COMPARISON OF FOREST AREA DATA IN THE CHESAPEAKE BAY WATERSHED

Tonya W. Lister and Andrew J. Lister[1]

Abstract.—The Chesapeake Bay, the largest estuary in the United States, has been designated by executive order as a national treasure. There is much interest in monitoring the status and trends in forest area within the bay, especially since maintaining forest cover is key to bay restoration efforts. The Chesapeake Bay Land Cover Data Series (CBLCD), a Landsat-based, multi-temporal change detection raster geographic information system (GIS) product was developed by the U.S. Geological Service (USGS) to monitor land cover change in the bay. The objective of this study was to assess relationships between the CBLCD dataset and Forest Inventory and Analysis (FIA) estimates of land use in order to provide a better understanding of the nature of the CBLCD and its potential for use in assessing forest cover dynamics. Data were summarized at different geographic scales, and differences between datasets were highlighted with the goal of providing information that will help users of the CBLCD interpret findings. Our analyses suggest there is a strong, positive relationship between the CBLCD forest information and that from the FIA data. Misclassifications can be explained by analyses created by integrating the CBLCD data with the FIA data and standard FIA reporting tools.

INTRODUCTION

The Chesapeake Bay Watershed, which includes parts of Maryland, Delaware, Pennsylvania, New York, Virginia and West Virginia, contains the largest estuary in the United States and provides habitat for thousands of species of plants and animals. The forests in this area provide many ecological services including protecting drinking water, serving as buffers against sedimentation and nutrient enrichment for estuarine species, and providing economic and other benefits for humans. However, the forests in the Chesapeake Bay Watershed have experienced a 2 percent net loss of forest land since the 1980s (Lister and Perdue 2011). Claggett et al. (2004) predict significant perturbations of the bay region's forests in the coming decades, due largely to development pressures.

[1] Research Foresters (TWL and AJL), U.S. Forest Service, Northern Research Station, 11 Campus Blvd, Newtown Square, PA 19073. TWL is corresponding author: to contact, call 610-557-4033 or email at tlister01@fs.fed.us.

In recognition of the importance of and current and future threats to the bay's natural resources, Congress and several of the states containing parts of the watershed established the Chesapeake Bay Program partnership in the early 1980s. This partnership was based in part on an agreement to protect and restore the Bay's ecosystems and has since been reaffirmed and updated. One component of the agreement is that the signatories will work to establish a system to monitor the status of and trends in forest cover.

The Chesapeake Bay Land Cover Data Series (CBLCD) was developed to help with these efforts to track and monitor forest land in the bay (Irani and Claggett 2010). The CBLCD is a 30-m Landsat-based, multi-temporal geographic information system (GIS) change detection dataset produced by the U.S. Geological Survey (USGS 2006). The data series was created by merging the National Land Cover Dataset (NLCD) from the USGS (Homer et al. 2007) with land cover data from the National Oceanic and

Atmospheric Administration's (NOAA) Coastal Change Analysis Program (C-CAP) (NOAA 2007) and improving the classification of urban and agricultural land. The CBLCD can be used to characterize land cover conversions for several dates between 1984 and 2006. It is not clear, however, how the land cover classification system used by the CBLCD, which is based on that of Anderson et al. 1976, agrees with that used by other agencies, such as the U.S. Forest Service's Forest Inventory and Analysis (FIA) Program. The FIA Program uses a network of ground plots to characterize land use on all lands along with tree and site information associated with forest land.

The goal of the current study was to assess relationships between the CBLCD dataset and FIA estimates of land use in order to provide a better understanding of the nature of the CBLCD and its potential for use in assessing forest cover dynamics.

MATERIALS AND METHODS

The original classification system used for the 2006 CBLCD image (USGS 2006) was reclassified into eight categories: open water/perennial ice and snow, developed open space, developed low-high intensity, barren land, forest including woody wetlands, shrub/scrub, grassland/herbaceous, pasture/hay/cultivated crops, and emergent herbaceous wetlands.

