INTERNATIONAL JOURNAL OF DIGITAL EARTH, 2016
VOL. 9, NO. 10, 963–980
http://dx.doi.org/10.1080/17538947.2016.1158876

Selection and quality assessment of Landsat data for the North American forest dynamics forest history maps of the US

Karen Schleeweis[a], Samuel N. Goward[b], Chengquan Huang[b], John L. Dwyer[c], Jennifer L. Dungan[d], Mary A. Lindsey[e], Andrew Michaelis[f], Khaldoun Rishmawi[b] and Jeffery G. Masek[g]

[a]Forest Inventory and Analysis, US Forest Service, Rocky Mountain Research Station, Ogden, UT, USA; [b]Department of Geographical Sciences, University of Maryland, College Park, MD, USA; [c]U.S. Geological Survey Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD, USA; [d]NASA Ames Research Center, Moffett Field, CA, USA; [e]Climate Program Office, NOAA, Washington, DC, USA; [f]NASA Ames Research Center, University Corporation Monterey Bay, Moffett Field, CA, USA; [g]NASA Goddard Space Flight Center, Greenbelt, MD, USA

ABSTRACT

Using the NASA Earth Exchange platform, the North American Forest Dynamics (NAFD) project mapped forest history wall-to-wall, annually for the contiguous US (1986–2010) using the Vegetation Change Tracker algorithm. As with any effort to identify real changes in remotely sensed time-series, data gaps, shifts in seasonality, misregistration, inconsistent radiometry and cloud contamination can be sources of error. We discuss the NAFD image selection and processing stream (NISPS) that was designed to minimize these sources of error. The NISPS image quality assessments highlighted issues with the Landsat archive and metadata including inadequate georegistration, unreliability of the pre-2009 L5 cloud cover assessments algorithm, missing growing-season imagery and paucity of clear views. Assessment maps of Landsat 5–7 image quantities and qualities are presented that offer novel perspectives on the growing-season archive considered for this study. Over 150,000+ Landsat images were considered for the NAFD project. Optimally, one high quality cloud-free image in each year or a total of 12,152 images would be used. However, to accommodate data gaps and cloud/shadow contamination 23,338 images were needed. In 220 specific path-row image years no acceptable images were found resulting in data gaps in the annual national map products.

ARTICLE HISTORY
Received 19 November 2015
Accepted 23 February 2016

KEYWORDS
Landsat; forest cover change; time-series mapping; forest disturbance; North American forest dynamics; nasa earth exchange

1. Introduction

Time-series analysis of Landsat satellite images offers opportunities and challenges for large area forest cover change monitoring. The Landsat mission offers the longest continuous data record at a spatial resolution suitable for land-based ecosystem studies (Cohen and Goward 2004; Goward and Masek 2001). With the opening of the Landsat data archive and free data availability, the number of studies mining large portions of the archive is rising (Woodcock et al. 2008; Wulder et al. 2012). Image selection is an important step in these land-based Landsat studies with consequences for the quality of derived scientific data (Townshend et al. 2012). Images from outside the growing season, poorly georegistered, with biased radiometric calibration, and/or contaminated by clouds and their shadows will degrade measurement of forest or other vegetation dynamics. Each Landsat

CONTACT Karen Schleeweis kgschleeweis@fs.fed.us Forest Inventory and Analysis, US Forest Service, Rocky Mountain Research Station, 507 25th Street, Ogden, UT 84401, USA.

study has a specific scope, goal and method that may imply potentially competing requirements for image quantity and quality. Therefore, image selection approaches are not necessarily transferable between projects (Franks et al. 2009; Yang et al. 2001). However, strategies for image selection should always exclude data that add more noise than signal to the time-series and track and report the quality of inputs.

Here we consider the major decisions and potential consequences that factor into Landsat image selection for forest cover change and describe the automated image selection process and processing stream used by the North American Forest Dynamics (NAFD) project to create wall-to-wall annual maps of forest history for the contiguous US (CONUS). The science goal of the NAFD project, a core project of the North American Carbon Program, is to characterize the annual and spatial variability in estimates of forest canopy changes in the CONUS between 1986–2010 (Goward et al. 2008; Masek et al. 2013). To this aim, the NAFD project used the L5 and L7 archives to create Landsat time-series stacks (LTSS) (Huang et al. 2009) used to create annual maps of US forest disturbance using the vegetation change tracker (VCT) algorithm (Huang, Goward, et al. 2010). The NAFD forest history maps are available for public download through the ORNL DAAC (Goward et al. 2015).The goals of the NAFD image selection and processing stream (NISPS) include (1) to fully automate the image selection process, (2) to minimize errors in forest disturbance products related to the quality of image inputs, (3) to minimize errors and gaps in forest disturbance products due to lack of imagery and (4) to develop a quality assessment approach to evaluate image selection results.

2. Background

2.1. Seasonality

The timing and duration of temporal windows defining biological growing seasons is an important consideration in many forest cover change studies. Studies such as NAFD that use an approach based on the concept of a steady forest spectral signal during leaf-on periods have only a short window of useable images each year. As latitude increases these windows get shorter. The ideal leaf-on analysis uses images from the stable 'mature' green period of the growing season, minimizing the number of images selected from the variable 'shoulders' of the green-up and senescence periods (Song and Woodcock 2003; Zhang et al. 2001). Some forest cover change algorithms are better at minimizing errors due to small variations in solar geometry and phenology within growing-season image dates (Hansen et al. 2014; Healey et al. 2005; Huang, Goward, et al. 2010).

Recently, more comprehensive approaches have been developed to take advantage of phenological changes in different land-cover types and make use of imagery acquired throughout the entire year (Brooks et al. 2014; Zhu, Woodcock, and Olofsson 2012).

2.2. Clouds

Clouds Pose a significant obstacle to terrestrial monitoring using passive optical wavelength sensors such as those on Landsat (Kaufman 1987). Over the years, various Earth Resources Observation and Science (EROS) Center efforts have been made to assess scene-level cloud cover (CC) in Landsat images. These CC scores, provided as part of the online metadata, allow users to select images with a preference for minimal CC.

The EROS automatic cloud cover assessment (ACCA) algorithms have changed with time and have had varying levels of success. Visual assessment techniques used with the Landsat Multispectral Scanner sensors were replaced with the first ACCA algorithm, a simple three-rule approach developed for Landsat 4 and 5 Thematic Mapper (TM) data. The visual approach overestimated and the first ACCA algorithm underestimated scene-level CC estimates (Hollingsworth et al. 1996). A new ACCA algorithm was derived for Landsat 7 Enhanced Thematic Mapper (ETM+), (Goward et al. 2006; Irish 2000b; Irish et al. 2006). This L7 ACCA was revised to work with Landsat 4 and 5

TM images at EROS and used to replace the original ACCA scores in the TM metadata. Unfortunately, until corrected in 2009, large programming errors in this version of ACCA made the generated CC scores highly inaccurate (John Dwyer, personal communication, http://landsat.usgs.gov/percentage_of_cloud_cover_calculated.php). In January 2009, the ACCA for L5 was corrected and all newly ingested L5 scenes when converted from raw computer format to Level-0 were processed using the corrected ACCA. L5 images processed before January 2009 still have the erroneous ACCA scores. In the absence of a suitable cloud identification algorithm from EROS, the burden of cloud identification and masking becomes the users. Second party cloud assessments must be run on every image under consideration in order to aid with image selection. Cloud identification and masking algorithms have been developed that provide both global and quadrant-based CC estimates and spatially explicit masks of clouds and their shadows (Huang, Thomas, et al. 2010; Zhu and Woodcock 2012). There is a lack of comparisons between available cloud identification and masking algorithms in the literature.

2.3. Image georegistration

Errors can accumulate and become compounded in time-series analyses if the sub-pixel co-registration of multi-temporal images is not adequate (Townshend et al. 1992). Early time-series land studies depended upon the image libraries from the North American Landscape Characterization project (Sohl and Dwyer 1998), Global Land Survey (GLS) (Gao, Masek, and Wolfe 2009; Gutman et al. 2008) or Food and Agriculture Organization of the United Nations remote sensing survey (FAO 2006) projects, or to precisely georegister and orthorectify chosen images themselves, a costly process. Under the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) project, the automated registration and orthorectification package (AROP) process was developed and provided, for the first time, an inexpensive public option for handling large volume Landsat data processing (Gao, Masek, and Wolfe 2009; Masek et al. 2006). In 2008, EROS began publically offering standardized Level-1 terrain-corrected (L1T) products. The L1T product has a nominal geolocation error of less than one 30-m pixel in the CONUS (Irish 2000a; Lee et al. 2004). Though robust, where images have a high proportion of water, large amounts or specific types of CC, or central pivot irrigation, L1T generation can fail because of the lack of identifiable reference points in the imagery (Roy et al. 2010). In these cases, images are offered at with systematic correction (L1G) with a stated geometric accuracy of 250 m for low-relief areas near sea level (http://landsat.usgs.gov/descriptions_for_the_levels_of_processing.php), inadequate for change detection.

