Chapter 8

Mapping Existing Vegetation Composition and Structure for the LANDFIRE Prototype Project

Zhiliang Zhu, James Vogelmann, Donald Ohlen, Jay Kost, Xuexia Chen, and Brian Tolk

Introduction

Overview

The Landscape Fire and Resource Management Planning Tools Prototype Project, or LANDFIRE Prototype Project, required the mapping of existing vegetation composition (cover type) and structural stages at a 30-m spatial resolution to provide baseline vegetation data for the development of wildland fuel maps and for comparison to simulated historical vegetation reference conditions to develop indices of ecological departure. For the LANDFIRE Prototype Project, research was conducted to develop a vegetation mapping methodology that could meet the following general requirements:

- Cover types (species composition) must be characterized at a scale suitable for subsequent mapping of wildland fuel and fire regime condition class (FRCC). The vegetation map unit classification used for mapping cover types must be based on existing national systems, such as the United States National Vegetation Classification System (NVCS; Grossman and others 1998). The alliance (a community with multiple dominant species) or association (a community with a single dominant species) levels of this standard must provide a clearly defined list of map units that can be used as a basis for mapping vegetation classes that are both scaleable and representative of suitable units for modeling historical fire regimes (see Long and others, Ch. 6 for details on the LANDFIRE vegetation map units).
- The mapping of existing vegetation structure must be based on the relative composition of forest, shrub, and herbaceous canopy cover and average forest, shrub, and herbaceous canopy height. Although structural stages are discrete map units describing unique combinations of canopy cover and canopy height by life form, mapping individual canopy cover and height variables as continuous variables is desired to provide additional information for mapping and modeling vegetation and flexibility for setting threshold values.

The task of mapping existing vegetation is interconnected with several major tasks performed in the LANDFIRE Prototype Project. The mapping of existing vegetation requires attribute tables developed from the LANDFIRE reference database (LFRDB) (Caratti, Ch. 4), satellite imagery acquisition and processing, the development of a vegetation map unit classification system (see Long and others, Ch. 6), the development of a biophysical settings stratification (Frescino and Rollins, Ch. 7), and the modeling of environmental gradient layers (Holsinger and others, Ch. 5). The design and testing of the vegetation mapping methodology have substantial influences on the outcome of the overall project because accuracies of subsequent products (such as maps of wildland fuel) are a function of the accuracy of mapped vegetation types and structure. In this chapter, we discuss the design features of the existing vegetation mapping component of LANDFIRE and present

In: Rollins, M.G.; Frame, C.K., tech. eds. 2006. The LANDFIRE Prototype Project: nationally consistent and locally relevant geospatial data for wildland fire management. Gen. Tech. Rep. RMRS-GTR-175. Fort Collins: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

results of the prototype. We conclude the chapter with recommendations for the national implementation of a consistent vegetation mapping effort.

Technical Problems

Significant technical limitations exist regarding achieving desired accuracies in the mapping of vegetation types and structure variables over broad areas. In the LANDFIRE Prototype, accuracies were affected by the spatial resolution, geographic extent, and information content defined by the project's objectives. The U.S. Geological Survey (USGS) Gap Analysis Program demonstrated the feasibility of mapping many existing vegetation cover types at the regional scale; however, methodologies have been inconsistent between regions (Eve and Merchant 1998). In addition, the mapping of forest canopy cover using imagery and regression techniques has been routinely performed for the operational mapping of vegetation structure variables (Huang and others 2001). Beyond that, however, literature reporting success stories regarding the mapping of vegetation structure using imagery is scant.

We conducted a prototype study to test a methodology for mapping vegetation cover types and structure variables. The three central objectives of the study were to:

- test an adaptable approach for mapping existing vegetation types and canopy structure at a 30-m resolution for the entire prototype area;
- develop digital maps of existing vegetation types and structural stages and conduct an accuracy assessment for the vegetation deliverables; and
- document research findings and limitations to the consistent mapping of existing vegetation composition and structure.

Specifically, this study tested a vegetation mapping protocol that met the design criteria and guidelines of the LANDFIRE Project (Keane and Rollins, Ch. 3). Further, this study investigated the limitations of using data contained within the LANDFIRE reference database (Carrati, Ch. 4) as training data and the applicability of satellite and ancillary data in meeting LANDFIRE's objectives. For vegetation modeling and wildland fuel mapping, the LANDFIRE Prototype Project required a structural stage map classified on the basis of mapped canopy cover (closed and open) and canopy height (high and low) by forest, shrub, and herbaceous life forms. We attempted to generate continuous maps of vegetation height and cover to maximize the utility of these products in a variety of applications.

As described in Rollins and others (Ch. 2), the LANDFIRE Prototype Project was conducted in two mapping zones: Zone 16, located in the central highlands of Utah and covering approximately 4 million ha of forest ecosystems (57 percent of the total land cover) and 2.5 million ha of shrub and herbaceous ecosystems (35 percent of the total land cover); and Zone 19, located in the northern Rocky Mountains of western Montana and northern Idaho and covering approximately 5.4 million ha of forest ecosystems (47 percent of the total land cover) and 5 million ha of shrub and herbaceous ecosystems (44 percent of the total land cover).

Literature Review of Vegetation Mapping

Similar to other natural science problems, the regional-scale mapping of vegetation types and structure variables carries unique technical and organizational challenges (Gemmell 1995). Spatial variations of vegetation types and structure are generally not characterized by unique spectral signatures, as captured by conventional broadband optical sensors (Kalliola and Syrjanen 1991; Keane and others 2001). Although significant improvements can be made by using specialized sensors, such as hyperspectral spectrometer or canopy lidar, data from such sensors having desired spatial resolutions are not available at national or regional scales. The associated enormous data volumes and high costs (in time and labor) make these technologies impractical for large-area applications at the present time.

Various techniques exist for modeling and estimating vegetation type and canopy structure (particularly percent forest cover); these include physics-based canopy reflectance models, empirical models linking ground-referenced data to satellite imagery, spectral mixture analysis, neural networks, and direct measurement using lidar and interferometric synthetic aperture radar. Each of these approaches has limitations in large-area applications, such as those related to cost and consistency. However, recent applications using the classification and regression tree (CART) approach (Breiman and others 1984) have been found to overcome many such limitations, provided sufficient amounts of field and geospatial data are available. Recent studies (Friedl and others 2002; Huang and Townshend 2003; Mahesh and Mather 2003; Yang and others 2003) have demonstrated the utility of CART techniques in mapping land cover, estimating species distribution, modeling percent forest canopy cover, and computing imperviousness at a 30-m grid resolution for large areas and even for the United States. Although CART techniques require relatively little human decision-making during algorithm executions, it is important to note that, ultimately, the knowledge scientists have acquired through studying vegetation patterns and attributes enhances the development mapping models to produce the most accurate results possible. Computer classifiers, regardless of their sophistication, are no substitute for scientists' understanding of the patterns, attributes, and conditions of existing vegetation and associated ecological processes.

Environmental data layers (such as elevation) are important predictor variables for characterizing vegetation patterns and attributes and for stratifying the distribution of vegetation along environmental gradient lines (Balice and others 2000). The use of spectral bands in combination with topographic data (for example, digital elevation models (DEM), slope, and aspect) is common in many land cover and vegetation mapping applications. However, topographic data capture only a part of the overall environmental factors that determine the establishment, growth, distribution, and succession of plant species and associations. The incorporation of a more complete set of environmental gradient layers into the mapping of existing vegetation should lead to increased predictive power and thematic accuracy (Keane and others 2002; Rollins and others 2004). Keane and others (2002) discuss techniques for deriving an entire set of climate, soil, and ecological gradient layers using interpolated weather observations in conjunction with topographic and soil databases and also describe the advantages of using such biophysical gradients in combination with remote sensing and field data to map vegetation, wildland fuel, and general ecosystem conditions.

In addition to the development and use of gradient variables, Keane and others (2001, 2002), Keane and Rollins, Ch. 3, and Rollins and others (2004) also suggest an approach for developing site-specific biophysical settings maps by mapping stable, late-seral communities as a function of certain climate, topographic, soil, and ecological gradients. This mapped "potential" vegetation can be used as a stratification tool in mapping actual vegetation distribution by constraining the distribution of cover types to those geographic strata where growth of the cover types' dominant species is ecologically possible.

Methods

The LANDFIRE Prototype Project involved many sequential steps, intermediate products, and interdependent processes. Please see appendix 2-A in Rollins and others, Ch. 2 for a detailed outline of the procedures followed to create the entire suite of LANDFIRE Prototype products. This chapter focuses specifically on maps of vegetation composition and structure, which served as important precursors to maps of wildland fuel and ecological departure in the LANDFIRE Prototype Project. Figure 1 outlines the technical approach used in LANDFIRE Prototype vegetation mapping and illustrates the data flow between several technically challenging tasks. Details of these tasks are described below.

Satellite Data Acquisition and Processing

The LANDFIRE Project partnered with the Multi-Resolution Land Characterization (MRLC) Consortium (Homer and others 2004) to facilitate the acquisition and processing of Landsat imagery. The consortium has completed the acquisition and processing of a full set of Landsat imagery for the United States with a minimum of three cloud-cover dates (circa 2001) for each pixel corresponding to phenological cycles of leaf-on, leaf-off, and spring green-up. Huang and others (2002) describe the steps involved in processing the MRLC satellite imagery, including terrain-corrected geometric registration and radiometric calibration using at-satellite reflectance models, calculations of normalized difference of vegetation index (NDVI), and tasseled cap transformations. The MRLC Consortium-sponsored development of the National Land Cover Dataset (NLCD) includes general land cover map units such as forest, agriculture, water, and urban areas mapped at a 30-m resolution (Homer and others 2004). The acquisition and processing of satellite imagery and the mapping of NLCD land cover map units were conducted for mapping zones, which were loosely delineated along major ecological regions. The LANDFIRE central Utah highlands and northern Rockies prototype areas were examples of these MRLC map zones.

