ISPRS Journal of Photogrammetry and Remote Sensing 146 (2018) 108–123

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies

Limin Yang[a,*], Suming Jin[b,*], Patrick Danielson[c], Collin Homer[d], Leila Gass[e], Stacie M. Bender[f], Adam Case[i], Catherine Costello[g], Jon Dewitz[d], Joyce Fry[c], Michelle Funk[h], Brian Granneman[c], Greg C. Liknes[f], Matthew Rigge[b], George Xian[d]

[a] *Consultant, Stinger Ghaffarian Technologies, Contractor to U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd Street, Sioux Falls, SD 57198, USA*
[b] *ASRC Federal InuTeq, Contractor to U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd Street, Sioux Falls, SD 57198, USA*
[c] *Stinger Ghaffarian Technologies, Contractor to U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd Street, Sioux Falls, SD 57198, USA*
[d] *U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd Street, Sioux Falls, SD 57198, USA*
[e] *U.S. Geological Survey, Western Geographic Science Center, Menlo Park, CA 94025-3561, USA*
[f] *USDA Forest Service, Northern Research Station, 1992 Folwell Ave., Saint Paul, MN 55108, USA*
[g] *U.S. Geological Survey, Geosciences & Environmental Change Science Center, Denver, CO 80225, USA*
[h] *U.S. Geological Survey, Eastern Geographic Science Center, Reston, VA 20192, USA*
[i] *Innovate! Inc., Contractor to U.S. Geological Survey, Earth Resources Observation and Science Center, 47914 252nd Street, Sioux Falls, SD 57198, USA*

ARTICLE INFO

Keywords:
United States National Land Cover Database 2016
Multi-temporal land cover and change
Methodology
Implementation strategies

ABSTRACT

The U.S. Geological Survey (USGS), in partnership with several federal agencies, has developed and released four National Land Cover Database (NLCD) products over the past two decades: NLCD 1992, 2001, 2006, and 2011. These products provide spatially explicit and reliable information on the Nation's land cover and land cover change. To continue the legacy of NLCD and further establish a long-term monitoring capability for the Nation's land resources, the USGS has designed a new generation of NLCD products named NLCD 2016. The NLCD 2016 design aims to provide innovative, consistent, and robust methodologies for production of a multi-temporal land cover and land cover change database from 2001 to 2016 at 2–3-year intervals. Comprehensive research was conducted and resulted in developed strategies for NLCD 2016: a streamlined process for assembling and pre-processing Landsat imagery and geospatial ancillary datasets; a multi-source integrated training data development and decision-tree based land cover classifications; a temporally, spectrally, and spatially integrated land cover change analysis strategy; a hierarchical theme-based post-classification and integration protocol for generating land cover and change products; a continuous fields biophysical parameters modeling method; and an automated scripted operational system for the NLCD 2016 production. The performance of the developed strategies and methods were tested in twenty World Reference System-2 path/row throughout the conterminous U.S. An overall agreement ranging from 71% to 97% between land cover classification and reference data was achieved for all tested area and all years. Results from this study confirm the robustness of this comprehensive and highly automated procedure for NLCD 2016 operational mapping.

1. Introduction

1.1. History and recent activities in large area land cover database development

The late 1990s to early 2010s witnessed several pioneering and formative developments in global and national land cover datasets using coarse resolution (~1 km) remotely sensed data. The most significant are The International Geosphere-Biosphere Programme (IGBP) DISCover global land cover database (Loveland and Belward, 1997) developed by U.S. Geological Survey (USGS)/University of Nebraska (Loveland et al., 2000), University of Maryland Global land cover data (Hansen et al., 2000), the National Aeronautics and Space Administration (NASA) Earth Observing System's MODIS global land cover (by

* Corresponding authors.
E-mail addresses: lyang0117@yahoo.com (L. Yang), suming.jin.ctr@usgs.gov (S. Jin).

https://doi.org/10.1016/j.isprsjprs.2018.09.006
Received 21 May 2018; Received in revised form 7 August 2018; Accepted 5 September 2018

Friedl et al., 2002), Global Land Cover 2000 led by the European Commission's Joint Research Centre (Bartholomé and Belward, 2007), and GLOBCOVER 2005 -led by European Space Agency with International collaborators (Arino et al., 2008). The United States National Land Cover Database (NLCD) was initiated during this era to bring national large-area land cover classification to a medium resolution (30-m) in order to make land cover information more relevant to national applications (Vogelmann et al., 2001). Following the success of NLCD, international communities have continued to advance medium resolution (~30-m) land cover product development at global scales, including the 30-m Global Land Cover (Globeland30) by the China land cover team (Chen et al., 2014, 2015), Fine Resolution Observation and Monitoring of Global Land Cover (FROM_GLC) by Tsinghua University of China (Gong et al., 2013), Global Mangrove Forests Data (Giri et al., 2011), Global Forest Change 2000–2014 (Hansen et al., 2013), Global Landsat Tree Cover (Sexton et al., 2013a), Global Forest Cover Change of 1975, 1990, 2000, and 2005 (Kim et al., 2014), Global Human Settlement of 1975, 1990, 2000, 2014 (Pesaresi et al., 2016), Global Urban Footprint circa 2002 (Esch et at., 2013), Global Impervious Cover 2010 (Song et al., 2016), Global Inland Water Body circa 2000 (Feng et al., 2015), and Global Cropland Extent 2015 (Teluguntla et al., 2017). These activities were driven by an increasing demand for more accurate, higher spatial resolution, and up-to-date land cover datasets required by the global change and land management user communities. The successful completion of those products was attributable primarily to the improved technological and algorithm advancement for land cover characterization, and the ever-increasing availability of multi-temporal and multi-resolution remote sensing and geospatial datasets.

1.2. Emerging trends and challenges in large area land cover monitoring from remote sensing

Until the late 2000s, most large-area (national or global) comprehensive land cover and land cover change monitoring based on medium resolution images (~30-m) was conducted using a conventional change detection method between two points in time. Since 2009, the opening of the USGS Landsat data archive has enabled a new paradigm for advancing land change science (Woodcock et al., 2008; Wulder et al., 2018). The paradigm promotes a new approach to land change monitoring by extending from simple change detection at a bi-temporal scale to a multi-temporal scale (Jin and Sader, 2005; Latifovic and Pouliot, 2005; Kennedy et al., 2007; Huang et al., 2010; Zhu et al., 2012; Sexton et al., 2013b; Franklin et al., 2015). Such an approach can generate land cover products that depict more complex spatial and temporal land cover condition and changes caused by natural or anthropogenic driving forces. The multi-temporal datasets enable a better understanding of land cover dynamics and the implications of these changes on land resources management and ecosystem services.

There are several challenges for realizing this new paradigm as it relates to large-area land cover and change monitoring. From a thematic perspective, it has long been recognized that there is spectral variation within a single land cover type and spectral similarities among different land cover types (e.g., different types and practices of cultivated croplands, and forested wetlands versus upland forest). These have posed great challenges for spectral-based land cover classification. The challenges can be further compounded when a long time-series land cover and change product is targeted. From a temporal perspective, the quality of each individual land cover map in a time series has a direct impact on the accuracy of mapped land cover and change (no-change). The errors and inconsistency in multi-temporal time-series land cover and change maps due to differences in class definition, input data, and methods can lead to illogical and false land cover changes (Latifovic and Pouliot, 2005; Sexton et al., 2013b; Franklin et al., 2015). Consequently, they may yield unreliable estimates of land cover change rate and change trajectory and have a direct and negative impact on the accuracy of the product.

Another challenge is to accurately map various land cover types and changes over vast and complex landscapes subject to various land use and management practices. Some land cover patterns are spatially unique in shape and size, and changes occur at a confined spatial and temporal scale (e.g., a few pixels of water body or a stream), while other types are spatially clustered and confined within areas of certain geometry or terrains (e.g., forest cut or irrigated cropland, objects formed by a group of pixels). Mapping diverse land cover classes and changes in a large region requires spatially and temporally representative training data and a need to achieve a balance between maintaining the spatial coherence of certain land cover types while keeping single pixel level information for other types. In addition, mixed pixels are a challenge for spectral-based classification algorithms. Separating changes between land cover condition and land cover conversion over large and diverse landscapes often requires special treatment and strategies beyond the conventional spectral-only change detection. Under such conditions, geographic ancillary data and local knowledge about the landscape and natural environment, vegetation dynamics, and land use practices can all be used to improve the accuracy of either classification or the post-classification processes (Srinivasan and Richards, 1990; Brown et al., 1993; Jin et al., 2013; Chen et al., 2014). In essence, to achieve a high accuracy in large area time series land cover and change mapping, careful integration of multi-temporal, multi-spectral, and geospatial data and knowledge is necessary.

1.3. Review of U.S. National land cover database development

The United States NLCD had its beginnings in the mid-1990s with the formation of the Multi-Resolution Land Characteristics (MRLC) Consortium by the USGS, the U.S. Environmental Protection Agency (EPA), and the National Oceanic and Atmospheric Administration (NOAA). Additional MRLC partners beyond the three originals now include the U.S. Department of Agriculture (USDA) Forest Service (USFS), National Agricultural Statistical Service (NASS), the Bureau of Land Management (BLM), the National Park Service (NPS), the U.S. Fish and Wildlife Service (USFWS), and the Army Corps of Engineers (USACE) (Wickham et al., 2014). For NLCD, the 1992 product was the first land cover dataset at 30-meter resolution ever produced for the 48 conterminous states with a consistent, coast-to-coast methodology. By the 2001 release, NLCD had evolved to a database concept with multiple products including land cover, percent tree canopy, percent imperviousness, and database derivatives of Landsat imagery, elevation data and derivatives, other ancillary and intermediate datasets, and metadata and other supporting information (Homer et al., 2004). For the 2006 release, NLCD began quantifying land cover change over time (Fry et al., 2011). NLCD 2011, which was released in 2013, represents a decade of consistently produced land cover and impervious surface for the Nation across three periods: 2001, 2006, and 2011 (Fry et al., 2011; Homer et al., 2015). Overall, the USGS, in partnership with several federal agencies, has developed and released four NLCD product databases over the past two decades: NLCD 1992, 2001, 2006, and 2011. These databases provide spatially explicit and reliable information on the Nation's land cover and land cover change.

2. Requirements for a new generation NLCD 2016

Despite several successful data releases, there remains a fundamental need across government and private sectors for more timely, accurate, and relevant products. In addition, there is increasing demand for products that better represent shrub and grass ecosystems than past NLCD land cover classes. Hence, the NLCD team has responded to this need with an NLCD 2016 database design that produces accurate land cover change information more cohesively and consistently by correcting legacy errors in NLCD products; including additional products for shrub, grass, and bare ground and additional forest disturbance classes; and executing this in a production model faster than any previous NLCD.

3. Objectives and components of NLCD 2016

The aim of NLCD 2016 is to develop a nationally consistent multi-temporal land cover and land cover change database at 30-m spatial resolution from 2001 to 2016 at 2–3-year intervals. The database will enable the monitoring of a wide variety of land cover and land cover changes (e.g., urban development, vegetation disturbance and succession), and will facilitate scientific understanding of causes and consequences of these changes in a consistent and credible way. The database will also provide input data to a variety of environmental models and allow simulation of many natural and anthropogenic processes that are not directly observable.

