GTAC-10264-RPT1

Forest Service Tree Canopy Cover Mapping: 2016 Product Suite and Methods

Cover: Photo of a forested road.

Ruefenacht, B.; Heyer, J.; Johnson, V.; Goetz, W.; Bender, S.; Schleeweis, K.; Megown, K. 2022. Forest Service Tree Canopy Cover Mapping: 2016 Product Suite and Methods. GTAC-10264-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Geospatial Technology and Applications Center. 47 p.

Abstract

Tree canopy cover (TCC) is a fundamental component of the landscape; it influences physical phenomena such as fire behavior and ecological concerns such as habitat fragmentation. Percent tree canopy cover data and products built by the United States Department of Agriculture, Forest Service (USFS) are a 30-meter raster geospatial dataset that represents the vertically projected tree canopy cover. The USFS contributes these datasets and products to the National Land Cover Database (NLCD), which is updated on a 5-year cycle and released by the interagency Multi-Resolution Land Characteristics (MRLC) consortium. This report reviews the methods and techniques used to produce the 2016 product suite (released in 2019), which includes 2011 epoch 2 TCC and 2016 TCC.

Authors

Bonnie Ruefenacht is a senior specialist employed by RedCastle Resources at the Geospatial Technology & Applications Center (GTAC) in Salt Lake City, UT.

Josh Heyer is a geospatial analyst employed by RedCastle Resources, Inc. at GTAC in Salt Lake City, UT

Vicky Johnson was a geospatial analyst employed by RedCastle Resources, Inc. at GTAC in Salt Lake City, UT.

Wendy Goetz is a contract manager employed by RedCastle Resources at GTAC in Salt Lake City, UT.

Stacie Bender is an assistant program leader for the Resource Mapping, Inventory and Monitoring (RMIM) program at GTAC in Salt Lake City, UT.

Karen Schleeweis is an ecologist with Forest Inventory and Analysis (FIA) Information Management and Technology Team at the Rocky Mountain Research Station Inventory and Monitoring Program in Ogden, UT.

Kevin Megown is the program leader for the Resource Mapping, Inventory and Monitoring (RMIM) program at GTAC in Salt Lake City, UT.

Acknowledgements

The authors would like to acknowledge contributions from Greg Liknes, Joshua Reynolds, Mark Finco, Holly Russon, Jan Johnson, Nicholas Daniels, Steve Dale, and Zhiqiang Yang.

Table of Contents

Introduction

Tree canopy cover (TCC) maps (as geospatial depictions of TCC data) have broad applications and importance in forest and natural resource management. TCC maps are used for evaluating land cover change, monitoring forest health, assessing natural hazards such as fires, and estimating forest carbon pools. For instance, TCC was used to adjust snowpack model parameters (Rittger et al., 2020), improve forest inventory estimates (McRoberts et al., 2014), model surface temperatures in urban areas (Heris et al., 2021), refine insect and disease affected forest areas (Berryman & McMahan, 2019), and map conifer cover for the purpose of monitoring eastern redcedar encroachment (Filippelli et al., 2020). Also, the Intergovernmental Panel on Climate Change uses TCC estimates for reporting on greenhouse gas inventories (Intergovernmental Panel on Climate Change, n.d.). Considering the increased wildfire activity in the western United States (Dennison et al., 2014), TCC will continue to be a critical input for assessing fire fuels and developing fire behavior models (Nelson et al., 2013). Knowing the spatial extent of TCC is an important first step in managing forest resources and assessing forest health and watershed conditions.

As a member of the interagency Multi-Resolution Land Characteristics (MRLC) consortium, the United States Department of Agriculture, Forest Service (USFS) contributes TCC products to the National Land Cover Database (NLCD). The MRLC includes federal agencies working collaboratively to generate consistent and relevant land cover, characteristics, and condition information at the national scale for a wide variety of environmental, land management, and modeling applications. The MRLC coordinates and hosts the NLCD, which is comprised of products for several variables: land cover, tree canopy cover, imperviousness, and others.

Within the USFS, several key stakeholders and supporters collaborate on the TCC mapping project by contributing expertise, funding, and other resources. These stakeholders and supporters include representatives from the following groups: Forest Service Research and Development, the Forest Inventory and Analysis (FIA) program, the National Forest System, State and Private Forestry, and the Geospatial Technology and Applications Center (GTAC).

Historically, the United States Geological Survey (USGS) and the Earth Resources Observation and Science Center (EROS) produced TCC products for the NLCD (Huang et al., 2001; Homer et al., 2004; Homer et al., 2007). Beginning with the NLCD 2011 product suite, the USFS assumed responsibility for building the TCC component of the NLCD (Coulston et al., 2012; Homer et al., 2015). The TCC products within the NLCD 2011 product suite are referred to as the 2011 “epoch 1” TCC product suite and were released in 2014. For the NLCD 2016 product suite (publicly released in 2019), the USFS updated the 2011 TCC products (“epoch 2”) and generated a new set of TCC products for 2016 (Yang et al., 2018). Please note that all subsequent mentions of 2011 TCC products within this report refer to the “epoch 2” (updated) version of the 2011 products that were released in 2019.

Metadata and documentation accompany the TCC maps. The maps, metadata, and documentation together are called a “product suite” or “products”. TCC products are available for the conterminous United States (CONUS), southeastern portions of coastal Alaska (hereafter referred to as “coastal

Alaska”), Hawaii, Puerto Rico, and the U.S. Virgin Islands (PRUSVI). The coastal Alaska, Hawaii, and PRUSVI areas are collectively referred to as “outside of CONUS” or “OCONUS” for this project.

This report describes workflows and methods used to build the 2016 TCC product suite that consists of updated TCC products for 2011 and new TCC products for 2016, including versions of the TCC products built specifically for the NLCD 2016. For the NLCD 2016, there are three primary maps: TCC maps for the years of 2011 and 2016, and a TCC map that represents change between 2011 and 2016. The maps were generated using over 60,000 training plots with a probabilistic sample design to train machine learning models. Change in TCC was constrained to statistically significant change, as estimated through numerical optimization with FIA disturbance data.

Study Areas

The 2016 TCC product suite focuses on four areas: CONUS, coastal Alaska, Hawaii, and PRUSVI (figure 1).

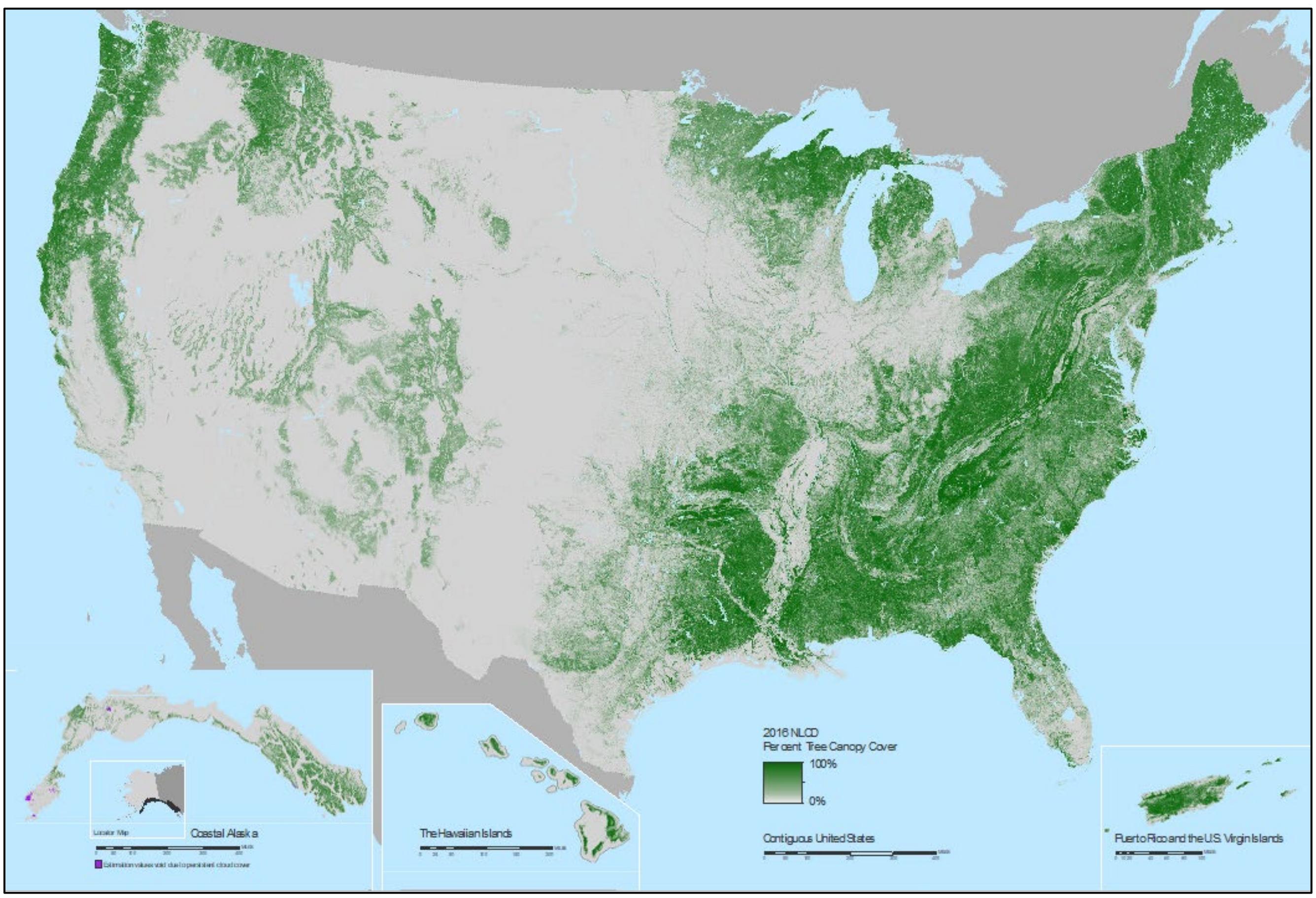

Figure 1.—Depiction of the tree canopy cover study areas for conterminous Unite States, coastal Alaska, Hawaii, Puerto Rico, and the US Virgin Islands.

Conterminous United States

CONUS consists of diverse ecological landscapes, including dry deserts, wet coastal forests, tropical coniferous forests, forested mountains, and temperate deciduous forests. CONUS covers 8.0 million km^2, of which 3.3 million km^2 are considered forested lands (United States Department of Agriculture Forest Service, 2017). Elevations range from 86 m below sea level in Death Valley, California to 4,418 m at Mount Whitney, California.

Coastal Alaska

The coastal Alaska study area mapped as part of the 2016 TCC product suite is a 219,000 km^2 region comprised of southeast and south-central coastal Alaska (figure 1) with elevations ranging from sea level to 5,489 m at Mount St. Elias. The 61,000 km^2 that is forest land in this study area is predominantly (75 percent) coastal rainforest with warm annual temperatures and high precipitation dominated by western hemlock (*Tsuga heterophylla*), mountain hemlock (*T. mertensiana*), Alaska yellow cedar (*Callitropsis nootkatensis*), and Sitka spruce (*Picea sitchensis*) (Cahoon et al., 2020). Other parts of this study area including the western Kenai Peninsula are cooler and drier and serve as transition zones between coastal rainforest and boreal forest. These areas are dominated by aspen (*Populus tremuloides*), paper birch (*Betula neoalaskana*), white spruce (*Picea glauca*), and black spruce (*P. mariana*) (Cahoon et al., 2020).

Hawaii

Hawaii is the most isolated archipelago on earth and is comprised of 137 islands in the Pacific Ocean stretching over 2,400 km. The majority (99 percent) of the land area is encompassed by eight main islands (figure 1) covering an area of 16,000 km^2 with about 6,000 km^2 of forest (Hauff et al., 2007). TCC maps are built for only these eight islands: Ni’hau, Kaua’i, O’ahu, Moloka’i, Lāna’i, Kaho’olawe, Maui, and Hawai’i.

The broad range of elevation (0 to 4,205 m) and climate creates a range of diverse ecosystems that include deserts, rain forests, and alpine (LaRosa et al., 2008). Waialeale Peak on the island of Kauai contains the rainiest spot on earth with a mean annual rainfall of 11.7 m while the deserts of Hawaii achieve annual rainfalls as low as 0.51 m. Vegetation types are diverse with an estimated 10,000 native species and over 8,000 introduced species (Little & Skolmen, 1989). Hawaii has more native tree species (approximately 300) than any other state. However, native Hawaii tree species are not numerous, and many are considered uncommon, rare, or endangered. Four native and five non-native generalized forest communities are recognized: 'ōhl'a--hāpu'u, koa--'ōhl'a, mamane--nalo, native dry forest, eucalyptus, mixed exotic hardwoods, guava, klawe--leucaena, and conifers (Little & Skolmen, 1989).

Puerto Rico and U.S. Virgin Islands

Puerto Rico is located 1,600 km southeast of Florida in the Caribbean Sea. Puerto Rico includes the main island of Puerto Rico and several smaller islands including Culebra, Vieques, and Mona (figure 1). The main island of Puerto Rico is 8,900 km^2 in area with approximately 5,000 km^2 area of forest land (Brandeis & Turner, 2013). Culebra, Vieques, and Mona are 30 km^2, 135 km^2, and 56 km^2, respectively. The sizes of the forested areas for Culebra, Vieques, and Mona are 27 km^2, 100 km^2, and 46 km^2, respectively (Brandeis & Turner, 2013). Puerto Rico is mainly mountainous with the highest peak being Cerro de Punta with an elevation of 1,340 m.

The US Virgin Islands (figure 1) are located 64 km east of Puerto Rico. The US Virgin Islands include St. Croix with an area of 219 km^2 with 110 km^2 of forest, St. Thomas at 90 km^2 with 67 km^2 of forest, and St. John at 53 km^2 with 49 km^2 of forest (Brandeis & Oswalt, 2007). The highest point occurs on St. Thomas at 474 m.

The islands of Mona and Culebra have only one forest type which is subtropical dry forest. The other islands either have subtropical moist forest (Puerto Rico, St. Thomas) as the predominant forest type and subtropical dry forest as the secondary forest type, or vice versa (Ewel & Whitmore, 1973). Puerto Rico has small amounts of other forest types including subtropical wet forest, subtropical rain forest, subtropical lower montane wet forest, and subtropical lower montane rain forest (Ewel & Whitmore, 1973). Subtropical dry forests of PRUSVI are found at elevations below 300 m where annual precipitation totals 600 to 1100 mm (Brandeis & Oswalt, 2007). The trees have low floristic diversity and generally do not exceed 15 to 20 m in height. Trees tend to be deciduous with sparse tree crowns and with small succulent or leathery leaves (Lugo et al, 2006; Brandeis & Oswalt, 2007). The transition from subtropical dry forest to subtropical moist forest happens gradually and, thus, the two types have many species in common. Trees of subtropical moist forests are typically evergreen and are found in areas with annual precipitation totals of 1100 to 2200 mm with a two-to-four-month dry period. Subtropical moist forests are more structurally developed and can have tree heights up to 30 m (Brandeis & Oswalt, 2007).

Methods

The major steps for creating the maps within the 2016 TCC product suite were 1) reference data collection, 2) preparation of predictor data layers, and 3) modeling and mapping.

Reference Data Collection

Reference data were collected using high-resolution imagery and the Canopy Cover Tool, developed by the USFS as an ArcMap™ extension (Goeking et al., 2012) (figure 2). To compute a plot’s estimated tree canopy cover percentage, a grid of 109 dots (with a spacing of 26.25 feet) are placed within a 144-foot radius circle that is centered on the plot. These 109 dots are rotated 15 degrees to avoid following linear anthropomorphic features in cardinal directions. Photo-interpreters indicate whether the dot falls upon tree canopy or not and the percent tree canopy cover is calculated for the plot. Interpreters also assigned a confidence level in their photo-interpretations for each plot. The plots used for reference data collection were located on the FIA grid (Reams et al., 2005).

