

Accuracy assessment of the vegetation continuous field tree cover product using 3954 ground plots in the south-western USA

M. A. WHITE*†, J. D. SHAW‡ and R. D. RAMSEY†

†Department of Aquatic, Watershed, and Earth Resources, 5210 Old Main Hill,
Utah State University, Logan, UT 84322-5210, USA

‡US Forest Service, Rocky Mountain Research Station, Forest Inventory and Analysis,
Ogden Forestry Sciences Lab, 507 25th St, Ogden, UT 84401, USA

(Received 17 September 2004; in final form 28 October 2004)

An accuracy assessment of the Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation continuous field (VCF) tree cover product using two independent ground-based tree cover databases was conducted. Ground data included 1176 Forest Inventory and Analysis (FIA) plots for Arizona and 2778 Southwest Regional GAP (SWReGAP) plots for Utah and western Colorado. Overall rms. error was 24% for SWReGAP and 31% for FIA data. VCF bias was positive at low observed tree cover but systematically increased thereafter until at greater than 60% observed tree cover, VCF tree cover was 40% (SWReGAP) to 45% (FIA) too low. Errors are unlikely to be related to habitat fragmentation or variation in canopy height but may be influenced by scaling discontinuities between ground and satellite resolutions.

1. Introduction

Tree cover, defined as the percentage of ground surface area covered by a vertical projection of live tree crowns, is an important variable for basic and applied science. Tree cover strongly influences processes in diverse fields such as carbon cycle simulation (Simioni *et al.* 2003), bird habitat (Chapman *et al.* 2004), and soil erosion (Hartanto *et al.* 2003). Clearly, synoptic and accurate representation of tree cover has the potential to be widely used.

A global 500 m spatial resolution, per cent tree cover product now exists (Hansen *et al.* 2002, 2003) as part of the Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation continuous fields (VCF) project. The tree cover product clearly exhibits generally realistic global patterns and is gaining wide acceptance as an alternative to categorical land-cover classifications.

VCF scientists have conducted a validation using finer resolution remote sensing products (Hansen *et al.* 2003) but such approaches may produce different values than physically measured ground observations (O'Brien 1989). Thus, some uncertainty exists regarding the relationship between VCF tree cover and ground conditions. In the UK, Disney and Lafont (2004) found that the VCF tended to overestimate measured tree cover but to our knowledge, no extensive ground-based accuracy assessment of the VCF tree cover product exists for the USA.

*Corresponding author. Email: mikew@cc.usu.edu

Here, using two independent ground datasets totalling 3954 individual plots, an accuracy assessment is presented of the VCF tree cover product for portions of the western USA, a region spanning semi-arid deserts, sparse dry woodlands, and cool mesic upland forests. Woody encroachment impacts on carbon sequestration, conflicts related to water resources and endangered species, and a wide range of tree cover make the western USA a scientifically and societally relevant location for such an assessment.

2. Data and methods

2.1 *Remotely sensed tree cover*

The version 1.0 North America 500 m VCF per cent tree cover product (data and technical manuals available from <http://www.landcover.org/data/modis/vcf/>) was obtained. Major elements of the fully automated VCF approach include the following three steps. First, the seven MODIS land reflectance bands and Normalized Difference Vegetation Index (NDVI) were used to develop 68 metrics, including numerous NDVI and channel-based measures of vegetation phenology and maximum, minimum and mean vegetation conditions. Second, 1995–1996 Advanced Very High Resolution Radiometer (AVHRR) brightness temperature was used to generate a global climate stratification. Third, a global database of more than 250 finer-resolution sites was used as tree cover training data (mostly 30 m and from years prior to the MODIS collection (DeFries *et al.* 1998)). Herein, each pixel was assigned to one of four tree cover strata with a prescribed per cent tree cover of 0%, 25%, 50% or 80% (for trees greater than 5 m in height). The fine resolution pixels were then aggregated to the MODIS resolution and used to train a regression tree (see Hansen *et al.* 2002 for details), which was subsequently modified by a stepwise procedure and bias adjustment. Thus the training data used as dependent variables were subject to uncertainties in the use of AVHRR brightness temperatures, input land-cover classification, and the 0%, 25%, 50% and 80% tree cover assigned to the four tree cover strata.

