Urban Ecosyst (2008) 11:81–90
DOI 10.1007/s11252-007-0040-9

Difficulties with estimating city-wide urban forest cover change from national, remotely-sensed tree canopy maps

Jeffrey T. Walton

Published online: 8 January 2008

Abstract Two datasets of percent urban tree canopy cover were compared. The first dataset was based on a 1991 AVHRR forest density map. The second was the US Geological Survey's National Land Cover Database (NLCD) 2001 sub-pixel tree canopy. A comparison of these two tree canopy layers was conducted in 36 census designated places of western New York State. Reference data generated by photo-interpreting the tree cover on 1994 and 2002 orthoimages were compared on a city-wide basis to estimates of the tree cover derived from the AVHRR and NLCD tree canopy maps. Comparison of the AVHRR-based estimate with the 1994 photo-interpreted reference showed low accuracy and high variability. The comparison between the NLCD-derived estimate and 2002 photo-interpreted values had higher accuracy but shows a consistent under prediction bias for all tree cover values. In comparing the two photo-interpreted reference sets, little actual change was detected for this study region. Assessing change using differences between the imagery-based AVHRR and NLCD urban tree cover datasets is shown to lead to erroneous results.

Keywords Urban tree canopy cover · Change detection · Accuracy assessment · NLCD 2001 tree canopy cover · AVHRR

Introduction

An important environmental indicator in urban areas, percent tree canopy cover is a basic element of knowledge needed to understand urban forest structure and to manage the urban forest resource effectively (Dwyer et al. 2000; Nowak 1994). Since percent tree cover in urban areas is often used as a basis for environmental planning and ecological management decisions, knowledge of how it is changing is of critical importance.

Using satellite remote sensing techniques, percent tree cover across the United States has been mapped using early 1990s, 1 km AVHRR data (Zhu 1994) and recently using circa 2001,

J. T. Walton (✉)
USDA Forest Service, Northern Research Station, SUNY-ESF, 5 Moon Library, Syracuse, NY 13210, USA
e-mail: jwalton@paulsmiths.edu

30 m Landsat data (Huang et al. 2001) as part of the National Land Cover Database 2001 (NLCD 2001) developed by the Multi-Resolution Land Characteristics Consortium (MRLC), a group of Federal agencies and cooperators (Homer et al. 2004; Homer et al. 2007).

With the increased availability of remote sensing derived map products and the continual development of new and faster analysis techniques, temporally repeated production of similar products is now becoming a reality. When two maps depict the same phenomenon, there is a natural tendency to compare them to assess change of the property being mapped. If both maps were developed using the same techniques and imagery, the comparison should be straightforward. However, if different techniques and/or imagery were used to map the same parameter, the comparison can be problematic. This study compares two tree canopy cover assessments in selected urban areas of western New York and northern Pennsylvania (Fig. 1) to test whether change can be detected between two urban tree canopy cover databases developed using different imagery and methods.

Methods

The first dataset of urban tree canopy cover was compiled by the United States Department of Agriculture (USDA) Forest Service in 2000 as part of a decadal Renewable Natural Resources Planning Act (RPA) assessment (Dwyer et al. 2000). This dataset was based upon a 1991 nation-wide forest density map (Zhu 1994) where multi-date Advanced Very High-Resolution Radiometer (AVHRR) imagery was used in 15 regional regression models to predict the forest density in each 1.1 km resolution pixel. The regression models were derived using 1 or 2 Landsat TM training sites in each region. The forest density map was summarized (by average pixel value) for all Census Designated Places (CDPs) in the US using 1990 US Census Bureau boundaries and tabulated by Dwyer et al. (2000). For this

Fig. 1 Mapping Zone 63 covers western New York and northern Pennsylvania

study, these values were re-tabulated using year 2000 census boundaries to be consistent with the second dataset.

