A NOVEL TIME SERIES BASED APPROACH TO DETECT GRADUAL VEGETATION CHANGES IN FORESTS

YASHU CHAMBER*, ASHISH GARG*, VARUN MITHAL*, IVAN BRUGERE*, MICHAEL LAU*, VIKRANT KRISHNA*, SHYAM BORIAH*, MICHAEL STEINBACH*, VIPIN KUMAR*, CHRIS POTTER** AND STEVE KLOOSTER**

Abstract. It is well-known that forests play a vital role in maintaining biodiversity and the health of ecosystems across the Earth. This important ecological resource is under threat from both anthropogenic and biogenic pressures, ranging from insect infestations to commercial logging. Detecting, quantifying and reporting the magnitude of forest degradation are therefore critical to efforts towards minimizing the loss of one of Earth's most crucial resources. Traditional approaches that use image-based comparison for detecting forest degradation are frequently domain- or region-specific, which require expensive training, and are thus not suited for application at global scale. More recently, time series based change detection methods applied on remote sensing datasets have gained much attention because of their scalability, accuracy, and monitoring capability at frequent regular intervals. In this paper, we propose a novel approach to identify regions where forest degradation occurs gradually. The proposed approach complements traditional domain- and region-specific approaches by providing information on where degradation is occurring, and during what time, at a global scale.

1. Introduction

Forests play a vital role in maintaining biodiversity and the health of ecosystems across the Earth. However, this important ecological resource is under threat of degradation by both anthropogenic and biogenic pressures. Forest degradation occurs due to a number of different causes ranging from insect infestations to logging. Such reduction in forest cover not only has implications on the global carbon cycle, but also causes adverse effects on the ecosystem which are often realized by decrease in biodiversity, increase in the frequency of floods, droughts, changes in rainfall patterns, etc. [15, 10, 16]. Thus, detecting, quantifying and reporting the magnitude of forest degradation is critical to efforts towards minimizing this loss of one of Earth's most crucial resources.

Remote sensing offers rich data sets that are very well-suited for monitoring forests around the globe, in a regular fashion across time. A large variety of techniques and tools have been developed for detecting changes in forest cover, and more generally land cover [4, 11]. However, detecting gradual forest degradation (as opposed to abrupt changes caused by fires etc.) is particularly challenging because the reduction in forest cover occurs very slowly and the amount of reduction observed across time is small compared to natural variations.

Traditional approaches that use image based comparison for detecting forest degradation are frequently domain-specific or region-specific [5] which require expensive training, and are thus not suited for application at global scale. More recently, time series based methods applied on remote sensing datasets have gained much attention to detect deforestation because of their scalability, accuracy, and forest monitoring capability at frequent intervals. However, even most of the current time series based approaches for detecting vegetation loss in forests are aimed at only certain types of changes (e.g. due to fires), which are characterized by sudden and severe vegetation loss [12].

A number of approaches have been proposed for identifying gradual changes in a time series. Kucera et al. [9] describe the use of the well-known CUSUM technique for land cover change detection. CUSUM follows a simple approach of determining deviation in the values of a time series

* University of Minnesota, <chamber,ashish,mithal,ivan,mwlau,krishna,sboriah,steinbac,kumar>@cs.umn.edu
**NASA Ames Research Center, chris.potter@nasa.gov, sklooster@gaia.arc.nasa.gov.

from an expected value, and the change score, giving the magnitude of change, is determined as the maximum cumulative deviation. Another approach presented recently by Verbesselt et al. [17], Breaks for Additive Seasonal and Trend (BFAST), decomposes a time series into trend, seasonal and residual components. The time series is divided into segments such that intra-segment trend is constant, while inter-segment trends are dissimilar. A trend breakpoint is associated with segment boundaries. The seasonal component is handled in a similar fashion.

In this paper, we present a novel approach to identify regions where forest degradation is occurring gradually (either due to biogenic or anthropogenic causes). The approach is robust, scalable and easy to apply across different regions and vegetation types. The proposed method represents an adaptation of CUSUM for the problem of gradual change detection. While CUSUM only identifies a time series as changed or not, the proposed approach also identifies the period of change, in addition to having a considerable improvement in performance.

We begin by describing the underlying remote sensing data and preprocessing procedure in Section 2. Sections 3 and 4 discusses the concepts behind the development of the new approach. We formally present our method in Section 5 followed by a discussion in Section 6. We then evaluate the performance of the proposed approach in Section 7 using independent validation data sets in two regions of the world where the degradation has entirely different causes. Finally, in section 8, we comparatively evaluate the proposed approach, CUSUM and BFAST for detecting gradual changes.

2. Data and Preprocessing

The time series data set used for this study is the Enhanced Vegetation Index (EVI), which is a product based on measurements taken from MODIS instrument on NASA's Terra satellite, and is available for download from the Land Processes Distributed Active Archive Center (LPDAAC) [1]. EVI essentially measures the "greenness" signal as a proxy for the amount of vegetation at a location. The spatial resolution of the dataset is 250 meters and the temporal resolution is 16 days (23 time steps per year), and covers the time period from from February 2000 to the present. The range of EVI is 0 to 1, where 0 indicates no vegetation and 1 indicates vegetation saturation. Figure 1 shows an example of an EVI time series.

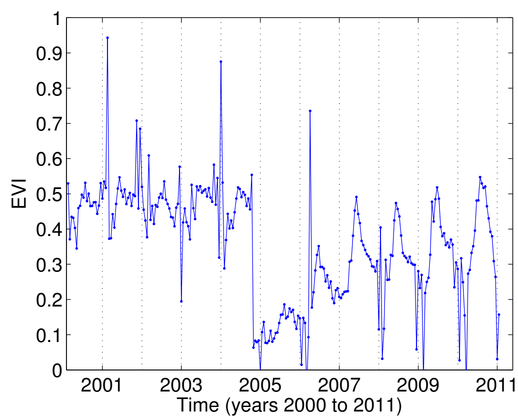


Figure 1. Example of an EVI time series (with noise in observations).

Remote sensing data sets are frequently subject to contamination due to clouds, haze, pixel geometry and other factors. We preprocess the EVI time series data set in order to remove undesired fluctuations in EVI (such as the sharp increases in Figure 1). This improves the efficiency of identifying signatures of interest. For smoothing purposes, we have used the Savitzky-Golay smoothing filter [13], which uses two parameters: polynomial *degree* desired for smoothing, and *frame size*. The smoothing filter fits a polynomial function of the indicated degree over a window equal to the frame size over each time step, the current time step being at the center of the window; the EVI value of the current time step is then replaced with the polynomial fit.

3. Detecting forest degradation: problem formulation and a CUSUM approach

A reduction in forest cover is often reflected as a decrease in the EVI value. In fact, many existing schemes compute difference in EVI (or related indices) between different years to identify changes. However, the values of vegetation indices such as EVI can have a high degree of variability due to seasonality (e.g. vegetation is greener in the summer than in the winter in the temperate zones), as well as due to natural variation in vegetation growth caused by environmental factors such as temperature and precipitation.

