Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 Keller Hall
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 12-011

A Data Mining Framework for Forest Fire Mapping

Anuj Karpatne, Xi Chen, Yashu Chamber, Varun Mithal, Michael Lau, Karsten Steinhaeuser, Shyam Boriah, Michael Steinbach, and Vipin Kumar

March 29, 2012

A Data Mining Framework for Forest Fire Mapping

Anuj Karpatne†, Xi C. Chen†, Yashu Chamber†, Varun Mithal, Michael Lau,
Karsten Steinhaeuser, Shyam Boriah, Michael Steinbach and Vipin Kumar
{anuj,chen,chamber,mithal,mwlau,ksteinha,sboriah,steinbac,kumar}@cs.umn.edu
University of Minnesota

Christopher S. Potter and Steven A. Klooster
chris.potter@nasa.gov,sklooster@gaia.arc.nasa.gov
NASA Ames Research Center

March 29, 2012

Abstract

Forests are an important natural resource that support economic activity and play a significant role in regulating the climate and the carbon cycle, yet forest ecosystems are increasingly threatened by fires caused by a range of natural and anthropogenic factors. Mapping these fires, which can range in size from less than an acre to hundreds of thousands of acres, is an important task for supporting climate and carbon cycle studies as well as informing forest management. There are two primary approaches to fire mapping: field- and aerial-based surveys, which are costly and limited in their extent; and remote sensing-based approaches, which are more cost-effective but pose several interesting methodological and algorithmic challenges. In this paper, we introduce a new framework for mapping forest fires based on satellite observations. Specifically, we develop spatio-temporal data mining methods for Moderate Resolution Imaging Spectroradiometer (MODIS) data to generate a history of forest fires. A systematic comparison with alternate approaches across diverse geographic regions demonstrates that our algorithmic paradigm is able to overcome some of the limitations in both data and methods employed by other prior efforts.

1 Introduction

Forest fires, which range in size from less than an acre to hundreds of thousands of acres, can be caused by both natural (such as lightning) or anthropogenic factors. Fires constitute a major component of terrestrial ecosystem disturbances every year, therefore accurate and low-cost fire mapping methods are important for understanding their frequency and distribution [26]. While monitoring fires in near-real time as they happen is critical for operational fire management, mapping historical fires in a spatially explicit fashion is also important for a number of reasons (recently highlighted in [16, 26]) including: (1) climate change studies – e.g., studying the relationship between rising temperatures and frequency of fires; (2) fuel load management – forest managers need a historical fire record when deciding where to conduct controlled burns; and (3) carbon cycle studies – quantifying how much CO_2 is emitted by fires is critical for emissions reduction efforts such as UN-REDD [32].

There are two primary approaches to mapping forest fires: field- and aerial-based studies, which allow detailed observation of land cover changes but are limited in their spatial extent and temporal frequency [20] because of their high cost; and remote sensing-based techniques, which offer the most cost-effective data for mapping fires because satellite observations—such as from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) instrument—are obtained globally with regular, repeated coverage. However, while numerous efforts have mapped fires at regional and local scales [5, 7, 10, 25, 29, 30], only two spatially explicit wall-to-wall efforts exist that regularly map fires at global scale: the MODIS Active Fire (AF) and

†These authors contributed equally to this work.

Burned Area (BA) products [18] (see Section 3). This discrepancy is largely due to a combination of data, algorithmic, and computational challenges associated with analyzing remote sensing data, as outlined below.

Fire mapping from remote sensing data is essentially a problem of change (or anomaly) detection. Such datasets have both temporal and spatial dimensions, and there are two primary ways to address the problem. On the one hand are approaches that focus on the temporal aspect, wherein fires are mapped based on time series analysis (e.g., [23, 28]). These types of methods usually take into consideration properties such as seasonality, variability and temporal coherence in a given time series. On the other hand are approaches that treat the data as a sequence of image snapshots, and image processing-based methods (e.g., [8, 17]) are used to detect burned areas. Such methods mainly leverage the spatial properties inherent in the data, for instance, the fact that burned pixels tend to cluster together. Recently, techniques have been developed for land cover change detection that utilize both spatial and temporal properties [15, 19, 22] to take advantage of autocorrelation structures present along *both* dimensions in the data.

Remote sensing data also poses several unique challenges for algorithm development, such as (1) the presence of noise and outliers, (2) inaccuracy and incompleteness of signals, (3) high natural variability and seasonality, (4) influence of climatic factors, (5) availability of multiple temporal scales. In the case of fire mapping, additional factors include potential obstruction of the signal due to smoke and the similarity of the signal relative to other types of changes and events.

In this paper, we introduce a new spatio-temporal data mining framework for forest fire mapping that is both robust and efficient. Specifically, the proposed approach is unsupervised in nature and exploits both temporal and spatial structure in the data to combine multiple sources of information. Using independently generated reference data for validation, we comprehensively evaluate our approach as well as alternate methods in a variety of geographies across different climatic zones. Finally, we examine specific properties of the BA algorithm and show how our approach is able to overcome some of its limitations.

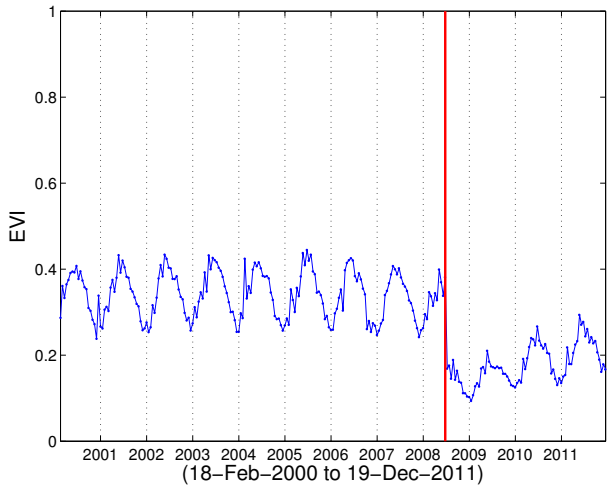
The remainder of the paper is organized as follows: In Section 2 we briefly describe the input datasets. Section 3 discusses related work, including the BA algorithm, and in Section 4 we introduce our new approach. Section 5 describes the experimental setup and validation data followed by results and discussion in Section 6. Section 7 provides concluding remarks and discusses directions for future work.

2 Data

Global remote sensing datasets are available from a variety of sources at different resolutions. The proposed fire mapping framework is based on two remotely sensed composite data products from the MODIS instrument aboard NASA's Terra satellite, which are available for public download [33]. Specifically, we use the Enhanced Vegetation Index (EVI) from the MODIS 16-day Level 3 1km Vegetation Indices (MOD13A2) and the Active Fire (AF) from the MODIS 8-day Level 3 1km Thermal Anomalies & Fire products (MOD14A2). EVI essentially measures "greenness" (area-averaged canopy photosynthetic capacity) as a proxy for the amount of vegetated biomass at a particular location (see Figure 1(a) for an example). AF is a basic fire product designed to identify thermal anomalies from the middle infrared spectral reflectance bands [18] and is used heavily in operational situations by fire-fighting agencies around the world. In order to separate forests from other land cover types, we use the MODIS Vegetation Continuous Fields (VCF) dataset (MOD44B), which provides the percent tree cover for every pixel. MODIS Level 3 products are provided on a global 1km sinusoidal grid in $10^\circ \times 10^\circ$ tiles. For this study, we focus on subsets of the data corresponding to geographical regions based on the available validation data (see Section 5).