The AVHRR forest density layer exhibited three major limitations: (a) the regression model regions were large and estimates of forest density were based on only a few training sites; (b) pixels that included medium-to-small water bodies sometimes resulted in high forest density values; and (c) areas with low forest cover such as the Mid-West and South West tended to show no tree cover (Zhu 1994). The summarization of this dataset using the 1990 CDPs added to the limitations. As reported in Dwyer et al. (2000), coastal urban areas tended to be under predicted and urban tree canopy cover in small CDPs may have been calculated from only few pixels thereby causing widely varying results.

The second tree cover dataset used for this comparison was a portion of the NLCD 2001 tree canopy layer developed in cooperation with the USGS EROS Data Center. In developing this tree cover layer, Landsat TM and ETM satellite imagery from 1999, 2000 and 2001 were classified using a supervised regression-based classification methodology (Huang et al. 2001; Yang et al. 2003). Generally, about forty 2,500 m × 2,500 m (approximate size) training sites were selected to represent the full spectrum of tree canopy density in the mapping zone with a concentration in rural, forested areas. High-resolution (approximately 1 m ground sample distance) imagery was classified into treed or non-treed areas for each of the training sites. The high-resolution classified imagery was then resampled to 30 m to match the spatial resolution of the Landsat imagery and to convert the binary classification to percent fraction of tree cover in each pixel. These lower resolution training data were used to generate a regression-based classification model, which outputs percent tree cover for each pixel of the mapping zone using three dates (spring, peak growing season, and late growing season) of Landsat imagery and ancillary data inputs. The resulting estimate was a 30 m ground resolution sub-pixel classification of tree canopy cover. Each pixel has a value from 0 to 100 representing the percent of the area covered by trees.

The NLCD 2001 tree canopy cover has a general accuracy goal of less than 15% mean absolute error (MAE). The approximate accuracy of the NLCD 2001-derived urban tree canopy cover in this zone is 93% (Walton 2005). Since the NLCD 2001 is a new dataset with portions only recently becoming available, its limitations have not been thoroughly investigated. Potential limitations for urban tree cover assessments could be caused by training sites being concentrated in rural areas and training classifications being developed using leaf-off high-resolution imagery in which urban trees are difficult to classify.

To determine AVHRR and NLCD urban tree cover estimates for each place, tree cover pixel values were averaged inside each of the polygonal, year 2000 CDP boundaries. The resulting mean tree cover in each of the CDP areas was the value used for comparison between the two sets of imagery and the photo-interpreted reference data.

The availability of contemporary orthoimagery for developing reference data limited the study area to a subset of MRLC mapping zone 63. Two sets of high-resolution aerial orthoimagery, one from 1994 and the other from 2002, were obtained from New York State through their publicly accessible data clearinghouse (http://www.nysgis.state.ny.us). The 1994 images were standard USGS digital orthoquads (DOQs), were color infrared (CIR), and had 1 m spatial resolution. The 2002 imagery consisted primarily of 1 ft (30 cm) true color images but also included some lower resolution 2 ft (61 cm) CIR or 3 ft (91 cm) panchromatic images. All available imagery had been acquired in the leaf-off season. Since urban tree cover classification comparison was the goal of this analysis, the study area was further reduced to include only urban areas (as defined by year 2000 CDPs) where contemporary images were available (Fig. 2). The potential urban areas ranged in size from large cities like Buffalo, Rochester and Binghamton to numerous small villages scattered

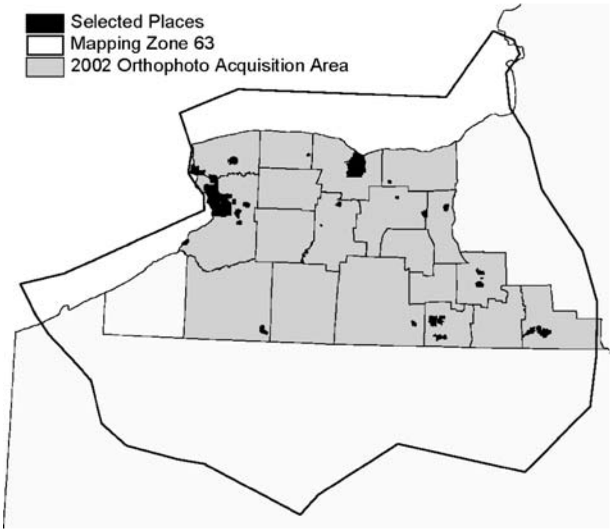

Fig. 2 Places selected for assessment of urban tree cover inside the 2002 New York orthophoto acquisition area

throughout the Allegheny plateau and Finger Lakes region of New York. Of these places, 36 CDPs were selected as the assessment units. The CDPs were ranked by areal size and stratified into three classes: (a) less than 10 km^2, (b) 10–20 km^2, and (c) greater than 20 km^2. Within each class 12 CDPs were randomly selected.

