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.

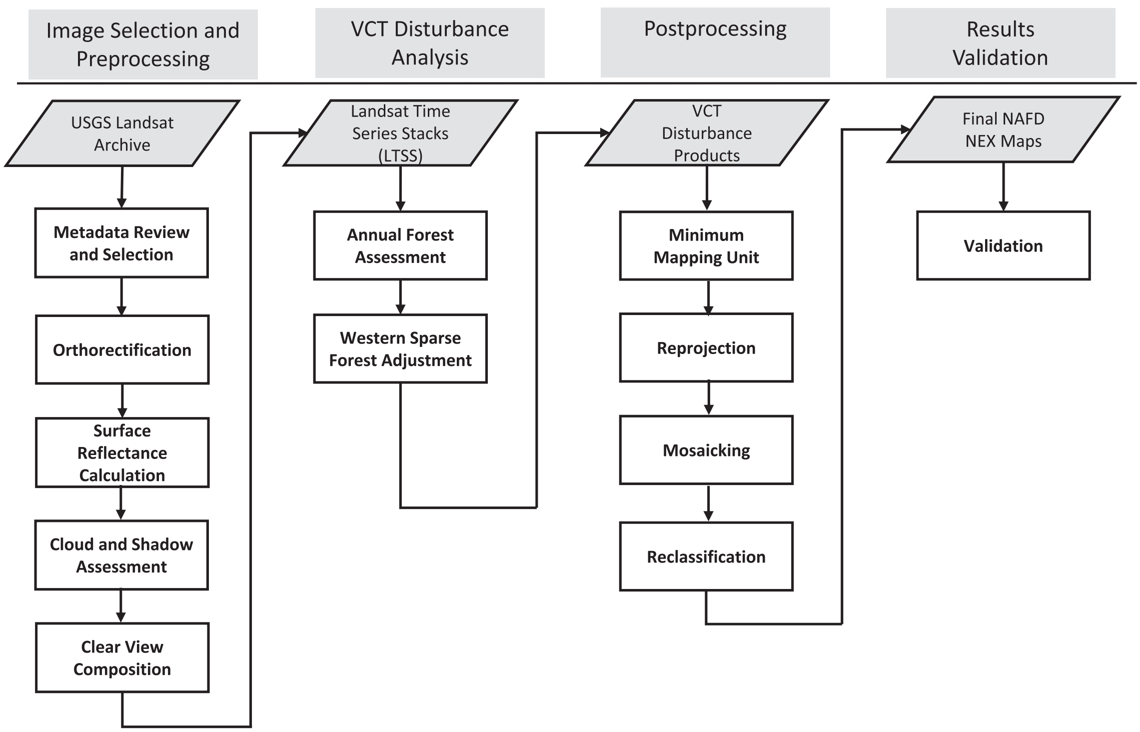

Fig. 1. The workflow for producing the NAFD-NEX (North American Forest Dynamics-NASA Earth eXchange) nation-wide annual forest disturbance maps.

3.2.1. Assessment of annual Landsat time series for disturbance mapping

NAFD I&II validation results revealed that many low magnitude disturbances, with partial canopy removal followed by rapid recovery, were not detected by the VCT with biennial time series (Thomas et al., 2011). Many of those low magnitude disturbances might be detected using annual LTSS, as case studies in the states of Washington and North Carolina revealed that annual LTSS resulted in more accurate disturbance products than biennial LTSS (Huang et al., 2015a; Huang et al., 2011). To quantify the improvements achievable using annual LTSS as compared with biennial LTSS, we evaluated annual and biennial results over four study areas selected from across the US, including North Carolina (WRS2 path 16/row 35), Mississippi (path 21/row 37), Minnesota (path 27/row 27), and Washington (path 47/row 27). For each study area, a biennial LTSS was created by removing one image every other year from the annual LTSS. The two LTSS were then analyzed using VCT to produce annual and biennial results. The two sets of results were compared to identify areas mapped as disturbances in the annual results but not in the biennial results, as well as areas mapped as disturbances in both but their disturbance years differed by > 5 years. From these difference areas we randomly selected 60 pixels for visual assessment. The year of disturbance, if one occurred, was determined for each selected pixel location by visually assessing the annual time series of Landsat imagery and high-resolution images available from Google Earth. The derived reference data were then used to assess the nature of the differences between the annual and biennial results. In this assessment, a disturbance year difference of up to 5 years was allowed to accommodate uncertainties in determining the disturbance year using biennial LTSS, which may arise from residual cloud contamination and other data quality problems (Thomas et al., 2011).

Table 1
Annual disturbance map class definitions at WRS-2 path/row level.

Value	Name	Definitions
0	background	Area not covered by the Landsat data
1	persisting nonforest	Area not covered by trees in all the years
2	persisting forest	Area covered by trees in all the years
4	persisting water	Water during the entire observing period
5	forest recovered from disturbance	Area covered by trees, but was recovered from previous disturbance
6	forest disturbance	Indicating the occurrence of a disturbance event over a forest area in the current year
7	post-disturbance nonforest	Area not covered by trees after disturbance

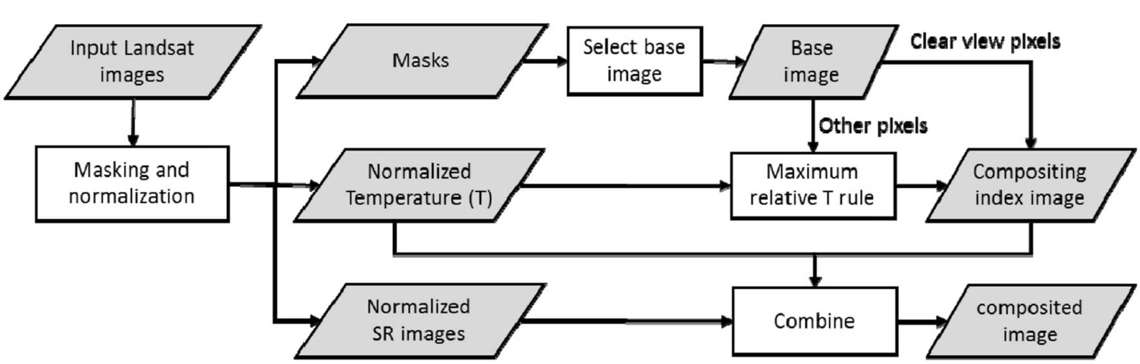

Fig. 2. A flowchart of the clear view compositing (CVC) algorithm implemented in the VCT.

3.2.2. VCT adjustment for western US low density forests

Due to dry environmental conditions, many regions in Western US are dominated by low density forests. These forests often have different understory and/or background cover and hence different spectral signatures, making it difficult to map them consistently across different regions using a single set of classification rules. In VCT, two indices through the time series are used to determine whether a pixel could possibly be “forest land”: forest z-score (IFZ) index and normalized difference vegetation index (NDVI) (Huang et al., 2010a). In the Eastern US where most of the forests have relatively dense canopy cover, a single set of IFZ and NDVI threshold values of 3 and 0.45 were found appropriate for separating forest from nonforested areas. But they were not suitable for the dry forest ecosystems in the Western US. Many low density forests would be mapped as nonforest using these threshold values. Since VCT was designed to detect disturbances over forestland only, this would result in substantial omission in disturbance detection.

During NAFD I&II, IFZ-NDVI threshold values for low cover/density conditions were determined iteratively through visual assessment of VCT results derived using different threshold values. In NAFD III, we developed a more systematic approach for calibrating the two threshold values to the local forest conditions within each WPR tile using field plot data provided by the USDA Forest Service Forest Inventory and Analysis (FIA) program. The FIA data were collected by field crew over a network of plots distributed across the U.S. with a nominal interval of approximately 5 km (Bechtold and Patterson, 2005; Reams et al., 2005; Smith, 2002). Forest is defined consistently across all regions as land that has or had (based on presence of stumps, snags, etc.) at least 10% crown cover and remain capable of producing a certain volume of timber (Smith et al., 2009a).

From all the FIA plots located within each WPR tile, a reference dataset was created by removing multiple condition plots as well as plots with disturbance events recorded by either the VCT or FIA, or both. Due to complications that may arise from spatial-temporal mismatches, use of such plots could make it difficult to interpret the results when comparing field plot data with satellite observations (Schroeder et al., 2014). The derived reference data were first used to assess the suitability of the initial IFZ-NDVI threshold values of 3 and 0.45, which were derived for eastern US closed canopy forests, for VCT analysis of western US dry forests. For each WPR tile located to the west of the 100°W longitude line (referred to as a western US WPR tile hereafter), a forest/nonforest map was produced using the IFZ and NDVI threshold values of 3 and 0.45, and its agreement with the FIA reference dataset was evaluated. Given the fact that precipitation is a major factor driving vegetation structure in dry environments, we further examined the relationships between the derived agreement levels and local precipitation.

Based on the above analysis, we developed a Monte Carlo approach to determine appropriate IFZ-NDVI threshold values for each of the western US WPR tiles. In this approach, the FIA reference dataset within each WPR tile was split randomly into two equal parts to create calibration and testing datasets. The calibration dataset was used to evaluate forest/nonforest classifications generated by VCT using randomly selected IFZ (between 0 and 30) and NDVI (between 0 and 1) values. This was repeated 20,000 times for each WPR tile. The pair of IFZ-NDVI threshold values that yielded the highest agreement between VCT results and the FIA calibration dataset were selected as “optimal” for that WPR tile. The VCT results generated using the selected “optimal” threshold values were evaluated using the testing dataset to quantify the improvement, if any, over the VCT classification derived using the initial eastern US closed canopy IFZ-NDVI threshold values.

3.3. Postprocessing

Once the VCT disturbance analysis was executed several additional evaluation and processing steps were needed to complete the development of the CONUS NAFD-NEX disturbance product. The first step was an initial quality analysis (IQA) designed to identify anomalies that may indicate gross errors. In the second step we adjusted the results to correct or reduce large errors identified through the IQA process. In the third step, a spatial and temporal minimum mapping unit (MMU) process was applied to the map results. The fourth postprocessing step merged the path/row results into a national map.

3.3.1. Initial quality analysis (IQA)

The key algorithms used in the NAFD-NEX processing flow: atmospheric correction (Masek et al., 2006), cloud and shadow masking (Huang et al., 2010b), and disturbance mapping (Huang et al., 2010a; Huang et al., 2011), have been tested extensively in previous studies. However, in some cases these methods may fail to address unidentified issues causing potentially significant errors in the derived map products. Assuming that results in neighboring WPR tiles or years should be similar in general, the IQA was designed to highlight anomalies that may be indicative of errors. It calculated a suite of WPR level metrics (Table 2) from the initial VCT results and used those metrics to create CONUS maps (see Fig. 9(a) in Section 4.3 for an example). These maps were visually examined for anomalies. For areas with identified anomalies the VCT maps were examined against input Landsat images at the 30-m resolution to determine whether those anomalies reflected real spatial-temporal variability of disturbance dynamics or were the result of large mapping errors.

