this letter, it is first proposed that the NDVI time series be modeled as a single but triply modulated cosine function, where the mean \( \mu \), the amplitude \( \alpha \), and the phase \( \phi \) are functions of time. Second, it is proposed that a nonlinear EKF be used to estimate these parameters as a function of time for each NDVI time series.

Using MODIS MOD43 data from two study areas, this letter shows that the \( \mu \), \( \alpha \), and \( \phi \) parameter streams over time are similar for the same class land cover types and dissimilar for different land cover types representing natural vegetation and settlement land cover types in the Limpopo province of Southern Africa. The parameter sequence can thus be used to determine the level of similarity between NDVI time series belonging to different land cover types.

II. DATA DESCRIPTION

A. Study Area

The Limpopo province is located in northern South Africa and is mostly covered by natural vegetation while a large number of informal settlements are rapidly expanding throughout the province. The proposed method was tested in two regions in the Limpopo province. The first study area (Region A) is centered around latitude 24°17'21.43" S and longitude...
Fig. 1. Mean NDVI time series for natural vegetation and settlement land cover in Region A.

29°39′42.96″ E and is 43 km southeast of the city of Polokwane. Region A covers a geographic area of approximately 190 km²; 42 natural vegetation and 42 settlement pixels were selected for analysis. Region B is centered around latitude 24°19′51.50″ S and longitude 29°18′04.07″ E and is 47 km southwest of the city of Polokwane. Region B covers a geographical area of 100 km²; 76 settlement and 52 natural vegetation pixels were selected. The study regions that were considered had settlements and natural vegetation areas in close proximity which ensured that the rainfall, soil type, and local climate were similar. Each of the MODIS pixels was evaluated using SPOT5 high-resolution data to ensure that none of them had experienced any land cover change during the study period.

B. MODIS Data

The NDVI time-series data were derived from eight daily composite MCD43 bidirectional-reflectance-distribution-function-corrected MODIS data with a spatial resolution of 500 m [17] for the period January 2001 to January 2008. Fig. 1 shows the mean NDVI time series for natural vegetation and settlement pixels in Region A.

III. METHODOLOGY

A. FFT Method

As discussed in Section I, the Fourier analysis of the NDVI time series has proved to be insightful because the signal can be decomposed into a series of cosine waves with varying amplitude, phase, and frequency. The DFT can be written in matrix form as

\[ Y = F_N Y \]  

where \( Y^T = [y_0 \ y_1 \ y_2 \ \cdots \ y_{N-1}] \) is the NDVI time series of length \( N \) in vector form, \( Y^T = [Y_0 \ Y_1 \ Y_2 \ \cdots \ Y_{N-1}] \) is the DFT of \( y \), and \( F_N \) is the DFT matrix in the form

\[ F_N(r, c) = \left[ \frac{1}{\sqrt{N}} e^{-\frac{2\pi i (r-1)(c-1)}{N}} \right] \]  

where \( F_N(r, c) \) is the value of row \( r \) and column \( c \) of the \( F_N \) matrix [11]. The first 30 FFT components of a typical seven-year natural vegetation NDVI time series are shown in Fig. 2. As expected, the majority of signal energy is contained in the mean and the annual component which relates to FFT component zero and seven, respectively. As proposed in [10], the similarity of any two arbitrary NDVI time series can be evaluated by computing their FFT and then comparing the first and seasonal FFT component of each FFT series. A distance metric based on the mean (\( \mu \)) and seasonal (\( \alpha \)) FFT component difference for any two FFT series can then be formulated as follows:

\[ D^\mu_{FFT} = \|Y_0^1 - Y_0^2\| \]  
\[ D^\alpha_{FFT} = \|2(Y_7^1 - Y_7^2)\| \]  

where \( D^\mu_{FFT} \) and \( D^\alpha_{FFT} \) are the Euclidean distances between the mean and annual FFT components, respectively, of two NDVI time series.