To determine the reference tree cover value, 300 points were randomly placed inside each CDP polygon. At each point both the 1994 and 2002 digital images were interpreted into one of four categories: tree canopy, impervious surface, natural water body, or 'other.' Since the deciduous trees in the images were leafless, the interpreter had to infer whether the point fell on tree canopy based on the branching and shadow pattern present. Percent tree cover was estimated using the proportion of the total number of points that were classified as tree canopy for each year. The sampling standard error was calculated as:

$$\mathrm{SE}_p = \sqrt{\frac{p \cdot (1-p)}{n}} \tag{1}$$

where p is the sample proportion of points that were tree covered and n is the total number of points (300) (Walpole and Myers 1985). From the sampling standard error, 95% confidence intervals (CIs) were established around the photo-interpreted values.

To assess if overall change (growth or loss) of tree cover had happened between 1994 and 2002, the difference between the two photo-interpreted proportions were statistically compared. Standard error for the difference was calculated as:

$$\mathrm{SE}_{p_1 - p_2} = \sqrt{\frac{p_1 \cdot (1-p_1)}{n_1} + \frac{p_2 \cdot (1-p_2)}{n_2}} \tag{2}$$

where subscript 1 and 2 represent the photo-interpreted proportion from 2002 and 1994 (Walpole and Myers 1985). From this standard error, 95% CIs were established around the differences.

With the two photo-interpreted and two satellite imagery-based datasets, four comparisons are possible:

1. AVHRR vs. 1994 photo interpretation (PI1994),
2. NLCD vs. 2002 photo interpretation (PI2002),
3. PI2002 vs. PI1994, and
4. AVHRR vs. NLCD.

The first two comparisons are accuracy assessments of the imagery products comparing how well the imagery classification matches the 'truth' as defined by the photo-interpreted reference data. The third comparison, of the reference datasets, gives an indication of whether change actually has occurred between the years. The fourth is the direct imagery comparison. For each comparison, several analyses were done.

A linear regression was used to test the fit between the two sets of data being compared. The slope of the regression was tested against the value of unity (the slope of the 1:1 correspondence line).

A non-parametric Wilcoxon signed rank test (Mendenhall 1979) was applied to determine if the distribution of one set of values in the comparison is significantly different from the other—in other words, was a bias detected between the two. Three comparative indices, mean absolute error (MAE), root mean square error (RMSE) and an index of agreement, were calculated (Willmott 1984; Walton 2005; Ji and Gallo 2006). While the MAE yields the average error magnitude, the RMSE is an estimator for the standard deviation of the error in predicted tree cover across the region. The index of agreement (d) gives a value between 0 and 1 that indicates the relative agreement between the two sets of data being compared. It is calculated from the potential error variance (PE) using:

$$\mathrm{PE} = \sum_{i=1}^{N} \left(\left| \mathrm{map}_i - \overline{\mathrm{ref}} \right| + \left| \mathrm{ref}_i - \overline{\mathrm{ref}} \right| \right)^2 \tag{3}$$

$$d = 1 - \frac{N \cdot \mathrm{RMSE}^2}{\mathrm{PE}} \tag{4}$$

where map_i are the predicted map values, ref_i are the reference values, and N is the number of places (Willmott 1984).

Earlier analysis of urban tree cover using AVHRR imagery has identified the possibility of trends relating to the size of the CDP and the amount of water inside the place boundary (Dwyer et al. 2000). All four comparisons were tested for error correlation with the size and amount of water using linear regression.