3.3.2. Correction of gross errors

The IQA procedure identified abnormally high disturbance rates across much of the North Central US in 1987 and 1988 that were confirmed to be false alarms (for further details see Section 4.3). Many false alarms occurred along the boundaries between forest and agriculture fields or grasslands. Others were in wetland areas of the North Central US and throughout Florida. The 1992 National Land Cover Dataset data set (NLCD 1992) (Vogelmann et al., 2001) was used to reduce these errors. In the North Central region any 1987 or 1988 disturbance pixels located within a 3 × 3 neighborhood of a planted/cultivated class in the NLCD 1992 dataset was recoded to nonforest. To reduce commission errors over wetland areas in both the North Central region and Florida, disturbance pixels that were classified as emergent herbaceous wetland in NLCD 1992 were recoded to nonforest. Here the 1992 NLCD was chosen over the later year products for its temporal proximity to the problem years.

3.3.3. MMU filtering

As with many satellite based map products, the initial VCT disturbance products had many small disturbance patches that likely were errors resulting from residual misregistration errors in the input images, difficulties in classifying mixed pixels, and other unidentified issues. MMU filtering is a common approach for reducing such errors (Saura, 2002). Visual assessment of the initial VCT disturbance products revealed that the nature of these errors appeared to be region dependent. Most of the patches that appeared to be erroneous were very small (< 5 pixels) in regions that had significant forest cover. However, in regions with very little forest cover (e.g., the Great Plains), many of the

Table 2
Stack level statistical summary metrics used in initial quality analysis (IQA).

Metric Names	Description
Annual Forest Disturbance Rate (AFDR)	Ratio of disturbed forest pixels to total forest pixels
Sum AFDR (ΣAFDR)	Sum AFDR: 1986–2010
Mean AFDR (μAFDR)	Mean AFDR: 1986–2010
Standard Deviation of AFDR (σAFDR)	Standard deviation of AFDR: 1986–2010
Coefficient of variation of AFDR (CVAFDR)	Coefficient of variation of AFDR: 1986–2010

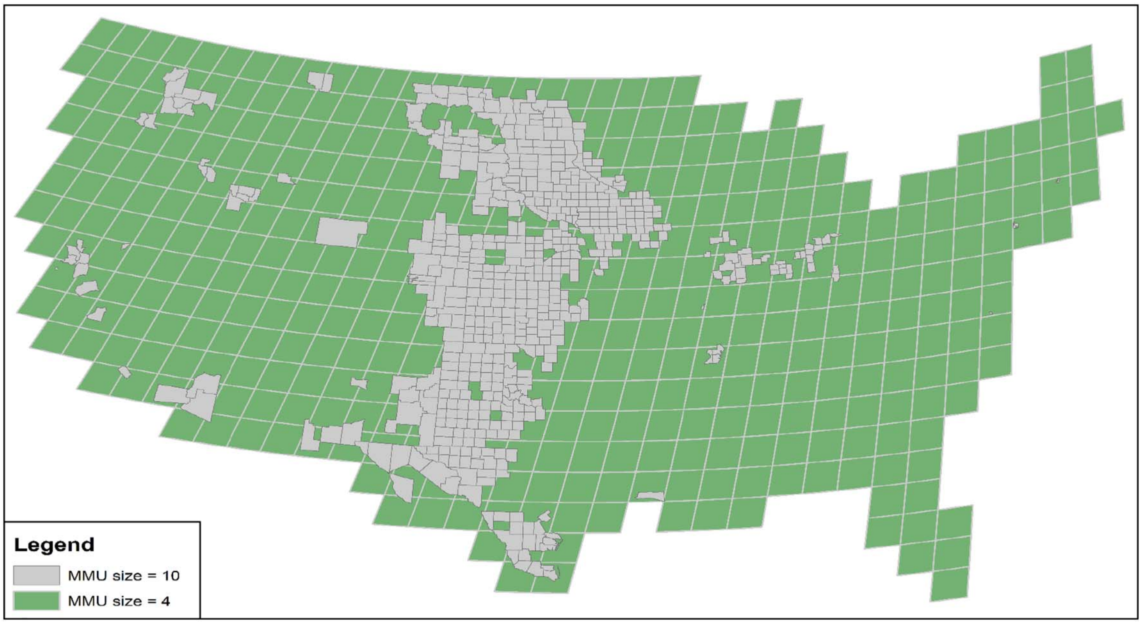

Fig. 3. Distribution of low forest cover (LFC) counties (shaded in gray) where the disturbance class was filtered using a 10-pixel (0.9 ha) minimum mapping unit (MMU). A 4-pixel (0.36 ha) MMU was applied to the rest of the conterminous US (CONUS) for this class. A 2-pixel (0.18 ha) was applied across CONUS for all static classes. The green quadrilaterals show the nominal non-overlap boundaries of the WRS-2 tiles needed to produce the NAFD-NEX product. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

patches were slightly larger (up to 10 pixels). To minimize the possibility of removing real disturbance patches by the MMU process, we used the NLCD 1992 data set (Vogelmann et al., 2001) to identify counties having < 6% forest cover or < 150 km^2 forest area as low forest cover (LCF) counties and applied an MMU of 10 pixels (0.9 ha) to the disturbance class in these counties. In all other counties, the disturbance class MMU was set to 4 pixels (0.36 ha) (Fig. 3).

Small errors in determining the disturbance year can lead to a disturbance event being classified into multiple artificially smaller patches. To account for these small temporal errors in adjacent disturbance pixels, we allowed a 1-year buffer in determining the patch size of a disturbance event (Huang et al., 2010a; Thomas et al., 2011), effectively allowing adjacent pixels whose disturbance years differed by ± 1 year to be considered as the same patch when determining the patch sizes for the MMU screening. This 1-year buffer was not used in LFC counties, because they had more speckled distribution of small spatial and temporal errors, which could aggregate into patches large enough to escape the 10-pixel MMU filter if the 1-year buffer were used. Disturbance patches smaller than the MMUs were recoded to the dominant no-change class, persisting forest or persisting nonforest, within the neighborhood of each patch. A flowchart of the entire MMU filtering process for disturbance pixels is shown in Fig. 4.

Because the no-change classes (e.g., persisting forest and persisting nonforest) were in general more reliable than the disturbance classes (Huang et al., 2015b; Huang et al., 2011; Thomas et al., 2011), a smaller MMU of 2 pixels was used. Patches smaller than the 2-pixel MMU were recoded to the dominant no-change classes surrounding those patches if they were adjacent to other no-change pixels. If they were surrounded by disturbance pixels, they were recoded to the dominant disturbance class in their neighborhood. This preference of no-change classes in the MMU process was designed to ensure that minimum or no new disturbances were introduced through the MMU process. It is important to note that whenever a pixel's class label was temporally adjusted based on the above rules, its class labels in all other years were adjusted properly to maintain temporal consistency as defined by the VCT algorithm across all years (Huang et al., 2010a).

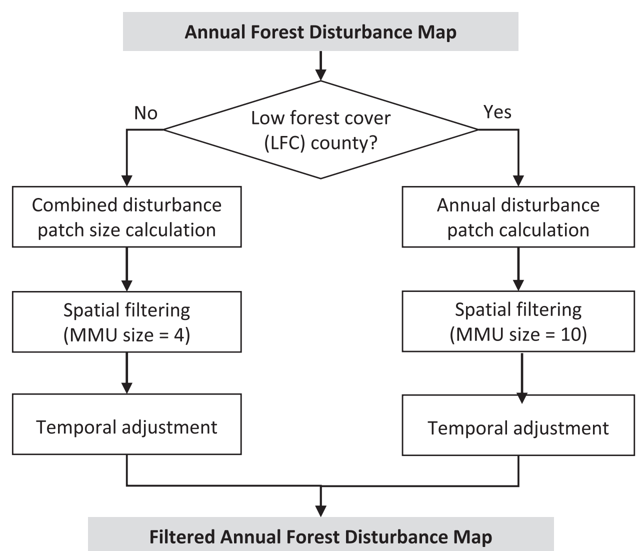

Fig. 4. A flowchart of the MMU algorithm used to filter disturbance pixels in the NAFD-NEX product.

3.3.4. *Map re-projection and mosaicking*

The initial VCT disturbance products had the same Universal Transverse Mercator (UTM) projection as the input Landsat images provided by USGS (Snyder, 1987). They were re-projected to the Albers Equal Area (AEA) projection using the inverse gridding approach and nearest neighbor resampling method. The AEA projection is more suitable for CONUS-wide map applications than the UTM projection and is widely adopted for Landsat-based products, including the Web-enabled Landsat Data (WELD) (Roy et al., 2010) and the series of NLCD land cover products (Homer et al., 2015; Homer et al., 2004; Vogelmann et al., 2001).

The WPR tile level VCT disturbance maps were mosaicked to produce a consistent national map for each year. Each NAFD CONUS map has the exact dimensions and corner coordinates as the NLCD

Table 3
Class definitions for national annual disturbance map mosaic.

Value	Name	Definitions
0	Data gaps	No acceptable data available.
1	Water	Area covered by water during the entire observing period, same as the persisting water class defined in Table 1.
2	No forest cover	Area not covered by trees in the current year but may or may not have tree cover in other years during the observing period, including the persisting nonforest class and post disturbance nonforest class defined in Table 1.
3	Forest cover	Area covered by trees in the current year but may or may not have tree cover in other years during the observing period, including the persisting forest class and forest recovered from disturbance class defined in Table 1.
4	Forest disturbance	Indicating the occurrence of a disturbance event over a forest area in the current year, same as the forest disturbance class defined in Table 1.
254	Outside of study area	Area outside the study area;

1992–2001 retrofit data product (https://www.mrlc.gov/nlcdrlc.php). The mosaicking process started by assigning pixels to the upper left corner of the national grid along the path direction. In the overlapping regions, pixels in the lower latitude path/rows were retained over those in higher latitude path/rows. When all the pixels in a single path were assigned, we then mosaicked the next path to the right, and, consequently, pixels in the next path would be retained over previous path.

3.3.5. Recoding and final NAFD-NEX product generation

The original classes of the VCT annual disturbance maps (Table 1) were deemed non-intuitive and somewhat confusion during internal reviews of the VCT map results. To address this concern, those classes were regrouped and renamed following the rules and definitions listed in Table 3.

The NAFD III explored Landsat data acquired from 1984 through 2011 to create annual disturbance maps. The results from 1986 to 2010 were used to produce the final NAFD-NEX maps. Those in the other three years were not used due to complications related to 1) lack of qualified acquisitions for many WPR tiles at the beginning of the time series (Schleeweis et al., 2016), and 2) need for consecutive low IFZ observations in determining forest cover before a disturbance event and consecutive high IFZ values for reliable detection of disturbance events (Huang et al., 2010a).

The annual maps were integrated to create two synoptic maps for 1986–2010 to depict the spatial–temporal dynamics of forest disturbances. In areas that had two or more disturbances mapped over the 25-year period, the first and second maps represented the first and last disturbance years, and hence are referred to as the first and last time integrated forest disturbance maps, respectively. These two maps are identical over areas where no disturbance (including water, nonforest, as well as "undisturbed" forest) or only one disturbance was detected over the 25-year period. The classes in these two maps were defined in Table 4.

