ISPRS Journal of Photogrammetry and Remote Sensing 144 (2018) 217–234

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Towards a polyalgorithm for land use change detection

Rishu Saxena[a,*], Layne T. Watson[b], Randolph H. Wynne[c], Evan B. Brooks[c], Valerie A. Thomas[c], Yang Zhiqiang[d], Robert E. Kennedy[e]

[a] Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[b] Departments of Computer Science, Mathematics, and Aerospace and Ocean Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[c] Department of Forest Resources and Environmental Conservation, Blacksburg, VA 24061, USA
[d] Forest Ecosystems & Society, Oregon State University, Corvallis, OR 97331, USA
[e] College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331, USA

ARTICLE INFO

Keywords:
Time series analysis
Remote sensing
Change detection
Scalable computing
Polyalgorithm

ABSTRACT

One way of analyzing satellite images for land use and land cover change (LULCC) is time series analysis (TSA). Most of the many TSA based LULCC algorithms proposed in the remote sensing community perform well on datasets for which they were designed, but their performance on randomly chosen datasets from across the globe has not been studied. A polyalgorithm combines several basic algorithms, each meant to solve the same problem, producing a strategy that unites the strengths and circumvents the weaknesses of constituent algorithms. The foundation of the proposed TSA based 'polyalgorithm' for LULCC is three algorithms (BFAST, EWMACD, and LandTrendR), precisely described mathematically, and chosen to be fundamentally distinct from each other in design and in the phenomena they capture. Analysis of results representing success, failure, and parameter sensitivity for each algorithm is presented. For a given pixel, Hausdorff distance is used to compare the distance between the change times (breakpoints) obtained from two different algorithms. Timesync validation data, a dataset that is based on human interpretation of Landsat time series in concert with historical aerial photography, is used for validation. The polyalgorithm yields more accurate results than EWMACD and LandTrendR alone, but counterintuitively not better than BFAST alone. This nascent work will be directly useful in land use and land cover change studies, of interest to terrestrial science research, especially regarding anthropogenic impacts on the environment.

1. Introduction

Land use change is described as changes in how humans use the surface of the Earth (e.g., for agriculture, plantations, pastures, managed woods, conservation, settlements, or leaving it alone as natural ecosystem). Changes in land use lead to changes in albedo, thereby directly affecting the temperatures of the surrounding area. Significant and lasting changes in land use and land cover (LULC) have more profound effects. The past century has seen an exponential growth in human activities such as deforestation and urbanization causing significant changes in land cover in several parts of the world (Hansen et al., 2013). Simultaneously, significant changes in the global climate have also been observed, driven in part by LULC change (LULCC) (e.g., Fall et al., 2010). LULCC also has impacts on a wide variety of other ecosystem services. Monitoring LULCC across the globe, therefore, has become the need *du jour*. Land use change detection comprises any methodology used for determining the occurrence and nature of change in LULC.

Earth observation satellites (EOS) such as Landsat capture images of the Earth's surface at regular intervals using multiple spectral frequencies. These images hold valuable information that, if harnessed well, can be immensely helpful in understanding, monitoring, and managing our natural resources, as well as studying LULCC. One way of analyzing these satellite images for LULCC studies is time series analysis (or, temporal trajectory analysis). For time series analysis, several images of the scene under consideration, taken over a period of time, are stacked together chronologically and subsequently analyzed. Commonly, the time series for each pixel is treated individually; the full image stack is thus a collection of many time series. The choice of spectral band(s) varies from application to application. The objective is to discover a 'trend' in how different relevant variables (indicators) evolve over time. In change detection analysis, when the trajectory of one or more of the variables departs from the normal, a change is detected. Time series analysis for LULCC studies has been receiving increasing attention in the last decade, specifically, after the Landsat data

* Corresponding author.
E-mail address: rishus@vt.edu (R. Saxena).

https://doi.org/10.1016/j.isprsjprs.2018.07.002
Received 11 January 2018; Received in revised form 29 June 2018; Accepted 6 July 2018

became freely accessible in 2008 (Woodcock et al., 2008). Several time series analysis algorithms have been proposed by different groups in the remote sensing community.

Despite a plethora of time series analysis algorithms available in remote sensing, design and selection of algorithms for LULCC detection in remote sensing appears to be almost always context specific. Most of the methods proposed to date seem to perform well on the type of data that they are designed for. Their performance on randomly picked datasets from across the globe has not been studied. The onus of choosing an appropriate algorithm that will perform well on their particular dataset falls on the user. Unfortunately, no single algorithm designed so far seems to work for all datasets (Cohen et al., 2017). For example, the Western Antarctica as well as the Greenland Ice Sheets are beginning to collapse due to global warming, the melting leading to continually receding snow covers at the respective locations. For these regions, using LULC algorithms based on periodicity assumptions is expected to lead to incorrect predictions and/or false alarms, although the nature and extent of this has not been studied yet. Even if there were no global warming, mild shifts in the 'phase' and 'amplitude' of seasons are known to take place (Petitjean et al., 2011). Time warping techniques (Petitjean et al., 2011) to deal with these issues may be helpful in some contexts, but their accuracy and scalability has not yet been satisfactorily investigated. Approaches based on periodicity and a moving window are possible, with additional computational costs.