Results

The variation of the AVHRR-derived values is large as shown by the RMSE (10.3%, Table 1). The MAE for the AVHRR comparison with PI 1994 is 8.2%. The index of agreement is 0.81. The individual AVHRR errors are both positive and negative with no demonstrated bias as shown by the inconclusive Wilcoxon signed rank sum test. A regression analysis of this comparison shows a slope (0.66, R^2=0.45, Fig. 3a) significantly different from 1 (p-value=0.0085). Thirteen of the AVHRR-derived tree cover percentages

Table 1 Comparative difference measures for the four urban tree cover comparisons

	Mean Absolute Error (MAE, %)	Root Mean Square Error (RMSE, %)	Index of agreement (*d*)
AVHRR vs. PI 1994	8.2	10.3	0.81
NLCD vs. PI 2002	6.0	7.7	0.93
PI 1994 vs. PI 2002	3.0	4.0	0.98
AVHRR vs. NLCD	9.6	12.6	0.76

MAE and RMSE are percent tree cover.

fall within the 95% confidence intervals of the 1994 photo-interpretation (Fig. 4). There was no strong trend exhibited relating increasing error in the AVHRR prediction to the size of the analysis region. Virtually none of the variation in the error was attributable to the

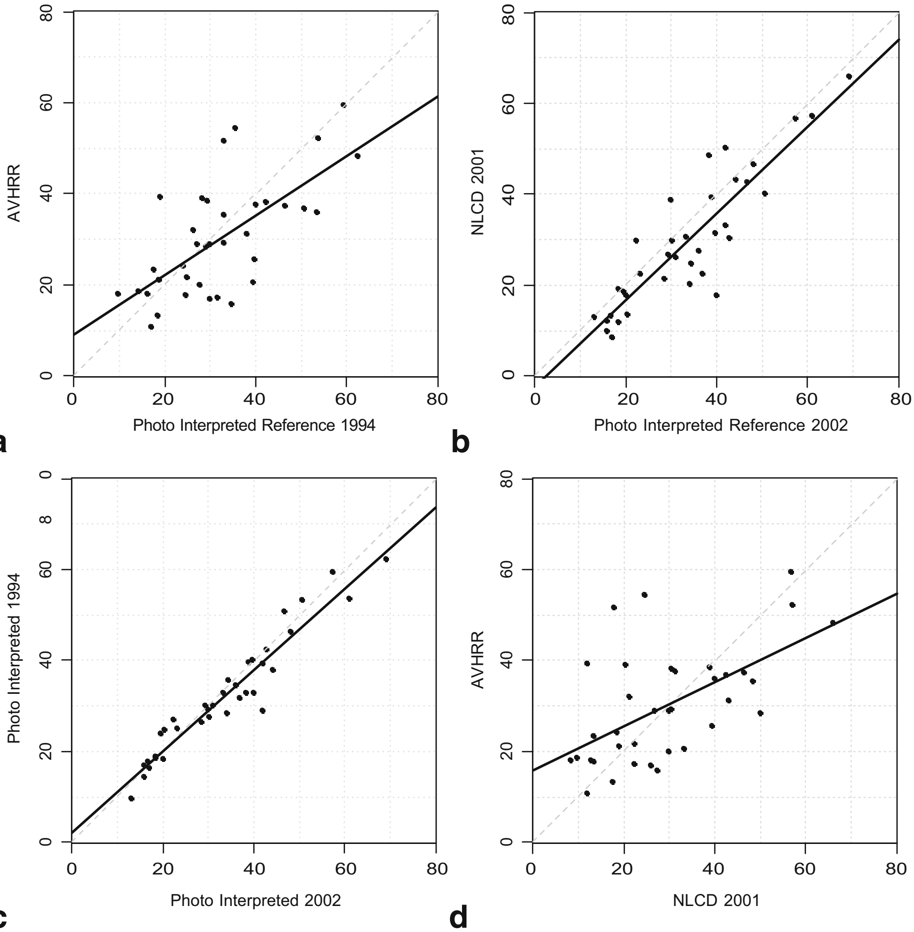

Fig. 3 Regression comparisons for the four cases: **a** slope=0.66, intercept=9.1, R^2=0.45; **b** slope=0.96, intercept=−2.5, R^2=0.80; **c** slope=0.90, intercept=2.2, R^2=0.92; and **d** slope=0.49, intercept=15.8, R^2= 0.33. Units on both axes of all four plots are percent tree cover

place size. Error increases as the amount of water cover in the CDP increases (slope=0.67, R^2=0.17, p-value=0.012).