3.4. Validation

The NAFD-NEX maps were validated using a reference dataset developed both for validating time-series Landsat based disturbance map products and for deriving disturbance rate estimates independent of the NAFD-NEX maps (Cohen et al., 2016). The classes in both the reference data and the NAFD-NEX maps were harmonized to a level where they could be compared directly before accuracy estimates were derived.

3.4.1. Reference data collection

Over 7200 geographically unique plot locations were selected across CONUS using a two-stage stratified random cluster design. During the first stage, the boundaries among the WPR tiles needed to cover CONUS were redrawn to create Thiessen Scene Area (TSA) polygons such that there were neither gap nor overlap between adjacent TSA polygons. These TSAs were divided into five strata based on a combination of ecoregions (Omernik, 1987) and forest type distribution (Ruefenacht et al., 2008). In each stratum, the number of TSAs to be selected was proportional to the product of the area of that stratum and the total forest area in that stratum. A total of 180 of 434 TSAs were selected. During the second stage, a simple random sampling method was used to select 40 plots, or 30-m Landsat pixels, within each selected TSA.

For each plot, reference classes throughout the entire observing period were determined through visual analysis of Landsat time-series observations and high-resolution imagery available in Google Earth along with other ancillary datasets using the TimeSync tool (Cohen et al., 2010). Years with visual evidence of change that impacted any portion of the 1-pixel plot area were noted with a corresponding change process, and a cumulative canopy cover loss estimate was derived for each process by relating photo interpreted tree canopy cover plots developed by Coulston et al. (2012) to Landsat spectral data using the Random Forests regression tree algorithm (Breiman, 2001). While readers are referred to the Cohen et al. (2016) study for more details on the derivation of the reference data, it is important to note that neither the input Landsat time-series observations nor the classification protocols used to derive the reference data and the NAFD-NEX product were exactly the same, which made the data harmonization discussed below necessary.

Table 4
Class definitions for the first and last time integrated disturbance maps.

Value	Name	Definitions
Stable classes		
1	Water	Area covered by water during the entire observing period
2	Nonforest	Area never covered by trees during the entire observing period
3	"Undisturbed" forest	Forested area that had no disturbance detected during the entire observing period. This includes areas that 1) had tree cover throughout the entire observing period, 2) were disturbed before the beginning of the time series, but returned to forest cover and were not disturbed again, and 3) were nonforest at the beginning of the observing period but had tree cover in later years (e.g., woody encroachment, agricultural abandonment, urban tree growth).
254	Outside of study area	Area outside the study area
Disturbance classes		
16–40	Disturbance year	Area disturbed in a particular year. Disturbance year = 1970 + class value. Only show the latest disturbance if an area was disturbed more than once during the observing period. The mapped disturbance event typically occurred between the acquisition dates of two Landsat images – one acquired in the disturbance year and the other in the immediately previous year, that were used in the vegetation change tracker (VCT) analysis.

3.4.2. Data harmonization

Several data harmonization rules were implemented to address several issues arising from the above described differences between the reference data and the NAFD-NEX product, including 1) uncertainties in determining the disturbance year arising from use of different Landsat images in the LTSS, 2) use of different approaches to characterize gradual changes, and 3) use of different definitions for forest.

Temporal uncertainty in detection year can be related to data quality problems in the Landsat time series, including residual cloud/shadow contamination, missing data, and SLC-off gaps (Huang et al., 2010a). These small temporal uncertainties can be exacerbated when comparing datasets created using 1) different Landsat time-series image acquisitions and 2) different compositing and cloud/shadow masking routines, as was the case with the reference and NAFD-NEX datasets. To minimize these methodology related temporal mismatches a 1-year buffer was allowed in determining the disturbance year match for abrupt events, a common practice in assessing time-series disturbance products (e.g. Cohen et al., 2017; Huang et al., 2010a; Thomas et al., 2011). Abrupt change events were defined by the reference data set as events where canopy loss occurred in 1–2 years.

Determining the onset year for gradual changes that occur over multiple years (e.g. forest decline or disease and insect related mortality) is difficult. Determining year matches for these disturbance events was more challenging, because these events were recorded differently between NAFD-NEX and the reference data. In NAFD-NEX, a gradual disturbance that occurred over N years was recorded as a single event with information on the onset year and the related recovery period (Huang et al., 2010a). The reference data provided information on the onset and the duration of the loss period of that event, but the loss process was reported in each of the N years (Cohen et al., 2016). To accommodate these differences, a NAFD-NEX disturbance year value was considered to agree with the reference data if it was within the disturbance year range of a gradual disturbance identified in the reference data and the event was treated as a single event.

Validation of the NAFD-NEX product was also complicated by the distinction between land cover and land use, a reoccurring issue in forest change studies using satellite imagery (Coulston et al., 2014; Woodall et al., 2016). Land areas not disturbed in a year were classified based on forest cover in the NAFD-NEX product (Tables 1 and 3), but in the reference data those areas were classified based on land use. Similar to the U.S. Forest Service definition of un-stocked forest land (Smith et al., 2009b), a reference pixel would be labeled as forest even if it had no tree cover after a disturbance. As a result, the reference data could not provide reliable information regarding whether and when a pixel had forest cover following a disturbance event. Therefore, we only focused on the disturbance class in validating the NAFD-NEX disturbance maps.

3.4.3. Accuracy assessment

The reference data were used to evaluate the NAFD-NEX product at two levels. At the first level, the canopy cover change information in the reference data was used to assess the sensitivity of the NAFD-NEX disturbance mapping algorithm suite to disturbance intensity. Given the fact that errors in disturbance products generated by many disturbance mapping algorithms, including VCT, are often affected by disturbance intensity (Cohen et al., 2017; Thomas et al., 2011), such a sensitivity analysis can provide a more complete understanding of the accuracies derived using reference data defined by a specific intensity threshold value for the disturbance class. To ensure that there would be adequate disturbance samples at different disturbance intensity levels, the disturbances in different years were grouped into a single disturbance class for both the NAFD-NEX products and the reference data. The disturbance samples in the latter were then reclassified using a cumulative canopy cover loss threshold varying from 0% (no reclassification) to 70% (a higher threshold would leave too few disturbance plots for reliable accuracy estimation, see Fig. 13 in Section 4.4) at 10% intervals. Finally, each thresholded reference dataset was used to calculate overall accuracy as well as commission and omission errors for the NAFD-NEX product, and the variability of these errors with the cumulative canopy cover change threshold was evaluated.

Based on the above sensitivity analysis, the disturbance intensity threshold that yielded roughly the same commission and omission errors for the disturbance class and hence likely represented the lower limit of canopy cover change detectable by the NAFD-NEX algorithm suite was identified. The second level assessment was conducted by using that threshold value to reclassify the reference disturbance plots and derive accuracy estimates for the annual disturbance maps. All accuracy measures, including the overall accuracy, class specific user's accuracy (1 – commission error) and producer's accuracies (1 – omission error), as well as biases arising from mapping error, were calculated following standard methods (McRoberts and Walters, 2012; Olofsson et al., 2014; Stehman, 2013).

4. Results

4.1. Effectiveness of the image selection and preprocessing stream

The NISPS was largely successful in providing cloud-free (i.e., cloud cover < 5%) growing season observations for the NAFD-NEX study. This was made possible in part by use of clear view compositing, which was capable of creating composites with minimum or no cloud cover using multiple partly cloudy images (Fig. 5). Only 62% of the 12,152 WPR-year combinations required by this study had cloud-free, growing season images without compositing (Fig. 6). Use of CVC composites increased the spatial-temporal coverage of qualified observations to near 90%. A detailed discussion of the NISPS results has been provided by Schleeweis et al. (2016).

4.2. Improved VCT disturbance analysis

4.2.1. Improvements from using annual Landsat time series

Use of annual LTSS allowed detection of more disturbances than using biennial LTSS. The assessment over study areas in North Carolina, Mississippi, Minnesota, and Washington revealed that > 50% of the 234 evaluated plots that had disturbances mapped in the annual results but not in the biennial results were true disturbances that were missed in the biennial results (Table 5). Most of them appeared to be partial canopy losses followed by rapid recovery, especially in the southern locations (WRS2 path 21/row 37 and path 16/row 35) where forests grow rapidly and are often thinned multiple times before the final harvest (Fig. 7).

About one third of the evaluated plots (82 out of 234) had true disturbances detected in both the annual and biennial VCT results, but the difference between their disturbance years was larger than the 5-year tolerance allowed for temporal mismatch (see Section 3.2.1). This was mainly due to inaccuracies in determining disturbance year in the biennial results. Only 10% (23 out of 234) of the selected samples had false positives in the annual VCT results.

4.2.2. Improved mapping of sparse forests in western US

The IFZ-NDVI threshold values derived based on eastern US forests were in general not suitable for mapping the variety of forests in the west. Agreements between the maps derived using those threshold values and FIA plot data were as low as 50%, although the agreement levels increased substantially as the local mean annual precipitation increased from below 300 mm to above 700 mm (Fig. 8, left). The lower agreements in the drier areas were generally due to forest class omission errors (i.e., the VCT incorrectly classified forested pixels as non-forest pixels), which were mostly associated with plots that had very low biomass (mostly < 60 tons/ha). The number of VCT forest class commission errors were generally low (Table 6).