A GIS was used to intersect the U.S. Census Bureau's Tiger Line 2010 GIS county boundary files with a GIS file of the outline of the Chesapeake Bay Watershed. Within each of the counties that were completely within the watershed, areas of each of the eight reclassified CBLCD land cover classes were tabulated. Within these same counties, the area of forest land use was calculated from the FIA data using the PC EVALIDator tool, and a linear regression analysis with accompanying R2 values was calculated to assess relationships. Next, the individual FIA plots were used to characterize the CBLCD. Locations of the 4,784 FIA plots within the watershed (Fig. 1) were used to generate 43-m radius circular buffers which were then intersected with the CBLCD. The reclassified CBLCD class with maximum area within the buffer was assigned to each FIA plot location. This new attribute was then incorporated into the PC EVALIDator reporting tool and was used as a categorical variable for conducting cross tabulations of FIA data (Miles 2009). Cross tabulations included estimates of forest area by combinations of CBLCD class and FIA major land use, stocking, stand size, and physiographic classes (Woudenberg et al. 2010). In addition, a land use matrix from the PC EVALIDator tool that compares FIA major land use (forest, nonforest, and water) with the equivalent classes from the CBLCD was generated.

RESULTS AND DISCUSSION

An assessment of the relationship between the county FIA forest use and the CBLCD forest cover area estimates is shown in Figure 2. Summary statistics that describe the fit of the linear relationship indicate a strong, positive, nearly 1:1 linear relationship between the two attributes for all counties in each of the states, with R^2 values greater than 0.80 in all cases. This suggests that there is no spatial trend in the strength of the relationship, or spatial variation in the CBLCD product. The land class matrix that was created to compare forest, nonforest, and water classes on FIA-sampled land with the equivalent classes from the CBLCD (Table 1) also indicates strong agreement, with 87 percent agreement between the FIA and CBLCD estimates. This is further evidence that there is a strong correspondence between forest information from the CBLCD and the land use information associated with the FIA data. The strength of the relationship is somewhat surprising given the differences in the definitions of the attributes. For example, FIA defines forest land using a combination of stocking, aerial extent and shape of the land cover patch, and an assessment of the ability of the land to provide forest regeneration. The CBLCD definition of forest land is based on that of the USGS National Land Cover Dataset (NLCD) (Homer et al. 2007) and includes tree canopy cover thresholds instead of