The NLCD 2016 database design was based on several guiding principles, including the need to (1) upgrade the original NLCD 2001 and correct base errors in the land cover product that persisted over several epochs from 2001 to 2011, (2) develop consistent and accurate land cover and change products from 2001 to 2016, (3) design procedures that capitalize on data-rich and technological advancement to automate NLCD production at a lower cost, (4) ensure that the new design works seamlessly to provide an integrated NLCD 2016 database for the nation, and facilitates and supports a more rapid update cycle for NLCD beyond 2016, and (5) develop additional fractional vegetation products for shrub, herbaceous, and bare ground and integrate them into the land cover product.

NLCD 2016 will consist of the following major components, all at 30-m resolution:

(1) A cloud-free time series Landsat imagery dataset from 2001 to 2016 at 2–3-year intervals for the conterminous U.S.
(2) A new set of 7-epoch NLCD land cover and change products from 2001 to 2016 at 2–3-year intervals.
(3) A percent imperviousness and change trajectory product for 2001, 2006, 2011, and 2016.
(4) A percent tree canopy cover product for 2011 and 2016.
(5) A suite of percent shrub, herbaceous, and bare ground products for 2016.

4. Research priorities in developing NLCD 2016

NLCD 2016 is envisioned as a new generation of land cover and land cover change product that further enhances information content, consistency, utility, and relevancy of NLCD. Here we highlight several research priorities we focused on throughout the development of the NLCD methodology: (1) integration of multi-source information for land cover characterization and change detection, (2) leveraging expert knowledge and ancillary data to support land cover and change modeling, (3) a hierarchical approach for land cover and change analyses, and (4) integration of a pixel-based and object-based land cover modeling. The following sections discuss each of the four research priorities.

4.1. Integration of multi-spectral, multi-temporal, and spatial information

The general procedure for NLCD 2016 land cover characterization and change detection consists of four steps: assemble training data, land cover classification, post-classification, and final integration. For each step we integrated spectral, temporal, spatial, and land cover trajectory data for help in building models. For assembling training data, we used multi-temporal Landsat images to generate spectral change maps, along with other land cover datasets [e.g., NLCD legacy data, USDA Cropland Data Layer (CDL), a National Wetland Inventory (NWI) dataset] to derive initial training data. We also created spatial objects via image segmentation and used them to refine the training data. At the classification step, we incorporated spectral, temporal, and spatial and terrain information as independent variables for the input to a decision tree classifier to create a pixel-based land cover map for each epoch. We also employed spatial objects to derive an object-based land cover map. At the post-processing and final integration step, we incorporated pixel-based and object-based land cover maps as well as ancillary data to improve the initial land cover label, and to produce the final land cover and change maps.

4.2. Expert knowledge and ancillary data

The important role of expert knowledge and ancillary data in characterizing land cover and changes has been demonstrated (Shafer and Logan, 1987; Srinivasan and Richards, 1990; Jin et al., 2013; Chen et al., 2014). Expert knowledge about land cover and change trajectories can be expressed as rules and/or attributes and built into a system to aid in land cover modeling and classification. Geospatial ancillary data have been recognized as a valuable source of information for large area land cover characterization (Srinivasan and Richards, 1990; Brown et al., 1993) and for land cover change analysis (Lu et al., 2007; Jin et al., 2013). For NLCD 2016, we have collected and prepared many ancillary datasets to address different issues related to particular land cover classes. The ancillary data were used not only for developing training data but also for land cover classification and post-processing. We used published products, as well as data and rules that we created based on expert knowledge. For example, we characterized the long-term dynamics of shrubland and herbaceous cover in the western United States based on historical records of fire extent and burn severity, the ecological conditions of the site (vegetation type, temperature, precipitation, and soil type), and the expected time of vegetation recovery after fire. We also determined the timing of forest disturbance and associated land cover types post-disturbance based on magnitude of spectral changes and the known trajectory of forest recovery at a given site.

4.3. A hierarchical approach to mapping land cover and change

A hierarchical approach offers advantages by adopting different training data creation, classification, and post-processing algorithms/models to handle different land cover classes separately in order to minimize the spectral and spatial confusion (Sulla-Menashe et al., 2011; Smith, 2013; Chen et al., 2014). For example, the NLCD production mapped urban developed areas and changes using a regression-tree algorithm, which was separate from the classification of all other land cover types (Yang et al., 2003a, 2003b; Xian and Homer, 2010). For NLCD 2016, we developed new models that were tailored to each particular land cover type. At the training data assembling step, we used different models to create training data for each land cover class. At the classification step, we employed a decision tree algorithm for land cover classification and a regression-tree algorithm for mapping urban and rangeland classes. At the post-classification stage, we built specific models for each land cover type to correct errors due to temporal and spatial inconsistency in class labels, and executed these models sequentially, with one class at a time.

4.4. Integration of pixel-based and object-based land cover modeling

Mapping diverse land cover classes and changes requires a balance between maintaining the coherence of multi-pixel object for certain land cover types and keeping single-pixel level information for other types. Several researchers have explored integration of pixel-based and object-based land cover mapping (Myint et al., 2011; Smith, 2013; Costa et al., 2014; Chen et al., 2014). We adopted that strategy for NLCD 2016 land cover processing. A general rule we followed was to use an object-based approach for mapping change of natural vegetation and agricultural classes (with relatively large patch size) to maintain the spatial coherence, to use a pixel-based approach for water, snow/

ice, and developed classes to retain their spatial details, and to integrate object-based and pixel-based approach for no change areas. This approach facilitates an optimized and consistent multi-temporal land cover and change product.

5. Implementation and methods

5.1. Methods for image pre-processing

One fundamental requirement for development of a long-term, consistent national land cover and land cover change database is a well-calibrated, spatially and temporally consistent Landsat imagery dataset for the Nation. Research on image pre-processing strategies has led to development of (1) a protocol for image selection, (2) a cloud, shadow, and scene anomaly mask, and (3) a cloud and shadow and gap fill procedure.

5.1.1. Landsat image selection

We created a script that searched through the Landsat image archive and downloaded all images that met the predefined threshold of cloud cover (< 20%). For each Landsat path/row, one cloud-free leaf-on image (hereafter referred to as a base image) was selected for each target year (2001, 2003, 2006, 2008, 2011, 2013, and 2016). For 2016, a leaf-off image was also selected. If there was no cloud-free image available for a target year, an alternate cloud-free image acquired from a year either before or after the target year was selected. In some cases when selected base images had clouds or some anomalies, additional image(s) (hereafter referred to as a fill images) were chosen and later used to fill cloud/shadow areas in the base image. Up to three fill images were selected as needed, and these images could be either leaf-on or leaf-off, preferably within two years of the base image date.

5.1.2. Cloud masking and filling

Clouds, cloud shadows, smoke from active fires, and other anomalies were delineated by drawing an area of interest around each anomaly in the affected base and fill images. New spectral data in masked areas were estimated using a cloud filling method (Jin et al., 2013). This method is based on the concept of the Spectral Similarity Group (SSG), which uses the fill image to find similar pixels in the base image. Pixels that have the same SSG from the reference image are projected (based on geographic coordinates) to the base image, and the mean values of those pixels from the base image are calculated and used to fill the cloud/shadow pixels.

5.2. Method and models for land cover characterization and land cover change modeling

Research conducted for NLCD 2016 land cover modeling focused on developing methods for (1) multi-temporal spectral change detection; (2) assembling training data for multi-temporal land cover modeling; (3) generating an initial set of multi-temporal land cover classification map based on training data and a decision-tree algorithm, (4) a post-classification strategy to improve the initial land cover and change maps, and (5) an integration process to generate the final NLCD 2016 land cover and change product. The following sections elaborate on each of the five elements in detail.

5.2.1. Methods for multi-temporal spectral change detection

In the past, NLCD has been produced either as one single dataset (NLCD 1992) or as an updated database from a baseline dataset (NLCD 2001, 2006 and 2011) using a bi-temporal change detection method named the Multi-Index Integrated Change Analysis (MIICA) (Jin et al., 2013). MIICA utilizes four spectral change indices to produce a change map with two classes of biomass increase and decrease for two-dates of Landsat imagery. Besides using the MIICA model, for NLCD 2016 we utilized multi-temporal Landsat images to detect and quantify long-term spectral change at 2–3-year intervals from 2001 to 2016, where either land cover type or condition change occurred. Several existing and newly developed spectral indices were utilized: (1) the Normalized Burn Ratio (NBR), (2) the Normalized Difference Vegetation Index (NDVI), (3) Change Vector (CV), (4) Relative Change Vector (RCV), and (5) the Normalized Spectral Distance (NSD). These spectral indices were used to detect spectral change over a multi-temporal domain from 2001 to 2016. In addition, a disturbance map produced by Vegetation Change Tracker (VCT) (Huang et al., 2010) was used to extend the multi-temporal spectral change back to 1986 to form a longer time series disturbance dataset. This 1986–2016 disturbance dataset plays a very important role in the following training data creation, classification, and post-processing.

5.2.2. Methods for multi-temporal land cover modeling

Mapping multi-temporal land cover and changes requires consistent and robust method. We designed a two-step processing procedure: (1) assembling spectrally, spatially, and temporally consistent training data for each epoch year (every 2–3 years from 2001 to 2016), and (2) conducting land cover modeling/classification for each epoch. First, a set of models was developed to assemble a training dataset for each land cover type of each epoch year starting from 2016. The training data were created based on Landsat images and derived indices, multiple spectral change products from change detection, trajectory analysis, and a variety of ancillary data. The ancillary data included NLCD legacy data 2001, 2006, and 2011, the NOAA Coastal Change Analysis Program land cover data, USDA CDL and an accumulated Cultivated Crop Layer, hydric soil and NWI dataset. In addition, image objects derived from each individual Landsat image were used to refine pixel-based training data to mitigate noise in the classification.

After all training data were prepared, the second step was to conduct a land cover classification for each epoch year using a decision-tree classifier called C5 (Breiman et al., 1984; Quinlan, 1993). The C5 classifier employs an information gain ratio method in tree development and pruning, and has many advanced features including boosting and cross-validation. Specifically, the classification process implemented for NLCD 2016 involves (1) drawing training samples (2% of all available training data per path/row), validation samples (1%), and a minimum number of 5,000 samples per class, (2) executing the C5 classification algorithm to generate a set of rules based on the training data, and (3) applying the decision rules to generate a land cover classification map. Our C5 classifications used four set of independent variables (1986–2016 disturbance year map at 2–3-year intervals, Landsat image of the year, compactness of Landsat image segmentation polygons, and a DEM and its derivatives). The classification was conducted for every epoch year twice (a full version and a light version). The light version excluded the urban and wetland classes from the dependent variables and the 1986–2016 disturbance year from the independent variables. Two versions of classifications were then integrated with ancillary data and the object-based information. A total of seven initial land cover maps were generated in a back-in-time order with the past NLCD legacy data year taking precedence (i.e. 2016, 2011, 2006, 2001, 2003, 2008, and 2013).

5.2.3. Methods and models developed for post-classification

There were errors in the initial land cover classification maps and inconsistency in land cover change sequence due to the quality of input data and limitation of the automated classification. A post-classification process was implemented to correct the errors in land cover maps for each year as well as the temporal inconsistency in the time series land cover maps from 2001 to 2016. The post-processing focused on checking the spatial coherence of land cover labels for each epoch, temporal consistency of land cover labels over time, and logic of land cover change trajectory. The process utilized information from spectral and spatial data, temporal change trajectory, expert-knowledge, and ancillary data to refine the initial land cover and change labels through a set of rules.

Table 1
Example of main issues and data used for rule-based post-processing.