The Monte Carlo approach was effective in selecting more

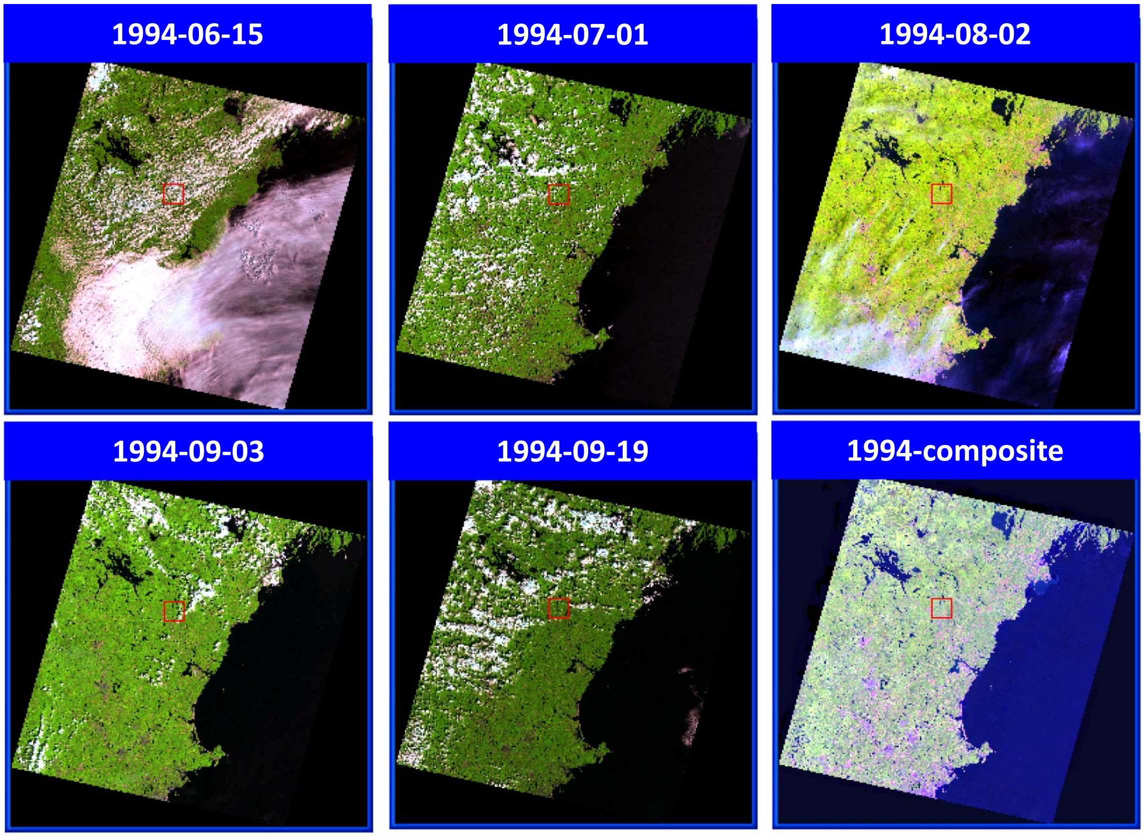

Fig. 5. A near cloud free composite was created using the CVC algorithm and five partly cloudy images acquired during the summer growing season of 1994 over eastern New Hampshire/southern Maine (WRS2 path 12/row 30). No qualified 1994 image would be available for this tile without using this composite. The acquisition year, month, and day of each image are shown as yyyy-mm-dd.

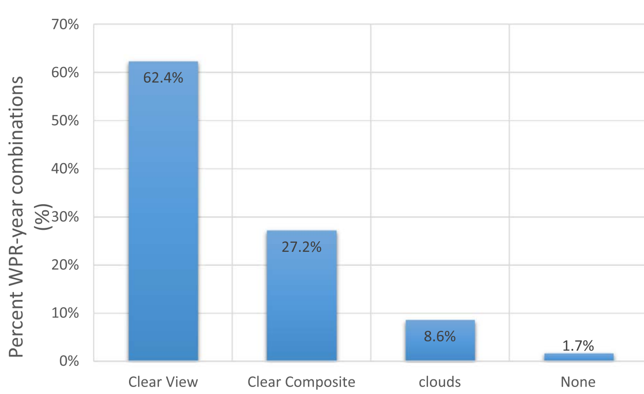

Fig. 6. Composition of different image types used for the 12,152 WPR-year combinations (434 WPR tiles × 28 years) required by the NAFD-NEX study. Clear view: clear view (cloud cover < 5%) was provided by individual Landsat acquisitions; Clear composite: clear view was achieved used the CVC algorithm; Clouds: partly cloudy (5% ≤ cloud cover < 50%) images (individual acquisitions or CVC composites) were used; None: no image was used due to lack of individual acquisitions or composites with cloud cover ≤50%.

appropriate IFZ and NDVI threshold values for the VCT algorithm for the WPR tiles that had mean annual precipitation below 500 mm/year. Compared with the original threshold values optimized for closed canopy forests in eastern US, the fine-tuned parameters resulted in substantial increases in the agreement level between VCT results and the FIA reference, with many tiles having improvements of 20% or more in absolute agreement values (Fig. 8, right).

Table 5
Distribution of randomly selected samples of disturbances mapped by annual Landsat time series stacks (LTSS) but not by biennial LTSS in Mississippi (path 27/row 37), North Carolina (path 16/row 35), Minnesota (path 27/row 27) and Washington (path 47/row 27). The value in each cell is the number of pixels evaluated.

Disagreement type	WRS2 path/row				Total for all sites
	21/37	16/35	27/27	47/27	
Errors in biennial results					
Disturbance missed by biennial LTSS	36	37	37	19	129
Disturbance year in biennial results off by > 5 years	15	21	12	33	82
Errors in annual results					
Commission error by annual LTSS	8	1	10	5	23
Total number of samples evaluated	59	59	59	57	234

4.3. Improvements from postprocessing

4.3.1. Effectiveness of the IQA and error correction methods

The IQA approach identified several types of anomalies that were confirmed to be gross errors through follow-up visual assessments. The first were large commission errors resulting from unmasked clouds due to a failure of the cloud masking algorithm for certain images. For example, the 1999 disturbance map over the WRS2 path 12/row 27 tile appeared to be an anomaly when compared with its spatial neighbors (Fig. 9(a)). A visual check of the full resolution map revealed that the anomaly was due to large disturbance patches resulted from a failure in the original VCT algorithm to flag clouds in that Landsat scene

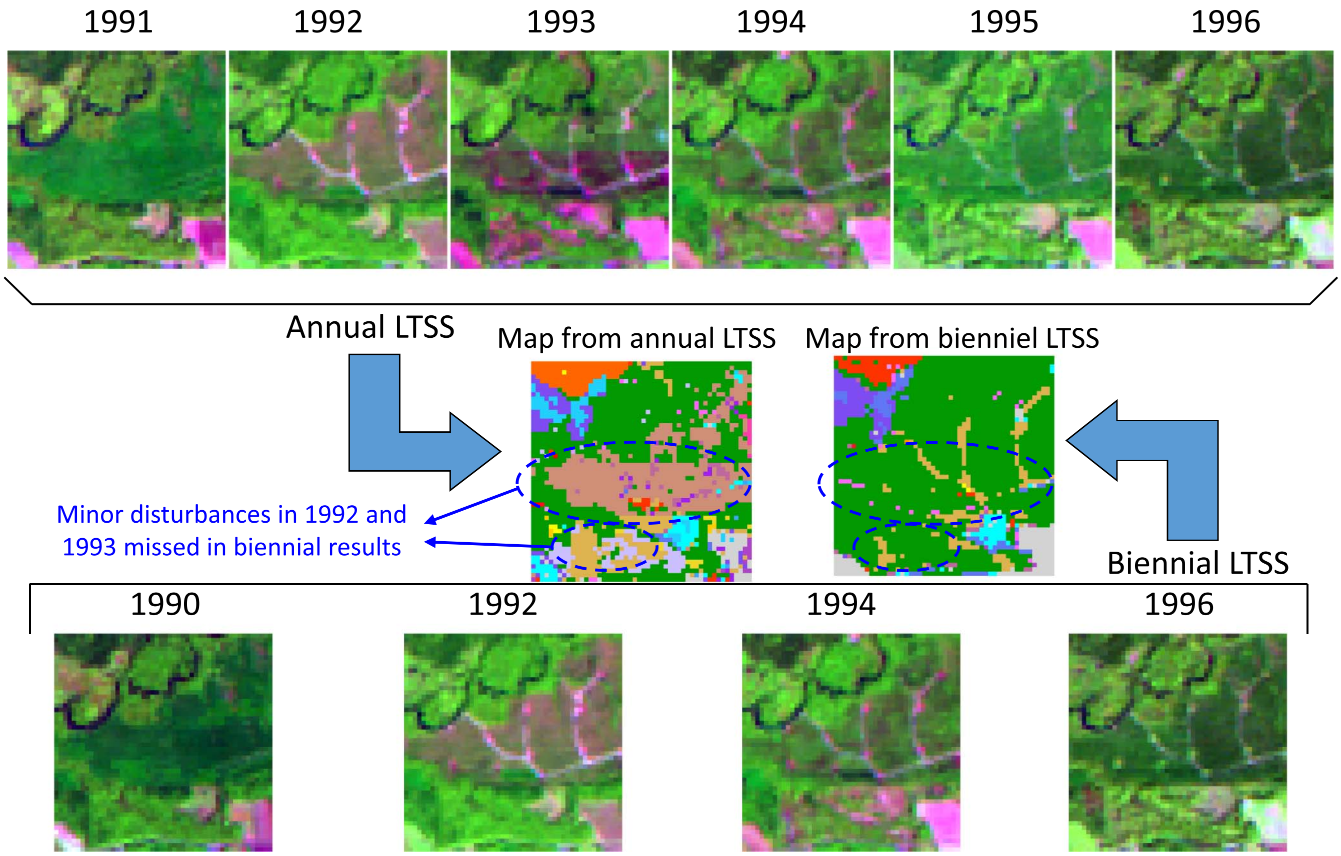

Fig. 7. A comparison of VCT disturbance maps (middle row) produced using annual (top row) and biennial LTSS (bottom row) for an area in Alabama (WRS2 path 21/row 37) shows the amount of visually identifiable disturbances captured by the annual results but not the biennial results.

(Fig. 9(b) and (c)). We subsequently reexamined the VCT cloud masking algorithm to identify and address the problem, which removed most of the false disturbances in the final product (Fig. 9(d)).

The second type of anomalies was abnormally high disturbance rates across much of the North Central US in 1987 and 1988, including North Dakota, South Dakota, Nebraska, Kansas, Minnesota, Iowa, Missouri, Wisconsin, Illinois, Michigan, Indiana, and Ohio. A comprehensive visual assessment revealed that a large portion of the mapped disturbances were false alarms driven by a widespread drought that occurred in that region (Andreadis et al., 2005). Because crops and other herbaceous vegetation are more sensitive to microclimate than are forests (De Frenne et al., 2011), this drought might have caused a multi-year shift in the spectral boundary between forests and nonforest vegetation cover types. Many of these false alarms were distributed along the boundaries between forest and agriculture fields or grassland (Fig. 10(a&b)). The 3 × 3 kernel filtering based on the NLCD dataset

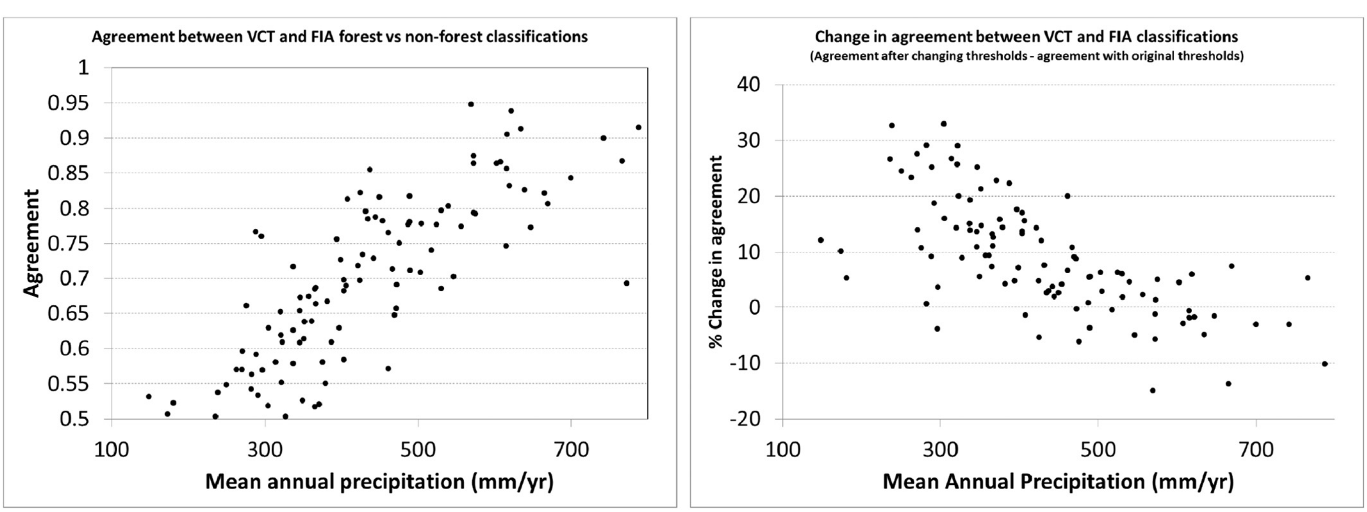

Fig. 8. Agreement between FIA plot data and VCT forest/non-forest classification produced using parameters tuned for eastern close canopy forests (IFZ = 3, NDVI = 0.45) appeared to be correlated with precipitation for WPR tiles in western US (left). The Monte Carlo approach identified more appropriate parameters that resulted in substantial increases in the agreement level for most WPR tiles having mean annual precipitation below 500 mm/year (right).

