Rocky Mountains Cooperative Ecosystem Studies Unit (RM-CESU)
RM-CESU Cooperative Agreement Number: H1200040001

FINAL REPORT

TITLE OF PROJECT: Mapping Whitebark Pine Distribution throughout the Greater Yellowstone Ecosystem

DATE OF SUBMISSION: October 26, 2006

NAME OF PARK/NPS UNIT: Greater Yellowstone Network, Inventory and Monitoring Program

NAME OF UNIVERSITY PARTNER/PRINCIPAL INVESTIGATOR

Lisa Landenburger, Research Associate
Rick L. Lawrence, Associate Professor
Montana State University
Department of Land Resources and Environmental Sciences
College of Agriculture
P.O. Box 173492
Bozeman, MT 59717-3490
Phone: 406-994-5409
Fax: 406-994-5122
lisa_landenburger@usgs.gov
rickl@montana.edu

NPS KEY OFFICIAL:

Robert Bennetts
Greater Yellowstone Network I & M Program
1648 South 7th Ave
Bozeman, MT 59717
Phone: 406-994-2281
Fax: 406-994-4160
E-mail: Robert_Bennetts@nps.gov

MAPPING WHITEBARK PINE DISTRIBUTION THROUGHOUT THE GREATER YELLOWSTONE ECOSYSTEM

INTRODUCTION

Whitebark pine seeds have long been reported as an important food source for grizzly bears in the GYE and are an important element of suitable grizzly bear (*Ursus arctos*) habitat (Lanner and Gilbert, 1994). Whitebark pine also serves as a keystone species in that its presence increases the biodiversity of both plant and animal communities throughout the ecosystem (Tomback and Kendall, 2001). The overall health and status of whitebark pine is currently threatened by infestation by mountain pine beetle (*Dendroctonus ponderosae*) and the spread of whitepine blister rust (*Cronartium ribicola*). The mapping of whitebark pine distribution is integral to the success of the long-term monitoring of whitebark pine since, before we can study, understand, and mitigate the mechanisms driving destructive agents of whitebark pine, we must first know its distribution across the landscape.

The Interagency Grizzly Bear Study Team (IGBST) initiated an effort to map the distribution of whitebark pine (*Pinus albicaulis* Engelm.) throughout the Greater Yellowstone Ecosystem (GYE) in the fall of 2003. This research initiative was sponsored by the USGS Interdisciplinary Science Support Activities (ISSA) program with the larger aim of promoting the use of remote sensing throughout the US Geological Survey. This funding opportunity allowed the IGBST to purchase necessary software and to compile existing ground data generated by U.S. Forest Service and National Park Service units throughout the ecosystem. It also provided an impetus to forge a constructive collaboration between the USGS, Northern Rocky Mountain Science Center, and the Remote Sensing Laboratory at Montana State University. More recent funding was appropriated to the project by the GYE Interagency Whitebark Pine Monitory Program and the USGS Northern Rocky Mountain Science Center.

METHODS

Our study area for this mapping endeavor covers an area spanning parts of Montana, Wyoming, and Idaho referred to as the Greater Yellowstone Ecosystem (Fig. 1). Landsat 7 Enhanced Thematic Mapper Plus (ETM+) satellite imagery was used as the primary mapping data source for reasons of cost and computational efficiency. Each ETM+ image covers a 170 km by 185 km area and cost approximately $600 per image, making mapping at regional scales highly cost effective. Seven ETM+ scenes covering the GYE were provided with geometric and radiometric corrections by the EROS Data Center, Sioux Falls, South Dakota. Although complete coverage was initially provided for July 2002 (summer scenes) and September 1999 (fall scenes), the summer scenes were eliminated from the analysis due to significant cloud coverage in the southern portion of the ecosystem. Also, since cambial and shoot growth of whitebark pine typically occur in late-May to mid-June (Weaver, 2001), the July images would not have added any phenological contrast in spectral signatures among the various conifer species in the GYE.

Reference data consisting of known whitebark pine locations with accurate ground coordinates were necessary to "train" the software spectral pattern-recognition algorithms to identify other areas of potential whitebark pine. Substantial time and effort was allocated to compile and generate reliable reference data that represented the complete variation in spectral response and terrain attributes associated with whitebark pine on the landscape.

Our initial expectations were to rely heavily on ground truth data collected by U.S. Forest Service and National Park Service personnel in conjunction with their standard timber-stand exams, vegetation plots, soil surveys, and other field activities where ground information was collected. Reliance on data from outside sources was necessary since the scope of this project precluded adequate time and resources for extensive field work. The agencies responded well to our requests for data, and we were able to compile a substantial pool of vegetation data that collectively constituted a fairly sufficient representation of the spatial complexities of the ecosystem. These data were populated with attributes of varying degrees of descriptive detail regarding forest coverage due to the different data sources and the various mandates associated with each source dataset. We consequently had to employ the "least-common-denominator" approach and apply a binary response variable: either whitebark or non-whitebark. The presence of whitebark for our study was indicated by a dominant component of whitebark pine in forest canopy.

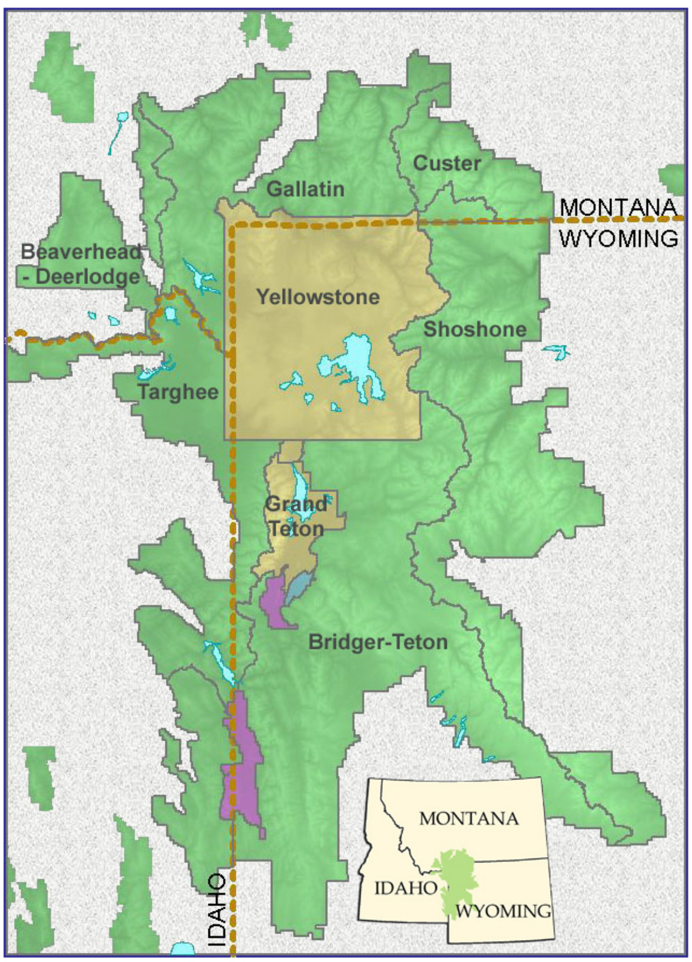


Figure 1. Study area: the Greater Yellowstone Ecosystem.