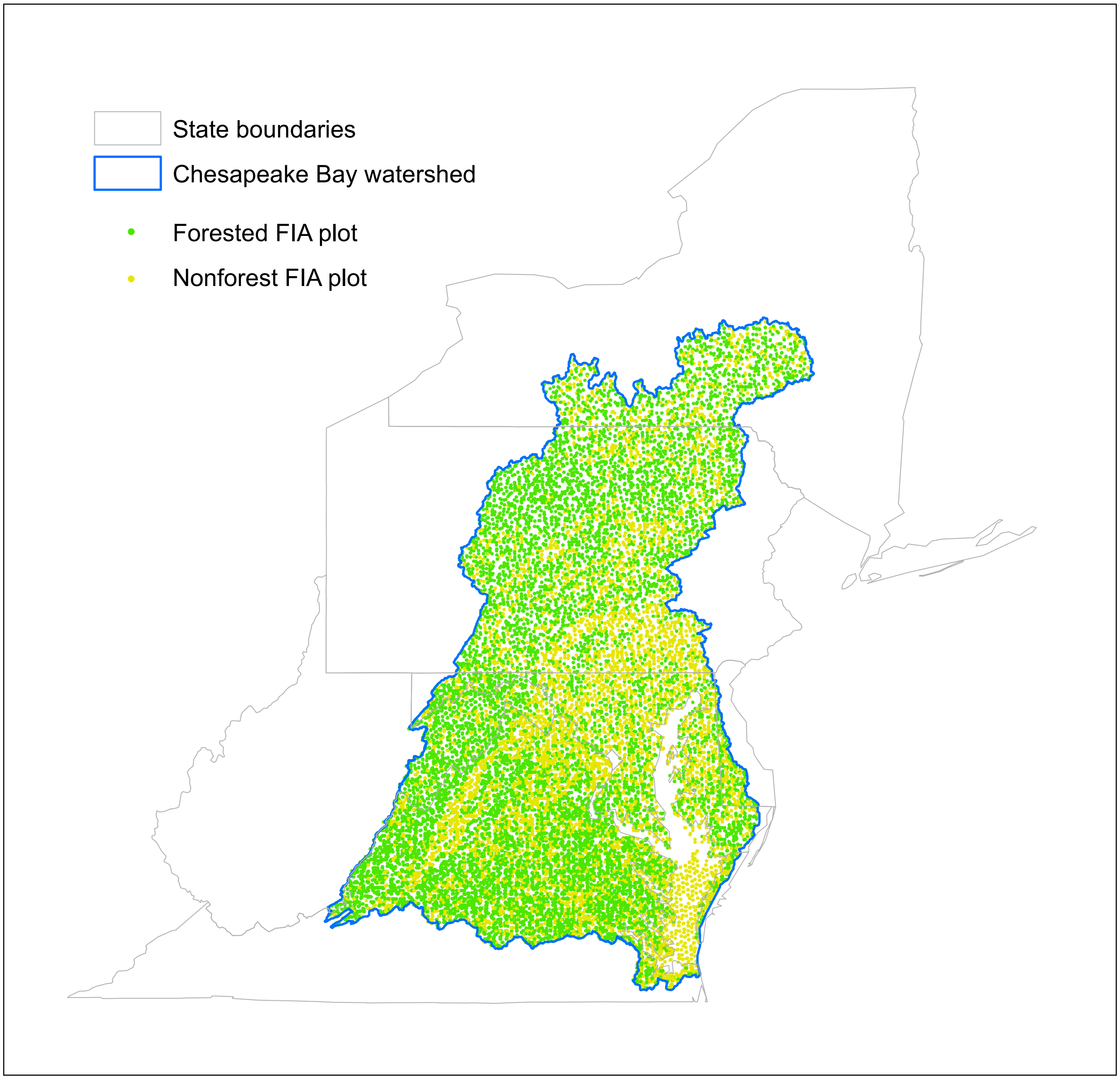

Figure 1.—Distribution of FIA plots within the Chesapeake Bay. Plot locations are approximate.