2.4. Data gaps

Combining needs for low to no CC and limited seasonal time intervals can severely constrain the number of observations available and leave gaps in time-series data. Methods used to handle the paucity of clear view 'growing season' Landsat images can be loosely grouped into four approaches (with some overlap).

The oldest 'best image' approach uses low CC and a short seasonal window to select a single image from multiple years around a target year. This approach has been used to develop five-to ten-year time-interval data sets (Beuchle et al. 2011; Fry et al. 2011; Gutman et al. 2008, 2013; Lunetta and Sturdevant 1993; Masek et al. 2006; Townshend et al. 1995; Tucker, Grant, and Dykstra 2004).

A 'best pixel compositing' approach, originally applied to course resolution (>250 m) data (Gatlin, Sullivan, and Tucker 1984) and then adapted to Landsat observations (Roy et al. 2010), creates a new composited image by selecting from all images within the target season, the value for each pixel that best meets some selection criteria. Multiple algorithms have been developed that focus on best practices for building composite Landsat data products using a best pixel approach (Griffiths et al. 2013; Hermosilla et al. 2015; White et al. 2014). A modification to this approach uses all images within a single year's growing season to generate time-series pixel-based spectral metrics, without creating composites (Hansen et al. 2013).

An 'image filling' approach replaces in areas identified as cloud, shadows or bad data in a base image with clear portions of additional images. For example, the GLS 2005 and 2010 coverage was generated using this scene-based approach combining up to three images, with a maximum of 20% CC per path-row, to fill data gaps due to the failure of the scan line corrector (SLC) in post-2003 Landsat 7 imagery (Gutman et al. 2008). A within growing season 'image filling' scene-based approach was developed for and applied in this study (Zhao, Huang, Goward, Schleeweis, et al., forthcoming). During the VCT LTSS forest change analysis, masked image areas remaining after clear-view compositing are filled with values from inter-year linear interpolations (Huang, Goward, et al. 2010).

It is important to note that any approach can fail to provide a clear view if there are not sufficient good observations in the archive. However, a fourth approach has been developed to predict reflectance values where they cannot be obtained from the sensor record, using all images available in the Landsat archive and time-series phenology statistics (Zhu, Woodcock, and Olofsson 2012). Approaches that are able to take advantage of the large amount of across-scan overlap in adjacent scenes have considerably more observations per pixel available (for some portions of the WRS 2 – footprint) than approaches that use a scene-based selection approach to compositing (Hansen and Loveland 2012; Kovalskyy and Roy 2013).

3. The NISPSmethod

3.1. Work flow

In 2008, EROS began to offer its millions of Landsat images for free download (Woodcock et al. 2008). Still, few have the dedicated resources to store, unpack and process high volumes of Landsat imagery. The NASA Earth Exchange (NEX) computing environment liberated this NAFD study from many of the computing resource constraints often faced in large-scale land science research (Nemani et al. 2011). NEX, housed at NASA Ames Research Center, was designed to offer more efficient use of Earth observations for NASA Earth science technology, research and applications. NEX provides a collaborative sharing network allowing users to tap into the NASA Advanced Supercomputing facility including its High-End Computing Capability. Currently, NEX holds 1.5 million+ Landsat 5–7 images for use and has the ability to continuously ingest imagery processed by EROS. During the course of the project, the NEX data processing infrastructure was housed in a heterogeneous architecture totaling more than 190,000 cores of Westmere, Sandy Bridge and Ivy Bridge CPUs, high performance Lustre file systems interconnected with InfiniBand technologies in a dual-plane hypercube and high speed networking. The infrastructure uses an advanced scheduling system to efficiently allocate resources and execute processing and computing pipelines. Due to the flexibility of this infrastructure, the NAFD project was able to fully leverage different portions of the NEX architecture on an as-needed basis, depending on the NISPS pipeline components' specific needs and overall usage patterns.

The NISPS used Landsat 5 and 7 imagery from 1984 to 2011, over 434 WRS-2 path-rows[1] to create the 1986–2010 NAFD forest history maps. Though NEX can continuously ingest all Landsat images as they are processed by EROS, it does not mirror the complete archives. In June 2011, we queried the EROS Landsat metadata archive for images in these 434 path-rows acquired between May 1 and November 1 with an ACCA score of less than 50% CC. This step identified more than 22,000 images of interest that were not held by NEX. The images were ordered and processed by EROS and automatically ingested into NEX. In February 2012, we repeated these steps and made a 'base list' of the images held on NEX that met the initial query thresholds. Aware of limited project hours, upcoming changes to the metadata formats of all Landsat images in the EROS archive (July 2012) that our codes were not designed to handle, we decided to 'freeze' the initial base list images to a separate directory on NEX were they could not be overwritten. The NISPS was then only run these images.

The NISPS was built in three modules for independent testing and execution (Figure 1). In an effort to be efficient and conserve computing space and time, the work flow was designed to sequentially cull imagery using image quality criteria before each large processing step. For example, path-row specific growing-season ranges and ACCA CC scores were used to select imagery from the archive before unzipping images and parsing metadata in module 1. Images that were not labeled L1T in the metadata were culled before generating masks and vegetation indices for each top-of-atmosphere reflectance image (TOA). Images were then culled based on the VCT TOA CC estimates before processing to surface reflectance (SR) in module 2, a time and space intensive step. Both individual image characteristics and criteria relative to the 'stack' of images were used. For example, decision rules were used to select a single image in years when more than one cloud-free image

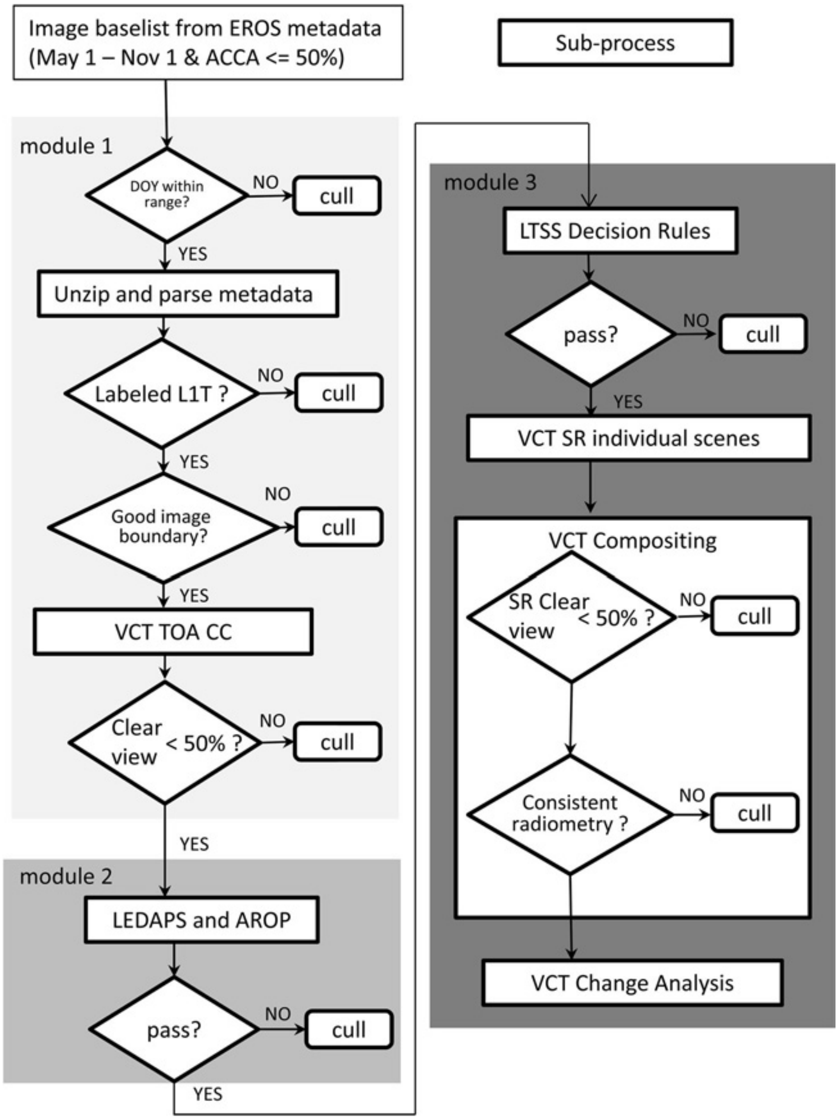

Figure 1. NISPS work flow.