The LANDFIRE Prototype Project had access to the following data layers from the MRLC catalogue for the Utah and northern Rockies prototype areas: 10 spectral bands for each of the 3 Landsat seasonal acquisitions (6 original spectral bands excluding the thermal band, 3 tasseled cap transformation bands, and 1 NDVI band) and land cover classes mapped to Anderson's Level 1 land cover classification (Anderson and others 1976). Using these data as a starting point, we mapped forest, shrub, and herbaceous cover types and structure attributes. These maps formed the foundation for mapping wildland fuel and fire regime characteristics (Holsinger and others, Ch. 11; Keane and others Ch. 12).

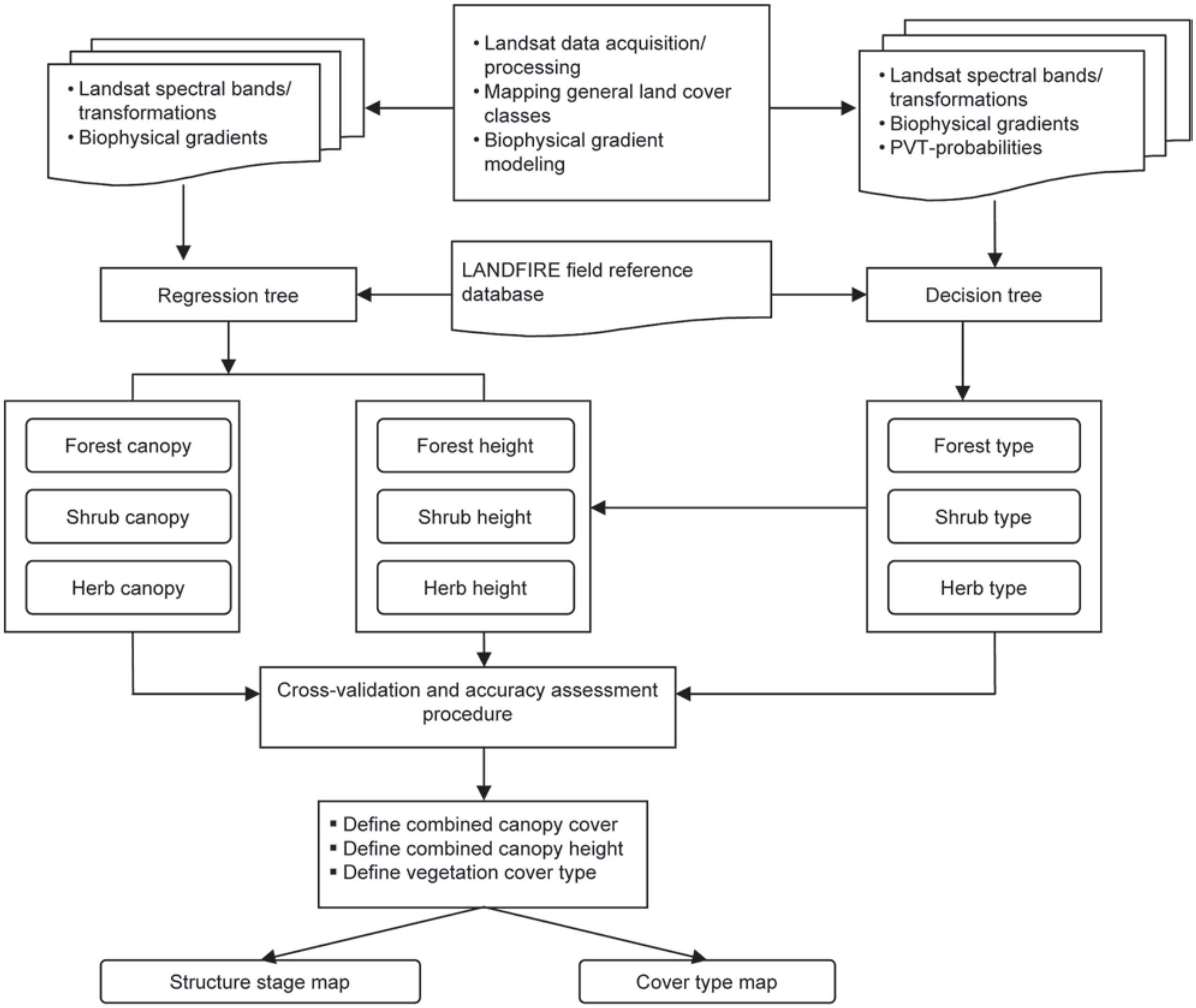

Figure 1—Flow diagram of the methodology used for mapping cover type and vegetation structure in the LANDFIRE Prototype Project.

Use of Biophysical Gradient Variables and Potential Vegetation Maps

In addition to the spectral predictor variables discussed above, the LANDFIRE existing vegetation mapping task incorporated two ancillary data sets that functioned differently in the mapping process. One was a suite of biophysical gradient layers developed as a set of intermediate LANDFIRE products with input from weather, topographic, and soil databases (Holsinger and others, Ch. 5: table 6). Table 1 lists the biophysical gradient variables used in the prototype for mapping existing vegetation; these represent a winnowed set of the entire suite of variables produced for the LANDFIRE Prototype. Biophysical gradients were used in the mapping process to provide a geographic context for the ecological processes that control establishment, growth, and distribution of vegetation communities.

The second data set was a potential vegetation type (PVT) map with attributes describing the probability of specific cover types existing in each PVT. This database was derived by calculating the distribution of cover types within individual PVTs by intersecting the plots contained in the LFRDB with the PVT map (Keane and Rollins, Ch. 3; Frescino and Rollins, Ch. 7). Conceptually, by using the PVT and cover type probability

Table 1—Biophysical and topographic layers used in the LANDFIRE vegetation mapping process.

Symbol	Description	Unit	Source data
SRAD	Daily solar radiation flux	$KW/m^2/Day$	Weather and topographic data
Tmin	Daily minimum temperature	C°	Weather and topographic data
Tmax	Daily maximum temperature	C°	Weather and topographic data
Tnight	Daily average nighttime temperature	C°	Weather and topographic data
Dday	Degree days	C°	Weather and topographic data
PPT	Daily precipitation	cm	Weather and topographic data
RH	Relative humidity	%	Weather and topographic data
PET	Potential evapotranspiration	kgH_2O/yr	Weather and topographic data
AET	Actual evapotranspiration	kgH_2O/yr	Weather, topographic, and soil data
GSWS	Growing season water stress	-Mpa	Weather, topographic, and soil data
PSI	Soil water potential	-Mpa	Weather, topographic, and soil data
KDBI	Keetch-Byram drought index	Index	Weather database
SWF	Soil water fraction	%	Weather, topographic, and soil data
Sdepth	Soil depth to bedrock	cm	Soil and topographic data
LAI	Potential leaf area index	Index	Landsat spectral data
DEM	Digital elevation model	m	National Elevation Database
Slope	Slope	%	National Elevation Database
Aspect	Aspect	Azimuth	National Elevation Database
POSIDX	Topographic position index	Index	National Elevation Database

information in the mapping of vegetation cover types, we implemented a stratification that constrained cover types to the geographic areas where cover types were ecologically possible. Sites (pixels) where certain cover types were not likely to occur would have low probabilities; therefore, these cover types were less likely to be predicted for these pixels. Each cover type was associated with a probability distribution map. The probability layers were implemented in the mapping process much in the same way as the biophysical gradient layers and satellite imagery.

Vegetation Map Unit Classification

Two different approaches were used in the development of the vegetation map unit classification systems for the prototype mapping zones. For the central Utah mapping zone, we formulated the map unit classificaton based on an overall understanding of the presence of vegetation alliances and associations (Long and others, Ch. 6). For the northern Rocky Mountains prototype area, we examined and summarized the LFRDB to form the basis for the vegetation map unit classification. Brohman and Bryant (2005) have described these approaches as the “top-down” and the “bottom-up” approaches, respectively. Long and others (Ch. 6) discuss the criteria and factors used in developing the LANDFIRE vegetation map unit classification systems, the lessons learned in applying them, and recommendations for a national approach to vegetation map unit development.

We were concerned with two technical issues when evaluating the map unit classifications of existing cover types for the prototype: 1) whether each cover type was sufficiently represented by an adequate number of field-referenced data from the LFRDB and, if not, how such “rare map units” should be treated and 2) whether some cover types (such as the Juniper cover type versus the Pinyon – Juniper cover type) would be floristically or ecologically difficult to separate in spectral, biophysical, and geographical domains. The technical issues were considered in the context of four guidelines defined at the beginning of the LANDFIRE Prototype Project: a map unit, whether it is a cover type or a fuel model, must be identifiable, scalable, mappable, and modelable (Keane and Rollins, Ch. 3). Because the prototype study areas were the first mapping zones to be mapped under the LANDFIRE design criteria and guidelines, we were unsure whether the map unit classification systems could perform consistently across different geographic areas.

Reference Data

Caratti (Ch. 4) describes in detail the compilation of the LFRDB for the prototype. The compilation of the LANDFIRE reference database relied on the coordination of three separate and independent efforts: 1) the cooperation and support from the U.S. Forest Service (USFS) Forest Inventory and Analysis (FIA) database collected nationwide on permanent inventory plots (Smith 2002); 2) the collection and processing of existing field data from all land management units such as Bureau of Land Management districts or national parks; and 3) the acquisition of new, supplementary field data from areas where there were no or not enough existing data (for example, various western rangelands in the United States do not currently have adequate field data collection programs).