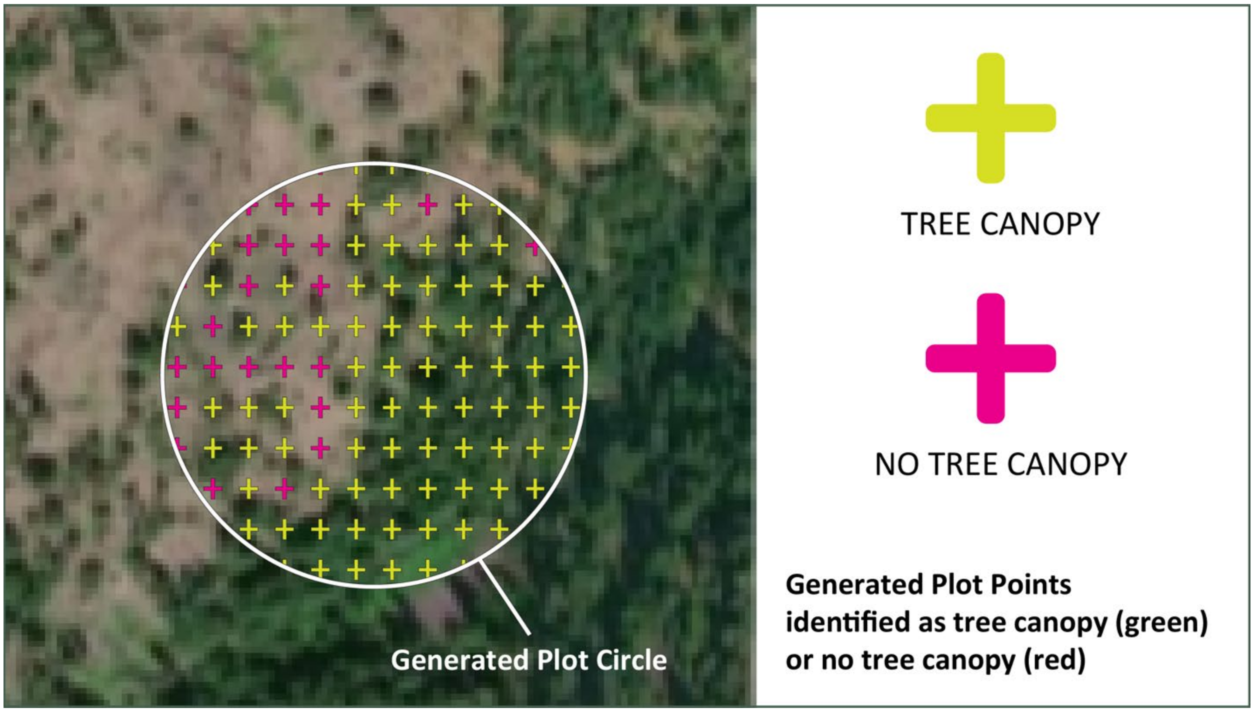

Figure 2.— A depiction of the Canopy Cover tool used to develop tree canopy cover reference data.

Conterminous United States

Reference data for 63,011 plots across CONUS were collected by FIA regional analysts through photo-interpretation of imagery from the National Agriculture Imagery Program (NAIP) (Farm Service Agency, n.d.). NAIP from 2010 was used to photo-interpret 66 percent of the reference data. The remainder used NAIP from 2009 for photo-interpretation. To the extent possible, these reference data were utilized in production efforts for the 2016 TCC product suite. The reference data were examined for obvious changes in TCC between the years of 2011 and 2016. The photo-interpreted levels of confidence were also reviewed.

To identify reference data in areas where there were obvious TCC changes between 2011 and 2016, two primary datasets were used: Monitoring Trends in Burn Severity (MTBS) burned area perimeters from 2009 – 2016 (Eidenshink et al., 2007) and a change layer developed using Landsat imagery Normalized Difference Vegetation Index (NDVI) (Kriegler et al., 1969) from two time periods (2007-2010 and 2013-2016). Intersecting the MTBS data and the change layer with the reference data identified 2,112 plots as having experienced change between the nominal years of 2011 and 2016. For these plots, analysts collected reference data using 2015/2016 NAIP and the Canopy Cover tool. Analysts were not able to confidently photo-interpret 113 of the 2112 plots due to shadows in imagery. These plots were discarded from the reference dataset.

The confidence scores of reference data were also reviewed. Data with low confidence scores (476 plots) were removed from the reference dataset. After updating the reference dataset, 62,535 and 62,442 reference data values were used for modeling 2011 and 2016 TCC for CONUS, respectively.

Coastal Alaska

High-resolution NAIP imagery with state-wide coverage is not available for Alaska like it is for CONUS. For TCC reference data collection in coastal Alaska, a mix of imagery was used. For the Chugach and Tongass National Forests, natural resource digital orthophotos with a resolution of 30 cm to 1 m were provided by Region 10 of the USFS (Bellante et al., 2020). For areas outside of the national forests, imagery from the Alaska's Statewide Digital Mapping Initiative (SDMI) Best Data Layer (BDL) was used where available. The BDL is a 'best available' imagery base layer from the SDMI archive (Statewide Digital Mapping Initiative, n.d.). It contains imagery with a wide range of resolutions and acquisition dates from several sources contributed by 17 agencies. High-resolution images with a resolution of 1 m or less from the BDL were used for reference data collection in coastal Alaska where they were available. In areas where no USFS natural resource digital orthophotos or BDL imagery existed, Digital Globe's WorldView-3 imagery provided by the State of Alaska Open Data Geoportal web map services was used (Alaska Geospatial Council, n.d.).

In 2014, analysts photo-interpreted TCC reference data using the Canopy Cover tool on 2,035 FIA plots. These reference data were reviewed to determine their suitability as reference data for the 2016 TCC product suite. There were originally 325 out of the 2,035 plots that had low confidence scores in the photo-interpretation. The reinterpretation changed 11 plots from low confidence to high confidence and one plot changed from high confidence to low confidence. Thus, there were 315 plots with low confidence photo-interpretation scores, and these were discarded. There were 454 plots in offshore water, and they were discarded. This left 1,266 photo-interpreted plots for use in the 2011 modeling.

For the nominal year of 2016, 1,581 plots were examined for use as reference data. These plots included all the 2,035 original plots minus the 454 offshore water plots. The 315 low confidence plots from 2011 were not initially discarded because there was a chance that some of these plots would be flagged for reinterpretation and their confidence scores could change. Plots that had visually discernable changes in TCC or plots where higher resolution imagery was available when it was not available previously were reinterpreted (221 plots). Out of these 221 plots, 143 plots had their confidence scores changed from low to high and 12 plots had their confidence scores change from high to low. Overall, there were 184 low confidence plots (315-143+12=184) and they were discarded. For the plots that were not reinterpreted, the photo-interpreted 2011 TCC estimate was used. The final number of plots used for training models in the 2016 production was 1,397.

Hawaii

High-resolution imagery for 2011 and 2016 from Worldview-1, Worldview-2, Worldview-3, and GeoEye-1 with spatial resolutions of 50 cm, 46 cm, 30 cm, and 50 cm respectively, were used for photo-interpretation (Maxar, n.d.). In 2015 analysts collected TCC reference data on 1,709 plots in Hawaii. These reference data were assessed for suitability as reference data for the 2016 TCC product suite. Plots with nonzero TCC were evaluated for discernable visible changes that occurred since 2011. There were 50 plots with changes, and they were photo-interpreted. None of the 50 plots received low confidence scores. There were 100 plots with low confidence photo-interpretation scores for 2011 and 2016. There was an additional plot that was discarded for 2016 because of a technical issue with the plot. A total of 1,609 plots were used for 2011 TCC and 1,608 plots were used for 2016 TCC.

Puerto Rico and U.S. Virgin Islands

High-resolution imagery from 2011 and 2016 from Worldview-1, Worldview-2, Worldview-3, and GeoEye-1 with spatial resolutions of 50 cm, 46 cm, 30 cm, and 50 cm respectively, were used for photo-interpretation (Maxar, n.d.). In 2015, analysts collected TCC reference data for 1,246 plots in PRUSVI. All plots were assessed for suitability as reference data for the 2016 TCC product suite. Plots were evaluated for discernable visible changes that occurred since 2011 and, if changed, were photo-interpreted (75 plots). Of the 75 plots that were photo-interpreted for 2016, there were 22 that had low confidence photo-interpretation scores. These plots were discarded. No plots received low confidence scores for 2011. A total of 1,246 plots were used for 2011 TCC and 1,224 were used for 2016 TCC.

Predictor Layers

Predictor layers included Landsat median composites and image derivatives such as NDVI, normalized difference moisture index (NDMI) (Gao, 1996), tasseled cap transformation (Kauth & Thomas, 1976; Crist & Cicone, 1984; Baig et al., 2014), and exponentially weighted moving average (EWMA) layers (Brooks et al., 2012). Topographic data including elevation, aspect, slope, and gradients were also used. Additionally, PRISM (Parameter-elevation Regressions on Independent Slopes Model) climate data were used for mapping TCC in PRUSVI (Daly et al., 2000). Details about these predictor datasets are in the following sections.

Landsat Median Composites

Landsat 5 (L5) Thematic Mapper and Landsat 8 (L8) Operational Land Imager imagery were used in developing the median composites. Landsat 7 Thematic Mapper imagery was not included because of the scan line corrector (SLC) failure that occurred in May 2003. Mapping TCC with Landsat 7 SLC images produces striped images.

A customized ERDAS IMAGINE® 2013 (ERDAS, 2013) tool was used to produce the median composites for all the 2011 TCC study areas and for 2016 TCC CONUS (Ruefenacht, 2016). Google Earth Engine (GEE) (Gorelick, 2017) was used to collect Landsat imagery and produce median composites for 2016 TCC in Hawaii, coastal Alaska, and PRUSVI. Cloud and shadow masks were created using the FMASK algorithm (Zhu & Woodcock, 2012) and applied to each associated Landsat scene. Masked areas were excluded from the median calculations. For a given WRS-2 path/row, median values were calculated on a pixel-wise basis from the selected Landsat scenes. Since each of the six bands of the Landsat scenes were considered separately, the six values (bands) in a composite pixel could come from different scenes representing different dates. The compositing procedure returned two images in addition to the median composite. One image contained the number of input pixels that were used in the median calculation. The other image contained the pixel-wise date used for the median value. This image included 12 layers, allowing up to two dates per band for each pixel. Two dates were stored for median values calculated using an even number of pixels.

The original 2011 median composites were created in 2013. Because of computing limitations at the time, the overlap areas for the 2011 median composites were computed individually. To merge the overlapping median composites, the median composite pixel used for the overlap area came from the median composite with the highest number of input pixels. This was done based on the assumption that a median calculated from more values was better. When the median composites were created for 2016, technology had improved, allowing the median composite process to include overlapping areas. The 2016 TCC median composites for the OCONUS study areas were created using GEE. For these median composites, there weren't any computing limitations and all input pixels were used for median calculations.

All the L5 scenes used to create the 2011 median composites were converted to top-of-atmosphere (TOA) reflectance using the radiometric calibration coefficients and equation 2 in Chander et al. (2009) and to surface reflectance using dark-object-subtraction (DOS) (Chavez, 1988). Scenes were converted from floating point to 8-bit unsigned integer using an 8-bit min-max linear stretch. Scenes were projected to GRS1980, NAD83, Albers Conical Equal Area and snapped to a grid based upon the NLCD 2001 Land Cover grid (Homer et al., 2007).

All imagery used to create the 2016 CONUS median composites were obtained from the EROS Science Processing Architecture on Demand Interface website (EROS Science Processing Architecture – Land Satellites Research and Development, n.d.). L8 scenes were processed to TOA using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Schmidt et al., 2013) and were projected to GRS1980, NAD83, Albers Conical Equal Area and snapped to a CONUS grid based upon the NLCD 2001 Land Cover (CONUS) grid (Homer et al., 2007).

All the OCONUS 2016 median composites created in GEE were exported from GEE as 16-bit integer, 6-band geotiffs. Additionally, geotiff images with pixel-wise information on the number of L8 scenes used to calculate the median, and a date image representing which L8 scene(s) were used for the median calculation were exported. Following export, the geotiffs were projected to the GRS1980 NAD83, Albers Conical Equal Area projection and snapped to a grid based upon the NLCD 2001 Land Cover grid (Homer et al., 2007).

For 2011, composites were processed to surface reflectance. For 2016, the composites were processed to TOA reflectance. An investigation examining the effects of surface reflectance and TOA on mapping TCC concluded that the mapping of TCC was not influenced by the type of correction. More information about this investigation can be found in Appendix A.

Each median composite was visually inspected, and two types of problems were found: 1) holes in the datasets caused by clouds and shadows, and 2) seam line issues. Because the cloud and shadow mask algorithms use thermal data to aid in the identification of clouds, areas that are cooler than the surroundings (such as isolated mountain ranges adjacent to desert valley floors) are incorrectly classified by the algorithms to be cloudy when the areas are cloud free. Extremely bright areas, such as building rooftops or permanent ice fields, can also be interpreted as clouds. If all Landsat scenes used for the median composite had these issues, holes would appear in the median composite. It is also feasible that all the Landsat scenes used for the median composite were perpetually cloudy, which also caused holes in the median composite. For these types of situations, either the cloud masks were edited to remove false clouds or additional Landsat scenes were obtained that were cloud-free in the problem areas. The median composites were recreated after the holes were addressed.

Seam lines occasionally appeared in the median composites due to different dates of adjacent Landsat scenes. These problems were addressed by selecting additional Landsat scenes with narrower date ranges. In situations where limited Landsat scenes were available for scene selection due to persistent cloud cover, seam lines were left in the median composites. The effect of seam lines in the median composite on TCC maps is addressed in the Modeling and Mapping section.

Due to varying climates and latitudes, each study area had different date ranges used for imagery acquisition and minor differences regarding the compositing procedures. More details are provided in the sections below on the image selection process and compositing methods for each study area.

Conterminous U.S. median composites

Conterminous U.S. 2011 median composites

The objective for L5 scene selection for the 2011 median composites was to select 15 scenes for each of the 436 Worldwide Reference System-2 (WRS-2) path/rows covering CONUS. Since the majority of the reference data were photo-interpreted using imagery from 2010, L5 scenes were primarily selected from this year as well. The Landsat Image Viewer tool was used for image selection. Appendix B provides more information about this tool. The filters used for L5 scene selection were growing season window and cloud cover. Growing season windows were derived from NDVI graphs generated from 1 km Advanced Very High-Resolution Radiometer (AVHRR) data (figure 3). The growing season and cloud cover constraints are shown in table 1. Because of narrow growing season windows and cloud cover constraints, it was difficult to obtain all 15 scenes from a single year and, thus, the year range was expanded to 2007 through 2011. (Detailed information about the dates of imagery used to model TCC in specific areas of the CONUS is available upon request; please see the Summary section for contact information.) A total of 6,540 L5 scenes were selected and downloaded using the USGS Visualization Viewer (GloVis) (United States Geological Survey Earth Resources Observation and Science Center, 2001).

