Comparing the USDA Forest Service Tree Canopy Cover Datasets 2011, 2016, 2019 and 2021 with Photointerpretation of High-Resolution Google Earth Imagery: Guidance for Use in in Community Scale Forest Management

November 01, 2023

Eric J. Greenfield and Karen Schleeweis, USDA Forest Service Forest Inventory and Analysis

Introduction

Sound and comprehensive forest management requires consistent measurements over time to determine progress toward desired outcomes in community forest management plans. One general measure important to forestry is percent tree canopy cover of an area of interest (i.e., parcel, park, institution, local government, watershed). The percent tree canopy cover of an area in one year can be compared to the same metric at a later or earlier time to determine one aspect of progress in a management plan. i-Tree (https://www.itreetools.org/) is a free and publicly available set of software tools for assessing and managing community trees and forests developed by the USDA Forest Service, Davey Tree and other partners, and the tools use percent tree canopy cover to model the environmental benefits of trees and forests including carbon storage and sequestration, pollution removal, and stormwater impacts. Percent tree canopy cover's value as a reliable and consistent metric in forest management is dependent on its accuracy and precision. That dependency additionally has an impact on the accuracy and precision of the derived i-Tree benefit values of carbon storage and sequestration, pollution removal, and stormwater impacts that rely on measurements of percent tree canopy cover.

While field sampling and photointerpretation provide more detailed methods of measuring tree canopy cover, wall-to-wall mapped land cover datasets furnish a quick and inexpensive set of benchmarks for forest management. High-resolution land cover (1-meter or less of resolution) is available for many locations and is becoming more common, but it lacks consistent spatial and temporal coverage across the United States. A publicly available land cover dataset that does have spatial and temporal consistency throughout the United States is the National Land Cover Database (NLCD), provided by a consortium of agencies of the federal government within the Multi-Resolution Land Characteristics Consortium (MRLC), which hosts its data at https://www.mrlc.gov/. In addition to the United States Geological Service (USGS) produced Anderson-classed land cover (LC) and percent impervious surface cover, the NLCD provides percent tree canopy cover (TCC) through a partnership with the USDA Forest Service (USFS). There are two different TCC data sets available from these sources, a filtered and masked cartographic layer designed for visualization covering each of the years 2011 to 2021 and the more detailed and recently released (April 2023) 2021.4 USFS Science Tree Canopy Cover data (USFS TCC) available at annual time steps between 2008-2021 (https://data.fs.usda.gov/geodata/rastergateway/treecanopycover/). The USFS TCC datasets were developed with a spatially consistent and temporally coherent methodology, enabling comparisons through time. The NLCD, including the USFS TCC, is produced with a 30-meter resolution, a scale appropriate for strategic extent small-scale landscapes, but a spatial resolution that is often too coarse for the precision necessary in large-scale heterogenous urban and community landscapes.

This brief paper provides guidance for using the most recent USFS science TCC data in local community-scale applications. It illustrates a comparison between the percent tree canopy cover for six cities of the United States generated from estimations of TCC based on the USFS science maps versus estimations of TCC based on photointerpretation of high-resolution Google Earth imagery from the same vintage. Please note and consider the results of this and previous analyses (some referenced below) as you evaluate the use of USFS TCC in forest management plans or other assessments.

Background

This work follows and builds upon previous evaluations of 30-meter resolution NLCD TCC products relative to addressing community-scale data needs for tree canopy cover measures. In Greenfield and Nowak (2009), estimates of TCC using zonal aggregations of NLCD percent tree canopy cover maps generally underestimated tree canopy relative to design based TCC estimation using finer scale field measurements and photointerpretation. Those findings led to a national-scale analysis of the NLCD 2001 TCC product, where we found a national average underestimation of 9.7% TCC from zonal statistics (pixel counting) compared with point-based estimation from photointerpretation (Nowak and Greenfield 2010). Based on these results, many of our subsequent analyses were conducted using design-based samples and response data from photointerpretation to derive more accurate results from finer-scale measurements (e.g., Nowak and Greenfield 2012, Roman et al 2017, and Nowak and Greenfield 2020). These design-based estimates also allow reporting of confidence intervals and uncertainty in the estimates based on the standard error of proportion, critical information for decision-making. In addition, we have continued to evaluate raster TCC datasets periodically to evaluate their utility relative to community-scale measurements (Riitters et al 2023).