Because the LFRDB was compiled from various sources collected for different purposes, information gleaned from the LFRDB was highly variable in terms of sampling design. The FIA data represented the most consistent information for forest cover types and canopy height. Rangeland field data usually contained cover type labels, but structure information was rare. In addition, reference data for mapping forest canopy cover were generated by calculating the number of forest cells within a 30-m cell using either high-resolution satellite data (spatial resolution of 1-m or better) or digital orthophotographs (Homer and others 2004).

Quality-control procedures were conducted as a part of the existing vegetation mapping process to detect problems and errors inherent in field-referenced data derived from disparate sources. We assumed that these procedures would identify most existing data problems but would not identify and eliminate all problems. These procedures were as follows:

Detecting outdated field data—Many field plots measured in years past were considered useful if the dominant species had not changed. A substantial number of plots, however, had undergone major disturbances such as fire or logging. We therefore computed the differences between the 1992 and 2001 Landsat NDVI values to flag field plots with conditions that had potentially changed during that 10-year period.

Detecting field data with erroneous geographic coordinates—We identified major geo-coding problems such as coordinates located on roads or located out of mapping areas. We visually examined plot locations overlaid with road networks and general land cover maps (such as NLCD maps).

Detecting field data with major coding errors—We detected such problems by overlaying field data on raw satellite imagery and by sorting variables according to major cover types. For example, if a field plot coded as *sagebrush* was located in the center of an otherwise intact *forest* polygon, or if a shrub plot had a height value taller than that of forest plots, such plots were flagged.

Reducing spatially clumped field plots—The LFRDB contains field data that come from different sources and are collected with different objectives, which occasionally results in spatially clumped plot information. In order to produce a spatially well-distributed and balanced data sample, we sub-sampled clumps of the available data to result in a more even distribution of field data.

The use of these quality-control procedures resulted in the exclusion of a number of available field plots from either the mapping or validation processes. This led to a total of 6,177 field plots (1,809 FIA forest plots and 4,368 non-FIA forest and rangeland plots) for Zone 16 and 7,735 field plots (1,993 FIA forest plots and 5,742 non-FIA forest and rangeland plots) for Zone 19 to be used for subsequent training or accuracy assessment. These numbers differ slightly from other applications of the LFRDB in LANDFRIE mapping because, based on objectives, each mapping effort implemented its own quality control procedure. Although all of the plots contained LANDFIRE cover type labels, only subsets of plots from the LFRDB had attributes of canopy height and canopy cover (table 2). In addition, ten percent of the field data points available for each of the cover type and structure mapping tasks were withheld from the mapping process for the purpose of accuracy assessment (Vogelmann and others, Ch. 13).

Mapping Algorithms

Classification and regression tree algorithms have demonstrated robust and consistent performance and advantages in integrating field data with geospatial data layers (Brown de Colstoun and others 2003; Friedl and Brodley 1997; Hansen and others 2000; Joy and others 2003; Moisen and others 2003, Moore and others 1991; Rollins and others 2004). Nonparametric CART approaches recursively divide feature space into many subsets in a hierarchical fashion to achieve the best overall model performance (lowest error and highest R^2, derived using a cross-validation technique). For this study, we adopted the classification tree algorithm to map vegetation types as discrete map units and the regression tree algorithm to map canopy cover and canopy height as continuous variables using two related

Table 2—Numbers of field reference plots in each mapping zone used in either mapping or accuracy assessment and corresponding to various map products. Forest canopy cover mapping relied on imagery of high spatial resolution instead of field reference plots.

	Mapping zone	Number of cover types	Cover type plots	Canopy cover plots	Canopy height plots
Forest	16	10	1,809	N/A	1,809
	19	14	1,993	N/A	1993
Shrub	16	14	1,595	2,120	1,698
	19	15	1,788	1,788	989
Herbaceous	16	7	300	2,263	1,311
	19	8	597	597	282

commercial applications: See5 (classification trees) and Cubist (regression trees) developed by Quinlan (1993). The mapping models were trained on the compiled data set of spectral bands and biophysical ancillary variables listed in table 1 and cover type and structure variables from the LFRDB.

Vegetation Database Development

Training vegetation mapping models—The creation of the CART-based algorithms for mapping existing vegetation involved several steps: 1) exploration of general data such as correlation analyses and plotting of cover types from the LFRDB against predictor layers, 2) iterations of CART algorithm runs to determine the adequacy of training data and other biophysical layers, 3) visual evaluation of classification and regression trees and final output maps, 4) generation of cross-validation statistics as an initial indicator of map accuracies, and 5) development of vegetation maps by applying the final mapping models. As mentioned above, we withheld data from 10 percent of available field reference plots for accuracy assessment and used the rest of the field plots for training the CART algorithms. We ran classification tree or regression tree classifiers, depending on whether the mapped theme was categorical or continuous, and generated 10-fold cross-validation statistics. Results of the cross-validation were used to determine the quality of training data and the performance of the predictor layers, but not to assess the final accuracy of resulting maps.

Determination of rare and similar map units—Although the LANDFIRE Prototype Project vegetation map unit classifications were developed to meet specific design criteria and guidelines (Keane and Rollins, Ch. 3; Long and others, Ch. 6), two technical questions arose during the mapping of existing vegetation: how to treat 1) rare cover types and 2) spectrally and biophysically similar cover types. We considered a cover type to be rare if it was supported with fewer than 30 reference plots, and those plots were not concentrated in one general location. We retained a rare map unit in the overall mapping process if the resulting spatial pattern made sense (such as when a riparian cover type followed river patterns) and if retaining the map unit did not result in a significant drop in accuracy. Otherwise, the rare map unit would be omitted. Additionally, we decided, based on differences in historical disturbance regimes, to keep cover types that were biophysically and spectrally similar (such as Pinyon – Juniper) separate, even though merging the cover types would significantly improve overall map accuracy.

Stratifications by life form—During the mapping of these vegetation attributes, the question arose as to whether the cover types and structural stages should be constrained by their respective forest, shrub, and herbaceous life forms; that is, we questioned whether a given pixel could be assigned more than one life form for cover type, height, and canopy designations. Multiple life form assignments provided flexibility for the characterization of wildland fuel. Such flexibility would also benefit other potential applications of LANDFIRE data, such as insect and disease or biomass studies. In the process of LANDFIRE vegetation mapping, we therefore modeled each pixel independently for each of the three life forms (forest, shrub, and herbaceous; fig. 1).

Product Validation Plan and Accuracy Assessment

The LANDFIRE accuracy assessment is described in detail in Vogelmann and others (Ch. 13). We tested the approach in which ten percent of the field data points available for cover type mapping were withheld from the mapping process for the purpose of accuracy assessment but found that the approach did not work well because of the uneven availability of field data in support of different cover types in the map unit classification. For several cover types in each of the mapping zones, the amount of data withheld in the 10 percent sample was too low to be statistically meaningful. As the result, we reported overall accuracies for cover types using the results of 10-fold cross-validations. For structure variables, we used a set of independent plots to assess statistical accuracy using regression techniques. This afforded us the opportunity to examine the behaviors of mapping structure variables versus those of categorical variables. Forest canopy cover, mapped with fine-resolution imagery as training data, would be assessed with both a sample of withheld reference points generated from the fine-resolution imagery as well as field estimates obtained from the use of digital cameras equipped with fisheye lenses.

Results

Maps of Cover Type and Structural Stage

We applied the vegetation mapping approach described above to the central Utah and northern Rockies prototype areas. Spectral imagery, biophysical gradients, PVTs, and probabilities were used together with field plot data to produce maps of forest, shrub, and herbaceous cover types, as well as canopy cover and canopy height by life form.

Accuracy of LANDFIRE Prototype Vegetation Mapping

We reported accuracy assessments using a cross-validation approach for cover types by life form (table 3) and by withholding field data for the structure variables by life form (table 4). For cover types, only overall accuracies were reported. For structural stages, R^2 values were variable and ranged from relatively consistent (for forest canopy cover and height) to relatively inconsistent (for shrub and herbaceous canopy cover and height). This variability indicates that forest structure may be mapped reasonably as a continuous variable, whereas consistency and accuracy would be questionable when mapping shrub and herbaceous structure as continuous variables. However, when evaluated as two-class variables (either as closed and open canopy cover or high and low canopy height), results showed that the same shrub and herbaceous structure can perform as consistently and accurately as categorical variables.

Discussion

Analysis of Mapping Consistency for Vegetation Types and Structure

In general, we found that the approach described above for mapping existing vegetation characteristics effectively met LANDFIRE requirements, which was a difficult objective to achieve due to the large number of vegetation map units, reliance on existing field-referenced data, the task of characterizing vegetation structure, and the requirement for a nationally consistent methodology. For the moderately detailed vegetation map unit classification, mapping accuracies of 60 percent or better were achieved at a 30-m spatial resolution.

We explored the mapping of more than two map units for structure variables. For example, we mapped herbaceous height to three map units (0 to 0.5 m, >0.5 to 1 m, and >1 m), shrub height to four map units (0 to 0.5 m, >0.5 to 1 m, >1 to 3 m, and >3 m), and forest height to four map units (0 to 5 m, >5 to 10 m, >10 to 25 m, and >25 m). The tests yielded independent overall accuracies of 73, 61, and 82 percent for herbaceous, shrub, and forest height, respectively. From these results, we concluded that grouping continuous values of the structure variables into several discrete map units would be an acceptable and rational alternative methodology for national implementation of the LANDFIRE methods. Use of this alternative methodology would require the development of a consistent national structural stage map unit classification.

