Remote Sensing of Environment 209 (2018) 312–326

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Development of Landsat-based annual US forest disturbance history maps (1986–2010) in support of the North American Carbon Program (NACP)

Feng Zhao[a,*], Chengquan Huang[a], Samuel N. Goward[a], Karen Schleeweis[b,a], Khaldoun Rishmawi[a], Mary A. Lindsey[a], Elaine Denning[a], Louis Keddell[a], Warren B. Cohen[c], Zhiqiang Yang[c], Jennifer L. Dungan[d], Andrew Michaelis[d]

[a] Department of Geographical Sciences, University of Maryland, College Park, MD 20740, USA
[b] Rocky Mountain Research Station Forest Service, USDA, Ogden, UT 84401, USA
[c] Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA
[d] NASA Ames Research Center, Moffett Field, CA 94035, USA

ARTICLE INFO

Keywords:
CONUS
Landsat
Forest disturbance
NAFD-NEX
VCT

ABSTRACT

In Phase III of the North American Forest Dynamics (NAFD) study an automatic workflow has been developed for evaluating forest disturbance history using Landsat observations. It has four major components: an automated approach for image selection and preprocessing, the vegetation change tracker (VCT) forest disturbance analysis, postprocessing, and validation. This approach has been applied to the conterminous US (CONUS) to produce a comprehensive analysis of US forest disturbance history using the NASA Earth Exchange (NEX) cloud computing system. The resultant NAFD-NEX product includes 25 annual forest disturbance maps for 1986–2010 and two time-integrated maps to provide spatial-temporal synoptic view of disturbances over this time period. These maps were derived based on 24,000+ scenes selected from 350,000+ available Landsat images at 30-m resolution, and were validated using a visual assessment of Landsat time-series images in combination with high-resolution and other ancillary data sources over samples selected using a probability based sampling method. The validation revealed no major biases in the NAFD-NEX maps for disturbance events that resulted in at least 20% canopy cover loss. The average user's and producer's accuracies for the disturbance class were 53.6% and 53.3%, respectively, with the individual year's user's accuracy varying from 42.8% to 73.6% and producer's accuracy from 39.0% to 84.8% over the 25-year period. The NAFD-NEX disturbance maps are available from a web portal of the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL-DAAC) at https://doi.org/10.3334/ORNLDAAC/1290.

1. Introduction

North America is currently a net carbon source to the atmosphere, largely because of fossil fuel emissions, with approximately 30% of these emissions offset by vegetation growth (CCSP, 2007). North American forests are thought to be a long-term sink for atmospheric carbon, with much of the sink attributed to either forest regrowth from past agricultural clearing or to woody encroachment (CCSP, 2007). However, the magnitude of the North American forest sink is uncertain, because disturbance and regrowth dynamics are not well characterized or understood (Goward et al., 2008).

Disturbance events, including harvest, fire, insect and storm damage, and disease, strongly affect carbon dynamics in many ways, including biomass removal, emissions from decaying biomass, and changes in productivity (Birdsey et al., 2013; CCSP, 2007). Uncertainties related to disturbance and subsequent regrowth make it difficult to predict if North American forests and woodlands will continue to absorb atmospheric carbon for the foreseeable future (CCSP, 2007). To reduce these uncertainties, there is a need to develop temporal and spatial assessments of forest disturbance and regrowth dynamics for the North American continent (Birdsey et al., 2009; Houghton and Hackler, 2000; Hurtt et al., 2002).

The Landsat series of satellites offer the highest resolution, longest historical archive (~ 40 years) of systematically collected remotely sensed data (Goward, 2006; Roy et al., 2014; Wulder et al., 2016). Recently, a variety of disturbance mapping algorithms have been developed, and many disturbance products have been produced from these observations (e.g. Hansen et al., 2013; Hermosilla et al., 2015b;

* Corresponding author.
E-mail address: fengbjfu@gmail.com (F. Zhao).

https://doi.org/10.1016/j.rse.2018.02.035
Received 9 January 2017; Received in revised form 16 February 2018; Accepted 20 February 2018
Available online 19 March 2018

Hermosilla et al., 2016; Huang et al., 2010a; Jin et al., 2013; Kennedy et al., 2010; Potapov et al., 2011; Vogelmann et al., 2012). The North American Forest Dynamics (NAFD) study, a core project of the North American Carbon Program (NACP), was designed to improve understanding of the North American carbon budget through use of Landsat time series and related US Forest Inventory Data (FIA) (Goward et al., 2008). The primary goal of NAFD was to reduce spatial and temporal uncertainty in estimating US forest dynamics over a quarter century. NAFD began in 2003 with a prototype study and continued through two sampling phases. The results from these studies have been reported (Masek et al., 2013) and the data products are available at the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) (Goward et al., 2012).

In the previous phase I and phase II (referred to as NAFD I&II hereafter), Landsat data were sampled spatially (50 path/rows) and temporally (biennially) to reduce study costs of purchasing the Landsat data. These previous results showed that our disturbance estimates were nearly 20% lower than our reference validation results, indicating that both spatial and temporal sampling had limited the accuracy of the map products (Thomas et al., 2011). To address these limitations, the NAFD Phase III (referred to as NAFD III hereafter) study was developed to produce a wall-to-wall, annual analysis of the conterminous United States (CONUS), which was made feasible by the 2008 decision of the US Geological Survey to open the Landsat archive for no-cost access by data users (Woodcock et al., 2008; Wulder et al., 2012). The processing volume raised by this step was addressed through automation and a collaboration with NASA Earth Exchange computing facility (*NEX, http://nex.arc.nasa.gov*) (Nemani et al., 2011). The NAFD III wall-to-wall data set is referred to as the NAFD-NEX data. These data sets are now also available at the ORNL DAAC (Goward et al., 2015).