For the 2002 photo-interpreted and NLCD-derived tree cover values the MAE is 6.0% and the RMSE is 7.7% (Table 1). The index of agreement (d) between the NLCD tree cover value and the reference value is 0.93. Linear regression of predicted versus reference values had a coefficient of determination (R^2) of 0.80 (Fig. 3b). The slope (0.96) and intercept (−2.5) of the fit line were tested against the ideal line (slope=1 and intercept=0) and both are not significantly different from their ideal values (p-value=0.60 and 0.40). Seventeen of the 36 NLCD-derived tree cover values fell within the 95% confidence intervals of the photo interpretation (Fig. 5). Thirty of the 36 places (83.3%) have negative errors and the Wilcoxon signed rank sum test indicates the overall error is negative at a confidence level of 99%. Comparing error among various sized places, the slope of the regression (R^2=0.02) was not significantly different from zero (p-value=0.39). Additionally, NLCD error was compared to the percent water cover of the CDP; a very slight increase in error as the amount of water cover increased (slope=0.03, R^2=0.00093) was shown to be not significantly different from zero (p-value=0.86).

Comparing the photo-interpreted datasets (PI 1994 and PI 2002) using the difference in the two proportions shows that the 95% confidence interval contains zero for all but one CDP (Fig. 6). The MAE for the photo interpretation comparison is 3.0%, with a RMSE of 4.0%, and an index of agreement of 0.98 (Table 1). (The term 'mean absolute difference', MAD, should be used here rather than MAE since the two reference datasets are being compared.) A Wilcoxon signed rank sum test fails to indicate a difference between the two reference datasets. A regression for this comparison (Fig. 3c) shows a good fit with the 1:1 ideal line (R^2=0.92) with the slope (0.90) not significantly different from 1 (p-value=0.025).

Direct comparison of the tree cover estimates from the two imagery products (AVHRR and NLCD) has a MAE of 9.6%, a RMSE of 12.6%, and an index of agreement of 0.76 (Table 1). The Wilcoxon signed rank sum test does not show a difference between the two datasets of this comparison. A regression between the two sets of values yields a slope of 0.49 (R^2=0.33, Fig. 3d) that is significantly different from 1 (p-value=0.00013).

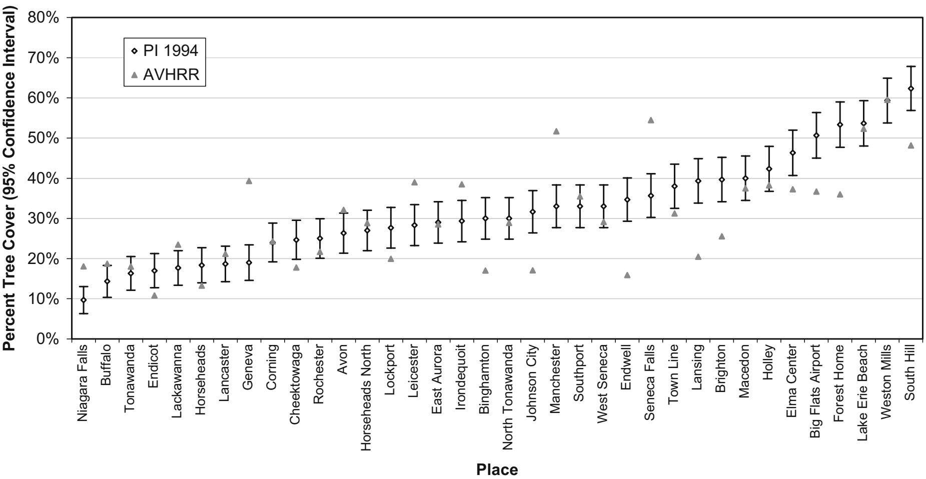

Fig. 4 1994 photo-interpreted (with 95% confidence) and AVHRR-derived percent tree cover for the 36 places sorted by reference tree cover (PI 1994)