Table 3—Cross validations (10 percent withheld, ten-fold repetitions) conducted separately by mapping zones and by forest, shrub, and herbaceous life forms.

Life form	Mapping zone	Number of classes	Cross validation
Forest	16	10	67
	19	14	64
Shrub	16	14	62
	19	15	68
Herbaceous	16	7	60
	19	8	56

Table 4—Accuracy assessments conducted separately for two structure variables by life forms and map zones. Overall accuracy (OA) was obtained by using holdout withheld field plots (n) that were set aside based on quality and distribution of the total available field plot data (N). Structure variables are treated as both continuous variables measured with the R2 statistic and two-class categorical variables for overall accuracy (OA). The Ttwo canopy cover classes of canopy cover are closed (≥40%) and open (<40%); for canopy height they classes are high (≥10m, 1m, 0.24m) and low (<10m, 1m, 0.24m) for forest, shrub, and herbaceous life forms, respectively.

Life form	Map zone	Canopy cover n/N	Canopy cover R^2	Canopy cover Overall accuracy	Canopy height n/N	Canopy height R^2	Canopy height Overall accuracy
Forest	16	1,272/20,000	0.78	0.92	220/2204	0.58	0.88
	19	1,200/20,000	0.88	0.89	127/5,541	0.56	0.78
Shrub	16	125/1,253	0.41	0.74	107/1,073	0.36	0.85
	19	119/1,788	0.59	0.79	81/989	0.65	0.86
Herbaceous	16	18/182	0.37	0.71	15/280	0.04	0.86
	19	126/597	0.58	0.69	75/182	0.63	0.70

Consistency in field sampling and data collection affects the consistency of mapping vegetation characteristics. Of the three types of reference data used in mapping existing vegetation, cover type and canopy height values can generally be identified or measured consistently in the field. Canopy cover, on the other hand, can be difficult to measure in the field. This issue does not affect the measurement of forest canopy cover values because training data are derived from high-resolution (1 m or better) imagery by calculating numbers of high-resolution forest pixels within each 30-m Landsat pixel. The use of inconsistently estimated canopy cover values as training data, however, can potentially affect the mapping of shrub and herbaceous canopy percent cover (as happened during the prototype). Shrub and herbaceous canopy results from the two prototype mapping zones were reasonable (table 3), but difficulties in consistently estimating canopy cover in the field indicated that we needed to further research new or alternative methods for mapping shrub and herbaceous canopy cover.

The results of this study may be attributed, in part, to the use of ecologically significant ancillary data layers, which accounts for a moderate but nonetheless significant increase in accuracy (ranging from 1 to 9 percent). The development of biophysical gradient layers and PVT probabilities follows a standardized process for all mapping zones. However, for any given area, satellite reflectance can vary significantly for the same cover type with different canopy cover percentages (either due to land management practices or regeneration stages) or appear similar for different vegetation types or different structural stages during certain seasonal periods. Different cover types or structural stages, however, should respond consistently to the effects of biophysical gradient variables such as soil depth or potential evapotranspiration (PET); this addition of information from the biophysical gradient variables increases the likelihood that these map units will be discriminated by mapping algorithms. For example, one might expect Engelmann spruce (*picea engelmannii*) to grow in relatively deep soil on cool, north-facing sites with low PET, regardless of whether it is found in Zone 16 or Zone 19. Therefore, the incorporation of biophysical and PVT data in the mapping process should contribute to enhanced consistency and thematic accuracy in mapped existing vegetation across the United States.

Even though the existing vegetation maps shown in figures 2 and 3 characterize the vegetation composition of all life forms, it should be noted that each life form was mapped independently, by design, for cover type and structure. Modeling life forms independently preserves the possibility of more than one mapped life form per pixel (in other words, allows for probabilities of multiple canopy layers within a pixel) to improve fuel mapping and enhance the range of the data's ecological applications. However, mapping approaches should be carefully considered when comparing or merging these separate data sets. For example, a final map of cover types may look different depending on the order of precedence between forest, shrub, and herbaceous cover and the threshold values used in defining the life forms (for example, a pixel with 10 percent or greater forest canopy cover may be considered as forested land). It is important that precedence and thresholds be applied uniformly between mapping zones for consistency.

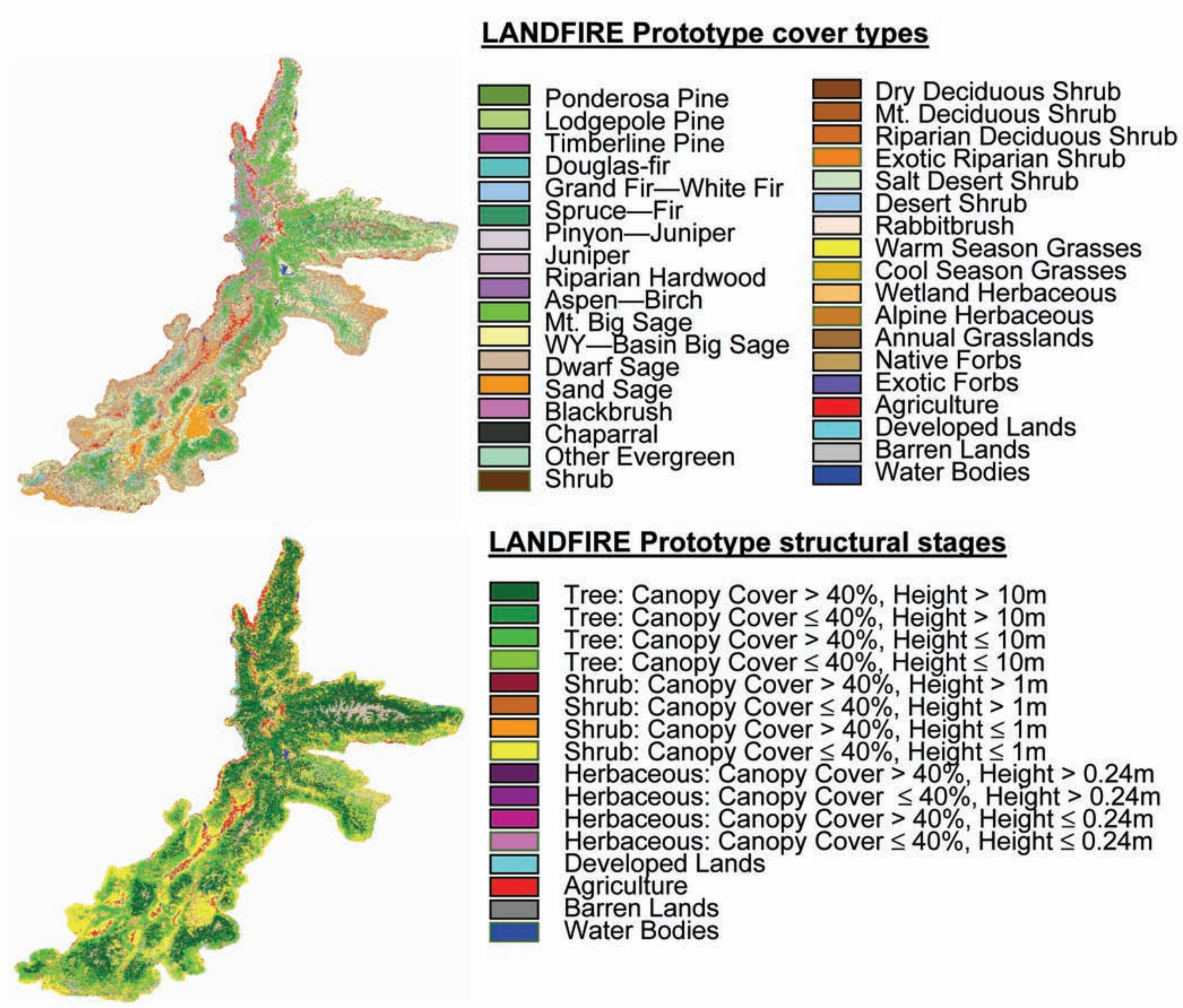

Figure 2—LANDFIRE Prototype cover type (top) and structural stage (bottom) maps for Zone 16. The cover type map is compiled from separate forest, shrub, and herbaceous cover type maps, whereas the structural stage map is grouped from continuous maps of height and cover for display purposes.

Factors that Affect Mapping Accuracies

Several factors should be considered when examining the accuracy estimates for maps of cover types and structure. First, the mapping and accuracy assessment of cover type and structure variables by life form were conducted based on field-referenced databases of different sizes and data collected throughout the study areas using a variety of sampling strategies. As would be expected, vegetation mapping was sensitive to the availability of field data. Test results showed that the number of field-referenced plots used for mapping and accuracy assessment affected not only the level but also the consistency of mapping accuracies, with fewer plots related to greater variability in accuracy estimates and more plots to more robust accuracy estimates (fig. 4). Data for herbaceous vegetation were limited in availability relative to the overall size of the field-referenced data set and hence affected herbaceous mapping accuracy. To improve uncertainties related to shrub and herbaceous cover and height, we determined that these variables should be mapped as categorical map units.