Table 1. Decision rules to select images for each year of the LTSS. decisions are made in sequential order until a selection is made.

Decision rules for intra-year image selection
1. If ≥1 image with ≤5% CC & SLC-ON:
1.1. If 1 image available: select
1.2. If >1 image available: select image with minimum CC
1.3. If >1 image has minimum CC value: select for DOY nearest to center of DOY range for this WRS2 path/row.
2. If 1 image with ≤50% CC AND SLC-on available: select image.
3. If ≥ 1 image with ≤50% CC: select image(s) for cloud clearing
3.1 If only 1 image available: select
3.2 If at least 1SLC-On image available: select all, use SLC-on minimum cc image as base
4. If 0 images with ≤50% cc available: now consider additional images from ± 8 days outside original DOY range for path/row
4.1 Repeat steps 1–3
5. No acceptable imagery: move to next year in time-series

was available (Table 1). Any image with more than 2% of the view obscured by cloud and shadow or missing/bad data was filled with good data images within the same growing season.

The second NISPS module processed nearly 65,000 L1T Landsat images to SR requiring roughly 19,365 CPU hours (on individual cores with 3 GB of memory). In this module, images were converted to SR using the atmospheric correction algorithm implemented in LEDAPS version 2.0 (Masek et al. 2006; Vermote et al. 1997) and georegistration was checked with AROP developed for LEDAPS, (http://www.ars.usda.gov/services/software/download.htm?softwareid=364, version 2.2.5).

The third NISPS module, which operates on SR images and generates VCT forest change analysis maps, generates roughly 27 TB of data for a national run. Four national runs and a handful of regional subsets were needed for algorithm development and testing. Using the NEX platform allowed the NAFD project to effectively address the volumes of data needed for annual wall-to-wall CONUS mapping.

Sequential culling using image quality criteria resulted in fewer images available to pass or fail the next filter. Therefore, counts of failures for each criterion are specific to the NISPS. Some projects have employed a criteria weighting approach to rank and select images (Franks et al. 2009). For NAFD, the costs for designing and using a weighted rather than sequential approach for time-series image selection were estimated to be larger than the benefits.

3.2. Quality assessment

The quality assessment (QA) protocol of the NISPS facilitated evaluation of image selection and processing outcomes. Though the selection and processing of the stream is automated the QA relies on analyst input. We had learned through earlier phases of the NAFD project that there is no substitute for opening images/maps after each production step to verify their quality. The VCT had been run on 50+ path-rows in the CONUS for a previous NAFD study we were aware that this run of 434 path-rows (Masek et al. 2013). To check for errors and anomalies after each image processing step (radiance to TOA, TOA cloud-cover masking, AROP verification, TOA conversion to SR, SR cloud-cover masking, SR clear-view compositing (CVC)) a random selection of images were evaluated by analysts. For example, the spectral trajectories for B3, B4, B5, and B7 of random cloud-free undisturbed evergreen pixels from TOA and SR LTSS were graphed and visually inspected for anomalies. Cloud locations and cover estimates for TOA and SR VCT cloud masks were compared to each other and visual assessments of L1T images. Individual pixels in four quadrants per image were examined to look for pixel shifts before and after running AROP. Finally, random SR composites were examined for edge effects along fill areas, appropriate fill pixel selection, and overall acceptability of the composites. To test the effect of different criteria values for each culling step in the NISPS, QA maps showing the quantity, location, and timing of image 'gaps' in annual national coverage were generated after each step.

3.3. Image criteria for selection

3.3.1. Growing season

To determine 'mature' leaf-on periods and related Julian day of year (DOY) ranges for individual path-rows we employed Normalized Difference Vegetation Index (NDVI) time profiles derived from the *Advanced Very High Resolution Radiometer* (AVHRR) 1 km observation record for the US (Eidenshink 2006; Reed et al. 1994) and compared them with the Terra Moderate Resolution Imaging Spectroradiometer (MODIS) decadal mean daily record. DOY ranges were manually selected to exclude the two shoulders of the mature growing period. Shoulders were defined by their higher standard deviations in the AVHRR record and higher average slopes in the MODIS decadal NDVI profile (Figure 2). Both MODIS and AVHRR NDVI values were subset to forest cover only. In path-rows with less than 5% forest cover, the DOY range was interpolated from the nearest forested path-row neighbor with considerations of continental and coastal climate effects and widely varying elevations.

Choosing specific date ranges based on 16-day composited AVHRR NDVI images reduces temporal resolution and may introduce error into the definition of the growing season (Thayn and Price 2008). After the DOY range was selected, a buffer of ±8 days outside the selected DOY range for a path-row was used to include 'extra' images in the NISPS. These 'extras' were only used if there were no acceptable images within the original DOY range.

3.3.2. Cloud cover

Three CC estimates were used in the NISPS, from ACCA (L5 and L7) scores and from the VCT cloud-mask module on TOA and SR images (Huang, Thomas, et al. 2010). If any of these estimates were above 50% CC the image was culled. Actual VCT output masks include cloud and shadow locations over all imaged areas in a scene. However, the VCT estimates describe the amount of clear view or conversely, bad data, in the scene. These estimates are calculated over land only to improve scene selection in coastal areas. This NAFD analysis was CONUS specific, so the estimates are for only US land area. Scene area affected by SLC issues (L7) and dropped scan lines in early L5 images are also counted in the VCT estimates of bad data per scene. We compared ACCA CC and VCT CC estimates to better understand the impact of using the ACCA estimates to generate base lists from the EROS archive.

3.3.3. Image boundaries

All images used in a LTSS were clipped to the intersection of image boundaries (Huang et al. 2009). If one image in the time-series was anomalously small or irregularly shaped it would result in large portions of useable data being clipped from the entire time-series stack. Images with Upper Left

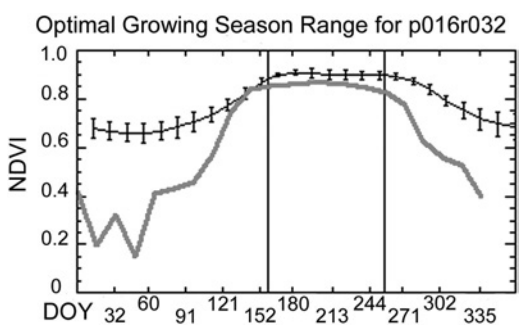

Figure 2. The grey line depicts the MODIS decadal mean NDVI for deciduous forest which cover 39% of the nominal scene boundary in this location (Kim et al. 2011). The black line is the AVHRR twenty year mean NDVI for forest cover. Standard deviation marks the variability among years in the AVHRR annual record (Eidenshink 2006) at the first of each month. Black vertical lines denote the beginning and end of the acceptable DOY range for leaf-on imagery in this location.

and Lower Right coordinates greater than two standard deviations from the mean for all images in the stack were discarded, because they did not overlap enough with other images in the LTSS. This filter, applied in the first NISPS module, culled three L7 images.

3.3.4. Orthorectification

We Found discrepancies between online metadata values for processing level and those in the fields of the metadata text file that were packaged with each EROS processed image. Images not processed to L1T according to the associated metadata text file were culled early in the first NISPS module. In the second NISPS module, images that were converted to SR also had their georegistration double-checked using AROP. AROP compares tie points between an image and its equivalent in the orthorectified GLS 2000 reference image (Gao, Masek, and Wolfe 2009). AROP thresholds used included a minimum acceptable correlation coefficient for a matching tie point of 0.60, a maximum acceptable RMSE for a tie point of 0.75 pixels, and the minimum required number of control points for a registration attempt of 10. If enough tie points in any of the SR L1T image quadrants, could not be matched to their locations in the GLS 2000 reference image, the image was culled from the NISPS.

3.3.5. Radiometry

Criteria for acceptable image radiometery were defined using all years of imagery in the SR LTSS for a path-row. Abnormal radiometry for an individual image was defined as having a mean forest SR value in band 3 or band 4 that was either two standard deviations greater than the average value using all images in the stack or that were outside of a predetermined acceptable range (whichever of the two ranges was narrower). This filter was applied during VCT CVC in the third NISPS module.

4. Results

4.1. An automated streamlined process

During previous NAFD studies to produce sample-based national estimates of forest disturbance (Masek et al. 2013), image selection and pre-processing were the most resource intensive steps in the mapping process, demanding large amounts of analyst hours and subjective decisions. This NISPS was designed to be flexible, repeatable and objective. These improvements reflect the concurrent modernization of the Landsat archive holdings, increases in computing and storage capacity and related need for programmatic image processing, tracking, and QA.