Land cover theme	Issues	Data and method used for Post-processing
Water	*Confusion with wetland*	• *The frequency of water exists over seven epochs* • *The extent of water at a given epoch* • *The maximum extent of water over seven epochs* • *Ecoregion map*
	Mixed pixels near border of water body	• *The frequency of water exists over seven epochs* • *The extent of water at a given epoch*
	Differentiate ephemeral water change or meaningful water change	• *Region stratification* • *Water change pattern analysis*
	Confusion in area with deep shadows caused by terrain or buildings	• *NLCD legacy data* • *Digital Elevation model (DEM) derivatives*
Permanent snow/ice	*Seasonal and inter-annual variability of snow/ice cover*	• *The frequency of snow over seven epochs* • *The extent of snow at a given epoch* • *The minimum extent of snow over seven epochs*
	Missing areas of snow/ice under deep terrain shadows	• *NLCD legacy data* • *DEM derivatives*
Forest	*Forest succession in an illogical order*	• *Change land cover label according to ecological succession process*
	Confusion with agriculture class due to phenology	• *Use NASS ancillary data and long-term disturbance year product to separate disturbed forest areas from agricultural area*
	Confusion between forest regrowth classes with permanent shrub/grass lands	• *Time series of disturbance data & long-term land cover strata* • *Ancillary data such as NLCD legacy data* • *Shrub continuous fields from Shrub project* • *Existing Vegetation Type (EVT)* • *Monitoring Trends in Burn Severity (MTBS) data*
	Confusion among coniferous, mixed, and deciduous forest classes	• *Both leaf-on and leaf-off images of 2016* • *Forest disturbance data* • *NLCD legacy data*
Cultivatedcropland	*Spectral confusion with non-agricultural herbaceous or shrubland or forest classes*	• *USDA CDL and Cultivated cropland data* • *NLCD legacy data* • *Image object-based information* • *Long-term land cover strata*
	noise within the agricultural fields	• *Image segmentation and objects*
	Change of cropland and CRP land	• *Persistence of spectral and land cover changes over time*
Shrubland	*Confusion between disturbed forest class and shrub*	• *Persistence of shrubland* • *Shrubland maps based on continuous field modeling products*
	Challenges in monitoring shrubland changes by spectral data alone	• *Use fire ecological succession models* • *Recovery rate stratified by land cover type (sage vs. non-sage), region (mainly based on perspiration), fire severity*
Herbaceous	*Confusion between disturbed forest and herbaceous*	• *Persistence of herbaceous* • *Shrubland and herbaceous maps based on continuous field modeling products*
	Confusion with hay/pasture	• *Regional mask and disturbance data layer* • *NLCD legacy data*
Wetland	*Change between emergent wetland and woody wetland*	• *Reference to classification results and disturbance date to determine if it is forest/shrub or grass*
	Woody wetland confusion with upland forest class	• *Use NWI, Wetland Potential Index, NLCD legacy data*
	Inland wetland and coastal wetland change pattern difference	• *Region stratification*
Developed	*Confusion with non-developed classes*	• *Use imperviousness data to determine developed classes*

The post-processing was conducted for each land cover type in a hierarchical order: (1) water, (2) wetlands, (3) forest and forest transition classes, (4) permanent snow and ice, (5) agricultural lands, and (6) persistent shrubland and herbaceous. Different sets of rules and models were developed for each of the six land cover categories. Table 1 lists major issues and solutions applied for land cover post-processing. Two examples are presented here to illustrate the concept: one for water and one for the forest classes. For water, several models were used to reduce the confusion between water and other land cover classes. For instance, some coniferous forest was misclassified as water because of its spectral darkness and/or impact of deep terrain shadows. This error was corrected if the slope of the site is greater than 2% and the land cover from NLCD legacy data is coniferous forest. We also identified water change patterns such as water patch size, water frequency, and water change consistency over time to separate spurious change from real change. Other misclassified water pixels near the borders of water bodies were relabeled as wetland if the presence of water at the location is not persistent over time. Another strategy for water post-classification is to use ecoregion boundaries as strata to separate real land cover change (from water to herbaceous or hay/pasture) in the Prairie Pothole Region (PPR) of the U.S. from the ephemeral changes in the eastern coastal wetlands caused by tides.

We also developed a series of post-processing models for forest and disturbed forest classes. For example, we integrated object-based land cover information based on Landsat image segments of each epoch year and the polygons of disturbance-year for improvement of land cover labels. We also used a disturbance dataset, a rangeland cover map, and other ancillary data to distinguish climax shrub and herbaceous classes from the spectrally similar yet different land cover classes related to forest disturbance and recovery. In addition, we mitigated confusions between forest and cropland using CDL data, confusions between upland and lowland forest using NWI and NLCD legacy data. We employed a leaf-off image to separate coniferous and deciduous forest and improved accuracy of mixed forest class.

Our post-processing methods also included a check on temporal consistency of time series land cover maps, and applied models to correct the illogical changes in the land cover label sequence. For example, for a pixel of interest, if the multi-temporal land cover map shows only one year out of seven mapped as forest, and the year is neither the beginning (2001) nor in the end year (2016), and all other years are mapped as cropland, then the forest label for that one year will be changed to cropland. Another example is that if there are no spectral changes identified over the entire time period from 2001 to 2016, and the land cover of all NLCD legacy datasets is forest, then the final label of this pixel will be forest for all seven epochs. Table 1 lists the examples of main issues and procedures developed for post-processing for major land cover classes.

5.2.4. The final integration process

After all post-processing was completed, the last step was to integrate all intermediate datasets from the previous steps into a final product. The main goal of the final integration was to create coherent spatial-temporal objects across all years from 2001 to 2016 so that all spatial-temporal objects followed a temporal change (or no change) trajectory in the same way, and all pixels within an object belonged to the same land cover class for an epoch year. The final integration checked and corrected spatial inconsistencies of all land cover and change patches, and also resolved class label issues pertinent only to local environments (e.g., in coastal area only). The premise for spatial consistency checking is that many land cover types (e.g., cropland, forest and forest disturbance, wetland, herbaceous, shrubland, and hay/pasture) often appear as a spatial object rather than as an individual isolated pixel. For those land cover types we applied models to check for differences between pixel-based and the object-based land cover labels. If the difference occurs, a rule-based model was applied to reconcile the differences using the majority label of the object-based land cover map. In contrast, for water and developed classes, we kept pixel-based land cover labels in order to retain all small land cover and the related changes. As a result of this process, spatially and temporally consistent land cover maps were obtained.

6. Methods and models for continuous fields products

6.1. Impervious surface

The methodology for generating the NLCD 2016 urban impervious surface product was built upon previous experience and evolved from mapping for a single date (Yang et al., 2003a) and bi-temporal change (Yang et al., 2003b; Xian and Homer, 2010) to multiple dates (Homer et al., 2015) across 2001–2016. Additional models and procedures were developed to employ the NLCD 2011 impervious surface product as the baseline for training data and Landsat imagery pairs acquired in 2011 and 2016 as the primary predictive variables for identifying impervious changed areas.

The procedure includes four steps: (1) training data development, (2) impervious surface modeling, (3) comparison of initial model outputs, and (4) final editing and product generation. In step one, either a NASA Defense Meteorological Satellites Program (DMSP) Nighttime Lights data layer (for 2011) or NOAA's Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night band (DNB) Nighttime Lights data (for 2016) was superimposed on the published NLCD 2011 impervious surface product to exclude low density impervious areas outside urban and suburban centers. This ensures only urban core areas are used to produce statistically stable training data. Two training datasets, one having a relatively large urban extent and one having a relatively small extent, were assembled through imposing two separate thresholds of nighttime lights imagery on the 2011 impervious data layer. Next, each of the two training datasets were used separately to build regression tree models for predicting percent impervious surface using 2011 Landsat imagery as predictive variables (Xian and Homer, 2010). These two sets of regression tree models were created and used to produce two 2011 initial impervious surface maps. Similarly, the same two training datasets were used with 2016 Landsat images to create two sets of regression tree models and two 2016 initial impervious surface maps. Using the same training data for both sets of regression trees ensures stable and comparable outcomes from multiple predictive scenarios. In step three the two pairs of initial impervious maps were compared to remove false estimates due to high reflectance from nonurban areas and to retain 2011 impervious values unchanged from 2011 to 2016. For changed areas, an updated 2016 impervious surface map was generated. The last step of the procedure was a clean-up to correct mapping errors by both hand editing and automated processes. In some areas hand editing was used to remove false impervious estimates in areas such as mines and barren land, and to add missed developed areas where very low imperviousness exists such as in city parks and golf courses. To map earlier years of NLCD, two impervious product pairs were compared with the preceding published year to retain areas of changed impervious values. Additionally, improved road data layers for 2011 through 2016 were created and used in impervious feature identification to improve the accuracy of the impervious product.

6.2. Tree canopy cover

The NLCD Tree Canopy Cover (TCC) products provide estimates of percent tree canopy cover (0–100%) for each 30-m pixel. The U.S. Forest Service became the lead agency for the TCC product for the 2011 NLCD. The product included three geospatial data layers: (1) an analytical TCC product based on raw model output, (2) standard errors for each pixel, and (3) a cartographic product for which non-tree areas with non-zero canopy cover were masked out of the analytical product. The basic approach is described by Coulston et al. (2012). The 2016 TCC product will include a re-map of 2011 TCC and the 2016 TCC product using consistent methods for both dates. The analytical vs. cartographic distinction will not be used for the 2016 product suite; rather, a "best" pixel product and an accompanying metadata layer that indicates the origin of the pixel (modeled tree canopy cover vs. masked based on ancillary information) will be provided. A change layer (2016 vs. 2011) will also be produced.

6.2.1. Training datasets

To support the 2011 TCC mapping effort, percent canopy cover was estimated using a dot grid overlaid on images from the National Agriculture Imagery Program (NAIP) or other sources of high-resolution imagery where NAIP imagery was unavailable. The land cover at each dot was interpreted as tree or not tree, and then a percent canopy cover value was computed for the overall grid. Canopy cover was determined in this manner at locations representing approximately 20% of the plot network used by the Forest Inventory and Analysis (FIA) program. Over 63,000 locations were interpreted for the 2011 TCC mapping effort. For the 2016 TCC product, resources were not available to re-interpret tree canopy cover for all of the locations used in the 2011 product. Rather, 3% of the original locations were re-interpreted using newer NAIP imagery based on the occurrence of wildfires or large NDVI changes detected in Landsat-derived time series.

6.2.2. Canopy cover models

Percent canopy observations, derived from the aforementioned photo-interpreted dot grids, were used to train models using Landsat spectral data and derived indices as predictor variables. To produce the 2011 tree canopy data layer, Landsat 5 TM images from 2009 to 2011 were used. Landsat 8 OLI images from 2014 to 2016 were used for the 2016 product. Image data were summarized temporally for each pixel in two ways: (1) growing season median values were calculated (Ruefenacht, 2016), and (2) harmonic regression coefficients were produced from the full three-year time series. Ancillary variables, such as elevation derivatives, were also used as predictor variables. Random forests regression was used as the modeling framework and has been found to produce better results than other methods tested (Freeman et al., 2016).