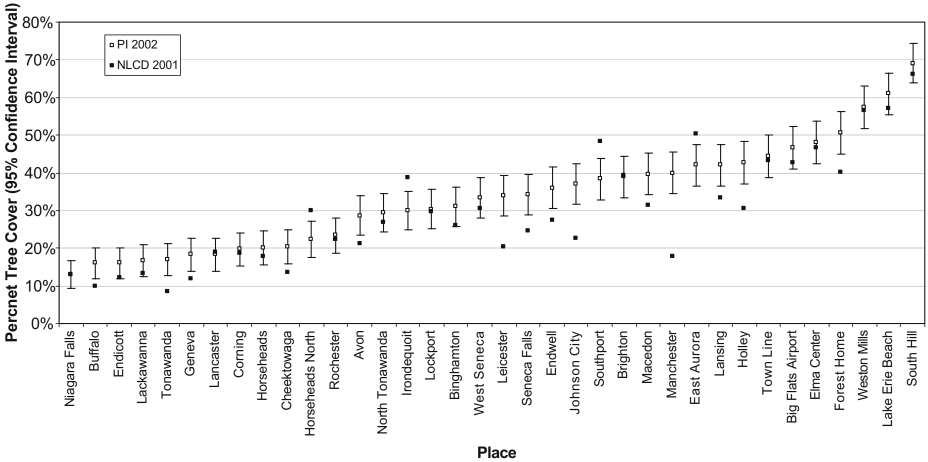

Fig. 5 2002 photo-interpreted (with 95% confidence) and NLCD-derived percent tree cover for the 36 places sorted by reference tree cover (PI 2002)

Discussion

Comparison of the imagery-derived tree cover values with the photo-interpreted 95% CI can be useful to identify areas that widely differ from the reference value, but identifying if a particular prediction is erroneous based on whether it fell inside the CI can be problematic (Walton 2005). The main concern is that the CI width could easily change just by adding or removing photo interpretation points. Also the confidence level (95%) affects the size of the CI. The 17 cities for NLCD (Fig. 5) and 13 for AVHRR (Fig. 4) within the CI noted above are specific to the size of the sample and the confidence level.

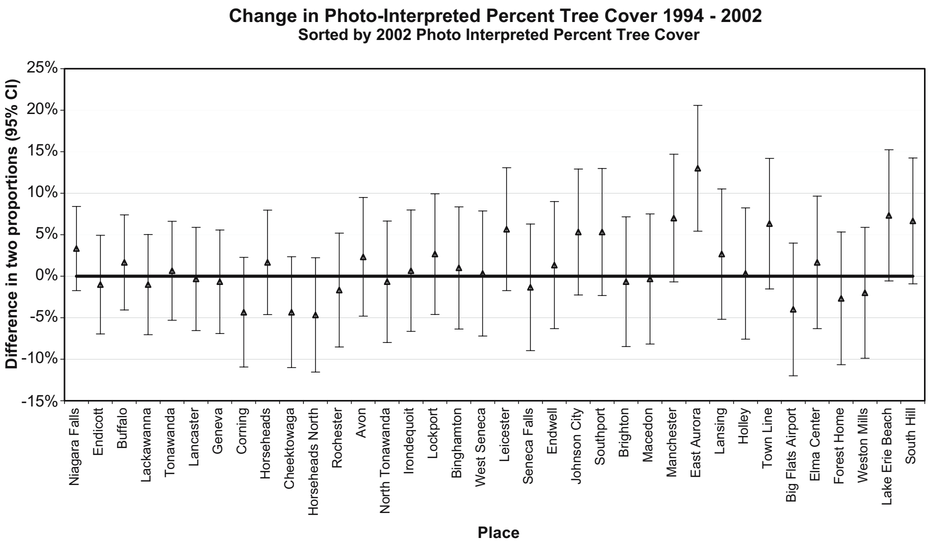

Fig. 6 Actual Change: Difference of the 2002 and the 1994 photo-interpreted percent tree cover (PI2002–PI1994) with 95% confidence intervals sorted by reference tree cover (PI 2002)