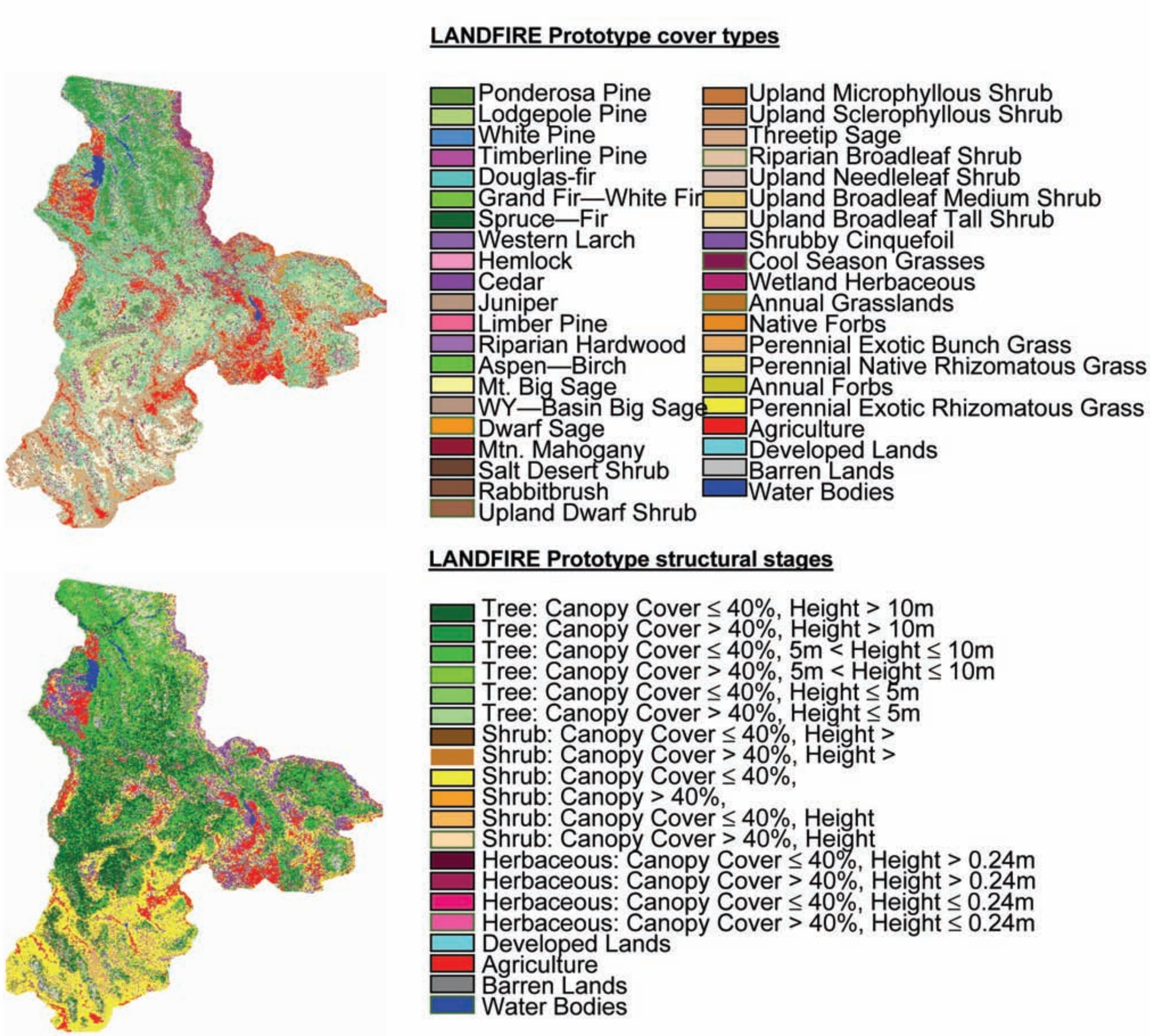

Figure 3—LANDFIRE Prototype cover type (top) and structural stage (bottom) maps for Zone 19. The cover type map is compiled from separate forest, shrub, and herbaceous cover type maps, whereas the structural stage map is grouped from continuous maps of height and cover for display purposes.

Second, field-referenced data, with which mapping models were trained and accuracy assessed, were collected from different sources, for different objectives, and with different techniques. Even though these plot data were quality-screened and standardized through an extensive effort (Caratti and others, Ch. 4), it was inevitable that the differences and errors in field data carried over into map quality and accuracy assessment. For example, certain reference data for forest canopy cover were derived using digital ortho-photographs, viewing forest cover synoptically from above the canopy. On the other hand, field estimates for shrub and herbaceous canopy cover were made using visual estimation from close-range, oblique positions that limited objectivity and consistency. We did not experience these problems when determining forest, shrub, and herbaceous *height*, which was usually directly measured and had a high degree of user-confidence.

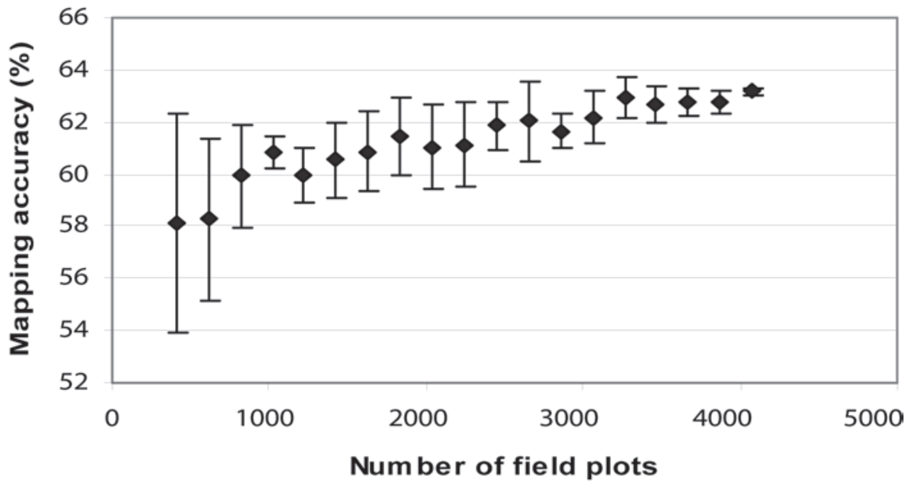

Figure 4—Cross-validation accuracy estimates obtained for the mapping of forest cover types as a function of the number of forest field plots. More plots contributed to better accuracy and consistency (smaller standard deviation) to a certain point, after which the relationship became flat.

Third, as discussed above, rare map units and ecologically and biophysically similar map units affected mapping accuracies. For example, if the Juniper cover type was merged with the combined Pinyon – Juniper cover type, forest cover type accuracy increased by more than 10 percent. The rationale for keeping such similar cover types separate is that, even though they occupy similar ecological niches and have similar site characteristics, separating them increases the utility of the LANDFIRE wildland fuel and fire regime products.

Utility of Biophysical Gradient Data for Vegetation Mapping

Although the use of DEM data for improving mapping results has been widely documented, the effects of a whole host of biophysical gradient layers and PVT-probability data layers is largely untested at the scale and scope of this study. These data layers provide information that supplements satellite imagery. Plant distribution patterns and conditions are strongly linked to a multitude of environmental factors (for example, temperature, soil, weather patterns, day length, soil properties, and rainfall), and the accurate characterization of these variables should, at least in theory, improve mapping results. In addition, spatial information that indicates where particular vegetation types can and cannot exist across a wide region (that is, PVT-probability data layers) should be similarly useful. Figure 5 compares cross-validation results using mapping models with and without the additional biophysical gradients listed in table 1 and using PVT-probabilities as predictor variables. Figure 6 displays mean and standard deviation values of a subset of the biophysical variables intersected with vegetation cover types from field plot data collected in the central Utah prototype area. These figures show that the incorporation of certain biophysical gradients and PVT-probabilities in mapping models contributes to increased mapping accuracy and consistency. These results are consistent with the findings of Keane and others (2002) and Rollins and others (2004).

Vegetation Patterns in Areas of Major Disturbances

Wildfires, insect and disease outbreaks, and forest clear cuts are some of the major disturbances to ecosystems captured by the satellite sensor in terms of their spectral properties. How well did our mapping capture and reflect these changes in vegetation conditions? We evaluated our mapping methods' effectiveness in this regard by looking at known areas of wildland fire, bark beetle infestation, and clear-cuts in the prototype mapping zones.

We evaluated two wildland fires areas that burned in Bryce (summer 2001) and Zion (fall 2001) national parks to determine what differences might exist between pre-fire and post-fire vegetation maps when mapped with the same pre-fire models. Pre- and post-fire map comparisons showed distinct differences between both

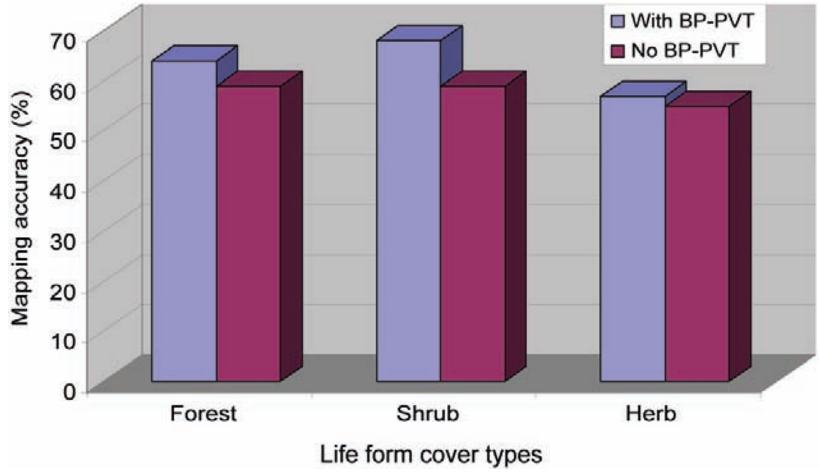

Figure 5—Cross-validation accuracy estimates obtained in the Zone 16 prototype area, by life form, with and without the 15 biophysical gradients and PVT-probabilities in the mapping models. An average of 8 percent increase in cross-validation accuracy was obtained by incorporating the selected biophysical gradients and PVT-probabilities that together describe the habitats of the cover types to be mapped.

