Rangeland Ecology & Management 70 (2017) 644–655

Contents lists available at ScienceDirect

Rangeland Ecology & Management

journal homepage: http://www.elsevier.com/locate/rama

Integrating Remotely Sensed Imagery and Existing Multiscale Field Data to Derive Rangeland Indicators: Application of Bayesian Additive Regression Trees

Sarah E. McCord [a,*], Michaela Buenemann [b], Jason W. Karl [c], Dawn M. Browning [c], Brian C. Hadley [d]

[a] Jornada Experimental Range, Las Cruces, NM 88003, USA
[b] Department of Geography, New Mexico State University, Las Cruces, NM 88003, USA
[c] US Department of Agriculture—Agricultural Research Service, Jornada Experimental Range, Las Cruces, NM 88003, USA
[d] US Department of the Interior Bureau of Land Management, National Operations Center, Lakewood, CO 80225-0047, USA

ARTICLE INFO

Article history:
Received 2 July 2016
Received in revised form 24 January 2017
Accepted 22 February 2017

Key Words:
Bayesian additive regression trees
BLM AIM
monitoring
rangelands
remote sensing

ABSTRACT

Remotely sensed imagery at multiple spatial scales is used increasingly in conjunction with field data to estimate rangeland indicators (e.g., vegetation cover) and meet the growing need for landscape-scale monitoring and assessment of rangelands. Remote sensing studies that produce rangeland indicators often require intensive and costly field-data collection efforts to produce accurate model predictions. Existing monitoring data, such as those collected by the Bureau of Land Management's Assessment, Inventory, and Monitoring (AIM) program, are potentially useful sources of field data in remote sensing modeling studies. Given their data-hungry nature, common regression tree—based modeling approaches may be inadequate for reliably predicting rangeland indicators with the smaller sample sizes of AIM data than typically used for remote sensing studies. Current literature suggests that Bayesian models, such as Bayesian additive regression trees (BART), may provide a suitable alternative to traditional regression tree—based modeling approaches to overcome the sample size limitation of the AIM data. In this study, we used 182 AIM field plots together with both high (RapidEye) and moderate (Landsat OLI) spatial resolution satellite imagery to predict bare ground and bare soil, total foliar, herbaceous, woody, and shrub cover indicators on rangelands in a 14 625-km^2 area of northeastern California. We demonstrate that a BART model performed similarly to other regression tree approaches when field data and high spatial resolution imagery predictions were combined to predict indicator values using the medium spatial resolution Landsat image. The BART models also provided spatially explicit uncertainty estimates, which allow land managers to more carefully evaluate indicator predictions and to identify areas where future field data collection might be most useful. This study demonstrates that existing field data and freely available, remotely sensed imagery can be integrated to produce spatially explicit and continuous surface estimates of rangeland indicators across entire landscapes.

Published by Elsevier Inc. on behalf of The Society for Range Management.

Introduction

Land managers and management agencies need information on the condition of rangelands at multiple spatial scales[1] (e.g., local administrative units to nationwide) to support adaptive management (MacKinnon et al., 2011). National monitoring efforts, such as the Bureau of Land Management's (BLM) Assessment, Inventory, and Monitoring (AIM) program, collect data on core rangeland indicators at these multiple spatial scales to provide consistent information about the condition of rangelands in the United States (Taylor et al., 2014). These core rangeland indicators, collected with consistent methods (Herrick et al., 2009), include fractional cover of bare ground and plant species, canopy gap sizes, vegetation height, non-native invasive species, and plant species of management concern (MacKinnon et al., 2011). Together, these indicators provide a suite of monitoring data critical for understanding the effects of disturbances occurring across spatial scales and for a range of management concerns (MacKinnon et al., 2011).

Collecting sufficient field data to provide information about rangelands at all spatial scales[1] and continuously across extensive geographic areas, however, is unfeasible due to logistical and budgetary constraints. Where plot-based field data are insufficient to provide information for landscape-scale decision making, remote sensing is a

Mention of a proprietary product does not constitute a guarantee or warranty of the products by the us government or the authors and does not imply its approval to the exclusion of other products that may be suitable.

* Correspondence: Sarah E. McCord, US Department of Agriculture—Agricultural Research Service, Jornada Experimental Range, P.O. Box 30003, MSC 3JER, New Mexico State University, Las Cruces, NM 88003-8003, USA. Tel.: +1 575 646 2961, fax: +1 575 646 5889.
E-mail address: smccord@nmsu.edu (S.E. McCord).

[1] For this paper, scale refers to spatial resolution unless otherwise specified.

http://dx.doi.org/10.1016/j.rama.2017.02.004
1550-7424/Published by Elsevier Inc. on behalf of The Society for Range Management.

promising supplemental information source (Booth and Tueller, 2003; Hunt et al., 2003; MacKinnon et al., 2011; Karl et al., 2012). Remote sensing technologies have the potential to provide low-cost information that is spatially explicit, continuous, and available at multiple spatial and temporal scales. In addition, remotely sensed data can provide information that may not be easily measured by field data collection methods, such as landscape metrics (Ludwig et al., 2000, 2007; Booth and Tueller, 2003; Hunt et al., 2003; MacKinnon et al., 2011) and retrospective monitoring for change detection (Washington-Allen et al., 2006).

Estimates of vegetation and bare ground fractional cover are core pieces of information for rangeland management that can be derived from remotely sensed data (MacKinnon et al., 2011). Field data, however, are often required to train and test multiscale fractional cover models to produce indicator estimates based on remotely sensed imagery. Many studies that provide remotely sensed cover indicators rely on independent field sampling efforts, including specialized sample designs and field data collection campaigns, to obtain training and testing data (Mishra and Crews, 2014; Sant et al., 2014; Xian et al., 2015). These studies use a range of field data collection methods, such as ocular estimation (Homer et al., 2012; Xian et al., 2015), line-point intercept (Karl, 2010; Mishra et al., 2014), and photo interpretation (Sant et al., 2014). These field data collection methods vary in precision and may not be consistent with other field data collection efforts or national rangeland monitoring programs (Karl et al., 2014). While double sampling approaches (Karl et al., 2014) may be used, this may not be sufficient to overcome inconsistencies between field methods and fundamental differences in definition between fractional cover indicators reported in these studies (Toevs et al., 2011). Furthermore, alternate methods of data collection decrease inefficiency as double sampling is time consuming, and specialized field data may not be usable for other monitoring purposes.

The cost of implementing a remote sensing—based monitoring program could be reduced by using existing field data and adapting modeling approaches to meet existing data constraints. One potential source of existing field data is the data collected through BLM AIM (MacKinnon et al., 2011). These data have been collected annually across the western United States since 2011 by seasonally trained and calibrated field crews using core data collection methods (Herrick et al., 2009). The broad spatial extent and consistency of these data may provide a high-quality training and testing data source for use in remote sensing studies. The amount and location of these field data, however, are variable and may not be sufficient to model rangeland indicators from moderate spatial resolution imagery. Thus, a modeling approach that includes existing field data together with imagery of multiple spatial resolutions could be used to model fractional cover of rangelands.

Using high spatial resolution imagery to train successively coarser spatial resolution imagery is a well-described method for predicting fractional cover across large study areas (Laliberte et al., 2007; Homer et al., 2012; Gessner et al., 2013). This multi-image modeling approach can improve prediction performance while reducing the amount of field data required to predict cover across large spatial extents. Producing rangeland cover indicator estimates and uncertainty estimates using existing field data and multiscale imagery, however, may require novel modeling approaches. Regression tree models are common for predicting landscape-scale rangeland cover indicators using multiscale data (Laliberte et al., 2007; Homer et al., 2012; Gessner et al., 2013; Mishra et al., 2014). These models, however, do not provide spatially explicit prediction uncertainty estimates and at the same time require large amounts of training data across the range of possible conditions to produce robust estimates. Applying a Bayesian framework to multiscale regression tree approaches may provide a potential solution to both the small field data sample size problem (McCarthy, 2007) and the need for a spatially explicit estimate (Wilson et al., 2011). Such a framework, however, has not been developed and tested for its utility to monitor rangeland conditions.

In our study, we addressed this problem by using existing AIM data for training and testing a multiscale Bayesian additive regression tree (BART) model of rangeland fractional cover for a study area in northeastern California. We had three objectives: 1) to use existing field and remotely sensed data to produce spatially explicit and continuous rangeland indicator estimates of bare ground, bare soil, total foliar cover, herbaceous cover, woody cover, and shrub cover; 2) to provide spatially explicit and continuous uncertainty estimates for the fractional cover indicator estimates; and 3) to determine if the sample size available via the AIM dataset was adequate to build robust BART models.