Most existing methods have handled seasonality by comparing vegetation index values at (or around) the same date in different years. Natural variability is much harder to handle since it can result in too many false positives. The problem becomes even more acute for many non-forest covers such as shrubs, since natural variability tends to be much larger in these cases. Although our focus is on identification of degradation in forests, it is not possible to completely exclude non-forests from any study due to the unavailability of highly precise forest maps [6].

Given an EVI time series dataset, we are interested in identifying time series such as the one shown in Figure 2, where there is a perceptible decrease in the signal, along with determining the approximate period of decrease (as shown by the vertical lines in Figure 2). Identifying a time series with a gradual decrease in vegetation is challenging due to a number of reasons: distinguishing vegetation loss from natural seasonal variations; differentiating between a spurious decrease due to noise or environmental factors and a genuine decrease from degradation on ground; correctly determining the period of decrease (start and end time steps) especially when there is a high degree of variability in the time series. There could also be a phenological change during the decrease period or across the decrease, and the algorithm must be able to handle such cases and extract the decrease period appropriately.

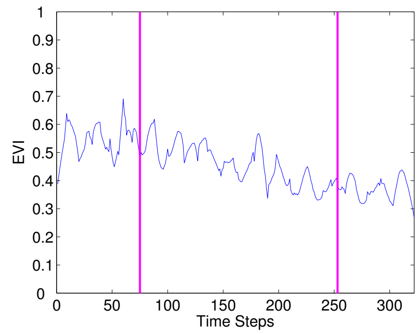


FIGURE 2. Example of a decreasing time series. Vertical lines enclose a gradually decreasing segment.

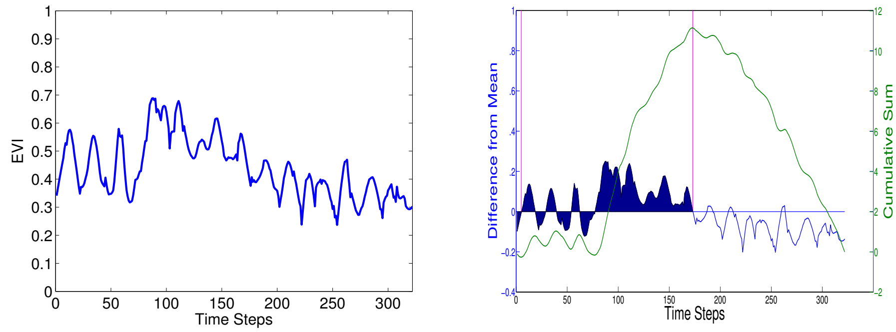
3.1. **Notation.** Table 1 defines notation used in this paper.

n	The number of time steps in a time series.
S	The number of time steps corresponding to one year of data (we also call this the season length). For biweekly data $S = 23$, and $S = 12$ for monthly data.
t_1	First time step.
t_i	ith time step.
v_{t_i}	Data value at the time step t_i
$v_{t_i}...v_{t_j}$	All values between time steps t_i and t_j.
T	A sample time series $= v_1 v_2 v_3 ... v_i ... v_n$
$v_{t_i...t_j}$	$mean(v_{t_i}...v_{t_j})$
Δ_i	$v_{t_{i-S+1}...t_i} - v_{t_{i+1}...t_{i+S}}$
Δ-series	$\Delta_S \Delta_{S+1} \Delta_{S+2} ... \Delta_{n-S}$

TABLE 1. Notation for time series change detection.

3.2. **CUSUM Method for detecting decreasing time series.** CUSUM is a well-known change detection algorithm that was originally developed in the domain of process control. It is one of the earliest change detection algorithms developed, proposed by Page [14]. One of the defining features of CUSUM is its ability to detect *small* and *gradual* changes in the process. The basic CUSUM scheme has an expected value μ for the process. It then compares the *deviation* of every observation to the expected value, and maintains a running *statistic* (the cumulative sum) CS of deviations from the expected value. If there is no change in the process, CS is expected to be approximately 0; if CS exceeds a user-defined threshold at any time step, the time series is flagged as changed.

There are multiple ways in which a change score can be assigned to a time series, the simplest of which is to use $\max\{CS_1, CS_2, \ldots, CS_n\}$. However, this score can be sensitive to noise [3]. Kucera



(a) A sample time series with decreasing period between time steps 100 and 210.

(b) Corresponding difference series, D, (blue) and cumulative sum series, Q, (green).

FIGURE 3. Measures computed by CUSUM.

et al. [9] developed a CUSUM approach for land cover change detection which uses a more robust technique to compute the change score. Specifically, a bootstrap procedure is used to determine the confidence of CS by determining the degree to which such a score can occur by chance. The bootstrap procedure involves randomly permuting the input time series to obtain a distribution of change scores $\mathcal{R}$ (CUSUM is run on each randomization). The confidence of the drop is determined by the relative frequency of CS being greater than the randomized distribution, i.e. $\frac{|CS > \mathcal{R}|}{|\mathcal{R}|}$.

Kucera et al. [9] take the expected value μ as the mean of the entire time series. Other measures may also be used to compute μ such as the value of the first time step, or the mean of the first S values. The advantage of using the mean value across a periodic cycle over a single time step is that the mean value is independent of the fluctuations in a time period (or seasonal variation in case of the MODIS EVI time series).

We illustrate some drawbacks of the scoring mechanism of CUSUM described above:
(1) *Change point of drop and period of decrease not identified.* This method only identifies a score corresponding to the maximum deviation in cumulative sum time series, and does not give the period of change. Figure 3 shows the scoring process using CUSUM. It identifies the maximum value in the cumulative sum series as the score. Thus, no change point of drop or period of drop is identified.
(2) *Computed score may not be associated with the decreasing period.* Computed score is the maximum cumulative deviation from the expected value, which may or may not depict the amount of EVI lost during the decrease period. This can again be noticed from Figure 3b.

4. Adapting CUSUM for Gradual Degradation

In the original CUSUM approach, the expected value is always fixed in a time series regardless of the way it is computed. In this paper, we propose a different strategy: If we take the expected value at any time step t_{i+1} as the value at time step t_i, then the deviation of values at each time step from its expected value would give the amount of drop or rise from its previous value. However, such a model would be dependent on the intra-periodic variation and the resulting deviation could be due to the natural periodicity of the time series. In order to make this process independent of periodicity, averaging over a periodic cycle can be used. Therefore, instead we take the mean value of the current periodic cycle as the expected mean value for the next periodic cycle. The deviation between mean values of successive periodic cycles is also equivalent to the drop in EVI across a time step t_i that marks the boundary between these periodic cycles, i.e. the one that ends at t_i, and the other that begins at t_{i+1}. We refer to this type of differencing (computing drop from previous periodic cycle in succession) as *Successive Differencing* (SD), which is different from computing the deviation from a fixed expected value as done in CUSUM, which we refer to as *Fixed Differencing* (FD).