Aside from the inconsistent descriptive content, the data also exhibited varying degrees of spatial inaccuracy. Since much of the data was collected before GPS units were readily available, various methods had been used in estimating ground locations, resulting in disparate degrees of spatial accuracy. A significant component of the data collected via GPS also had considerable error (up to +/- 300 m) due to selective availability and the lack of post-differential correction. Due to the 30-m resolution of Landsat imagery, a comparable locational accuracy of the reference data was a necessary requirement. In order to verify the spatial reliability of the data, random checks were performed by overlaying each datum on top of geo-referenced digital orthographic quad (DOQ) photos. From this, we determined that a substantial number of data points, both field-estimated and GPS-acquired, lacked the necessary spatial accuracy. For example, data points allegedly representing whitebark pine stands were often found either at the very edge of forest stands, in open alpine meadows, or even within lakes. Preliminary analyses suggested that these inconsistencies, if left uncorrected, resulted in poor predictive power of models (refer to Appendix A). It became apparent that all WBP data points, as well as all non-WBP conifer reference points occurring in potential WBP terrain, had to be examined. Consequently, 7000+ reference points were examined against underlying DOQs. Points with questionable spatial accuracy were flagged. With the aid of higher resolution aerial photos from government archives and/or the expertise of government personnel with substantial field experience, points were shifted over distances ranging from ten to a few hundred meters to their most probable location. Points that could not be corrected with a high degree of certainty (29% of the compiled data) were eliminated from the analysis.

Aerial photographs were also used to generate an additional 8000+ reference data points by way of photo interpretation. The use of aerial photos to verify and supplement existing data added substantially to the overall integrity and quantity of our reference data collection. Combining field based data with photo-interpreted data resulted in a total of 15,110 reference points used in this analysis. From this final compilation of data points, 85% was randomly selected for use in the classification “training” process, while the remaining 15% was reserved for accuracy assessment.

Spectral and spectrally derived predictor variables (also referred to as explanatory or independent variables) used in this analysis included (1) at-satellite reflectances scaled to 8-bit values for the six ETM+ reflective bands (the thermal band was not provided), (2) re-scaled at-satellite tasseled cap brightness, greenness, and wetness values (Huang et al. 2002), (3) principal component data values for all six bands, and (4) normalized difference vegetation index (NDVI), where NDVI = (near infrared – red)/(near infrared + red). In addition, terrain parameters considered to have strong predictive powers for whitebark pine occurrence were used, including elevation, slope, aspect and latitude. Slope and aspect layers were derived from USGS 30-m National Elevation Dataset Digital Elevation Models. Aspect was transformed by taking the cosine in radians and stretching it to an 8-bit value by adding 1 and multiplying the sum by 200. For computational reasons, latitude was generated from a 1-km regular grid and then re-sampled to 30 m (Δlatitude $\cong$ 0.00011 degrees per km).

In this study logistic regression and a rule-based method, classification tree analysis (CTA), were both used for comparison to generate classified images depicting the distribution of whitebark pine. Logistic regression has been shown to be an appropriate statistical tool when the response variable is binary in nature (Bricklemyer et al., 2002). The goal of logistic regression is to determine the best fitting model to describe the relationship between a dichotomous characteristic (i.e., presence/absence of whitebark pine) and a set of independent predictor variables. A forward/backwards stepwise logistic regression algorithm in S-Plus was employed to generate the coefficients of a linear equation predicting the logit transformation of the probability of whitebark pine presence:

$$\text{logit (p)} = b_o + b_1X_1 + b_2X_2 \ldots + b_nX_n$$

where p is the probability of presence of whitebark pine, the logit of p is expressed as a linear combination of the *n* explanatory variables X_n, and the regression coefficients (b_n) are a measure of the predictive capability of the independent variables (Dallal, 2001). Probability of presence can be calculated from the logit since the logit is defined as the natural log of the odds:

$$\text{logit(p)} \equiv \ln (p / (1 - p)) \equiv \ln \text{(probability of presence / probability of absence)}$$

Solving for the probability yields:

$$p = (1 / (1 + e^{-\text{logit(p)}})) = (1 / (1 + e^{-(b0 + b1X1 + b2X2 \,.\,.\,. + bnXn)}))$$

The model was evaluated using the Akaike's Information Criterion (AIC) (Burnham and Anderson, 1998). The resulting logistic equation was then applied to the input predictor variables. The probability of whitebark pine (WBP) presence was thereby calculated for each pixel in the image. The resulting map is an image with pixel values ranging from 0 to 1 where a value of 1 indicates a predicted probability of 100% that WBP is present. A threshold criteria of p > 40, based on maximizing overall class accuracy, was used to segregate WBP from non-WBP.

Classification Tree Analysis (CTA) is also well suited to modeling response variables by producing predicted target classifications based on a series of if-then conditions (tree nodes). CTA has been reported to be an effective tool for classification of remotely sensed data in conjunction with ancillary data (Lawrence and Wright, 2001). CTA examines the input reference observations (populated with predictor variable values) and recursively partitions the data based on binary splits of individual predictor variables such that deviance in the response variable is minimized (Breiman et al., 1984). By following the paths of the resulting tree, one can determine a series of rules predicting classes. These rules were applied to the input spectral and ancillary predictor variables for the entire study area. The resulting image maps the response variable, in our case presence or absence of whitebark pine.

In this study two distinct CTA splitting algorithms were used. The first algorithm was a standard class probability splitting rule employed in the S-Plus decision-tree function. In this method the partitioning algorithm essentially splits at every possible value of every predictor and chooses the split that minimizes deviance (a measure of class impurity) while maximizing node homogeneity. If all observations were classified correctly at a terminal node the deviance at that node would be zero. Tree pruning to avoid over-fitting followed standard cross-validation techniques. The second method used in this study was the entropy splitting algorithm employed in See5, a proprietary software package produced by Rulequest. In this software program, the decision tree grows by applying a gain ratio criteria to recursively parse the training observations into homogeneous subsets (Quinlan, 1993; Huang et al., 2001). One distinct advantage to the See5 program was the option for boosting, a technique reported to significantly reduce the training error and to boost or enhance the classification accuracy (Freund and Schapire, 1999; Schapire, 1999). Boosting generated a user-specified number of classification trees such that each successive tree attempted to correct misclassification of the previous tree (Lawrence et al., 2004). At each iteration, the training samples were re-assigned weights with misclassified data given greater weight. The final predicted classification was based on a plurality vote from the complete set of classification trees. See5 provides a default of 10 boosts and a maximum of 99 boosts, both of which were evaluated in this study.

Classifications were conducted separately on three sets of Landsat images covering the study area (Fig. 2): the middle-path (path 38, rows 28-30), the east-path (path 37, rows 29-30), and the west-path (path39, rows 28-29). Classification was initially performed on the middle path yielding high accuracy rates that justified using the classification results of the middle-path in areas of path-overlap to identify supplemental training samples for the classification of the east- and west-paths (Parmenter et al., 2003). This method ensured a smooth and seamless transition across the final merged classified image. 4,000 random points in each overlap area were generated and populated with the corresponding classification codes from the middle-path results. These points were then added to the training samples for the east- and west-paths respectively. Accuracy was assessed using the reserved 15% of the reference data.