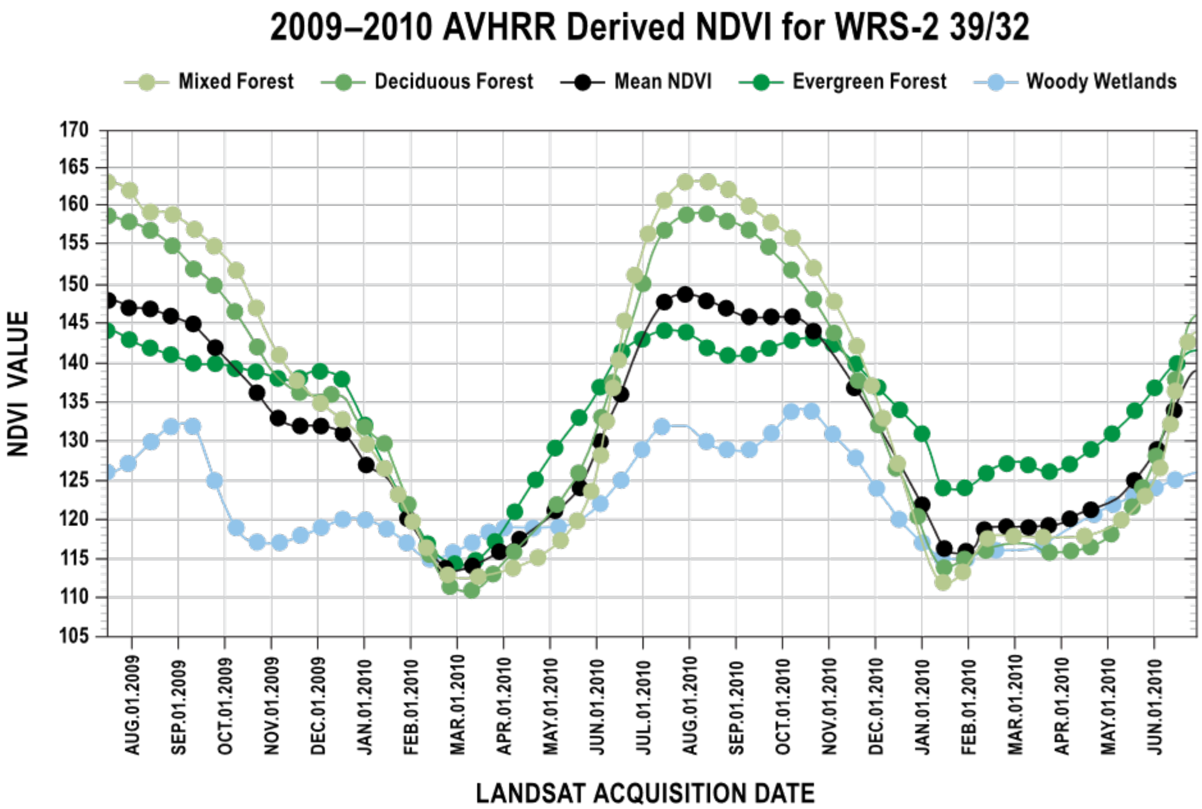

Figure 3.—Example of Normalized Difference Vegetation Index (NDVI) graph produced by the Landsat Image Viewer.

Table 1.—Constraints used for Landsat scene selection.

Constraint	Value	Comments
Max Cloud Clover	70%	Cloud cover between 0 and 70%
Max Distance from Peak NDVI	10%	Based on NDVI curve for dominant land cover class (woody wetlands and mixed, deciduous, and evergreen forest) for each WRS-2 path/row

Conterminous U.S. 2016 median composites

Similar Landsat scene selection concepts developed for 2011 median compositing were used for 2016 median compositing, but the tools used to select the scenes were different. Instead of using the Landsat Image Viewer tool, GEE was used to determine appropriate Landsat scene selection date ranges based on growing seasons windows, which were determined from NDVI graphs created from L8 TOA collections (figure 4). For 2016 TCC, there was no minimum or maximum number of scenes required for each WRS-2 path/row. The number of L8 scenes per WRS-2 path/row varied from a high of 33 scenes for a WRS-2 path/row in Arizona to a low of one L8 scene for a WRS-2 path/row in Michigan. A total of 2,212 L8 scenes were selected. The customized ERDAS IMAGINE® 2013 tool was used to produce the median composites.

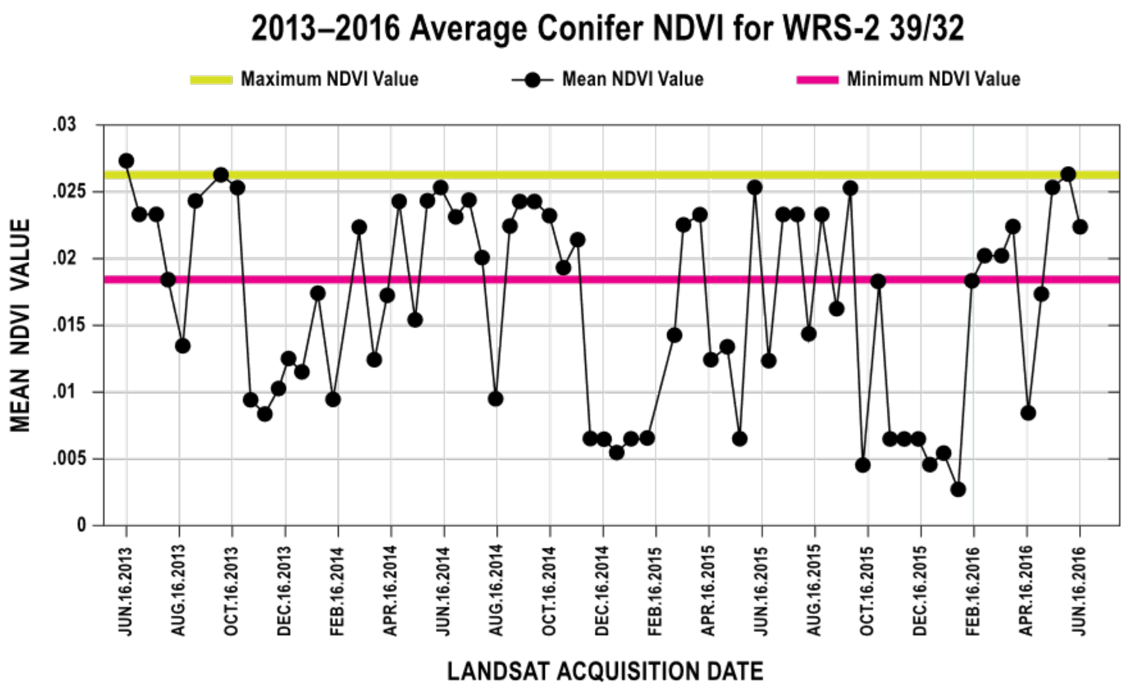

Figure 4.—Google Earth Engine generated Normalized Difference Vegetation Index (NDVI) graph for Landsat Operational Land Imager (OLI) for coniferous forest type for WRS-2 39/32. The green line depicts the maximum NDVI value, and the red line represents the minimum NDVI value (set to 70 percent of the maximum NDVI value) used to select Landsat scenes.

Coastal Alaska median composites

Coastal Alaska 2011 median composites

There are many factors that affect the coverage and acquisition of Landsat 5 imagery such as business priorities, mission operation decisions, technical capabilities, and clouds (Goward et al., 2006). These factors influenced the number of L5 scenes available for the 2011 median composite. Two of the WRS-2 path/rows covering coastal Alaska had no L5 scenes available, and three had a single L5 scene. The remaining 46 WRS-2 path/rows covering coastal Alaska had 11 to 49 L5 scenes available. There is considerable overlap of the WRS-2 path/rows for Alaska and, thus, there was enough overlap from adjacent WRS-2 path/rows to compensate for WRS-2 path/rows with no or a low number of L5 scenes. Given the limited number of available scenes, the limited number of days in the growing season, high latitude, and excessive cloud cover, it was difficult to obtain imagery from a single year (2010) and, thus, the year range was expanded to 2006 through 2011.

Peak greenness was identified using the day with maximum NDVI from 2006 through 2011 obtained from the Geographic Information Network of Alaska (GINA) (Zhu et al., 2013). Table 2 shows the selected dates. One to nine L5 scenes for the 49 WRS-2 path/rows were downloaded from GloVis for a total of 202 L5 scenes covering coastal Alaska. The FMASK algorithm erroneously classified snow and ice fields as clouds. Manual editing of the cloud and shadow masks was required to fix these problems before the composites could be created.

Table 2.—Date constraints used for Landsat scene selection for coastal Alaska derived from day of maximum Normalized Difference Vegetation Index (NDVI) data provided by the Geographic Information Network of Alaska. The left table depicts the acquisition dates for developing 2011 maps and the right depicts the acquisition dates for 2016 maps.

Year	Date Range
2006	3 Jul – 12 Aug
2007	8 Jul – 19 Aug
2008	1 Jul – 12 Aug
2009	1 Jul – 6 Aug
2010	10 Jul – 26 Aug
2011	1 Jul – 18 Aug
2013	8 Jul – 20 Aug
2014	1 Jul – 28 Aug
2015	12 Jun – 27 Aug
2016	11 Jun – 2 Sep
2017	11 Jun – 30 Jul

Coastal Alaska 2016 median composites

For the 2016 composite, the years 2014 through 2016 were initially considered for L8 scene selection, but persistent clouds in some areas warranted the use of additional scenes from 2013 and 2017. GINA was used to determine the maximum NDVI date and date ranges to use for L8 scene selection (table 2). A total of 1,017 L8 scenes were accumulated for coastal Alaska.

GEE was used to generate the 2016 median composite. The median composite was inspected for cloud and shadow contamination. If a cloud and shadow mask was unsatisfactory, an NDVI mask was applied. If cloud or shadow masked pixels had a NDVI > 0.5, the pixels were used in compositing and the median composite was recreated. If the NDVI mask did not fix the observed problems, the Landsat scene was discarded, and the median composite was recreated.

Hawaii median composites

Hawaii 2011 median composites

Ten WRS-2 path/rows cover the eight main Hawaii islands. L5 stopped collecting imagery over Hawaii in 1991 and, thus, the earliest year Landsat imagery was available was 2013 with L8. Because Hawaii has a year-round growing season, all scenes for the years 2013 through 2015 were reviewed regardless of L8 acquisition date. There were no scenes that were completely unusable. However, there were areas with persistent cloud cover in all scenes. There were 121 L8 scenes selected and downloaded from GloVis.

Hawaii 2016 median composites

Using GEE, all L8 scenes available from 2016 through 2018 were collected and reviewed for the 10 WRS-2 path/rows covering the eight main islands. There was a total of 517 L8 scenes used for the 2016 median composite.

Puerto Rico and US Virgin Islands median composites

Puerto Rico and US Virgin Islands 2011 median composites

Six WRS-2 path/rows cover PRUSVI. L5 stopped collecting imagery over PRUSVI in 2001 and, thus, the earliest year Landsat imagery was available for PRUSVI was 2013 with L8. L8 scenes for the 2011 composite were acquired during the dry season (mid-October through mid-March) from 2013 through 2015. These scenes were inspected for cloud cover and any other scene quality issues. There were areas with persistent cloud cover in all scenes, but there were no scenes that were completely unusable. Seven to eight scenes were obtained for each of the WRS-2 path/rows for a total of 46 scenes for PRUSVI.

Puerto Rico and US Virgin Islands 2016 median composites

Using GEE, L8 scenes for 1 Jan 2016 through 20 Sep 2017 were assembled for the 2016 composite. L8 scenes were not obtained after 20 Sep 2017 because Hurricane Maria made landfall at this time. The PRUSVI 2016 TCC map assumes pre-hurricane TCC conditions. Analysis of NDVI across the calendar year revealed that NDVI was stable across time and, thus, L8 scenes were acquired for all months. There was a total of 240 L8 scenes used for the median composite.

Landsat Derivatives and Ancillary Data Layers

For each median composite image, NDMI, NDVI, and Tasseled Cap Transformation images were created to use as predictors in the TCC modeling and mapping. Exponentially Weighted Moving Average (EWMA) images, utilizing algorithms developed at Virginia Polytechnic Institute and State University (Brooks et al., 2012) and generated for the TCC project by collaborators at Oregon State University, were derived from Landsat time-series using advanced harmonic regression techniques that help to characterize temporal variability. EWMA images for CONUS and coastal Alaska were generated for 2011 using L5 scenes from 2009 to 2011, and for 2016 using L8 from 2014 to 2016. L5 imagery was not available for Hawaii and PRUSVI, so EWMA images generated for 2011 used L8 scenes from 2013 to 2015, and images generated for 2016 used L8 scenes from 2016 to 2017. EWMA time series images were produced for short-wave infrared-1 (SWIR1), short-wave infrared-2 (SWIR2), and NDVI with each having three layers (intercept, cosine, and sine) and a spatial resolution of 30 m.

Different topographic data with various resolutions were available for each study area. For CONUS, the 30 m USGS national digital elevation model (Gesch et al., 2018) was used. The coastal Alaska modeling process used the 30 m ASTER GDEM 2 (Tachikawa et al., 2011). The Hawaii modeling process used a 10 m digital elevation model (DEM) provided by the National Oceanic and Atmospheric Administration (NOAA) (Department of Commerce et al., 2007). The PRUSVI modeling process used a 10 m DEM also provided by NOAA (Taylor et al., 2008). The 10 m datasets were resampled to 30 m using cubic convolution resampling. Slope, aspect, and sine and cosine of aspect layers were produced for all elevation datasets.

PRISM climate data were acquired and used as predictor data in mapping TCC for PRUSVI. Mean monthly and mean annual precipitation averages for a 32-year period from 1963 to 1995 were obtained from the International Institute of Tropical Forestry (International Institute of Tropical Forestry, n.d.).

Modeling and Mapping

The 2016 TCC product suite includes three sequential versions of TCC maps: analytical, cartographic, and NLCD. The different versions are designed to serve different user communities with a wide range of geospatial and statistical analysis resources available to them. The 2011 and 2016 analytical maps are TCC as predicted from the models, along with associated standard errors for 2011 and 2016. The 2011 and 2016 cartographic maps have a cleaner visual appearance and were created by masking water bodies, non-tree croplands, and non-tree pixels from the 2011 and 2016 analytical maps. The NLCD version includes a 2011 TCC map, a 2016 TCC map, and a change layer representing TCC change between the nominal years of 2011 and 2016. The three components in the NLCD version meet the criterion of ‘2011 TCC + change in TCC = 2016 TCC’ for all pixels. Information about the development of these maps is provided below.

Analytical Modeling and Maps

The same procedures were used to create both 2011 and 2016 analytical maps. Reference data were intersected with predictor data, creating modeling datasets. Random forest models were created using the tuneRF function in R, which attempts to find the optimal number of variables (mtry) to use at each split of the decision tree (Breiman, 2001; Liaw & Wiener, 2002; R Core Team, 2017). For the random forest models, the corr.bias option was set to TRUE. The number of decision trees used was 500 and, thus, when the random forest model was applied to the predictor data, each pixel had 500 TCC estimates. Mean TCC and standard error values were calculated from these 500 TCC estimates and assigned to the analytical TCC and standard error layers. No thresholds or masking was applied to the analytical maps. Thus, TCC values exist for non-tree areas such as water, roads, and grasslands in the analytical versions of the maps. Details on these procedures are provided in the following sections.

Conterminous U.S. Analytical Maps

To control for the ecological variability in CONUS, the reference data were divided into 68 zones (figure 5) (Homer & Gallant, 2001). Reference data were intersected with predictor data to create modeling datasets for each zone.