One goal of the NISPS was to reduce algorithm error from multiple sources of noise in the image inputs. Inconsistency of the archive and processing changes are not always well documented. For example, downloadable images were processed with different mixed calibration parameter files and versions of LPGS, but the user would not readily know (Dwyer, personal communication). Small drifts in the reflectance measurements of the ETM+ sensor (Angal et al. 2010) led to a L-0 to L1-T reprocessing in 2012 (John Dwyer, personal communication). We found that between queries of the EDC archive images, some images became newly available, others became unavailable for ordering and that values in the online metadata fields changed or did not match those in the accompanying .mtl files (i.e. from L1T to L1G and vice versa). As of this writing, EROS is working towards a new collection management scheme to, among other things, provide clearly defined radiometric and geometric quality (http://landsat.usgs.gov/about_LU_Vol_9_Issue_6.php).

Consistent radiometric calibration of the Landsat family of sensors provides earth scientists with a rich resource of data (Markham and Helder 2012). The VCT algorithm uses Forest Z-score indices, essentially a distance from the mean SR value for forested pixels in a single year (Huang, Goward, et al. 2010; Huang, Thomas, et al. 2010). These intra-year metrics were originally conceived to help mitigate inter-year variations from varying seasonality and solar zenith angle, which can muddle true change signals. They may also have helped mitigate effects from sensor drift not yet accounted for when this approach was used in early 2012. It is important to consider, but difficult to quantify, the implications of Landsat archive inconsistencies on forest cover change study results and data products. As Landsat moves to a more 'collections' based approach it is also important that best scene approaches be highly repeatable. Currently, the NISPS approach is being updated and used for a Canadian forest mapping study over 339 Landsat WRS-2 locations (1984–2015). The NISPS will be modified to use only SR VCT modules. TOA VCT modules and ACCA scores will no longer be used in the NISPS.

4.2. Balancing image quality and quantity

The NISPS aimed to reduce errors in the time-series analysis introduced by low image quality and by gaps in the annual image stacks. For each year of the analysis (1984–2011), only one image meeting all of the NISPS image criteria was needed per path-row (434), 12,152 images in total. The NISPS selected ~23,000 images for use in the CONUS annual forest disturbance maps using criteria based on time of year, CC, geometric accuracy, and image radiometry (Table 2). Only 67% of the path-row annual time-series maps were created using a single cloud-free image (less than 5% clouds). Another 15,225 images were used to fill clouds and missing data in years when a single growing-season SLC-on image with less than 5% clouds was not available. Over the 28-year period, no satisfactory imagery was available in a particular path-row-year 220 times (1.8% of total), leaving gaps in the annual time-series of forest maps.

It is important to note that the timing and location of image gaps and counts from the NISPS are relative to this study. The order of the sequentially applied NISPS filters determines the number of images available for the next filter. The EROS archive lists 365,565 images for the 434 path-rows used in NAFD. Only 33% of these images were inside the growing-season range used by NAFD.

The NISPS used the same seasonal DOY range for each year. This could be problematic since the NDVI decadal average is a constant and the peak leaf-on range can vary. For example, when there is a persistent severe drought or strong El Niño, the growing-season range can shift for a period of years. The 1987–1989 northern Midwest multi-year drought started before the NDVI record used in our study (1989). Using the full AVHRR NDVI record (Allen et al. 2010) starting in 1982 confirmed that the NISPS date ranges for path-rows affected by the drought were within the peak growing seasons for 1987 and 1988. Though the NISPS growing-season range passed this test, using annual NDVI records to select a year specific growing season for each path-row would alleviate this potential problem.

Table 2. Counts of images that failed and passed each NAFD selection and processing step are listed. The order of the sequentially applied quality filters determines the number of images available for the next filter/processing step.

L5-L7 Images from 434 CONUS path-rows	Culled	Total
Create 'Base List' from EROS archive		
Outside of path/row specific DOY range AND Greater than 50% ACCA CC		91,565
NA	10,591	
Radiances converted to Top of atmosphere (TOA)		80,974
Not L1T	1.60%	
Greater than 50% VCT TOA CC/bad data:	18.80%	
Other	0.20%	
TOA Imagery converted to SR		64,552
Failed AROP orthorectification test:	1.42%	
Culled via decision rules on TOA:	7.50%	
Imagery into SR LTSS stacks for VCT change analysis		26,142
Failed SR radiometry test:	0.40%	
Decision rules on SR:	10.15%	
Imagery used in forest history analysis		23,418
Single image used to create annual map		8193
Images used to create cloud cleared annual map		15,225

Note: The shading separates different modules of the processing stream and relates to the shading in the flow chart (Figure 1).

ACCA scores, known to underestimate scene wide CC (Hollingsworth et al. 1996), were used to save processing time and space and should have resulted in a biased sample of scenes from the Landsat archive. After the ACCA threshold of equal to or less than 50% was applied to the base list of growing-season images, 24% of the archive remained on the base list. Before and after the ACCA filter, the proportion of L5 images (pre 2009) on the base list was roughly 23%. The inaccurate nature of the pre-2009 L5 ACCA scores was not explicitly clear from Landsat documentation available at the inception of the NISPS base list. The 'freezing' of base list images (see Section 3.1) prohibited removing the impact of the ACCA filter from the NISPS.

It is impossible to know exactly how many images within the growing season and with acceptable CC were in the archive, but not selected due of erroneous pre-2009 L5 ACCA scores. Post January 2009, L5 ACCA scores have a better relationship with VCT TOA CC estimates than pre-2009 L5 ACCA (Figure 3(a)). Outliers to the left of the 1:1 lines suggest that there are additional images in the archive with low CC that were not included in the NISPS, especially pre-Jan 2009. VCT SR cloud clover and L7 ACCA scores have a tighter relationship with fewer outliers than do VCT TOA and L7 ACCA scores (Figure 3(b)). Not using ACCA to filter the base list for the NISPS

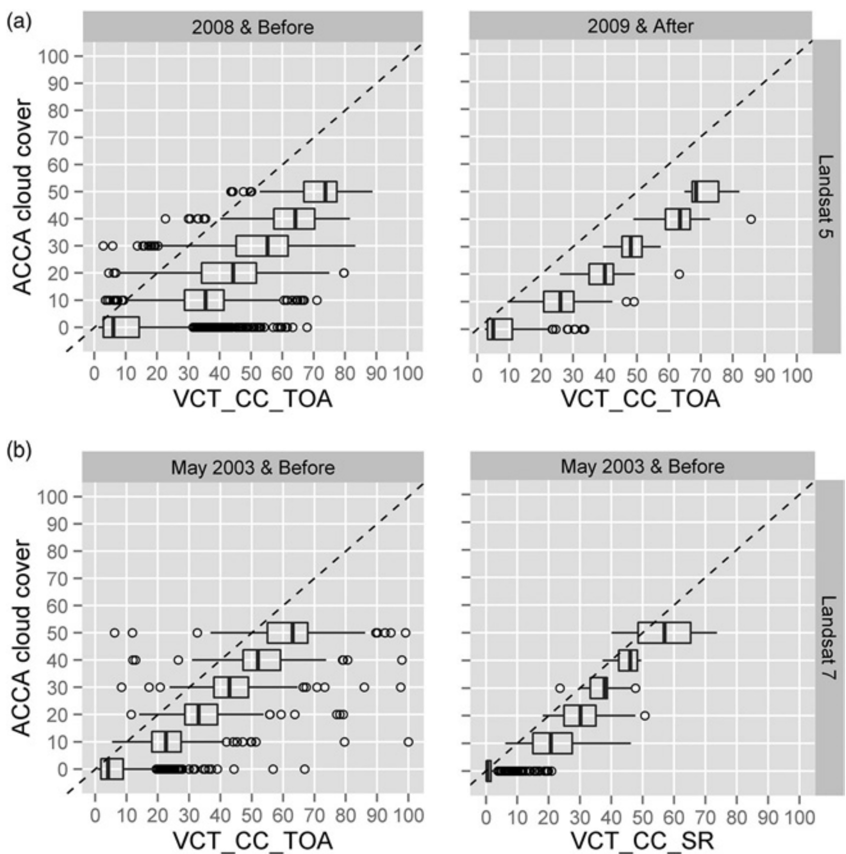

Figure 3. (a) Comparing VCT TOA CC estimates with L5 ACCA algorithm CC estimates from pre-2009 correction ($n = 9929$) and post 2009 corrections ($n = 534$). (b) Comparing VCT TOA ($n = 2425$) and VCT SR ($n = 477$) CC estimates with L7 ACCA algorithm. To account for methodological issues in the comparisons, only images within NAFD growing season, with greater than or equal to 99% US land area in the scene, an ACCA CC of less than or equal to 50%, SLC on and from US ground stations are included. ACCA CC at time of analysis was given only in increments of 10 resulting in a binned distribution. Boxplot shows median line, first and third quartile and 95% confidence interval of median. Black dots are outliers from the 95% confidence interval. The 1:1 line is dashed.

might have yielded better quality images and filled in some of the specific locations and years where there are image coverage gaps.