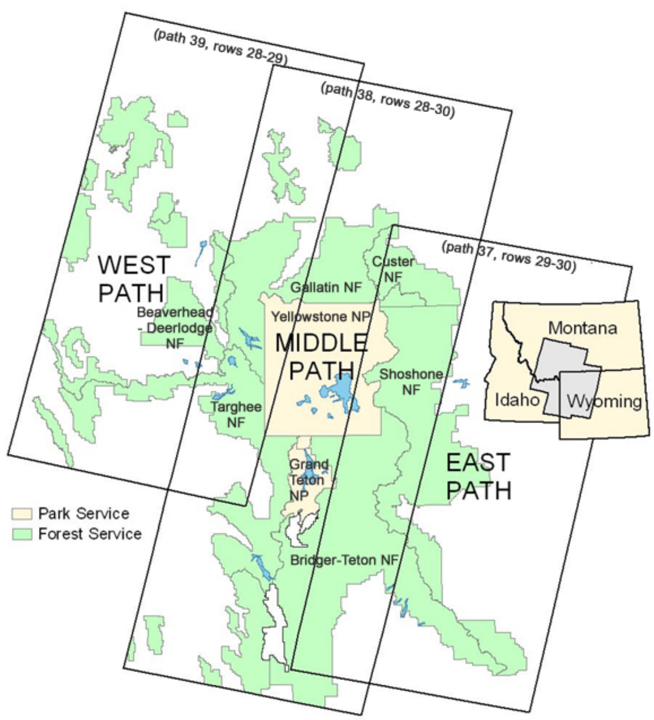


Figure 2. Study area with divisions based on east, west and middle paths of Landsat ETM+ satellite imagery.

An independent field-based ground validation was conducted in the summer of 2005. The primary objective was to obtain ground validation data void of any potential biases inherent in the reference dataset. It was also our objective to analyze the sensitivity of the analysis to varying densities of WBP, the effects of elevational gradients on map accuracy, and the variation of accuracy associated with different data sources. Sampling strategy was dictated by the size of the study area (approximately 57,000 km^2), fiscal limitations, time constraints, and inaccessibility of sizeable roadless wilderness areas within the ecosystem. A subset for field investigation was selected from the total number of sites predicted as WBP. These field sites were stratified using distance from nearest road ($\leq$ 4 miles) as well as wide geographic coverage. Additional GPS data collected by the Bridger Teton National Forest, the GAP project, and the Inter-agency Whitebark Pine Monitoring Program were included to augment the field data collected.

RESULTS

A total of 15,110 training data points, excluding random points generated in the image overlap areas, were compiled for this analysis. WBP sampling points comprised 31% of the reference dataset and 69% were non-WBP (Fig. 3). Photo-interpreted points comprised 54% and agency supplied field data comprised 46%. Five different predictive models were run on the middle-path to determine which yielded the best results and hence was the most appropriate for the entire study area. User's, producer's and overall class accuracies conducted for the five statistical methods (Table 1) proved Classification Tree Analysis (CTA) with boosting to perform the best.

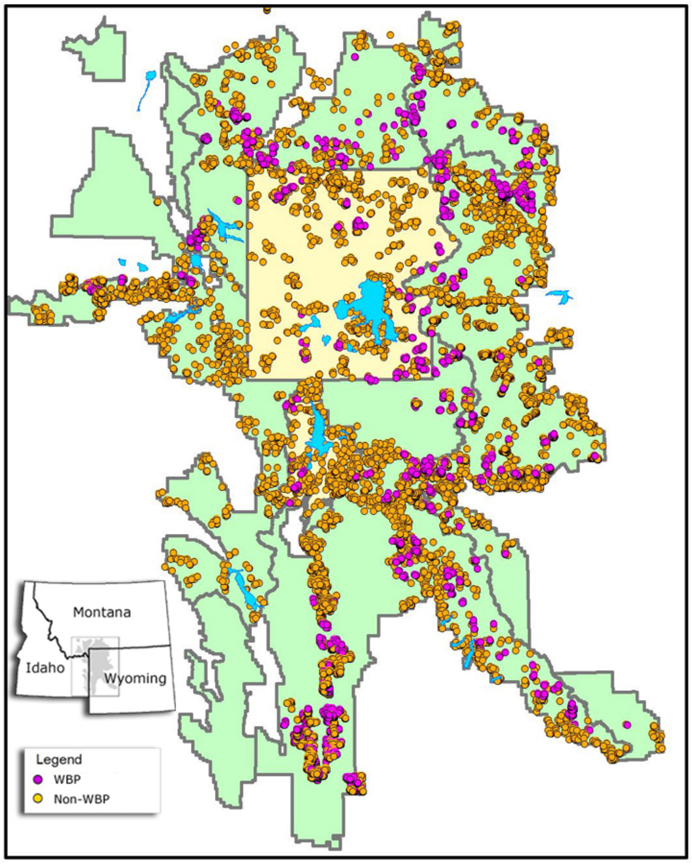


Figure 3. Distribution of reference data across the GYE.

Table 1. Comparative accuracies for classification of middle path Landsat ETM+ imagery.

Statistical Method	Producer's Accuracy		User's Accuracy		% Overall Accuracy
	% WB	% NWB	% WB	% NWB	
Logistic regression / S-Plus	89.6	92.6	85.6	94.8	91.6
CTA / S-Plus	90.7	93.8	87.8	95.3	92.8
CTA / See5, boost = 0	90.2	95.2	90.8	94.9	93.5
CTA / See5, boost = 10	92.5	97.0	94.2	96.1	95.5
CTA / See5, boost = 99	93.6	97.0	94.1	96.6	95.8

CTA, with or without boosting, yielded consistently higher accuracies compared with logistic regression. However, CTA employing the SEE5 maximum boosting algorithm improved overall accuracy by 2.3% with respect to See5 CTA with no boosting, and by 3% compared to the S-Plus single decision-tree results. The user's accuracy, referred to as errors of commission, indicates the probability that a given pixel classification matches what is actually on the ground. Maximum boosting increased the user's class accuracy for whitebark by 3.3% compared to See5 with no boosting, and produced an increase of 6.3% compared to the single decision-tree of S-Plus.

For further comparison, Landsat images from the east and west paths were each classified using See5 CTA without boosting and with maximum boosting (99 trials). Classification with maximum boosting of the east and west paths provided substantially similar accuracies, except that the producer's accuracy for the west path resulted in more errors of omission for WBP and fewer errors of omission for non-WBP (Table 2). Nonetheless, boosting consistently improved classification accuracies. The producer's estimated class accuracy for WBP showed an improvement ranging from 0.6 to 6.0% and the overall accuracy showed an increase between 1.2 and 2.9%. Although, the user's class accuracy for WBP actually decreased by 1.4% with boosting for the west path, it was increased by 2.3 and 2.9% for the middle and east paths respectively (Table 2).

Table 2. Comparative accuracies for classification with no boosting and maximum boosting.

	Image Path	Producer's Accuracy		User's Accuracy		% Overall Accuracy
		% WB	% NWB	% WB	% NWB	
NO BOOSTING	middle	90.2	95.2	90.7	94.9	93.5
	east	94.0	91.6	86.2	96.5	92.5
	west	83.0	98.4	95.1	94.0	94.2
BOOSTING = 99	middle	93.6	97.0	94.1	96.6	95.8
	east	94.6	95.9	92.7	97.0	95.4
	west	89.0	97.8	93.7	96.0	95.4

All five statistical methods used in this analysis strongly indicated terrain parameters elevation, aspect, and latitude as significant predictors of whitebark pine. The final logistic regression equation for predicted whitebark pine was:

$$\text{LOGIT}(p) = -81.51013 + 0.00348(\text{ELEVATION}) - 0.01919(\text{ASPECT}) + 0.00110(\text{LATITUDE}) - 0.35120(\text{TASSLED CAP BRIGHTNESS}) + 0.01950(\text{SLOPE}) + 0.33386(\text{TM BAND 7}) - 0.19694(\text{TM BAND 5}) + 0.28711(\text{TM BAND 4})$$

CTA analysis via S-Plus created a decision tree with 23 rules (terminal nodes) for classifying WBP (Fig. 4) with a misclassification error rate of 6.5%. Predictor variables appearing in the CTA decision-tree construction included all of the variables resulting in the logistic regression equation minus slope and TM band 7.