Table 6
Confusion matrix of the initial VCT forest/nonforest classification derived using the integrated forest z-score (IFZ) and NDVI parameters tuned for eastern US closed canopy forests evaluated using Forest Inventory and Analysis (FIA) plot data over a study area in southern Utah/northern Arizona (WRS-2 path 37/row 34). The low overall agreement was mainly due to misclassification of almost half forest plots (mostly with very low canopy cover) as nonforest.

FIA/VCT	Forest	Nonforest	Total	Producer's Accuracy
Forest	59	386	445	13%
Nonforest	3	422	425	99%
Total	62	808	870	
User's Accuracy	95%	52%		Overall accuracy = 55%

removed most of the false alarms along forest-agriculture/grassland boundaries in the North Central region (Fig. 10(c)).

Substantial errors were also found in the wetland rich North Central region and Florida, which resulted from difficulties to separate wetland temporal variations from forest disturbances using Landsat data. These errors were removed using the NLCD emergent herbaceous wetland class as a mask. Overall, the NLCD-based filtering and masking methods were effective in reducing the second and third error types. In Minnesota, for example, the WRS2 path 27/row 29 tile had an abnormally high annual disturbance rate of over 7% in 1987 according to the initial VCT disturbance products (Fig. 11). After postprocessing, the rate in that year was reduced to below 4%, which was more comparable to the rates in other years. It should be noted that although the NLCD 1992 had its own classification errors (Stehman et al., 2003), visual analyses suggested that the postprocessing methods aided in the removal of false disturbances without introducing noticeable errors.

4.3.2. Final NAFD-NEX disturbance maps

The final NAFD-NEX product covered 99.9% of the CONUS land area. Two additional WPR tiles would be needed to cover the remaining 0.1%, but they were not included due to their extremely small land area within CONUS. Due to lack of qualified image acquisitions, small amounts of additional data gaps were present in the individual year disturbance maps for all years except 1999–2002, 2008, and 2010. The approximate distribution of those gaps have been documented by Schleeweis et al. (2016). The first and last time-integrated disturbance maps did not have those gaps (Fig. 12), because they were filled by valid observations available in other years.

The two time integrated maps provided spatial-temporal synoptic views of disturbances occurred across CONUS over the entire observing period (Fig. 12). They show how pervasive forest disturbance were in many regions across CONUS when examined over a quarter century, including the Southeast, the Pacific Northwest, the Northern Rockies, and Maine, although the amount of forests subject to disturbance in any given year may not be that much in most regions (Huang et al., 2009b; Huang et al., 2015b; Huang et al., 2011). The last disturbance map can also be used to infer age since disturbance when examined in 2010 or a later year. Either one of the two time integrated maps can be used to identify nonforest, "undisturbed" forest, as well as areas had at least one disturbance between 1986 and 2010, but neither should be used to calculate annual disturbance rates. It is not uncommon that an area could be disturbed two or more times over a quarter century, especially in the Southeast where forests could reestablish in just a few years. But each time integrated map can represent only one disturbance event for any given location. Even when the two maps are used together, they cannot represent more than two disturbance events over the same location. The individual disturbance maps need to be used to calculate annual forest disturbance rates.

4.4. Accuracies of the NAFD-NEX product

The accuracies of the NAFD-NEX products varied with the cumulative canopy cover loss threshold used to define disturbance in the reference data. The disturbance class derived by aggregating disturbances in different years into a single class had a commission error of 25% and omission error of 62% when evaluated using all disturbance plots. While the overall accuracy did not change by > 4% points (81.2%–84.5%) across the entire range of cumulative canopy cover loss threshold values, increases in this threshold value resulted in substantial increases in the commission error and decreases in the omission error of the disturbance class (Fig. 13). The two errors were closest when the threshold was slightly over 20%, suggesting that the NAFD-NEX products had the least bias in representing disturbances that resulted in cumulative canopy cover loss equal to or greater than this threshold value (Stehman, 2013). Therefore, we used the 20% cumulative canopy cover loss threshold to define disturbance in deriving accuracy estimates for the NAFD-NEX product.

When disturbance year information was not considered, the NAFD-NEX product had an overall accuracy of 84.5% in representing disturbances that resulted in at least 20% cumulative canopy cover loss. The user's and producer's accuracies of the aggregated disturbance class were 67.0% and 62.9% respectively. When disturbance year match was considered, the average user's and producer's accuracies for the disturbance class were 53.6% and 53.3% respectively. Over time, the user's accuracy varied from 42.8% to 73.6% and the producer's accuracy from 40.0% to 84.8% (Fig. 14). Except for the 1986 map which had substantially higher user's accuracy than its producer's accuracy, in all other years the two accuracies were not statistically different from each other at the 95% confidence interval for disturbances that resulted in at least 20% cumulative canopy cover loss.

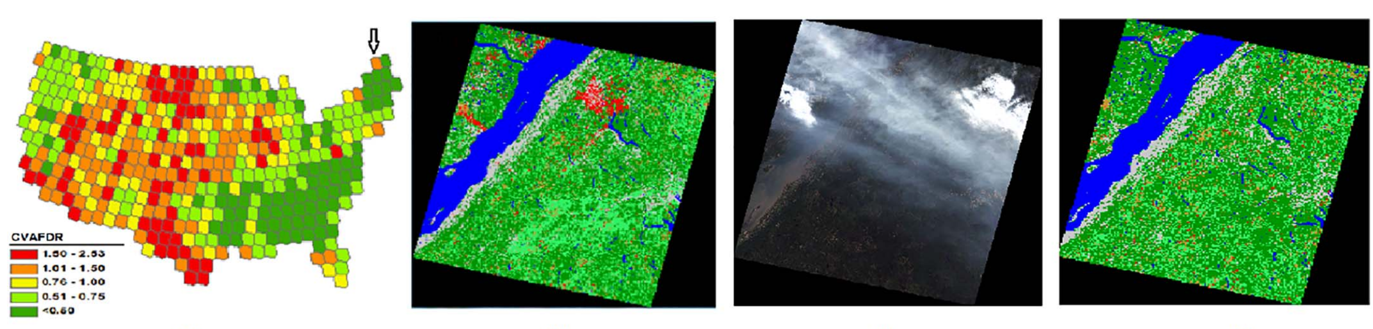

Fig. 9. The WRS2 path 12/row 27 (pointed to by the arrow) appeared to be an anomaly according to the 1999 CONUS-wide coefficient of variation of annual forest disturbance rate (CVAFDR) map derived through the IQA process (a). Further analyses revealed that the initial disturbance map had large patches of false disturbances (b), which were caused by a failure of the cloud masking algorithm to flag some of the clouds visible in the 1999 Landsat image (c, with Landsat bands 3,2,1 shown in red, green, and blue). Most of the false disturbances were removed after the problem with the cloud masking algorithm was identified and fixed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

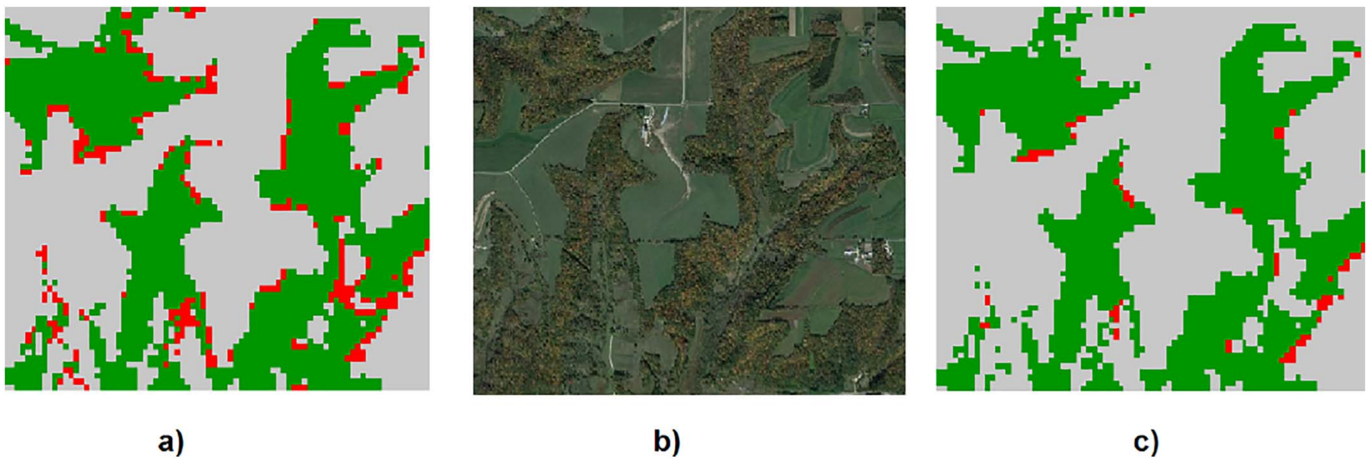

Fig. 10. A detailed assessment of an initial VCT disturbance map (a) against an aerial photograph (b) over southwest Wisconsin (WRS2 path 25/row 29) shows that most false disturbances in the north central region were small and were distributed along boundaries between forest and nonforest vegetation. After post-processing using a NLCD (National Land Cover Dataset) based filter, most of the false disturbances were removed in the final NAFD-NEX disturbance map (c).

5. Discussion

A comprehensive record of forest disturbance history is essential for understanding the impact of forest disturbance on regional carbon budget (Thornton et al., 2002; Williams et al., 2014). Dating back to the 1970s, the Landsat record provides the only data source for reconstructing forest disturbance history for up to 40 + years at regional to global scales. A number of algorithms have been developed for mapping disturbances using time-series satellite observations (e.g. Hansen et al., 2013; Hermosilla et al., 2016; Huang et al., 2010a; Kennedy et al., 2010; Zhu and Woodcock, 2014). Use of these algorithms to produce consistently high-quality disturbance products over large regions, however, requires adequate computing resources and a well-designed processing flow to address issues related to inputs, algorithm robustness, quality analysis and improvement, and product validation. This paper discusses how these issues were addressed in generating the NAFD-NEX product using the VCT algorithm.