Because AVHRR comparison with the photo-interpreted values shows errors that are highly variable, the AVHRR-based urban tree cover assessment should not be used at the city scale if other more reliable data are available. However, at the regional scale or larger AVHRR should provide reasonable estimates of urban tree cover because there is no apparent bias in the estimate. While the NLCD estimates are more precise (generally lower overall errors than the AVHRR), the under estimate across the entire range of tree cover shown in mapping zone 63 requires more testing in other mapping zones to see if it is replicated elsewhere. The comparison of the 1994 and 2002 photo-interpreted reference datasets reveals a very good fit (Fig. 3c) indicating very little change in urban tree canopy cover for this region. One place with a high possibility for change, East Aurora, is shown to have a large difference (13% increase) and its 95% confidence interval does not contain zero (Fig. 6).

Based on the errors primarily in the AVHRR, but also the bias in the NLCD estimates, direct comparison of the two imagery datasets will not yield any definitive indications of change in the urban tree cover for a particular place. In the case of East Aurora (the place most likely to exhibit urban canopy cover change based on the photo-interpreted reference data), the AVHRR tree cover value is 28.5% while the NLCD value is 50.2% indicating an increase by 21.7% in the tree cover; whereas, the photo interpretation change estimate is about half that value, only 13% (Fig. 6). Places having the highest losses of tree canopy as shown by the AVHRR and NLCD estimates, such as Geneva (−27.4%), Manchester (−33.9%), and Seneca Falls (−29.8%), do not have as much loss (or even exhibit an increase in tree cover) when the photo-interpreted reference values are compared (Geneva, −0.7%; Manchester, +7.0%; Seneca Falls, −1.3%). Though Geneva has 27% water cover inside the CDP boundary, Manchester and Seneca Falls have little water (0 and 3%, respectively) to contribute to the large differences.

Urban tree cover estimates may be consistently low from mapping zone to mapping zone. This study shows a tendency for under prediction in urban areas of zone 63. Statistically estimated tree canopy cover values based on field plot data from Philadelphia and Baltimore (15.7 and 21.0%, respectively; Nowak et al. 2007; USDA Forest Service, unpublished data) compared to NLCD data for the same cities in mapping zone 60 (10.1 and 12.8% respectively; USDA Forest Service, unpublished data) also show an NLCD underestimate for those densely developed urban areas.

As shown, CDP-level change of urban tree cover is difficult to detect over this 10-year period in this region. Most of the places in this study are older, well-established cities or villages where urbanization and development are relatively static. Urban tree canopy cover change may be detected over this time frame in other areas of the US where urban growth is now happening.

Results of urban forest canopy cover change analyses have often been used to garner support, both social and financial, for urban forestry. Forest cover loss due to development can be dramatically illustrated with time series maps showing urban tree canopy cover over the last 3 decades. The recent release of the NLCD 2001 tree canopy estimate (Homer et al. 2007) used in conjunction with the RPA 2000 report (Dwyer et al. 2000) allows for the potential of urban tree canopy change to be calculated for every CDP in the US. As illustrated here, a direct comparison of these two imagery-based urban tree canopy cover estimates can lead to erroneous results.

Conclusion

Assessing change using differences between the AVHRR and NLCD urban tree cover datasets could not be done. The AVHRR dataset's accuracy varies considerably from place to place. Perhaps when a second dataset of higher reliability becomes available, urban forest change assessments can consistently be made on a national level.