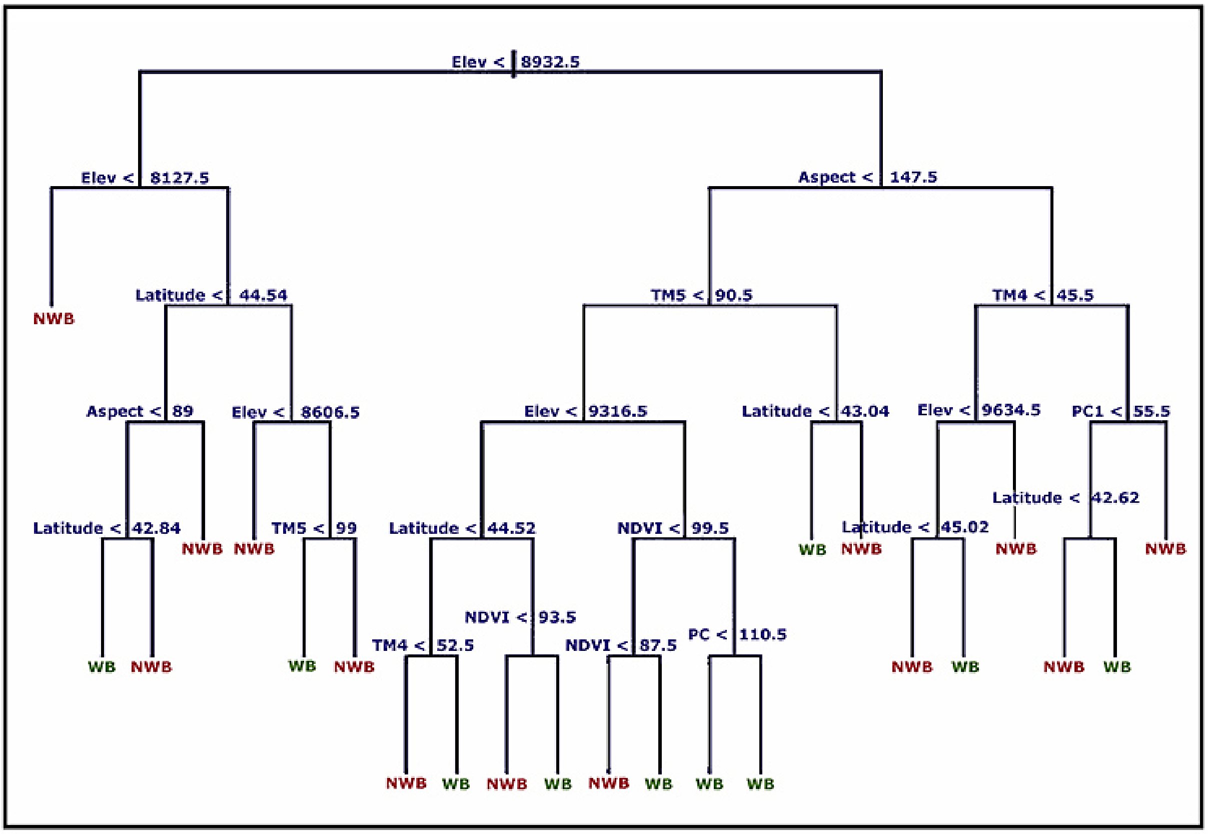


Figure 4. Classification Tree results from S-Plus CTA analysis.

CTA using See5 with no boosting produced a very large tree (too large to be shown here) with 105 decision rules and 15 levels of branching. All 20 predictor variables except for TM band 7 and tassled cap wetness were evident in the See5 classification results. Examining the upper branches (first 5 levels) of the tree revealed elevation, aspect, latitude, tasseled cap greenness, TM bands 1, 4, and 5, and principal component 5 as the significant predictor variables for whitebark pine. Although these are similar to the resulting variables of the S-Plus and logistic regression classifications, there are some differences. Notably, slope, which was part of the logistic regression equation, was not a factor in the S-Plus decision rules and did not appear in the See5 classification tree until branching level 6 and lower. Also, TM band 7, a significant predictor in the logistic regression results, was not called upon in either of the CTA procedures. The overall misclassification error rate using See5 with no boosting was 3.3%.

Although results from all five classification methods shared high accuracy rates, close visual inspection of the non-boosting classification images revealed some non-forested areas in the higher elevations (above 2,900 m) that were misclassified as WBP. Boosting reduced this over-estimation. It is possible that since boosting focuses on reducing classification errors in areas of inherent ambiguity (Freund and Schapire, 1999), that the boosting algorithm more readily segregated extremely sparse high-elevation krumholtz WBP from non-forest better than the other algorithms.

Classification results for all three satellite paths, generated via See5 CTA with maximum boosting, were merged into a final contiguous and seamless map of the entire study area (Fig. 5).

Classification results for the final merged image yielded an overall accuracy of 95.7% and a user's class accuracy for WBP of 92.9% . The KHAT statistic, calculated at 0.90, indicates that a given classification is 90% better than a classification resulting by chance (Lillesand and Kiefer, 2000).