6.2.3. Final product assembly

Modeled tree canopy cover for each path/row was qualitatively examined for anomalies, and when present, path/rows were modeled iteratively using slight changes to input parameters or input data until a satisfactory output was produced. A series of three masks were then applied to each path/row: (1) a percent canopy "threshold" mask (per-pixel standard errors were used to determine if any pixel with non-zero canopy could actually have zero canopy), (2) an agricultural mask (based on the consistent presence of cultivated crops on a pixel as indicated by the Cropland Data Layer) and (3) a water mask. The masked products were mosaicked by MRLC mapping zone, and then subjected to another round of qualitative reviews to check for anomalies. A companion layer that indicates change in tree canopy cover between 2011 and 2016 will be generated for the NLCD product suite for the first time as part of the 2016 release. Standard errors output from tree canopy models will be used to assess confidence in tree canopy cover change between the two years. Change areas that meet a threshold level of confidence will be portrayed in the final change product.

6.3. Shrub, herbaceous, and bare ground

We developed an approach that generated the continuous field products of shrub, herbaceous and bare ground for the western U.S. as a new database component of NLCD 2016. The approach requires use of multi-resolution remote sensing data and field measurements to quantify land cover and vegetation components in the rangeland ecosystem (Homer et al., 2012; Xian et al., 2013). Specifically, nine components in the Western U.S. are being estimated for every Landsat 30-m pixel, including: continuous fractional cover of herbaceous, annual herbaceous, bare ground, litter, shrub, sagebrush (*Artemisia spp*.), big sagebrush (*Artemisia tridentata spp*.), and sagebrush and shrub height. Each of these components was predicted by the following steps: (1) ground collection of ocular estimates of component cover corresponding with vegetation patches visible on WorldView 2 (WV–2) (2-m resolution) imagery, (2) using these data as training for regression tree models to predict each component across each WV–2 footprint, (3) downscaling the 2 m products to 30-m resolution, and (4) using the downscaled data as training data for a regression tree model that predicts each fractional component using Landsat images (Xian et al., 2013).

Most mapped components had significant correlations with independent validation data collected from the field, and the root mean square errors of predicted factional component of all vegetation and cover type is less than 10% in study areas of the northwest U.S. and are less than 20% for study areas in the southern U.S. (Xian et al., 2013). While some uncertainties remain with height estimates, the method provides an unbiased and cost effective approach to quantify shrubland components at a regional scale, and offers valuable information to improve thematic land cover classification in arid and semiarid areas. For NLCD 2016 development, the estimated components were cross-walked to NLCD land cover classes of shrubland, herbaceous, and barren based on a set of decision rules determined by indicators of vegetation cover and ecological regions (Rigge et al., 2017). The three cross-walked land cover classes are incorporated into the NLCD 2016 land cover modeling process.

7. Automation and documentation of the implementation procedures

Given the objectives, magnitude, and large spatial and temporal dimensions that NLCD 2016 targets, it was not feasible to implement the project without a high degree of processing automation. A large integrated central working database and server system was procured to allow collaborative and centralized processing and scripting. For NLCD 2016 production, a series of scripts were written to automate the process. Scripting improved efficiency and reduced the possibility of human errors during production. Scripts were developed for image preprocessing, land cover modeling, and impervious modeling processes using Practical Extraction and Report Language (PERL) and Python. With these two languages custom code was generated to gather input data and call software packages such as ERDAS IMAGINE®, ArcGIS®, and Trimble eCognition® to process the data, execute models, and generate all output from the modeling processes. All processing procedures and models for NLCD 2016 were documented so that they are transparent and can be tracked and reproduced. The documents include: (1) NLCD 2016 image selection and pre-processing, and NLCD 2016 land cover and change modeling and post-classification, (2) a line-by-line interpretation of each individual model, and (3) a dictionary that lists all files from model output, including file names and a description of each file.

8. Results from pilot studies and preliminary assessment

8.1. Results

All procedures described previously were implemented and tested in twenty WGS-2 path/row across the Conterminous United States. The test areas cover diverse landscapes with many land cover types and changes over the period of 2001–2016, including forest areas in the Southeastern U.S. and the Pacific Northwest, the wetlands in southeast coastal areas and the PPR region in north central U.S., the agricultural lands and grasslands in the central U.S., the shrubland and grassland in western U.S., and the urban development in and around cities. The results from the pilot studies are presented by focusing on several major land cover categories and changes, namely, forest and changes caused by disturbances, wetland distribution and changes, agricultural land and conversion, shrubland and changes caused by fire disturbance, and urban development and changes.

8.1.1. Forest and changes

Fig. 1 shows a subset area of all land cover and change maps in west-central Georgia in the southeastern U.S. (path 22, row 39) for 2001, 2003, 2006, 2008, 2011, 2013, and 2016. The area is dominated by forests with some hay/pasture lands. Timber production is very important throughout the region where vast expanses of industrial-scale pine plantations occur. Intensive land use of commercial tree farming resulted in a cyclic change of land cover associated with a sequence of tree planting, growth, maturation and cutting over time. Forest harvesting and regeneration are the major driving forces of land cover change. The cleared land typically remains in a mechanically disturbed state for a brief period and is then replanted into trees; however, in some cases, the land is left for natural regeneration.

The land cover change maps in Fig. 1 clearly show the trajectory of this change sequence from 2001 to 2016. The forest clear-cut is shown in purple, the regeneration of shrub/scrub after cutting was mapped in light blue, and the young trees or transitional forest at the later stage of regrowth was mapped in green. The success of mapping this land cover and changes is attributed to the spectral and temporal information derived from Landsat images and other data.

8.1.2. Wetland and changes

The area of wetlands in the U.S. wetland increased by an average of 32,000 acres (12,900 ha.) annually from 1998 to 2004 (Dahl 2006). In contrast, the wetland area declined by an estimated 62,300 acres (25,200 ha) between 2004 and 2009 (Dahl 2011). For this study, two examples of wetland changes are presented. One is in the PPR of north central U.S., which is characterized by many small lakes and potholes. Fig. 2 shows a subset in this area where changes of wetland and water extent are highly sensitive to seasonal and inter-annual variability of weather conditions. For instance, in a dry year (2008), many water bodies and potholes dried out and the land cover changed from water to either wetlands or hay/pasture; in contrast, in a very wet year (2011), many wetland areas were covered by water. The dynamics of the wetlands is clearly shown in the time series land cover and change maps. Another example of mapped wetland changes is in the coastal zone of the southeastern US (not shown by a figure). The area is characterized with mixed forest, woody wetland, and agricultural lands. The main changes are from the forest (woody) wetland to mixed woody and herbaceous wetlands as the result of tree cutting. These

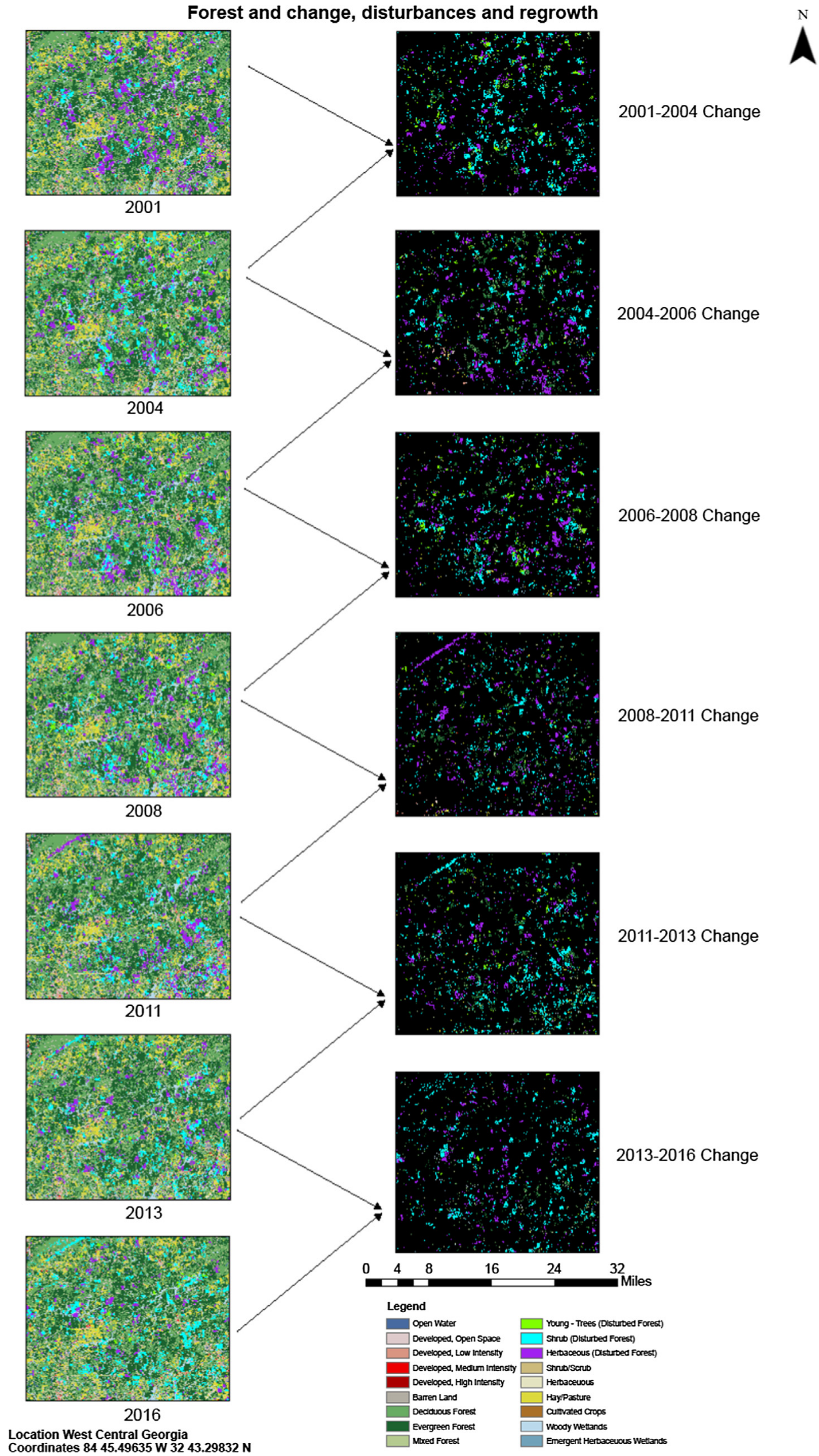

Fig. 1. Forest and changes trajectory in west-central Georgia of the southeastern U.S. from 2001 to 2016. Note the forest clear-cut area (in purple), the regeneration of shrub/scrub after cutting (in light blue), and the young trees or transitional forest at the later stage of regrowth (in green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