The geographic and temporal distributions of the number of images on the NISPS base list varied considerably (Figure 4). These distributions are a function of all available Landsat images constrained by: (1) NISPS-defined growing season, (2) only L5 or L7, (3) not processed with the National Land Archive Production System (NLAPS), and 4) having a less than 51% ACCA score. Mostly available in 1984 and 1985, L4 and L5 images generated by NLAPS (~35,000 images in the archive) were excluded from consideration in the NISPS because during this and previous phases of the NAFD project they were found to have more technical issues and less support from EROS and LEDAPS. Roughly half of the more than 22,000 images we ordered from EROS were unavailable or unusable because they were processed by NLAPS or had 'missing headers' (Rachel Headley, personal communication).

The QA map showing the geographic and temporal distribution of image coverage helped identify trends and anomalies in the Landsat archive (Figure 4). For example, image availability increased predictably from north to south as the growing-season lengthens. Image availability increases dramatically in 1999 when L7 began acquiring images. The quantity of images decreased when L7 stopped acquiring data at the onset of the SLC issues (2003).

The location and timing of the 126 path-row years where no images were available from EROS for the NISPS base list can also be seen in Figure 4 as black filled polygons. There are more gaps early in the time-series when fewer images were available. It is possible that the inaccurate pre-2009 L5 ACCA score filtered out images that would have filled one or more of these gaps. Other gaps in the number of images available are not as quickly explained. For example, the reason for image gaps over the South Central portion of the US (covering west Texas) in 1984–1985 and again in 1988–1990 is not readily apparent (Figure 4). Early coverage for the L4 and L5 satellites, before the launch of the Tracking and Data Relay Satellite System, was acquired from either the NASA Goddard Space Flight Center portable ground antenna or the eastern Canada ground station. In either case the 'reach' of the antennae did not extend as far as west Texas, limiting the data acquired for the region (Rachel Headley and Terry Arvidson, personal communication).

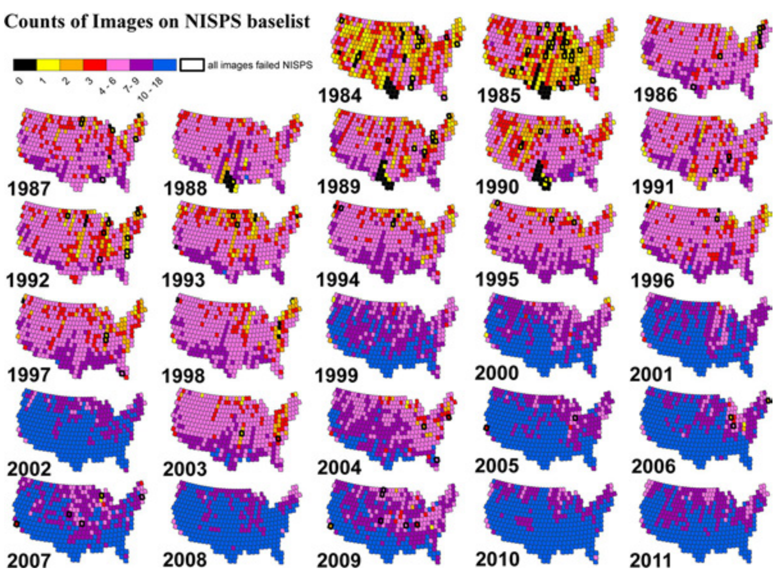

Figure 4. Black polygons on the map depict path rows where, through time, no images were available on the NISPS baselist. Polygons outlined in black represent path rows were no imagery passed the NISPS image quality control filters.

The location and timing of the 94 image coverage gaps due to culling of NISPS base list images using NAFD image selection criterion are mapped as black outlined polygons in Figure 4. Throughout the second and third NISPS modules a higher proportion of the pre-2009 L5 and SLC-off L7 images compared to post January 2009 L5 and SLC-on L7 images (using absolute and relative counts for each of the four image categories) was culled. It is likely more of the pre-2009 L5 images failed NISPS image quality criteria due to high CC not filtered out by the inaccurate ACCA score. Also, more SLC-off images, which lack 22% of image data, would be expected to fail. For example, wherever fewer ground control points (GCPs) are identifiable it is more difficult to accurately and precisely georegister an image. Of the images remaining in the NISPS, 39% of SLC-off L7 and 15% of pre-2009 L5 images failed the VCT TOA threshold of at least 50% clear US land view, compared with 12% each of post January 2009 L5 and SLC-on L7.

Of the images that passed to the second NISPS module, 880 failed the independent AROP georegistration accuracy test (Figure 5). Though not mentioned on the EROS website, a USGS presentation advises '… L1T product does not guarantee a sub-pixel registration – check. Precision fit and verification statistics [are] reported in metadata … ' (Dwyer 2007). Comparisons between image metadata fields, reporting horizontal, vertical and overall RMSE statistics and the number of GCPs used, generated during successful L1T terrain correction for images that passed and failed AROP did not yield a reliable relationship to substitute for the AROP verification step. This analysis was conducted after the 'reprocessing' of L7 in 2012 excluding images with greater than 5% water cover and SLC-off images. The SLC-off L7 category of images had the highest relative rate of AROP failures (6.5%). SLC-on L7, post January 2009 L5 and pre-2009 L5 images had relatively few failures georegistration, 0.2%, 0.3% and 0.7%, respectively.

In this wall-to-wall study, a single growing-season image was used if it met NAFD quality standards and was less than 5% bad data over US land area. Of the 8,193 single images used in the LTSS VCT change analysis, 612 did not meet these criteria, but were used because no other acceptable image was available during the growing season (Figure 6(a)). In 73 of these cases, mostly in 2003 when L7 stopped acquiring due to the SLC problem, only a single SLC-off image was available.

Where no single image met NAFD quality standards and more than one image was available in the growing season a clear-view composite was generated for the LTSS change analysis. The percentage of bad data in the images used for these composites varied (Figure 6(b)). Clear-view compositing increased the number of growing-season path-row image years with less than 0–5% bad US land views by 44% (Figure 6(c)).

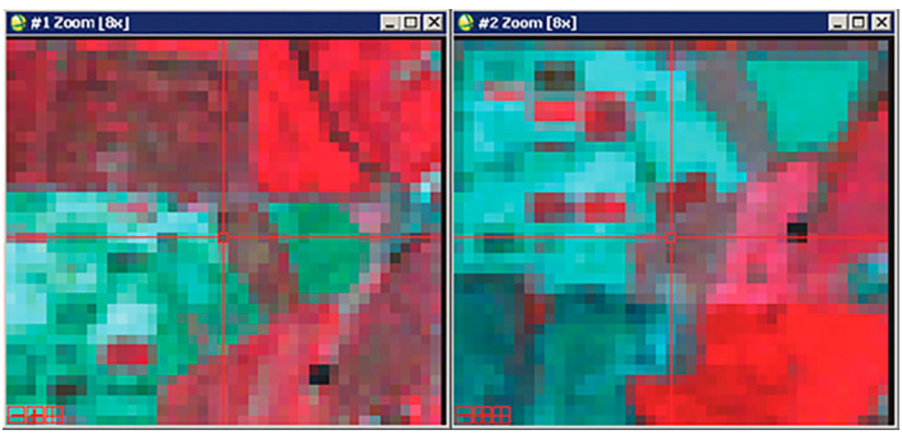

Figure 5. A cloud-free L1T image (LT50350321986193XXX03) that failed the AROP verification test (on the right) compared to the same GLS 2000 image (on the left). The failed image is 3–4 pixels off in the East–West plane and 13–14 pixels off in the North–South plane compared to the GLS image.