B. New EKF Method

It is proposed that an NDVI time series for a given pixel can be modeled by a triply modulated cosine function given as

\[ y_k = \mu_k + \alpha_k \cos(\omega k + \phi_k) + v_k \]  

where \( y_k \) denotes the observed value of the NDVI time series at time \( k \) and \( v_k \) is the noise sample at time \( k \). The noise is additive but with an unknown distribution. The cosine function is based on a number of parameters (that are not directly observable), namely, the frequency \( \omega \), the nonzero mean \( \mu \), the amplitude \( \alpha \), and the phase \( \phi \). The frequency can explicitly be computed as \( \omega = 2\pi f \), where \( f \) is based on the annual vegetation growth cycle. Given the eight daily composite MCD43 MODIS data, \( f \) was calculated to be 8/365. The values of \( \mu_k \) and \( \phi_k \) are functions of time and must be estimated given \( y_k \) for \( k \in 1, \ldots, N \). The estimation of these parameters is nontrivial and requires a nonlinear estimator. According to the EKF formulation, for every increment of \( k \) (the discrete
time), a state vector $x_k$ is defined containing the parameters to be estimated in the form $x_k = [\mu_k \quad \alpha_k \quad \phi_k]^T$. The relation between $x_k$ and $x_{k-1}$ is given by $v$, which is a known but possibly nonlinear function. The state vector $x_k$ is related to the observation vector $y_k$ via a nonlinear measurement function $h$. Both these models are possibly nonperfect, so the addition of process $w_k$ and measurement $v_k$ noise is used. This is expressed as [18]

$$x_k = v(x_{k-1}) + w_k$$  \hspace{1cm} (6)  

$$y_k = h(x_k) + v_k.$$  \hspace{1cm} (7)

The state vector parameters may be estimated over time $k$ by recursive iteration [18] based on the observation data $y_k$ up to time $k$. In the observation equation (7), $y_k$ is the predicted measurement. Function $h$ is used to compute a measurement given the predicted state, and $v_k$ is the observation noise vector.

The estimated values for $x_k$ = $[\mu_k \quad \alpha_k \quad \phi_k]^T$ over time $k$ effectively result in a time series for each of the three parameters. The next step was to define the metrics to measure class separability corresponding to a land cover type. As shown in [10], substantial separability can be achieved when comparing mean and annual FFT components of an NDVI time series of different land cover types. In the proposed EKF framework, the mean FFT component (i.e., FFT component 0) corresponds to $\mu$, and the annual FFT component (i.e., FFT component 7 for the seven-year NDVI time series) corresponds to $\alpha$. Hence, it is proposed that, within the EKF framework, a separability or distance metric between two NDVI time series be defined as

$$D_{\mu}^{EKF} = \max\{\mu_{k,1} - \mu_{k,2}\}, \quad 1 \leq k \leq N \quad (8)$$

$$D_{\alpha}^{EKF} = \max\{\alpha_{k,1} - \alpha_{k,2}\}, \quad 1 \leq k \leq N. \quad (9)$$

$D_{\mu}^{EKF}$ is the maximum distance between the first ($\mu_1$) and second ($\mu_2$) parameter sequence over time $k$. $D_{\alpha}^{EKF}$ is calculated in a similar manner finding the maximum distance between the annual amplitude parameter sequence.

For the present case, it was assumed that the state vector $x$ does not change significantly when time is advanced by one; hence, $v = 1$, and the process model is linear. The measurement model, however, contains the cosine term and, as such, is evaluated via the standard Jacobian formulation, thereby linearizing the nonlinear measurement model around the current state vector [18].

IV. RESULTS

Taking the FFT of two NDVI time series and comparing the Euclidean distance between their first and annual FFT components, respectively, produce a scalar quantity in each case. When using the EKF to estimate the $\mu$ and $\alpha$ parameter sequence for each NDVI time series, the difference between the parameter sequences fluctuates over the seven-year period. This is shown in Fig. 3, where the $\mu$ sequence estimated using the EKF is shown along with the FFT mean component for two typical NDVI time series belonging to each of the two classes in Region A. The $\mu$ sequence for the settlement and natural vegetation time series clearly varies in similarity (Fig. 3). This is to be expected as land cover classes tend to be more similar during certain parts of the season than others. This characteristic was exploited by only considering the maximum distance between each parameter sequence as given in (8) and (9).

To evaluate the performance of the proposed EKF algorithm, an all-to-all comparison was made between the NDVI time series of each pixel in the natural vegetation class with each pixel in the settlement class.