The VCT algorithm was designed to map forest disturbance using annual or biennial LTSS. Previous studies revealed that many minor disturbances resulting in partial canopy loss followed by rapid recovery were not detectable using biennial LTSS (Huang et al., 2009b; Thomas et al., 2011). A multi-site analysis in this study confirmed that substantially more disturbances were mapped by using annual instead of biennial LTSS, and most of them were true disturbances that only resulted in partial canopy removal. As the total area of US forests subject to partial disturbances is multiple times larger than that of forests affected by stand clearing events (Birdsey and Lewis, 2003; Cohen et al., 2016), and post-disturbance recovery varied with disturbance intensity and type of treatment, these low magnitude disturbances may have large impact on future carbon uptake by US forests and their resilience to high severity fire and climate stresses (Dore et al., 2016). Therefore, use of annual LTSS in NAFD III represented a major improvement over previous NAFD phases for understanding the carbon dynamics of US forestland.

Despite the sub-monthly repeat cycle of each individual Landsat satellite, which in theory should be able to provide multiple growing season images per year for any given location, assembling the annual LTSS needed to produce the NAFD-NEX product was not straightforward. Some required observations were not available due to lack of acquisitions, data quality problems with available acquisitions (e.g., missing data, unusable observations due to cloud/shadow or other data quality problems), as well as issues arising from problematic radiometric and/or geometric corrections. With the CVC algorithm, the NISPS provided cloud free growing season images or image composites for nearly 90% of the WPR-year combinations of all the LTSS needed to produce the NAFD-NEX product. Further reduction of clouds or missing data in the assembled LTSS might be achievable by using other compositing methods designed to explore all available Landsat acquisitions regardless of cloud cover (e.g. Hansen et al., 2013; Hermosilla et al., 2015a; Roy et al., 2010; Zhu et al., 2013). However, cautions should be applied when using extremely cloudy images. As discussed earlier, lack of adequate clear-view land features may cause geometric correction

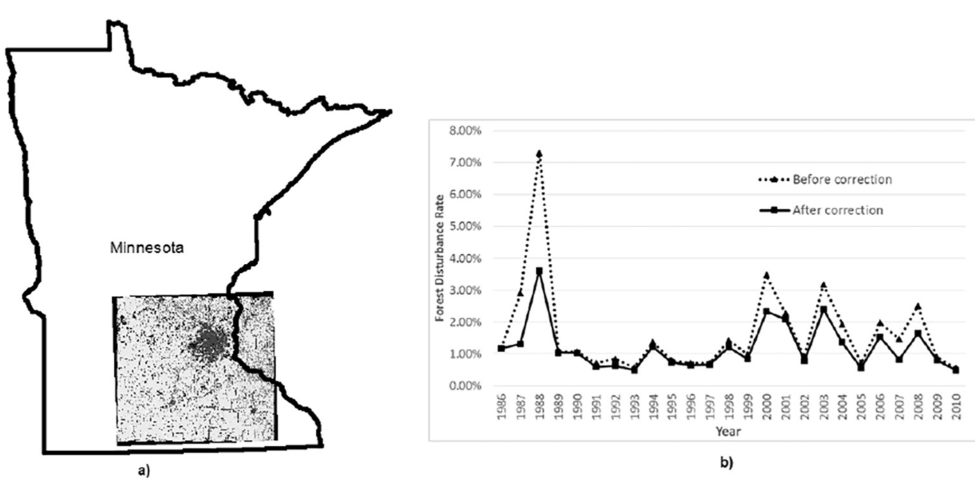

Fig. 11. Annual forest disturbance rates over eastern Minnesota (WRS2 path 27/row 29, a) were reduced substantially through post-processing in several years (b).

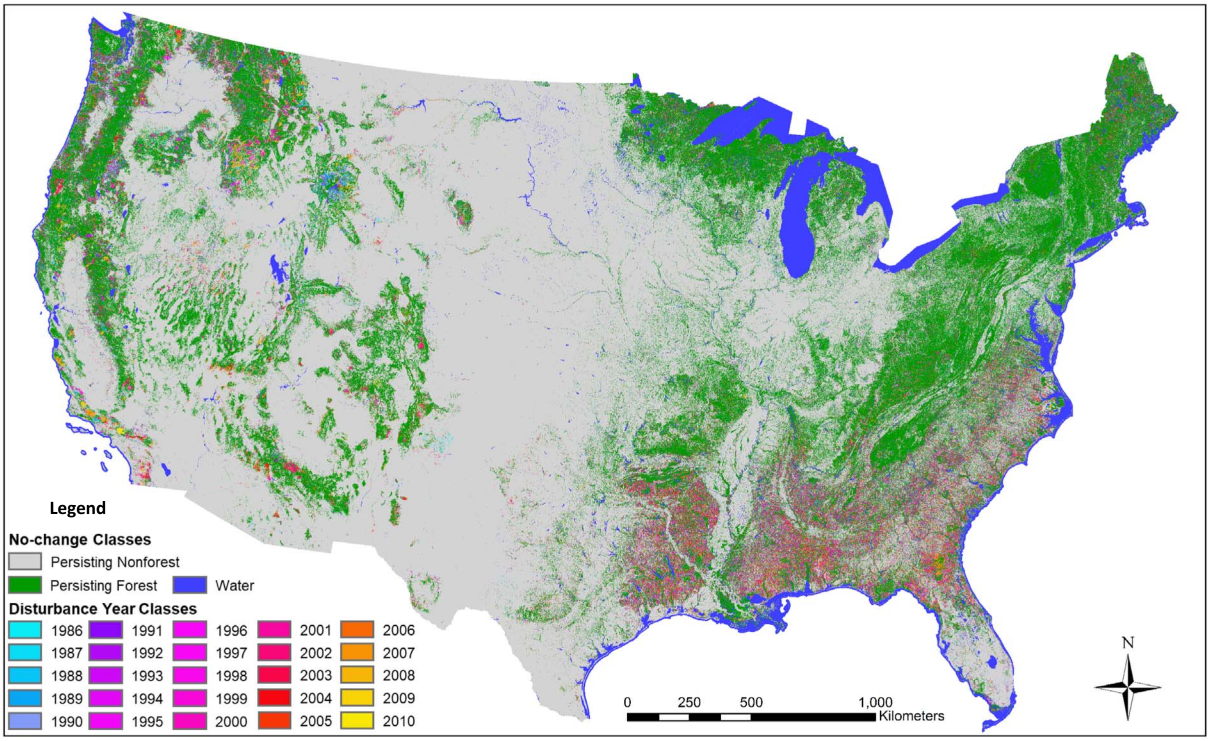

Fig. 12. NAFD-NEX last time integrated forest disturbance map.

algorithms like the Automated Registration and Orthorectification Package (AROP) to fail due to difficulties in identifying ground control points (GCP) (Gao et al., 2009). The impact of excessive clouds on the performance of such algorithms has yet to be evaluated.

Forest disturbance mapping in western US was complicated by the difficulty to separate low canopy cover forests from nonforest surfaces over the vast arid environments. The 10% crown cover threshold used by the FIA to define forest did not map into a single set of clearly defined spectral boundaries in the Landsat data. The Monte Carlo method developed in this study provided a practical approach for finding an optimal spectral boundary to maximize the agreement between the NAFD-NEX product and FIA plot data on a per WPR basis. Without this approach, both forest area and forest disturbance area would be substantially underestimated for many regions in western US if the IFZ and NDVI threshold values developed for eastern US close canopy forests were used without adjustment.

Timely error analysis is crucial for disturbance mapping, as false changes caused by erroneous inputs and/or algorithm failure can be many times larger than the typically small percentage of forest land subject to disturbance in any given year. Many of these errors could be identified by an experienced image analyst (e.g., Fig. 9(b) and 11(a)). For large area studies that result in large quantities of map products, however, it may not be feasible to have every single map examined by an image analyst. The IQA developed in this study provided a semi-automated approach for screening large quantities of map products. It successfully identified several types of large errors, which were

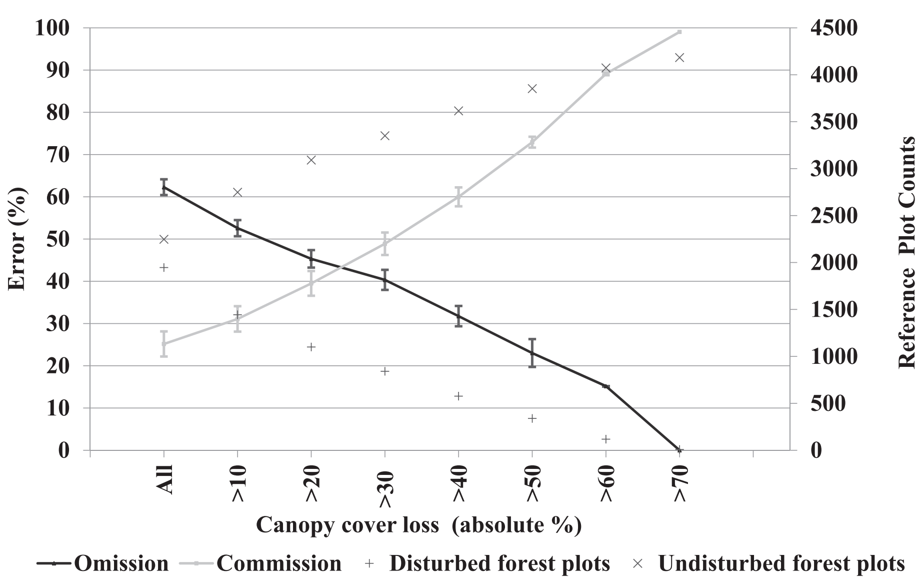

Fig. 13. Evaluation of disturbances mapped by the NAFD-NEX product without considering disturbance year match using reference data defined based on different canopy over change threshold values. The numbers of disturbed and undisturbed forest plots changed drastically as the threshold value increased, and the commission and omission errors reached an approximate balance when the threshold value was about 20%.