3 Related Work

Fire-related data products broadly fall into two categories: *active fire* products, which capture the location and intensity of fires burning at the time of observation (the prototypical example being AF, see Section 2); and *burned area* products, which map areas that were burned by fires based on historical observations. In the following, we discuss two approaches for mapping burned areas, which require more sophisticated analysis methods and are the topic of this paper.



(a) A sample EVI time series from the fire.

(b) A photograph of the fire in progress. *Source: Associated Press*

Figure 1: The Basin Complex fire, which was started by lightning near Big Sur, CA in June 2008, consumed more than 160,000 acres before it merged with another fire and over $120M was spent fighting it.

3.1 The V2DELTA Algorithm

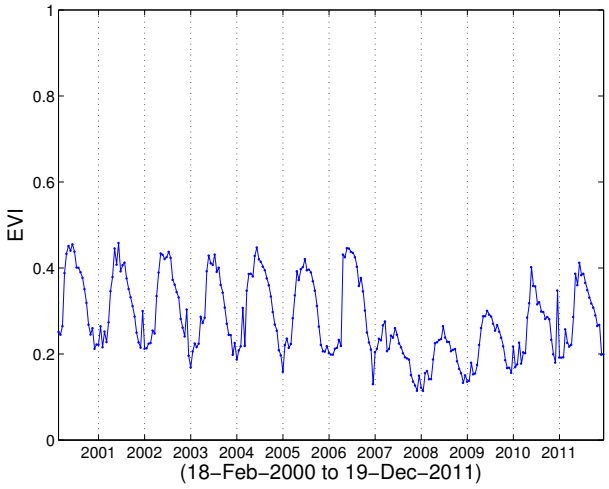
[24] presented a time series change detection algorithm that incorporates *natural variation* into the change detection framework. The algorithm, called V2DELTA, identifies abrupt forest disturbances using EVI as an input. More specifically, the V2DELTA algorithm compares a drop in EVI with the variability in a fixed "training" window, thereby providing a mechanism to ascribe significance to any given drop. This relies on the assumption that EVI values in the initial window were not affected by a land cover change, thus enabling the algorithm to differentiate abrupt changes from naturally occurring vegetation changes.

While V2DELTA identifies a broad class of disturbances [24], it fails at distinguishing fires from other land cover changes caused by climatic factors (e.g., droughts). Figure 2(a) shows a sample time series which was incorrectly identified as burned – the loss in EVI was due to logging. Moreover, the fire mapping task poses specific challenges that affect the algorithm's efficacy, especially for time series that exhibit high variability. Figure 2(b) shows one such example where a fire was not detected by V2DELTA due to variability and an irregular seasonal cycle.

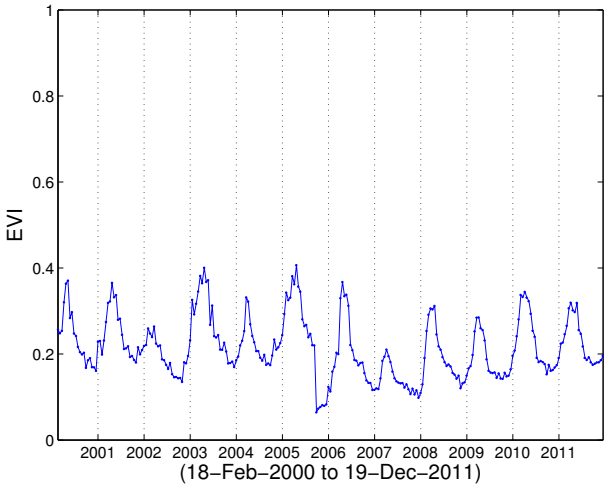
3.2 The BA Algorithm

The burned area approach (henceforth called BA) recently presented by [15] is a state-of-the-art methodology in the earth science research community for identifying regions burned by fire. The overall approach can be viewed as a semi-supervised Bayesian classification method with two classes: *burned* and *unburned*. The technique builds on key concepts and ideas developed over several years by [14] and others [11, 13, 21, 27]. The key steps of the algorithm are outlined below:

1. Representative sets of samples for the *burned* and *unburned* classes are constructed. The sample pixels for each class are discovered using conservative heuristics which label pixels as unburned or burned if they pass a set of conditions.
2. The *burned* class is further enriched with closely related pixels from the dataset, while the *unburned* class is refined by pruning pixels that are geographically close to *burned* training pixels.
3. A statistic that estimates the daily loss in vegetation (ΔVI) is computed for all training pixels.
4. The conditional probability distribution of the vegetation loss statistic is estimated for both the *burned* as well as the *unburned* class, i.e., $P(\Delta VI|burned)$ and $P(\Delta VI|unburned)$.
5. Bayes' Rule is applied to obtain the posterior probability of a pixel belonging to the *burned* class.



(a) A pixel in California incorrectly identified as burned by V2DELTA.



(b) A 2005 California fire event not detected as burned by V2DELTA because of high seasonal variability in the initial window.

Figure 2: Illustrative examples of limitations in V2DELTA.

There are two major limitations in the BA framework. First, the burned and unburned training data is very expensive to obtain manually, and is difficult to generate automatically. Second, due to the very nature of forest fires (they happen rarely in both spatial and temporal dimensions), the burned and unburned classes are often highly imbalanced. Since the Bayesian classifier is biased toward the majority class, the recall will suffer when the classes are imbalanced.

The BA algorithm is run on a regular basis using the latest spectral reflectance and AF inputs. The output is released by the University of Maryland as a product called MODIS Direct Broadcast Monthly Burned Area Product (`MCD64A1`).

4 Proposed Approach

Due to the nature of forest fires, there is inherently a huge class imbalance between the positive and negative classes. Owing to this fact, estimating two separate distributions for the rare and majority classes is likely to be affected by problems arising from small sample size and over-fitting. In contrast, a one-class approach of building a null distribution over the normal population and detecting rare events as anomalies in the distribution is more robust to noise and outliers as it utilizes a larger sample size. Note that "normal" here denotes the *unburned* population which may be unchanged or exhibit changes unrelated to fire (logging, drought, etc.). Our fire mapping framework is based on the idea of characterizing fire events as anomalies (or events forming a rare class) in relation to the distribution of unburned pixels (events forming the majority class). It is also necessary to ensure that true anomalies are excluded from the null distribution to avoid distorting it due to the rare anomaly class. To achieve this, we aim at creating a pool of candidate fire events and construct our null distribution by discarding any such candidates.