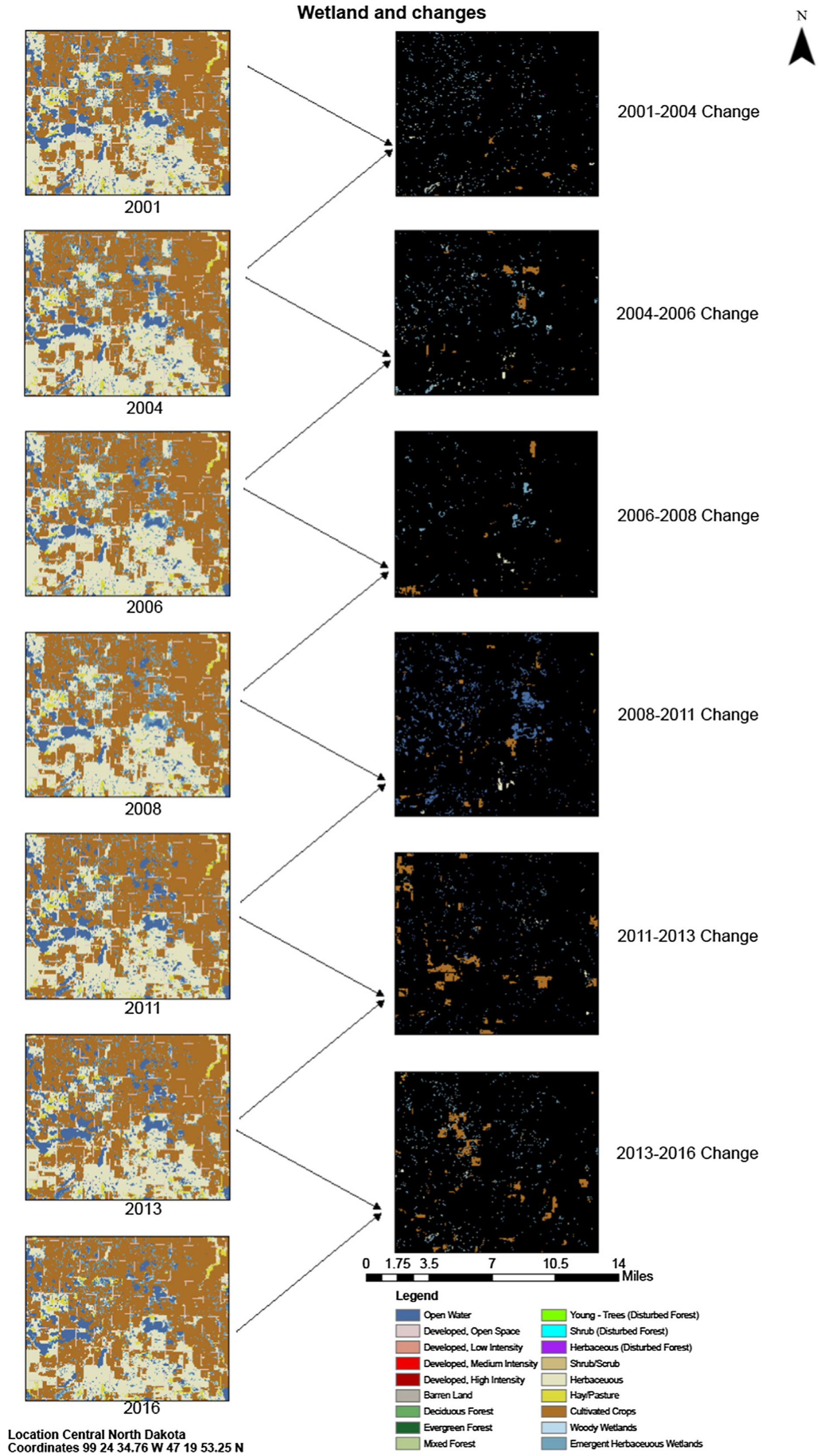

Fig. 2. Wetland and change trajectory from 2001 to 2016 in central North Dakoda of the U.S. Note the changes of wetland and water extent over time caused by seasonal and inter-annual variability of weather conditions. Many water bodies and potholes dried out in 2008 and the land cover changed from water to either wetlands or hay/pasture. In contrast, many wetland areas were covered by water in a very wet year in 2011.

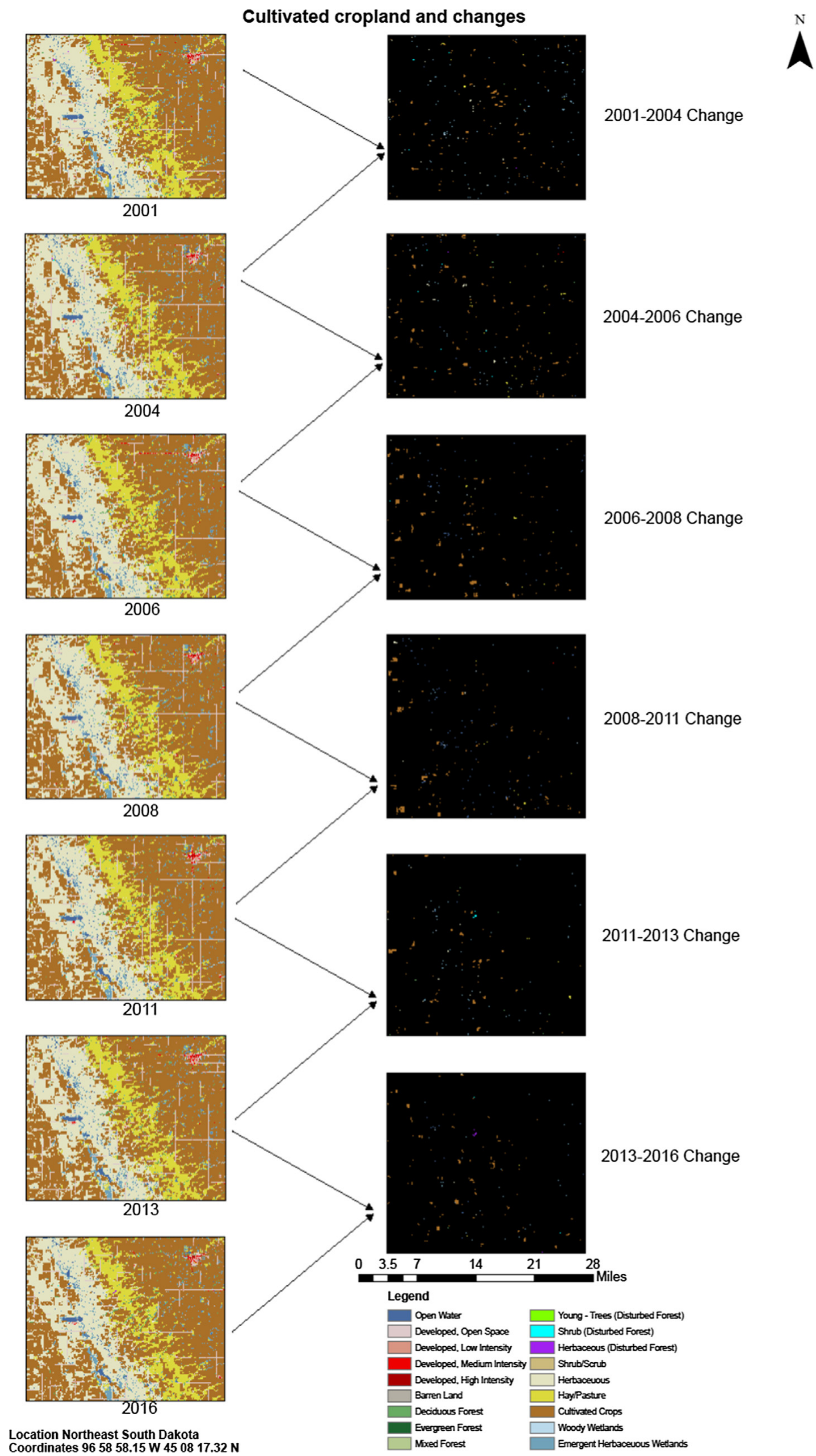

Fig. 3. Cultivated cropland and changes from 2001 to 2016 in northeast South Dakoda of the central U.S. The area is predominantly by cultivated cropland with a stable land use practice.

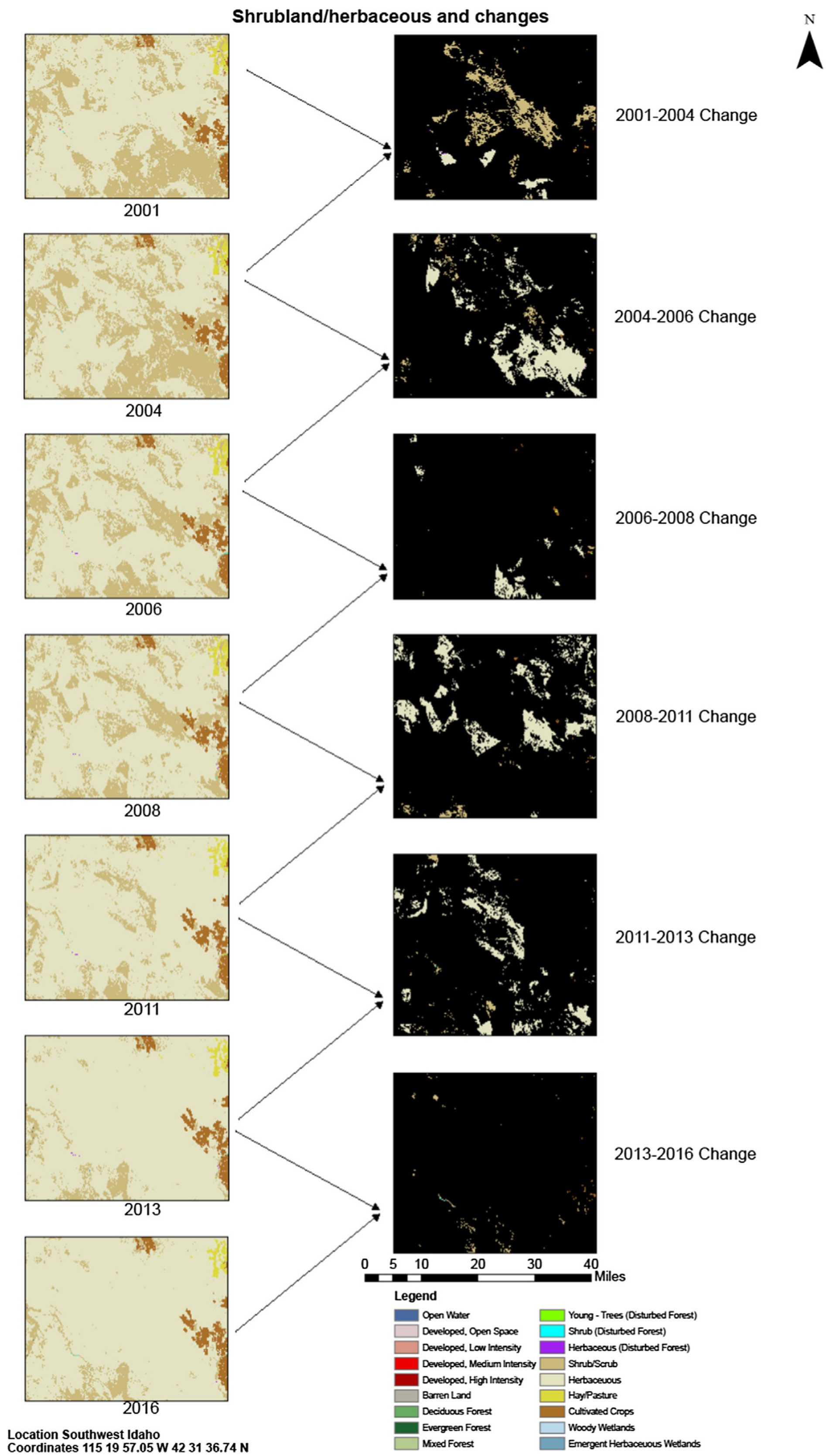

Fig. 4. Shrubland/herbaceous/bare ground and changes from 2001 to 2016 in southwest Idaho of the U.S. Some shrubland in the area was stable and changed little over time, but other areas were burned and changed to herbaceous or bare ground after the fire.