Bayesian Additive Regression Trees

Regression Tree—Based Models

Regression trees using remote sensing and ancillary data to predict rangeland cover indicators are nonparametric machine learning algorithms, which recursively partition a dataset through simple regression models into increasingly smaller groups of similar response variables (Homer et al., 2012; Gessner et al., 2013; Mishra and Crews, 2014). Regression trees are built such that the same series of splitting decisions applied to the predictor variables (i.e., regression planes) can be repeated to predict a continuous dependent variable via nonlinear relationships (Breiman et al., 1984; Strobl et al., 2009). A regression tree starts with all observations and predictor variables (i.e., the "root"). Observations in the root are split into subgroups (i.e., "branches") using splitting rules for the predictor values. Internal nodes are branched again until a terminal node, called a "leaf," is produced. This predicts a single response. Simple regression trees use linear regression to determine fixed leaf attributes and splitting rules. In contrast, regression model trees apply a regression plane to the internal nodes and leaves of regression model trees in each prediction instance (Homer et al., 2013). Some implementations of regression model trees (e.g., Cubist, RuleQuest, 2012) use generalized rules derived from regression trees. This results in a series of stratified piecewise regressions that reduce over fitting tendencies by weighting the process to have fewer parameters, thereby mitigating outlier impacts. Regression model trees have been shown to be more robust than simple regression trees and are thus more widely applied to prediction problems such as remote sensing image product development (Walton, 2008).

Although single regression model trees show improvement over simple regression trees, there is greater benefit from considering ensembles (also called committees) of model trees. Random Forest (RF, Breiman, 2001) is a common implementation of ensemble model trees (Strobl et al., 2009). In RF models, a large number of diverse trees is grown using a random subset of the training data and a random subset of explanatory variables for each tree. To make a prediction, each tree in the ensemble is applied to an observation and supplies a prediction. The prediction with the most "votes" among all the trees is assigned to the observation. The advantage of an ensemble of model trees is that a reasonably accurate model can be improved can be successively fine-tuned by additional trees, thereby making the model more resilient to both weak training data sets and weak explanatory variables than simple regression or regression model trees (Strobl et al., 2009; RuleQuest, 2012). Because RF uses only a small subset of the training data, cross-validation is built into the method. It has been argued that an independent testing dataset may not be required, which has appeal when only a small training dataset is available (Walton, 2008; Strobl et al., 2009). RF models are also thought to be resilient (but not immune) to overfitting due to the weak correlation between trees in the model (Breiman, 2001). However, large tree structure (Segal et al., 2004) and spatial autocorrelation of the data (Mascaro et al., 2014) may still induce model overfit, especially with small training data sets (De 'ath and Fabricius, 2000).

While regression model tree approaches have been successfully applied across disciplines, they can fail on two accounts. First, RF does not satisfactorily handle missing data (Walton, 2008). When missing data

occur for an independent variable within a dataset, RF compensates by imputing the mean, median, or mode of that missing observation. The assumptions required for such imputation can reduce prediction accuracies of these two models (Walton, 2008). It is possible to build RF models that recognize missing values as not useful for prediction, similar to other regression tree models (e.g., Gu and Wylie, 2015). However, this is not a standard implementation of RF. A second shortcoming of RF is it does not provide a measure of estimate uncertainty. RF predicts by a combination of pruning parameters, minimum cases rules, and ensemble of trees "committee," but the outcome of the vote is binary and therefore the level of agreement among trees regarding the prediction is not known (Walton, 2008). This lack of a prediction-based uncertainty has implications for model interpretation in that while overall model error can be described, uncertainty associated with an estimate as it relates to the population is not known.

Bayesian regression tree models, of which BART is a common implementation, provide point-level prediction uncertainty (Wilson et al., 2011; Pratola et al., 2014), which has appeal for interpreting and using cover estimates from remotely sensed imagery. This is because a Bayesian interpretation of uncertainty differs from that of RF-based models (Clark, 2007). In a Bayesian uncertainty framework, the true population value is assumed to be unknown and a "credible interval" describes the range of values within which the true population parameter is likely to fall. Thus, Bayesian uncertainty describes the uncertainty of the population parameter. For example, an estimated 95% credible interval for bare ground means that the true value of bare ground has a 95% likelihood of being within 5% and 25%. This allows for expression of uncertainty about the model inputs, which are considered random, in addition to uncertainty in the model prediction. In contrast, non-Bayesian models (such as RF) typically describe model uncertainty as a confidence interval for the prediction, which for the earlier example would be interpreted as 95% percent of the times the exercise (or experiment) is repeated, the true population parameter of bare ground would fall between 5% and 25%. A critical assumption of both RF and BART models is that the training and testing data adequately represent the true population parameter. Thus, in a management context, a Bayesian confidence interval is a simpler uncertainty metric to interpret for decision making and has the added benefit of acknowledging the uncertainty introduced by the model parameters in the final credible interval (Ellison, 2004).

Bayesian Additive Regression Trees

BART implements an ensemble regression tree model in a Bayesian framework where a collection of comparatively weakly performing model trees is aggregated to create one strong model (Chipman et al., 2010). Any individual model tree is considered to be weak because that tree itself is assumed to be a poorly performing model with small individual influence on the overall BART model performance. While the additive result across all trees may not necessarily produce a model with the highest training accuracy, it can produce a robust model that predicts well to new areas.

The BART approach provides many of the same strengths as RF including resistance to overfit, applicability to nonparametric data, and good performance using small training data sets. However, BART may provide advantages over RF because it offers improved treatment of missing data and produces prediction uncertainty estimates. Missing data are handled via a "missing in attributes" approach, which either builds the concept of "missingness" into the splitting rules themselves (i.e., models the missing data) or adds a splitting rule to handle missing data (Bleich et al., 2014). Similar to the Cubist regression tree approach, the predictive power of a missing attribute value is zero and the model accommodates accordingly (Gu and Wylie, 2015). Uncertainty estimates in BART are produced by making predictions from the model's posterior distribution. Furthermore, BART models have shown lower root mean square errors (RMSEs) than RF (Chipman et al., 2010).

The BART model is given by:

$$\boldsymbol{y} \approx \sum_{\boldsymbol{i}=1}^{\boldsymbol{m}} \boldsymbol{T_i^M}(\boldsymbol{x_i}\ldots\boldsymbol{x_k})+\boldsymbol{\mathcal{E}},\ \boldsymbol{\mathcal{E}}=\boldsymbol{N_n}\left(0,\boldsymbol{\sigma^2 In}\right), \quad (1)$$

where T^M denotes the tree structure and m is the number of distinct trees formed by a set of k predictor variables x; M describes the set of leaf parameters at the terminal nodes of the quantity b_t such that $M_t = \{\mu_{t,1}, \mu_{t,2}, \ldots \mu_{t,b_t}\}$; and the error term $\mathcal{E}$ is drawn from the normal distribution N_n whose standard deviation is a function of the variance times an inverse gamma distribution (Kapelner and Bleich, 2016). Internal nodes, or tree branches, are specified by a predictor variable splitting rules, $x_j > c$, where x_j is the splitting variable and c is the splitting value.

In Bayesian models, the posterior distribution is the conditional probability of a prediction occurring, which is the product of the model's prior distributions (i.e., "priors") and a selected likelihood function. In BART, the set of priors governs the tree structure and leaf parameters and is specified by "hyperparameters" (Kapelner and Bleich, 2016). There are two classes of hyperparameters: those specified by the user and those derived from the data. The structure of each prior and associated hyperparameters is discussed in detail in Chipman et al. (2010). The setting of priors and hyperparameters can be a daunting task. However, while the use of covariates as informative priors may improve model performance, uninformative or uniform priors often are sufficient to achieve desired model results and may be adequate (Bleich et al., 2014). Bayesian applications in ecology frequently employ uniform priors in the absence of meaningful covariate-derived priors (McCarthy, 2007). Thus, the default BART hyperparameters are often suitable (Pratola et al., 2014).

Methods

Study Area

The study area encompassed 14 625 km^2 of land administered by the BLM in the Modoc Plateau region of northeast California and northwest Nevada (Fig. 1; central coordinates: 41.153°N, 120.145°W). Within the study area we considered only lands defined as shrublands and

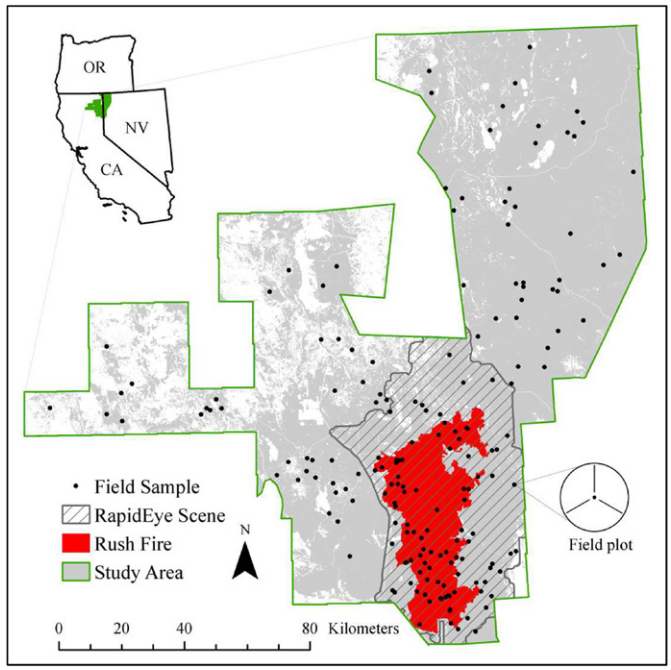

Figure 1. Map of the study area and extent of RapidEye imagery. A diagram of a circular field plot with three transects radiated from plot center is also shown.

grasslands in the National Land Cover Dataset (Homer et al., 2015). Vegetation communities of the study area are typical of the Great Basin and include a range of ecosystems from desert playas to mountain big sagebrush (*Artemisia tridentata* Nutt. ssp. *vaseyana* [Rydb.] Beetle). Riparian vegetation makes up a small fraction of the landscape. The study area is in the transition between semiarid or steppe (BSk) and Mediterranean (Csa, Csb) Köppen climates (Peel et al., 2007). Annual temperatures range from −10°C to 40°C, and 70% of the annual precipitation (317 mm) occurs in late fall and during the winter (Desert Research Institute, 2015).