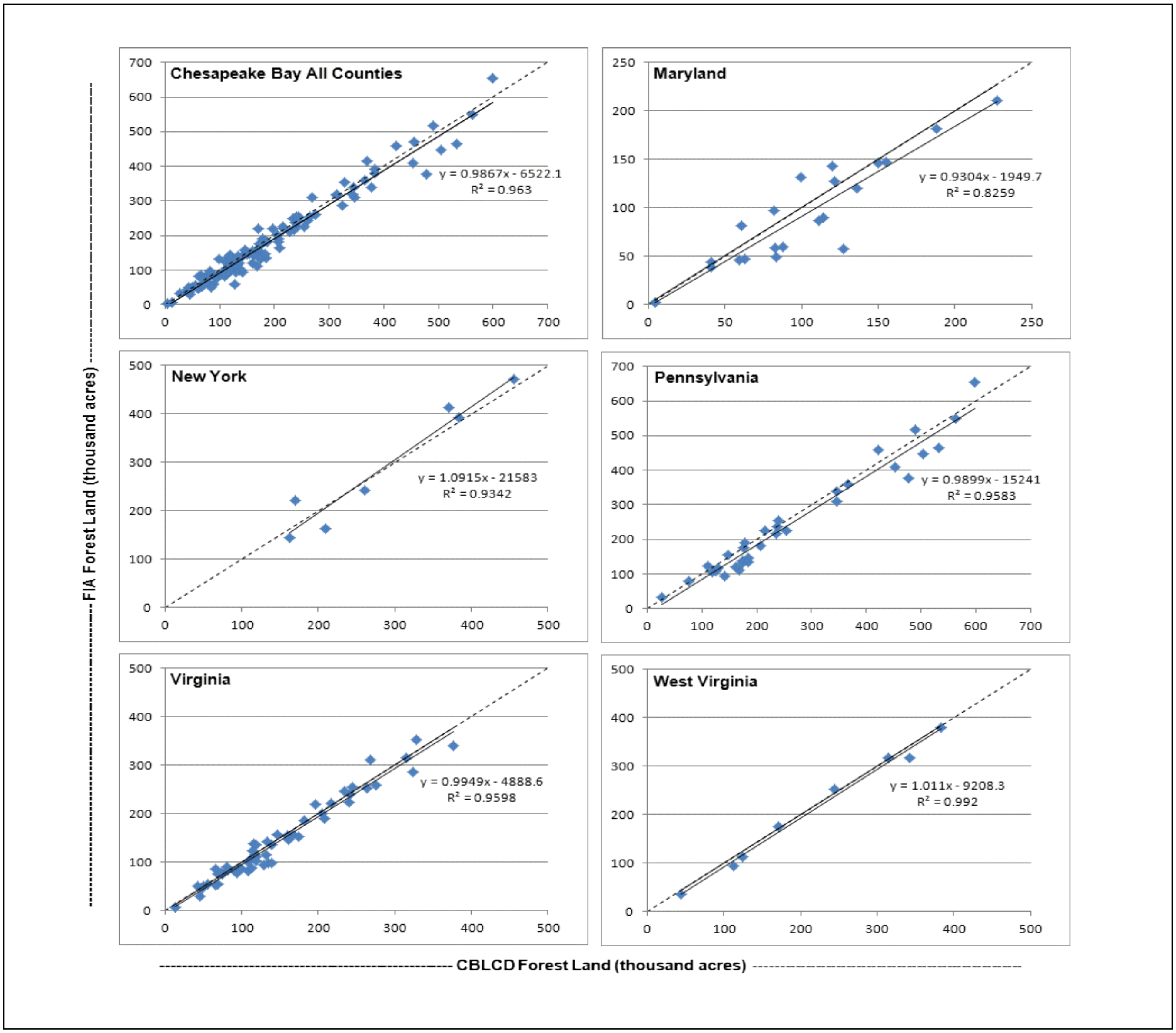

Figure 2.—Comparison of county-level FIA forest land estimates with county-level CBLCD within the Chesapeake Bay Watershed. Dashed line represents a 1:1 linear relationship.

Table 1.—Land use matrix comparing FIA to collapsed Chesapeake Bay Land Cover Data series (CBLCD) cover class data in the Chesapeake Bay Watershed

	CBLCD (acres)			
FIA (acres)	**Forest**	**Nonforest**	**Water**	**Total**
Forest	21,636,051	1,948,338	21,198	23,605,587
Nonforest	2,981,308	13,679,780	60,660	16,721,748
Water	202,872	97,682	457,811	758,365
Total	24,820,231	15,725,800	539,669	41,085,700

stocking and different aerial extent requirements for inclusion in the forest class. Although the distinction between land cover and land use can sometimes be stark (e.g., a clearcut with no trees can still be a forest land use), it is clear that in the case of the CBLCD and the FIA data, the majority of the land cover, as detected by the satellite, is also forest land use as detected by a ground observer on an FIA plot. Generally, it must be assumed that forest cover maps like the CBLCD are surrogates for forest land use maps, since forest land use maps are challenging to produce and are not widely available.

The area of forest land from the CBLCD classes assigned to the FIA plots was 5 percent higher than that derived from the FIA forest classes from the same plots (Table 1). This is an expected result given that the CBLCD estimates forest cover and FIA estimates are land use based. For example, the CBLCD may assign a forest cover class to treed areas where the underlying use is not forest, but rather a treed, low density residential development. Accuracy statistics from the NLCD 2001 dataset in Massachusetts and Rhode Island show a similar (almost 8 percent) overestimate of forest, with most misclassifications occurring where the NLCD forest class was actually developed for agricultural land use (Hollister et al. 2004).