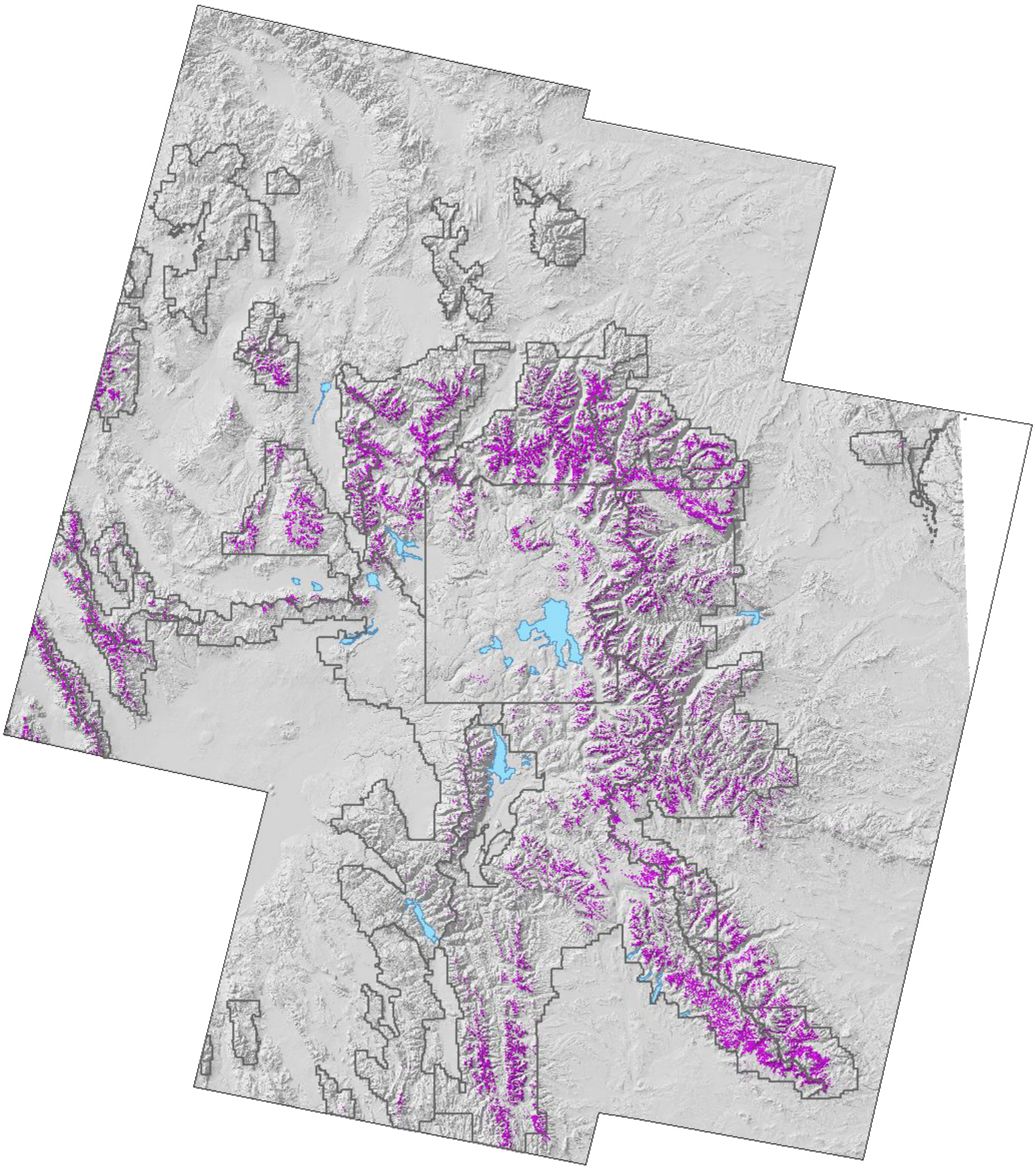


Figure 5. Image of predicted WBP distribution throughout the GYE.

The resulting distribution of 2,184 field validation data points was well distributed throughout the study area, but tended to cluster, making them potentially inadequate for testing overall map accuracy (Fig. 6). They were used, rather, for evaluating variations in accuracy related to WBP density, elevation, and geographic area. A series of accuracy assessments using the field validation data was conducted to determine a threshold value of percent WBP needed in the upper canopy to delineate "presence" versus "absence" (Table 3). Each assessment assumed a different threshold value for WBP. The defining threshold was determined as that value which returned the highest accuracy of the predictive model. A threshold value of 15% optimized the collective user's accuracy for both WBP and non-WBP (83.1% and 83.3%, respectively) while maintaining a 91.4% producer's accuracy for WBP, and an optimal 83.2% for overall classification. Based on these results, and for the applications of this model, the threshold for presence of WBP was determined to be 15%.

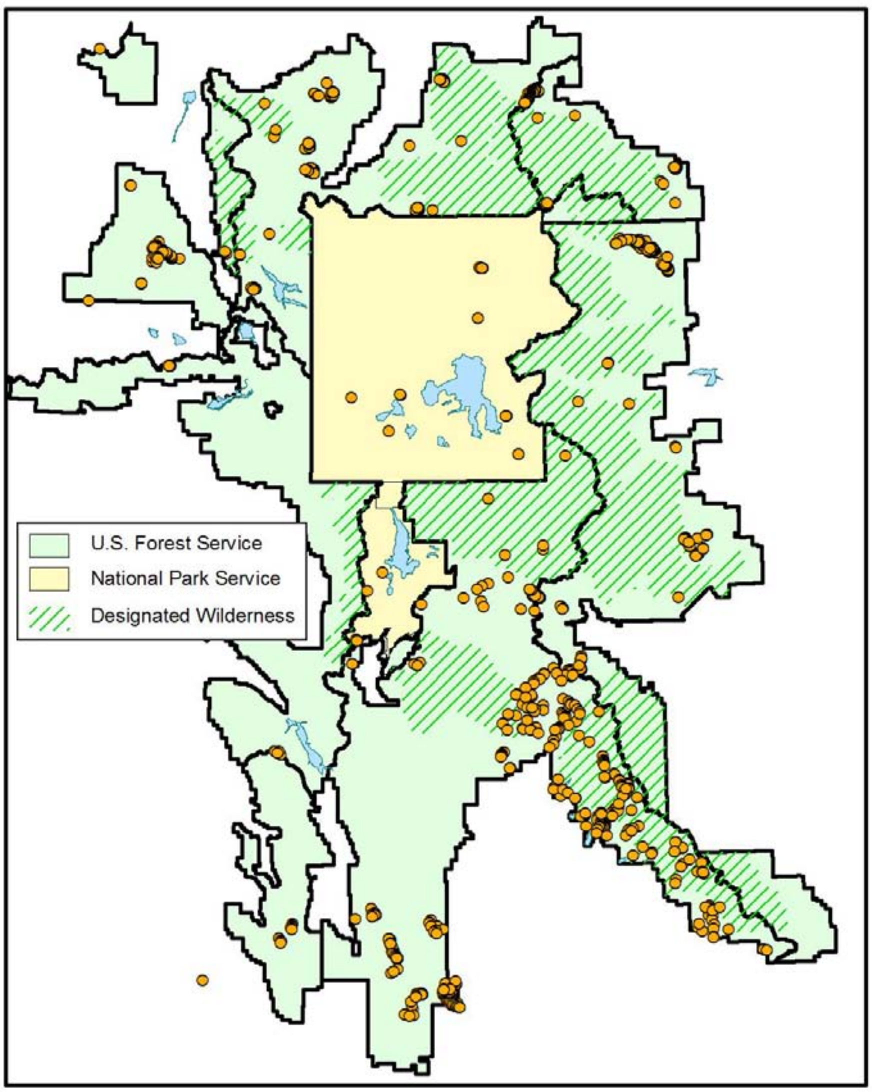


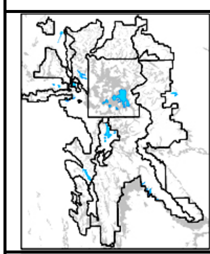
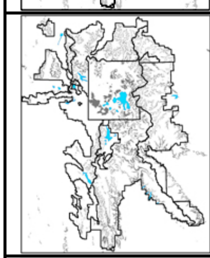
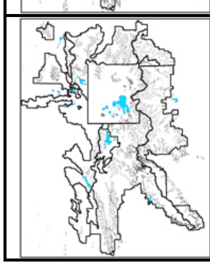
Figure 6. Distribution of validation points across the study area.

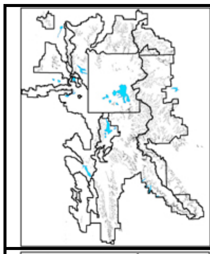
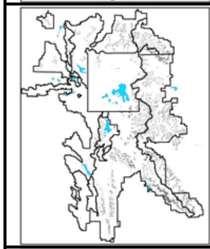
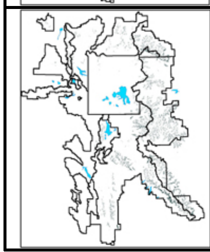
Table 3. Accuracy assessments at different Presence / Absence thresholds for WBP.