The study area is managed under the BLM's multiple-use mandate to promote livestock grazing, wildlife habitat, mining activities, and recreation (USDI BLM, 2012). Rangelands in the study area face additional management concerns including declining greater sage-grouse *(Centrocercus urophasianus)* populations, increasing density of wild horses and burros, wildfire, drought, juniper encroachment, and the spread of invasive annual grasses (USDI BLM, 2012).

Field Data Collection

Field data were collected as part of the AIM program in a multiresource monitoring effort across three BLM field offices in northeastern California. The objective of this monitoring effort was to gather data regarding the condition of BLM rangelands in the study area relative to grazing impacts, greater sage-grouse habitat suitability, and the extent of invasive species. In August 2012, the Rush Fire burned 1 277 km^2 of sage-steppe vegetation in the southeastern portion of the study area, necessitating additional data in that area to monitor postfire recovery (USDI BLM, 2012). As a result, a nested sample design was implemented with higher sampling densities in and adjacent to the Rush Fire burn (1 plot per 38 km^2) and a lower-intensity (1 plot per 80 km^2) sampling network for the entire study area (see Fig. 1). For both sets of samples, the generalized-random tessellation stratified algorithm was used to draw the sample locations to ensure good spatial distribution of sample locations (Stevens and Olsen, 2004).

Field data were collected at 182 locations between May and August of 2014 by BLM field crews. A field plot was composed of three transects radiating from the plot center with each transect offset of the center by 5 m to avoid areas disturbed by trampling from the field crew (see Fig. 1). Midway through the study, the BLM elected to standardize plot sizes across all AIM monitoring projects. Before this change, 69 plots were measured using 50-m transects (2.2-acre plot area). After this change, the remaining 113 plots were measured using 25-m transects (0.7-acre plot area).

At each plot, field crews collected line-point intercept data by dropping a pin 50 times at regularly spaced intervals along each transect and measuring the number of plant species hits and the ground cover type intersecting the pin (Herrick et al., 2009). Field crews were trained and calibrated to BLM AIM data quality standards. From these measurements, plot-level fractional cover indicators were calculated, including bare ground cover, bare soil cover, total foliar cover, herbaceous cover, total woody cover, and shrub cover (see Herrick et al., 2009 for details on calculating indicators). Bare ground consisted of all between-plant cover (bare soil, rocks, litter, and water). Bare soil was the exposed soil with no plant, litter, or rock cover. Total foliar cover, the inverse of bare ground, was the proportion of the plot covered by any plant canopy. Herbaceous cover included nonwoody plant species recorded at any layer of the line-point intercept pin hit (see Karl et al., 2017). Woody cover was cover provided by all woody plant species (i.e., trees and, shrubs, and subshrubs) at any pin hit and shrub cover by only shrub plant species at any pin hit.

Imagery Acquisition and Processing

Two remote sensing image products were acquired for this study: 30-m, 12-bit Landsat 8 Operational Land Imager Surface Reflectance (US Department of the Interior US Geological Survey Earth Resources Observation and Science Center, 2015) and 7-m, 8-bit RapidEye Level 1B imagery (Blackbridge, 2015). Four cloud-free Landsat 8 scenes (Paths 43 and 44, Rows 31 and 32) from between May 26 and June 26, 2014 were mosaicked together using histogram matching in ENVI 5.1 to cover the entire study area (see Fig. 1). We spectrally subset the Landsat mosaic to only those six bands informative for describing terrestrial land cover characteristics: blue (0.45 − 0.51 μm), green (0.53 − 0.49 μm), red (0.64 − 0.67 μm), near-infrared (0.85 − 0.88 μm), and the two short-wave infrared bands (1.57 − 1.65 μm and 2.11 − 2.29 μm). RapidEye imagery, collected June 1 − 4, 2014, was mosaicked converted to at-surface reflectance using the FLAASH Module in ENVI 5.0 to convert the multispectral RapidEye imagery from calibrated radiance to scaled percent reflectance according to the generalized atmospheric parameters available in FLAASH. All five RapidEye bands were included in the final mosaic: blue (0.44 − 0.51 μm), green (0.52 − 0.59 μm), red (0.63 − 0.69 μm), red edge (0.69 − 0.73 μm), and near-infrared (0.76 − 0.85 μm).

Following the preprocessing, we conducted a principal component analysis (PCA) on each of the Landsat and RapidEye mosaics to reduce data dimensionality (Jensen, 2005) and minimize spatial autocorrelation (Dormann et al., 2007). The first three principal component (PC) bands explained > 99% of the total spectral variance in both the Landsat and RapidEye image mosaics, and initial exploration of the BART model showed that these three PCs were of relative equal importance and the remaining components had minimal influence in improving model performance. Thus, we selected these three PCs from each image mosaic as the BART predictor variables in this study.

To prepare the PC bands for BART model training, we intersected each band raster with the field plot point locations buffered by 30 m or 50 m. We selected the pixels from each PC band of both image mosaics, which intersected each field plot buffer and aggregated these pixel values using a weighted average based on the percent of each pixel within the field plot. This created a one-to-one relationship between the field and remote-sensing data. The average numbers of aggregated pixels per field plot in the RapidEye and Landsat image mosaics were 45 and 3, respectively.

BART Implementation

Prediction of fractional cover for the indicators occurred in a three-stage process, each of which generated a unique BART model (Kapelner and Bleich, 2016): RapidEye-only model, Landsat-only model, and a combined Landsat-RapidEye model (Fig. 2). This multistage process has been shown to be a successful technique for scaling from field data to high spatial resolution imagery like RapidEye to moderate spatial resolution imagery such as Landsat (Homer et al., 2012; Mishra et al., 2014). In the first stage, we used data from the 101 field plots located within the RapidEye scene to develop the RapidEye-only model using the RapidEye PCs as the predictor variables. In the second stage, we used all 182 field plots in the study area to develop the Landsat-only model using the Landsat PCs as the predictors. In the third stage, we used 1 000 randomly selected 30 × 30 m pixels from the RapidEye indicator predictions together with measured indicator values from the 182 field plots to develop the Landsat-RapidEye model with the Landsat PCs as the predictor variables. The assumption of this approach was that the RapidEye scene (which accounted for 85% of the training data) captured the environmental gradients of the larger study area. This final BART model with both field and RapidEye-derived training data produced the most accurate results; thus, we used this hybrid model to predict fractional cover indicators for the entire study area (see Kapelner and Bleich, 2016).

For all stages of model training, testing, and prediction, we used the *bartMachine* (Kapelner and Bleich, 2016) and *raster* packages (Hijmans and van Etten, 2012) in the R statistics program (version 3.2.2, R Core Team, 2015). Each model was built using 500 iterations with 200 trees, excluding a 200-iteration burn-in period. These parameters

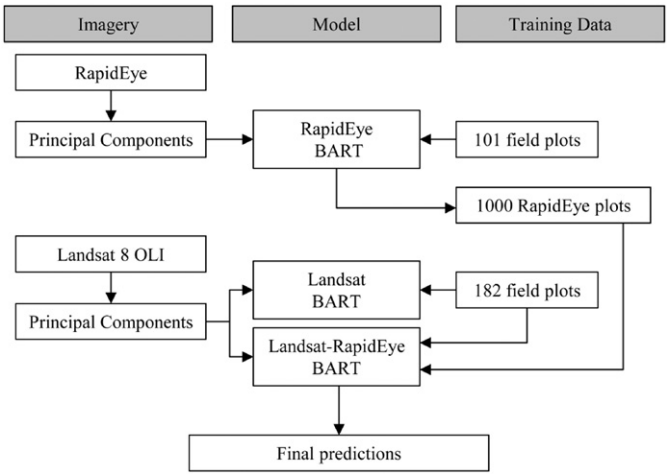

Figure 2. Flowchart illustrating how this study predicted fractional cover indicators using two kinds of imagery, three models, and two types of training data.

were the minimum number of iterations and trees needed to ensure model convergence. We used BART's default 5-fold cross-validation with 500 iterations to test each model. The number of trees and iterations was selected on the basis of preliminary exploration, which showed that variance among individual tree predictions converged at 200 trees with 500 iterations.