The initial state parameters, as well as the observation and process noise estimates, were determined offline based on known training data from the study areas. The training data were random selection of 5% of the total number of pixels per region. The initial state parameters were calculated using the FFT mean and annual components of the training data as

$$\mu_1 = \frac{1}{Z} \sum_{i=1}^{Z} \frac{Y_{i}^{0}}{Z}$$  \hspace{1cm} (10)  

$$\alpha_1 = \frac{1}{Z} \sum_{i=1}^{Z} \frac{2 \parallel Y_{i}^{7} \parallel}{Z} \quad (11)$$

$$\phi_1 = \frac{1}{Z} \sum_{i=1}^{Z} \frac{\angle Y_{i}^{7}}{Z} \quad (12)$$

where $Z$ is the total number of training time series and $Y_{i}^{n}$ is the $n$th FFT component of time series $i$. The observation noise was determined as

$$\sigma_v = \frac{1}{Z} \sum_{i=1}^{Z} \text{std}(e_i) \quad (13)$$

$$e_i = \parallel \hat{y}_i - y_i \parallel. \quad (14)$$

Here, $\sigma_v$ is the estimated standard deviation of the observation noise; $\text{std}(e_i)$ is the standard deviation of a vector containing the difference between the original time series $y_i$ and a filtered version $\hat{y}_i$ calculated as

$$\hat{y}_i = F_N^{-1} \hat{Y}_i.$$  \hspace{1cm} (15)
shown in Tables I and II, respectively. The values of process noise standard deviation for Regions A and B, are all other components were set to zero.

<table>
<thead>
<tr>
<th>Region</th>
<th>(\mu_1)</th>
<th>(\alpha_1)</th>
<th>(\phi_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.3008</td>
<td>0.0835</td>
<td>0.2700</td>
</tr>
<tr>
<td>B</td>
<td>0.3447</td>
<td>0.1185</td>
<td>0.1708</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>(\sigma_v)</th>
<th>(\sigma_\mu)</th>
<th>(\sigma_\alpha)</th>
<th>(\sigma_\phi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>(3.8 \times 10^{-2})</td>
<td>(8 \times 10^{-5})</td>
<td>(8 \times 10^{-5})</td>
<td>(1.5 \times 10^{-2})</td>
</tr>
<tr>
<td>B</td>
<td>(4.4 \times 10^{-2})</td>
<td>(9 \times 10^{-5})</td>
<td>(9 \times 10^{-5})</td>
<td>(1.7 \times 10^{-2})</td>
</tr>
</tbody>
</table>

Fig. 4. Probability distribution functions of \(D_\alpha\) using the FFT and EKF method in Region A, where \(p(D_\alpha|s_f)\) is the distribution of the \(D_{\alpha,\text{FFT}}\) values for an NDVI time series of the same class, \(p(D_\alpha|d_f)\) is the distribution of the \(D_{\alpha,\text{FFT}}\) values for an NDVI time series of different classes, \(p(D_\alpha|s_k)\) is the distribution of the \(D_{\alpha,\text{EKF}}\) values for an NDVI time series of the same class, and \(p(D_\alpha|d_k)\) is the distribution of the \(D_{\alpha,\text{EKF}}\) values for an NDVI time series of different classes.

\[\hat{Y}^i(k) = \begin{cases} Y^i(k), & k = \{0, 7\} \\ 0, & 1 \leq k \leq 6 \\ 0, & 8 \leq k \leq N \end{cases} \quad (16)\]

and \(F_N^{-1}\) denotes the inverse DFT operation. \(\hat{Y}^i\) is thus a copy of \(Y^i\) but with only the mean and seasonal FFT components; all other components were set to zero.

The initial state parameters, as well as the observation and process noise standard deviation for Regions A and B, are shown in Tables I and II, respectively. The values of \(\mu_1\), \(\alpha_1\), and \(\phi_1\) in Table I were calculated using (10)–(12), respectively. In Table II, the observation noise variance \(\sigma_v\) was calculated using (13) and (14), while the process noise variance \(\sigma_\mu\), \(\sigma_\alpha\), and \(\sigma_\phi\) were estimated by maximizing the class separability on the training data for each region.

Once determined, the parameters were kept fixed for all numerical results relating to the specific region.

The distance metrics \(D_{\alpha,\text{EKF}}\) and \(D_{\mu,\text{EKF}}\) were recorded, and the consequent distributions of the values of these distance metrics were calculated. The distribution and values of the FFT distance metrics \(D_{\alpha,\text{FFT}}\) and \(D_{\mu,\text{FFT}}\) were also calculated for comparison. Fig. 4 shows the distribution of the FFT and EKF \(D_\alpha\) values for Region A.