If T is a given time series, n is the number of time steps in T, and S is the periodicity of T, we can define SD and FD methods as:

FD: $\Delta_i^c = v_{t_i} - \mu \quad \forall \quad i \in 1 \cdots n$

SD: $\Delta_i^s = \Delta_i \quad \forall \quad i \in S \cdots n - S \qquad (Refer \quad to \quad Table\,1 \quad for \quad notation)$

For MODIS EVI time series, Δ_i^s's are computed as difference between the mean values of two successive years (two consecutive sets of 23 values since $S = 23$).

FD gives deviation in values relative to the fixed expected value, while SD gives the drop relative to the previous year. Also, in the first equation, individual data values are used for differencing while in the second equation, averaged value over a seasonal cycle is being subtracted. Computing drop from a previous value in succession can provide trend information in a time series, which is what successive differencing does. Also, subtracting averaged values instead of individual values make the trend information more robust to seasonal variations and noisy outliers. On the other hand, Δ_i^c's fluctuates with the seasonal variations, even when there is no decrease in vegetation. Also, it does not provide trend information which is vital for identifying decreasing period in a time series.

Using Successive Differencing. As we have seen above, successive differencing using mean value of annual segments can be used to determine trends in a time series. Therefore, we use SD instead of FD for our approach. Below, we mention some possibilities in which SD could be used:

Consider a method for detecting gradual decrease that tries to identify the window in a time series that has the largest drop: given a time series, identify two years, i.e. two sets of 23 consecutive time steps, y_1 and y_2, such that the difference between the mean EVI of y_1 and y_2 is maximum in the time series. This is similar to identifying the window where the sum of Δ_i^s is maximum. This method will work well for consecutively decreasing time series. However, there are some disadvantages of this method when applied to a time series with high variation. The primary disadvantage is that this method loses information about the time steps in between y_1 and y_2. For example, given the time series shown in Figure 4, this method would identify the decrease as having occurred between years 2 and 11 even though it is clear that the time series increased significantly after year 7. Specifically, if the time series rises and falls in between then such a time series is highly variable and it should not be considered as changed. Another disadvantage is that if there are large spikes in a year due to noise that distorts the mean EVI for that year in an otherwise stable time series, this time series will be given a high score (drop from y_1 to y_2) by this method, even though this change is spurious.

To overcome drawbacks of the above method, yet another method to detect gradual decrease could be to compute the difference between successive yearly sets (Δ_i^s), and determine the longest continuous window of positive Δ_i^s. This method again has a major disadvantage that if there is a spurious rise in time series due to noise, the drop window will fall short of that false rise and thus could be determined much smaller than it actually is (e.g. for the time series in Figure 4, this method would incorrectly detect end of degradation in year 5).

Building on the concepts described in this section, we propose a novel time series change detection method, *Persistent-*Δ *Approach*, or $PDELTA$. It uses successive differencing as the base to compute Δ_i's. The key property of successive differencing is that as long as there is a decrease in the time series from one year to the next, Δ_i would be positive. If the decrease is at an almost constant rate, the Δ_i would be almost constant. As soon as Δ_i becomes zero, it means that there has been no vegetation change from past year to the present year. But it could be too soon to say that the change in vegetation has stopped since this could be due to some noisy time steps and it's possible that after very few time steps, Δ_i's become positive and stay positive for a couple of years or more. Thus the change didn't really stop, but continued after a short time. Therefore, the primary objective of PDELTA is to determine the window of *maximum reliable drop*. This method tolerates natural variation which may cause small increases in individual years during an extended period of degradation. For example, in Figure 4, the technique correctly identifies the period between years 2 and 6 when degradation has occurred since it accounts for the perturbations in the intervening years. However, if this rise in the time series violates a *reliability condition*, the approach differentiates it from the natural variation and does not consider this in the changed phase. The next section describes the PDELTA method in detail.

5. Description of PDELTA

As mentioned in the previous section, the main objective of the PDELTA approach is to identify the window of maximum reliable drop in a time series. It need not be a continuous drop, and some amount of intermittent rise can be allowed as long as a decreasing trend is persistent. The amount of intermittent rise allowed is controlled by a simple, but strong condition (*Reliability Condition*) which essentially limits the amount of every intermittent rise during an extended period of degradation. The remainder of this section describes the approach in detail.

As a first step, we compute Δ_i, as described in the previous section, at each time step beginning at the end of the first year (since we don't have sufficient information to compute Δ_i during the first year) and terminating before the start of the last year (again due to insufficient information during the last year). Let the series composed of Δ_i's, $S \leq i \leq n - S$, be Δ-series (Delta-series) where S are the number of time steps in a year, and n are the number of time steps in the entire time series. Next, we identify those time steps in the time series that have the characteristic to become the extremes of the drop window. For this effect, we compute a Γ-series (Gamma-series) from Δ-series, with each time step represented by γ_i, using the following transformation:

$$\gamma_i = \begin{cases} 1 & \Delta_i > 0 \\ -1 & \Delta_i \leq 0 \end{cases}$$

We say that those ith time steps are candidates for drop start for which γ_{i-1} is -1 and γ_i is 1 (transition to 1). Similarly, those ith time steps are candidates for drop end for which γ_{i-1} is 1 and γ_i is -1 (transition from 1). The bottom plot in Figure 4 show an example of the Δ-series, Γ-series scaled by a factor of 0.07, and candidate start and end time steps (b_1, e_1, etc.).

Let there be K candidate start and stop time steps identified in time series T, which are denoted by b_k and e_k respectively, $\forall k \in 1 \ldots K$. Between every b_k and e_k there is a decrease in EVI values (decreasing trend), and between every e_k and b_{k+1} there is an increase in EVI values (increasing trend). For each b_k we are interested in identifying the farthest e_l ($1 \leq k \leq l \leq K$) such that the time series pattern within these limits in general has a decreasing trend, even if there are mild rises in between. If a drop starts at b_k, then an intermediate rise occurs between every e_l and b_{l+1} ($k \leq l < K$). In order to ensure that the decreasing trend is followed across these intermediate rises, we test for a drop reliability condition at every e_l which must be satisfied before allowing the rise between e_l and b_{l+1}. This reliability condition is given below.

Reliability Condition (RC): It states that after the commencement of a drop at a time step b_k, the rise occurring at a certain time step e_l ($e_k \leq e_l < e_K$) would not be considered as drop termination if this rise is bounded by a fraction ($x\%$) of the drop that has already occurred, and, if there is an *overall decrease* in the EVI values during the time steps in the *limited future* of e_l. The *limited future* is defined as the time steps following e_l during which the rise does not exceed $x\%$. The motivation for this is that if the current drop window has accumulated a large sum of Δ_i, a greater room for rise is allowed as long as the time series in general still has a decreasing trend.

As long as this condition is satisfied, e_l can be extended. As soon as this condition fails at a certain candidate stop time step, we stop at that e_l, and the maximum reliable drop that began at time step b_k is terminated at e_l. This becomes a candidate drop window (cw_p) for time series T. Since there could be more such windows in the same time series that begin at other candidate start time steps, b_j, $j \neq k$, we repeat the above process for the remaining candidate start time steps. Thus, we would have at most K candidate drop windows cw_p, $\forall p \in 1 \cdots K$.