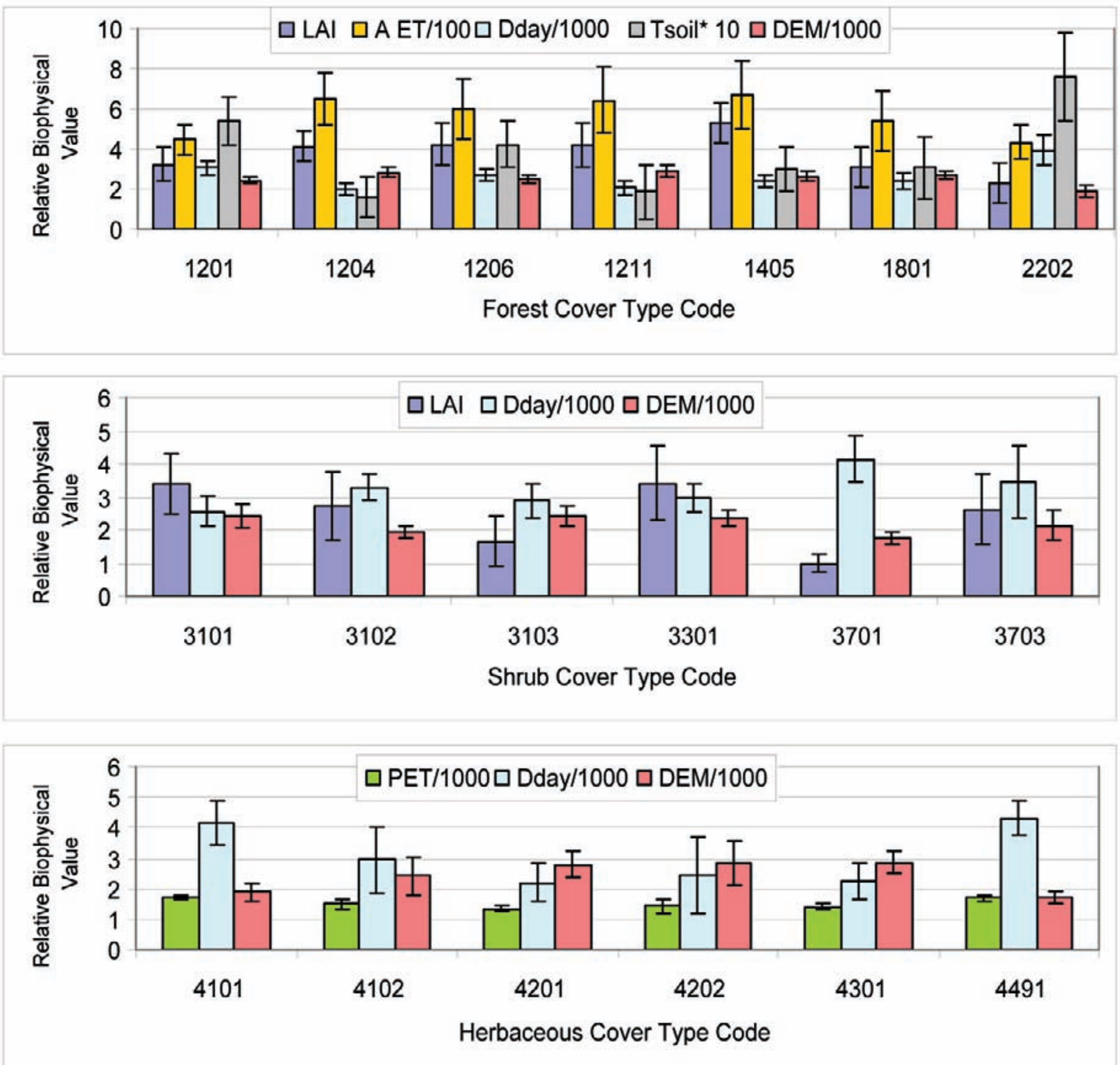

Figure 6—Mean and standard deviation values of selected biophysical variables found effective in mapping cover type against various forest (top), shrub (middle), and herbaceous (bottom) cover types of field-reference data. Most of the biophysical variables were divided or multiplied by a constant for display purposes. Refer to table 1 for definitions and descriptions of the biophysical variables. Refer to Long and others, Ch. 6 appendix 6-A for vegetation cover type coding protocol.

vegetation cover types and structural stages. The Bryce fire, a prescribed fire, showed general shifts from forest to shrub map units and, regarding structure, showed a shift toward increased low-height shrubs. The Zion fire, a wildfire, revealed a shift from predominately deciduous forest types to low shrubs.

Using bark beetle survey data obtained from the Dixie National Forest, we conducted simple zonal statistical analyses. Results indicated that the mapped species composition corresponded fairly well to that of those species identified in the survey data for the years 1998-2000. (Note that the level of actual disturbance varied within the survey data and was not differentiated in this study.) Structure information was not available in the survey data, but mapped structure data indicated that most bark beetle infestations occurred in areas identified as high forest cover (greater than 40 percent canopy cover) and height (greater than 10 m), indicating old-growth forest.

Similarly, we compared clear-cut areas, identified using modeling and masking methods, with mapped vegetation cover type and structure variables. Shrubs and a high percentage of grasses were dominant in clear-cut areas. Structural stages indicated a trend from forests with high canopy cover and canopy height to a high percentage of low cover (less than 40 percent), low height (less than 1 m) shrubs. Herbaceous cover was identified as being high cover (greater than 40 percent) with mixed heights.

Field Data Quality and Quantity Requirements

The acquisition of field-referenced data posed a significant challenge to the LANDFIRE Prototype effort, both logistically and technically. Caratti (Ch. 4) describes the logistical efforts and complications associated with conducting a national field data campaign. Specifically, technical challenges encountered during the mapping process, such as uneven amounts and disparate quality of field data used to meet various vegetation mapping objectives, were tied to the fact that the LFRDB was based on data from varying sources and collected with different objectives. As discussed above, such issues necessitated the careful implementation of a quality-control and quality-assurance (QA/QC) process prior to the training of the mapping algorithms for existing vegetation types and structure. “Lessons learned” from the QA/QC process follow:

- Accuracy and consistency are a function of the amount of available field-referenced data. Greater amounts of field-referenced data contribute to enhanced confidence in mapping accuracy (fig. 4), whereas limited field-referenced data are correlated to reduced confidence in mapping accuracies of affected cover types.
- The use of data from different sources requires that special attention be given to those cover type map units that are not supported with sufficient numbers of field plots. Both prototype mapping zones had map units with only a few field reference data points for training. As discussed above, the question of how to define and treat rare map units arose during the prototype, and we defined rare map units as those having less than 30 field reference plots scattered spatially within a mapping zone. Options for the treatment of these rare map units included keeping the map units in maps, omitting them, or omitting them and then “burning” the few field plots to the map in a post-process and merging them with floristically similar cover types. For the prototype, we chose to retain the rare map units in the models and resulting map products to inform the development of the LANDFIRE vegetation map unit classification system. For national implementation, rare map units that cannot be supported with a sufficient number of field plots will not attain target-level accuracies. We recommend omitting such map units from the mapping of existing vegetation cover types.
- Spatial distribution and a valid probability-based sampling design increase the consistency and accuracy of the map products. Compared with field-referenced data from various agency sources, the use of FIA forest inventory plots for mapping forest cover types and structure produced more consistent and accurate mapping results because the sampling design for FIA data produced training data that were spatially well-distributed across the landscape. Further, FIA data required very little additional processing time and were easy to use; in contrast, non-FIA field data required extensive processing time, related to QA/QC and re-selecting/re-sampling, to derive suitable data sets (in terms of spatial distribution and data quality) from available data points. For example, in Zone 19, a Bureau of Land Management study produced more than 4,800 field plots, mostly describing sagebrush, Douglas-fir, and lodgepole pine vegetation communities, in a relatively small area of approximately 1,152 km^2, near Salmon, ID. Spatially, this data set equated to approximately one plot for every 24 ha, versus a mapping zone average of one plot for every 835 ha. The inclusion of this data set in the

training process overwhelmed the mapping models and overrode areas with sparse plot coverage of different cover types. We therefore determined that the application of locally limited or concentrated data collected using various sampling designs to an entire mapping zone could have adverse effects on the accuracy of final products. For this reason, forest mapping in LANDFIRE National should employ FIA data exclusively. Rangeland mapping in LANDFIRE National, however, will require extensive QA/QC processing steps to transform available field-referenced data to a more suitable data set.

- As noted above, the following critical steps should be taken prior to the development of the mapping models: 1) examine field-referenced data, 2) conduct QA/QC procedures to detect spatial errors as well as information content-related errors, 3) correct these errors if necessary, and 4) derive a final, refined, error-free data set for training and accuracy assessment. This is a time-consuming yet necessary process that will contribute to increased consistency and confidence of map products.

Effects of the Vegetation Map Unit Classification System

Determining accuracy objectives and the appropriate extent of mapping areas are among the factors that need to be considered when defining a workable national vegetation map unit classification system. If floristically or ecologically overlapping cover types (such as Juniper and Pinyon -- Juniper or Upland Microphyllous and Upland Sclerophyllous) are to be mapped for LANDFIRE National, then guidelines must be developed for defining how the mapping accuracy of such overlapping map units is to be assessed.

Next, although our use of the NVCS was a reasonable starting point for vegetation map unit classification and the approach worked fine for each individual mapping zone, vegetation cover types were not always comparable between the two prototype mapping areas, however, as is evidenced by the legends in figures 2 and 3. As a result, accuracy estimates for the two prototype mapping zones could not be compared in a straightforward fashion, particularly for shrub cover types.