We present a stratified framework for mapping forest fires, where the multiple strata signify varying degrees of observability in the available datasets and confidence in detecting them. To obtain the highest stratum of detected fire events, we employ multiple complementary scoring mechanisms using both EVI and AF data. This stratum is expanded by including very similar events in close proximity to form the middle stratum. The lowest stratum is generated by including loosely similar events in a spatial window around the other two strata. Thus, our framework consists of the following two modular tasks: (1) selecting candidate fire events in three strata (all will be excluded from the null distribution), and (2) determining the anomaly scores of the lowest strata candidates using context-based anomaly detection. Figure 3 shows a flowchart of this approach; each task is described in detail below.

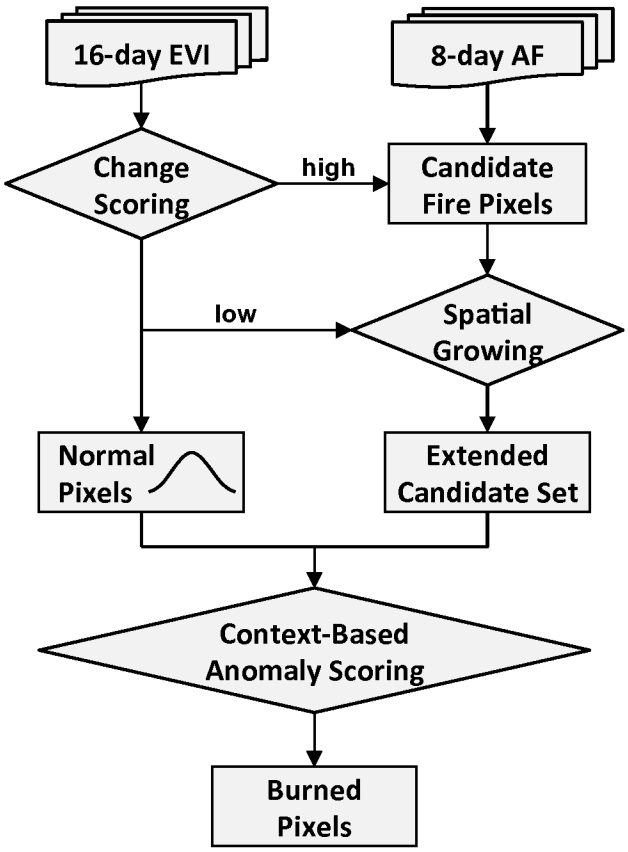


Figure 3: Flowchart illustrating the proposed framework for mapping forest fires.

4.1 Candidate Fire Detection

Forest fires lead to the burning of vegetation (trees and shrubs) and the emission of large amounts of thermal energy close to the land surface. Thus, a forest fire typically exhibits simultaneous changes (at a given location and time) in both the greenness signal (marked by a rapid decrease) as well as the thermal anomaly signal observed by satellite. We utilize these properties to generate candidate fire events, by performing an integrated analysis of abrupt changes in both the EVI and AF datasets. Furthermore, forest fire events exhibit particular spatio-temporal properties which can be harnessed to improve candidate fire detection. The process for detecting candidate forest fires consists of generating stratified sets of candidate pixels based on confidence of change exhibited in the EVI and AF signals.

4.1.1 Initial Candidate Pixels

There are certain vegetation characteristics associated with a forest fire which are manifested in the EVI signal. The algorithm begins operation by looking for such fire events using an initial set of candidates associated with the most confident stratum. Since AF captures information about thermal anomalies, it is arguably a good choice for initial candidates, providing a direction to search for candidate fires in its vicinity of space and time. It restricts the search space by pruning out potential false positive examples caused by other land cover changes such as conversion from forest to farm, or changes attributed to climatic factors such as droughts etc. Once the candidates have been initialized with events having high AF values (≥ 7), we use an array of scoring mechanisms on the EVI data to determine the significance of the given event as a forest fire. The intent behind using multiple scoring mechanisms is to cover multiple facets of information about the forest fire, where each score captures a distinct characteristic of the EVI signature at a forest fire event. Specifically, we introduce the following scoring mechanisms:

K-month Delta (KD) K-month Delta (KD) is an extension of the V2DELTA algorithm discussed in Section 3.1. Similar to V2DELTA, we model the inter-annual variation as a Gaussian distribution. The KD score at time step t is the z-score of the drop in EVI over K months after t as compared to the same K months in years before t. The key innovation of KD is that the mean and the variance are estimated based on the data in a 4 year window preceding t using *bootstrapping*, which is more robust than V2DELTA (V2DELTA uses a fixed window and measures inter-annual variation using fewer samples). K is set to 12 months in the proposed approach.

KD is a robust scoring mechanism that accounts for the natural variability present in the vegetation which is specific to the region under study and thus offers ample generalizability. However, KD underperforms in regions having vegetation types with high variability, since KD discounts the K-drop with the high variability in such regions, a limitation it inherits from its predecessor algorithm V2DELTA. In addition, a typical signature of forest fires is instantaneous loss in vegetation (regardless of the variability in the time series), but KD may not capture some abrupt drops because of the K-month look-ahead time. These limitations can be overcome by integrating KD score with other scoring mechanisms which serve orthogonal functionalities of being local and instantaneous in nature.

Local Instant Drop (LID) Forest fires occur as instantaneous drops in the vegetation, most often observed in consecutive EVI observations, and do not persist for a long time period. In addition, the loss in EVI due to a forest fire is generally much higher than the normal variations in the EVI signal attributed to seasonality and noise. Local Instant Drop (LID) leverages these local characteristics about the instantaneous nature of forest fire events. Specifically, LID normalizes the instantaneous drop with fluctuations around the drop in a small temporal neighborhood to obtain the LID score. Furthermore, LID accounts for noise and outliers occurring in the temporal locality of a candidate fire, as well as the seasonal context in EVI data to improve the robustness of the scoring algorithm.

Even though LID provides an efficient mechanism for detecting forest fires, it relies on the assumption that fires exhibit single-time step drops in EVI, which are distinct from other fluctuations in their temporal vicinity. Since EVI is a 16-day product, errors in registration and EVI quality issues may sometimes result in a forest fire drop extending its influence over more than two consecutive EVI values. Moreover, LID is affected by large noisy fluctuations in the EVI values around a candidate fire event, since it discount the instantaneous drop with temporally local variations. This demonstrates the need to couple LID with other scoring mechanisms which are robust to sporadic fluctuations in the temporal neighborhood of a candidate fire event.

Near Drop (ND) Near Drop (ND) simply measures the change in the average EVI value over n ($n = 3$) time steps, before and after a candidate forest fire event. It captures the instantaneous nature of the drop in the EVI which persists for more than one time step. Furthermore, averaging reduces the effect of random fluctuations in the temporal neighborhood of a candidate fire, and hence complements the shortcomings of LID. However, since ND is easily affected by noise and outliers, it is well-suited to be a necessary filter condition to be satisfied by a candidate fire event detected by other scoring mechanisms.