To identify the best drop window, we compute a score for each candidate window using one of the methods described in Section 6. The maximum scoring window is determined as the representative window of the time series. Currently, we identify a single representative window of a time series because we are interested in identifying time series that have undergone change at least once. Thus,

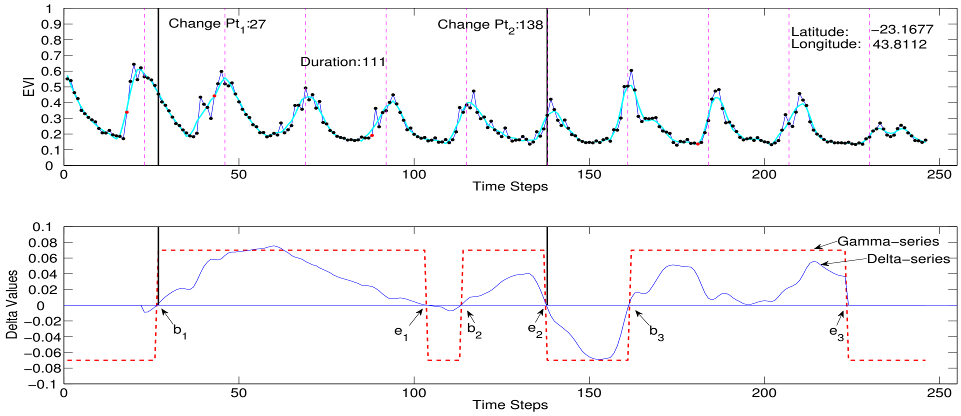


FIGURE 4. The top plot shows the original time series (line connecting the dots), smoothened time series, and the identified changed period (between solid vertical lines). The bottom plot shows the corresponding Δ-series (continuous curved line) as well as the Γ-series (broken line) scaled by a factor of 0.07.

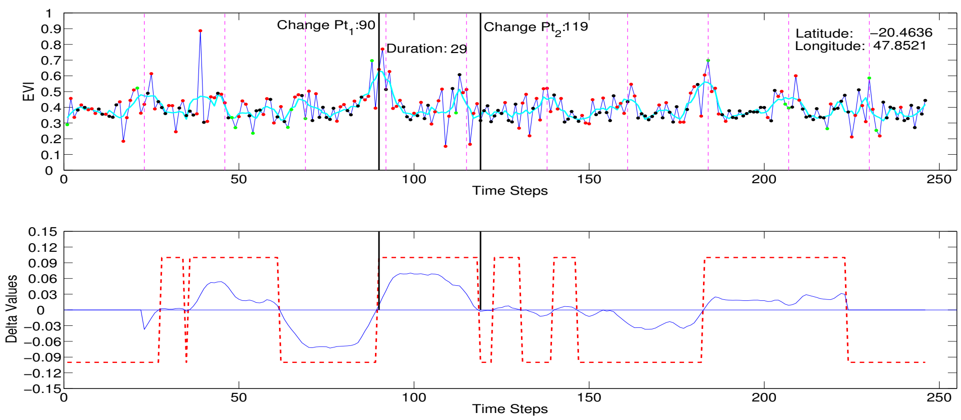


FIGURE 5. A time series in Madagascar showing a spurious rise in EVI.

.

the significance of change in each time series is given by the score computed for their corresponding representative change window. The higher the value, the more severe the change.

The example time series in Figure 4 captures the effectiveness of this approach. The top plot shows the EVI time series, the smoothed time series, and the two vertical solid lines identifying the drop period. It can be noticed that the time series gradually decreased over many years and then stabilized. The bottom plot shows the corresponding Δ-series in solid curved line and the Γ-series scaled by a factor of 0.07 by a broken line. Notice that the first drop in Δ-series below zero is included in the maximum reliable window since the reliability condition is not violated.

Let us also consider a case of a spurious rise as shown in Figure 5. In such cases the drop window resulting from this rise will be small since the subsequent period will not be able to satisfy the reliability condition. Furthermore, the score of the identified drop window according to our methodology (as would be described shortly) would be low. Hence these type of drops would be easily differentiated from the genuine drops.

6. Scoring Mechanisms and Discussion on Reliability Condition

Scoring Mechanisms. Once the candidate change windows are determined in a time series, the next step is to quantify the change in each window. Some of the methods used are: (i) **Drop in**

EVI, which is the difference between the mean annual EVI just before the start of the drop and the mean annual EVI just after the termination of the drop; (ii) **Length of the Drop Window**: The length can be a powerful indicator of the confidence of the change. A decrease of longer duration, even with a small *Drop in EVI*, can be of high confidence. Beetle infestation in Colorado (Figure 7) is a good example of this scenario; (iii) **Total Loss in EVI**: If we assume that the drop didn't take place in an actually changed time series, we can suppose that the EVI pattern, P_1, representing the year before the start of the drop would have continued. In such a scenario, the total loss in EVI would be the area enclosed between two time series, one that should have been had no vegetation loss occurred, and the other which is the actual time series in which there is a loss in vegetation.

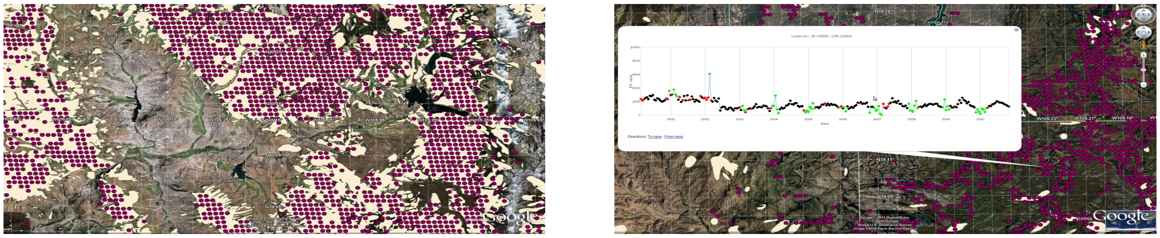
Here, we also introduce the concept of a *third* change point (the first being the drop start time step and the second being the drop end time step). After the drop occurs, if one wants to determine how long it takes for the time series to have a *significant* recovery, the third change point can be used. The third change point is positioned at a time step after the second change point such that the EVI values have risen a significant percent (say, 50%) of what it has dropped during the drop window. The position of the third change point can also be an indicator of the confidence of the drop. If the third change point is realized after many time steps following the drop window, the drop is trustworthy because the vegetation stays low for a long time. On the other hand, if the third change point occurs soon after the second change point, it might mean that either the vegetation indeed recovered very quickly, or the drop was actually spurious and short-lived.

The different scoring schemes could also be incorporated into a single *cost function*. The new cost function could be constructed such that it gives a minimum cost to a drop window that receives a high score from all the above mentioned schemes, as well as high cost to a drop window getting a low score from each scheme. A single cost threshold could be set which differentiates a genuine change from a spurious one. This cost function must be designed with care such that it's applicable globally, and we leave this for investigation in future work.