As discussed above, another challenge encountered during the application of the two vegetation map unit classification approaches (as discussed above in the *Vegetation Map Unit Classification* section) was answering the question of how to treat rare map units. There were no guidelines for consistently defining and treating rare map units. Moreover, there was no answer as to whether dropping rare map units, instead of using the alternative options discussed above, might affect the utility of LANDFIRE vegetation maps in other future natural resource management projects.

Recommendations for National Implementation

Because of the size and complexity of this research effort, many questions concerning LANDFIRE's national implementation are as of yet unanswered. The field data compilation effort will be an expensive and time consuming task, and a pressing need exists regarding the study of links between mapping performance, resource expenditure, and methods of field data collection. Ecological relationships between mapped potential vegetation and existing vegetation need to be investigated. Further research must be conducted to quantify the relative contributions of the different approaches and data sets used in the prototype. Performance consistency must be tested between adjacent western mapping zones, as well as in one or more prototype areas located in the eastern United States. Repeatability of the methods used in the prototype, both temporally and spatially, must also be evaluated. Furthermore, it is not clear whether the LANDFIRE Prototype methodology will suffice for other vegetation metrics, such as quantifying woody or non-woody biomass; a study in this area could yield information leading to enhanced applications of LANDFIRE vegetation maps. Nevertheless, the LANDFIRE Prototype Project provides sufficient information on which to base several recommendations regarding the national implementation of LANDFIRE.

Ways to Ensure Consistent National Vegetation Mapping

As noted above, several tasks related to existing vegetation mapping for the prototype effort may be standardized and potentially automated to facilitate LANDFIRE's national implementation. These tasks include: 1) the creation of a national vegetation map unit classification system that is mappable using spectral and biophysical/ecological data and is supported with adequate field-referenced data; 2) the consistent acquisition and processing of a multi-seasonal Landsat database; 3) the application of QA/QC procedures to the LFRDB to ensure a robust field-referenced database that can be used for a wide variety of applications; 4) the consistent modeling of biophysical data layers and probabilities of existing vegetation species or types associated with potential vegetation types; and 5) the continued application of CART as the primary mapping algorithms to ensure objectivity and flexibility when using high volumes of field data and predictor variables. We discuss these points in detail below.

Need for a Mappable Vegetation Map Unit Classification System

The vegetation map unit classification system used for the national implementation of LANDFIRE must meet a number of key criteria including the following: 1) the system must be nationally consistent, ecologically logical and hierarchical, acceptable to a wide array of users and groups, and must meet existing Federal Geographic Data Committee (FGDC) standards; 2) vegetation map units must be mappable using operational methodology to achieve reasonable accuracies; and 3) the map unit classification system must include vegetation map units that have high relevance with respect to the core LANDFIRE products. The Ecological Systems classification (Comer and others 2003) developed by NatureServe meets these objectives. This system represents the hierarchical merging of NVCS alliances into a nationally available suite of vegetation map units. Unlike alliances, which have proved exceedingly difficult to map accurately, most Ecological Systems classes are mappable, assuming an adequate number of field plots exist for training purposes. In addition, the Ecological Systems classification was developed by plant ecologists, lending credibility to the approach and resulting in a greater level of acceptance throughout the user community. We anticipate that a few additional "target alliance" map units will be added to the LANDFIRE National map unit classification legend on a case-by-case basis. These will be added only when it is determined that a particular map unit not specifically identified by the Ecological Systems classification has special relevance to LANDFIRE.

Need for National Field-referenced Data Collection and Processing

Many LANDFIRE tasks rely on a comprehensive, consistent, and extensive field-referenced database. The database serves as a reference for the development, testing, and accuracy assessment of all LANDFIRE vegetation, biophysical settings, and wildland fuel data layers and of all vegetation and fire regime simulation models. Field data from existing projects should be incorporated into this database whenever available and should include but not be limited to data sets such as FIREMON fire monitoring databases, USFS Landscape Ecosystem Inventory Systems databases, and the National Park Service fire monitoring databases. In addition, the USFS FIA Program's forest inventory plot database proved a useful source for the majority of forest data. Where data are lacking, supplemental field data collection is required to fill informational needs on rangeland map units. This assortment of field-referenced data should be collectively scrutinized for quality assurance, regularly updated, and maintained as a comprehensive LANDFIRE field-referenced database.

Need for Nationally Consistent Imagery Database

The availability of a quality Landsat imagery catalog is a key prerequisite for national implementation of the approaches developed for the LANDFIRE Prototype Project. Among all predictor variables, it is satellite imagery that usually captures the most current vegetation conditions, and, when used repeatedly over time, identifies changes in vegetation conditions and distributions. Thus, we recommend that LANDFIRE National continue to play an active role in the MRLC Consortium. This membership ensures the continued development of suitable multi-seasonal Landsat image catalogs, optimal levels of image processing (geometric, radiometric, and atmospheric rectification and calibration) for the rest of the country, and mapping zone-based image compilation for national vegetation mapping. In addition, LANDFIRE National should support studies that examine and compare the characteristics of other mid-resolution sensors with those of Landsat. Even though the LANDFIRE Project does not currently require any additional Landsat imagery, the potential benefits of using different satellite data for future updating should be considered.

Need for Nationally Consistent Set of Biophysical Gradient Layers

Biophysical gradients have effects similar to that of Landsat imagery on the spatial and information integrity of existing vegetation maps. Many of the biophysical layers are physiologically and ecologically related to the establishment, distribution, and conditions of plant species, and the incorporation of these gradient layers into the mapping process contributes to increased accuracies. For the national implementation of LANDFIRE, we recommend that a set of biophysical gradient layers similar to those listed in table 1 be used to map vegetation in all mapping zones. In addition, we recommend that further research be conducted to quantify the contribution of the individual biophysical variables to mapping accuracy. Furthermore, research should be conducted to minimize residual coarse-resolution imprints in 30-m biophysical data resulting from the coarser resolution weather and soil databases used to produce these data. The development of standard minimum mapping units in modeling simulations has shown promise in standardizing the process and eliminating coarse imprints.

The Need to Continue with Research and Improvements

Although results of the LANDFIRE Prototype Project indicate that the general approach should effectively meet target accuracy and consistency requirements for national implementation, there are areas where continued research and improvements are needed. One ongoing research effort involves the development of a new and more consistent approach to mapping shrub and herbaceous canopy cover. Current research is testing ways to effectively correlate calibrated Landsat-based NDVI to shrub and herbaceous canopy cover (Liu and others 2004). Other research areas include more efficient use of the individual biophysical gradient layers, more effective mapping of riparian vegetation, and a national accuracy assessment strategy.

Conclusion

The mapping of existing vegetation with complete national coverage at a 30-m spatial resolution is a core requirement of the LANDFIRE Project. National data at this 30-m resolution do not currently exist. As a result, the prototype research was needed to answer questions related to the mapping and characterizing of cover types and structure variables. LANDFIRE's existing vegetation products are expected to provide data not only for use in wildland fire management, but also for use in many other natural resource and environmental applications. Findings from the LANDFIRE Prototype effort are summarized as follows:

If supported with an adequate amount of field-referenced data, target accuracies of 60 percent or better are achievable for a mid-level vegetation map unit classification at the regional scale. The addition or subtraction of floristically or ecologically similar cover types has significant effects on resulting accuracies. Of the three major life forms, herbaceous cover types are the most difficult to map because these species adapt to many general biophysical characteristics and have few unique spectral signatures. Relationships between the floristic complexity of the vegetation map unit classification and mapping accuracies indicate that the national vegetation map unit classification will need to be designed carefully to include adequate flexibility.

For LANDIFRE, vegetation structure is defined by canopy cover and canopy height of forest, shrub, and herbaceous life forms. These structure attributes can be mapped consistently as categorical variables. Mapping these attributes as continuous variables, particularly for shrub and herbaceous height and cover, is inconsistent and, thus, is not recommended for national implementation of the LANDFIRE prototype methods.

Field data collection and processing are the most critical factors in ensuring that LANDFIRE maps of existing vegetation are objective and accurate. The detection and correction of errors existing in field-referenced data are time-consuming but absolutely necessary tasks, particularly for field data from sources other than FIA (as these other data sets tend to be locally limited and have various sampling designs). The objective of repeated field data processing and quality control is to derive a refined, high-quality field data set.

The incorporation of LANDFIRE biophysical gradient layers and cover-type probabilities associated with potential vegetation types into the mapping models contributes to a significant increase in mapping accuracy. In addition, the use of the biophysical and ecological stratifications that describe the environmental effects on species establishment and growth also contributes to enhanced mapping consistency.

For further project information, please visit the LANDFIRE website at www.landfire.gov.

The Authors

Zhiliang Zhu is a Research Physical Scientist with the DOI USGS Center for Earth Resources Observation and Science (EROS). Zhu's research work focuses on mapping and characterizing large-area land and vegetation cover, studying land cover and land use change, and developing remote sensing methods for the characterization of fuel and burn severity. His role in the LANDFIRE Prototype Project has been to design and test a methodology for the mapping of existing vegetation cover types and vegetation structure and to direct research and problem-solving for all aspects of the methodology. He received his B.S. degree in Forestry in 1982 from the Nanjing Forestry University in China, his M.S. degree in Remote Sensing in 1985, and his Ph.D. degree in Natural Resources Management in 1989, both from the University of Michigan.

James Vogelmann is a Principal Scientist with the Science Applications International Corporation (SAIC), contracting with the DOI USGS Center for Earth Resources Observation and Science (EROS). Vogelmann's research work focuses on large-region land cover characterization and change assessment using remote sensing and ancillary sources of spatial data. His roles in the LANDFIRE Prototype Project have been to assess different methods for mapping vegetation types, to serve on the LANDFIRE Vegetation Working Group, and to help direct project research activities. He received his B.A. degree in Botany from the University of Vermont in 1978 and his Ph.D. degree in Plant Sciences from Indiana University in 1983.