Threshold	Producer's Accuracy		User's Accuracy		Overall Accuracy
	WBP	Non-WBP	WBP	Non-WBP	
≥ 5%	84.6%	76.7%	90.9%	64.5%	82.5%
≥ 10%	89.3%	71.8%	85.7%	78.0%	83.2%
≥ 15%	91.4%	69.8%	83.1%	83.3%	83.2%
≥ 20%	91.6%	67.1%	80.8%	84.0%	81.8%
≥ 25%	91.9%	64.7%	78.2%	85.3%	80.5%

A Natural Breaks classification based on the Jenk's optimization was applied to partition the validation points into six discrete elevation classes of minimal variance (Cromley, 1996). Accuracy assessments of field validation data were conducted at each of these elevation ranges (Table 4). Inspection of the error matrices associated with the mid-to-high elevations (ranges 3 – 6) consistently indicates lower accuracies for non-WBP than WBP. This implies that WBP was over-predicted at these elevations. At lower elevations (ranges 1 and 2), however, where WBP tends to represent a relatively low presence in mixed coniferous stands, the classification tended to under predict WBP presence.

Table 4. Accuracy assessments at successive elevation ranges.

	Elevation Range	Producer's Accuracy		User's Accuracy		Overall Accuracy
		WBP	Non-WBP	WBP	Non-WBP	
	2277 – 2553 m (7470 – 8376 ft) (145 data pts)	0%	100%	NA	95.9%	95.9%
	2553 – 2691 (8377 – 8829 ft) (377 data pts)	87.0%	83.0%	81.9%	87.8%	84.9%
	2692 – 2805 (8830 - 9203 ft) (454 data pts)	92.8%	76.0%	91.4%	79.3%	88.3%

	2806 – 2900 (9204 - 9512 ft) (491 data pts)	92.1%	47.8%	78.4%	74.8%	77.6%
	2901 – 3025 (9513 - 9924 ft) (576 data pts)	94.4%	43.7%	80.4%	76.0%	79.7%
	3026 – 3104 (9925 – 10,185 ft) (141 data pts)	85.9%	76.2%	89.5%	69.6%	83.0%

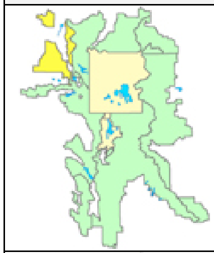
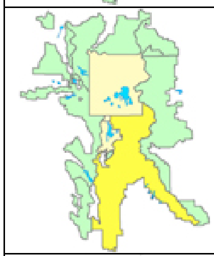
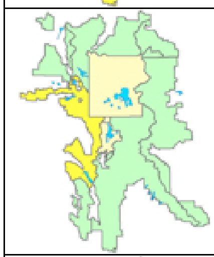
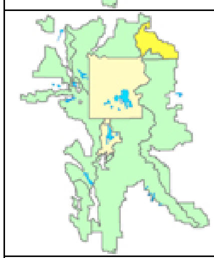
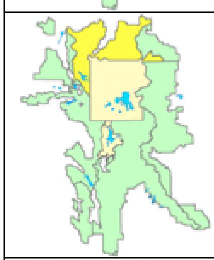
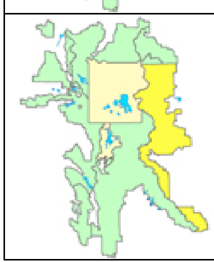
We also conducted separate accuracy assessments based on the field validation data for each of the national parks and national forests included in the study area (Table 5). Caribou-Targhee National Forest and the national parks are included for completeness, but are of limited value as no WBP present stands were sampled in those jurisdictions. Substantial differences in accuracy existed amongst the various administrative units. The Shoshone National Forest on the eastern side of the study area (Fig.1) had the lowest overall accuracy (74.5%) while the Beaverhead National Forest on the northwest corner of the study area (Fig. 1) had the highest overall accuracy (94.6%). All other accuracies were in the 83% to 88% range. Evaluating all 2,142 field validation points with an assumed WBP “presence” threshold of 15% (refer to Table 3) yields an overall accuracy of 83.2%. A relatively low producer’s accuracy of 67.1% for non-WBP point to a high rate of errors of omission which indicates that presence of WBP is over predicted.

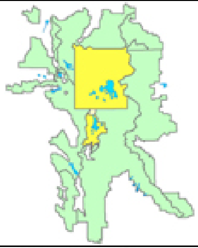
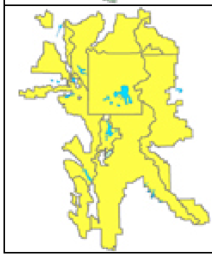
DISCUSSION

Our classification resulted in very high accuracy rates, demonstrating considerable success in detecting WBP as a dominant and/or mixed component in forest canopies. This was especially notable considering Landsat is generally not expected to be adequate for classification at the species level, and previous attempts had not been successful. There are several factors that we believe were important in our classification success, although it is not possible to quantify the impact of the factors individually.

We used classification algorithms that have been recently developed and applied to remotely sensed data (Lawrence et al., 2006). Classification tree analysis generally has resulted in improved accuracies when

Table 5. Accuracy assessment by administrative unit.

	Administrative Unit	Producer's Accuracy		User's Accuracy		Overall Accuracy
		WBP	Non-WBP	WBP	Non-WBP	
	Beaverhead N. F. (142 data pts)	95.0%	92.7%	97.0%	88.4%	94.4%
	Bridger-Teton N. F. (675 data pts)	86.1%	81.0%	80.1%	86.8%	83.4%
	Caribou-Targhee N. F. (30 data pts)	NA	100.0%	NA	100.0%	100.0%
	Custer N. F. (158 data pts)	87.0%	50.0%	95.5%	24.0%	84.2%
	Gallatin N. F. (558 data pts)	95.6%	61.8%	89.1%	81.0%	87.6%
	Shoshone N. F. (580 data pts)	92.5%	44.7%	73.4%	78.4%	74.5%

	Yellowstone N. P. & Grand Teton N. P. (41 data pts)	NA	92.3%	NA	94.7%	87.8%
	Entire Study Area (2184 data pts)	91.4%	69.8%	83.1%	83.3%	83.2%

compared to other classification accuracy, and boosting algorithms have been commonly reported to increase classification algorithms by 10% or more compared to non-boosted classification trees, although increased accuracy is not guaranteed (Lawrence et al., 2004). We used a simple boosting algorithm (Quinlan, 1993); it is possible that recent advances in boosting and related bagging algorithms might have further improved accuracies (Lawrence et al., 2004; Lawrence et al., 2006).

We believe that focusing on a single class might have improved our accuracies compared to previous classifications. Classification of multiple land cover types in a single classification necessarily entails trade-offs; an approach that improves accuracy of one class might decrease accuracy of another class. We were able to select from among multiple algorithms and approaches and select the one that would improve WBP accuracy without concern for other species.

We were also able to assemble an extensive reference data set as a result of a high level of cooperation from national forests and national parks within our study area, as well as the availability of excellent aerial photography coverage. These data required extensive review and filtering to make them acceptable for use in a remote sensing study, and again high levels of cooperation for local land managers was extremely valuable in this process.

A fourth factor that might have been important in our success compared to other species-level studies with Landsat data was the spatial distribution of whitebark pine. Exploratory data analysis of our results indicated that elevation was the most important predictor variable for the occurrence of whitebark pine. Whitebark pine distribution is heavily controlled by elevation and it can occupy nearly pure homogeneous stands in harsh, dry, windy mountainous terrain, although it typically co-exists with other conifers in moister and more protected high-elevation sites (Arno and Hoff, 1989). This elevation control on distribution likely reduced species confusion with other pines, which typically exist at lower elevations within the study area.