All models were developed and tested using BART's default hyperparameters ($\beta = 2, k = 2, \nu = 3, q = 0.9$) and uniform priors ($^1/_p$) at $\alpha = 0.90$ (Bleich et al., 2014). Bayesian applications in ecology frequently use uniform priors in the absence of meaningful covariate priors (McCarthy, 2007), and thus the default BART hyperparameters and settings are often suitable (Pratola et al., 2014; Kapelner and Bleich, 2016). Furthermore, the application of the BART model predictions in this study will be used for making management decisions; thus, the advantages to implementing BART consistently between indicators and over time outweigh the small performance benefits that might be gained from custom parameters, priors, and hyperparameters.

BART outputs can be classified as those associated with the predictions and those associated with the model specification itself. For each indicator prediction, the BART model provides a per-pixel fractional cover estimate and a 95% credible interval. The credible interval, drawn from the BART posterior probability, is the interval within which there is a 95% chance the true population parameter is found (Clark, 2007). BART model specification outputs include in-sample (training) pseudo-R^2, in-sample RMSE, out-of-sample (testing) RMSE, and model overfit. In BART, pseudo-R^2 is calculated as:

$$pseudo{-}R^2 = \frac{1 - \boldsymbol{SSE}}{\boldsymbol{SST}}, \quad (2)$$

where *SSE* is the sum of squared errors in the training data and *SST* is the total sum of squares or the sample variance of the response multiplied by $n - 1$. Pseudo-R^2 is not an ideal statistic because it is a relative measure of fit, sensitive to the number of predictors and distribution of the training data (Maindonald and Braun, 2006). Therefore we also report in-sample RMSE as an absolute measure of fit for accuracy comparison between models. The RMSE describes the differences between the predicted and observed responses of a model:

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i)^2}{\boldsymbol{n}}}, \quad (3)$$

where $\hat{y}_i$ is the predicted response, y_i is the observed response, and n is the number of samples. To calculate in-sample RMSE, y_i is drawn only from the set of training observations. For out-of-sample RMSE, y_i is drawn from a set of withheld testing data. The ratio of out-of-sample RMSE to in-sample was used as a measure of model overfit (*sensu* Mascaro et al., 2014).

Model Evaluation

We evaluated the performance of each BART model using five criteria: pseudo-R^2, in-sample RMSE, out-of-sample RMSE, model overfit, and the relationships among predicted values for the different indicators. The first four criteria are defined earlier. The fifth criterion relates to the idea that logical relationships exist between fractional cover indicators on the ground based on how the indicators are defined (e.g., total foliar cover plus bare ground should sum to 100%) and that deviations from these relationships in the model predictions provide insights into model performance (Homer et al., 2012). To assess BART model performance, we considered disagreement of modeled relationships with the following specific logical relationships: 1) total foliar cover plus bare ground should sum to 100%, 2) total foliar cover should be greater than or equal to the individual and summed predictions of woody cover, shrub cover, and herbaceous cover; and 3) bare ground should be greater than or equal to bare soil.

Sampling Sufficiency

One concern with using field data not originally intended to test and train remote sensing models is that the number of field plots may be lower than the sample sizes used in similar landscape-scale fractional cover prediction studies (e.g., Karl, 2010; Sant et al., 2014; Xian et al., 2015). To evaluate sampling sufficiency for this work and examine the influence of sample size on BART model performance, we ran sample size simulations of 100 iterations and evaluated the variability in RMSE and model overfit as a function of sample size. For the RapidEye-only BART model, sample size simulations explored field plot sample sizes between 5 and 101 at one-plot increments, with a 5-fold cross-validation. For the Landsat-only BART, we examined the impact of running BART with field plot sample sizes between 5 and 182 at two-plot increments. Because an infinite number of random samples could be drawn from the RapidEye image to improve the Landsat-RapidEye model performance, sample sufficiency was not examined in this model. The goal of these simulations was to determine if model performance began to converge within the field plot sample sizes available for this study or if more samples were needed to build a stable BART model.

Results

Model Evaluation

Pseudo-R^2 values varied by model and indicator (Table 1). Of the three models, the RapidEye-only BART model gave the lowest mean pseudo-R^2 values (0.64), then the Landsat-only BART model (0.66). The best mean pseudo-R^2 was produced by the Landsat-RapidEye BART model (0.73). Bare soil consistently gave the highest pseudo-R^2, ranging from 0.79 in the RapidEye-only and Landsat-only models to 0.85 in the Landsat-RapidEye model. Woody and shrub models produced the lowest pseudo-R^2, averaging 0.51 and 0.46, respectively, across all models.

Both out-of-sample RMSE and model overfit varied by model and by indicator (see Table 1). A Tukey's honest significant difference test showed the RapidEye-only and Landsat-only models' out-of-sample RMSEs were not statistically different ($P = 0.98$), but both models had significantly higher out-of-sample RMSE values than the Landsat-RapidEye model ($P < 0.001$). There was no statistically significant trend in model overfit by indicator, although shrub out-of-sample RMSEs were consistently among the lowest out-of-sample RMSE values in each model. This is likely because shrub cover was low across large portions of the study area due to fire history. For all indicators and

Table 1
Results of prediction and cross-validation for all indicators in the RapidEye, Landsat, and Landsat-RapidEye models.

Indicator	Pseudo-R^2	In-sample RMSE (%)	Out-of-sample RMSE (%)	Model overfit
RapidEye				
Bare ground	0.77	8	12	1.50
Bare soil	0.79	7	13	1.86
Total foliar	0.77	8	13	1.63
Herbaceous	0.69	11	15	1.36
Woody	0.44	8	9	1.13
Shrub	0.37	8	9	1.13
Mean	0.64	8	12	1.43
Standard deviation	0.19	1	2	0.29
Landsat				
Bare ground	0.73	8	11	1.38
Bare soil	0.79	7	11	1.57
Total foliar	0.73	8	11	1.38
Herbaceous	0.72	11	14	1.27
Woody	0.53	11	14	1.27
Shrub	0.45	9	11	1.22
Mean	0.66	9	12	1.35
Standard deviation	0.14	2	2	0.13
Landsat-RapidEye				
Bare ground	0.80	6	7	1.22
Bare soil	0.86	5	6	1.17
Total foliar	0.80	6	7	1.20
Herbaceous	0.78	6	8	1.33
Woody	0.60	5	7	1.40
Shrub	0.55	4	5	1.25
Mean	0.73	5	7	1.25
Standard deviation	0.13	1	1	0.10

RMSE, root mean square error.

models, the overfit ratio of out-of-sample RMSE to in-sample RMSE ranged from 1.13 to 1.86. Model overfit decreased with each BART model (see Table 1), from the RapidEye-only model (model overfit ratio of 1.43) to the Landsat-only model (model overfit ratio of 1.35) to the Landsat-RapidEye model (model overfit ratio of 1.25). There was no consistent overfit trend between indicators.

Indicator Estimates

For each of the six rangeland cover indicators, we used the BART models to produce maps of fractional cover and 95% credible interval estimates from the Landsat-RapidEye models. Bare ground (average cover: 46.0%), total foliar cover (average cover: 38.4%), and herbaceous cover (average cover: 34.1%) had the highest average cover across the study area while woody cover (average cover: 13.1%), shrub cover (average cover: 12.0%), and bare soil (18.9%) had the lowest average cover (Table 2). The spatial distribution of the indicators varied across the study area, with higher proportions of total foliar and nonwoody vegetation in the southeastern portion of the study area (Fig. 3) corresponding to the area of the Rush Fire. Higher bare ground, bare soil, woody, and shrub cover values were found in the northeastern and western portions of the study area.

Table 2
Mean prediction and 95% credible interval width for each indicator summarized across all per-pixel predictions in the study area.

Indicator	Mean percent cover	σ	Mean CI width	σ
Bare ground	46.0	20.7	8.6	3.8
Bare soil	18.9	9.6	2.6	4.3
Total foliar	38.4	17.6	8.8	3.9
Herbaceous	34.1	9.2	3.0	5.1
Woody	13.2	5.3	2.5	4.3
Shrub	12	4.0	1.7	2.9

CI, confidence interval.

Credible interval estimates also varied spatially and by indicator. Total foliar and bare ground had the widest average credible interval widths (8.8% and 8.6%, respectively), whereas herbaceous cover had the largest variation in credible interval width (3%, $\sigma = 5.1$). Shrub cover had the smallest average credible interval width at 1.7. Credible interval widths were highest in the northeastern part of the study area for all indicators, although there were patches of increased credible interval widths within the fire boundaries in the south-central portion of the study area (Fig. 4). These areas correspond to areas with lower sampling intensity and undersampled vegetation community types.