Integrating Multiple Scoring Mechanisms It is evident from the previous discussion that each scoring mechanism has its own specific limitations in detecting forest fires. However, it can be observed that the scoring mechanisms serve complementary purposes and capture multiple characteristics of the fire event, which are distinct in nature. Hence, they offer us a possibility for developing a fire detection framework which utilizes the orthogonal aspects of each scoring mechanism in an integrative manner. Since a forest fire exhibits large instantaneous drop, a large ND score (≥ 500) is a filtering criterion that must be met by all candidate forest fires. A large KD score (≥ 4) is representative of a forest fire event when coupled in conjunction with a moderate LID score (≥ 1), which helps in rejecting other land cover changes that show high KD score but are not associated with fires (low LID score). A large LID score (≥ 4) is in itself a good indicator of forest fire events. Events which satisfy these scoring criteria are considered as candidate forest fires (**highest stratum**) arising from the initial choice of candidates and are further used in subsequent steps of our fire mapping framework.

4.1.2 Spatial Growing

The AF signal often fails to detect forest fire events which do not register a thermal anomaly because of smoke or satellite overpass timing. Thus, the initial candidate pixels might suffer from low coverage. To overcome this challenge, we exploit the inherent spatio-temporal autocorrelation of forest fire events to increase coverage. Since events corresponding to the same forest fire occur in close proximity of space and time, we exploit this property by searching for candidate fire events around the initial candidates classified as forest fires by the scoring mechanism above. In the current framework, we consider the 24 spatial neighbors in a 5×5 spatial grid around the initial candidates, with a temporal constraint of being within one time step from the change time of the initial candidate fire event. We then apply our scoring mechanism on the new pool of candidate events with exactly the same scoring criteria as we used for detecting initial candidate forest fire events. We iteratively grow in a spatial neighborhood to exhaustively detect candidate forest fire events. They represent candidate forest fire events (**middle stratum**) which have sharp fire characteristics in the EVI signal but were not initial candidates because of the absence of AF signal.

Because of the presence of noise in the EVI signal as well as cases where EVI loss is small, there are a number of forest fire events which do not exhibit strong characteristics of fire in our scoring mechanisms and thus go undetected. However, we can exploit the autocorrelation property of forest fires to discover such events.

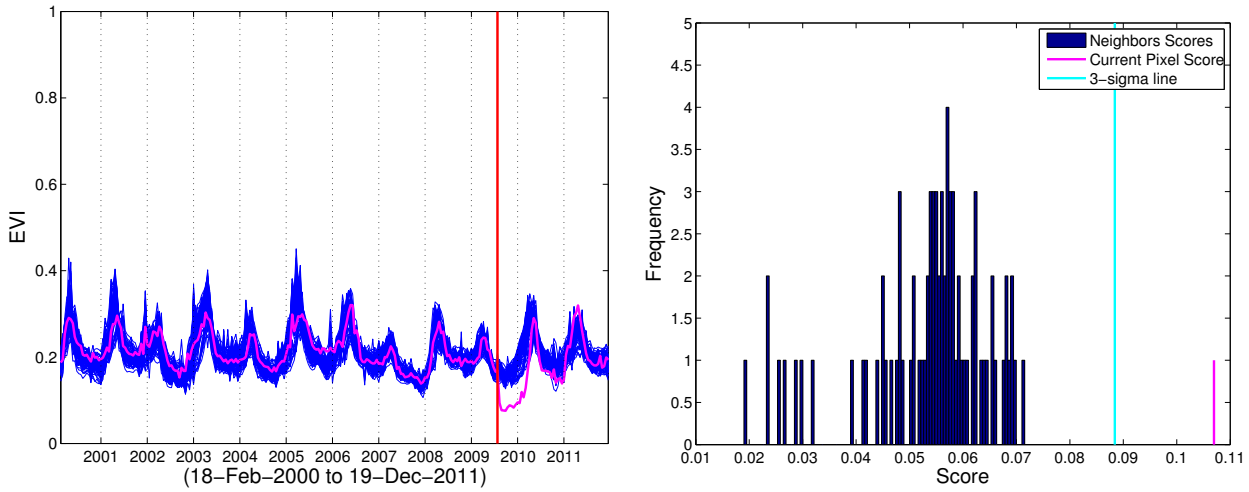
We exploit this property to create a pool of candidate forest fire events with a relaxed scoring criteria indicating a lower confidence. We accept events to be part of this pool (**lowest stratum**) if they exhibit a positive ND score and either a moderately large LID score (≥ 1.5), or a moderately large KD score (≥ 1.5) in conjunction with a moderate LID score (≥ 1). Thus, we iteratively grow in a spatial neighborhood (5×5 grid) to exhaustively include any probable candidate fire events. This improves our coverage of forest fires but is likely to include false positives, as the lowest stratum may contain spurious events detected due to the presence of noise. Such spurious events are targets for removal by the context-based anomaly detection scheme.

4.2 Context-based Anomaly Detection

Since forest fires form a rare class in the complete dataset, they exhibit anomalous behavior compared to the null distribution built with similar vegetation types. Hence, events from the lowest stratum that are similar to the null distribution do not obtain a high anomaly score and are pruned. Thus, we employ a context-based anomaly detection scheme on each of the lowest stratum candidate fire events to discriminate fire events from spurious events. We utilize the EVI information of the candidate fire over a time period of 3 years before its detected time, by performing a time series similarity search in a large spatial neighborhood (81×81). We utilize a tight similarity search criterion (pairwise distance ratio ≤ 0.2 at every time step, after accounting for not more than 5 outliers) and a large spatial neighborhood to ensure a rich sample of unburned locations in the null distribution, that are not spatially autocorrelated with the candidate fire event and which show similar vegetation characteristics. Further, the similarity search is only performed over EVI values which are high quality and are free from smoke and clouds, so as to avoid outliers from affecting the similarity measure.

In order to populate the null distribution with only normal (unburned) events, we exclude every detected candidate fire from consideration while building the null distribution. Also, since forest fires exhibit spatio-temporal autocorrelation, locations in the immediate vicinity of the candidate fire event under consideration may be potentially burned (even if not contained in the lowest stratum) and thus may adversely affect the null distribution as they are likely to exhibit similar EVI characteristics as the candidate fire event. Hence, a spatial buffer (of size 40×40) around the candidate fire event is excluded while building the null distribution. Figure 4(a) shows a candidate fire time series and the spatio-temporal context obtained by similarity search.