Ninety-two percent of the land FIA plots classified as forest is also classified as forest by the CBLCD (Table 1). The majority of the remaining 8 percent is classified as agricultural and shrubland (Fig. 3). These types of differences may be due to the presence of marginal agricultural lands at various stages of succession in the area. It is likely that the CBLCD classification procedures had difficulty distinguishing old or fallow fields from forest. We had expected to see this hypothesis reinforced by an analysis of the areas of CBLCD classes within areas classified as FIA forest, partitioned by stocking class and stand diameter class (Fig. 4) since one might expect areas

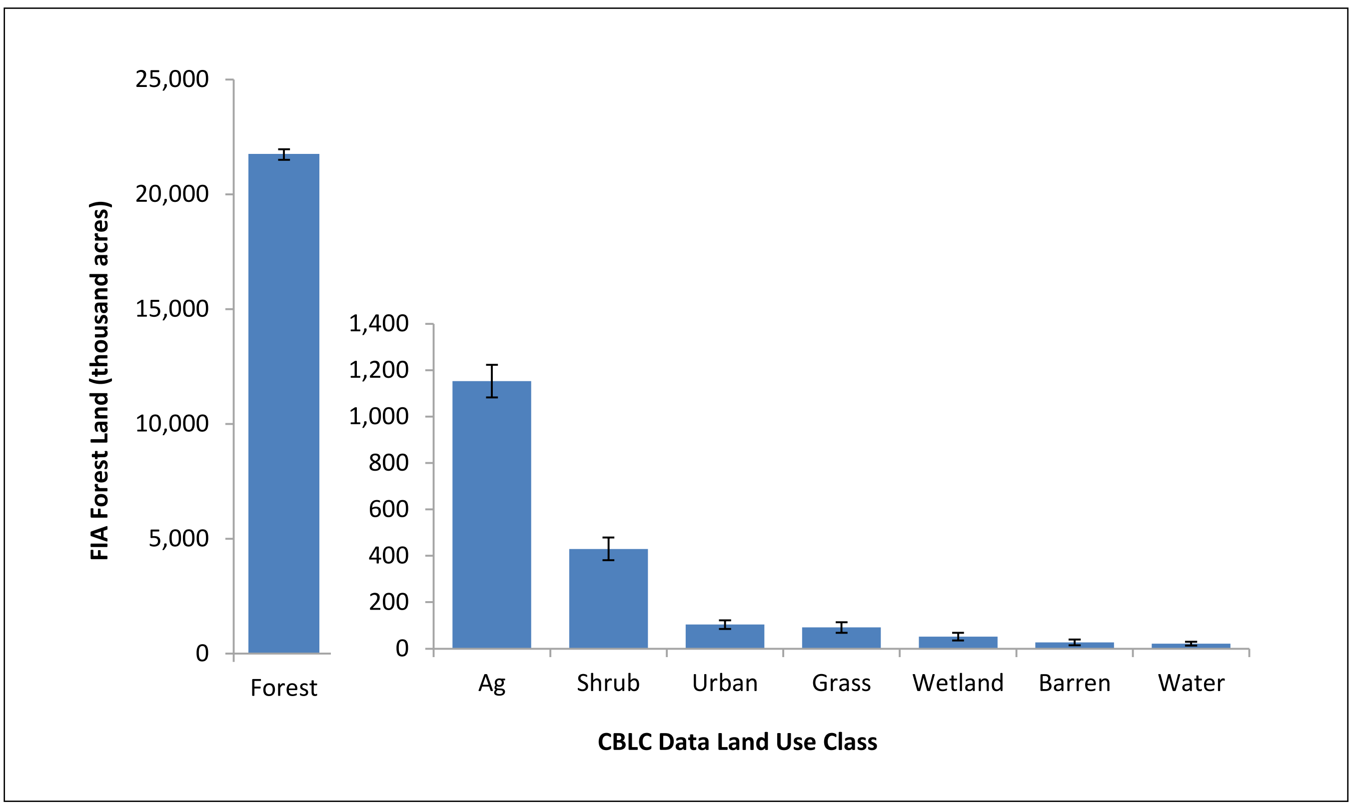

Figure 3.—Composition of the FIA forest land estimate by corresponding CBLCD land use class. Error bars represent 68 percent confidence intervals.

with a higher stocking class or larger trees to generate more of a "forest signal" in the CBLCD classification procedure. Although the data do not suggest that forested plots with low stocking classes are more likely to be misclassified, the data do suggest that there is generally a large proportion of medium and higher stocked plots in the forest class. Two notable exceptions to this pattern, however, are the barren and the grassland CBLCD classes, which also show relatively high areas of medium and higher stocked