2.2 *Intensively measured ground data*

Tree cover data for Arizona from the Forest Inventory and Analysis (FIA) project were obtained. The data consisted of 1778 674.5 m² plots coded with forest types and measured on a regular grid of forested areas, each with four circular 168.6 m² subplots. Within each subplot, tree cover was measured along four slope-corrected 7.6 m line intercept samples (LIS) started 0.3 m from the subplot centre and oriented orthogonally along the cardinal directions (US Department of Agriculture 2002). Trees with a diameter at breast height of more than 0.0254 m were measured. From the overall dataset, 1176 plots were extracted for which the following two conditions were met: (1) complete subplot information was available, and (2) a condition proportion of 1.0 existed, indicating that all four subplots sampled a single forest type (i.e. stand) and that no barren or water coverage existed.

2.3 *Extensively measured ground data*

Next, the Utah and western Colorado Southwest Regional GAP (SWReGAP) field dataset was obtained. From this, the 2778 plots containing non-zero tree cover were extracted. The SWReGAP data were collected entirely independently from the FIA

data and with a different method. In contrast to the regular sampling interval and time consuming LIS methods of the FIA project, SWReGAP sampled mostly along secondary road networks and used a visual estimation of tree cover for plots of at least 8100 m². Field technicians worked in teams and developed a consensus estimate for each plot. Most observations were recorded in 5% increments. Species-level decisions were used to assign individual plants to tree or shrub categories but no height or diameter distinctions were used. The two datasets thus represent two extremes of ground sampling philosophy: intensive discrete measurements using established forest mensuration techniques sampled along a fixed landscape grid (FIA) versus a non-intensive co-calibrated visual estimation technique sampled along a non-random road network (SWReGAP).

2.4 Analysis

A tiered analysis for the FIA and SWReGAP accuracy assessments was used. First, the overall rms. error and bias (mean of VCF minus FIA or SWReGAP) was calculated. Second, rms. error and bias was calculated for observed tree cover from 0–10%, 11–20%, 21–30%, 31–40%, 41–50%, 51–60%, and greater than 60%. Third, the overall relationship between observed (FIA/SWReGAP) tree cover and the predicted (VCF) minus observed tree cover residuals was assessed. Finally, this same relationship was assessed for individual FIA forest types and SWReGAP ecological systems for which more than 50 plots were available.

3. Results and discussion

In the overall assessment, VCF tree cover rms. error was 31% for the FIA data and 24% for the SWReGAP data (table 1). An overall negative bias existed such that VCF tree cover was, on average, 19% lower than FIA tree cover and 14% lower than SWReGAP tree cover (table 1). For both assessments, the rms. error increased with increasing observed tree cover. Above 60% observed tree cover, VCF tree cover was approximately 40% too low (table 1).

The residuals between the predicted (VCF) and observed (FIA/SWReGAP) tree cover decreased systematically as a function of observed tree cover (figure 1). For both the FIA and SWReGAP assessments, the residual plot is bounded on the lower end by a linear feature in which observed tree cover was non-zero but VCF tree

Table 1. Accuracy assessment of the vegetation continuous fields per cent tree cover product with the ground-based Forest Inventory and Analysis (FIA) and Southwest Regional GAP (SWReGAP) datasets. The first column shows range of ground-measured tree cover. All data are per cent.