The anomaly score of the candidate fire event is then computed by considering a statistic which discriminates forest fire events from normal events. The ideal statistic should be instantaneous in nature and offer high separability between the fire events and unburned events. The ND score computed over EVI values free from marginal observation serves this purpose in measuring the instantaneous separability of fire events from similar unburned events. Thus, we employ ND score (with $n = 6$ time steps) as the distribution statistic. A candidate fire event in the lowest stratum is then pruned if it fails to manifest an ND score which is greater than $\mu + 3\sigma$, where μ and σ are the mean and standard deviation of the null distribution. The set of events obtained after pruning non-anomalous events from the candidate set is considered to be the final set of forest



(a) A candidate EVI time series for pruning (magenta) along with its spatial (normal) neighbors.

(b) The distribution of the ND scores of candidate and its neighbors.

Figure 4: Example of context-based anomaly detection.

fire events detected by our algorithm. Figure 4(b) shows an example of a candidate fire event in relation to the null distribution; since the candidate lies at least $+3\sigma$ from μ, it was correctly retained as a fire pixel.

5 Evaluation Setup

We examine the performance of AF, BA and the proposed approach in several regions around the world, including the states of California (United States), Georgia (United States), and Victoria (Australia). These geographic areas represent diverse regions with widely differing variability, land cover types, geography and noise characteristics. The following describes the validation data used in this study and provides a brief overview of the evaluation methodology.

5.1 Validation Data

For each region considered in our evaluation, we obtained fire validation data from government agencies responsible for monitoring and managing forests and wildfires. The validation data is in the form of fire perimeter polygons, each of which is associated with the time of burning. Table 1 summarizes the regions studied in this paper and the respective sources of validation data. Although government agencies make their best effort in documenting historical fires, fire perimeter datasets are neither complete nor without error due to finite resources available to any agency. However, inaccuracies and incompleteness are represented only in a small portion of the validation data, and these datasets are still useful for quantitatively comparing methods across large spatial regions. The AF, BA and EVI datasets are georeferenced by the latitude and longitude values of their pixel centers. We consider an event to be positive if the corresponding pixel lies inside a polygon. Similarly, an event is considered to be unburned (forming the negative class) only if the entire pixel is outside a polygon. The remaining pixels (which partially overlap polygon boundaries) are discarded from the evaluation framework to avoid ambiguity.

Since our primary focus is on detecting forest fires, we utilize the MODIS Vegetation Continuous Fields (VCF) dataset (MOD44B) which contains the percentage tree cover information. We only consider pixels with high percentage tree cover (≥ 20) in our evaluation scheme, a procedure commonly used in the earth science domain to separate forests from non-forests [6, 15].

Region	References	Positives	Negatives
California (US)	[1, 9]	4597	597208
Georgia (US)	[2, 4]	2003	425528
Victoria (Australia)	[3, 12]	17190	604391

Table 1: Regions studied in this paper and their respective sources of historical fire validation data.

		Predicted	
		Fire	No Fire
Validation Data	Fire	TP	FN
	No Fire	FP	TN

Table 2: Confusion matrix.

5.2 Evaluation Methodology

In this paper, we use *precision* and *recall* as evaluation metrics for quantitatively comparing the performance of AF, BA and the proposed approach. These two well-known metrics are used to evaluate the performance of algorithms in information retrieval, machine learning and data mining [31]. Each algorithm provides a set of positive and negative events that it detects, which is validated using fire perimeter polygons to obtain the number of true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN) for each algorithm, as shown in Table 2. Note that a TP event means that the pixel lies inside a polygon and the time of change agrees with the polygon. The precision and recall values for each algorithm are then given by:

$$\text{Precision, } p = \frac{TP}{TP + FP} \quad \text{Recall, } r = \frac{TP}{TP + FN}$$

6 Experimental Results & Discussion

Tables 3, 4 and 5 contain validation results of our proposed fire mapping framework along with other algorithms discussed in this paper for the three different geographic regions. In the following we discuss their performance as well as relative strengths and weaknesses in detail, beginning with some general observations.

First, we note that our method initially achieves high precision by using strict criteria (rows 1-2), then increases recall through spatial growing using lower thresholds (row 3), and finally recovers the majority

Algo	Precision	Recall	TP	FP	FN
L1	0.994	0.817	3754	21	843
L2	0.994	0.894	4111	23	486
L3	0.781	0.957	4399	1236	198
L3 pruned	0.985	0.936	4302	65	295
BA highQ	0.983	0.929	4270	73	327
BA lowQ	0.981	0.929	4270	81	327
AF7	0.593	0.915	4208	2884	389
V2DELTA	0.595	0.833	3828	2602	769

Table 3: Evaluation results for California.

Algo	Precision	Recall	TP	FP	FN
L1	0.953	0.437	875	43	1128
L2	0.964	0.631	1263	47	740
L3	0.723	0.817	1637	626	366
L3 pruned	0.957	0.718	1439	64	564
BA highQ	0.468	0.193	386	438	1617
BA lowQ	0.699	0.674	1350	581	654
AF7	0.146	0.645	1291	7530	712
V2DELTA	0.293	0.514	1030	2491	973

Table 4: Evaluation results for Georgia.

Algo	Precision	Recall	TP	FP	FN
L1	0.997	0.376	6472	17	10718
L2	0.997	0.676	11614	38	5576
L3	0.784	0.805	13834	3814	3356
L3 pruned	0.988	0.709	12191	152	4999
BA highQ	0.993	0.693	11908	85	5283
BA lowQ	0.993	0.697	11977	90	5213
AF7	0.844	0.533	9168	1691	8022
V2DELTA	0.855	0.629	10815	1836	6375

Table 5: Evaluation results for Victoria.

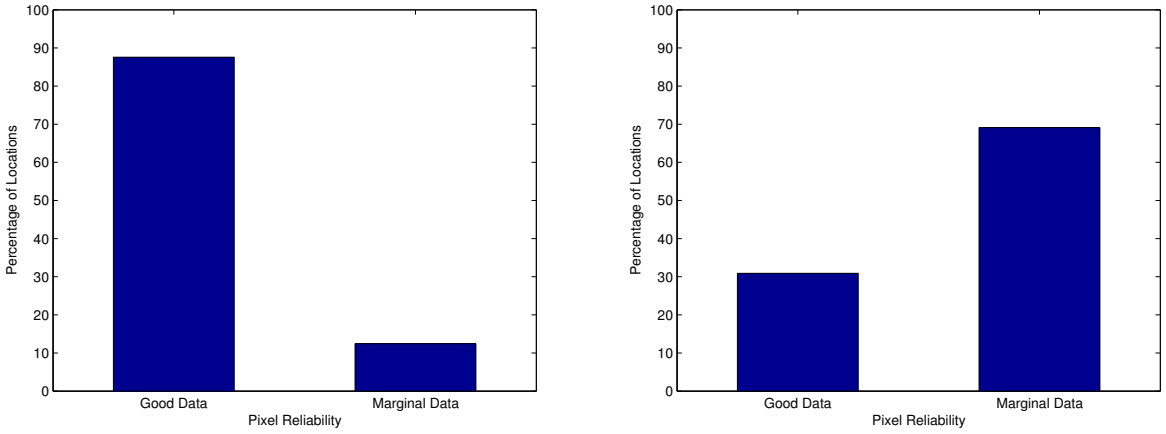
of the precision by discarding false positives using an anomaly detection approach (row 4). These results demonstrate that the general paradigm represented by our framework (Section 4) behaves as expected.