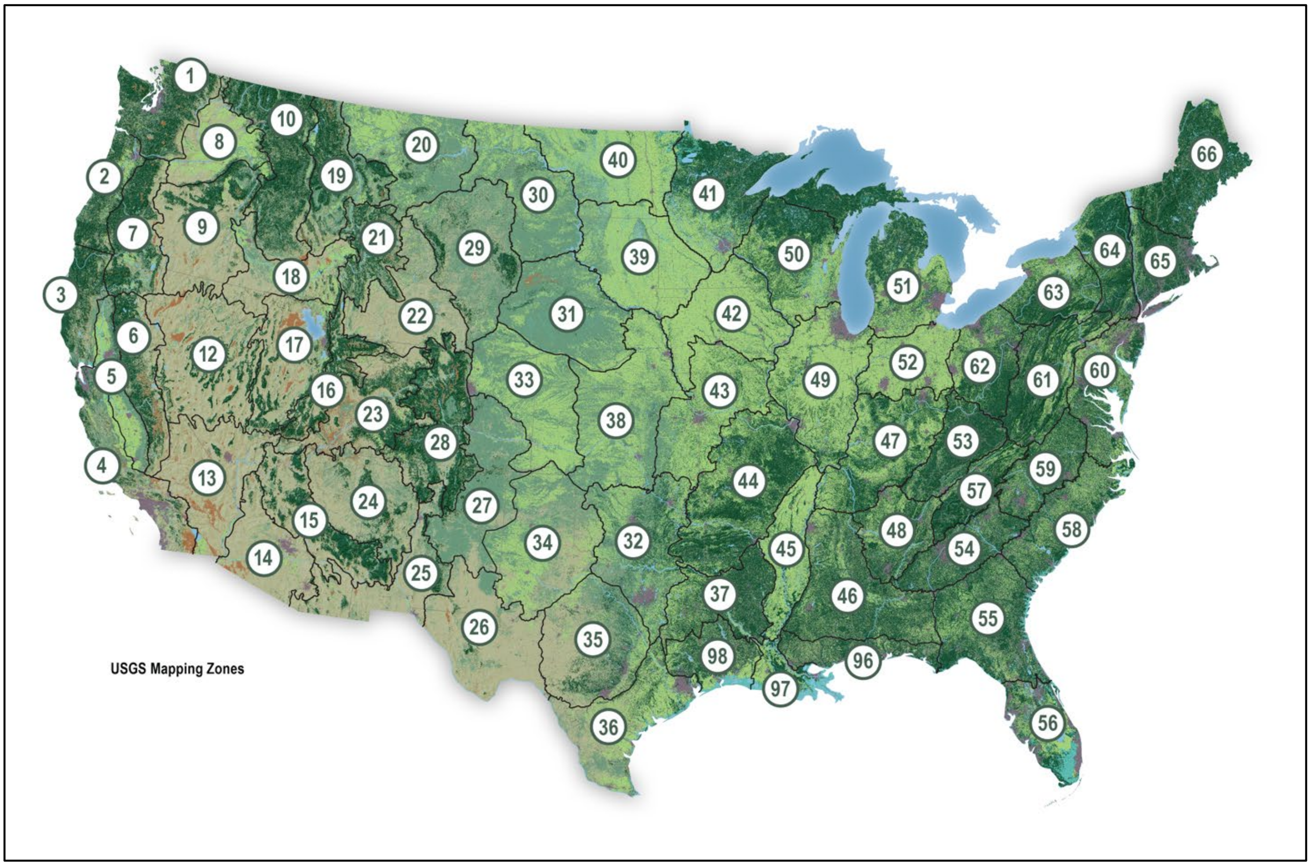

Figure 5.—United States Geological Survey (USGS) mapping zones used in the mapping process.

All reference data were used for model building. The following model performance metrics were derived from the out-of-bag data (table 3 and figures 6-7). For the 68 zones for 2011 TCC, the maximum root mean square error (RMSE) was 22.1 and the minimum RMSE was 2.4 with a mean of 13.4 and standard deviation of 5.1. The percent variance explained had a high of 86.9 and a low of 13.7 with a mean of 66.2 and standard deviation of 17.0. For the 68 zones for 2016 TCC, the model metrics suggested a weaker relationship between reference data and model predictions. The maximum RMSE was 23.9 and the minimum RMSE was 2.5 with a mean of 14.7 and standard deviation of 5.2. The percent variance explained had a high of 85.1 and a low of 9.4 with a mean of 62.5 and a standard deviation of 15.8. Zone 14, for both 2011 and 2016 TCC, was the zone with the lowest percent variance explained (13.7 and 9.4, respectively). This zone was southern Arizona, where 69 percent of the reference data recorded zero TCC and the mean TCC was 1.5.

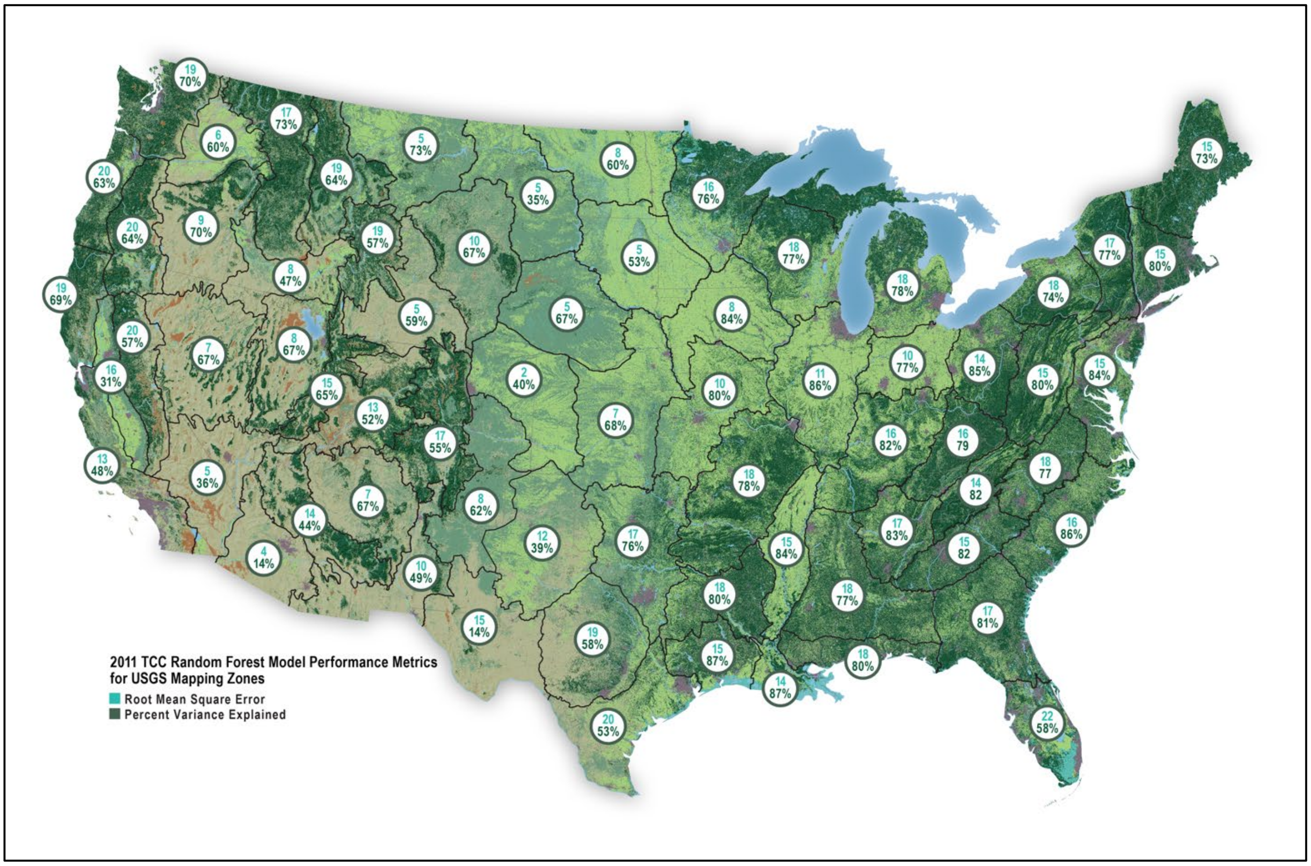

Figure 6.—2011 Tree Canopy Cover (TCC) model performance metrics derived from out-of-bag data. Information on root mean square error (RMSE) and percent variance explained was produced.

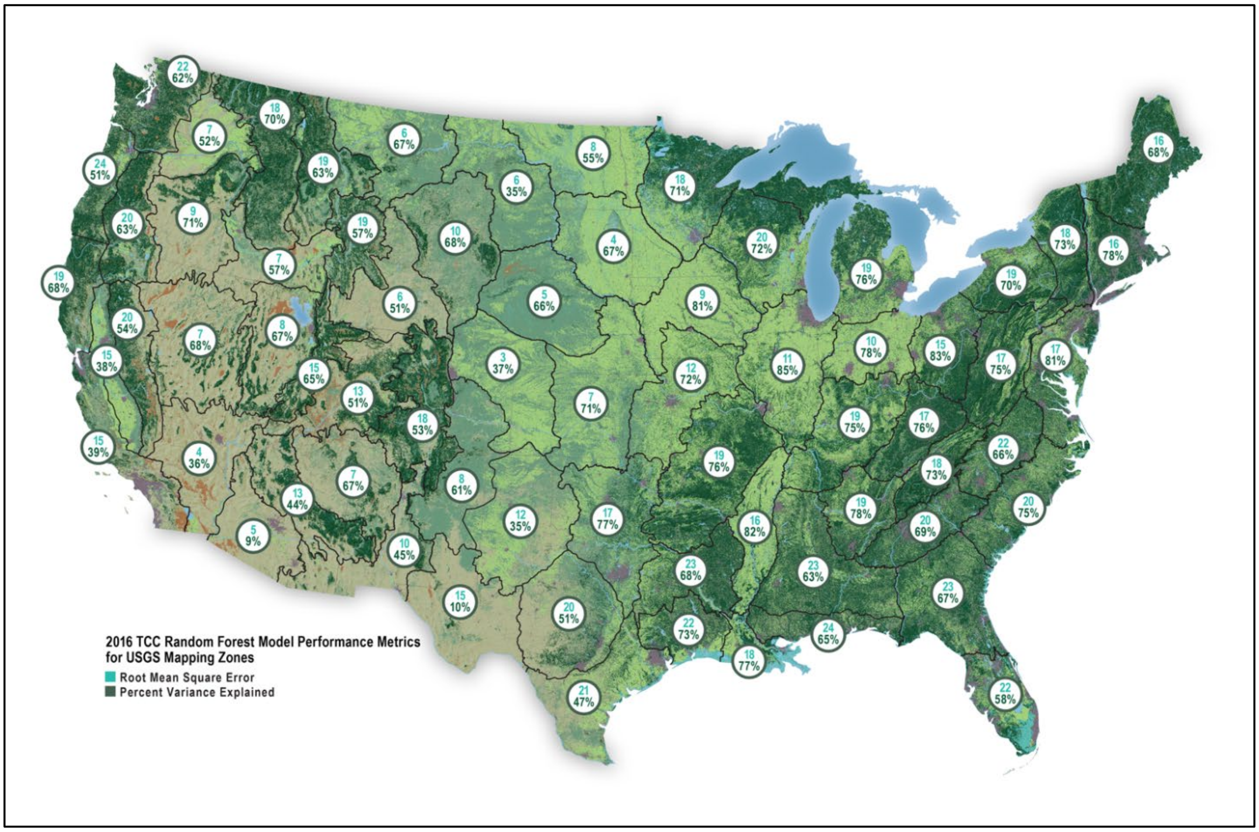

Figure 7.—2016 Tree Canopy Cover (TCC) model performance metrics derived from out-of-bag data. Information on root mean square error (RMSE) and percent variance explained was produced.

Table 3.—Model performance statistics calculated on the out-of-bag data from the random forest tree canopy cover models. Root mean square error (RMSE) and percent variance explained are shown here for the years 2011 and 2016 for the Conterminous US (CONUS) United States Geological Survey (USGS) mapping zones.

Mapping Zones	RMSE (2011)	% Variance Explained	RMSE (2016)	% Variance Explained
1	18.9	70.2	21.7	62.2
2	20.5	63.1	23.9	50.5
3	18.8	68.7	19.4	67.7
4	13.4	48.3	14.5	39.4
5	16.3	30.9	15.4	38.1
6	20	56.9	20.1	54.2
7	19.6	64.2	20.1	62.9
8	6.2	60.1	6.7	51.6
9	9.3	69.8	9.2	71.1
10	17	73.1	18.2	69.7
12	6.6	66.8	6.6	67.5
13	4.5	35.6	4.2	35.6
14	4.4	13.7	4.5	9.4
15	13.7	43.9	13.1	43.8
16	14.8	64.7	14.8	64.7
17	8.4	66.6	8.2	66.9
18	7.6	46.5	6.6	56.9
19	19.3	63.7	19.2	62.7
20	5.4	72.5	5.9	67.2
21	18.8	56.8	18.7	57.4
22	5.1	59	5.6	50.8
23	12.6	51.3	12.7	50.8
24	6.6	67	6.7	66.1
25	9.7	48.5	9.5	44.6
26	14.8	14.3	15.1	9.8
27	7.5	62.3	7.6	61.3
28	17.4	55.2	18	52.6
29	10	67.4	9.7	68.2
30	5.1	34.7	5.6	34.9
31	5.1	67	4.9	66.3
32	16.9	76.2	16.5	76.8
33	2.4	39.9	2.5	36.6
34	11.6	38.8	11.7	35
35	18.7	58.1	20.1	50.5
36	19.6	52.9	20.7	46.8
37	17.6	80.4	22.6	68.3
38	7.4	67.6	7	71.1
39	4.9	53.1	4.1	66.5
40	7.7	60	8.3	54.6
41	16	76.2	17.7	71.3
42	7.9	83.7	8.8	81
43	10	79.6	11.8	72.4
44	18.1	78.4	19.2	75.6
45	15	84.3	16.1	81.8
46	18.4	76.5	23.2	63.4
47	16.2	82.2	18.9	75.3
48	16.7	83.3	19.1	77.9
49	10.5	85.6	10.6	85.1
50	17.6	77.4	19.7	72
51	18.1	78	18.8	76.1
52	10.4	76.5	10.1	78
53	16.1	78.5	17.1	76
54	15	82.2	19.8	69.4
55	17.1	80.9	22.7	66.7
56	22.1	58.1	21.9	57.8
57	14.3	81.9	17.6	72.7
58	15.5	85.7	20.4	75.4
59	18.4	76.7	22.4	66.1
60	15.4	84.3	17.1	80.6
61	15.4	80.2	17.3	74.9
62	14	84.8	15	82.5
63	17.9	74.3	19.4	69.8
64	16.5	77.1	17.9	72.9
65	14.8	80.1	15.8	77.5
66	14.6	73.4	16.1	68.3
96	18.2	79.5	23.7	65.2
97	13.6	86.9	18	77.1
98	15	86.9	21.6	72.8

For each zone, the associated random forest model was applied to predictor datasets that were clipped to the individual Landsat WRS-2 path/rows boundaries that intersected the zone. It was possible for a WRS-2 path/row to be mapped multiple times because it intersected multiple zones. There were 84 WRS-2 path/rows that were mapped only once, 136 that were mapped twice, 161 that were mapped three times, 44 that were mapped four times, 9 that were mapped five times, and 2 that were mapped six times. The duplicate TCC images were combined into a single image by selecting the pixel from the available pixels with the lowest standard error, weighted by the amount of area that the intersecting zone overlapped with the corresponding WRS-2 path/row. This ensured that pixels mapped using the zone model whose zone occupied most of the WRS-2 path/row had a higher probability of being selected.

All model outputs (TCC maps in draft or intermediate forms) were carefully inspected by several experienced remote sensing analysts. Anomalies in individual WRS-2 path/row Landsat composites (visual artifacts) were present and were visible in the TCC draft map. Because 81 percent of the WRS-2 path/rows were mapped multiple times for different mapping zones, it was often possible to replace the anomalous artifacts for the WRS-2 path/row with another model output that was free of artifacts. However, there were instances where all WRS-2 path/row model outcomes had Landsat spectral anomalies. In these cases, the Landsat data were removed as predictors for the WRS-2 path/row, and the draft TCC map was generated by using just the EWMA data and topographic data as predictors for the WRS-2 path/row, resulting in a more visually representative output.

To analyze the effect of using EWMA data without Landsat data on the modeling of TCC, zones 41, 50, and 51 were combined into one zone and modeled and mapped with different combinations of predictor datasets. As table 4 shows, the random forest model using EWMA and topographic data had a slightly weaker relationship between reference and predictor data (RMSE 15.1, Percent Variance Explained 83.6) than the model using Landsat median composites, derivatives, and topographic data (RMSE 14.3, Percent Variance Explained 83.7). Additionally, difference images were created from the draft TCC maps generated using the models described in table 4. The majority (83 percent) of the TCC values were within $\pm$5 percent from each other. All differences had 97 percent of the TCC values within $\pm$20 percent from each other. These results suggest that for areas modeled without Landsat median composites and associated derivatives, there was minimal difference or impact.

Table 4.—Model performance metrics calculated on the out-of-bag data from the random forest tree canopy cover models created using different types of predictor data for USGS mapping zones 41, 50, and 51 for 2011.

Predictor Layers for Combined Zones (41, 50, and 51)	RMSE	% Variance Explained
Landsat Median Composite	21.8	65.8
EWMA	17.7	77.5
Landsat Median Composite w/ Derivatives & Topographic Data	14.3	83.7
EWMA w/ Topographic Data	15.1	83.6
All Predictors	14.7	84.5

There were also instances where remote sensing analysts determined that the TCC draft map values for a WRS-2 path/row were not representative of the actual TCC for the area and, instead, overlapping TCC values from other zones were more representative. In these instances, the image for the WRS-2 path/row with the more representative TCC values were used. After extensive review, the individual mapping zones were merged.