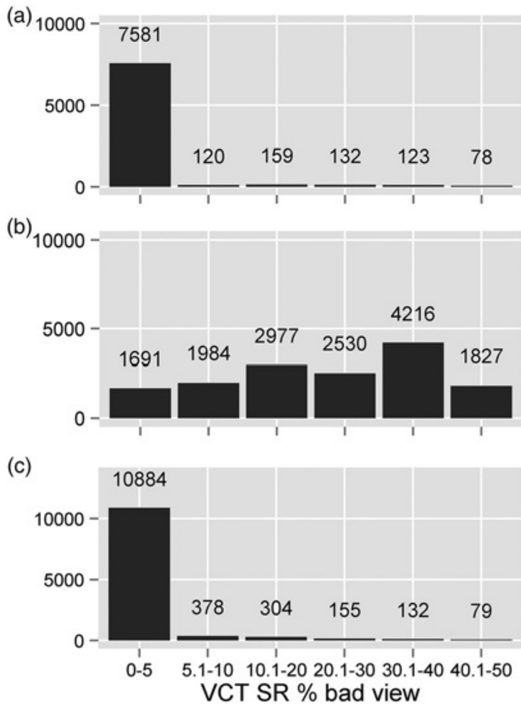

Figure 6. Quality and Quantity of images used for LTSS VCT national mapping. (a) 8,193 single images were used in the LTSS VCT time-series analysis. (b) Where no single image met NAFD quality standards and more than one image was available in the growing season a clear-view composite (CVC) was created. This histogram shows the percentage of bad US land view for images used in CVC. (c) A histogram of the quality, in terms of the percentage bad US land view in final images (single and CVC) used in the wall-to-wall NAFD VCT change analysis.

All land-cover change approaches will fail if there are not sufficient observations, though the relationship between number of observations and magnitude of the measurement error are algorithm specific and often not well characterized. Although the CONUS has the richest archive of Landsat data, the number of available scenes and their quality, in terms of percentage of bad data, still suffer from a paucity of annual growing-season images, particularly early in the time-series and when only one Landsat satellite was operating (Figure 7). This can affect the VCT change analysis in several ways. First, wherever a path-row tile did not have at least one qualified image in a year, VCT could not produce a disturbance map for that tile in that year. This resulted in substantial data gaps in the final wall-to-wall maps in some years. Second, while VCT has built-in mechanisms to flag bad pixels (Huang, Thomas, et al. 2010) and replace them with temporally interpolated values (Huang, Goward, et al. 2010), missing observations during or immediately before or after a disturbance event may result in inaccurate or even erroneous change results, including incorrect change magnitudes, larger uncertainty in disturbance year determination, or completely missing that disturbance event. Further, unflagged poor quality observations in two or more consecutive years over a forest area may result in false disturbance detection. The impact of data inadequacy on change results derived by VCT or other disturbance mapping algorithm needs to be assessed more quantitatively.

The VCT annual results for the first and last year of the time-series are not used in NAFD national results. Due to the high quantity of image gaps and images with high proportions of bad data in the first 2 years of this study (see Figure 4), the NAFD project also dropped 1985 from statistics and will not distribute the annual maps for this year.

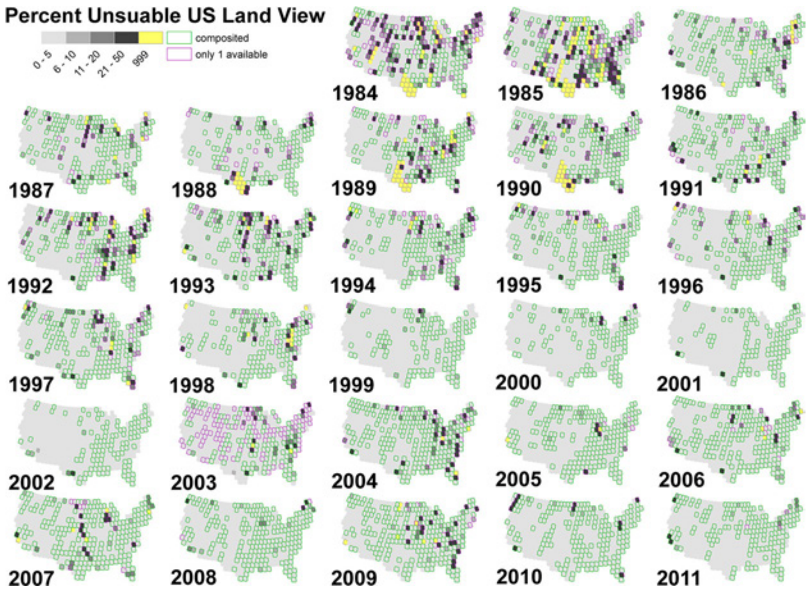

Figure 7. Percentage of unusable data (due to clouds/shadows, dropped scan lines, and bad pixels) in images (per path-row) used for change analysis. Path-rows outlined in green use clear-view composites. Path-rows outlined in purple would have been composited if more than one acceptable image was available. Path-rows filled with yellow had no available growing-season imagery on the NISPS base list.

4.3. Quality assessment protocol

The QA protocol used in the NISPS provided decision support and directed limited project resources between choosing criteria thresholds, solving processing errors, and testing algorithm updates. Though maps of the Landsat archive exist (Franks et al. 2009; Kovalskyy and Roy 2013; Wulder et al., forthcoming) they have focused on different aspects of the archive. The NISPS QA maps provide a novel perspective of the growing-season CONUS annually and help call attention to image coverage, CC, and georegistration issues in the Landsat archive. It is important to note that this project would not have been successful without the open dialogue, updates, advice and input from Landsat scientists at EROS. Routinely opening and inspecting a selection of images after each processing step, helped identify issues with the original VCT TOA and SR CC masks, leading to algorithm revision. Routinely inspecting outputs of each processing step increased confidence that the approach was reducing, not increasing noise from image inputs. Quality assessment of the forest history maps, algorithm modifications, and map post-production steps will be described in a separate manuscript (Zhao, Huang, Goward, Rishmawi, et al., forthcoming).

5. Conclusions

This Paper highlights the decisions, consequences and discoveries often buried within time-series science data products. The NISPS used the Landsat 5 and 7 image archives to map annual wall-to-wall CONUS forest history (1986–2010) to assess the role of forest dynamics in the carbon balance support of forest carbon science. With tools like the NEX environment, computing space and processing capacity are no longer prohibitive constraints on time-series land science using large volumes of satellite data.

Though selection criteria vary for different project goals and methods, the criteria used in the NISPS highlight issues that affect time-series studies relying on Landsat imagery and metadata including inadequate georegistration, unreliability of the early L5 ACCA algorithm, growing-season imagery gaps in space and time, and paucity of clear views. Surprising holes in the archive and a general scarcity of imagery in the beginning of the time-series led to gaps in the NAFD science data products. It is important to note that all projects that have used L5 ACCA scores to select imagery acquired prior January 2009 will have suffered the same unpredictable sampling of the archive. Though EROS is in the process of replacing the Landsat 5 and 7 ACCA algorithms, it is estimated that it will take years before the older erroneous CC estimates are replaced. Gaps and historical inconsistencies in the radiometric and geometric processing in the archive serve as a reminder of the importance of the continuity of the Landsat family of sensors (Irons, Dwyer, and Barsi 2012; Loveland and Dwyer 2012; Wulder et al. 2008). Greater attention needs to be paid to how limited quantity and quality of image inputs affect the magnitude, location and timing of measurement errors. As time-series forest science products are increasingly used to support national and international monitoring programs and political and financial frameworks, this need and the importance of the Landsat legacy will continue to grow.

Note

1. Complete wall-to-wall coverage of the CONUS consists of more than 434 path-rows. Path-rows were excluded if they did not have at least 5% land area in the US, or if they contained only redundant data due to the overlap of neighboring path-rows.

Acknowledgements

We acknowledge the invaluable guidance, data and knowledge provided by Rachel Headley at USGS EROS, and Petr Votava at NASA Ames Research Center. We also are grateful to Louis Keddell and Elaine J. Denning for analysis and support throughout the wall-to-wall mapping effort.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study contributes to the North American Carbon Program, with grant support from NASA's Carbon Cycle Science and Applied Sciences Programs [NNX11AJ78G]. Previous NASA NACP grants [NNG05GE55G] and [NNX08AI26G] were critical in developing the foundations of the current NISPS.

ORCID

Karen Schleeweis http://orcid.org/0000-0002-4673-7265
Chengquan Huang http://orcid.org/0000-0003-0055-9798

References

Allen, Craig D., Alison K. Macalady, Haroun Chenchouni, Dominique Bachelet, Nate McDowell, Michel Vennetier, Thomas Kitzberger, et al. 2010. "A Global Overview of Drought and Heat-Induced Tree Mortality Reveals Emerging Climate Change Risks for Forests." *Forest Ecology and Management* 259 (4): 660–684. http://dx.doi.org/10.1016/j.foreco.2009.09.001.

Angal, Amit, Xiaoxiong Xiong, Tae-young Choi, Gyanesh Chander, and Aisheng Wu. 2010. "Using the Sonoran and Libyan Desert Test Sites to Monitor the Temporal Stability of Reflective Solar Bands for Landsat 7 Enhanced Thematic Mapper Plus and Terra Moderate Resolution Imaging Spectroradiometer Sensors." *Journal of Applied Remote Sensing* 4 (1): 043525–043612.