Donald Ohlen is a Environmental Scientist for the Science Applications International Corporation (SAIC) at the DOI USGS Center for Earth Resources Observation and Science (EROS), 47914 252nd Street, Sioux Falls, SD 57198; phone: (60. Ohlen's research work and interest focus on land cover mapping for fire science applications, including the characterization of satellite data for fuel mapping and post-fire burn mapping. He earned his B.S. (1976) and M.S. (2000) degrees in Geography from South Dakota State University.

Jay Kost is a Research Physical Scientist with the Science Applications International Corporation (SAIC), contracting with the DOI USGS Center for Earth Resources Observation and Science (EROS). Kost's work focuses primarily on mapping existing vegetation and vegetation structure (percent canopy and height) for the LANDFIRE Project using decision and regression tree models. Optimization of these models and high map accuracy results are paramount in his work and improvements in methodology and results are continually pursued. He received his B.S. in Electronic Engineering Technology in 1987 from Minnesota State-Mankato and his M.S. degree in Space Studies from the University of North Dakota, Grand Forks. In addition, Kost has completed four years of post-graduate study in the Atmospheric, Environmental, and Water Resources Ph.D. program at South Dakota State University, Brookings.

Xuexia (Sherry) Chen is an Environmental Scientist with the Science Applications International Corporation (SAIC), contracting with the DOI USGS Center for Earth Resources Observation and Science (EROS). Chen's research work focuses on vegetation mapping of canopy cover and canopy structure. Her role in the LANDFIRE Prototype Project has been to develop a hierarchical methodology for the mapping of existing vegetation cover types and vegetation structure and to explore new technologies for accuracy and efficiency improvements in LANDFIRE products. She received her B.S. degree in Geography in 1997 and her M.S. degree in Environmental Sciences in 2000, both from Peking University, China. Chen received her Ph.D. degree in Atmospheric, Environmental, and Water Resources in 2004 from the South Dakota School of Mines and Technology.

Brian Tolk is a Research Scientist with the Science Applications International Corporation (SAIC), contracting with the DOI USGS Center for Earth Resources Observation and Science (EROS). Tolk's research work focuses on the mapping and characterization of large-area land and vegetation cover and on the use of close-range remote sensing methods to aid and improve LANDFIRE mapping techniques. His role in LANDFIRE has been to map land cover and structure variables for the prototype zones, implement a data management scheme, and produce promotional products for the project. He received his B.A. degree in Geography from Augustana College, Sioux Falls in 1990 and his M.A. degree in Geography from the University of Nebraska, Lincoln in 1996.

Acknowledgments

We acknowledge the funding provided by the USDOI Office of Wildland Fire Coordination and the USFS Fire and Aviation Management. In addition, the cooperation of the USFS FIA office in Ogden, Utah, is gratefully acknowledged. We also wish to acknowledge the help and support provided by the staff at the USDA Forest Service Fire Sciences Laboratory in Missoula, Montana and by the staff at Utah State University in Logan, Utah. The research was performed in part by the Science Application International Corporation under U.S. Geological Survey Contract 1434-CR-97-40274 and 03CRCN0001.

References

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. U.S. Geological Survey Professional Paper 964, USGS Center for Earth Resources Observation and Science, Sioux Falls, SD. 28 p.

Balice, R.G.; Oswald, B.P.; Yool, S.R. 2000. Fuels inventories and spatial modeling of fire hazards in the Los Alamos region. In: The Joint Fire Science Conference and Workshop: Crossing the Millennium: Integrating Spatial Technologies and Ecological Principles for a New Age in Fire Management. Boise, ID. Moscow, ID: University of Idaho Press.

Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. 1984. Classification and Regression Trees. Pacific Grove, CA: Wadsworth & Brooks/Cole Advanced Books & Software. 358 p.

Brohman, R.; Bryant, L., eds. 2005. Existing Vegetation Classification Mapping and Technical Guide. Gen Tech. Rep. WO-67. Washington, DC: U.S. Department of Agriculture Forest Service, Ecosystem Management Coordination Staff. 305 p.

Brown de Costoun, E.C.; Story, M.H.; Thompson, C.; Commisso, K.; Smith, T.G.; Irons, J.R. 2003. National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. Remote Sensing of Environment. 85:316-327.

Comer, P.; Faber-Langendoen, D.; Evans, R.; Gawler, S.; Josse, C.; Kittel, G.; Menard, S.; Pyne, M.; Reid, M.; Schulz, K.; Snow, K.; and Teague. J. 2003. Ecological Systems of the United States: A Working Classification of U.S. Terrestrial Systems. Arlington, Virginia: NatureServe. 83 p.

Eve, M.; Merchant, J. 1998. A national survey of land cover mapping protocols used in the GAP Analysis Program. [Online] Available: http://www.calmit.unl.edu/gapmap/report.htm [May 12, 2006].

Friedl, M.A.; Brodley, C.E. 1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment. 61:399-409.

Friedl, M.A.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A.; Baccini, A.; Gao, F.; Schaaf, C.; McIver, D.K.; Hodges, J.C.F. 2002. Global land cover mapping from MODIS: Algorithms and early results. Remote Sensing of Environment. 83(1-2): 287-302.

Gemmell, F.M. 1995. Effects of forest cover, terrain, and scale on timber volume estimation with Thematic Mapper data in a rocky mountain site. Remote Sensing of Environment. 51(2): 291-305.

Grossman, D.H.; Faber-Langendoen, D.; Weakley, A.S.; Anderson, M.; Bourgeron, P.; Crawford, R.; Goodin, K.; Landaal, S.; Metzler, K.; Patterson, K.; Pyne, M.; Reid, M.; Sneddon, L. 1998. International classification of ecological communities: terrestrial vegetation of the United States. Volume I. The national vegetation classification system: development, status, and applications. Arlington, VA: The Nature Conservancy. 127 p.

Hansen, M.; Defries, R.S.; Townshend, J.R.G.; Sohlberg, R. 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. International Journal of Remote Sensing. 21:1331-1364.

Homer, C.G.; Huang, C.; Yang, L.; Wylie, B.K.; and Coan, M. 2004. Development of a 2001 national land-cover database for the United States. Photogrammetric Engineering & Remote Sensing. 70(7): 829-840.

Huang, C., L.; Yang, L.; Wylie, B.; Homer, C. 2001. A strategy for estimating tree canopy density using Landsat 7 TM+ and high resolution images over large areas. In: Proceedings of the Third International Conference on Geospatial Information in Agriculture and Forestry, Denver, CO. [Online]. Available: http://landcover.usgs.gov/pdf/canopy_density.pdf [5-12-2006]

Huang, C.; Wylie, B.; Yang, L.; Homer, C.; and Zylstra, G. 2002. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. International Journal of Remote Sensing. 23(8): 1741-1748.

Huang, C.; Townshend, J.R.G. 2003. A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover. International Journal of Remote Sensing. 24(1): 75-90.

Joy, S.M.; Reich, R.M.; Reynolds, R.T. 2003. A non-parametric, supervised classification of vegetation types on the Kaibab National Forest using decision trees. International Journal of Remote Sensing. 24(9):1835-1852.

Kalliola, R.; and Syrjanen, K.S.; 1991. To what extent are vegetation types visible in satellite imagery? Annals Botany Fennici. 28: 45-57.

Keane, R.E.; Burgan, R.; van Wagtendonk, J.V. 2001. Mapping wildland fuels for fire management across multiple scales: integrating remote sensing, GIS, and biophysical modeling. International Journal of Wildland Fire. 10: 301-319.

Keane, R. E.; Rollins, M. G.; McNicoll, C.; Parsons, R. A. 2002. Predictive landscape modeling using gradient-based sampling, remote sensing, and ecosystem simulation. Gen. Tech. Rep. RMRS-GTR-92. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Liu, Y.; Zha, Y.; Gao, J.; Ni, S. 2004. Assessment of grassland degradation near Lake Qinghai, West China, using Landsat TM and in situ reflectance spectra data. International Journal of Remote Sensing. 25(20): 4177-4189.

Mahesh, P.; Mather, P.M. 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment. 86(4): 554-565.

Moisen, G.G.; Frescino, T.S.; Huang, C.; Vogelmann, J.; Zhu, Z. 2003. Predictive modeling of forest cover type and tree canopy height in the central Rocky Mountains of Utah. Proceedings of the 2003 meeting of the American Society for Photogrammetry and Remote Sensing, Anchorage Alaska. Washington D.C.: American Society for Photogrammetry and Remote Sensing.

Moore, D.M.; Lees, B.G.; Davey, S.M. 1991. A new method for predicting vegetation distributions using decision tree analysis in a geographical information system. Environmental Management. 15(1):59-71.

Quinlan, J.R. 1993. C4.5: programs for machine learning. San Mateo, CA: Morgan Kaufman Publishers.

Rollins, M.G.; Keane, R.E.; Parsons, R.A. 2004. Mapping fuels and fire regimes using remote sensing, ecosystem simulation, and gradient modeling. Ecological Applications. 14(1): 75-95.

Smith, W.B. 2002. Forest inventory and analysis: a national inventory and monitoring program. Environmental Pollution. 116: S233-S242.

Yang, L.; Huang, C.; Wylie, B.; Homer, C.; and Coan, M. 2003. An approach for mapping large-area impervious surfaces: synergistic use of Landsat 7 ETM+ and high spatial resolution imagery. Canadian Journal of Remote Sensing. 29(2): 230-240.