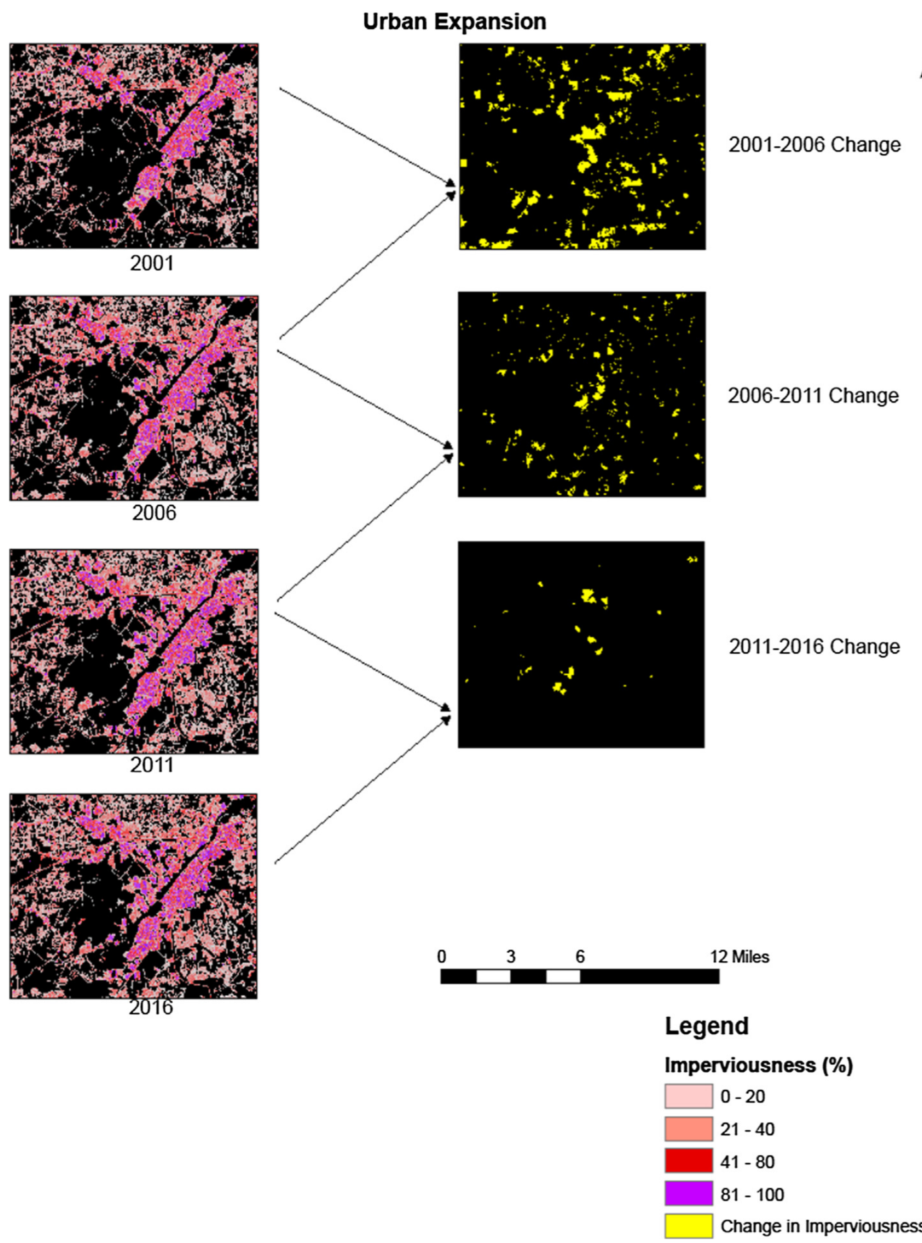

Fig. 5. Developed areas and changes from 2001 to 2016 near the city of Atlanta, Georgia. The changes of imperviousness were observed both within existing developed areas (intensity change) and in non-developed areas (conversion to developed), especially from 2001 to 2006 period.

areas often recovered to woody wetland quickly due to regrowth in this hot and humid environment.

8.1.3. Cultivated cropland and changes

Fig. 3 shows the land cover in a predominantly cultivated cropland area in South Dakota with a stable land use practice and land cover condition. The three dominant land cover types are cultivated cropland, hay/pasture, and herbaceous. The time series land cover maps show that very limited land cover changes occurred in the area over the 2001–2016 period. The limited changes shown in the land cover change maps are the increase of cultivated cropland due to change of land use and change between wetland and surface water extent caused by variation in weather and climate condition.

8.1.4. Shrubland/herbaceous/bare ground and changes

It is a great challenge to monitor land cover condition and changes of shrubland, herbaceous and bare ground in the semi-arid western US using remote sensing data (Homer et al., 2012). The sometimes subtle differences in spectral signature and the transient changes of vegetation in response to weather conditions make it very difficult to accurately detect and map these land cover type and changes. Our method relied on a disturbance dataset, the time series Landsat images, and the understanding of ecological conditions of the region, and was applied to regions where new NLCD cross-walked shrub, herbaceous and bare ground products are available. Fig. 4 shows the distribution and changes of one shrubland and herbaceous area in southwest Idaho. Some shrubland in the area was stable and changed little over time, but some were burned and then changed to herbaceous or bare ground after the fire (shown in change maps). Some shrubland areas burned early (before 2001) and regrew back to shrubland in 2004, then later changed to herbaceous after 2011. Over time, a loss of shrubland in the area is observed, which indicates the potential impact of disturbances on this ecosystem.

8.1.5. Developed areas and changes

Land cover condition and changes in and around urban and

Table 2
Agreement between map and reference labels for NLCD 2001 for path 18, row 35. Sample size is reported in the column and row labeled n. Producer's agreement (Prod) and User's agreement (User) are rounded to the nearest whole number. Agreement was defined as a match between the land cover map and the reference labels of training data. OA is overall agreement defined as a match between the map and reference labels. OA = 0.90.

LC class	11	22	23	24	31	41	42	43	44	45	46	52	71	81	82	90	95	User	n
11	2515	1	0	0	1	2	0	0	0	0	0	0	0	0	0	3	0	0.99	2522
22	0	1027	90	2	2	1	0	0	0	1	1	0	25	302	3	4	0	0.70	1458
23	0	200	396	32	6	0	0	0	0	0	0	0	4	92	12	0	0	0.53	742
24	7	23	113	190	3	0	0	0	0	0	0	0	3	20	4	1	0	0.52	364
31	2	41	20	2	60	2	1	0	0	0	13	0	14	111	10	0	0	0.22	276
41	2	47	1	0	3	136,491	7426	5142	0	0	1	0	0	358	10	271	0	0.91	149,752
42	0	4	0	0	0	2217	4421	447	0	1	0	0	0	2	1	6	0	0.62	7099
43	0	0	0	0	0	254	135	106	0	0	0	0	0	0	0	0	0	0.21	495
44	0	0	0	0	0	1	0	0	138	1	0	0	0	2	0	3	0	0.95	145
45	0	2	0	0	0	2	1	0	0	138	46	0	0	48	2	8	0	0.56	247
46	0	13	1	0	9	4	0	0	0	5	522	0	0	189	1	5	0	0.70	749
52	5	12	0	0	7	0	0	0	0	0	0	0	14	56	9	1	0	0.00	104
71	0	24	0	0	3	0	0	0	0	0	0	0	1027	211	0	0	0	0.81	1265
81	0	619	28	0	39	48	2	0	0	1	38	0	805	42,479	273	67	0	0.96	44,399
82	0	13	5	0	3	6	0	0	0	0	0	0	6	352	561	4	0	0.59	950
90	2	3	0	0	0	412	31	21	0	0	0	0	3	63	5	250	0	0.32	790
95	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0.00	2
Prod	0.99	0.51	0.61	0.84	0.44	0.98	0.37	0.02	1.00	0.94	0.84	n/a	0.54	0.96	0.63	0.40	n/a		
n	2534	2029	654	226	136	139,441	12,017	5716	138	147	621	0	1901	44,285	891	623	0		211,359

Land cover classes: Open Water (11), Perennial Ice/Snow (12), Developed-Open Space (21), Developed-Low Intensity (22), Developed-Medium Intensity (23), Developed-High Intensity (24), Barren Land (31), Deciduous Forest (41), Evergreen Forest (42), Mixed Forest (43), Transitional Forest/Young Tree (44), Transitional Forest/Shrub (45), Transitional Forest/Herbaceous (46), Shrub (52), Grassland/Herbaceous (71), Pasture/Hay (81), Cultivated Cropland (82), Woody Wetland (90), Herbaceous Wetland (95).

Table 3
Agreement between map and reference labels for NLCD 2006 for path 18, row 35. The notations in the table are the same as in Table 2. OA = 0.89.

LC class	11	22	23	24	31	41	42	43	44	45	46	52	71	81	82	90	95	User	n
11	2228	2	1	2	0	1	1	0	0	0	0	0	0	0	0	1	0	0.99	2236
22	2	890	90	4	4	0	0	0	0	0	0	0	32	365	5	7	0	0.64	1399
23	1	201	390	34	4	0	0	0	0	0	1	0	7	108	12	3	0	0.51	761
24	8	12	146	199	8	0	0	0	0	0	0	0	4	19	4		0	0.50	400
31	3	55	35	7	71	6	1	0	0	0	18	0	12	95	9	2	0	0.23	314
41	5	16	0	0	2	140,829	8042	5335	9	0	6	0	0	63	1	261	0	0.91	154,569
42	2	2	0	0	0	2052	4178	434	8	0	0	0	0	1	0	5	0	0.63	6682
43	0	0	0	0	0	268	125	99	0	0	0	0	0	0	0	1	0	0.20	493
44	1	0	0	0	1	30	25	2	323	1	13	0	0	20	1	4	0	0.77	421
45	0	1	0	0	4	0	0	0	0	57	64	0	0	84	5	10	0	0.25	225
46	0	8	3	0	3	0	0	0	0	3	812	0	0	174	29	11	0	0.78	1043
52	1	6	0	0	6	0	0	0	0	0	0	0	10	35	0	1	0	0.00	59
71	0	37	3	0	2	0	0	0	0	0	0	0	1012	345	3	0	0	0.72	1402
81	0	657	45	2	23	3	0	1	0	2	48	0	787	40,828	438	83	0	0.95	42,917
82	0	26	10	1	1	0	0	0	0	0	5	0	23	340	723	3	0	0.64	1132
90	2	9	0	0	0	431	26	17	1	1	14	0	3	86	1	236	0	0.29	827
95	0	2	1	0	0	2	0	0	0	0	0	0	1	2		2	0	0.00	10
Prod	0.99	0.46	0.54	0.80	0.55	0.98	0.34	0.02	0.95	0.89	0.83	n/a	0.54	0.96	0.59	0.37	n/a		
n	2253	1924	724	249	129	143,622	12,398	5888	341	64	981	0	1891	42,565	1231	630	0		214,890

suburban areas due to development activities were mapped by multi-temporal imperviousness modeling. It is one of the best land cover types mapped spatially and temporally in terms of consistency and accuracy. Fig. 5 shows one urban area and its changes of imperviousness near the city of Atlanta, Georgia. The notable changes of imperviousness in this area are mapped correctly in both timing and location. The changes took place both within existing developed areas (intensity change) and in non-developed areas (conversion to developed), and the urban expansion over time can be clearly observed, especially from 2001 to 2006 period.