Sampling Sufficiency

Sampling sufficiency simulation results indicated that enough AIM field data were available to train and test a BART model in our study area. In-sample RMSE and model overfit ratios converged to stable values at sample sizes at least 50 plots lower than the number of available field plots (Figs. 5 and 6). In the RapidEye-only model, convergence occurred at approximately 50 field plots for in-sample RMSE and the model overfit ratio (see Fig. 6). In the Landsat-only model, the convergence for both metrics also occurred at approximately 100 field plots for both pseudo-R^2 and model overfit (see Fig. 6). Sampling sufficiency convergence patterns varied slightly by indicator. For both the RapidEye-only and Landsat-only models, herbaceous cover results converged before any of the other indicators at approximately 25 (RapidEye) plots and 50 (Landsat-only) plots for both in-sample RMSE and model overfit ratio.

Discussion

The results of this study demonstrate that a multiscale BART model can be successfully applied to predict rangeland cover indicators using remotely sensed imagery and available field data. Although success of multiscale regression tree approach has been shown in prior studies (Homer et al., 2012; Sant et al., 2014; Xian et al., 2015), this study demonstrated that existing field data collected for other studies can be successfully applied as training data in a rangeland remote sensing study (see Figs. 5 and 6). Integrating samples derived from RapidEye-only model predictions and AIM field samples improved model performance over using just AIM data with the Landsat-only model (−4% RMSE, see Table 1). Furthermore, the BART approach provided prediction-level uncertainty estimates that previous approaches often did not (see Fig. 4). Variations among the RapidEye-only, Landsat-only, and Landsat-RapidEye BART models' performance may have been due to changes in spatial resolution, the density of field samples, or differences between indicators.

Model Evaluation

By all evaluation measures, the Landsat-RapidEye model was the most robust model of the three models evaluated in this study. Pseudo-R^2 values were the highest, while RMSE values and model overfit were the lowest for all indicators in the Landsat-RapidEye model (see Table 1). The pseudo-R^2 values for bare ground, herbaceous, and bare soil were greater than the R^2 values reported in other studies implementing multiscale regression tree models (Homer et al., 2012, 2015; Mishra et al., 2014). RMSE values for all indicators in our study were comparable with those produced in other studies, although our study covered a smaller study area size and less vegetation community variability. While the Landsat-RapidEye RMSEs were similar to other research studies, shrub pseudo-R^2 values in this study were lower than R^2 of previous research. However, these results are not directly comparable due to differences in metrics. Woody and total foliar cover as defined in this study were not comparable with previous work that employed similar regression model trees, due to differences in the definition of these indicators. While some studies report training and testing accuracies

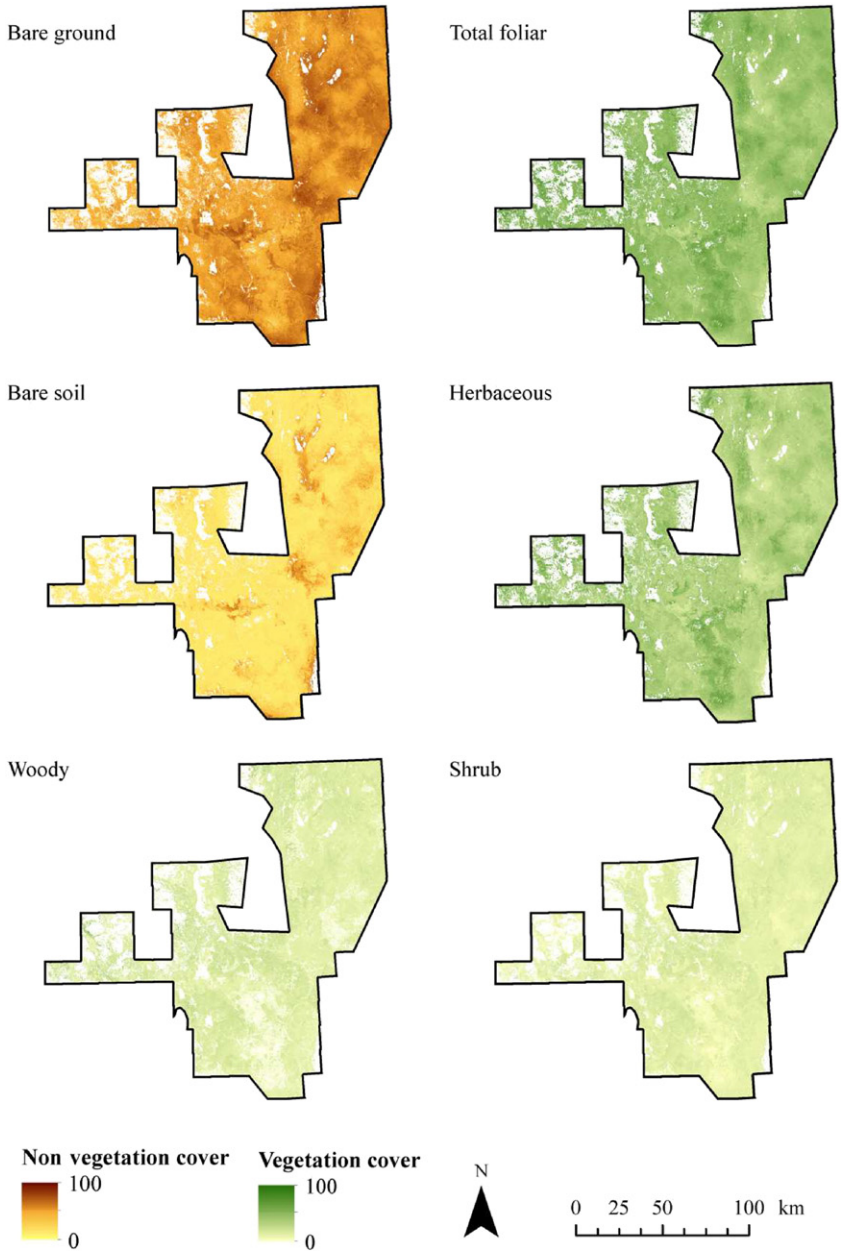

Figure 3. Percent fractional cover of each indicator: bare ground, total foliar, bare soil, herbaceous, woody, and shrub.

(e.g., Boyte and Wylie, 2015), measures of model overfit metrics have not been commonly reported in multiscale rangeland fractional cover remote sensing (Homer et al., 2012; Mishra et al., 2014; Sant et al., 2014); thus, it is not apparent the extent to which model overfit has occurred elsewhere or how the model overfit results of this study compare with previous works.

Spatial Autocorrelation

Multicollinearity and spatial autocorrelation among predictor variables with linear regression techniques can result in inflated accuracy statistics and erratic model coefficients (i.e., sensitivity to small changes in input data or model structure). While regression tree approaches are resistant to multicollinearity effects (Wylie et al., 2007), spatial autocorrelation in the input variables can still lead to model overfit. Tendencies of the BART models toward model overfit may also be related more to spatial autocorrelation than to the number of training samples or predictor variables (Telford and Birks, 2005; Mascaro et al., 2014). Our sample sufficiency analyses (see Figs. 5 and 6) showed that the model overfit ratio became smaller (i.e., closer to one) with increasing sample size. Overfit ratios converged to their minimum values at sample sizes below what was available for this study. Preliminary implementations of the BART models for this study included the original sensor band values and additional topography, texture, and vegetation indices as these predictors have been used in other remote sensing studies (Marsett et al., 2006; Homer et al., 2012; Mishra et al., 2014). However, these models produced out-of-sample RMSE values that were, on average, 29% higher relative to models that used just the first three PC bands. This higher out-of-sample RMSE may have been due to multicollinearity among the PCA bands and the additional predictor variables.

Spatial filtering or spatial autoregression has been used to remove autocorrelation from the predictor variables to reduce model overfit in regression trees (Telford and Birks, 2005; Dormann et al., 2007). Alternatively, geographic coordinates can be added to the regression tree as input variables (Mascaro et al., 2014) or measures of spatial autocorrelation could be incorporated into the model, as it is done in regression

Figure 4. Credible interval width of each percent cover indicator: bare ground, total foliar, bare soil, herbaceous, woody, and shrub.

kriging (Karl, 2010). Future research should examine the influence of spatial autocorrelation on multiple spatial scale model performance, model overfit, and selection of predictor variables when incorporating spatially autocorrelated predictor variables in a remote sensing model (*sensu* Gu et al., 2016).