Outside Conterminous U.S. Analytical Maps

The amount of reference data for OCONUS was less than three percent of the amount of reference data used for CONUS. The area of OCONUS is also three percent the size of CONUS. Because of these size differences in both area and amount of reference data, the three OCONUS study areas were each mapped as their own unit and were not sub-divided into zones as was done for CONUS.

The VSURF (Variable Selection Using Random Forest) package in R was used to identify final predictor variables. VSURF uses a stepwise approach to identify variables that are most related to the response data (Genuer et al., 2019). Tables 5 through 7 show the selected predictors used for modeling and mapping each study area. VSURF was not used for CONUS because CONUS only had 28 predictor variables, whereas the OCONUS study areas had more available predictor variables. The mean of squared residuals and the variance explained calculated from out-of-bag samples are summarized for each study area in table 8.

Table 5.—Variables selected by Variable Selection Using Random Forest (VSURF) for final modeling for Coastal Alaska.

2011 Predictor Layers	**2016 Predictor Layers**
L5 Median Composite - Blue	L8 Median Composite - Blue
L5 Median Composite - Green	L8 Median Composite - Green
L5 Median Composite - SWIR 1	L8 NDVI
L5 NDVI	L8 NDMI
L5 Tasseled Cap - Brightness	L8 Tasseled Cap - Brightness
L5 Tasseled Cap - Greenness	L8 Tasseled Cap - Greenness
EWMA NDVI - Intercept	EWMA NDVI - Intercept
EWMA NDVI - Sine	EWMA NDVI - Sine
EWMA NDVI - Cosine	EWMA SWIR 1 - Intercept
EWMA SWIR 1 - Sine	EWMA SWIR 1 - Cosine
EWMA SWIR 1 - Cosine	Elevation
EWMA SWIR 2 - Intercept	Elevation

Table 6.—Variables selected by Variable Selection Using Random Forest (VSURF) for final modeling for the Hawaiian Islands.

2011 Predictor Layers	2016 Predictor Layers
L8 Median Composite - Red	L8 Median Composite - Green
L8 Median Composite - SWIR 2	L8 Median Composite - Red
L8 NDVI	L8 Median Composite - SWIR 1
L8 NDMI	L8 Median Composite - SWIR 2
EWMA NDVI - Intercept	L8 NDVI
EWMA NDVI - Cosine	L8 NDMI
EWMA SWIR 1 - Cosine	L8 Tasseled Cap - Greenness
Elevation	EWMA NDVI - Intercept
	EWMA NDVI - Cosine
	Elevation

Table 7.—Variables selected by Variable Selection Using Random Forest (VSURF) for final modeling for Puerto Rico and the U.S. Virgin Islands.

2011 Predictor Layers	2016 Predictor Layers
L8 Median Composite - Green	L8 Median Composite - Green
L8 Median Composite - Red	L8 Median Composite - Red
L8 Median Composite - SWIR 1	L8 Median Composite - SWIR 2
L8 NDVI	L8 NDVI
EWMA NDVI - Intercept	L8 NDMI
EWMA SWIR 1 - Intercept	L8 Tasseled Cap - Greenness
EWMA SWIR 1 - Sine	EWMA NDVI - Intercept
EWMA SWIR 2 - Intercept	EWMA SWIR 1 - Intercept
Mean Monthly Precipitation - March	EWMA SWIR 2 - Sine
Mean Monthly Precipitation - April	Mean Monthly Precipitation - February
Mean Monthly Precipitation - October	Mean Monthly Precipitation - March
Mean Monthly Precipitation - December	Mean Monthly Precipitation - October
Elevation	Mean Monthly Precipitation - December
N/A	Mean Annual Precipitation
N/A	Elevation

Table 8.—Model performance metrics calculated on the out-of-bag data from the random forest tree canopy cover models. Root mean square error (RMSE) and percent variance explained are shown here for the years 2011 and 2016 for the OCONUS study areas.

Study Area	RMSE	% Variance Explained
2011 Coastal Alaska	17.8	69.8
2016 Coastal Alaska	18.9	70.8
2011 Hawaii	23.2	68.3

Study Area	RMSE	% Variance Explained
2016 Hawaii	21.5	72.1
2011 Puerto Rico & US Virgin Islands	20.6	66.1
2016 Puerto Rico & US Virgin Islands	20.5	66.8

All draft TCC maps were carefully inspected using the best high-resolution imagery available for each study area. Because the OCONUS study areas were mapped individually as three separate mapping units, OCONUS mapping units did not have seamline issues as was observed in CONUS. For CONUS, there could be adjacent WRS-2 path/rows mapped using different reference and modeling datasets creating artificial seamlines and other anomalies. The main issue for the OCONUS study areas were areas of missing data caused by persistent cloud cover. These missing data were left as-is, resulting in data gaps over very small portions of the area in the OCONUS TCC maps (table 9).

Table 9.—Percent of the land area that were gaps due to persistent cloud cover.

Study Area	% (AND KM^2) of area that were gaps
2011 Coastal Alaska	0.53 (1.3 MILLION KM^2)
2016 Coastal Alaska	0 (0 KM^2)
2011 Hawaii	0.04 (9,000 KM^2)
2016 Hawaii	0.01 (1,000 KM^2)
2011 Puerto Rico & US Virgin Islands	0.49 (46,000 KM^2)
2016 Puerto Rico & US Virgin Islands	0.03 (2,000 KM^2)

Cartographic Maps

The random forest models were applied to all pixels within the study areas, resulting in TCC values in places where trees do not exist (such as impervious surfaces and water). To address this issue and improve the visual appearance of the TCC maps for cartographic purposes, thresholds and masks were applied to the analytical maps to produce the cartographic maps. An overall description of the thresholding process is below.

No reference dataset is perfect, and plots labeled zero percent canopy cover – when there is actually tree cover at the time of the imagery – are particularly problematic when a random forest model is applied to all pixels. False zeroes and false TCC values greater than zero are possible. While plots located on roads, water bodies, or barren areas are easier to interpret as zero percent TCC and have less error, certain conditions can be confusing to correctly label, depending on the condition of the NAIP imagery available and the condition of the vegetation present, such as areas with sparse tree vegetation, regenerating stands with seedlings and saplings, standing dead trees, and areas where shrub lifeforms are present (Frescino & Moisen, 2012). Even among interpreters, variability can be high (Stehman et al., 2022).

To understand the uncertainty and error in the zero percent TCC values predicted by the random forest model, the reference data were bootstrapped 100,000 times with 200 random samples withheld with each iteration. The associated random forest model was applied to the withheld samples. T-tests were performed on the observed zero TCC reference data and the associated predictions from the bootstrapped TCC values (t-value = (abs(observed – predicted))/standard error). High t-values indicate either a large difference between observed zero TCC values and the predicted TCC values, or a relatively small standard error as compared to the observed and predicted differences. Usually, the t-value at the 95th percentile was selected as the threshold to adjust the standard errors. If a predicted TCC value was less than the adjusted standard error (tcc – (threshold*se) < 0), then the predicted TCC value was forced to zero under the assumption that the predicted non-zero TCC value was an error. Once the analytical map had both (1) the predicted non-zero TCC values that were assumed to be in error subsequently forced to zero percent through the thresholding process, and (2) masks applied, the output was the cartographic map. The masks are described in more detail by geography in the Cartographic Maps section.

The final cartographic maps for all geographic extents were single-layer integer 8-bit TCC images with 255 set as the background value.

Conterminous U.S. Cartographic Maps

For CONUS, the reference data were divided into USGS mapping zones as described in the Analytical Modeling and Maps section. Threshold values as described above were calculated for each mapping zone. Eighty-one percent of the 436 WRS-2 path/rows belonged to more than one mapping zone and, thus, had more than one threshold. The threshold chosen for a WRS-2 path/row belonged to the predominant intersecting mapping zone. Threshold values ranged from 0.48 to 2.57 for TCC 2011 and 0.50 to 1.88 for 2016 TCC. Thresholded TCC maps were examined to determine whether non-treed areas were being included or treed areas were being excluded. Problems with the thresholds were corrected by choosing more appropriate thresholds from other intersecting mapping zones.

Agriculture masks were developed from the Cropland Data Layer (CDL) datasets from the years 2010, 2011, 2014, 2015, and 2016 (USDA National Agricultural Statistics Service Cropland Data Layer, 2010, 2011, 2014, 2015, 2016). For 2011, areas were labeled as zero TCC if they were identified as non-tree agriculture in either the 2010 or 2011 CDL dataset. For 2016, areas were labeled as zero TCC if they were identified as a non-tree agriculture class in the 2016 or in both the 2014 and 2015 CDL datasets.

A water and snow mask was derived from the 2011 NLCD land cover product’s open water and perennial ice and snow land cover types (Homer et al., 2015). The 2011 NLCD emergent herbaceous wetland class was also used as a mask in southern Florida where there was confusion between vegetation and water.

Outside Conterminous U.S. (Alaska, Hawaii, Puerto Rico and U.S. Virgin Islands) Cartographic Maps

The OCONUS areas were not sub-divided into mapping zones. Each study area was mapped as a single unit. Thus, there was only one t-value threshold for each study area. To provide optional thresholds, thresholds were generated for t-values at the 93rd, 95th, and 97th percentiles (table 10). An image was produced for each of these thresholds. Each of these images were compared visually by analysts to high-resolution imagery. The image that visually represented the landscape the best was selected, and the threshold used for the selected image was used to mask non-zero TCC values that were likely in error from the analytical TCC maps, to create the cartographic TCC maps.

The 2011 NLCD open water and perennial ice/snow classes (Homer et al., 2015) was used to mask coastal Alaska. The 2011 NLCD open water class (Homer et al., 2015) was used to mask Hawaii and PRUSVI.

Table 10.—Threshold values used to mask non-tree areas from the analytical tree canopy cover images for 2011 and 2016 for the outside the conterminous US (OCONUS) study areas. Values in bold were the final threshold values selected to create the non-tree masks.

Study Area	93% Threshold	95% Threshold	97% Threshold
2011 Coastal Alaska	1.04	1.16	1.41
2016 Coastal Alaska	0.97	1.14	1.47
2011 Hawaii	1.19	1.50	2.03
2016 Hawaii	1.26	1.64	2.64
2011 Puerto Rico & US Virgin Islands	1.49	1.52	1.59
2016 Puerto Rico & US Virgin Islands	1.60	1.67	1.75

NLCD TCC and Change Maps

A request from NLCD user communities and a requirement for the NLCD TCC maps was to have cohesive maps in which the TCC maps and associated change map add up (that is, 2011 TCC + change in TCC = 2016 TCC).

Preparation for producing the NLCD maps included creation of an average TCC data layer and generation of an initial change/no-change data layer. An average TCC data layer was created by averaging the 2011 and 2016 cartographic maps. An initial change/no-change data layer was created by subtracting the cartographic 2016 TCC values from the cartographic 2011 TCC values. Pixels with absolute differences of less than 10 percent were identified as no change. All other pixels were initially assigned as unknown with respect to change or no-change.

The change map is intended to represent a similar amount of TCC loss as obtained from FIA disturbance data. FIA disturbance data were obtained using the EVALIDator tool (Forest Inventory and Analysis National Program, 2019). The forest disturbance data were aggregated to estimates for CONUS, coastal Alaska, Hawaii, and PRUSVI (table 11). From the initial change/no-change data layer, 900,000, 20,000, 2,000, and 1,100 changed pixels were randomly selected for CONUS, coastal Alaska, Hawaii, and PRUSVI, respectively. The TCC values and associated standard errors of the selected changed pixels were extracted from the 2011 and 2016 analytical maps. This information served as inputs to the Excel numerical optimization tool (Klugman et al., 2008), which returned coefficients that would achieve the amount of change as constrained by FIA disturbed area estimates (table 12). For example, using the selected change/no-change pixels, if a 2016 TCC value plus its standard error multiplied by the coefficient was less than the 2011 TCC value subtracted by its standard error multiplied by the coefficient [i.e., (2016 TCC + coefficient * 2016 SE) < (2011 TCC – coefficient * 2011 SE)], the pixel was recorded as having experienced loss between the nominal years of 2011 and 2016 (figure 8, left panel). The Excel numerical optimization tool iteratively determined the coefficients, stopping when the computed amount of change across the study area was closest to the optimization target (FIA disturbed area estimate for the study area).

Table 11.—Percentage of the study area estimated as disturbed between the nominal years of 2011 and 2016, based on Forest Inventory and Analysis disturbance data obtained using EVALIDator tool.

Study Area	Estimate of Disturbed Forest Lands
CONUS	12.41%
Coastal Alaska	3.99%
Hawaii	14.00%
PRUSVI	10.25%

Table 12.—2011 and 2016 multipliers calculated from forest change disturbance estimates.

Project Area	2011 Multiplier Value	2016 Multiplier Value
CONUS	0.731	0.573
Coastal Alaska	1.492	0.078
Hawaii	0.062	1.115
PRUSVI	0.323	0.830

To create the final binary change/no-change data layer, the coefficients and the 2011 and 2016 standard errors from the analytical maps were multiplied, creating adjusted standard errors (ASE). If a TCC value ± ASE for one year fell outside of the other year's TCC ± ASE range and no overlap of the ranges occurred, pixels were recorded as changed, either gain or loss (figure 8, left and right panels). All other pixels were labeled as no-change (figure 8, center panel). The resulting binary change/no-change data layer was filtered to eliminate clumps of change that were less than five pixels (approximately 1 acre) in size. For more details on the spatial filtering of the change map, see Appendix D.

The final map set for the NLCD 2016 product suite included an updated NLCD 2011 TCC map, a NLCD 2016 TCC map, and a change map for each study area. To create a change map, all pixels labeled as change in the binary change/no-change image had the cartographic 2016 TCC value subtracted from cartographic 2011 TCC value. All binary change/no-change pixels labeled as no-change were set to zero in the change map. Figure 9 shows an example of the change map overlaid on a 2017 NAIP image in southern Arkansas, with red areas indicating loss of TCC and green areas indicating gain of TCC between the nominal years of 2011 and 2016.

To produce the NLCD 2011 TCC map, if a pixel was labeled as change in the binary change/no-change image, the cartographic 2011 TCC value was used. If the pixel was labeled as no-change, the average TCC value was used. The same logic was used to create the NLCD 2016 TCC map, except the cartographic

2016 TCC values were used for changed pixels. The final package for the NLCD 2016 product suite included NLCD TCC 2011 map, NLCD TCC 2016 map, the change map, and accompanying metadata.