8.2. Assessment of results from the pilot studies

The NLCD 2016 prototype method was heavily scrutinized in subjective ways as the process was developed. An objective assessment of the method was also needed and conducted to evaluate the results. It should be noted that the goal of this assessment is not a formal accuracy assessment of NLCD 2016 as those reported for the past NLCD products (e.g., Stehman et al., 2003; Wickham et al., 2013), as such an assessment can only be made after the final NLCD 2016 products are generated. Instead, the purpose of this assessment was to evaluate the consistency and robustness of the developed models for generating initial land cover maps in different landscapes and across temporal domains. Here we conducted the assessment for all pilot areas and for all epoch years. The reference data used to evaluate the classification were drawn from a training data pool, and was independent from the data used for land cover classification. The assessment parameters are overall agreement, producer's agreement and user's agreement between mapped and reference (training) data. While the assessment results reported in this paper are not the accuracy of the final NLCD 2016 land cover products, they are likely to be conservative (the same as or lower than the accuracy of the final product) because the final products should have a higher overall accuracy after all additional post classification procedures are applied. We note that the assessment reported

Table 4
Agreement between map and reference labels for NLCD 2011 for path 18, row 35. The notations in the table are the same as in table 2. OA = 0.88.

LC class	11	22	23	24	31	41	42	43	44	45	46	52	71	81	82	90	95	User	n
11	2419	2	1	1	0	3	0	0	0	0	0	0	1	0	0	1	0	0.99	2428
22	0	980	102	3	2	0	0	0	0	0	0	0	34	351	10	5	0	0.66	1487
23	0	248	473	39	16	3	0	0	0	0	3	0	10	81	15	2	0	0.53	890
24	9	12	176	218	7	0	0	0	0	0	3	0	4	19	9	0	0	0.48	457
31	2	56	50	7	74	3	1	0	0	2	24	0	11	148	26	2	0	0.18	406
41	4	10	2	0	1	141,267	8440	5344	5	1	1	0	0	4	1	274	0	0.91	155,354
42	0	0	0	0	0	2181	4117	400	0	0	1	0	0	0	1	11	0	0.61	6711
43	0	0	0	0	0	253	130	118	0	0	0	0	0	0	0	0	0	0.24	501
44	0	3	0	0	0	2	0	0	376	2	0	0	0	12	3	1	0	0.94	399
45	0	10	1	0	2	1	1	0	2	256	19	0	0	149	10	13	0	0.55	464
46	0	12	4	2	2	0	0	1	0	1	832	0	0	99	23	2	0	0.85	978
52	0	6	1	1	7	0	0	0	0	0	0	0	11	56	4	0	0	0.00	86
71	2	34	5		1	0	0	0	0	0	0	0	700	391	4	0	0	0.62	1137
81	0	594	41	2	37	0	0	0	0	3	32	0	757	27,282	374	68	0	0.93	29,190
82	0	25	20	1	3	0	0	0	0	0	5	0	27	252	879	3	0	0.72	1215
90	10	10	0	0	0	358	36	13	1	7	3	0		63	6	232	0	0.31	739
95	2	22	3	0	1	19	7	0	0	2	34	0	19	139	14	40	0	0.00	302
Prod	0.99	0.48	0.54	0.80	0.48	0.98	0.32	0.02	0.98	0.93	0.87	n/a	0.44	0.94	0.64	0.35	n/a		
n	2448	2024	879	274	153	144,090	12,732	5876	384	274	957	0	1574	29,046	1379	654	0		202,744

here is for land cover only. As for the percent imperviousness, percent tree canopy cover, and percent shrub/herbaceous/bare ground products, a separate evaluation will be made after the final products of these database components become available.

8.2.1. Overall, producer's and user's agreement of land cover maps for two sample path/row

For path 18 row 35, the land cover overall agreement of individual epoch year was high and consistent, 90% for 2001, 89% for 2006, 88% for 2011, and 88% for 2016 (Tables 2–5). High user's agreement (≥70%) were realized for water (11), deciduous forest (41), disturbed forest (44, 46), and hay/pasture (81) for all four epoch years, and for cropland (82) in 2011 and 2016 when agreement was defined as a match between the map and the reference label. In contrast, low user's agreement (< 35%) were observed for barren (31), mixed forest (43), and wetland (90, 95) classes. The high producer's agreement (≥70%) were also very consistently shown in several classes, including water (11), high intensity developed (24), deciduous forest (41), forest disturbed classes (44, 45, 46), and hay/pasture (81) for all epoch years. Low producer's agreement (< 35%) were recorded for mixed forest (43) and wetland (90). We noted that the low agreement of the mixed forest class was due to a problem of the training data, which was a very weak class mapped in the legacy NLCD.

For path 41 row 30, the land cover overall accuracies of individual epoch years were also high and consistent: 89% for 2001, 88% for 2006, 88% for 2011, and 88% for 2016. High user's accuracies (≥70%) were realized for water (11), low intensity developed area (22), coniferous forest (42), shrubland (52), herbaceous (71), and cropland (82) for all four epoch years. In contrast, low user's accuracies were observed for barren (31) and deciduous forest (41). High producer's accuracies (≥70%) were also very consistently shown in several classes, including water (11), low and medium intensity developed (22, 23), coniferous forest (42), disturbed forest (44, 45), shrubland (52), herbaceous (71), and cropland (82) for all epoch years. No real low (< 35%) producer's accuracies were recorded for any land cover classes.

8.2.2. Overall agreement of land cover maps for all path/row of pilot study area

Additional assessment was made for all 20 path/row from our pilot studies (Table 6). The results show that the overall agreement of the 20 path/rows and all seven-epoch year ranges from 71 to 97%. Also noted is an interesting geographic pattern showing difference in quality of land cover maps: in the eastern U.S., the overall agreement ranges from 70% to mid-80%, while in central and western U.S., it ranges from 80% to upper 90%. This pattern is likely attributable to the fact that the land

Table 5
Agreement between map and reference labels for NLCD 2016 for path 18, row 35. The notations in the table are the same as in table 2. OA = 0.88.

LC class	11	22	23	24	31	41	42	43	44	45	46	52	71	81	82	90	95	User	n
11	2261	1	2	1	0	2	0	0	0	0	0	0	0	0	0	2	0	0.99	2269
22	0	1131	97	4	0	17	4	0	0	0	0	0	20	307	7	1	0	0.71	1588
23	2	250	449	34	1	9	0	0	0	0	0	0	6	113	25	1	0	0.50	890
24	4	16	163	250	0	0	0	0	0	0	2	0	5	47	9	1	0	0.50	497
31	2	59	45	12	12	4		1	0	8	9	0	23	146	22	1	0	0.03	344
41	1	30	2	0	0	149,936	7194	4968	0	226	0	0	0	55	12	255	0	0.92	162,679
42	0	4	0	0	0	2732	5944	709	0	46	0	0	1	4	0	10	0	0.63	9450
43	0	0	0	0	0	449	255	319	0	0	0	0	0	0	0	0	0	0.31	1023
44	0	0	0	0	0	2	1		176	18	0	0	0	2	2	0	0	0.88	201
45	0	7	0	0	0	139	84	8	0	971	4	0	0	42	19	7	0	0.76	1281
46	0	8	2	1	0	0	0	0	0	9	778	0	0	38	18	8	0	0.90	862
52	1	3	1	0	0	0	0	0	0	0	0	0	25	111	8	0	0	0.00	149
71	0	10	1	0	0	0	0	0	0	0	0	0	266	259	3	0	0	0.49	539
81	0	587	43	1	6	8	4	1	0	29	13	0	1409	35,987	852	70	0	0.92	39,010
82	0	14	8	1	0	0	0	0	0	2	0	0	18	325	1193	6	0	0.76	1567
90	3	13	1	0	0	376	39	13	0	20	0	0	2	45	6	227	0	0.30	745
95	0	10	1	0	0	35	9	3	0	22	8	0	13	125	21	44	0	0.00	291
Prod	0.99	0.53	0.55	0.82	0.63	0.98	0.44	0.05	1.00	0.72	0.96	n/a	0.15	0.96	0.54	0.36	n/a		
n	2274	2143	815	304	19	153,709	13,534	6022	176	1351	814	0	1788	37,606	2197	633	0		223,385

Table 6
Overall Agreement (%) between map and reference labels for 20 WRS-2 path /row within the conterminous U.S. Agreement was defined as a match between mapped land cover class and a reference label from training data.

		Year						
		2001	2003	2006	2008	2011	2013	2016
WGS-2 Path/row	14/29	72.5	73.6	72.6	73.1	71.0	71.6	74.3
	16/37	80.8	80.5	80.6	80.3	80.0	77.8	77.7
	18/37	75.6	73.0	76.2	75.8	75.3	71.6	72.9
	19/37	78.9	78.4	78.1	77.2	76.5	74.7	76.2
	20/37	78.2	77.9	74.6	74.1	76.4	72.6	74.6
	25/28	74.4	75.2	74.1	75.8	74.5	74.9	75.5
	24/36	83.2	83.1	82.8	82.9	82.7	81.9	83.2
	27/38	86.7	87.8	85.7	87.5	82.0	83.2	77.1
	30/27	92.3	92.5	92.3	92.9	93.0	94.0	95.0
	31/27	89.0	89.6	88.9	90.8	90.3	92.5	91.7
	32/27	89.0	88.4	87.5	85.4	89.4	91.3	89.1
	32/37	93.2	92.9	92.7	92.6	92.2	90.6	92.0
	35/33	91.2	90.7	90.7	91.1	90.5	90.6	91.0
	38/35	97.2	97.0	96.8	96.8	96.8	96.3	96.6
	41/30	89.5	89.2	88.4	88.0	88.3	87.4	87.8
	42/34	94.5	94.3	94.3	94.3	94.2	93.9	94.0
	44/28	94.6	94.2	94.4	94.5	94.6	94.0	94.0
	43/34	92.6	92.6	92.0	92.0	91.5	91.3	91.6
	45/28	93.2	92.5	92.8	93.6	93.2	93.1	93.5
	46/28	88.1	88.5	88.3	88.2	86.5	88.4	86.9

cover and land use in the eastern U.S. is more complex with smaller patch size and more fragmented land than those in the central and western U.S.

9. Discussion

The NLCD 2016 approach reported here provides a solid foundation to meet the project objectives, although some method dependencies and limitations should be noted. The developed methods rely on relatively perfect imagery to achieve a high accuracy. If the image quality and acquisition time are not optimized, they can have a negative effect on land cover model outcome. Also, some modeling process rely more on ancillary data and/or expert knowledge to determine land cover condition and changes, which have uncertainties regarding spatial, temporal and thematic accuracies. For example, it is often difficult to determine how long exactly it will take for trees to completely recover after disturbances occur at a given location and time; likewise, the decision rules used for separating hay/pasture from cultivated croplands based on general land use practice may not always hold.

The NLCD 2016 product is a comprehensive and flexible database with a medium spatial resolution designed and suitable for wide range of applications. We note that this database is for conterminous U.S. only. For users engaged in land cover and change studies at a global scale, other datasets are needed (e.g. those cited in Section 1.1 of this paper). The NLCD land cover product has always followed one predefined classification scheme. For users who use a different scheme or require more detailed classes, they need to either cross-walk the NLCD classes or to use other available datasets, such as the USDA NASS Cropland Data Layer (for detailed cropland types), the LANDFIRE Existing Vegetation Type (for detailed vegetation types). Likewise, for applications that require products with higher spatial resolution, users should explore other regional data sets, such as the NOAA Costal Change Analysis Program land cover and change product with 1–5-m resolution.