Spatial Resolution

Differences in performance between the RapidEye-only, Landsat-only, and Landsat-RapidEye models may be due, in part, to differences in the spatial resolution of the model input data. Although RMSE values decreased from RapidEye-only to Landsat-only to Landsat-RapidEye, the differences in spatial resolution between the models were especially apparent when examining model overfit. Model overfit can be the result of too much variability within the predictors (Balac et al., 2013). In the RapidEye-only model, the increased spatial complexity associated with increased spatial resolution may be influencing model overfit. The RapidEye-only model, which had the highest density of field samples (1 per 38 km^2) also had the greatest degree of model overfit (see Table 1). This could be due to the aggregation of an average of 45 RapidEye pixels per training plot to create a direct relationship between the PC bands and the field-sampled area. In contrast, the Landsat pixels are more similar in size to the training plots; thus, fewer pixels were aggregated for each field plot. This commonality in scale between the spatial resolution of the plot and the Landsat pixel may contribute to lower model overfit.

BART Limitations

Despite the advantages of BART and successful implementation demonstrated here, its computationally intensive nature and the technical expertise required to implement it relative to other more traditional image classifiers may constrain its implementation for land management applications. Additionally, BART is a relatively new regression tree approach (Chipman et al., 2010) and its application to remote sensing in ecology and rangeland management is novel. Further

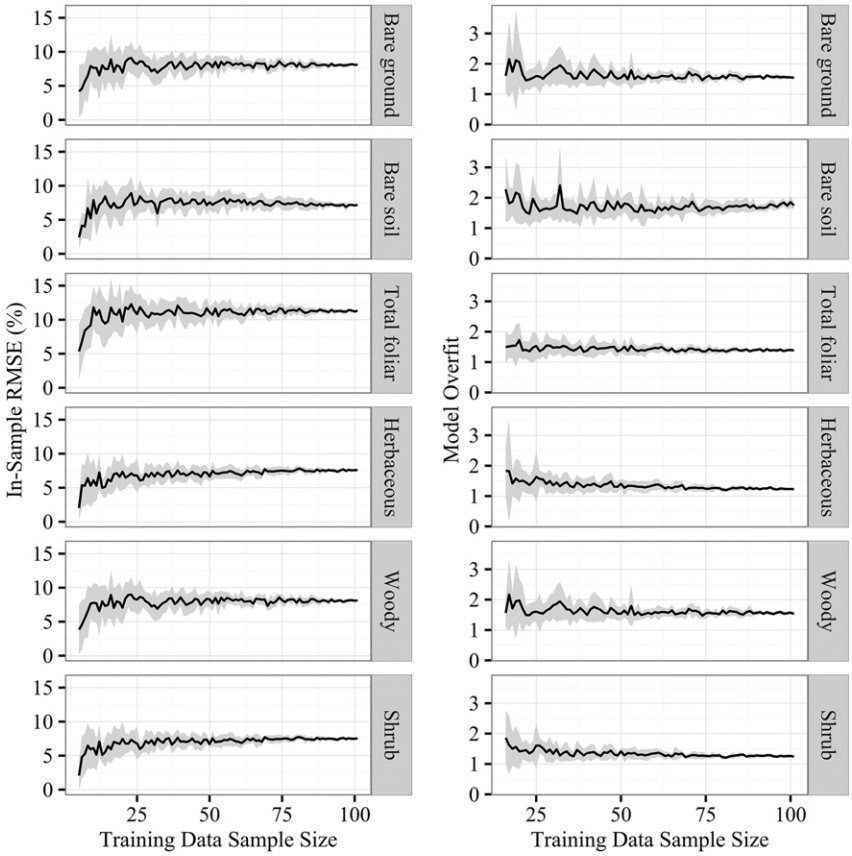

Figure 5. Relationship between field training data sample size and RapidEye-only BART model performance as indicated by root mean square error and model overfit. The range between minimum and maximum values is shown in gray.

research and exploration into its performance in a variety of ecosystem types are warranted.

Indicator Estimates

Overall, indicator predictions behaved as expected relative to one another. In the summation of bare ground and total foliar predictions, 99.99% of the predictions were between 90% and 110% cover and 82.8% of the predictions between 95% and 99% cover. This is an improvement upon Homer et al. (2012), who performed a similar analysis, albeit with a lower sampling density (1 plot per 129 km^2 vs. 1 plot per 80 km^2 in this study), and reported that 93% of the predictions were between 90% and 110% cover and 73% of the predictions between 95% and 100% cover. This indicates, however, that our BART predictions may be slightly underestimating cover across the landscape. Bare ground and total foliar cover had identical pseudo-R^2 and RMSE values in each model (0.80 in the Landsat-RapidEye model), a result expected from a robust model given that the two indicators are inverse of one another.

Woody and shrub models had lower pseudo-R^2 and RMSE values than the other indicators. In addition, 16% of the shrub predictions exceeded the predicted woody cover. This suggests that these models were not as robust as the other indicator models. Reduced woody and shrub model performance is seen in similar studies in rangelands (Gessner et al., 2013; Xian et al., 2015) and may be due to the complex structure of woody canopy and variations in the spectral signatures of woody plants, which results in a more dimensional response than that of bare soil or herbaceous cover. For the purposes of monitoring juniper encroachment and managing other woody species dynamics using the current modeling approach, additional training data may be required in those areas, or different seasonal imagery may be necessary to better distinguish between woody vegetation and herbaceous vegetation in rangeland systems.

The spatial distribution of the six rangeland indicators (see Figs. 3 and 4) show that the BART models were able to predict changes in vegetation related to difference in fire history and topography. For instance, increased herbaceous cover and decreased woody cover in the southeastern portion of the study area coincides with boundaries of the 2012 Rush Fire. In unburned areas, bare soil is higher in playas and lowland areas and lower on high-elevation sites. Similarly, increased woody vegetation and total foliar cover in unburned areas correspond with increasing elevation. This is to be expected as higher-elevation rangelands typically receive more precipitation and thus typically produce greater amounts of vegetation (Lomolino, 2001).

Sampling Sufficiency

Sampling sufficiency analysis showed that the sample sizes available through the AIM dataset were sufficient to develop BART-based fractional cover models. Previous fractional cover studies (Karl, 2010; Xian et al., 2015) have relied on higher field sampling densities than those used in this study. These sample sufficiency results indicate that fractional cover products may be produced at lower costs by not only using freely available data but also applying a model that required fewer training samples overall. The data available through AIM, however, may not capture the spatial variability necessary to train finer cover classes adequately (e.g., shrub, herbaceous cover); thus, targeted supplemental data collection where possible may be beneficial to improve model performance or where the timing of AIM data collection is incompatible with modeling efforts.

The amount of data required to develop and test a BART model may also be dependent on spatial resolution. The RapidEye-only model required at least one sample plot per 76 km^2 to converge to a stable pseudo-R^2, while the Landsat-only model required one sample plot for every 146 km^2. This suggests that the amount of data required to train

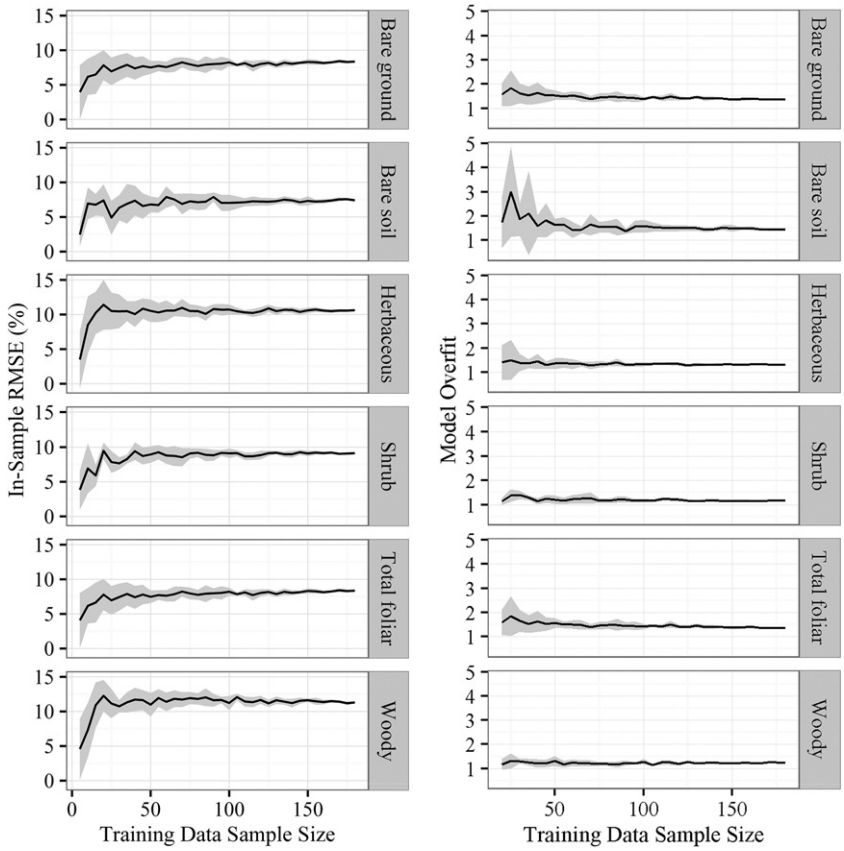

Figure 6. Relationship between field training data sample size and Landsat-only BART model performance as indicated by root mean square errors and model overfit. The range between minimum and maximum values is shown in gray.

high spatial resolution imagery may be greater than the amount required for moderate spatial resolution imagery. Consequently, when considering sample sizes, the spatial resolution of the predictor variables and number of pixels to be predicted should be considered in addition to the spectral variability, vegetation complexity, and size of the study area.