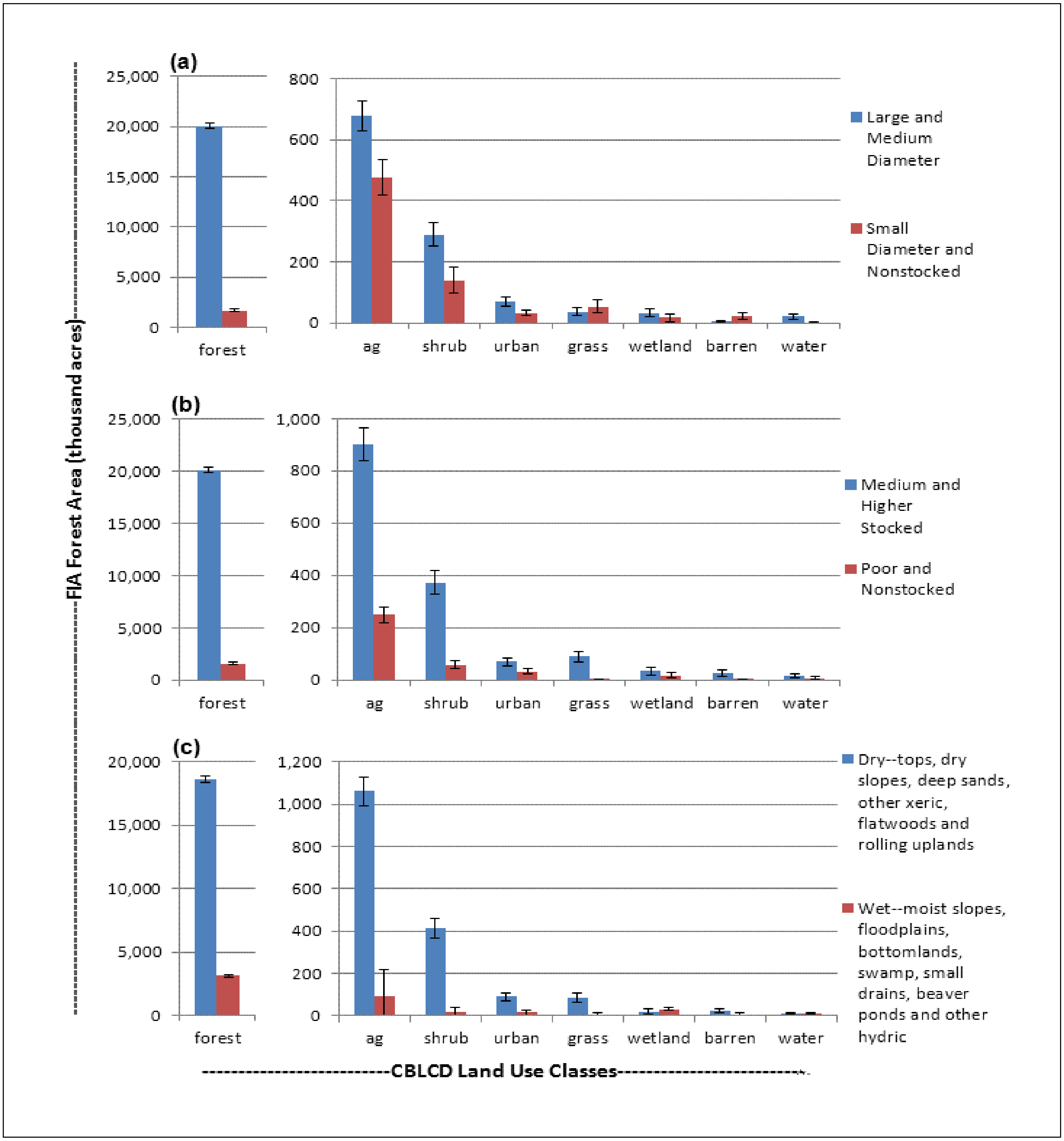

Figure 4.—Forest land area of select FIA forest class variables including stocking class (a), stand size class (b), and physiographic class (c), within each CBLCD land use class. Error bars represent 68 percent confidence intervals.

stands. This could be explained by the relatively large areas of small diameter stands within these classes (Fig. 4b), suggesting that the high stocking values are due to large numbers of small trees or seedlings. These areas may confuse the classification algorithm as they may resemble barren sites or grasslands.