Second, by comparing the fourth and eighth row in Tables 3-5, we find that our framework consistently outperforms the V2DELTA algorithm [24]. Although V2DELTA achieves comparable recall in some cases, it tends to generate a large number of false positives which is not surprising given that it is a more general change detection method. In fact, it was this shortcoming that prompted the development of the bootstrapping procedure for more robust modeling of the historical EVI patterns and incorporation of the AF signal as an additional source of information. This point is illustrated in Figure 2(a), which shows the time series for a pixel that was incorrectly identified by V2DELTA as burned. However, with a combination of robust time series modeling and the help of AF, our new framework is able to correctly dismiss such pixels as unburned.

6.1 Qualitative and Quantitative Comparison with Burned Area

Looking at rows four and five in Tables 3-5, we note that our proposed framework generally performs comparable to, and in some cases better than, the BA algorithm. However, BA should not be viewed as a competitor but rather as an alternate approach – both methods use different datasets and have (complementary) strengths and weaknesses. For instance, a close examination of both false positive and false negative pixels reveals that the two approaches often make different mistakes in the same geographical region.

Two particular limitations of the BA algorithm are illustrated by the results for Georgia (Table 4). (1) We note that AF has extremely low precision in Georgia; since BA depends on AF to generate training data, we hypothesize that its performance is significantly diminished in cases where the precision of AF is poor. (2) We observe a peculiar case where both precision and recall are *higher* for the "low-quality" BA product. We postulate that this counterintuitive observation can be explained by the fact that values in the immediate (temporal) neighborhood of a fire tend to be of lower quality (due to obstruction from smoke); since BA only uses high quality observations within a limited window surrounding the fire date, the algorithm does



(a) TP in BA highQ

(b) FN in BA highQ but TP in BA lowQ

Figure 5: A comparison of histograms for pixel reliability of BA suggests that the algorithm is sensitive to data quality.

Data-set	Precision (3σ)	Recall (3σ)
AF + Noise(N)	0.9936 ± 0.0016	0.8166 ± 0.0006
	0.9936 ± 0.0015	0.8754 ± 0.0003
	0.7479 ± 0.0786	0.9571 ± 0.0011
AF + Noise(5N)	0.9933 ± 0.0015	0.8165 ± 0.0008
	0.9934 ± 0.0016	0.8754 ± 0.0008
	0.7558 ± 0.0631	0.9571 ± 0.0011
AF + Noise(10N)	0.9935 ± 0.0015	0.8165 ± 0.0006
	0.9935 ± 0.0017	0.8753 ± 0.0004
	0.7411 ± 0.0849	0.9570 ± 0.0004

Table 6: Performance of the proposed framework in California given "noise" in the form of additional (false positive) AF signals. Each group of three rows corresponds to the top three rows in Table 3.

not have sufficient information and is thus unable to correctly classify such pixels. To test this hypothesis, we generated two histograms (Figures 5(a) and 5(b)) of the pixel reliability extracted from the Quality Assurance (QA) fields: the first is for the true positives of the high-quality product (left panel), and the second for the false negatives of the high-quality product which are true positives in the low-quality version (right panel). Indeed, we find that the events BA highQ correctly identified were by and large of high quality, while BA lowQ was able to take advantage of the additional information from lower-quality inputs and thus find a significantly larger number of burned pixels.

6.2 Robustness to Active Fire

The AF algorithm is essentially an anomaly detection scheme applied to the thermal band measured by MODIS; as such, it is intended to have high recall but not necessarily high precision. From our results, we further note that the overall performance of BA is tied closely to that of AF. For example, in California and Victoria (Tables 3 and 5, respectively) AF performs quite well, and so does BA. However, in Georgia (Table 4) AF incurs 7530 false positives resulting in very poor precision, which in turn severely limits the performance of BA.

In contrast, although our framework also takes the AF signal as an input, the spatial growing is able to

overcome its limitations. To better understand the effect of AF on our approach, we conducted an additional set of experiments wherein we randomly inserted varying levels of false positives into the AF signal, with the number of instances being added ranging from one to ten times the number of actual AF instances; each experiment was repeated 30 times. The results, shown in Table 6, are nearly identical to the main result (Table 3) and suggest that our framework is extremely robust to the presence of false positive AF signals, unlike BA as seen in Table 4.

6.3 Limitations of the Proposed Approach

The forest fire mapping framework proposed in this paper faces limitations in a number of scenarios, leading to both false positives and false negatives. These include situations where (1) the vegetation rapidly recovers after a fire or if there are multiple fires in short succession (FN), (2) the loss in vegetation is insignificant (FN), and (3) the vegetation has high natural variability (FP and FN). Each of these scenarios poses distinct challenges for our current fire detection framework. Additionally, if a fire polygon does not contain any pixel with an AF signal, it will not be detected by our framework (FN). However, we observed that such instances happen only in small polygons, and hence its effect on the performance of the proposed framework is insignificant.

We briefly present a discussion of the false positives and false negatives of the proposed approach. We had observed that a number of pixels exhibiting strong fire characteristics were missed by the fire perimeter polygons in Victoria and California (Figure 6(a)). Such pixels were incorrectly counted as false positives in our evaluation scheme affecting the performance. In addition, we encountered a small number of false positives which exhibited other types of land cover changes (Figure 6(b)), and were incorrectly included in the highest stratum of detected points. This limitation can be overcome by applying context-based anomaly detection on the entire spectrum of detected events, not just the lowest strata.