Uncertainty

One distinct advantage of this study in comparison with many other rangeland remote sensing studies (Homer et al., 2012; Gessner et al., 2013; Mishra et al., 2014; Xian et al., 2015) is the focus on providing spatially explicit uncertainty estimates in addition to indicator predictions. Bayesian uncertainty frameworks provide a credible interval in which the true population parameter falls, thereby estimating both model error and uncertainty from other sources (e.g., training data, predictor variables). Spatially explicit credible interval ranges provide an additional metric for understanding model performance as it varies over space. Previous fractional cover studies (Mishra et al., 2014; Sant et al., 2014; Homer et al., 2015) reported overall map accuracy, but did not report spatial measures of prediction uncertainty, even when the model used (e.g., Cubist) has the ability to report such measures of model uncertainty. This spatial understanding of uncertainty is useful for land managers and researchers because it helps to interpret landscape patterns as well as determine the level of confidence placed in the fractional cover product overall as well as in specific areas, and thus these uncertainty products should be promoted more rigorously by the modelling community. In this study credible interval widths were higher within the fire boundaries and in the northeastern portion of the study area (Figure 4). The increase in uncertainty in post-burn areas may be due to an uneven vegetation response post-fire while the general increase in uncertainty in the northern portion of the study area may be the result of insufficient training data. Understanding the cause of the uncertainty is important for interpretation of the data and further exploration is needed to understand the uncertainty patterns in the prediction maps in this study.

Land managers and others interpreting and making decisions with indicator prediction maps can incorporate the level of uncertainty in the indicators at the spatial scale and in the exact area where the decision is being made as an additional line of evidence to support a specific decision. If the level of uncertainty in a particular region of the study area is deemed unacceptable (e.g., due to insufficient training data), future field data collection efforts may be focused in areas where predictions are uncertain (McCarthy, 2007). Subsequent field data could then be used to update the BART model, thereby creating a feedback mechanism between field data collection and remote sensing models to inform adaptive management. As a result, model predictions improve over time to provide better information to managers (Holst, 1992; Williams et al., 2011).

Understanding the role of error propagation (i.e., uncertainty accumulation; Hunsaker et al., 2001) in remote sensing models is also important (Wang et al., 2005). In this study, we estimated only uncertainty and not error propagation, but others (e.g., Wilson et al., 2011) have proposed using alternate hierarchical Bayesian models to incorporate the uncertainty of each dataset more explicitly than BART. With the current implementation of BART in *bartMachine*, error is propagated as a global parameter based on the dataset variability (Chipman et al., 2010). However, sampling and non-sampling errors can vary on a local scale as a result of temporal and spatial variation in field plots and image scenes. BART currently allows the user to specify hyperparameters that govern which predictors may by more informative or have less uncertainty overall, but BART does not allow a likelihood function to be assigned to training data points individually. Thus the uncertainty associated with individual data points cannot be

explicitly included in the model. In a multi-scale model, such as the Landsat-RapidEye model, error could also be propagated from the field to the RapidEye-only predictions and then to the Landsat predictions. This error propagation could be modeled by assigning, for example, higher uncertainty to the modeled training data than to the field data in the Landsat-RapidEye model. Future research should explore ways to explicitly include error propagation in BART models.

Management Implications

In this study, we proposed and tested a multi-scale BART model that could be used to provide efficient and robust rangeland indicator information to land managers using existing field datasets. Data collected as part of rangeland monitoring programs with consistent methods such as AIM are a promising source of training data for both prospective and retrospective remote sensing studies. Existing field data from monitoring programs could not only decrease the costs of remote sensing efforts, but help ensure consistency in indicator definitions between remote sensing studies and studies based solely on field data (e.g., Karl et al., 2017; Herrick et al., 2010). Although the sample design of existing monitoring programs may not be sufficient to capture the range of training and testing data needed for remote sensing studies, these existing designs could be easily supplemented at targeted location by using consistent field data collection methods. Similar to other regression tree approaches, using a multiscale modeling approach that scales predictions from RapidEye to Landsat improved BART model performance. This multiscale approach can be used when high spatial resolution imagery is available to supplement field data. In contrast to other regression tree approaches, however, the BART model also provided uncertainty estimates which are valuable ancillary information for land managers who use BART indicator predictions for adaptive management. The added uncertainty information, together with efficiencies gained from using available field data, provide a valuable framework for using remotely sensed imagery to derive cover indicators in other rangeland ecosystems.

Acknowledgments

RapidEye imagery was provided by the BLM's National Operations Center. We would also like to thank the BLM Applegate and Eagle Lake field offices, especially Dereck Wilson, Colleen Dulin, Elias Flores, and Michael Dolan. The AIM field crews who collected the field data for this study were instrumental in the success of this project: Alexander Traynor, Amanda Smith, Ben Faucher, Daniel Wright, Henry Palamores, Jemma Williams, Kelsey Gilstrap, Robert Lee, and Ryan Watson.

References

Balac, N., Sipes, T., Wolter, N., Nunes, K., Sinkovits, B., Karimabadi, H., 2013. Large scale predictive analytics for real-time energy management. Big data. Santa Clara, CA, USA: 2013 IEEE International Conference, pp. 657–664.

Blackbridge, A.G., 2015. Blackbridge.Available at:. http://blackbridge.com/rapideye/mosaics/index.html Accessed 25 September 2015.

Bleich, J., Kapelner, A., George, E.I., Jensen, S.T., 2014. Variable selection for BART: an application to generegulation. The Annals of Applied Statistics 8 (3), 1750–1781.

Booth, D.T., Tueller, P.T., 2003. Rangeland monitoring using remote sensing. Arid Land Research and Management 17 (4), 455–467.

Boyte, B.S.P., Wylie, B.K., 2015. Near-real-time cheatgrass percent cover in the Northern Great Basin, USA, 2015. Rangeland Ecology & Management 38 (5), 278–284.

Breiman, L., 2001. Random forests. Machine Learning 45 (1), 5–32.

Breiman, L., Friedman, J., Ohlsen, R., Stone, C., 1984. Classification and regression trees. Wadsworth International Group, Belmont, CA, USA.

Chipman, H.A., George, E.I., McCulloch, R.E., 2010. BART: Bayesian additive regression trees. Annals of Applied Statistics 6 (1), 266–298.

Clark, J.S., 2007. Models for ecological data. first ed. Princeton University Press, Princeton, NJ, USA.

De 'ath, G., Fabricius, K.E., 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81 (11), 3178–3192.

Desert Research Institute, 2015. Period of record monthly climate summary: Cedarville, California.Available at:. http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca1614 Accessed 25 September 2015.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30 (5), 609–628.

Ellison, A.M., 2004. Bayesian inference in ecology. Ecology Letters 7 (6):509–520 Available at:. http://doi.wiley.com/10.1111/j.1461-0248.2004.00603.x Accessed 6 November 2013.

Gessner, U., Machwitz, M., Conrad, C., Dech, S., 2013. Estimating the fractional cover of growth forms and bare surface in savannas. A multi-resolution approach based on regression tree ensembles. Remote Sensing of Environment 129, 90–102.

Gu, Y., Wylie, B.K., 2015. Developing a 30-m grassland productivity estimation map for central Nebraska using 250-m MODIS and 30-m Landsat-8 observations. Remote Sensing 171 (2015), 291–298.

Gu, Y., Wylie, B.K., Boyte, S.P., Picotte, J., Howard, D.M., Smith, K., Nelson, K.J., 2016. An optimal sample data usage strategy to minimize overfitting and underfitting effects in regression tree models based on remotely-sensed data. Remote Sensing 8 (943), 20–26.

Herrick, J.E., Lessard, V.C., Spaeth, K.E., Shaver, P.L., Dayton, R.S., Pyke, D.A., Jolley, L., Goebel, J.J., 2010. National ecosystem assessments supported by scientific and local knowledge. Frontiers in Ecology and the Environment 8 (8), 403–408.

Herrick, J.E., Van Zee, J.W., Havstad, K.M., Burkett, L.M., Whitford, W.G., 2009. Monitoring manual for grassland, shrubland and savanna ecosystems. Volume 1. Quick start. volume ii: design, supplementary methods and interpretation. USDA-ARS Jornada Experimental Range, Las Cruces, NM.

Hijmans, R.J., van Etten, J., 2012. Raster: geographic analysis and modeling with raster data. R package version 2.0–12.

Holst, J.D., 1992. Unforeseeability factor: federal lands, managing for uncertainty, and the preservation of biological Diversity. Public Land and Resources Law Review 13, 113.