Estimating the value of TCC over a geographic area can be done in many ways. Overall, the goal is to make strong inferences about the value of a population, in this case mean TCC over a given area. The ever-expanding dominance of mapped data in forestry and environmental resource applications, makes useful the re-examination of the underlying assumptions and needs for validity associated with different estimation approaches. Since the goal of this analysis is to inform users of the difference or similarity of the source of TCC data, we have limited the estimation paradigm and methods. Previous assessments compared both different sources of TCC used in different paradigms of estimation, a design-based estimate of TCC (or statistical sample using PI TCC values) to a model-assisted estimate of TCC (or summarized zonal map statistics using the NLCD TCC values) (Nowak and Greenfield 2010). Each estimation method has its strengths and limitations, and for a variable like tree canopy cover, which is adequately related to remote sensing variables (from analyst interpreted as well as machine interpreted images), different estimation techniques can yield meaningful results. It is also worth noting that field-based sample surveys do remain uniquely important for several forest parameters that do not adequately relate to remotely sensed data.

Methods

This analysis included data collection of a unique TCC data source over six U.S. cities and a comparison for each city at multiple points in time of TCC estimated from these unique data and from the newly released USFS TCC maps.

The sample-design used for estimation was based on a random point design with 1,000 samples per city and followed procedures available elsewhere (Nowak and Greenfield 2020). Because of limited resources, we employed three staff members to conduct photo interpretation of the 6,000 total samples (1,000 from each city) from June to November 2022. We selected the six cities based on the relative proportion of the United States land area within forest, grassland, and desert (Nature Conservancy 2018), with three cities within forested areas (Syracuse, NY, Charlotte, NC, and Seattle, WA), two within grassland areas (Des Moines, IA, and St. Louis, MO) and a single one within desert areas (Boise, ID). We used 2020 census geography (US Census 2022) to delineate the six city boundaries. The photointerpretation land cover classes used included bad imagery (not interpretable), tree canopy cover, impervious land cover, water, and all other land cover. We interpreted land cover class values using Google Earth and its historical imagery for the years 2011, 2016, 2019 and 2021; and if imagery for those specific years was unavailable, the closest earlier year was selected (e.g., if no 2011, 2010 was used). Additionally, another photo interpreter audited each set of the 1000 city points using a random sample of 10% (100) of those interpreted points. If agreement was over 90%, the set of points was accepted and used for the analysis. All six data sets were deemed acceptable after audit.

For estimation using the USFS TCC data, we downloaded the most recent science dataset (version 2021.4) from (https://data.fs.usda.gov/geodata/rastergateway/treecanopycover/) where previous releases of the USFS TCC and documentation are also archived. While the science version of the USFS TCC data do include a per pixel standard error, users should note that this is a per pixel standard error of model performance, not the type of sampling error or uncertainty needed to construct a confidence interval or make inferences about the uncertainty of the TCC value itself. Comparison of estimates generated with other NLCD TCC, such as the cartographic version, or other estimation methods, were out of the scope of this analysis.

We generated one estimate and its standard error of the proportion from the common i-Tree method for evaluations of TCC – an estimate of TCC over the given area using a design-based sample and TCC data from photointerpretation of high-resolution (one meter or less) Google Earth imagery, and we used zonal statistics on the 30-meter USFS TCC for the same area at or near the same point in time for the second estimate. However, and perhaps most importantly, pixel counting (zonal statistics) does not yield a measure of the uncertainty or confidence around the estimated value. Furthermore, due to the strength and convincing nature of visual symbols, the user may assume that the specific spatial patterns in the map share the same "accuracy" as the more general mean and are more than a spatially modeled representation of the truth. Meanwhile statistical design-based estimates, or inference from a sample, do not overpromise on spatial precision, do deliver quantifiable measures of uncertainty around the estimated variable, but do not provide a convincing wall-to-wall image.