Accounting for Variability in a Time Series. Though we have briefly mentioned variability in a time series before, here we discuss it in the context of quantifying the EVI loss. Natural variability occurs in EVI values from one year to the other due to changes in environmental conditions such as temperature, precipitation, cloud cover, etc., or imprecision in measurement. Such a change in a time series should not be regarded as a loss in vegetation. Furthermore, a true loss in vegetation would be over and above the natural variability because a loss in vegetation equal to the natural variability would actually be a common signature of that vegetation. A way to model the natural variability is to take a mean of pairwise distance between EVI values of annual segments either during the first few years (before the first change point) or immediate previous few years before the first change point. The *city-block* distance measure (L_1 norm) works well for computing this variability. This variability is subtracted from the *Drop in EVI* in order to reflect a true drop in vegetation on the ground. Similarly, the EVI values of the pattern P_1 should be lowered by this variability before computing the *Total Loss in EVI*. For evaluation presented in this paper, we have used *Total Loss in EVI* as the scoring scheme with compensation for variability.

Potential Reliability Condition Augmentation. In the previous section, we described a method to determine the start and end time steps of a candidate drop window. As it would appear, these change points are dependent on the reliability condition used. We have described a simple reliability condition that limits the amount of rise in the time series once the drop has begun. If the rise is greater than a threshold, the drop is terminated by positioning a second change point before this rise and a new drop window is initiated at the next candidate start time step. Finally, the best drop window is selected as the representative drop window of the time series.

The advantage of using the above reliability condition is that it is simple, as well as it has only one parameter. But this condition may handle some time series changes differently from what one might prefer. For instance, if there is a time series that drops during the third year, stays almost constant



(a) A snapshot showing locations identified as changed by the proposed approach (red circles) overlaid on the validation data polygons.

(b) Snapshot showing a large region in Colorado where there was a drop in EVI due to Fire in the year 2002. Most of this region is not covered by the polygons.

Figure 6. Snapshots showing regions in Colorado where vegetation loss was detected.

for the next five years, and then drops again during the eighth and ninth years, then the start and end change points would be identified around the second year and the ninth year respectively. It would overlook that the time series is not decreasing at all for many intermittent years. It might be more desirable to include these two drops in separate drop windows instead of one. But such cases are not a limitation, as the proposed reliability condition can be adapted to handle these cases. It can be taken as a base condition over which other conditions are added, which may or may not be region specific. In the context of the above example problem, one possible adaptation could be to add a condition that the drop window must have recurrent drops, say every y years or so. This wouldn't allow a period of stagnation for more than y years. Note that this modification would result in the use of another parameter, y, thus deviating further from simplicity.

7. Evaluation

Evaluation of a scheme for detecting changes in forest cover is challenging due to the lack of high-quality ground truth. The most reliable methods for generating ground truth (e.g. ground surveys) are very expensive and are thus only available for small regions.

In the absence of such gold standard ground truth, less reliable labels generated by some other scheme or via aerial surveys can still be used for validation, but care must be taken to check if "false positives" (i.e. changes found by the scheme but not in the validation data) are indeed false, since they could have been missed by the scheme used to generate labels. Similarly, one needs to check if "false negatives" (i.e. changes noted in the validation data but not found by the scheme) are indeed changes on the ground as they could be incorrectly classified as changed by the other scheme.

We evaluated our approach on two regions; Northern Colorado and Southern Madagascar for which moderate quality labels are available in the form of polygons covering degraded areas of the forests. These regions are interesting because they have completely different vegetation types, and the degradation is caused by different mechanisms; specifically, insect damage in Colorado, and logging in Madagascar. During analysis in both regions we emphasize the strength of this algorithm in detecting changes that are difficult to identify, as well as the capability of this approach to capture many changes that are missed by the validation data.

7.1. **Evaluation on Colorado Forests.** The first region of analysis is forests in northern Colorado (the region bounded by 39°N—41°N; 108°W—104.5°W). The US Forest Service and its partners [2] maintain data sets which map the regions of forest cover that have degraded between years 2002 and 2008 in northern Colorado. The objective was to detect regions of forest degradation using our approach and evaluate it against the above validation data (which has been transformed to polygons). We present our analysis as follows:

True positives are points detected by our approach that also lie in the polygons. As seen in Figure 6a, there is a very good overlap of the detected points with the validation polygons. The algorithm is also able to correctly identify the period of degradation. Colorado is a difficult region

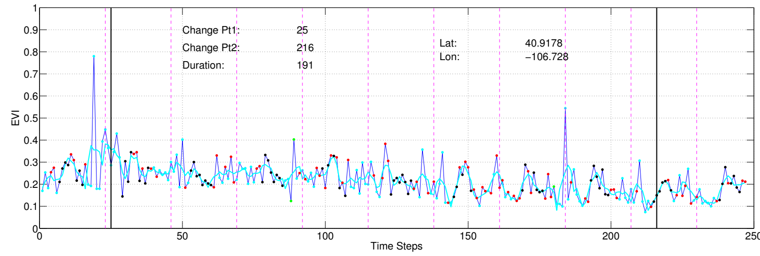
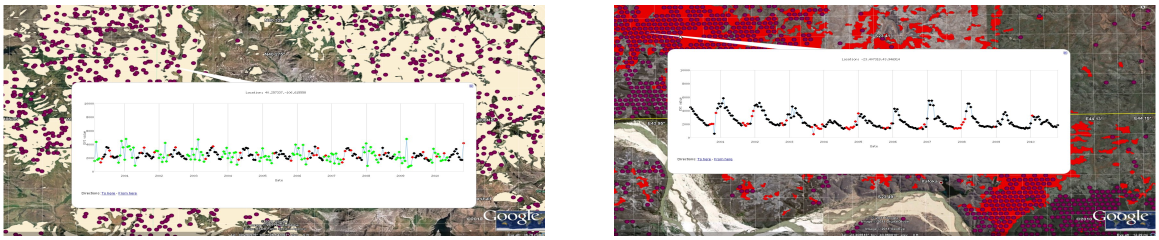


FIGURE 7. A typical gradual drop in Colorado due to beetle infestation.



(a) This figure shows an example time series in Colorado that has no perceptible change but which fall inside the ground truth polygon.

(b) Time series of a region in Madagascar showing gradual decrease in EVI starting from year 2001, which lie inside the ground truth polygon.

FIGURE 8. Snapshots of false negatives in Colorado and true positives in Madagascar.

because changes here are often very gradual, sometimes to the extent that there is no visible change in EVI signal upon manual inspection. Nevertheless, our approach identified a significant number of points. Figure 7 shows the typical EVI time series in this region.