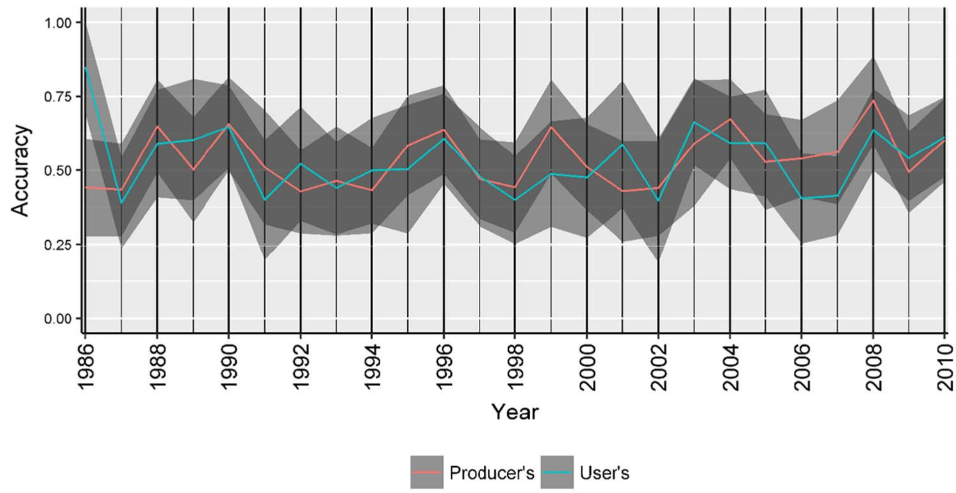

Fig. 14. Producer's and user's accuracies and their 95% confidence intervals (represented by the gray shade) of the NAFD-NEX annual disturbance maps evaluated based on the reference disturbance data defined using the 20% absolute cumulative canopy cover loss (per forest disturbance event) threshold value.

corrected or greatly reduced by improving the cloud masking algorithm and by using properly designed filtering methods and the NLCD 1992 land cover product. While these methods were designed specifically for the NAFD study, the guiding principles based on which they were developed could be useful for other disturbance studies.

A major feature of the reference dataset developed in this study was the derivation of a canopy cover change estimate as a measure of disturbance intensity for every disturbance event identified in the reference dataset. With this dataset we were able to both derive accuracy estimates for the NAFD-NEX product and to analyze the sensitivity of those estimates to the disturbance intensity threshold used to define disturbance. Although it has been observed that the performance of a disturbance mapping algorithm is often affected by the intensity of the disturbance events being mapped (Cohen et al., 2017; Healey et al., 2017; Thomas et al., 2011), it was difficult to determine the sensitivity of an algorithm to low intensity disturbances. The reference dataset derived through this study revealed that by using annual LTSS, the VCT algorithm was able to detect many low intensity disturbances. The derived NAFD-NEX product had little bias in representing disturbances that resulted in 20% or more cumulative canopy cover loss.

The sampling design of the reference data collection method was in general adequate for assessing the annual NAFD-NEX disturbance maps at the national scale. By defining disturbance using the 20% cumulative canopy cover loss threshold value, the number of disturbance plots in the reference data ranged from 33 to 52 per year. Splitting these plots among the five strata used in the sampling design would result in an average of 10 or less plots per year per stratum, which would not be enough for deriving reliable accuracy estimates. The number of disturbance plots needs to be increased at least by several times in order to derive accuracy estimates at the individual year level for the five regions used in the sampling design (Olofsson et al., 2014). This could be achieved by greatly increasing the overall sample size in each region if no stratification is used within each region, which would greatly increase the effort of the validation work. Alternatively, the number of disturbance plots could be increased without substantially increasing the overall sample size (and hence not increasing the cost of the validation work) by using a stratified random method that favors disturbance plots within each region (Cochran, 1977).

Use of the NEX cloud computing system was crucial to the production of the NAFD-NEX product. Many algorithm components, including preprocessing and VCT disturbance analysis, were CPU and memory intensive. Use of such intensive algorithms to analyze tens of thousands of Landsat images selected from a significant portion of the entire Landsat archive would not be possible by using the computing capabilities a typical research lab could afford. The NEX system made it convenient to access the Landsat images needed for this study and provided ample disk space for storing the inputs, intermediate products, and final results. Once the major algorithms were set up on the NEX system, it took the system < 10 days to produce the VCT map products for CONUS. Such short processing cycles were crucial to the successful development of the NAFD-NEX product, as multiple re-processings were needed to identify and address the various issues encountered in this study.

6. Summary

A workflow has been developed for producing the NAFD-NEX product using the NEX cloud computing system. This product consists of 25 CONUS-wide 30-m maps depicting forest disturbances occurred from 1986 to 2010 on an annual basis. The workflow had four major components, including the NISPS for image selection and preprocessing, VCT disturbance analysis, postprocessing, and validation. The NISPS was effective in selecting images needed by the NAFD-NEX study. It provided cloud free (i.e., cloud cover < 5%) growing season images for near 90% of the 12,152 WPR-year combinations required by this study, of which 30% were image composites derived using the CVC algorithm. Compared with previous phases of NAFD studies, the VCT disturbance analysis had two major improvements. One was use of annual instead of biennial LTSS, which significantly increased the success in mapping low severity disturbances without incurring excessive commission errors. The other was to use FIA data to calibrate the VCT algorithm to allow more consistent representation of low cover forests over the vast dry environments in western US. During postprocessing, a semi-automated IQA procedure was developed and used to identify large errors in the initial VCT disturbance products, which were then corrected or reduced through algorithm improvement and other triage methods. The corrected maps were then reprojected, mosaicked, and recoded to produce the final NAFD-NEX product. Based on a reference dataset derived by visual analysis of Landsat time series and available GoogleEarth and other high-resolution images to identify disturbance events and by using aerial photo based canopy cover estimates to model the cumulative canopy cover loss for those events, the NAFD-NEX product had an overall accuracy of 84.5% in representing disturbances that resulted in at least 20% cumulative canopy cover loss. The average user's and producer's accuracies of the individual disturbance year classes were 53.6% and 53.3% respectively. The NAFD-NEX product is available from an ORNL DAAC web portal at https://doi.org/10.3334/ORNLDAAC/1290.

Acknowledgement

The NAFD (North American Forest Dynamics) study, a core project of the North American Carbon Program (NACP), was pursued with grant support from NASA's Terrestrial Ecology, Carbon Cycle Sciences, and Applied Sciences Programs. The UMD group is supported under NASA Grant NNX11AJ78G S01. Associated NAFD III activities have been carried out by Warren Cohen of USFS Pacific Northwest Research Station (Product Validation), Jeffery Masek of NASA Goddard Space Flight Center (Regrowth Dynamics), Gretchen Moisen of USFS Rocky Mountain Research Station (Causal Factors Attribution), and Ramakrishna Nemani of NASA Ames Research Center (NEX computing). We are grateful to the reviewers and editors for their comments and suggestions that led to major improvements to the original manuscript.

References

Andreadis, K.M., Clark, E.A., Wood, A.W., Hamlet, A.F., Lettenmaier, D.P., 2005. Twentieth-century drought in the conterminous United States. J. Hydrometeorol. 6, 985–1001.

Bechtold, W.A., Patterson, P.L., 2005. The Enhanced Forest Inventory and Analysis Program - National Sampling Design and Estimation Procedures. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC, pp. 85.

Birdsey, R.A., Lewis, G.M., 2003. Current and historical trends in use, management, and disturbance of U.S. Forestlands. In: Kimble, J.M., Heath, L.S., Birdsey, R.A., Lal, R. (Eds.), The Potential of U.S. Forest Soils to Sequester Carbon and Mitigate the Greenhouse Effect. CRC Press, New York, pp. 15–33.

Birdsey, R., Bates, N., Behrenfeld, M., Davis, K., Doney, S.C., Feely, R., Hansell, D., Heath, L., Kasischke, E., Kheshgi, H., Law, B., Lee, C., McGuire, A.D., Raymond, P., Tucker, C.J., 2009. Carbon cycle observations: gaps threaten climate mitigation policies. EOS Trans. Am. Geophys. Union 90, 292–293.

Birdsey, R., Pan, Y., Houghton, R., 2013. Sustainable landscapes in a world of change: tropical forests, land use and implementation of REDD+: part II. Carbon Manage. 4, 567–569.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

CCSP, 2007. In: King, A.W., Dilling, L., Zimmerman, G.P., Fairman, D.M., Houghton, R.A., Marland, G., Rose, A.Z., Wilbanks, T.J. (Eds.), The First State of the Carbon Cycle Report (SOCCR): The North American Carbon Budget and Implications for the Global Carbon Cycle. National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, NC, pp. 242.

Cochran, W.G., 1977. Sampling Techniques, 3rd edn. Wiley, New York.

Cohen, W.B., Yang, Z.G., Kennedy, R., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync - Tools for calibration and validation. Remote Sens. Environ. 114, 2911–2924.

Cohen, W.B., Yang, Z., Stehman, S.V., Schroeder, T.A., Bell, D.M., Masek, J.G., Huang, C., Meigs, G.W., 2016. Forest disturbance across the conterminous United States from 1985–2012: the emerging dominance of forest decline. For. Ecol. Manag. 360, 242–252.

Cohen, W., Healey, S., Yang, Z., Stehman, S., Brewer, C., Brooks, E., Gorelick, N., Huang, C., Hughes, M., Kennedy, R., Loveland, T., Moisen, G., Schroeder, T., Vogelmann, J., Woodcock, C., Yang, L., Zhu, Z., 2017. How similar are forest disturbance maps derived from different Landsat time series algorithms? Forests 8, 98.

Coulston, J.W., Moisen, G.G., Wilson, B.T., Finco, M.V., Cohen, W.B., Brewer, C.K., 2012. Modeling percent tree canopy cover: a pilot study. Photogramm. Eng. Remote. Sens. 78, 715–727.

Coulston, J.W., Reams, G.A., Wear, D.N., Brewer, C.K., 2014. An analysis of forest land use, forest land cover and change at policy-relevant scales. Forestry 87, 267–276.

De Frenne, P., Brunet, J., Shevtsova, A., Kolb, A., Graae, B.J., Chabrerie, O., Cousins, S.A., Decocq, G., De Schrijver, A.N., Diekmann, M., Gruwez, R., Heinken, T., Hermy, M., Nilsson, C., Stanton, S., Tack, W., Willaert, J., Verheyen, K., 2011. Temperature effects on forest herbs assessed by warming and transplant experiments along a latitudinal gradient. Glob. Chang. Biol. 17, 3240–3253.

Dore, S., Fry, D.L., Collins, B.M., Vargas, R., York, R.A., Stephens, S.L., 2016. Management impacts on carbon dynamics in a Sierra Nevada mixed conifer Forest. PLoS One 11, e0150256.

Gao, F., Masek, J., Wolfe, R., 2009. An automated registration and orthorectification package for Landsat and Landsat-like data processing. J. Appl. Remote. Sens. 3, 033515. http://dx.doi.org/10.1117/1111.3104620.

Goward, S.N., 2006. Future of land remote sensing: what is needed. In: Joint Agency Commercial Imagery Evaluation (JACIE) Workshop. Patuxent National Wildlife Research Center, Laurel, MD.

Goward, S.N., Masek, J.G., Cohen, W., Moisen, G., Collatz, G.J., Healey, S., Houghton, R., Huang, C., Kennedy, R., Law, B., Turner, D., Powell, S., Wulder, M., 2008. Forest disturbance and North American carbon flux. EOS Trans. Am. Geophys. Union 89, 105–106.

Goward, S.N., Huang, C., Masek, J.G., Cohen, W.B., Moisen, G.G., Schleeweis, K., 2012. NACP North American Forest Dynamics Project: forest disturbance and regrowth data. https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1077In: Maryland, U.O. (Ed.), ORNL DAAC, Oak Ridge, Tennessee, USA.