Elevation, however, also created issues for our classification. Whitebark pine can be completely out-competed by subalpine fir and Engleman spruce in localized areas of higher moisture, for example along drainages and poorly drained sites. Our model over predicted WBP by over 20% in these high elevation sites (Table 4). Adding a hydrologic index as a predictor variable might improve accuracy for these sites. Bedrock geology, geomorphology, and soil types also impact WBP distribution (Hansen-Bristow et al., 1990) and might be evaluated for future classifications. Also, since timber stand exams were rarely conducted in elevations where WBP typically occurs, it could be useful to collect additional non-WBP reference points at high elevation to assist the training process in segregating high elevation non-WBP from WBP.

We also noted differences in accuracies across our study area associated with different administrative units (Table 5). The reasons for these spatial differences in accuracy were unclear but we believe there were multiple possibilities. These variations might have been a function of differences in the quality of reference data from each jurisdiction. Lower quality data might have less accurately represented the spatial and spectral variability in that location and resulted in a model that did a poorer job in predicting WBP locally. Another possibility was that these differences represented a broader spatial trend in accuracy across our study area. Accuracies tended to be highest in the west side of the study area and lowest on the east side when evaluated by administrative units. This same trend, however, was not present when accuracy was evaluated by Landsat path (Table 1).

The results of this study are potentially valuable for several on-going efforts, including: (1) GYE Interagency Whitebark Pine Monitory Program from which probabilistic samples will be derived from the WBP map resulting from this study (2) expansion of efforts to conduct a habitat-based grizzly bear Population Viability Analysis (USFWS, 1993, Boyce et al., 2001), which is currently restricted to areas inside the recovery zone; (3) updates to data layers for the Yellowstone Grizzly Bear Cumulative Effects Model (Weaver et al., 1986, Dixon, 1997); (4) modeling the potential effects of decline in major food sources or global climate change; (5) use in habitat selection models evaluating the effects of motorized recreation on denning and active grizzly bears; and (6) use in two studies examining GYE carnivore population dynamics that are sponsored by the US Geological Survey, National Park Service, and the Wildlife Conservation Society. Other efforts that might benefit include: (1) monitoring the distribution of whitebark pine blister rust in the GYE as part of key foods monitoring required by the grizzly bear recovery plan (USFWS, 1993) and conservation strategy (USFWS, 2003), (2) use by state wildlife and federal land agencies for planning and evaluation of management efforts, and (3) distribution through National Biological Information Infrastructure (http://www.nbii.gov), making this data layer available to the public.

ACKNOWLEDGEMENTS

We would like to give special thanks to Bryan Bailey, principal Remote Sensing Scientist with EROS Data Center, whose assistance was paramount in funding this study. Acknowledgments are also given to the many people throughout the Greater Yellowstone Ecosystem who were willing to share their time, expertise, and hard-earned field data, and who gave access to their air photo archives. These include: Dennis Barron, Dean Burnham, Liz Davy, Dale Dawson, Andy Norman, and Jim Ozenberger from the Bridger-Teton National Forest; Jeff Dibenedetto and DeeDee Arzy from the Custer National Forest; Mark Novak, Joan Roe, Sally Senger, Julie Shea, Steve Swain, and Dan Tyers from the Gallatin National Forest; Chip Fisher with the Helena National Forest, Steve Haynes, Dirk Shaupe, and Klara Varga from Grand Teton National Park; Mary Maj from the Greater Yellowstone Coordinating Committee; David Tart with the USFS Region 4 office, Melissa Jenkins and Judy Warwick from the Targhee-Caribou National Forest; Kent Houston and Ken Ostrom from the Shoshone National

Forest; Roy Renkins and Ann Rodman from Yellowstone National Park. We would also like to give special thanks to Don Despain (USGS, NRMSC) who helped greatly in the air photo interpretation, Steve Cherry (Montana State University) who shared his expertise as a statistician, Collin Homer (EROS Data Center) who provided us with source imagery as well as sound advise, and Maury Nyquist who was very helpful in getting us started on this project.

REFERENCES

Arno, S. F., and R. J. Hoff. 1989. Silvics of whitebark pine (Pinus albicaulis) U.S. Department of Agriculture, General Technical Report INT-253. Ogden, Utah. 11 pp.

Boyce, M. S., B. M. Blanchard, R. R. Knight, and C. Servheen (2001). Population viability for grizzly bears: a critical review. *International Association of Bear Research and Management*. Monograph Series no.4.

Brieman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone (1984). *Classification and regression trees*. Belmont, California, Wadsworth International Group, 358 pp.

Bricklemyer, R. S., R. L. Lawrence, and P. R. Miller (2002). Documenting no-till and conventional till practices using Landsat ETM+ imagery and logistic regression. *Journal of Soil and Water Conservation*, 57:267-271.

Burnham, K. P., and D. R. Anderson (1998). *Model selection and inference: a practical information-theoretic approach*. New York, Springer-Verlag New York Inc., 349 pp.

Dallal, G. E. (2001). The little Handbook of Statistical Practice. Available on-line only, www.tufts.edu/~gdallal/LHSP.HTM

Dixon, B. G. (1997). *Cumulative effects modeling for grizzly bears in the Greater Yellowstone Ecosystem*. M.S. Thesis, Montana State University, Bozeman, Montana, USA.

Cromley, R.G., A comparison of optimal classification strategies for choroplethic displays of spatially aggregated data. International Journal of Geographic Information Systems, 1996. 10(4): p. 405-424.

Hansen-Bristow, K., Montagne, C., and Schmid, G. (1990). Geology, geomorphology, and soils within whitebark pine ecosystems. In Proceedings – Symposium on whitebark pine ecosystems: ecology and management of a high-mountain resource. U.S. Department of Agriculture, General Technical Report INT-270. Ogden, Utah. p. 62-71.

Huang, C., L. Yang, B. Wylie, and C. Homer (2001). A strategy for estimating tree canopy density using Landsat 7 ETM+ and high resolution images over large areas. *Proceedings of the Third International Conference on Geospatial Information in Agriculture and Forestry*, Denver, Colorado, 5-7 November, CD-ROM 1 disk.

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

Felicetti, L. A., C. C. Schwartz, R. O. Rye, M. A. Haroldson, K. A. Gunther, D. C. Phillips, and C. T. Robbins (2003). Use of sulfur and nitrogen stable isotopes to determine the importance of whitebark pine nuts to Yellowstone grizzly bears. *Canadian Journal of Zoology*, 81:763 – 770.

Freund, Y., and R. E. Schapire (1999). A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 14:771-780.

Lanner, R. M., and B. K. Gilbert (1994). Nutritive value of whitebark pine seeds, and the questions of their variable dormancy. In W. C. Schmidt and F. K. Holtmeier, eds. *Proceedings international workshop on subalpine stone pines and their environment: the status of our knowledge*. U.S. Forest Service General Technical Report INT-GTR-309, p. 206-211.

Lawrence, R., and A. Wright (2001). Rule-based classification systems using classification and regression tree (CART) analysis. *Photogrammetric Engineering and Remote Sensing* 67:1137-1142.

Lawrence, R., A. Burn, S. Powell, and M. Zambon (2004). Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment* 90:331-336.

Lawrence, R.L., S. Wood, and R. Sheley, 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). *Remote Sensing of Environment* 100:356-362.

Lillybridge, Terry R.; Kovalchik, Bernard L.; Williams, Clinton K.; Smith, Bradley G. 1995. *Field guide for forested plant associations of the Wenatchee National Forest*. Gen. Tech. Rep. PNW-GTR-359. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 335 p. In cooperation with: U.S. Department of Agriculture, Forest Service, Pacific Northwest Region, Wenatchee National Forest. [29851]

Lillesand, T. M., and R. W. Kiefer (2000). *Remote sensing and image interpretation*. John Wiley & Sons, Inc., New York, p.574.