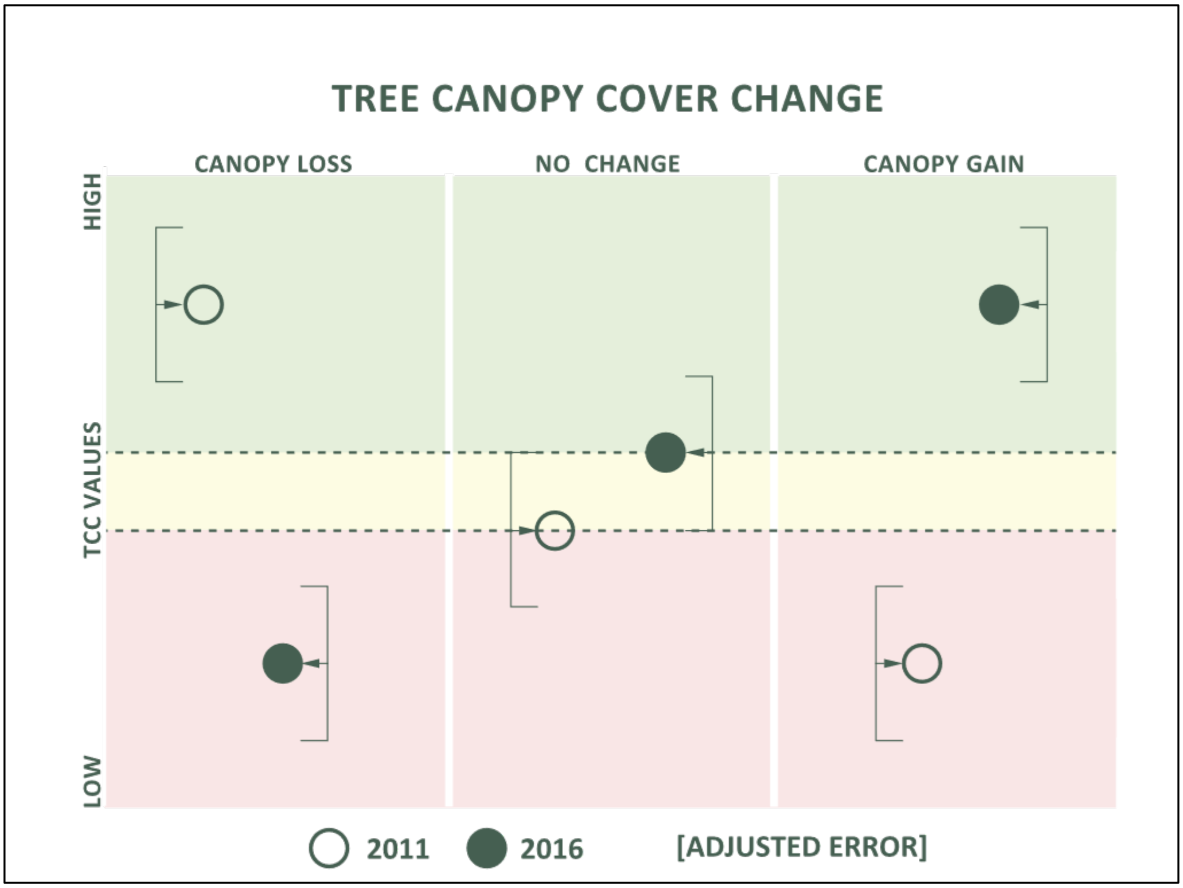

Figure 8.—If there is no overlap between coefficient-adjusted standard error bars in the canopy loss and canopy gain scenarios, pixels are classified as 'Change'. The coefficient-adjusted standard error bars in the no change scenario overlap so pixels are classified as 'No Change'.

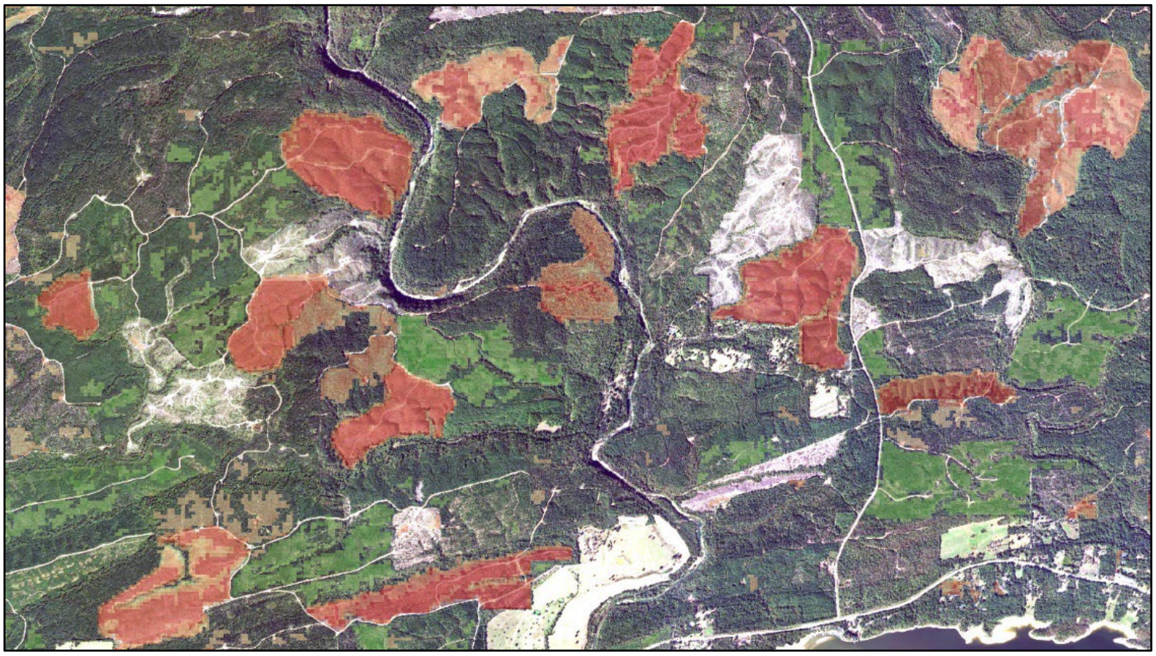

Figure 9.—Example of change map overlaid on a 2017 NAIP image in southern Arkansas, with red areas indicating loss of tree canopy cover and green areas indicating gain of tree canopy cover between 2011 and 2016.

Summary

Three versions of maps for two time periods (2011 and 2016) were produced for the 2016 TCC product suite (figure 10).

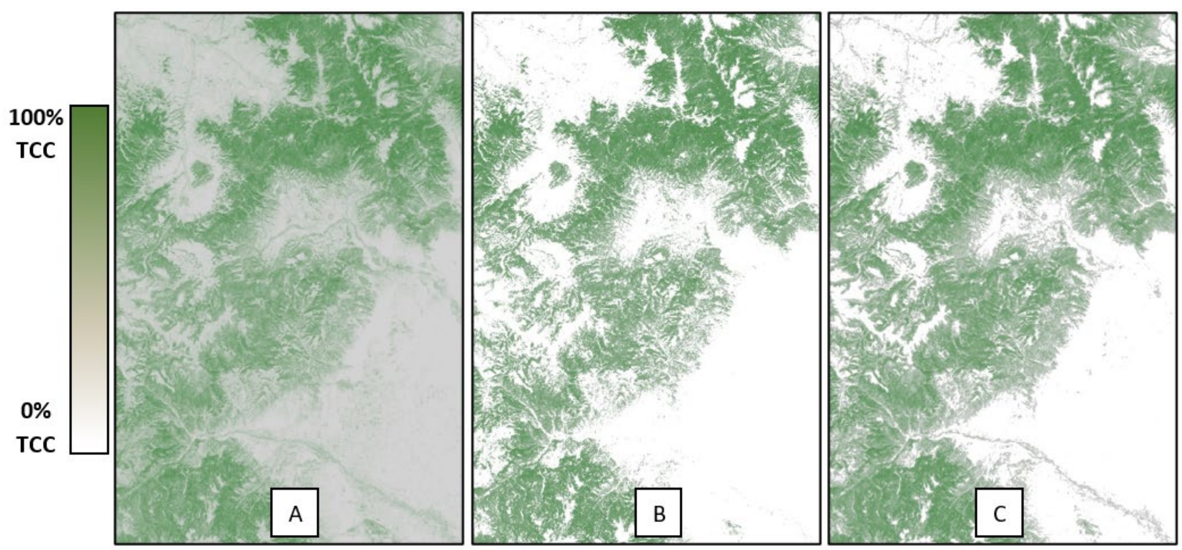

Figure 10.—An example of the three Tree Canopy Cover (TCC) map versions from Colorado is shown. The Analytical map is on the left (A), the Cartographic map is in the middle (B), and the National Land Cover Database (NLCD) map is on the right (C).

The analytical map, which is the initial model predictions and outputs, is best suited for users who will carry out their own detailed statistical and uncertainty analyses and place lower priority on visual appearance. The analytical map for each time period consists of two components: TCC and model standard error. The standard error represents the model uncertainty associated with the corresponding pixel in the TCC map. The analytical maps for both 2011 and 2016 are available as downloads from the USFS TCC web site or an image service at:

- USFS Tree Canopy Cover Maps and Datasets (Download)
- USFS Enterprise Data Warehouse (Image Service)

The cartographic maps, which are masked versions of the analytical maps, are best suited for users who prioritize visual appearance of the maps for cartographic and illustrative purposes. A cartographic map is comprised of a single TCC layer for each time period (no standard error). The cartographic maps for both 2011 and 2016 are available as downloads from the USFS TCC web site or an image service at:

- USFS Tree Canopy Cover Maps and Datasets (Download)
- USFS Enterprise Data Warehouse (Map Service)

The NLCD maps are the result of further processing of the analytical and cartographic maps. The NLCD version is the only version out of the three that identifies canopy cover change between the 2011 and 2016 maps. NLCD maps are best suited for users who require a coordinated three-component data stack where each pixel's values meet the criterion of '2011 TCC + change in TCC = 2016 TCC' (figure 11). The NLCD TCC map set is comprised of four components: 2011 TCC map, 2016 TCC map, and two change maps between the years. The change maps are signed 8-bit and unsigned 8-bit. The signed 8-bit layer has values ranging from -100 to 100 with negative values representing canopy loss, positive values representing canopy gain, and zero values representing no change. The unsigned 8-bit layer has values ranging from 0 to 200. Increases in tree canopy cover are represented by 1 to 100 values, decreases are represented by values 101 through 200, and zero values represent no changed. The NLCD maps are available as a download from the MRLC consortium or an image service at:

- Multi-Resolution Land Characteristics (MRLC) Consortium (Download)
- USFS Enterprise Data Warehouse (Map Service)

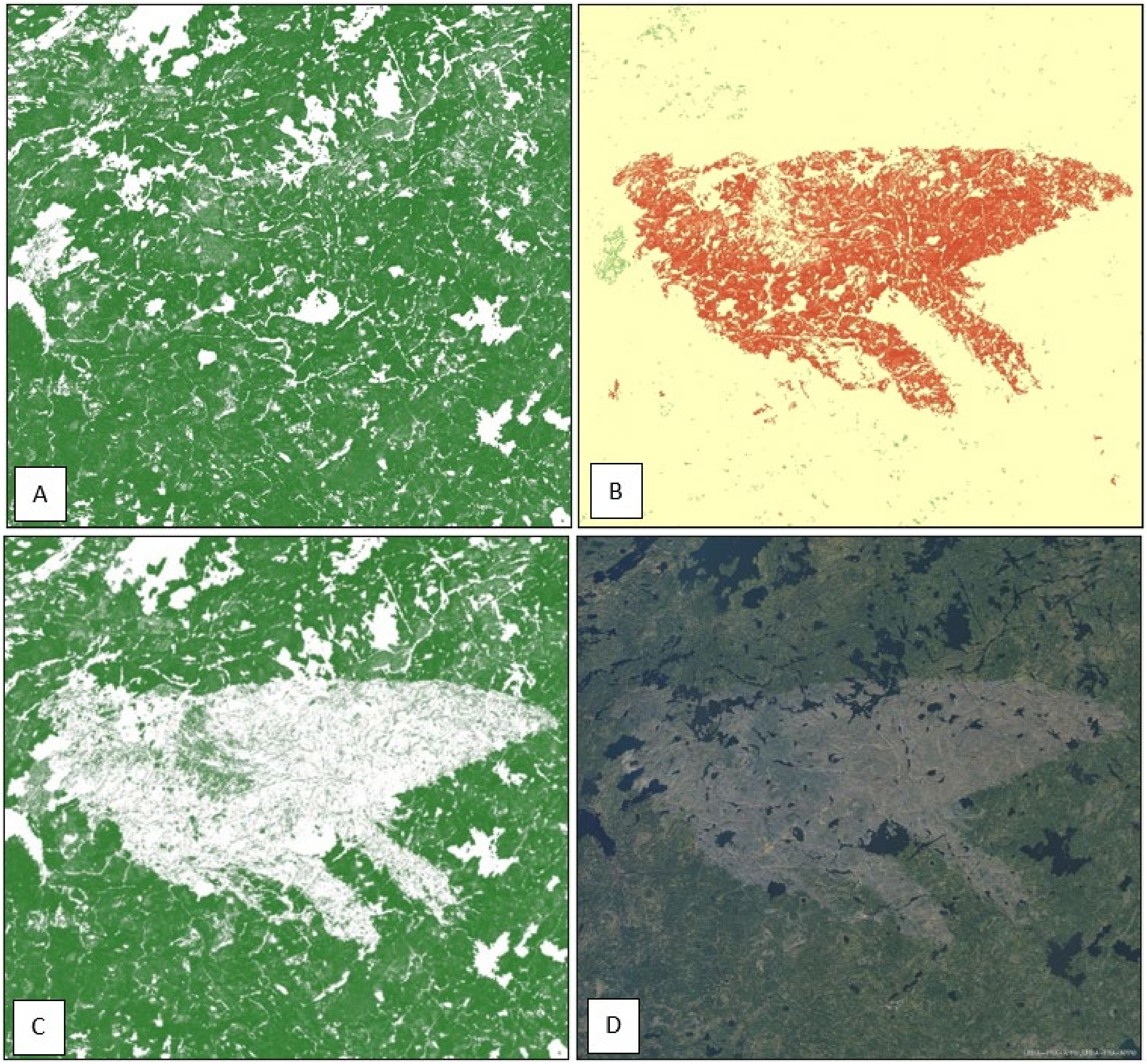

Figure 11.—Example of the National Landcover Database (NLCD) maps with associated change layer. The figure in the top left (A) shows the 2011 NLCD map, the figure in the top right (B) shows the change layer (wildfire event) with red areas indicating loss of tree canopy cover and green areas indicating gain of tree canopy cover between 2011 and 2016, the figure in the lower left (C) shows the 2016 NLCD map, and the figure in the lower right (D) shows the 2015 NAIP imagery.

For more information about the TCC project, please visit the USDA Forest Service Tree Canopy Cover web site at https://data.fs.usda.gov/geodata/rastergateway/treecanopycover/ or send an email to SM.FS.TCC@usda.gov.

References

Alaska Geospatial Council. (n.d.). *State of Alaska Open Data Geoportal*. https://gis.data.alaska.gov/

Baig, M.H.A., Zhang, L., Shuai, T., & Tong, Q. (2014). Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. *Remote Sensing Letters*, 5(5), 423-431. https://doi.org/10.1080/2150704X.2014.915434

Bellante, G.; Boucher, T.; Charnon, B.; Goetz, W.; Homan, K.; Pan, C.; Pugh, N.; Megown, K; Schrader, B.; Schulz, B.; Wittwer, D. (2020). Kenai Peninsula Existing Vegetation Map Project. (GTAC-10209-RPT1). U.S. Department of Agriculture, Forest Service, Geospatial Technology and Applications Center, Salt Lake City, UT. 39 p.

Berryman, E., & McMahan, A. (2019). Chapter 7 – Using tree canopy cover data to help estimate acres of damage. In: K.M. Potter, B.L. Conkling (Eds.), Forest health monitoring: national status, trends, and analysis 2018 (pp. 125-141). (General Technical Report SRS-239) U.S. Department of Agriculture, Forest Service, Southern Research Station, Knoxville, TN. https://www.srs.fs.usda.gov/pubs/chap/chap_2019_berryman_001.pdf

Brandeis, T.J., & Oswalt, S.N. (2007). *The status of U.S. Virgin Islands' forests, 2004* (Resource Bulletin SRS-112). U.S. Department of Agriculture, Forest Service, Southern Research Station, Knoxville, TN. https://www.srs.fs.usda.gov/pubs/rb/rb_srs122.pdf

Brandeis, T.J.; & Turner, J.A. (2013). *Puerto Rico's forests, 2009* (Resource Bulletin SRS-191). U.S. Department of Agriculture, Forest Service, Southern Research Station, Knoxville, TN. https://foresthistory.org/wp-content/uploads/2017/01/Puerto-Ricos-Forests-2009.pdf

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brooks, E.B., Thomas, V.A., Wynne, R.H., & Coulston, J.W. (2012). Fitting the multitemporal curve: a fourier series approach to the missing data problem in remote sensing analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 50(9), 3340-3353. https://doi.org/10.1109/TGRS.2012.2183137

Cahoon, S.M.P., Kuegler, O., & Christensen, G.A. (Tech Eds.) (2020). *Coastal Alaska's forest resources, 2004–2013: Ten-year Forest Inventory and Analysis report* (General Technical Report PNW-GTR-979). U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR. https://www.fs.fed.us/pnw/pubs/pnw_gtr979.pdf

Chander, G., Markham, B.L., & Helder, D.L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5), 893-903. https://doi.org/10.1016/j.rse.2009.01.007

Chavez, P.S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24(3), 459-479. https://doi.org/10.1016/0034-4257(88)90019-3

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 Engineering and Remote Sensing*, 78(7), 715–727. https://doi.org/10.14358/PERS.78.7.715.