A polyalgorithm is an effective strategy to unite the strengths and circumvent the weaknesses of multiple algorithms that are also individually designed to solve the same problem. The concept of polyalgorithm was introduced by Rice and Rosen (1966). A polyalgorithm uses a combination of several basic methods in a framework. Each of these basic methods is applicable to the same problem, with only their performance and/or success being different for different datasets (inputs). The construction of this framework involves experimenting with an increasingly heterogeneous set of situations to evolve a robust algorithm that is capable of choosing a correct subset of algorithms suitable for a given input, and has performance metrics to integrate their outputs. The details of algorithm selection and processing stay hidden from the user. Polyalgorithms have been designed in the past for solving various problems, for example, nonlinear systems of equations (Rice and Rosen, 1966, Rice, 1969, 2014, 1967), matrix computations on parallel architectures (Li, 1996, Häfner et al., 1999), and certain chemical models (Gomeni and Gomeni, 1979).

This work lays the foundation for a polyalgorithm for LULCC detection. Three currently existing, fundamentally different from each other, change detection algorithms are utilized. A similar work in this direction is Zhan et al. (2002), wherein a framework is developed to evaluate five different algorithms on the input dataset, compare them based on certain scores, and then return the best results. Similar approaches are also gaining ground recently in the field of classification algorithms (Dietterich et al., 2001, Kittler et al., 1998, Wozniak et al., 2014). Most recently, in Healey et al. (2017), multiple change detection algorithms are utilized to build a decision trees based ensemble algorithm for LULCC.

The rest of this paper is organized as follows: Section 2 presents background on state-of-the-art change detection algorithms available in remote sensing, puts them in the context of the general time series literature, and explains the choice of algorithms used in this work. Section 3 defines the notation. Sections 4–6 describe three different trend and change detection algorithms — EWMACD, BFAST, and LandTrendR; experimental results demonstrating the successes, failures, and sensitivity to parameters for each algorithm are presented. Prospects for a viable polyalgorithm are discussed in Sections 7 and 8 concludes with an assessment and future work.

2. Background

Most of the LULC algorithms proposed in the remote sensing literature can be divided into two categories: *bitemporal analysis* and *temporal trajectory analysis*. Bitemporal analysis was more popular before 2008 (when the availability of satellite data to the public was very limited) and forms the classical way of analyzing images — these algorithms analyze changes occurring between two images (dates). The more preferred bitemporal algorithms rely on image differencing (Banner and Lynham, 1981; Hame, 1986; Coppin and Bauer, 1994; Cohen and Fiorella, 1998; Cohen et al., 1998; Serneels et al., 2001), and linear transformations (Richards, 1984, Neilsen et al., 1998, Fung and LeDrew, 1987, Fung, 1990). Other strategies used to design bitemporal algorithms include image rationing (Jensen, 1983), image regression (Joyce and Burns, 1981), and composite analysis (Thomson et al., 1980). Detailed surveys of these algorithms can be found in Coppin et al. (2004), and Lu et al. (2004). Multi-Index Integrated Change Analysis (MIICA) (Jin et al., 2013), a recent popular algorithm, utilizes two Landsat image pairs and four different derived spectral indices for change detection. The interested reader is referred to Campbell and Wynne (2011) for a further decent categorization of these algorithms.

For time series analysis based change detection algorithms, the popular strategy has been to design pixel based algorithms, wherein the time series for one pixel at a time is analyzed. One strategy is to segment the time series into piecewise linear segments. Specifically, the time span is partitioned into intervals where each interval corresponds to a sustained trend in observed values. The boundaries of these intervals correspond to points of change or the start of a new trend. The number of intervals depends on how many changes in trends occurred for that time series. This approach is adopted, for example, in Cohen et al. (2010), Kennedy et al. (2007), de Jong et al. (2012, 2013), Verbesselt et al. (2010a,b). In Moisen et al. (2016), on the other hand, seven shapes that can possibly occur in time series spectral data are identified. Constrained regression is done using splines that can generate these shapes. Some methods leverage the fact that climate related phenomena such as vegetation, temperatures, and the like are expected to follow a periodic pattern and utilize models based on Fourier series (trigonometric polynomials) (Brooks et al., 2014, Zhu and Woodcock, 2014). In Vegetation Change Tracker (VCT) (Huang et al., 2010), another popular method, for *each* image, cloud, shadow, and water are first masked using histograms. A derived index based on the mean and standard deviation of observed values of multiple bands in that image is calculated. One (that with the best derived index) image per year is selected for further processing. Any masked values appearing in these selected images are filled in by interpolating two temporally nearest available values in the previous and subsequent years. Then a suite of decision rules is used to detect and classify forest disturbances. Wavelets were utilized in Cai and Desheng (2015).

Data mining approaches have been proposed for classification and change detection (Goodwin et al., 2008, Mougel and Folcher, 2012, Petitjean et al., 2012, 2010, Vintrou et al., 2012, 2013). In Goodwin et al. (2008), a decision tree classifier was used to detect an outbreak of mountain pine beetle. This algorithm was originally implemented only on a subset of all available Landsat images (specifically, one scence per year was chosen from a 14 year span). In Petitjean et al. (2010), sequential pattern mining was proposed for finding trends (and changes) in land cover; all images were utilized. Dynamic time warping (DTW) was proposed in Petitjean et al. (2011) for comparing and analyzing remote sensing time series as well as characterizing change. Other recent related references include Anees et al. (2016), Benedek et al. (2015) and Bouziani et al. (2010).

Improved trend approximation can be obtained if, instead of treating each pixel independently, information from nearby pixels (spatial information) is also utilized. One such approach is VeRDET (Hughes et al., 2017), which utilizes two-dimensional total variation regularization (TVR) (Rudin et al., 1992;Goldstein and Osher, 2009) to modify the images so that they have small-scale spatial patches (reduced spatial heterogeniety). TVR is then used again for fitting a piecewise linear polynomial to the time series of each pixel. In Petitjean