Results and Discussion

In the forested cities of Charlotte, Seattle, and Syracuse the difference between the two measures of percent tree canopy cover ranged from a high of 10.94 percentage points (PI higher than USFS TCC) in Charlotte 2016 to 1.95 percent in Syracuse 2021 (see Table 1).

The difference between the PI and USFS TCC in the grassland cities of Des Moines and St. Louis ranged from a high in Des Moines of 10.97 percent in 2016 to a low in St. Louis of 4.70 percent in 2016.

In the desert city of Boise, USFS TCC underestimated percent tree canopy cover compared to PI and ranged from a high of 14.53 in 2016 to a low of 13.22 in 2021.

Consistent with previous assessments of PI versus 30-meter resolution raster NLCD TCC comparisons, the USFS TCC underestimates tree canopy cover relative to PI methods for these six cities.

This analysis compared estimates of tree canopy cover by varying only the input source of the estimate. Model-assisted inference or, calculating zonal statistics from geospatial data or maps, is a common practice. There is a growing body of best practices related to map accuracy and interpretation that suggest pixel counting might have limited utility in accurately representing the characteristics of a landscape (e.g., Olofsson et al 2014 and Olofsson et al 2021). A notable limitation of this practice is that it does not yield uncertainty or confidence intervals with which to interpret the validity of the estimate of mean TCC for the given area. Further, the infrequency of offering users measures of uncertainty accompanying the mapped data may help propagate the misconception, due to the strength and convincing nature of visual symbols, that the specific spatial patterns of mapped TCC share the same "accuracy" as the more general mean TCC for the given area, and worse, that mapped TCC do not have an associated uncertainty. Mapped values of TCC have variance and bias (together described as uncertainty), but it is often not quantifiable, though efforts are being made in this direction (McRoberts et al 2022). Still, the spatial patterns evident in maps of TCC have many uses, which are not served by statistical design-based estimates, or inference from a sample. Perhaps a middle road between these two approaches in the future is model-assisted design-based inference or, more simply, hybrid estimation (Stahl et al 2016). In this approach, wall-to-wall spatial maps can be used to assist statistical estimates, avoiding the potential pitfalls of pixel counting, allowing for inference about populations using only a part, or sample of the whole, and allowing for the efficient calculation and communication of uncertainty.

Table 1. Comparison between PI (with reported standard error) and USFS TCC estimates of percent canopy cover for 2011, 2016, 2019, and 2021 (all values in percent).

		Forested Cities			Grassland Cities		Desert City
		Charlotte, NC	Seattle, WA	Syracuse, NY	Des Moines, IA	St. Louis, MO	Boise, ID
2011	TCC	39.33%	15.31%	28.36%	23.65%	16.53%	10.49%
	PI	46.80%	17.30%	30.60%	30.70%	21.60%	24.00%
	SE	1.58%	1.20%	1.46%	1.46%	1.30%	1.35%
	Delta	7.47%	1.99%	2.24%	7.05%	5.07%	13.51%
2016	TCC	39.26%	14.93%	27.30%	22.83%	15.20%	9.97%
	PI	50.20%	18.20%	30.80%	33.80%	19.90%	24.50%
	SE	1.58%	1.22%	1.46%	1.50%	1.26%	1.36%
	Delta	10.94%	3.27%	3.50%	10.97%	4.70%	14.53%
2019	TCC	38.95%	14.83%	26.96%	22.53%	14.66%	10.20%
	PI	49.30%	18.30%	29.50%	33.40%	20.10%	24.70%
	SE	1.58%	1.22%	1.44%	1.49%	1.27%	1.36%
	Delta	10.35%	3.47%	2.54%	10.87%	5.44%	14.50%
2021	TCC	39.15%	14.97%	27.25%	22.48%	14.61%	10.38%
	PI	48.40%	18.60%	29.20%	33.00%	19.40%	23.60%
	SE	1.58%	1.23%	1.44%	1.49%	1.25%	1.34%
	Delta	9.25%	3.63%	1.95%	10.52%	4.79%	13.22%

The goal in this analysis was to focus on comparisons of TCC generated from high resolution data sources (1m or smaller imagery) vs moderate resolution data sources (30m USFS TCC) over six US cities through time. A secondary focus was to highlight some points in the discourse on how, as a community, we use maps, generate estimates, and judge the validity of those values. We offer this approach to estimation using the USFS TCC maps to show a path away from simple pixel counting and towards more dialogue on accuracy (error) and precision (bias).