Crist, E.P., & Cicone, R.C. (1984). A physically-based transformation of thematic mapper data—The TM tasseled cap. *IEEE Transactions on Geoscience and Remote Sensing*, GE-22(3), 256-263. https://doi.org/10.1109/TGRS.1984.350619.

Daly, C., Taylor, G.H., Gibson, W.P., Parzybok, T.W., Johnson, G.L., & Pasteris, P.A. (2000). High-quality spatial climate data sets for the United States and beyond. *Transactions of the American Society of Agricultural Engineers*, 43(6), 1957-1962. https://prism.oregonstate.edu/documents/pubs/2001asae_spatialClimate_daly.pdf

Dennison, P.E., Brewer, S.C., Arnold, J.D., & Moritz, M.A. (2014). Large wildfire trends in the western United States, 1984–2011. *Geophysical Research Letters*, 41(8), 2928-2933. https://doi.org/10.1002/2014GL059576

Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Center for Coastal Monitoring and Assessment, & Biogeography Branch. (2007). *Digital Elevation Models (DEMs) for the main 8 Hawaiian Islands.* NOAA's National Ocean Service, National Centers for Coastal Ocean Science.

Eidenshink, J., Schwind, B., Brewer, C., Zhu, Z., Quayle, B., & Howard, S. (2007). A project for monitoring trends in burn severity. *Fire Ecology*, 3(1), 3-21. https://doi.org/10.4996/fireecology.0301003

ERDAS. (2013). ERDAS IMAGINE 2013. Hexagon Geospatial. Madison, Alabama.

EROS Science Processing Architecture – Land Satellites Research and Development. (n.d.). *EROS Science Processing Architecture On Demand Interface - Land Satellites Data System Science Research and Development*. http://espa.cr.usgs.gov/index

Ewel, J.J., & Whitmore, J.L. (1973). *The ecological life zones of Puerto Rico and the U.S. Virgin Islands* (Research Paper ITF-018). U.S. Department of Agriculture, Forest Service, Institute of Tropical Forestry, Rio Piedras, Puerto Rico. https://www.srs.fs.usda.gov/pubs/rp/rp_itf018.pdf

Farm Service Agency. n.d. *National Agriculture Imagery Program*. U.S. Department of Agriculture, Farm Service Agency. https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/

Filippeli, S.K., Vogeler, J.C., Falkowski, M.J., & Meneguzzo, D.M. (2020). Monitoring conifer cover: leaf-off lidar and image-based tracking of eastern redcedar encroachment in central Nebraska. *Remote Sensing of Environment,* 248(2020), 1-15. https://doi.org/10.1016/j.rse.2020.111961

Forest Inventory and Analysis National Program. (2019). *EVALIDator User Guide (Version 1.8.0.01)*. U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis.

Frescino, T.S., & Moisen, G.G. (2012). Comparing alternative tree canopy cover estimates derived from digital aerial photography and field-based measurements. In: W. McWilliams, F.A. Roesch (Eds.), *Monitoring across borders: 2010 joint meeting of the Forest Inventory and Analysis (FIA) symposium and the southern mensurationists (pp. 237 – 244).* (General Technical Report SRS-157). U.S. Department of Agriculture, Forest Service, Southern Research Station. Ashville, NC. https://www.srs.fs.usda.gov/pubs/41019

Gao, B. (1996). NDWI – a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment,* 58(3), 257-266.

Genuer, R., Poggi, J.M., & Tuleau-Malot, C. (2015). VSURF: an R package for variable selection using random forests. *The R Journal*, 7(2), 19-33. https://journal.r-project.org/archive/2015/RJ-2015-018/RJ-2015-018.pdf

Gesch, D.B., Evans, G.A., Oimoen, M.J., & Arundel, S. (2018). The National Map Seamless Digital Elevation Model Specifications. In D.F. Maune & A. Nayegandhi (Eds.), *Digital Elevation Model Technologies and Applications: The DEM Users Manual* (3rd ed., pp. 83-100). American Society for Photogrammetry and Remote Sensing.

Goeking, S.A., Liknes, G.C., Lindblom, E., Chase, J., Jacobs, D.M., & Benton, R. (2012). A GIS-based tool for estimating tree canopy cover on fixed-radius plots using high-resolution aerial imagery. In: R. Morin, S. Randall, G.C. Liknes (Comps.), *Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012 December 4-6, Baltimore MD (pp. 237-241)*. (General Technical Report NRS-P-105). U.S. Department of Agriculture, Forest Service, Northern Research Station. Newtown Square, PA. https://www.fs.fed.us/nrs/pubs/gtr/gtr_nrs-p-105.pdf

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202(2017), 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Goward, S., Arvidson, T., Williams, D., Faundeen, J., Irons, J., & Franks, S. (2006). Historical record of landsat global coverage: mission operations, NSLRSDA, and international cooperator stations. *Photogrammetric Engineering and Remote Sensing,* 72(10), 1155-1169. https://doi.org/10.14358/PERS.72.10.1155

Hauff, R., Wilkinson, M., & LaRosa, A.M. (2007). *Forest Health Highlights: Forest Resource Summary*. https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/fsbdev3_045553.pdf

Heris, M., Bagstad, K.J., Rhodes, C., Troy, A., Middel, A., Hopkins, K.G., & Matuszak, J. (2021). Piloting urban ecosystem accounting for the United States. *Ecosystem Services,* 48(2021), 1-18. https://doi.org/10.1016/j.ecoser.2020.101226

Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., & Wickham, J. (2007). Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogrammetry Engineering & Remote Sensing*, 73(4), 337- 341.

Homer, C., & Gallant, A. (2001). *Partitioning the conterminous United States into mapping zones for Landsat TM land cover mapping* [White paper]. US Geological Survey Report.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., & 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 Engineering and Remote Sensing*, 81(5), 345–354. https://www.ingentaconnect.com/content/asprs/pers/2015/00000081/00000005/art00002#

Homer, C.G., Huang, C., Yang, L., Wylie, B.K., & Coan, M.J. (2004). Development of a 2001 National Land Cover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, 70(7), 829–840. https://doi.org/10.14358/PERS.70.7.829

Huang, C., Yang, L., Wylie, B.K., & Homer, C. (2001). A strategy for estimating tree canopy density using Landsat 7 ETM+ and high resolution images over large areas. *Proceedings of the Third International Conference on Geospatial Information in Agriculture and Forestry November 5-7, 2001.* Denver, Colorado. https://digitalcommons.unl.edu/usgspubs/111/

Intergovernmental Panel on Climate Change. (n.d.). *Intergovernmental Panel on Climate Change.* United Nations. https://www.ipcc.ch

International Institute of Tropical Forestry. (n.d.) Climate data of Puerto Rico and U.S. Virgin Islands. Retrieved from Dr. Eileen Helmer.

Kauth, R.J., & Thomas, G.S. (1976) The tasseled cap - a graphic description of the spectral-temporal development of agricultural crops as seen by LANDSAT. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, IN*, 4B41-4B51. https://docs.lib.purdue.edu/lars_symp/159/

Klugman, S.A., Panjer, H.H., & Willmot, G.E. (2008). Loss Models: From Data to Decisions, (3rd ed.). John Wiley & Sons, Inc. https://onlinelibrary.wiley.com/doi/book/10.1002/9780470391341

Kriegler, F.J., Malila, W.A., Nalepka, R.F., & Richardson, W. (1969). Preprocessing transformations and their effects on multispectral recognition. *Proceedings of the Sixth International Symposium on Remote Sensing of Environment*, 97-131.

LaRosa, A.M., Tunison, J.T., Ainsworth, A., Kauffman, J.B., & Hughes, R.F. (2008). Fire and nonnative invasive plants in the Hawaiian Islands bioregion. In: K. Zouhar, J.K. Smith, S. Sutherland, M.L. Brooks (Eds.), *Wildland fire in ecosystems: fire and nonnative invasive plants* (pp. 225-242). (General Technical Report RMRS-GTR-42-vol. 6.) U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Ogden, UT.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2002(2), 18-22.

Little, E.L., & Skolmen, R.G. (1989). *Common forest trees of Hawaii (native and introduced)* (Agriculture Handbook No. 679.) U.S. Department of Agriculture, Forest Service, Washington, D.C.

Lugo, A.E., Medina, E., Trejo Torres, J.C., & Helmer, E. (2006). Botanical and ecological basis for the resilience of Antillean dry forests. In: R.T. Pennington, G.P. Lewis, J.A. Ratter (Eds.), *Neotropical*

Savannas and Seasonally Dry Forests (pp. 359-381). Taylor & Francis, Boca Raton, FL https://www.fs.usda.gov/treesearch/pubs/50777

Maxar. (n.d.). *Maxar – Global Enhanced GEOINT Delivery*. https://rdog.digitalglobe.com

McRoberts, R.E., Liknes, G.C., & Domke, G.M. (2014). Using a remote sensing-based, percent tree cover map to enhance forest inventory estimation. *Forest Ecology and Management,* 331, 12-18. https://doi.org/10.1016/j.foreco.2014.07.025

Nelson, K., Connot, J., Peterson, B., & Martin, C. (2013). LANDFIRE Refresh strategy: updating the national dataset. *Fire Ecology*, 9(2), 80-101. https://doi.org/10.4996/fireecology.0902080

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Reams, G.A., Smith, W.D., Hansen, M.H., Bechtold, W.A., Roesch, F.A., & Moisen, G.G. (2005). *The forest inventory and analysis sampling frame* (General Technical Report SRS-80 pp. 21-36). U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC. https://www.srs.fs.usda.gov/pubs/gtr/gtr_srs080/gtr_srs080-reams001.pdf

Rittger, K., Raleigh, M.S., Dozier, J. Hill, A.F., Lutz, J.A., & Painter, T.H. (2020). Canopy adjustment and improved cloud detection for remotely sensed snow cover mapping. *Water Resources Research,* 56(6), 1-20. https://doi.org/10.1029/2019WR024914

Ruefenacht, B. (2016). Comparison of three Landsat TM compositing methods: a case study using modeled tree canopy cover. *Photogrammetric Engineering & Remote Sensing*, 82(3), 199-211. https://doi.org/10.14358/PERS.82.3.199

Schmidt, G., Jenkerson, C.B., Masek, J., Vermote, E., & Gao, F. (2013). *Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description* (Report No. 2013-1057). U.S. Geological Survey. https://pubs.er.usgs.gov/publication/ofr20131057

Statewide Digital Mapping Initiative. (n.d.). *Alaska Statewide Digital Elevation Model Efforts.* http://www.alaskamapped.org/dem

Stehman, S.V., Mousoupetros, J., McRoberts, R.E., Næsset, E., Pengra, B.W., Xing, D. & Horton, J.A. (2022). Incorporating interpreter variability into estimation of the total variance of land cover area estimates under simple random sampling. *Remote Sensing of Environment*, 269(2022), 112806. https://doi.org/10.1016/j.rse.2021.112806

Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D.B., Oimoen, M.J., Zhang, Z., Danielson, J.J., Krieger, T., Curtis, B., Haase, J., Abrams, M., & Carabajal, C. (2011). *ASTER Global Digital Elevation Model Version 2 - summary of validation results*. National Aeronautics and Space Administration. http://pubs.er.usgs.gov/publication/70005960

Taylor, L.A., Eakins, B.W., Carignan, K.S., Warnken, R.R., Sazonova, T., & Schoolcraft, D.C. (2008). *Digital Elevation Models of Puerto Rico: Procedures, Data Sources and Analysis* (Report: NOAA Technical Memorandum NESDIS NGDC-13). NOAA National Centers for Environmental Information.

United States Department of Agriculture Forest Service. (2017). *Forest Inventory and Analysis Fiscal Year 2016 Business Report.* (Forest Service Publication No. FS-1075). U.S. Department of Agriculture, Forest Service. Washington Office. https://www.fs.usda.gov/sites/default/files/15817-usda-forest-service-fia-annual-report-508-update.pdf

United States Geological Survey Earth Resources Observation and Science Center. (2001). USGS Global Visualization Viewer (GloVis). https://glovis.usgs.gov/

USDA National Agricultural Statistics Service Cropland Data Layer. (2010). Published crop-specific data layer [Online]. Available at https://nassgeodata.gmu.edu/CropScape USDA-NASS, Washington, DC.

USDA National Agricultural Statistics Service Cropland Data Layer. (2011). Published crop-specific data layer [Online]. Available at https://nassgeodata.gmu.edu/CropScape USDA-NASS, Washington, DC.

USDA National Agricultural Statistics Service Cropland Data Layer. (2014). Published crop-specific data layer [Online]. Available at https://nassgeodata.gmu.edu/CropScape USDA-NASS, Washington, DC.

USDA National Agricultural Statistics Service Cropland Data Layer. (2015). Published crop-specific data layer [Online]. Available at https://nassgeodata.gmu.edu/CropScape USDA-NASS, Washington, DC.

USDA National Agricultural Statistics Service Cropland Data Layer. (2016). Published crop-specific data layer [Online]. Available at https://nassgeodata.gmu.edu/CropScape USDA-NASS, Washington, DC.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S.M., Case, A., Costello, C., Dewitz, J., Fry, J., Funk, M., Granneman, B., Liknes, G.C., Rigge, M., & Xian, G. (2018). A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146(2018), 108-123. https://doi.org/10.1016/j.isprsjprs.2018.09.006

Zhu, J., Miller, A.E., Lindsay, C., Broderson, D., Heinrichs, T., & Martyn, P. (2013). *MODIS NDVI products and metrics manual* (version 1.0. 2013). Geographic Information Network of Alaska, University of Alaska Fairbanks.

Zhu, Z., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118(2012), 83-94. https://doi.org/10.1016/j.rse.2011.10.028

Appendix A—Surface Reflectance versus Top of Atmosphere

The median composites used as predictor data for the development of 2011 TCC and 2016 TCC not only were derived from different satellite sensors (L5 and L8), but the satellite imagery was also processed differently. L5 imagery was processed to surface reflectance (SR). L8 imagery was processed to TOA. An investigation was conducted on three mapping zones (figure 5) examining how these processing differences influenced the mapping of TCC. The results in tables A1, A2, and A3 show very similar results for using SR for 2011 TCC and TOA for 2016 TCC.

Table A1.—Table comparing model performance metrics for top of atmosphere and surface reflectance for zone 28 for 2011.

	Zone	RMSE	% Var Explained
Top of Atmosphere (2011)	28	17.3	55.7
Surface Reflectance (2011)	28	17.4	55.2

Table A2.—Table comparing model performance metrics for top of atmosphere and surface reflectance for zones 2 and 55 for 2016.

	Zone	RMSE	% Var Explained
Top of Atmosphere (2016)	2	23.6	51.9
Surface Reflectance (2016)	2	23.9	50.7

Table A3.—Table comparing model performance metrics for top of atmosphere and surface reflectance for zones 2 and 55 for 2016.

	Zone	RMSE	% Var Explained
Top of Atmosphere (2016)	55	22.4	67.7
Surface Reflectance (2016)	55	22.6	67

Appendix B—Landsat Image Viewer Tool

The Landsat Image Viewer (figure B1) serves as a front end to GloVis. It was designed to enhance productivity of Landsat scene selection. The Landsat Image Viewer interface allows users to query and filter Landsat datasets for scene selection. The filters include date, WRS-2 path/row, sensor, maximum cloud cover, maximum distance from peak NDVI, and land cover types obtained from NLCD.