Beuchle, Rene, Hugh D. Eva, Hans-Jürgen Stibig, Catherine Bodart, Andreas Brink, Philippe Mayaux, Desiree Johansson, et al. 2011. "A Satellite Data set for Tropical Forest Area Change Assessment." *International Journal of Remote Sensing* 32 (22): 7009–7031.

Brooks, Evan B., Randolph H. Wynne, Varghese Thomas, Christine E. Blinn, and John W. Coulston. 2014. "On-the-Fly Massively Multitemporal Change Detection Using Statistical Quality Control Charts and Landsat Data." *Ieee Transactions on Geoscience and Remote Sensing* 52 (6): 3316–3332.

Cohen, W. B., and S. N. Goward. 2004. "Landsat's Role in Ecological Applications of Remote Sensing." *Bioscience* 54 (6): 535–545.

Dwyer, J. 2007. *USGS Landsat Data Continuity Mission (LDCM) Data Product Considerations*. Accessed July 2013. https://landsat.usgs.gov/documents/LDCM_Product_Considerations_Dwyer.pdf.

Eidenshink, Jeff. 2006. "A 16-Year Time Series of 1 km AVHRR Satellite Data of the Conterminous United States and Alaska." *Photogrammetric Engineering and Remote Sensing* 72 (9): 1027–1035.

FAO. 2006. "Global Forest Resource Assesment 2005." FOA Forestry Paper 147. Food and Agricultural Organization of the United Nations, Rome.

Franks, Shannon, J. Masek, R. Headley, John Gasch, and Terry Arvidson. 2009. "Large Area Scene Selection Interface (LASSI) Methodology of Selecting Landsat Imagery for the Global Land Survey 2005." *Photogrammetric Engineering and Remote Sensing* 75: 1287–1296.

Fry, J. A., G. Xian, S. Jin, J. A. Dewitz, C. G. Homer, Y. Limin, C. A. Barnes, et al. 2011. "Completion of the 2006 National Land Cover Database for the Conterminous United States." *Photogrammetric Engineering and Remote Sensing* 77 (9): 858–64.

Gao, Feng, Jeffrey Masek, and Robert E. Wolfe. 2009. "Automated Registration and Orthorectification Package for Landsat and Landsat-Like Data Processing." *Journal of Applied Remote Sensing* 3 (1): 033515–033520.

Gatlin, J. A., R. J. Sullivan, and C. J. Tucker. 1984. "Considerations of and Improvements to Large-Scale Vegetation Monitoring." *Ieee Transactions on Geoscience and Remote Sensing* 22 (6): 496–502. doi:WOS:A1984TN73800005.

Goward, Samuel, Terry Arvidson, Darrel Williams, John Faundeen, James Irons, and Shannon Franks. 2006. "Historical Record of Landsat Global Coverage: Mission Operations, NSLRSDA, and International Cooperator Stations." *Photogrammetric Engineering and Remote Sensing* 72 (10): 1155–1169.

Goward, S. N., C. Huang, F. Zhao, K. Schleeweis, K. Rishmawi, M. A. Lindsey, J. L. Dungan, et al. 2015. *NACP NAFD Project: Forest Disturbance History from Landsat, 1986–2010*. Oak Ridge, TN: ORNL DAAC. http://dx.doi.org/10.3334/ORNLDAAC/1290.

Goward, S. N., and J. G. Masek. 2001. "Landsat – 30 Years and Counting." *Remote Sensing of Environment* 78 (1–2): 1–2.

Goward, S. N., J. G. Masek, W. Cohen, G. Moisen, G. J. Collatz, S. Healey, R. Houghton, et al. 2008. "Forest Disturbance and North American Carbon Flux." *Eos, Transactions American Geophysical Union* 89: 105–116.

Griffiths, Patricia, Sebastian van der Linden, Tobias Kuemmerle, and Patrick Hostert. 2013. "A Pixel-Based Landsat Compositing Algorithm for Large Area Land Cover Mapping." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6 (5): 2088–2101.

Gutman, G., R. Byrnes, J. Masek, S. Covington, C. Justice, S. Franks, and R. Headley. 2008. "Towards Monitoring Land-Cover and Land-Use Changes at a Global Scale: The Global Land Survey 2005." *Photogrammetric Engineering and Remote Sensing* 74 (1): 6–10. doi:WOS:000252077700001.

Gutman, Garik, Chengquan Huang, Gyanesh Chander, Praveen Noojipady, and Jeffrey G. Masek. 2013. "Assessment of the NASA–USGS Global Land Survey (GLS) Datasets." *Remote Sensing of Environment* 134: 249–265.

Hansen, M. C., A. Egorov, P. V. Potapov, S. V. Stehman, A. Tyukavina, S. A. Turubanova, D. P. Roy, et al. 2014. "Monitoring Conterminous United States (CONUS) Land Cover Change with Web-Enabled Landsat Data (WELD)." *Remote Sensing of Environment* 140: 466–484.

Hansen, Matthew C., and Thomas R. Loveland. 2012. "A Review of Large Area Monitoring of Land Cover Change Using Landsat Data." *Remote Sensing of Environment* 122: 66–74. http://dx.doi.org/10.1016/j.rse.2011.08.024.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, and T. R. Loveland. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342 (6160): 850–853.

Healey, S. P., W. B. Cohen, Z. Q. Yang, and O. N. Krankina. 2005. "Comparison of Tasseled cap-Based Landsat Data Structures for use in Forest Disturbance Detection." *Remote Sensing of Environment* 97 (3): 301–310.

Hermosilla, Txomin, Michael A. Wulder, Joanne C. White, Nicholas C. Coops, and Geordie W. Hobart. 2015. "An Integrated Landsat Time Series Protocol for Change Detection and Generation of Annual gap-Free Surface Reflectance Composites." *Remote Sensing of Environment* 158: 220–234. http://dx.doi.org/10.1016/j.rse.2014.11.005.

Hollingsworth, Ben V., Liqiang Chen, Stephen E. Reichenbach, and Richard R. Irish. 1996. "Automated Cloud Cover Assessment for Landsat TM Images." In *Proceedings of SPIE's 1996 International Symposium on Optical Science, Engineering, and Instrumentation*, edited by M. R. Descour and J. M. Mooney, 170–179. Denver, CO: International Society for Optics and Photonics.

Huang, C., S. N. Goward, J. G. Masek, F. Gao, E. F. Vermote, N. Thomas, K. Schleeweis, et al. 2009. "Development of Time Series Stacks of Landsat Images for Reconstructing Forest Disturbance History." *International Journal of Digital Earth* 2 (2): 195–218. doi:10.1080/17538940902801614.

Huang, C., S. N. Goward, J. G. Masek, N. Thomas, Z. Zhu, and J. E. Vogelmann. 2010. “An Automated Approach for Reconstructing Recent Forest Disturbance History Using Dense Landsat Time Series Stacks.” *Remote Sensing of Environment* 114: 183–198.

Huang, Chengquan, Nancy Thomas, Samuel N. Goward, Jeffrey G. Masek, Zhiliang Zhu, John R. G. Townshend, and James E. Vogelmann. 2010. “Automated Masking of Cloud and Cloud Shadow for Forest Change Analysis Using Landsat Images.” *International Journal of Remote Sensing* 31 (20): 5449–5464.

Irish, R. R. 2000a. “Landsat 7 Science Data User’s Handbook, Report 430-15-01-003-0 National Aeruonautics and Space Administration.” *National Aeruonautics and Space Administration*. Accessed August 30, 2015. http://landsathandbook.gsfc.nasa.gov/.

Irish, Richard R. 2000b. “Landsat 7 automatic cloud cover assessment.” In *Proceedings of AeroSense 2000*, edited by S. S. Shen and M. R. Descour, 348–355. Orlando: International Society for Optics and Photonics.

Irish, Richard R., John L. Barker, Samuel N. Goward, and Terry Arvidson. 2006. “Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) Algorithm.” *Photogrammetric Engineering and Remote Sensing* 72 (10): 1179–1188.

Irons, James R., John L. Dwyer, and Julia A. Barsi. 2012. “The Next Landsat Satellite: The Landsat Data Continuity Mission.” *Remote Sensing of Environment* 122: 11–21. http://dx.doi.org/10.1016/j.rse.2011.08.026.

Kaufman, Yoram J. 1987. “The Effect of Subpixel Clouds on Remote Sensing.” *International Journal of Remote Sensing* 8 (6): 839–857.

Kim, D., R. Narashiman, J. O. Sexton, C. Huang, and J. R. G. Townshend. 2011. “Methodology to Select Phenologically Suitable Landsat Scenes for Forest Change Detection.” Proceedings of 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vancover, BC, July 24–29, 2613–2616. IEEE.