The new dimensions of NAFD III study required a new, automated data selection approach, a large step forward from the analyst-dependent, visual assessment based manual selection approach used in NAFD I&II. New methods for compositing to overcome cloud-contaminated observations, achieving consistency across diverse landscapes, handling the large data volumes, and identifying and addressing product quality issues under national production were also required. While several key components of NAFD algorithms have been published (Huang et al., 2009a; Huang et al., 2010a; Schleeweis et al., 2016), this paper provides an overview of the entire processing flow for producing the NAFD-NEX product, with a focus on the elements not detailed in previous publications as well as challenges unique to the NAFD III study, lessons learned, and implications for future research.

2. Overview of the NAFD-NEX processing flow

A major goal of the NAFD Phase III study was to compile annual CONUS-wide forest disturbance map using available Landsat images. This was achieved through four major steps: 1) image selection and preprocessing, 2) VCT disturbance analysis, 3) post-processing, and 4) results validation (Fig. 1). These steps reflected the requirements from acquiring images from the Landsat archive at USGS Earth Resources Observations and Science (EROS) to producing and validating national disturbance maps.

3. Methodology

3.1. Image selection and preprocessing

The quality of scene selection and pre-processing is critically important to reduction of measurement errors in remote sensing change detection studies (Townshend et al., 2012). The primary goal of this step was to assemble images that had best achievable geometric and radiometric qualities and minimum or no bad observations. In NAFD I&II, this was done through visual and manual analysis by image analysts (Huang et al., 2009a; Masek et al., 2013), which was extremely time consuming. In Phase III, we streamlined this process and developed the NAFD Image Selection and Processing Stream (NISPS) to provide an efficient, automated, and repeatable process with minimal analyst inputs for selecting, processing, tracking and assessing the large quantities of needed images (Schleeweis et al., 2016).

NAFD III was designed to examine Landsat images acquired from 1984 to 2011. A total of 434 World Reference System II (WRS2) path/rows (WPR) tiles were needed to provide a near complete coverage of CONUS (see Fig. 3 in Section 3.3.3). For each WPR tile, the VCT disturbance mapping algorithm only requires one cloud-free (defined as having $<$ 5% cloud cover throughout this study) growing season image per year. If cloud-free images were available for all years and WPR tiles then a total of 12,152 Landsat images would be needed for the NAFD III study. NISPS results showed that available CONUS Landsat scenes did not meet this requirement because of cloud contamination and limited scene acquisitions (Schleeweis et al., 2016). Where yearly, cloud-free imagery was not available, multiple cloud-contaminated, growing season images from that year were merged to produce, as much as possible, clear-view composites (CVC) that met the cloud-free definition. Only images with 50% or less cloud cover were used in this study, because we found that some EROS orthorectified images might suffer increased georegistration errors above 50% cloud cover.

The CVC algorithm was built upon the VCT cloud masking algorithm developed by Huang et al. (2010b). In addition to flagging cloud, shadow, other bad observations (e.g., missing scan lines, Landsat 7 SLC-off gaps, etc.), and clear-view pixels in an image, the VCT cloud masking algorithm uses a digital elevation model (DEM) dataset to normalize the brightness temperature band to minimize its dependency on surface elevation. From the available partly cloudy images selected by NISPS for a given WPR tile-year combination, the CVC algorithm first identifies the image with the most clear-view pixels according to the derived cloud masks as the composite base image (Fig. 2). Bad observations (i.e., clouds, shadow or SLC-off gaps) in the base image are filled with the clear-view pixels from the other available images using a maximum normalized temperature rule compositing rule. Because disturbed areas tend to have higher surface temperatures than undisturbed forests and other vegetated areas, this compositing method tends to preserve the disturbance signal over a disturbed area, and hence may enhance disturbance detection by VCT. To minimize among-image variations arising from differences in vegetation phenology and residual atmospheric effects, the reflectance values of all other input images are normalized to match the base image based on dark, dense forest (DDF) pixels (Huang et al., 2008). The quality flag of each selected pixel is used to create a new mask for the composited image, which tracks residual cloud/shadow and other bad observations in the composite.

3.2. VCT disturbance analysis

The Vegetation Change Tracker (VCT) is an automated algorithm designed for evaluating forest disturbance and post-disturbance recovery processes using spectral–temporal signals recorded in time-series Landsat observations (Huang et al., 2010a; Huang et al., 2009b; Huang et al., 2011). For a given Landsat time-series stack (LTSS), this algorithm first identifies forest samples in each image, based on which an integrated forest z-score (IFZ) is calculated for every pixel. The algorithm then tracks the changes of the IFZ time series to determine whether and when forest disturbances occur, and produces scene-level annual forest disturbance maps, which includes a disturbance class and 7 other classes (Table 1). Details of this algorithm have been published previously (Huang et al., 2010a; Huang et al., 2009b; Huang et al., 2011).

In NAFD III, we focused on two main issues that had impacted the automation and accuracy of the NAFD I&II results: 1) use of annual time series to reduce mapping errors seen in biennial results, and 2) development of an objective way to handle the wide range of forest reflectance signals across ecologically diverse regions, particularly in western US low density dry forest regions.