Acknowledgments The work of Jayme Rapp, who did the photo interpretation, and Brett Kelly, who has helped develop the NLCD tree canopy layer for zone 63, both students at the State University of New York—College of Environmental Science and Forestry (SUNY-ESF), is gratefully acknowledged. David J. Nowak (USDA Forest Service) and Stephen V. Stehman (SUNY-ESF) reviewed an early version of the manuscript and provided many helpful suggestions. This work was done in cooperation with USGS EROS Data Center and many other cooperators. Funding for this project was provided, in part, by the USDA Forest Service's RPA Assessment Staff and State & Private Forestry's Urban and Community Forestry Program. Data analysis for this project was done in part using the open source statistical software R (R Development Core Team 2007). The use of trade, firm, or corporation names in this article is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or Forest Service of any product or service to the exclusion of others that may be suitable.

References

Dwyer JF, Nowak DJ, Noble MH, Sisinni SM (2000) Connecting people with ecosystems in the 21st century: an assessment of our nation's urban forests. Gen Tech Rep PNW-GTR-490. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 483 p. URL: http://treesearch.fs.fed.us/pubs/12517 (last accessed: 11 May 2007)

Homer C, Huang C, Yang L, Wylie B, Coan M (2004) Development of a 2001 National Land Cover Database for the United States. Photogramm Eng Remote Sensing 70(7):829–840 URL: http://www.mrlc.gov/pdfs/July_PERS.pdf (last accessed: 11 May 2007)

Homer C, Dewitz J, Fry J, Coan M, Hossain N, Larson C, Herold N, McKerrow A, VanDriel JN, Wickham J (2007) Completion of the 2001 National Land Cover Database for the Coterminous United States. Photogramm Eng Remote Sensing 73(4):337–341 URL: http://www.mrlc.gov/pdfs/April_07_highlight.pdf (last accessed: 11 May 2007)

Huang C, Yang L, Wylie B, Homer C (2001) A strategy for estimating tree canopy density using Landsat 7 ETM+ and high-resolution images over large areas. *Third International Conference on Geospatial Information in Agriculture and Forestry*. Denver, CO. CD-ROM. URL: http://landcover.usgs.gov/pdf/canopy_density.pdf (last accessed 11 May 2007).

Ji L, Gallo K (2006) An agreement coefficient for image comparison. Photogramm Eng Remote Sensing 72 (7):823–833

Mendenhall W (1979) Introduction to probability and statistics, 5th edn. Wadsworth, Belmont, California

Nowak DJ (1994) Understanding the structure of urban forests. J For 92(10):42–46 URL: http://www.fs.fed.us/ne/syracuse/Pubs/Downloads/94_DN_UnderstandingStr.pdf (last accessed: 11 May 2007)

Nowak DJ, Hoehn RE, Crane DE, Stevens JC, Walton JT (2007) Assessing urban forest effects and values, Philadelphia's urban forest. *Resource Bulletin NRS-7*. U.S. Department of Agriculture, Forest Service, Northern Research Station, Newtown Square, PA, p 22URL: http://nrs.fs.fed.us/pubs/9659 (last accessed: 11 May 2007)

R Development Core Team (2007) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0, URL: http://www.R-project.org

Walpole RE, Myers RH (1985) Probability and statistics for engineers and scientists, 3rd edn. MacMillan, New York

Walton JT (2005) An investigation of national tree canopy assessments applied to urban forestry. Ph.D. Dissertation. State University of New York, College of Environmental Science and Forestry. Syracuse, NY. 95 p. UMI Publication Number: AAT 3200353

Willmott CJ (1984) On the evaluation of model performance in physical geography. In: Gaile GL, Willmott CJ (eds) Spatial statistics and models. D. Reidel Publishing Company, pp 443–460

Yang L, Huang C, Homer CG, Wylie BK, Coan MJ (2003) An approach for mapping large-area impervious surfaces: synergistic use of Landsat 7 ETM+ and high spatial resolution imagery. Can J Remote Sensing 29(2):230–240 URL: http://www.mrlc.gov/includes/pdfs/CJRSimpmapping2003.pdf (last accessed: 11 May 2007)

Zhu Z (1994) Forest density mapping in the lower 48 states: a regression procedure. Research Paper SO-280. U.S. Department of Agriculture, Forest Service, Southern Research Station, New Orleans, LA, p 11URL: http://www.srs.fs.usda.gov/pubs/rp/rp_so280.pdf (last accessed: 11 May 2007)