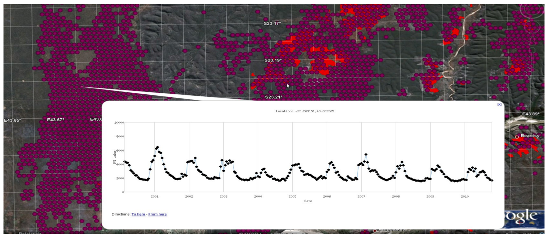
False positives are points that we detected as change but do not lie in any of the polygons. There could be several reasons for this: (i) the decrease in vegetation in these areas is caused by factors other than those considered in constructing the validation data; (ii) it is known that the polygons can be inaccurate (iii) these points are in fact not changed, but due to noise in EVI appear as changed and thus given a high score by the proposed approach. Our manual inspection shows that majority of false positives with high scores are due to (i) and (ii). For example, consider the region shown in Figure 6b. This region is not part of any polygon even though the change is quite apparent and is likely due to fire [7].

False Negatives are points that we did not detect as changed but which lie inside the polygons. Figure 8a, shows an example of such a time series. This time series does not show any change in the EVI signal. There are numerous time series in this region that show little perceptible change and are in the polygons. So either the vegetation loss here is too gradual to be detected by our approach, the change on the ground is not captured by the EVI signal, or the polygons are inaccurate.

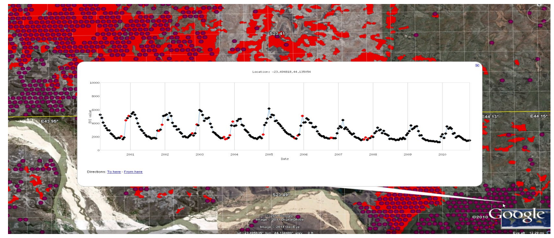
7.2. **Evaluation on Madagascar Forests.** The second region of analysis is southern Madagascar (the region bounded by 25.6°S—20°S). The validation data is obtained from Center for Applied Biodiversity Science (CABS) at Conservation International (CI), whose analysis is based on bitemporal Landsat image comparison between years 2001 and 2005 [8]. Hence, the validation data (or polygons) cover changes only between these two years.

True Positives and False Negatives: Figure 8b shows an example of a true positive. The image clearly shows the gradual drop starting in the year 2001. Most points in the validation polygons show similar behavior, although with varying decay rate and duration. Hence, effectively there are few false negatives since most points in the validation polygons can be found by our algorithm.

False Positives: Most of the false positives were observed due to the following reasons:



(a) A snapshot showing a large region in Madagascar where vegetation loss started to occur in the year 2001.



(b) This figure shows an area where vegetation loss occurred after year 2005.

Figure 9. False Positives in Madagascar having a decreasing EVI signal.

(1) Vegetation degradation occurred during the period of analysis (2000-2005) but the vegetation recovered during 2005, which causes it to be missed by the technique used in generating the validation data (Figure 9a). The entire cluster of points to the left in the figure has similar time series signature but lies outside the validation polygons. This illustrates the limitation of the technique that are based on the comparison of images taken on two different dates. (2) Significant vegetation degradation is visible only after 2005 hence was not included in the validation data set. Figure 9b shows an example of such a region. This type of identification also highlights the capability of our approach to find changes with high temporal precision in a continuous manner. This is opposed to the image-based methods where analysis is usually done on snapshots of images generally few years apart.

8. Comparative Evaluation on Synthetic Dataset

In this paper, we compare our proposed technique with two other approaches for gradual change detection, CUSUM and BFAST [17] using data sets with simulated noise and change characteristics. The BFAST technique is designed to detect long-term changes in satellite image time series. It decomposes a time series into trend, seasonal, and residual components, such that the intra-segment models are constant, while inter-segment models are dissimilar. BFAST identifies the optimal position of trend and seasonal breakpoints by minimizing the residual sum of squares (RSS), and the optimal number of breaks can be determined by minimizing an information criterion. Before estimating the breakpoints, the ordinary least squares residuals-based moving sum test is used to identify if any breakpoints are occurring in the time series. As output, BFAST provides the trend breakpoints and associated trends, seasonal breakpoints and associated seasonal models, and logical values indicating whether the time series is considered changed in the seasonal or trend components. We do not consider the seasonal component in this paper since we are looking for a decreasing trend.

Evaluating our algorithm against BFAST is not straightforward since BFAST looks not only for drops, but any type of trend change in a time series. Also, simply consulting the logical vector values that labels a time series as changed was not feasible for two reasons: (i) BFAST appears to be sensitive to noise and frequently finds different trends even in a stable time series and labels them as changed. (ii) BFAST would label a time series as changed if any type of trend change is present, notwithstanding the absence of a decreasing trend. In addition, BFAST also requires some parameter settings such as the minimum segment size and maximum number of breakpoints desired. These parameters are not mandatory, but not setting them makes it quite sensitive to noise, resulting in breaking even a single trend into multiple segments. Therefore, construction of the synthetic datasets had to be in consonance with the parameter values of BFAST.

We constructed two types of datasets, DS1 and DS2, the first containing three different trends (two trend breakpoints), and the other containing four different trends (three trend breakpoints). The maximum number of breakpoints set in BFAST for these two types of datasets were two and three respectively, and it was expected that BFAST would correctly identify all the given trends. Since we are interested in identifying the decreasing trend, our trend of interest among the ones returned by BFAST is the one which has the largest decrease across it. Its change score (which also represents the score of the time series) is computed in the same manner in which we compute the score for our proposed approach. Below, we describe the creation process of the synthetic dataset.

8.1. **Synthetic Data Generation.** The datasets, DS1 and DS2 are comprised of 1100 time series each, in which a gradual decrease phase was inserted in 80 time series for DS1, and 120 time series for DS2. In DS1, 40 time series also have an increasing trend. The remaining stable time series in both the datasets are identical (total: 980). Each time series has 322 time steps, with a seasonal period of 23 time steps (in order to mimic the MODIS EVI time series having 14 years of data). The seasonality in a time series is created using a function of the form:

$$F(x) = A * e^{\frac{-|x-m|}{B}}$$

where A controls the amplitude, x varies between time steps of a particular year, m controls the position of the peak in that year, and B controls the curve. The shape of $F(x)$ mimics a typical seasonal vegetation pattern of a forested region (or farming cycle) as reflected in an EVI time series.

Each time series has different types of noise added to it. We define these below, followed by the characteristics of the changed and stable time series.

Noise characteristics Two types of noise are introduced in the dataset. w_1 is white noise that is added to each time step in the time series. w_2 is outliers, that results in very high (upward spikes) or very low (downward spikes) values at certain time steps as compared to that of its neighbors.