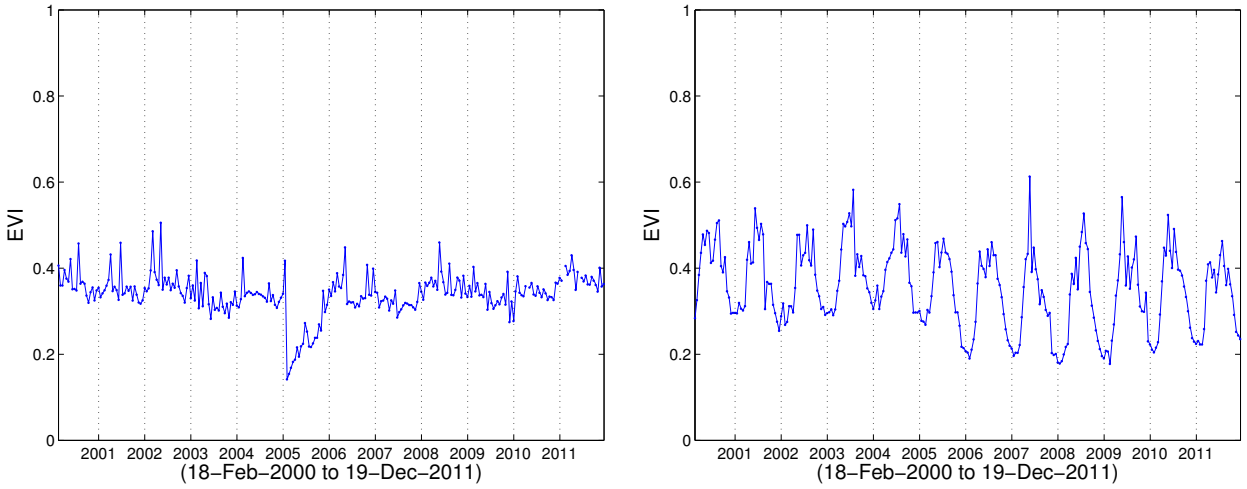
Most false negatives of the proposed approach consist of vegetation types with high natural variability and noisy EVI data, which got poorly scored by the proposed algorithms and went undetected (Figure 7(a)). Other false negatives can be attributed to weak characteristics of fire in their EVI signal, e.g. in California (Figure 8(a)) and Georgia (Figure 8(b)). Some other false negatives in Victoria exhibited atypical characteristics in their EVI signal which appeared spurious and were not detected by the proposed approach (Figure 7(b)). In addition, some of the false negatives were pruned during the anomaly detection phase due to lack of adequate number of similar (unburned) neighbors required to construct the normal distribution. This poses a challenge in performing context-based anomaly detection when the number of similar objects is small, thus offering a future direction for work.

We are in effect considering *less information* than BA, yet we are able to achieve comparable performance. Thus, there is reason to believe that some of the limitations above will be addressed by increasing the spatial and temporal resolution to 250m and daily, respectively. In particular, we expect that smaller fires (250m data has sixteen times more detail than 1km data) and pixels that exhibit rapid recovery (because of compositing, neighboring time steps can be up to a month apart in the current 16-day data) can be detected with higher resolution data. We have not used higher resolution data in this paper since these are not standard MODIS products and hence require extensive processing to generate.

7 Conclusion & Future Work

In this paper, we proposed a data mining framework for forest fire mapping. The proposed technique is unsupervised in nature and has the potential to be used globally, providing spatially explicit wall-to-wall coverage. We quantitatively showed that the technique performs comparably to the well-known BA algorithm in the states of California (US) and Victoria (Australia) and much better in Georgia (US); there are also complementarities between the two frameworks. We also showed that the proposed framework is highly robust to noise in one of its primary inputs, AF, which is known to have low precision.

While the current framework already performs relatively well in a variety of geographies, there are a number of interesting directions for future work. The data inputs used in this paper, EVI and AF, have temporal resolutions of 16 and 8 days, respectively. Both of these inputs can be computed on a daily basis, which is likely to provide much more information in many cases. This information can be exploited to identify the precise day of the fire and to ensure temporal coherence between neighboring pixels. Challenges



(a) Missed fire in Victoria.

(b) Land cover change in Georgia.

Figure 6: False positives of the proposed approach.

with daily data include increase in the noise level and additional effort required to generate a daily EVI product (since this is not a standard MODIS product). Another potential extension is to use BA (and similar products) as an *input*, taking advantage of complementarities that exist between the frameworks. Finally, the anomaly scoring that is currently applied to the lowest stratum events can also be extended to the middle and highest stratum to further increase the precision in these strata.

References

[1] California Department of Forestry and Fire Protection, Fire and Resource Assessment Program. `http://frap.fire.ca.gov`.

[2] USDA Forest Service and US Geological Survey, Monitoring Trends in Burn Severity Project, Burned Area Boundaries Dataset. `http://mtbs.gov`.

[3] Department of Sustainability and Environment, Victoria, Australia. Fire History Records of Fires on Public Land. `http://www.dse.vic.gov.au`.

[4] National Fire Decision Support Center, Wildland Fire Decision Support System. National Fire Perimeters. `http://wfdss.usgs.gov`.

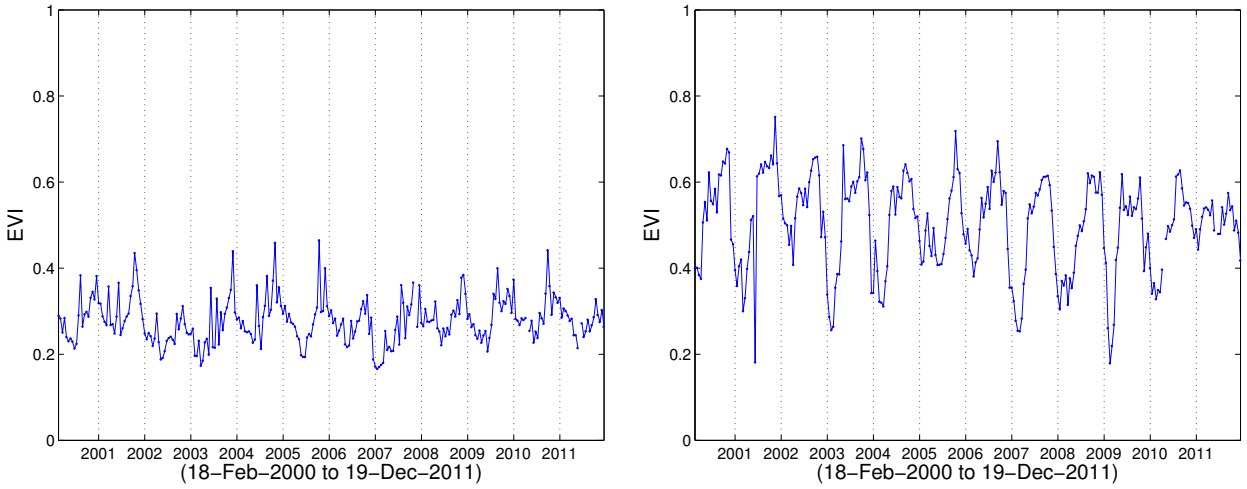
[5] K. V. S. Badarinath, A. R. Sharma, and S. K. Kharol. Forest fire monitoring and burnt area mapping using satellite data: a study over the forest region of Kerala state, India. *International Journal of Remote Sensing*, 32(1):85–102, 2011.

[6] S. Bandyopadhyay, P. Shyamsundar, and A. Baccini. Forests, biomass use and poverty in Malawi. *Ecological Economics*, 70(12):2461–2471, 2011.

[7] Y. Bergeron, S. Gauthier, M. Flannigan, et al. Fire regimes at the transition between mixedwood and coniferous boreal forest in northwestern Quebec. *Ecology*, 85(7):1916–1932, 2004.

[8] C. K. Brewer, J. C. Winne, R. L. Redmond, et al. Classifying and mapping wildfire severity: A comparison of methods. *Photogrammetric Engineering & Remote Sensing*, 71(11):1311—1320, 2005.

[9] California Department of Forestry and Fire Protection. Fire management for California ecosystems, 1995.