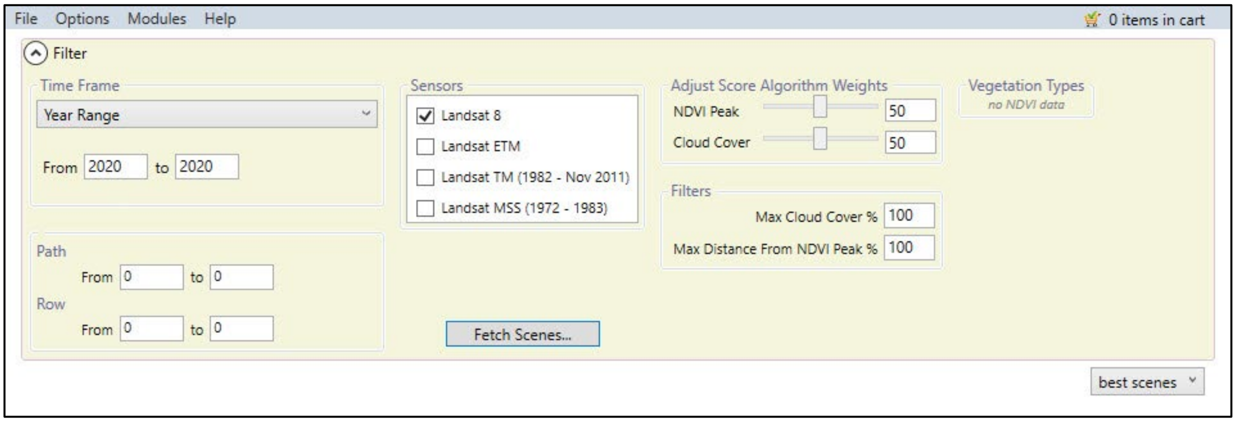

Figure B1.—The user interface for the Landsat Image Viewer.

The Time Frame pulldown had three options: Year Range, Date Range, Time Series. Year Range allows users to select single years to multiple years. Date Range allows users to set specific start and end dates within a single year. Time Series allows users to set specific start and end dates across multiple, consecutive years.

WRS-2 paths/rows are entered into the Path and Row boxes. If only a single path and row needs to be queried, enter the values in both From and to boxes. For multiple paths and rows, enter the range in the corresponding From and to boxes.

The type of Landsat datasets to be queried are selected by checking one or more of the boxes under the Sensors heading. If there is a conflict between the Time Frame entries and the Sensors selected (for example, 2009 and L8) nothing will be returned when the query is run.

Max Cloud Cover filters scenes based on the CLOUD_COVER variable contained within the metadata for each Landsat scene.

The Max Distance from NDVI Peak is calculated using NDVI curves obtained from GloVis. AVHRR 1 km data is used by GloVis to generate the NDVI graphs. Users can restrict the Landsat scene acquisition to dates where the NDVI values are within a user-defined percentage of the maximum NDVI value.

The Fetch Scenes button retrieves the scenes. Figure B2 shows an example of the interface when the Fetch button is clicked. The images displayed are organized according to the filter specifications from best to worst. The preferred image candidate set is displayed in the top viewer. Clicking on one of these images will show alternative images in the bottom viewer.

The alternative images can be sorted by cloud score or date. Either a single year or multiple years can be displayed in the bottom viewer. Metadata is displayed above each image and includes the cloud cover scores, scene quality scores, Julian dates, and full dates. Below each image there are three icons: a checkbox for selecting the scene, a magnifier to enlarge the image, and a graph icon which will bring up the NDVI curve for the path/row.

Clicking on the Fetch Scenes button will display the Vegetation Types box. Users can display the NDVI curves for land cover types found within the WRS-2 path/row. An example of an NDVI curve is shown in figure B3.

Once all scenes to be ordered are checked, click on the shopping cart icon. Clicking on the cart indicator opens the shopping cart dialog box (figure B4). The dialog box shows the scenes selected and provides options to download the full resolution scenes, generate a list of scene IDs for ordering later, or export the thumbnails to an alternate location. Selecting the option to download full resolution scenes prompts the user for their USGS GloVis/Earth Explorer login credentials. Once logged in to the USGS site the scene IDs are transferred to the USGS interface, and the user interacts directly with GloVis/EarthExplorer.

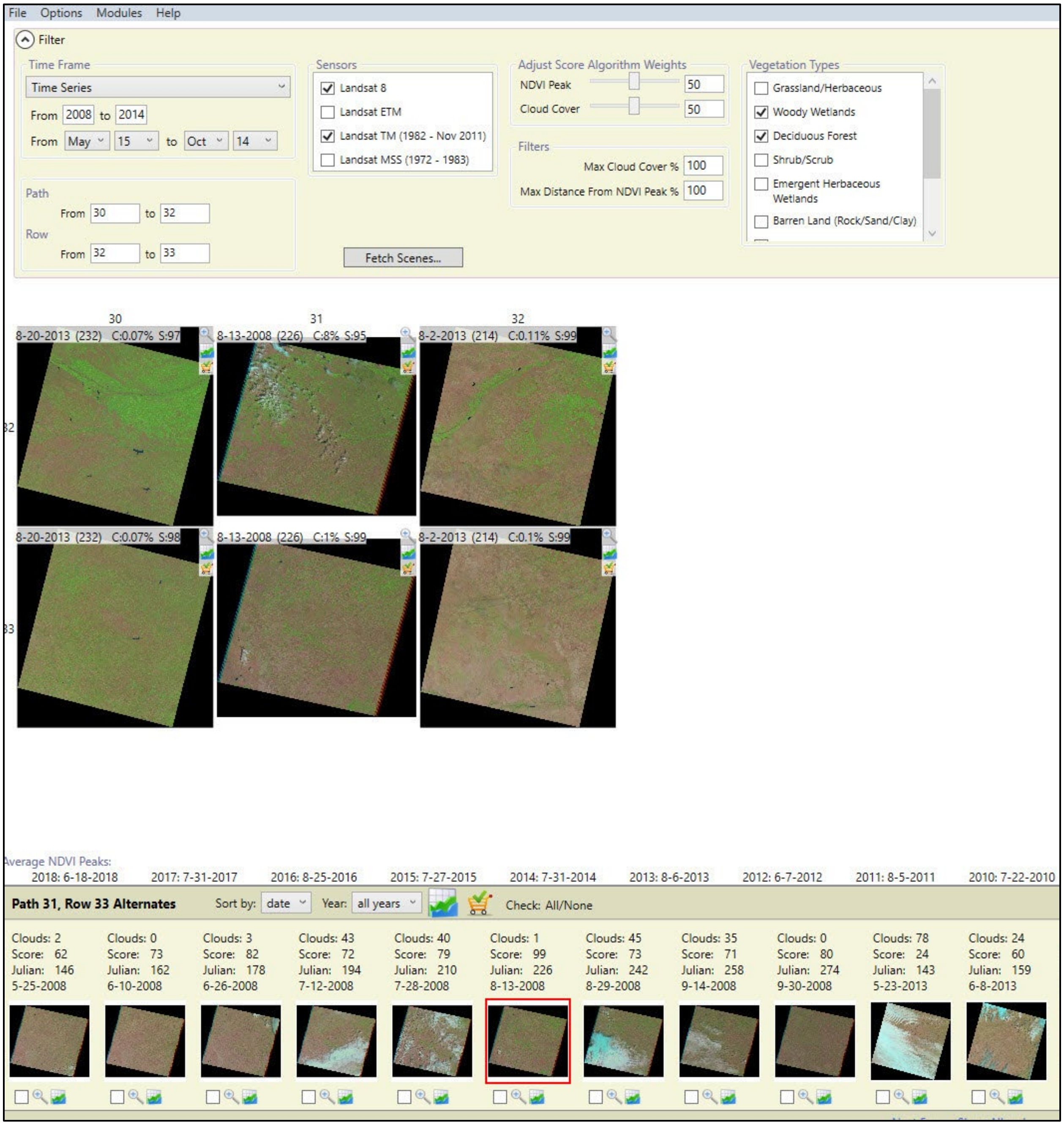

Figure B2.—Results of a Landsat Image Viewer time series query for paths 30 to 32 and rows 32 to 33 for both L5 and L8. This graphic shows a cropped view of the interface to better display the results.

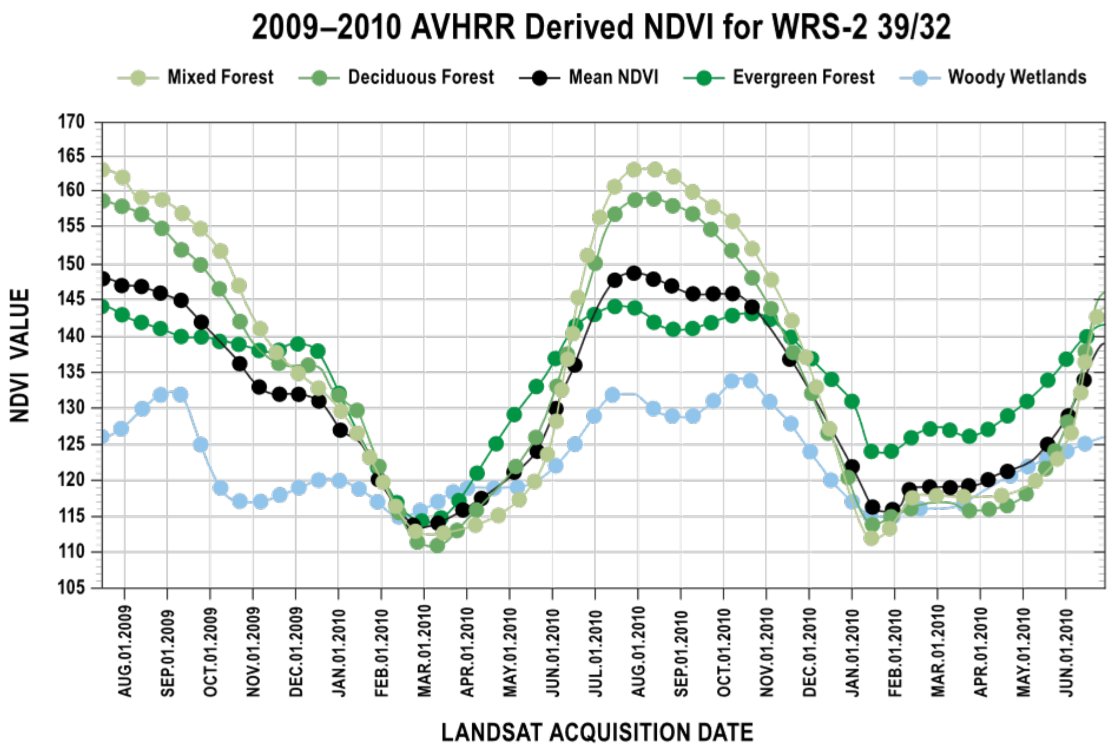

Figure B3.—NDVI graphs for NLCD land cover types for WRS-2 39/32.

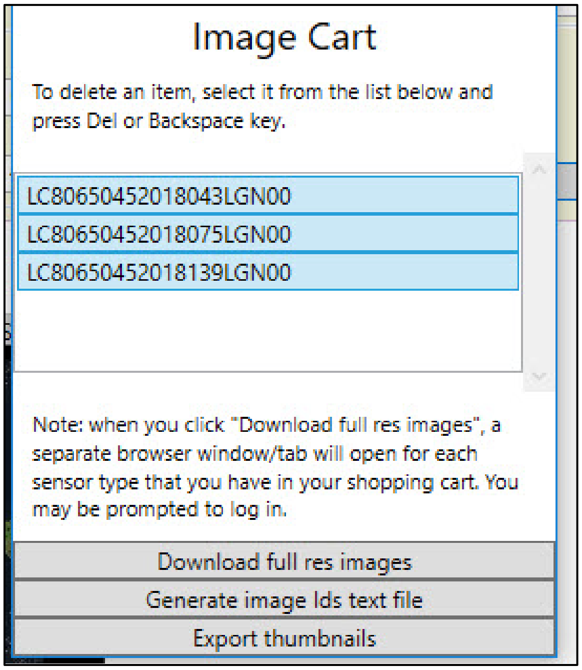

Figure B4.—Shopping cart dialog box for the Landsat Image Viewer tool.

Appendix C—List of Acronyms

ASE: Adjusted Standard Error

ASTER GDEM 2: Advanced Spaceborne Thermal Emission and Reflection radiometer Global Digital Elevation Model 2

AVHRR: Advanced Very High Resolution Radiometer

BDL: Best Data Layer

CDL: Cropland Data Layer

CONUS: Conterminous United States (lower 48 states)

DEM: Digital Elevation Model

DOS: Dark Object Subtraction

ERDAS: Earth Resources Data Analysis System

EROS: (United States Geological Survey) Earth Resources Observation and Science Center

EWMA: Exponentially Weighted Moving Average

FIA: Forest Inventory and Analysis program (within USDA Forest Service Research and Development)

FMASK: Function of Mask

FS: Forest Service

GEE: Google Earth Engine

GINA: Geographic Information Network of Alaska

GRS1980: Geodetic Reference System 1980

GTAC: Geospatial Technology and Applications Center

L5: Landsat 5 TM

L8: Landsat 8 Operational Land Imager

LEDAPS: Landsat Ecosystem Disturbance Adaptive Processing System

MRLC: Multi-Resolution Land Characteristics Consortium

MTBS: Monitoring Trends in Burn Severity

NAD83: North American Datum 1983

NAIP: National Agriculture Imagery Program

NASS: USDA National Agricultural Statistics Service

NDMI: Normalized Difference Moisture Index

NDVI: Normalized Difference Vegetation Index

NLCD: National Land Cover Database

NOAA: National Oceanic and Atmospheric Administration

O-CONUS or OCONUS: Outside CONUS

PRISM: Parameter-elevation Regressions on Independent Slopes Model

PRUSVI: Puerto Rico and United States Virgin Islands

RMSE: Root Mean Square Error

SDMI: Statewide Digital Mapping Initiative

SLC: Scan Line Corrector

SR: Surface reflectance

SWIR1: Short-Wave Infrared 1

SWIR2: Short-Wave Infrared 2

TCC: Tree Canopy Cover

TOA: Top Of Atmosphere

US: United States

USDA: United States Department of Agriculture

USFS: United States Forest Service

USGS: United States Geological Survey

VSURF: Variable Selection Using Random Forest

WRS-2: Worldwide Reference System 2

Appendix D—Masking and Minimum Mapping Unit

There were several masking procedures used to produce the final TCC maps. The details about each one of these procedures can be found in the body of this report, but for clarification purposes, each procedure is briefly documented below. The procedures are listed in the order that they were performed.

Masking used to produce the cartographic maps

The steps listed below describe the masking process used to produce the cartographic maps (see the Cartographic Maps section).

1. T-values were used to select thresholds to mask TCC values from the analytical maps where confidence in the TCC values was very low. If TCC values were less than zero when subtracted from the thresholds multiplied by the standard errors, the TCC values were forced to zero. Otherwise, the TCC values stayed the same.
2. TCC values occurring in water, perennial ice, or perennial snow land cover classes from NLCD 2011 were forced to zero.
3. For TCC values occurring in Florida, the emergent herbaceous wetland land cover class from NLCD 2011 was used to force TCC values to zero if they intersected this class.

4. TCC values occurring in non-tree cultivated areas as determined by the NASS CDL (National Agriculture Statistics Service Cropland Data Layers) program were forced to zero.

Masking used to produce NLCD change maps

The steps listed below describe the masking process used to produce the NLCD change maps (see the NLCD TCC and Change Maps section).

1. The TCC and standard error values from the TCC analytical maps were used to create a binary change/no-change mask, which means this change mask covered all pixels within the study areas.
2. Clumps of no-change or change that were less than five pixels in size were reclassed to the opposite class, to reduce noise in the NLCD maps.
3. All pixels labeled as change in the filtered binary change/no-change mask (output from step 2), had the cartographic 2016 TCC values subtracted from the cartographic 2011 TCC values.
4. Note: Theoretically, a pixel might be labeled as no-change in step 1, while its adjacent neighbors are all labeled as change (for instance, a small clump of trees that is left behind when creating a subdivision). The filter routine in step 2 will force this pixel to receive a change classification. However, the cartographic 2011 TCC value might be the same as the cartographic 2016 TCC value. When these values are subtracted (step 3) the result will be zero, which is also the value used to indicate no-change. This results in the pixel of no-change surrounded by pixels of change to reappear. Thus, even though small clumps were eliminated in step 3, they can reappear in the 2011 and 2016 NLCD maps, if the pixel had the same TCC value in 2011 and 2016.