Kovalskyy, V., and D. P. Roy. 2013. “The Global Availability of Landsat 5 TM and Landsat 7 ETM + Land Surface Observations and Implications for Global 30 M Landsat Data Product Generation.” *Remote Sensing of Environment* 130: 280–293. http://dx.doi.org/10.1016/j.rse.2012.12.003.

Lee, D. S., J. C. Storey, M. J. Choate, and R. W. Hayes. 2004. “Four Years of Landsat-7 on-Orbit Geometric Calibration and Performance.” *IEEE Transactions on Geoscience and Remote Sensing* 42 (12): 2786–2795. doi:10.1109/tgrs.2004.836769.

Loveland, Thomas R., and John L. Dwyer. 2012. “Landsat: Building a Strong Future.” *Remote Sensing of Environment* 122: 22–29. http://dx.doi.org/10.1016/j.rse.2011.09.022.

Lunetta, R. S., and J. A. Sturdevant. 1993. “The North American Landscape Characterization Landsat Pathfinder Project.” In *Proceedings of Pecora 12 Symposium, Land Information from Space-Based Systems*, edited by L. R. Pettinger, 363–371. Bethesda, MD: American Society of Photogrammetry and Remote Sensing.

Markham, Brian L., and Dennis L. Helder. 2012. “Forty-year Calibrated Record of Earth-Reflected Radiance from Landsat: A Review.” *Remote Sensing of Environment* 122: 30–40. http://dx.doi.org/10.1016/j.rse.2011.06.026.

Masek, J. G., S. N. Goward, R. E. Kennedy, W. B. Cohen, G. G. Moisen, K. Schleeweis, and C. Huang. 2013. “United States Forest Disturbance Trends Observed Using Landsat Time Series.” *Ecosystems* 16: 1087–1104. doi:10.1007/s10021-013-9669-9.

Masek, Jeffrey G., Eric F. Vermote, Nazmi E. Saleous, Robert Wolfe, Forrest G. Hall, Karl F. Huemmrich, Feng Gao, et al. 2006. “A Landsat Surface Reflectance Dataset for North America, 1990–2000.” *IEEE Transactions on Geoscience and Remote Sensing Letters* 3 (1): 68–72.

Nemani, R., P. Votava, A. Michaelis, F. Melton, and C. Milesi. 2011. “Collaborative Supercomputing for Gloal Change Science.” *Eos Transactions* 92 (13): 109–110.

Reed, Bradley C., Jesslyn F. Brown, Darrel VanderZee, ThomasR. Loveland, James W. Merchant, and Donald O. Ohlen. 1994. “Measuring Phenological Variability from Satellite Imagery.” *Journal of Vegetation Science* 5 (5): 703–714. doi:10.2307/3235884.

Roy, David P., Junchang Ju, Kristi Kline, Pasquale L. Scaramuzza, Valeriy Kovalskyy, Matthew Hansen, Thomas R. Loveland, et al. 2010. “Web-enabled Landsat Data (WELD): Landsat ETM+ Composited Mosaics of the Conterminous United States.” *Remote Sensing of Environment* 114 (1): 35–49.

Sohl, Terry L., and John L. Dwyer. 1998. “North American Landscape Characterization Project: The Production of a Continental Scale Three-Decade Landsat Data Set.” *Geocarto International* 13 (3): 43–51.

Song, C. H., and C. E. Woodcock. 2003. “Monitoring Forest Succession with Multitemporal Landsat Images: Factors of Uncertainty.” *IEEE Transactions on Geoscience and Remote Sensing* 41 (11): 2557–2567.

Thayn, J. B., and K. P. Price. 2008. “Julian Dates and Introduced Temporal Error in Remote Sensing Vegetation Phenology Studies.” *International Journal of Remote Sensing* 29 (20): 6045–6049.

Townshend, J. R. G., V. Bell, A. C. Desch, C. Havlicek, C. O. Justice, W. E. Lawrence, D. Skole, et al. 1995. “The NASA Landsat Pathfinder Humid Tropical Deforestation Project. Land Satellite Information in the Next Decade.” Paper presented at the American Society for Photogrammetry and Remote Sensing, Tysons Corner, Virginia.

Townshend, J. R. G., C. O. Justice, C. Gurney, and J. McManus. 1992. “The Impact of Misregistration on Change Detection.” *Ieee Transactions on Geoscience and Remote Sensing* 30 (5): 1054–1060.

Townshend, J. R., J. G. Masek, C. Q. Huang, E. F. Vermote, F. Gao, S. Channan, J. O. Sexton, et al. 2012. "Global Characterization and Monitoring of Forest Cover Using Landsat Data: Opportunities and Challenges." *International Journal of Digital Earth* 5 (5): 373–397. doi:10.1080/17538947.2012.713190.

Tucker, C. J., D. M. Grant, and J. D. Dykstra. 2004. "NASA's Global Orthorectified Landsat Data set." *Photogrammetric Engineering & Remote Sensing* 70 (3): 313–322.

Vermote, Eric F., Didier Tanré, Jean L. Deuze, Maurice Herman, and J.-J. Morcette. 1997. "Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: An Overview." *IEEE Transactions on Geoscience and Remote Sensing* 35 (3): 675–686.

White, J. C., M. A. Wulder, G. W. Hobart, J. E. Luther, T. Hermosilla, P. Griffiths, N. C. Coops, et al. 2014. "Pixel-Based Image Compositing for Large-Area Dense Time Series Applications and Science." *Canadian Journal of Remote Sensing* 40 (3): 192–212. doi:10.1080/07038992.2014.945827.

Woodcock, C. E., R. Allen, M. Anderson, A. Belward, R. Bindschadler, W. Cohen, F. Gao, et al. 2008. "Free Access to Landsat Imagery." *Science* 320 (5879): 1011. doi:10.1126/science.320.5879.1011a.

Wulder, Michael A., Jeffrey G. Masek, Warren B. Cohen, Thomas R. Loveland, and Curtis E. Woodcock. 2012. "Opening the Archive: How Free Data has Enabled the Science and Monitoring Promise of Landsat." *Remote Sensing of Environment* 122: 2–10. http://dx.doi.org/10.1016/j.rse.2012.01.010.

Wulder, M. A., J. C. White, S. N. Goward, J. G. Masek, J. R. Irons, M. Herold, W. B. Cohen, et al. 2008. "Landsat Continuity: Issues and Opportunities for Land Cover Monitoring." *Remote Sensing of Environment* 112 (3): 955–969. doi:10.1016/j.rse.2007.07.004.

Wulder, Michael A., Joanne C. White, Thomas R. Loveland, Curtis E. Woodcock, Alan S. Belward, Warren B. Cohen, Eugene A. Fosnight, et al. Forthcoming. "The Global Landsat Archive: Status, Consolidation, and Direction." *Remote Sensing of Environment*. http://dx.doi.org/10.1016/j.rse.2015.11.032.

Yang, Limin, Collin Homer, Kent Hegge, Chengquan Huang, Bruce Wylie, and Bradley Reed. 2001. "A Landsat 7 Scene Selection Strategy for a National Land Cover Database." Proceedings of IEEE 2001 International Geoscience and Remote Sensing Symposium, Sydney, July 9–13, 1123–1125. IEEE.

Zhang, Xiaoyang, John C. F. Hodges, Crystal B. Schaaf, Mark A. Friedl, Alan H. Strahler, and Feng Gao. 2001. "Global Vegetation Phenology from AVHRR and MODIS Data." Proceedings of IEEE 2001 International Geoscience and Remote Sensing Symposium, Sydney, NSW, July 9–13, 2262–2264. IEEE.

Zhao, F., C. Huang, S. N. Goward, K. Rishmawi, J. G. Masek, M. A. Lindsey, E. Denning, et al. Forthcoming. "Development of Annual Map Products of US Forest Disturbance History Using Landsat Data."

Zhao, F., C. Huang, S. N. Goward, K. Schleeweis, K. Rishmawi, J. G. Masek, M. A. Lindsey, et al. Forthcoming. "Development of Annual Map Products of US Forest Disturbance History Using Landsat Data."

Zhu, Zhe, and Curtis E. Woodcock. 2012. "Object-Based Cloud and Cloud Shadow Detection in Landsat Imagery." *Remote Sensing of Environment* 118: 83–94.

Zhu, Zhe, Curtis E. Woodcock, and Pontus Olofsson. 2012. "Continuous Monitoring of Forest Disturbance Using All Available Landsat Imagery." *Remote Sensing of Environment* 122: 75–91. http://dx.doi.org/10.1016/j.rse.2011.10.030.