(a) Vegetation with high variability (fire occurred in 2007).

(b) Atypical characteristic of fire (fire occurred in 2009).

Figure 7: False negatives in Victoria.

[10] M. Cochrane. *Tropical fire ecology: climate change, land use and ecosystem dynamics*. Springer, 2009.

[11] R. H. Fraser, Z. Li, and J. Cihlar. Hotspot and NDVI Differencing Synergy (HANDS): a new technique for burned area mapping over boreal forest. *Remote Sensing of Environment*, 74(3):362–376, 2000.

[12] G. R. Friend, M. Leonard, S. Troy, et al. Fire and biodiversity management in Victoria–integrating the science, planning and implementation process. In *Proceedings of the 3rd International Wildland Fire Conference*, pages 3–6, 2003.

[13] C. George, C. Rowland, F. Gerard, et al. Retrospective mapping of burnt areas in central Siberia using a modification of the normalised difference water index. *Remote Sensing of Environment*, 104(3):346–359, 2006.

[14] L. Giglio, G. R. van der Werf, J. T. Randerson, et al. Global estimation of burned area using MODIS active fire observations. *Atmospheric Chemistry and Physics*, 6(4):957–974, 2006.

[15] L. Giglio, T. Loboda, D. Roy, et al. An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sensing of Environment*, 113(2):408–420, 2009.

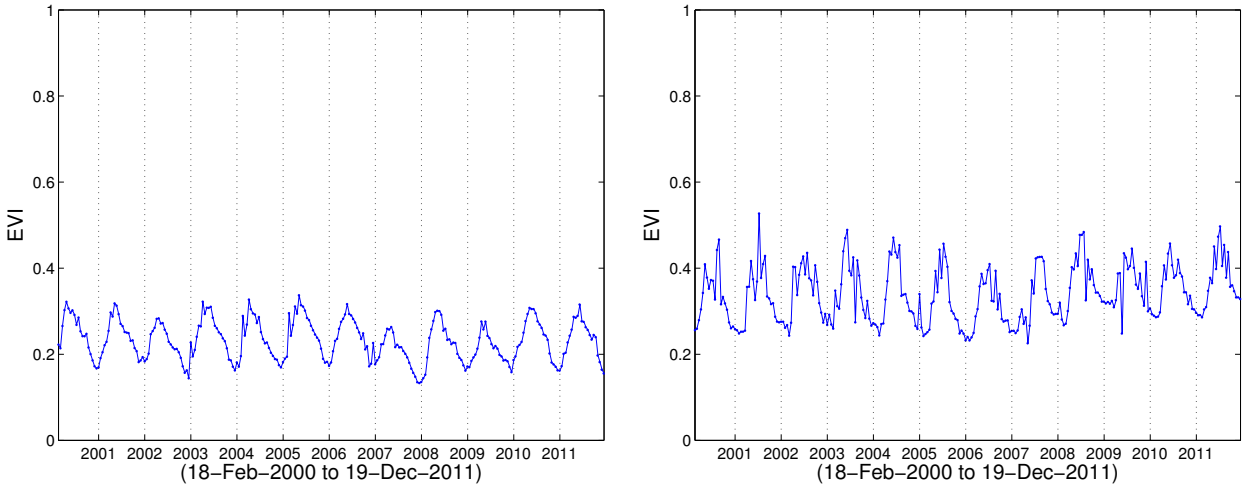
[16] J. Gillis. With deaths of forests, a loss of key climate protectors. *The New York Times*, October 1 2011.

[17] I. Z. Gitas, G. H. Mitri, and G. Ventura. Object-based image classification for burned area mapping of Creus Cape, Spain, using NOAA-AVHRR imagery. *Remote Sensing of Environment*, 92(3):409–413, 2004.

[18] C. O. Justice, L. Giglio, D. Roy, et al. MODIS-derived global fire products. In B. Ramachandran, C. O. Justice, and M. J. Abrams, editors, *Land Remote Sensing and Global Environmental Change*, pages 661–679. Springer, 2011.

[19] S. Lhermitte, J. Verbesselt, I. Jonckheere, et al. Hierarchical image segmentation based on similarity of NDVI time series. *Remote Sensing of Environment*, 112(2):506–521, 2008.

[20] D. Liu and S. Cai. A spatial-temporal modeling approach to reconstructing land-cover change trajectories from multi-temporal satellite imagery. *Annals of the Association of American Geographers*, 2011.



(a) FN in California (fire occurred in 2008).

(b) FN in Georgia (fire occurred in 2007).

Figure 8: False negatives due to weak fire characteristics.

[21] T. Loboda, K. O'Neal, and I. Csiszar. Regionally adaptable dNBR-based algorithm for burned area mapping from MODIS data. *Remote Sensing of Environment*, 109(4):429–442, 2007.

[22] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, et al. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment*, 105(2):142–154, 2006.

[23] V. Mithal, A. Garg, S. Boriah, et al. Monitoring global forest cover using data mining. *ACM Transactions on Intelligent Systems and Technology*, 2:36:1–36:24, 2011.

[24] V. Mithal, A. Garg, I. Brugere, et al. Incorporating natural variation into time series-based land cover change detection. In *Proceedings of the 2011 NASA Conference on Intelligent Data Understanding (CIDU)*, pages 45–59, 2011.

[25] M. Niklasson and A. Granström. Numbers and sizes of fires: Long-term spatially explicit fire history in a Swedish boreal landscape. *Ecology*, 81(6):1484–1499, 2000.

[26] Y. Pan, R. A. Birdsey, J. Fang, et al. A large and persistent carbon sink in the world's forests. *Science*, 333(6045):988–993, 2011.

[27] D. P. Roy, L. Giglio, J. D. Kendall, et al. Multi-temporal active-fire based burn scar detection algorithm. *International Journal of Remote Sensing*, 20(5):1031–1038, 1999.

[28] D. P. Roy, P. E. Lewis, and C. O. Justice. Burned area mapping using multi-temporal moderate spatial resolution data–a bi-directional reflectance model-based expectation approach. *Remote Sensing of Environment*, 83(1-2):263–286, 2002.

[29] R. Somashekar, P. Ravikumar, C. Mohan Kumar, et al. Burnt area mapping of Bandipur National Park, India using IRS 1C/1D LISS III data. *Journal of the Indian Society of Remote Sensing*, 37:37–50, 2009.

[30] B. Talon, S. Payette, L. Filion, et al. Reconstruction of the long-term fire history of an old-growth deciduous forest in Southern Québec, Canada, from charred wood in mineral soils. *Quaternary Research*, 64(1):36–43, 2005.

[31] P.-N. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison-Wesley, Boston, MA, 2006.

[32] United Nations. Collaborative Programme on Reducing Emissions from Deforestation and Forest Degradation in Developing Countries. http://www.un-redd.org/.

[33] US Geological Survey and NASA. Land Processes Distributed Active Archive Center (LP DAAC). https://lpdaac.usgs.gov/.