	FIA		SWReGAP	
	rms. error	Bias	rms. error	Bias
Overall	31	– 19	24	– 14
0–10%	15	5	12	4
11–20%	14	– 5	12	– 1
21–30%	20	– 12	15	– 9
31–40%	25	– 18	19	– 14
41–50%	32	– 26	24	– 19
51–60%	38	– 33	30	– 26
>60%	49	– 45	44	– 40

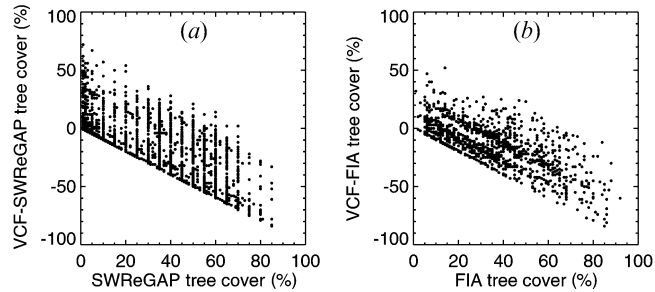


Figure 1. Residuals of predicted (VCF) tree cover minus observed tree cover as a function of observed tree cover. (a) Southwest Regional GAP: $y = -0.79x + 7.4$, $p < 0.01$. (b) Forest Inventory and Analysis: $y = -0.67x + 8.9$, $p < 0.01$. Negative values indicate VCF underestimation of observed tree cover.

cover was 0% or 1%. This occurred in 33% of SWReGAP plots and 13% of FIA plots, suggesting that the VCF failed to identify tree cover over a wide range of existing cover conditions.

In spite of significantly different ground methodologies, the FIA and SWReGAP residual plots are highly consistent, suggesting that possible errors in ground geolocation or sampling strategy are not the cause of the observed patterns. Land-cover fragmentation may produce differences between ground and satellite estimates. Subjective site selection could cause fragmentation-related bias for the SWReGAP assessment (figure 1(a)) but this is not the case for the fixed-grid FIA assessment points used here (figure 1(b)). Further, spatially coherent patterns of increasing residuals along a gradient from heavily forested stands to fragmented woodlands, as in the Mogollon Rim of central Arizona, did not occur.

Frequency distributions for the observed and predicted tree cover support a finding of systematic biases in the study region (figure 2). In the Utah and Colorado SWReGAP dataset, which includes numerous non forest-dominated ecological system, plots with tree cover less than 5% were common but VCF tree cover in this category was more than twice as abundant (figure 2(a)). At greater than 20% tree cover, the VCF distribution consistently underrepresented the frequency of SWReGAP tree cover. In the Arizona FIA dataset of forested regions, tree cover was rare below 5%, extensive between 5% and 45%, and progressively less common up to 90%. In contrast, VCF tree cover was strongly modal at less than 5%, 10–15%, and 25–30%. VCF tree cover between 20 and 25% and greater than 60% was rare

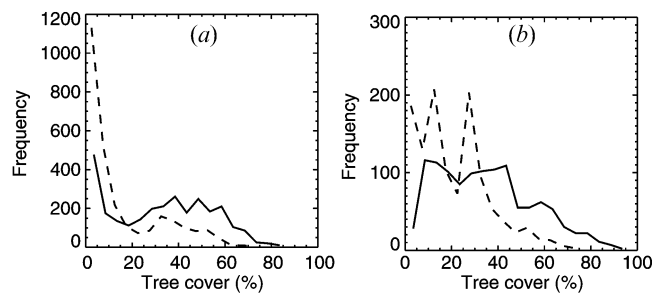


Figure 2. Frequency distributions of predicted (VCF, dashed line) and observed tree cover (solid line). (a) Southwest Regional GAP. (b) Forest Inventory and Analysis.

(figure 2(b)). The low 20–25% VCF tree cover accounts for the empty region in the centre of the residual distribution in figure 1(b).