Parmenter, A. W., A. Hansen, R. Kennedy, W. Cohen, U. Langner, R. Lawrence, B. Maxwell, A. Gallant, and R. Aspinall, (2003). Land use and land cover change in the Greater Yellowstone Ecosystem: 1975-1995. *Ecological Applications* 13:687-703.

Quinlan, J. R. (1993). C4.5: programs for machine learning. San Mateo, California: Morgan Kaufmann Publishers: 302 p.

Salford Systems (2002). *CART user's guide: An implementation of the original CART methodology*, San Diego, California, Salford Systems, p. 79-80.

Schapire, R. E. (1999). A short introduction to boosting, *Journal of Japanese Society for Artificial Intelligence*, 14:771-780.

Tomback, D. F., and K. C. Kendall (2001). Biodiversity losses: the downward spiral. In Tomback, D. F., S. F. Arno, and R. E. Keane, eds., *Whitebark pine communities: ecology and restoration*, Island Press, Washington D. C., p. 243-262.

U.S. Fish and Wildlife Service (USFWS) (1993). *Grizzly bear recovery plan*. U.S. Fish and Wildlife Service, Missoula, Montana, USA.

U.S. Fish and Wildlife Service (USFWS) (2003). *Final Conservation Strategy for the grizzly bear in the Yellowstone Ecosystem*. U.S. Fish and Wildlife Service, Missoula, Montana, USA.

Weaver, T. (2001). Whitebark pine and its environment. In Tomback, D. F., S. F. Arno, and R. E. Keane, eds., *Whitebark pine communities: ecology and restoration*, Island Press, Washington D. C., p. 41-73.

Weaver J, Escano R, Mattson D, Puchlerz T, Despain D. (1986). A cumulative effects model for grizzly bear management in the Yellowstone Ecosystem. In Contraras, G. P., and K. E. Evans , eds., *Proceedings- Grizzly Bear Habitat Symposium*. General Technical Report; 1985 Apr 30-1985 May 2; Missoula, MT. Ogden, UT: U.S.D.A, Forest Service, Intermountain Research Station, pp. 234-46.

APPENDIX A: Evaluating Effects of Correction to Source Reference Data

In the early stage of this study we determined that a substantial portion of the source reference data obtained from government agencies was collected pre-GPS days and hence, lacked the necessary spatial accuracy required by the analysis. These errors, if left uncorrected could potentially result in poor predictive power in the classification process. Consequently we invested several months of time and effort to verify and correct spatial locations prior to model construction (refer to paragraph 4 under Methods). Upon completing this corrective phase we compared classification results using corrected data with results using uncorrected data (supplemented with photo interpreted observations). We chose, for this evaluation, the middle Landsat scene of our study since (1) approximately 71.5% of this scene falls within the Grizzly Bear Recovery Zone, (2) the variable terrain in this scene was a good representation of the overall spatial complexities of the ecosystem, and (3) it offered a broad representation of the source reference data since it encompasses two National Parks and intersects five National Forests from which data was acquired (Fig A.1).

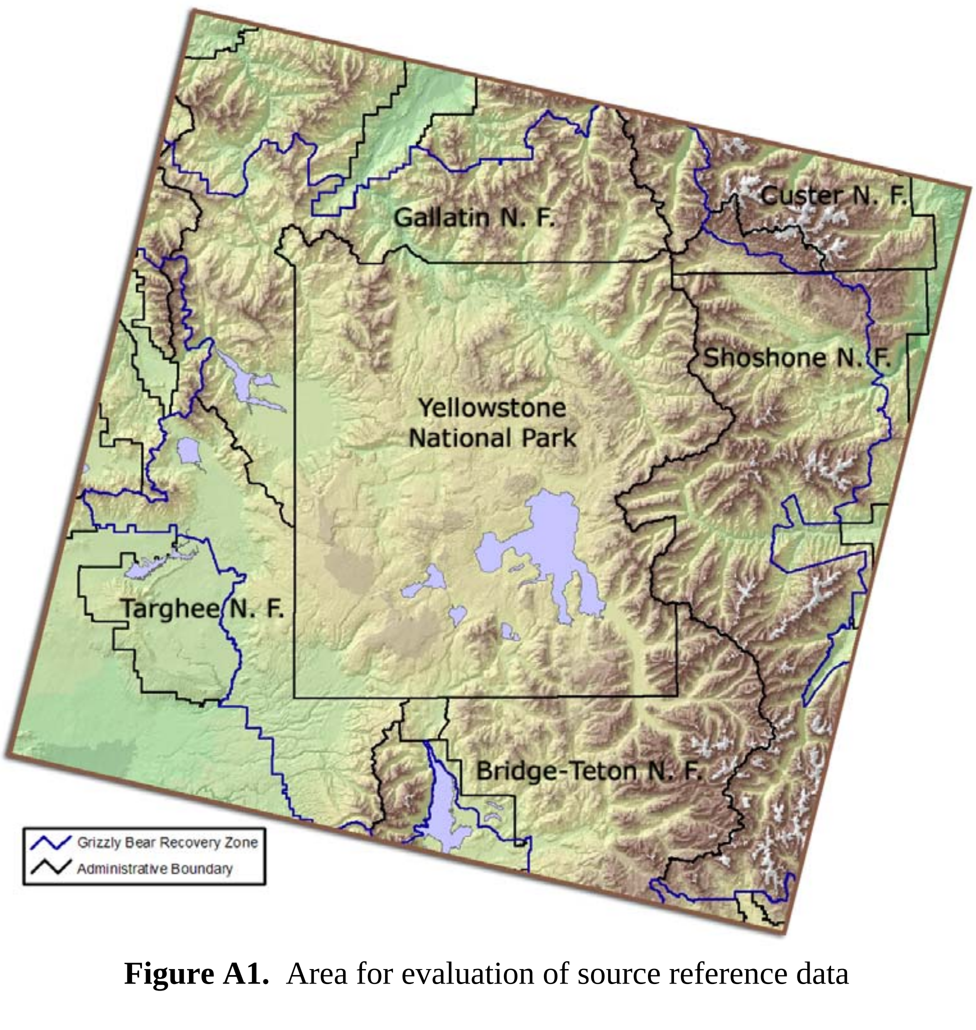


Figure A1. Area for evaluation of source reference data

Our uncorrected dataset consisted of 4,569 ground reference data points with whitebark observations comprising 18% of the total. The corrected data set consisted of 5,743 reference observations with 2,603 of these points collected via air photo interpretation. Approximately 31% of the initial raw data had to be eliminated due to poor spatial accuracy and insufficient information to modify coordinate locations. Classification tree analysis (CTA) using See5 with 10 boosting trials was applied to the raw uncorrected dataset and the modified dataset.

Results for the corrected dataset yielded an overall accuracy increase of only 4% from the results of the uncorrected data (Table A1). However, the producer's and user's class accuracies for whitebark showed substantial increases of 23.5% and 15.4% respectively. The KHAT statistic for the raw dataset was calculated as 0.683 versus 0.897 for the corrected data set, indicating an increased probability of 21.4% that an observed classification was better than one derived by chance.

Table A1. Accuracy results before and after correction to reference data.

Classification Tree Analysis (See5 with 10 Boosting Trials)	Producer's		User's		% Overall
	% WB	% NWB	% WB	% NWB	
Uncorrected dataset	68.0	96.3	79.8	93.3	91.3
Corrected dataset	91.5	97.5	95.2	95.4	95.3