10. Conclusions

This paper reports on a new strategy for developing a multi-temporal land cover and change database for the conterminous U.S., called NLCD 2016. This innovative approach provides a strategy to integrate multi-source data and expert knowledge for land cover characterization and change detection, uses a hierarchical approach for land cover and change analyses, and implements a hybrid approach to integrate pixel-based and object-based land cover modeling. The performance of the prototype NLCD 2016 database created using these strategies was tested in 20 WGS-2 path/rows within the conterminous U.S. An overall agreement from 71% to 97% between the land cover product and the reference data was achieved for all tested path/rows and all seven-epoch year of 2001, 2003, 2006, 2008, 2011, 2013 and 2016. These results indicate that the new method provides a robust, comprehensive, and highly automated procedure for NLCD 2016 operational mapping. This new approach will generate NLCD 2016 land cover and change products at 2–3 year intervals from 2001 to 2016, with additional database components and likely a higher accuracy than any previously-released NLCD. A full production of NLCD 2016 for the conterminous U.S. is now in progress, and is scheduled to be completed by the end of 2018.

Looking beyond NLCD 2016, new evolving land change requirements necessitate even more frequent cycles of land cover updates to understand land cover change on an annual basis. The developed NLCD 2016 approach provides a solid foundation to meet this future requirement. Further research and analysis will be required to determine the optimal course for future NLCD product direction and production. Towards this end, several existing and new initiatives at the national level (The Landscape Fire and Resource Management Planning Tools Project, The Gap Analysis Program, Landscape Change Monitoring System, and Land Change Monitoring, Assessment, and Projection) offer promising opportunities to help expand the future NLCD product vision through project collaboration. The Any change of future NLCD direction will carefully consider resources available, the credibility of new technical and data solutions, and the ability to maintain NLCD product accuracy and consistency.

Acknowledgements

This work has been performed under the USGS contract G13PC00028 and G15PC00012. The authors wish to thank Dr. Lei Ji and Mr. Thomas Adamson for conducting agency internal review. They also would like to thank all the anonymous reviewers for their constructive comments. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., Weber, J.L., 2008. GLOBCOVER the most detailed portrait of Earth. ESA Bull.-Eur. Space Agency 136, 24–31.
Bartholomé, E., Belward, A.S., 2007. GLC2000: a new approach to global land cover mapping from Earth observation data. Int. J. Remote Sens. 26 (9), 1959–1977.
Breiman, L., Friedman, J.H., Olshend, R.A., Stone, C.J., 1984. Classification and Regression Trees. California, Wadsworth International Group, Belmont, pp. 358.
Brown, J.F., Loveland, T.R., Merchant, J.W., Reed, B.C., Ohlen, D.O., 1993. Using multisource data in global land characterization: concepts, requirements, and methods. Photogrammetric Eng. Remote Sens. 59, 977–987.
Chen, J., Ban, Y., Li, S., 2014. China: Open access to Earth land-cover map. Nature 7523 2014, 514 434-434.
Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, Jon, 2015. Global land cover mapping at 30 m resolution: a POK-based operational approach. ISPRS J. Photogramm. Remote Sens. 103, 7–27.
Costa, H., Carrão, H., Bação, F., Caetano, M., 2014. Combining per-pixel and objectbased classifications for mapping land cover over large areas. Int. J. Remote Sens. 35 (2), 738–753.
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. Photogrammetric Eng. Remote Sens. 78 (7), 715–727.
Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenböck, H., Müller, A., Dech, S., 2013. Urban footprint processor – fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. IEEE Geosci. Remote Sens. Lett. 10 (6), 1617–1621.
Feng, M., Sexton, J.O., Channan, S., Townshend, J.R., 2015. A global, high-resolution (30-m) inland water body dataset for 2000: first results of a topographic-spectral classification algorithm. Int. J. Digital Earth 1–21.

Franklin, S.E., Ahmed, O.S., Wulder, M.A., White, J.C., Hermosilla, T., Coops, N.C., 2015. Large area mapping of annual land cover dynamics using multitemporal change detection and classification of landsat time series data. Can. J. Remote Sens. 41 (4), 293–314.

Freeman, E.A., Moisen, G.G., Coulston, J.W., Wilson, B.T., 2016. Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing and tuning processes and model performance. Can. J. For. Res. 46, 323–339.

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schnieder, A., Cooper, A., Baccini, A., Gao, F., Schaaf, C., 2002. Global land cover from MODIS: Algorithms and early results. Remote Sens. Environ. 83, 135–148.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., Wickham, J., 2011. Completion of the 2006 national land cover database for the conterminous United States. Photogrammetric Eng. Remote Sens. 77 (9), 858–863.

Giri, C., Ochieng, E., Tieszen, L.L., et al., 2011. Status and distribution of mangrove forests of the world using earth observation satellite data. Glob. Ecol. Biogeogr. 20, 154–159.

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of GLC: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 34 (7), 2607–2654.

Hansen, M., DeFries, R., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1km resolution using a decision tree classifier. Int. J. Remote Sens. 21, 1331–1365.

Hansen, M.C., Potapov, P.V., Moore, R., et al., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853.

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

Homer, C., Aldridge, C.L., Meyer, D.K., Schell, S., 2012. Multi-scale remote sensing sagebrush characterization with regression trees over Wyoming, USA: Laying a foundation for monitoring. Int. J. Appl. Earth Obs. Geoinf. 14 (1), 233–244.

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, 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. Photogrammetric Eng. Remote Sens. 81 (5), 345–354.

Huang, C.Q., Goward, S.N., Masek, J.G., et al., 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sens. Environ. 114, 183–198.

Jin, S., Sader, S.A., 2005. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. Remote Sens. Environ. 94, 364–372.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J.A., 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., Cohen, W.B., Schroeder, T.A., 2007. Trajectory-based change detection for automated characterization of forest disturbance dynamics. Remote Sens. Environ. 110, 370–386.

Kim, D.H., Sexton, J.O., Noojipady, P., Huang, C., Anand, A., Channan, S., Feng, M., Townshend, J.R., 2014. Global, landsat-based forest-cover change from 1990 to 2000. Remote Sens. Environ. 155, 178–193.

Latifovic, R., Pouliot, D., 2005. Multitemporal landcover mapping for Canada: Methodology and products. Can. J. Remote Sens. 31, 347–363.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, J., Yang, L., Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR Data. Int. J. Remote Sens. 21 (6/7), 1303–1330.

Loveland, T.R., Belward, A.S., 1997. The IGBP-DIS global 1 km land cover data set, DISCover–first results. Int. J. Remote Sens. 18 (15), 3289–3295.

Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. 28 (5), 823–870.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sens. Environ. 115 (5), 1145–1161.

Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A., Carneiro, F.S.M., Halkia, S., Julea, A.M., Kemper, M., Soille, P., Syrris, V., 2016. Operating procedure for the production of the global human settlement layer from Landsat data of the epochs 1975, 1990, 2000, and 2014.

Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A.J., Freire, S., Halkia, S., Julea, A.M., Kemper, T., Soille, P., Syrris, V., 2016. Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014. Publications Office of the European Union, EUR 27741 EN, 2016.

Quinlan, J.R., 1993. C4.5 Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, California.

Rigge, M., Gass, L., Homer, C., Xian, G., 2017. Methods for converting continuous shrubland ecosystem component values to thematic national land cover database classes. U.S. Geological Survey Open-File Report 2017-1119, 10p., https://doi.org/10.3133/ofr20171119.

Ruefenacht, B., 2016. Comparison of three Landsat TM compositing methods: a case study using modeled tree canopy cover. Photogrammetric Eng. Remote Sens. 82, 199–211.

Sexton, J., Song, X.P., Feng, M., et al., 2013a. Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. Int. J. Digital Earth 6 (5), 427–448.

Sexton, J., Urban, D.L., Donohue, M.J., Song, C., 2013b. Long-term landcover dynamics by multi-temporal classification across the Landsat-5 record. Remote Sens. Environ. 128, 246–258.

Shafer, G., Logan, R., 1987. Implementing Dempster's rule of hierarchical evidence. Artif. Intell. 33, 271–298.

Smith, G., 2013. Hybrid pixel- and object-based approach to habitat condition monitoring. In: Proceedings GI_ Forum – Creating the GI Society 3756. Wichmann-Verlag, Berlin, pp. 552–555.

Stehman, S.V., Wickham, J.D., Smith, J.H., Yang, L., 2003. Thematic accuracy of the 1992 National Land Cover Data (NLCD) for the eastern United States: Statistical methodology and regional results. Remote Sens. Environ. 86, 500–516.

Song, X.P., Sexton, J.O., Huang, C., Channan, S., Townshend, J.R., 2016. Characterizing the magnitude, timing and duration of urban growth from time series of Landsat-based estimates of impervious cover. Remote Sens. Environ.

Srinivasan, A., Richards, J.A., 1990. Knowledge-based techniques for multi-source classification. Int. J. Remote Sens. 11, 505–525.

Sulla-Menashe, D., Friedl, M.A., Krankina, O.N., Baccini, A., Woodcock, C.E., Sibley, A., Sun, G., Kharuk, V., Elsakov, V., 2011. Hierarchical mapping of northern Eurasian land cover using MODIS data. Remote Sens. Environ. 115, 392–403.

Teluguntla, P., Thenkabail, P.S., Xiong, J., Gumma, M.K., Congalton, R.G., Oliphant, A., Poehnelt, J., Yadav, K., Rao, M., Massey, R., 2017. Spectral matching techniques (SMTs) and automated cropland classification algorithms (ACCAs) for mapping croplands of Australia using MODIS 250-m time-series (2000–2015) data. Int. J. Digital Earth.

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., Van Driel, J.N., 2001. Completion of a 1990's National Land Cover Database for the conterminous United States. Photogrammetric Eng. Remote. Sens. 67, 650–652.

Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W.B., 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.

Wickham, J., Stehman, S.V., Gass, L., Dewitz, J., Fry, J.A., Wade, T.G., 2013. Accuracy assessment of NLCD 2006 land cover and impervious surface. Remote Sens. Environ. 130, 294–304.

Wickham, J.D., Homer, C., Vogelmann, J.E., McKerrow, A., Mueller, R., Herold, N.D., Coulston, J., 2014. The multi-resolution land characteristics (MRLC) consortium—20 years of development and integration of USA national land cover data. Remote Sens. 6 (8), 7424–7441.

Wulder, M., Coops, N., Roy, D., White, J., Hermosilla, T., 2018. Land cover 2.0. Int. J. Remote Sens. 39 (12), 4254–4284.

Xian, G., Homer, C., 2010. Updating the 2001 national land cover database impervious surface products to 2006 using landsat imagery change detection methods. Remote Sens. Environ. 114, 1676–1686.

Xian, G., Homer, C., Meyer, D., Granneman, B., 2013. An approach for characterizing the distribution of shrubland ecosystem components as continuous fields as part of NLCD. ISPRS J. Photogramm. Remote Sens. 86, 136–149.

Yang, L., Huang, C., Homer, C., Wylie, B.K., Coan, M.J., 2003a. An approach for mapping large-area impervious surfaces: Synergistic use of Landsat-7 ETM+ and high spatial resolution imagery. Can. J. Remote Sens. 29, 230–240.

Yang, L., Xian, G., Klaver, J.M., Deal, B., 2003b. Urban land cover change detection through sub-pixel imperviousness mapping using remotely sensed data. Photogrammetric Eng. Remote Sens. 69, 1003–1010.

Zhu, Z., Woodcock, C.E., Olofsson, P., 2012. Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sens. Environ. 122, 75–91.