Conclusion

While processes and methods may have improved to provide consistent percent tree canopy cover data throughout the United States in the USFS TCC layers, the coarse resolution of those datasets does not allow for the more detailed, accurate tree canopy cover percent values needed for appropriate evaluation of large-scale heterogenous urban and community landscapes. To obtain a more accurate estimate, we recommend finer-scale data gathering using photointerpretation or field sampling as outlined in the i-Tree tools protocols. Using methods that quantify the standard error or variability of the estimate helps interpret the validity of the estimates and adds additional information that can be important for resource management and monitoring.

Acknowledgements

This work was funded by a Resource Planning Act grant. I wish to thank State University of New York College of Environmental Science and Forestry students Bennett Caldwell and Nicholas Timoshenko for photointerpretation, and thanks to NRS FIA colleagues Daniel Crane and Alexis Ellis for their review, comments, and edits on this manuscript.

References

Greenfield, Eric; Nowak, David J.; Walton, Jeffrey T. 2009. Assessment of 2001 NLCD percent tree and impervious cover estimates. *Photogrammetric Engineering and Remote Sensing*. 75(11): 1279-1286.

McRoberts, R.E. Næsset, E., Saatchi, S. and Quegan, S. 2022. Statistically rigorous, model-based inferences from maps. *Remote Sensing of Environment*. 279: 113-28.

Nature Conservancy, 2018. tnc_terr_ecoregions. Retrieved July 2018 from http://maps. tnc.org/files/metadata/TerrEcos.xml

Nowak, David J.; Greenfield, Eric J. 2010. Evaluating the national land cover database tree canopy and impervious cover estimates across the conterminous United States: a comparison with photo-interpreted estimates. *Environmental Management*. 46(3): 378-390.

Nowak, David J.; Greenfield, Eric J. 2012. Tree and impervious cover in the United States. *Landscape and Urban Planning*. 107: 21-30.

Nowak, David J.; Greenfield, Eric J. 2020. The increase of impervious cover and decrease of tree cover within urban areas globally (2012-2017). Urban Forestry & Urban Greening

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

Olofsson, P., 2021, July. Updates to Good Practices for Estimating Area and Assessing Accuracy of Land Cover and Land Cover Change Products. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS* (pp. 1982-1985). IEEE.

Riitters, Kurt; Coulston, John W.; Mihiar, Christopher; Brooks, Evan B.; Greenfield, Eric J.; Nelson, Mark D.; Domke, Grant M.; Mockrin, Miranda; Lewis, David J.; Nowak, David J. 2023. Chapter 4: Land Resources. In: U.S. Department of Agriculture, Forest Service. 2023. Future of America's Forest and Rangelands: Forest Service 2020 Resources Planning Act Assessment. General Technical Report WO-102. Washington, DC.

Roman, Lara A.; Fristensky, Jason P.; Eisenman, Theodore S.; Greenfield, Eric J.; Lundgren, Robert E.; Cerwinka, Chloe E.; Hewitt, David A.; Welsh, Caitlin C. 2017. Growing Canopy on a College Campus: Understanding Urban Forest Change through Archival Records and Aerial Photography. *Environmental Management*

Ståhl, G.; Saarela, S.; Schnell, S; Holm, S.; Breidenbach, J.; Healey, S.P.; Patterson, P.L.; Magnussen, S.; Naesset, E.; McRoberts, R.E.; and Gregoire, T.G. 2016. Use of models in large-area forest surveys: comparing model-assisted, model-based and hybrid estimation. *Forest Ecosystems*. https://doi.org/10.1186/s40663-016-0064-9

United States Census Bureau, 2022. TIGER/Line Shapefiles. Retrieved May 2022 from https://www2.census.gov/geo/tiger/TIGER2022/

United States Forest Service Tree Canopy Cover Datasets, 2023. Retrieved May 2023 from https://data.fs.usda.gov/geodata/rastergateway/treecanopycover/