The shape of the residuals graph (figure 1) is not caused by large errors for specific FIA forest types or SWReGAP ecological systems, i.e. short canopy piñon/juniper woodlands did not have significantly different residuals than tall canopy ponderosa pine or deciduous aspen stands. For all cases in which more than 50 plots were available, all forest types and ecological systems showed the same pattern of increasingly negative residuals that is seen in the overall relationship (figure 1).

In spite of the overall consistency of this analysis, it is possible that scale discrepancies between the ground and VCF data may complicate this accuracy assessment. Due to sampling design, it was not possible to assess this possibility with the FIA or SWReGAP data. Instead, 30 m tree cover data for mountainous regions of Utah were obtained (www.landfire.gov). Data were produced with a regression tree using Landsat Thematic Mapper data trained to tree cover derived from 1 m digital orthophoto quadrangles. The 30 m data were resampled to 240 m and 510 m; the corresponding SWReGAP data were extracted; and a bias assessment was conducted at each resolution. For all comparisons, a pattern of residuals existed similar to those found in figure 1 but the slope of the relationship varied by resolution: -0.44 at 30 m; -0.57 at 210 m; and -0.73 at 510 m (similar to the slopes in figure 1). No consistent pattern of rms. error existed across resolutions. Consequently, although the two remotely sensed estimates of tree cover are not directly comparable, this analysis suggests that the increasingly negative bias shown in table 1 is probably affected by scaling issues and would likely be less severe, but still present, if the ground and satellite sensor data were sampled at compatible resolutions.

This assessment suggests that the VCF tree cover product should be used with caution in the south-western USA. Future steps to our analysis will include the expansion of this assessment to the entire intermountain USA, assessment of the VCF validation strategy using high resolution IKONOS imagery (Hansen *et al.* 2003), and a recalibration of the VCF product for intermountain USA tree cover.

Acknowledgements

M. A. White was supported by NASA grant NAG5-11282. We thank J. A. Blackard, Interior West Forest Inventory and Analysis, for extracting the VCF data at FIA plot locations.

References

- CHAPMAN, R.N., ENGLE, D.M., MASTERS, R.E. and LESLIE, D.M., 2004, Tree invasion constrains the influence of herbaceous structure in grassland bird habitats. *Ecoscience*, **11**, pp. 55–63.
- DEFRIES, R., HANSEN, M., TOWNSHEND, J.R.G. and SOHLBERG, R., 1998, Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, **19**, pp. 3141–3168.
- DISNEY, M. and LAFONT, S., 2004, Evaluation of the MODIS continuous tree cover map over the UK. *MODIS Vegetation Workshop II, 17 April 2004*, <http://www.nts.gov/ModisCon/>.
- HANSEN, M.C., DEFRIES, R., TOWNSHEND, J., SOHLBERG, R., DIMICELI, C. and CARROLL, M., 2002, Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment*, **83**, pp. 303–319.

- HANSEN, M.C., DEFRIES, R., TOWNSHEND, J., CARROLL, M., DIMICELI, C. and SOHLBERG, R., 2003, Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS vegetation continuous fields algorithm. *Earth Interactions*, **7**, Paper No. 10.
- HARTANTO, H., PRABHU, R., WIDAYAT, A.S.E. and ASDAK, C., 2003, Factors affecting runoff and soil erosion: plot-level soil loss monitoring for assessing sustainability of forest management. *Forest Ecology and Management*, **180**, pp. 361–374.
- O'BRIEN, R.A., 1989, Comparison of overstory canopy cover estimates on forest survey plots. USDA Forestry Service General Technical Report INT-GTR-417.
- SIMIONI, G., GIGNOUX, J. and LE ROUX, X., 2003, Tree layer spatial structure can affect savanna production and water budget: results of a 3-D model. *Ecology*, **84**, pp. 1879–1894.
- US DEPARTMENT OF AGRICULTURE, 2002, Interior West Resource Inventory, Monitoring, and Evaluation Program Forest Survey Field Procedures, Annual Ver.-1.6. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Ogden, Utah.