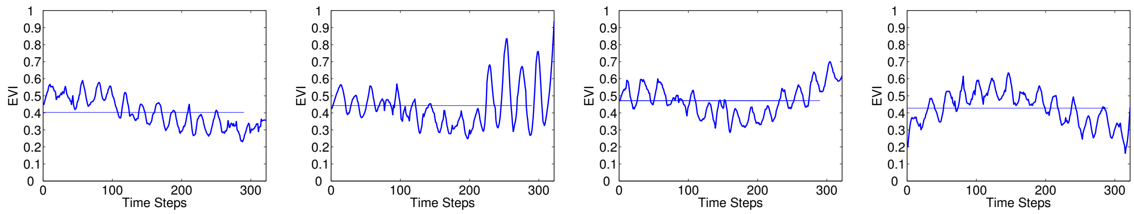
Characteristics of a changed time series There are three phases in these time series. (1) *beforePhase* is the period in the time series before a drop. Here, seasonal cycles (pattern during one year) are represented by $F(x)$. This phase may have an increasing trend or a stable trend. Noise w_1 and w_2 is added to each time step. All introductions in this phase are probabilistic as a Gaussian distribution within sufficient ranges specified in advance. This includes the values of w_1, w_2, duration of this phase, height of the data values during each year, duration of the increasing trend if any. (2) *changePhase* starts as soon as *beforePhase* ends. The majority of these time series have a decreasing trend. The base level of successive years in this phase is reduced gradually from starting of this phase till its end. The duration of this phase and the amount of drop introduced are probabilistic within a certain range. Noise w_1 and w_2 are added to this phase as well. In a small fraction of time series, an increasing trend is added during this phase instead of a decreasing trend to include more variety of time series. However, since these time series do not contain a decreasing trend, they are considered as false positives if detected by any algorithm. (3) *afterPhase* starts after the *changePhase* ends. Each year in this phase is also represented by $F(x)$ with w_1 and w_2 added.

Characteristics of a stable (unchanged) time series These time series have one phase with a constant level whose value is probabilistic within a certain range. Each seasonal cycle in these time series is also represented by $F(x)$ with noise w_1 and w_2 added.

8.2. **Evaluation Strategy.** PDELTA, CUSUM, and BFAST are applied to these datasets after preprocessing as described in Section 2. We compare the performance of our approach with CUSUM and BFAST separately. It is because these two approaches return different information about the drop, and we adjust our evaluation according to this information. For CUSUM, we combine the samples of time series from datasets DS1 and DS2 into a single dataset DS0.

8.3. **Comparison with CUSUM.** We evaluated the performance of PDELTA and CUSUM on the dataset DS0. There are four different types of time series in DS0 that have a decreasing trend (Figure 10). Each pattern has 50 different samples. Overall, this dataset has 200 decreasing time series, and 1020 stable time series (Total: 1220). The time series patterns included in DS0 are common in the real world datasets. Pattern one (Figure 10a) is an example of a stable forest that degrades over many years. Pattern two (Figure 10b) is an example of conversion of forests to farm lands, as could be noticed from the typical farming cycles during the later part of the time series. Figure 10c reflects some plantation following a deforestation. Figure 10d could depict a failed reforestation.

The precision-recall curve of the result of the two algorithms is shown in Figure 11a. CUSUM performed best on the samples of the pattern shown in Figure 10a. Admittedly, CUSUM performs



(a) Pattern one (b) Pattern two (c) Pattern three (d) Pattern four

FIGURE 10. Different decreasing patterns in D0 dataset. Each pattern has fifty samples (total two-hundred). The horizontal line shows the mean of the time series.

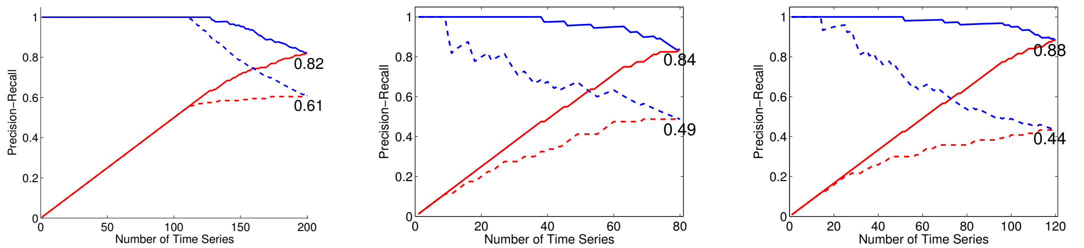
well for any gradually decreasing time series that accumulates a large score during the beginning time steps, which would happen if most of the beginning values are placed above the expected value μ. However, CUSUM would perform equally poorly on time series having an opposite signature. This highlights a major drawback of CUSUM. Consider the patterns two and three shown in Figures 10b and 10c, on which CUSUM performed poorly. In these time series, most of the values in the beginning are below μ, which is taken as the mean of the time series, implying that the majority of the later values are above μ. Such time series would never be able to accumulate a high cumulative sum since it incurs a large loss in the beginning due to negative deviations, and therefore would be given a low score. Many such scenarios could be constructed where there is a decreasing trend present but the time series never accumulates a high enough score for it to be significant.

We also investigated alternative ways of computing the expected value, but these variations either repeated some of the above drawbacks, or other drawbacks were discovered in them. For instance, by taking μ as the mean value of the first year, CUSUM performed poorly on patterns two, three, and four (Figure 10). Note that if we consider this variation, we are looking for the minimum cumulative sum (instead of the maximum) since an ideally decreasing time series would have a highly negative cumulative sum. The main disadvantage of this variation is that the cumulative sum is highly dependent on the first year values. If the first year values are noisy, it can drastically affect the algorithmic output. As an example, if μ is even slightly high due to noise, a stable time series could get a high score. In contrast, if μ is low, decreasing time series can go undetected (Figure 10d).

The precision-recall curve suggests that PDELTA performed considerably better than CUSUM on this dataset. Additionally, PDELTA also identifies the period of decrease.

8.4. **Comparison with BFAST.** Our evaluation with BFAST is based on two factors: (1) Precision-recall curves for PDELTA and BFAST, formed by ranking the time series in decreasing order of scores, when evaluated on DS1 and DS2 (Figures 11b and 11c). (2) Scatter plots of the deviation of the change points identified by the two approaches from the actual positions where the breakpoints were introduced in the synthetic datasets (Figure 12).

The scatter plots show that the distribution of the deviation of the change points identified by PDELTA is closer to zero than that of BFAST. BFAST segments a time series based on RSS and a Bayesian Information Criteria (BIC), which has no bias towards identifying decreasing periods. If identifying the decreasing trend as a separate segment minimizes RSS, BFAST will correctly identify the decreasing trend. Otherwise, it may combine a part of the decreasing trend with an adjacent trend. Also, BFAST appears to be sensitive to noise in a time series. It correctly detected the trends introduced for many time series (Figure 13a), but it segmented stable time series into different trends as well (Figure 13b). This suggests that BFAST might not be very suitable for highly variable time series, where noise levels can distort the ideal seasonal pattern enough. Such highly variable time series are characteristic of the tropical belt such as in Para (Brazil), Peru, Congo, etc.