Table III gives the Bayesian decision error for both the FFT and EKF distance metrics for each region defined as

\[P_{\text{FFT}}(\alpha) = \int_{D_\alpha = 0}^{D_\alpha^*} p(D_\alpha|d_f) + \int_{D_\alpha^* < D_\alpha < \infty} p(D_\alpha|s_f) \quad (17)\]

\[P_{\text{EKF}}(\alpha) = \int_{D_\alpha = 0}^{D_\alpha^*} p(D_\alpha|d_k) + \int_{D_\alpha^* < D_\alpha < \infty} p(D_\alpha|s_k) \quad (18)\]

\[P_{\text{FFT}}(\mu) = \int_{D_\mu = 0}^{D_\mu^*} p(D_\mu|d_f) + \int_{D_\mu^* < D_\mu < \infty} p(D_\mu|s_f) \quad (19)\]

\[P_{\text{EKF}}(\mu) = \int_{D_\mu = 0}^{D_\mu^*} p(D_\mu|d_k) + \int_{D_\mu^* < D_\mu < \infty} p(D_\mu|s_k) \quad (20)\]

where the value of \(D_{\alpha,\text{FFT}}^\text{opt}, x \in \{\alpha, \mu\}, y \in \{f, k\}\), is the optimal decision threshold minimizing the probability of error \(P_e\) in each instance. It can be seen that the Bayesian decision error in Region A using the EKF was reduced by 5% over the FFT method when considering \(D_\mu\) and by 1% when considering \(D_\alpha\) (Table III). In Region B, the Bayes error of the EKF method was increased with 2% over the FFT method when considering \(D_\mu\), but a significant reduction of 14% was achieved when considering \(D_\alpha\). Thus, overall, it may be concluded that the EKF formulation has a reduced probability of error when compared to the FFT-based approach on the same data. This implies that the EKF formulation offers improved separability of land cover classes for the study areas A and B.

The phase parameter \(\phi\) was found to provide negligible additional separability in the classes and thus was excluded from the results.

V. Discussion of Results

The overall improved separability of natural vegetation and settlement land cover types using the EKF based on a triply modulated cosine function model over FFT is evident for both regions A and B. In an effort to improve the results, a sum of sinusoid model was also considered, but preliminary results showed a negligible performance increase with a significant increase in the complexity as more parameters needed to be estimated. This corresponds to results shown in [10], where no significant added separability was achieved when considering more sinusoidal components other than the annual component.

Consistent with most EKF implementation, the tracking of state parameters is not instantaneous and does require a certain amount of observations. As this period is unknown, an initial
number of state parameter values need to be excluded when calculating $D^{EKF}_\alpha$ and $D^{EKF}_\mu$. The average square difference between the EKF-derived $\mu$ parameter and the FFT mean component is shown in Fig. 5; it can be seen that the variation seems to stabilize within the first two years, which relates to approximately 100 samples. It was also found that, when excluding the first 100 estimates of the state parameters, a stable distribution of $D^{EKF}_\alpha$ and $D^{EKF}_\mu$ was obtained. The first 100 state parameter estimates were consequently disregarded in all results shown.

The initialization procedure used to determine the initial EKF parameters, as shown in Section IV, was found to work well for each region. By using an initial training set and keeping the EKF initialization parameters constant for each region, the EKF is effectively adaptable for each region and requires minimal manual parameter selection.

VI. CONCLUSION

Previous research has found that class similarity can be evaluated by considering the difference in FFT components, in particular the mean ($\mu$) and annual ($\alpha$) FFT components, as they tend to carry the majority of signal energy [10], [12]. In this letter, the mean and annual frequency components were estimated for each time step using an EKF. Having iterative estimates of these components allows one to exploit the fact that the mean and annual frequency dissimilarity is more prevalent during certain parts of the seasonal cycle than other parts, an effect that is merely averaged out using the traditional FFT over the entire NDVI time series. The variance of the distribution of the distance metrics derived using the EKF $p(D_\mu|s_k)$ and $p(D_\alpha|s_k)$ is higher than that of the FFT method (Fig. 4). The reason for this is that the EKF is more sensitive to the variability of pixels of the same class that are separated geographically. Even though the variance of the EKF distance metric is higher, the total probability of the decision error was reduced for both of the regions considered in this letter. A further possible application of this method is toward land cover change detection. By following the changes of the cosine parameters through time and comparing them with neighboring pixels, a change detection method can be formulated, and this is currently being researched. Finally, note that the EKF is well suited to multiple observations so that the EKF can capitalize on all seven bands as opposed to just the NDVI time-series data to further improve estimates of the underlying process model. This possibility is also currently being researched.

REFERENCES