Another interesting finding from our comparisons is the amount of FIA forest land area misclassified as water by the CBLCD (Fig. 3). While this is a fairly serious misclassification, Figure 4c suggests that the physiographic classes associated with these misclassified areas tend to be wetter (coves, floodplains, bottomlands, swamps, ponds) than areas associated with other types of misclassifications. In these areas, the signal imparted by these wet sites likely led to the confusion of the classification algorithm.

Our main conclusion from this study is that, assuming FIA data are representative of the forest attributes that Chesapeake Bay land managers want to monitor, the strong relationship between CBLCD and FIA data indicates that the CBLCD can be a valuable tool for characterizing land cover in the Chesapeake Bay Watershed. Our analyses suggest that in aggregate (at the county scale) there is a strong, positive relationship between the CBLCD forest information and that from the FIA data. Misclassifications that occurred can be explained by analyses created by integrating the CBLCD data with the FIA data and standard FIA reporting tools. Results of this study shed light on the relationship between the two land classification systems and will provide managers with information that can be used to not only interpret land class changes in the Chesapeake Bay Watershed, but also to help improve future versions of the CBLCD product. Future work will involve looking at species assemblage data and their effects on relationships between CBLCD and FIA data, as well as the relationship between other land cover products and FIA data in this important region.

LITERATURE CITED

Anderson, J.R.; Hardy, E.E.; Roach, J.T.; Witmer, R.E. 1976. **A land use and land cover classification system for use with remote sensor data.** Geological Survey Professional Paper 964. A revision of the land use classification system as presented in U.S. Geological Survey Circular 671. Washington, DC: U.S. Government Printing Office. 41 p.

Claggett, P.; Jantz, C.; Goetz, S.; Bisland, C. 2004. **Assessing development pressure in the Chesapeake Bay Watershed: an evaluation of two land-use change models.** Environmental Monitoring and Assessment. 94(1): 129-146.

Hollister, J.W.; Gonzalez, L.M.; Paul, J.F.; August, P.V.; Copeland, J.L. 2004. **Assessing the accuracy of national land cover dataset area estimates at multiple spatial extents.** Photogrammetric Engineering and Remote Sensing. 70(4): 405-414.

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.** Photogrammetric Engineering and Remote Sensing. 73(4): 337-341.

Irani, F.M.; Claggett, P.R. 2010. **Chesapeake Bay Watershed land cover data series.** U.S. Geological Survey Data Series 2010-505.

Lister, T.W.; Perdue, J. 2011. **Maryland's forest resources, 2010.** Res. Note. NRS-124. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 4 p.

Miles, P. 2009. **EVALIDatorReports: reporting beyond the FIADB.** In: McWilliams, W.; Moisen, G.; Czaplewski, R., comps. 2009. 2008 Forest Inventory and Analysis (FIA) Symposium; 2008 October 21-23; Park City, UT. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 1 CD.

National Oceanic and Atmospheric Administration (NOAA). 2007. **C-CAP zone 65: 2001-2006-era land cover metadata.** Charleston, SC: National Oceanic and Atmospheric Administration, Coastal Services Center. Available at http://www.csc.noaa.gov/digitalcoast/data/ccapregional/. (Accessed July 15, 2012).

U.S. Geological Survey (USGS). 2006. **Chesapeake Bay Watershed land cover data series 2006 metadata document.** Reston, VA: U.S. Geological Survey. Available at http://www.pasda.psu.edu/uci/FullMetadataDisplay.aspx?file=CBWLCD_2006.xml#Metadata_Reference_Information. (Accessed July 15, 2012).

Woudenberg, S.W.; Conkling, B.L.; O'Connell, B.M.; LaPoint, E.B.; Turner, J.A.; Waddell, K.L. 2010. **The Forest Inventory and Analysis database: database description and users manual. Version 4.0 for phase 2.** Gen. Tech. Rep. RMRS-245. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 336 p.