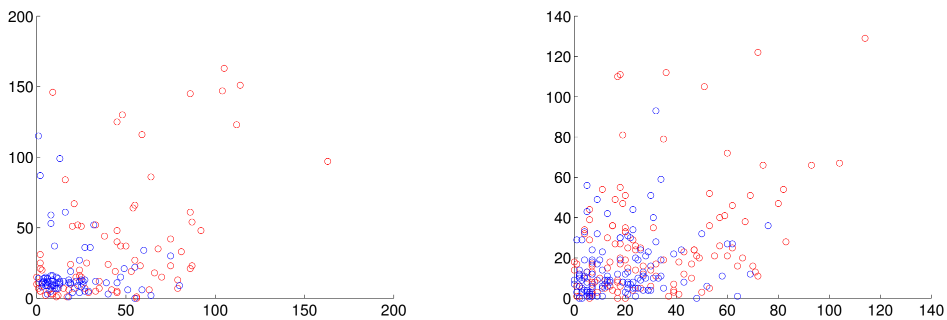


(a) DS0 - Precision-Recall for PDELTA (solid - precision: 0.82) and CUSUM (broken - precision: 0.61). Noise $w_1 = 200$, $w_2 = 10\%$.

(b) DS1 - PDELTA (0.84). BFAST (0.49)

(c) DS2 - PDELTA (0.88). BFAST (0.44)

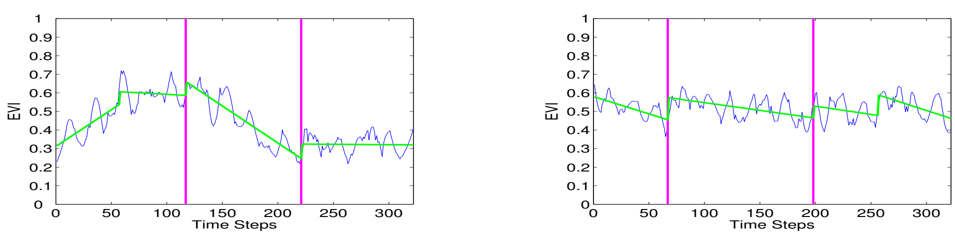
FIGURE 11. Precision curves (blue) and recall curves (red) for PDELTA (solid curves), CUSUM (dashed curves), and BFAST (dashed curves).



(a) DS1 - PDELTA (Blue), BFAST (Red).

(b) DS2 - PDELTA (Blue), BFAST (Red).

FIGURE 12. Scatter plots of the deviation of change points detected by PDELTA and BFAST from actual drop start (x-axis) and drop end (y-axis) time steps.



(a) Correctly detecting a decreasing trend.

(b) Many trends identified in stable time series.

FIGURE 13. Trends using BFAST. Vertical lines identify the period of maximum drop.

9. CONCLUSION

In this paper, we presented a globally scalable novel approach, PDELTA, for detecting a gradually decreasing EVI time series that can capture changes caused by a variety of sources. PDELTA can be considered an adaptation of CUSUM with the added capability of identifying the period of decrease and quantifying the magnitude of drop in a time series, while being more robust in the presence of noise and spurious changes. We demonstrated the efficacy of the proposed approach using independent validation data sets in Colorado and Madagascar. It was also shown that genuine changes were detected by our technique which were missed by other approaches, as well as points identified as changed by other approaches with no perceptible EVI signal were not detected. We comparatively evaluated our technique with CUSUM, and the state of the art BFAST technique. BFAST in its present form is computationally very expensive, whereas both PDELTA and CUSUM are quite fast. PDELTA can also identify reforested areas depicted by increase in vegetation simply by reversing a time series before applying this algorithm. Future extensions of this work include adapting PDELTA to detect more general types of changes (e.g. abrupt changes). Also, while this paper focuses on identifying the single most significant drop in a time series, PDELTA is able to identify multiple decreasing segments. Thus, other decreasing segments could also be identified as separate changes if the drop within them is also significant (multiple change detection). This aspect of PDELTA needs to be further explored and developed.

10. Acknowledgments

The research described here is supported by the Planetary Skin Institute, NSF Grant IIS-0713227, NSF Grant IIS-0905581, NASA grant NNX09AL60G and the University of Minnesota MN Futures Program. Access to computing facilities was provided by the University of Minnesota Supercomputing Institute.

References

[1] Land Processes Distributed Active Archive Center. `http://edcdaac.usgs.gov`.

[2] U.S. Department of Agriculture, Forest Service, Rocky Mountain Region, Forest Health Management. `http://www.fs.fed.us/r2/fhm`.

[3] S. Boriah. *Time Series Change Detection: Algorithms for Land Cover Change*. PhD thesis, University of Minnesota, 2010.

[4] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin. Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25(9):1565–1596, 2004.

[5] L. Eklundh, T. Johansson, and S. Solberg. Mapping insect defoliation in Scots pine with MODIS time-series data. *Remote sensing of Environment*, 113(7):1566–1573, 2009.

[6] G. M. Foody. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1):185–201, 2002.

[7] R. T. Graham. Hayman fire case study. Technical Report RMRS-GTR-114, U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 2003.

[8] G. J. Harper, M. K. Steininger, C. J. Tucker, D. Juhn, and F. Hawkins. Fifty years of deforestation and forest fragmentation in madagascar. *Environmental Conservation*, 34(4):325–333, 2007.

[9] J. Kucera, P. Barbosa, and P. Strobl. Cumulative sum charts-a novel technique for processing daily time series of modis data for burnt area mapping in Portugal. In *MultiTemp 2007: International Workshop on the Analysis of Multi-temporal Remote Sensing Images*, pages 1–6. IEEE, 2007.

[10] W. A. Kurz, C. C. Dymond, G. Stinson, G. J. Rampley, E. T. Neilson, A. L. Carroll, T. Ebata, and L. Safranyik. Mountain pine beetle and forest carbon feedback to climate change. *Nature*, 452(7190):987–990, 2008.

[11] D. Lu, P. Mausel, E. Brondízio, and E. Moran. Change detection techniques. *International Journal of Remote Sensing*, 25(12):2365–2401, 2003.

[12] V. Mithal, A. Garg, S. Boriah, M. Steinbach, V. Kumar, C. Potter, S. Klooster, and J. Castilla-Rubio. Monitoring global forest cover using data mining. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(4):36, 2011.

[13] S. J. Orfanidis. *Introduction to signal processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, 1995.

[14] E. S. Page. Continuous inspection schemes. *Biometrika*, 41(1/2):100–115, 1954.

[15] G. R. van der Werf, D. C. Morton, R. S. DeFries, J. G. J. Olivier, P. S. Kasibhatla, R. B. Jackson, G. J. Collatz, and J. T. Randerson. Co2 emissions from forest loss. *Nature Geoscience*, 2(11):737–738, 2009.

[16] P. J. van Mantgem, N. L. Stephenson, J. C. Byrne, L. D. Daniels, J. F. Franklin, P. Z. Fulé, M. E. Harmon, A. J. Larson, J. M. Smith, A. H. Taylor, and T. T. Veblen. Widespread Increase of Tree Mortality Rates in the Western United States. *Science*, 323(5913):521–524, 2009.

[17] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor. Detecting trend and seasonal changes in satellite image time series. *Remote sensing of Environment*, 114(1):106–115, 2010.