Goward, S.N., Huang, C., Zhao, F., Schleeweis, K., Rishmawi, K., Lindsey, M., Dungan, J.L., Michaelis, A., 2015. NACP NAFD Project: forest disturbance history from Landsat, 1986–2010. http://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1290In: U.O.M.-N. NEX (Ed.), ORNL DAAC, Oak Ridge, Tennessee, USA.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853.

Healey, S.P., Cohen, W.B., Yang, Z., Kenneth Brewer, C., Brooks, E.B., Gorelick, N., Hernandez, A.J., Huang, C., Joseph Hughes, M., Kennedy, R.E., Loveland, T.R., Moisen, G.G., Schroeder, T.A., Stehman, S.V., Vogelmann, J.E., Woodcock, C.E., Yang, L., Zhu, Z., 2017. Mapping forest change using stacked generalization: an ensemble approach. Remote Sens. Environ. 204, 717–728(in press), https://www.sciencedirect.com/science/article/pii/S0034425717304418.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2015a. An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. Remote Sens. Environ. 158, 220–234.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2015b. Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. Remote Sens. Environ. 170, 121–132.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Campbell, L.B., 2016. Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring. Int. J. Digital Earth 9, 1035–1054.

Homer, C., Huang, C., Yang, L., Wylie, B., Coan, M., 2004. Development of a 2001 national land cover database for the United States. Photogramm. Eng. Remote. Sens. 70, 829–840.

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., Megown, K., 2015. Completion of the 2011 national land cover database for the conterminous United States – representing a decade of land cover change information. Photogramm. Eng. Remote. Sens. 81, 345–354.

Houghton, R.A., Hackler, J.L., 2000. Changes in terrestrial carbon storage in the United States. 1: the roles of agriculture and forestry. Glob. Ecol. Biogeogr. 9, 125–144.

Huang, C., Song, K., Kim, S., Townshend, J.R.G., Davis, P., Masek, J., Goward, S.N., 2008. Use of a dark object concept and support vector machines to automate forest cover change analysis. Remote Sens. Environ. 112, 970–985.

Huang, C., Goward, S.N., Masek, J.G., Gao, F., Vermote, E.F., Thomas, N., Schleeweis, K., Kennedy, R.E., Zhu, Z., Eidenshink, J.C., Townshend, J.R.G., 2009a. Development of time series stacks of Landsat images for reconstructing forest disturbance history. Int. J. Digital Earth 2, 195–218.

Huang, C., Goward, S.N., Schleeweis, K., Thomas, N., Masek, J.G., Zhu, Z., 2009b. Dynamics of national forests assessed using the Landsat record: case studies in eastern U.S. Remote Sens. Environ. 113, 1430–1442.

Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., Vogelmann, J.E., 2010a. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sens. Environ. 114, 183–198.

Huang, C., Thomas, N., Goward, S.N., Masek, J., Zhu, Z., Townshend, J.R.G., Vogelmann, J.E., 2010b. Automated masking of cloud and cloud shadow for forest change analysis. Int. J. Remote Sens. 31, 5449–5464.

Huang, C., Schleeweis, K., Thomas, N., Goward, S.N., 2011. Forest dynamics within and around the Olympic National Park assessed using time series Landsat observations. In: Wang, Y. (Ed.), Remote Sensing of Protected Lands. Taylor & Francis, London, pp. 71–93.

Huang, C., Ling, P.-Y., Zhu, Z., 2015a. North Carolina's forest disturbance and timber production assessed using time series Landsat observations. Int. J. Digital Earth 1–23.

Huang, C., Ling, P.-Y., Zhu, Z., 2015b. North Carolina's forest disturbance and timber production assessed using time series Landsat observations. Int. J. Digital Earth 8, 947–969.

Hurtt, G.C., Pacala, S.W., Moorcroft, P.R., Caspersen, J., Shevliakova, E., Houghton, R.A., Moore, B., 2002. Projecting the future of the U.S. carbon sink. PNAS 99, 1389–1394.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., Xian, G., 2013. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. Remote Sens. Environ. 132, 159–175.

Kennedy, R.E., Yang, Z.G., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. Remote Sens. Environ. 114, 2897–2910.

Masek, J., Hall, F., Wolfe, R., Saleous, N., Huemmrich, F., Gao, F., Kutler, J., Lim, T.-K., Huang, C., Goward, S., Cohen, W., 2006. LEDAPS: A Satellite-based Disturbance Map for North America. NASA Joint Workshop on Biodiversity, Terrestrial Ecology, and Related Applied Sciences. College Park, MD.

Masek, J., Goward, S., Kennedy, R., Cohen, W., Moisen, G., Schleeweis, K., Huang, C., 2013. United States forest disturbance trends observed using Landsat time series. Ecosystems 16, 1087–1104.

McRoberts, R.E., Walters, B.F., 2012. Statistical inference for remote sensing-based estimates of net deforestation. Remote Sens. Environ. 124, 394–401.

Nemani, R., Votava, P., Michaelis, A., Melton, F., Milesi, C., 2011. Collaborative supercomputing for global change science. EOS Trans. Am. Geophys. Union 92, 109–110.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57.

Omernik, J.M., 1987. Ecoregions of the conterminous United States. Ann. Assoc. Am. Geogr. 77, 118–125.

Potapov, P., Hansen, M.C., Gerrand, A.M., Lindquist, E.J., Pittman, K., Turubanova, S., Wilkie, M.L., 2011. The global Landsat imagery database for the FAO FRA remote sensing survey. Int. J. Digital Earth 4, 2–21.

Reams, G.A., Smith, W.D., Hansen, M.H., Bechtold, W.A., Roesch, F.A., Moisen, G.G.,

2005. The Enhanced Forest Inventory and Analysis Program — National Sampling Design and Estimation Procedures. Southern Research Station General Technical Report SRS-80. F.S. U.S. Department of Agriculture, Southern Research Station, Asheville, NC, pp. 11–26.
Roy, D.P., Ju, J.C., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., Loveland, T.R., Vermote, E., Zhang, C.S., 2010. Web-enabled Landsat data (WELD): Landsat ETM plus composited mosaics of the conterminous United States. Remote Sens. Environ. 114, 35–49.
Roy, D.P., Wulder, M.A., Loveland, T.R.C.E.W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., Zhu, Z., 2014. Landsat-8: science and product vision for terrestrial global change research. Remote Sens. Environ. 145, 154–172.
Ruefenacht, B., Finco, M.V., Nelson, M.D., Czaplewski, R., Helmer, E.H., Blackard, J.A., Holden, G.R., Lister, A.J., Salajanu, D., Weyermann, D., Winterberger, K., 2008. Conterminous U.S. and Alaska forest type mapping using Forest inventory and analysis data. Photogramm. Eng. Remote. Sens. 74, 1379–1388.
Saura, S., 2002. Effects of minimum mapping unit on land cover data spatial configuration and composition. Int. J. Remote Sens. 23, 4853–4880.
Schleeweis, K., Goward, S.N., Huang, C., Dwyer, J.L., Dungan, J.L., Lindsey, M.A., Michaelis, A., Rishmawi, K., Masek, J.G., 2016. Selection and quality assessment of Landsat data for the North American forest dynamics forest history maps of the US. Int. J. Digital Earth 9, 963–980.
Schroeder, T.A., Healey, S.P., Moisen, G.G., Frescino, T.S., Cohen, W.B., Huang, C., Kennedy, R.E., Yang, Z., 2014. Improving estimates of forest disturbance by combining observations from Landsat time series with U.S. Forest Service Forest Inventory and Analysis data. Remote Sens. Environ. 154, 61–73.
Smith, W.B., 2002. Forest inventory and analysis: a national inventory and monitoring program. Environ. Pollut. 116, S233–S242.
Smith, B.W., Miles, P.D., Perry, C.H., Pugh, S.A., 2009a. Forest Resources of the United States: A Technical Document Supporting the Forest Service 2010 RPA Assessment. U.S. Department of Agriculture, Forest Service, Washington Office, Washington, D.C., pp. 337.
Smith, W.B., Miles, P.D., Perry, C.H., & Pugh, S.A. (2009b). Forest Resources of the United States, 2007. In (p. 336). Washington, DC: U.S. Department of Agriculture, Forest Service.
Snyder, J.P., 1987. Map Projections - A Working Manual. U.S. Geological Survey, Washington, DC, pp. 377.
Stehman, S.V., 2013. Estimating area from an accuracy assessment error matrix. Remote Sens. Environ. 132, 202–211.
Stehman, S.V., Wickham, J.D., Smith, J.H., Yang, L., 2003. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: statistical methodology and regional results. Remote Sens. Environ. 86, 500–516.
Thomas, N., Huang, C., Goward, S.N., Powell, S., Rishmawi, K., Schleeweis, K., Hinds, A., 2011. Validation of North American forest disturbance dynamics derived from Landsat time series stacks. Remote Sens. Environ. 115, 19–32.
Thornton, P.E., Clark, K.L., Falge, E., Ellsworth, D.S., Goldstein, A.H., Monson, R.K., Hollinger, D., Falk, M., Chen, J., Sparks, J.P., Law, B.E., Gholz, H.L., 2002. Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. Agric. For. Meteorol. 113, 185–222.
Townshend, J.R., Masek, J.G., Huang, C., Vermote, E.F., Gao, F., Channan, S., Sexton, J.O., Feng, M., Narasimhan, R., Kim, D., Song, K., Song, D., Song, X.-P., Noojipady, P., Tan, B., Hansen, M.C., Li, M., Wolfe, R.E., 2012. Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. Int. J. Digital Earth 5, 373–397.
Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., Driel, N.V., 2001. Completion of the 1990s national land cover data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogramm. Eng. Remote. Sens. 67, 650–662.
Vogelmann, J.E., Xian, G., Homer, C., Tolk, B., 2012. Monitoring gradual ecosystem change using Landsat time series analyses: case studies in selected forest and rangeland ecosystems. Remote Sens. Environ. 122, 92–105.
Williams, C.A., Collatz, G.J., Masek, J., Huang, C., Goward, S.N., 2014. Impacts of disturbance history on forest carbon stocks and fluxes: merging satellite disturbance mapping with forest inventory data in a carbon cycle model framework. Remote Sens. Environ. 151, 57–71.
Woodall, C.W., Walters, B.F., Russell, M.B., Coulston, J.W., Domke, G.M., D'Amato, A.W., Sowers, P.A., 2016. A tale of two forest carbon assessments in the eastern United States: forest use versus cover as a metric of change. Ecosystems 19, 1401–1417.
Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P.S., Vermote, E.F., Vogelmann, J., Wulder, M.A., Wynne, R., 2008. Free access to Landsat imagery. Science 320, 1011.
Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: how free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 122, 2–10.
Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global Landsat archive: status, consolidation, and direction. Remote Sens. Environ. 185, 271–283.
Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens. Environ. 144, 152–171.
Zhu, Z., Woodcock, C.E., Olofsson, P., 2013. Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sens. Environ. 122, 75–91.