Homer, C.G., Aldridge, C.L., Meyer, D.K., Schell, S.J., 2012. Multi-scale remote sensing sagebrush characterization with regression trees over Wyoming, USA: laying a foundation for monitoring. International Journal of Applied Earth Observation and Geoinformation 14 (1), 233–244.

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., Megown, K., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States–representing a decade of land cover change information. Photogrammetric Engineering & Remote Sensing 81 (5), 345–354.

Homer, C.G., Meyer, D.K., Aldridge, C.L., Schell, S.J., 2013. Detecting annual and seasonal changes in a sagebrush ecosystem with remote sensing-derived continuous fields. Journal of Applied Remote Sensing 7 (1), 073508.

Hunsaker, C.T., Goodchild, M.F., Friedl, M.A., Case, T.J., 2001. Spatial uncertainty in ecology. 1st ed. Springer Verlag, New York.

Hunt, E.R., Everitt, J.H., Ritchie, J.C., Moran, M.S., Booth, D.T., Anderson, G.L., Clark, P.E., Seyfried, M.S., 2003. Applications and research using remote sensing for rangeland management. Photogrammetric Engineering and Remote Sensing 69 (6), 675–693.

Jensen, J.R., 2005. Introductory digital image processing: a remote sensing perspective. third ed. Prentice Hall, Ann Arbor, MI, USA 318 pp.

Kapelner, A., Bleich, J., 2016. bartMachine: machine learning with Bayesian Additive Regression Trees. Journal of Statistical Software 70 (4).

Karl, J.W., 2010. Spatial predictions of cover attributes of rangeland ecosystems using regression kriging and remote sensing. Rangeland Ecology & Management 63 (3), 335–349.

Karl, J.W., Duniway, M.C., Nusser, S.M., Opsomer, J.D., Unnasch, R.S., 2012. Using very-large-scale aerial imagery for rangeland monitoring and assessment: some statistical considerations. Rangeland Ecology & Management 65 (4), 330–339.

Karl, J.W., McCord, S.E., Hadley, B.C., 2017. A comparison of cover calculation techniques for relating point-intercept vegetation sampling to remote sensing imagery. Ecological Indicators 73, 156–165.

Karl, J.W., Taylor, J., Bobo, M., 2014. A double-sampling approach to deriving training and validation data for remotely-sensed vegetation products. International Journal of Remote Sensing 35 (5), 1936–1955.

Laliberte, A.S., Fredrickson, E.L., Rango, A., 2007. Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands. Photogrammetric Engineering and Remote Sensing 73 (2), 197–207.

Lomolino, M., 2001. Elevation gradients of species-density: historical and prospective views. Global Ecology and Biogeography 10 (1), 3–13.

Ludwig, J.a., Bastin, G.N., Chewings, V.H., Eager, R.W., Liedloff, A.C., 2007. Leakiness: A new index for monitoring the health of arid and semiarid landscapes using remotely sensed vegetation cover and elevation data. Ecological Indicators 7 (2):442–454 Available at:. http://linkinghub.elsevier.com/retrieve/pii/S1470160X06000410 Accessed 7 August 2013.

Ludwig, J.A., Bastin, G.N., Eager, R.W., Ketner, P., Pearce, G., 2000. Monitoring Australian rangeland sites using landscape function indicators and ground- and remote-based techniques. Environmental Monitoring and Assessment 1670178, 167–178.

MacKinnon, W., Karl, J.W., Taylor, J., Karl, M., Spurrier, C., Herrick, J.E., 2011. BLM core terrestrial indicators and methods. BLM Technical Note, Denver, CO, USA.

Maindonald, J., Braun, J., 2006. Data analysis and graphics using R: an example-based approach. Cambridge University Press, New York, NY, USA 549 pp.

Marsett, R.C., Qi, J., Heilman, P., Biedenbender, S.H., Watson, M.C., Amer, S., Weltz, M., Goodrich, D., Marsett, R., 2006. Remote sensing for grassland management in the arid Southwest. Rangeland Ecology & Management 59 (5), 530–540.

Mascaro, J., Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Martin, R.E., Anderson, C., Higgins, M., Chadwick, K.D., 2014. A tale of two "forests": Random Forest machine learning aids tropical forest carbon mapping. PLoS One 9 (1), 12–16.

McCarthy, M.A., 2007. Bayesian methods for ecology. first ed. Cambridge University Press, New York, NY, USA 310 pp.

Mishra, N.B., Crews, K.A., 2014. Mapping vegetation morphology types in a dry savanna ecosystem : integrating hierarchical object-based image analysis with Random Forest. International Journal of Remote Sensing 35 (3), 1175–1198.
Mishra, N.B., Crews, K.A., Okin, G.S., 2014. Relating spatial patterns of fractional land cover to savanna vegetation morphology using multi-scale remote sensing in the Central Kalahari. International Journal of Remote Sensing 35 (6), 2082–2104.
Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrology and Earth System Sciences Discussions 4, 439–473.
Pratola, M.T., Chipman, H.A., Gattiker, J.R., Higdon, D.M., Mcculloch, R., Rust, W.N., 2014. Parallel Bayesian additive regression trees. Journal of Computational and Graphical Statistics 23 (3), 830–852.
R Core Team, 2015. R: a language and environment for statistical computing.Available at:. https://www.r-project.org/ Accessed 7 September 2015.
RuleQuest, 2012. Data mining with cubist.Available at:. http://www.rulequest.com/cubist-info.html Accessed 1 January 2015.
Sant, E.D., Simonds, G.E., Ramsey, R.D., Larsen, R.T., 2014. Assessment of sagebrush cover using remote sensing at multiple spatial and temporal scales. Ecological Indicators 43, 297–305.
Segal, M.R., Barbour, J.D., Grant, R.M., 2004. Relating HIV-1 sequence variation to replication capacity via trees and forests. Statistical Applications in Genetics and Molecular Biology 3 (1), 1–18.
Stevens, D.L., Olsen, A.R., 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99 (465), 262–278.
Strobl, C., Malley, J., Tutz, G., 2009. An introduction to recursive partitioning: rationale, application and characteristics of classification and regression trees, bagging and random forests. Psychological Methods 14 (4), 323–348.
Taylor, J., Kachergis, E., Toevs, G., Karl, J.W., Bobo, M., Karl, M., Miller, S., Spurrier, C., 2014. AIM-monitoring: a component of the BLM assessment, inventory, and monitoring strategy. AIM, Denver, CO, USA.
Telford, R.J., Birks, H.J.B., 2005. The secret assumption of transfer functions: Problems with spatial autocorrelation in evaluating model performance. Quaternary Science Reviews 24 (20-21), 2173–2179.
Toevs, G., Karl, J., Taylor, J., Spurrier, C., Karl, M., Bobo, M., Herrick, J.E., 2011. Consistent indicators and methods and a scalable sample design to meet assessment, inventory, and monitoring information needs across scales. Rangelands 33, 14–20.
USDI BLM, 2012. Rush Fire emergency stabilization and rehabilitation Environmental Assessment.Available at:. http://www.blm.gov/style/medialib/blm/ca/pdf/eaglelake/final_rush_fire_ea.Par.88873.File.dat/FINAL_RushFire Restoration_EA_RSN_11-27-12.pdf Accessed 15 March 2015.
USDI United States Geological Survey Earth Resources Observation and Science Center, 2015. Earth explorer.Available at:. http://earthexplorer.usgs.gov/ Accessed 4 May 2015.
Walton, J.T., 2008. Subpixel urban land cover estimation: comparing Cubist, Random Forests and support vector regression. Photogrammetric Engineering & Remote Sensing 74 (10), 1213–1222.
Wang, G., Gertner, G.Z., Fang, S., Anderson, A.B., 2005. A methodology for spatial uncertainty analysis of remote sensing and GIS products. Photogrammetric Engineering & Remote Sensing 71 (12), 1423–1432.
Washington-Allen, R.A., West, N.E., Ramsey, R.D., Efroymson, R.A., 2006. Applications of geospatial techniques a protocol for retrospective remote sensing-based ecological monitoring of rangelands. Rangeland Ecology & Management 59 (1), 19–29.
Williams, B.K., Eaton, M.J., Breininger, D.R., 2011. Adaptive resource management and the value of information. Ecological Modelling 222 (18), 3429–3436.
Wilson, A.M., Silander, J.A., Gelfand, A., Glenn, J.H., 2011. Scaling up: linking field data and remote sensing with a hierarchical model. International Journal of Geographical Information Science 25 (3, SI), 509–521.
Wylie, B.K., Fosnight, E.A., Gilmanov, T.G., Frank, A.B., Morgan, J.A., Haferkamp, M.R., AND Meyers, T.P., 2007. Adaptive data-driven models for estimating carbon fluxes in the Northern Great Plains. Remote Sensing of Environment 106, 399–413.
Xian, G., Homer, C., Rigge, M., Shi, H., Meyer, D., 2015. Characterization of shrubland ecosystem components as continuous fields in the northwest United States. Remote Sensing of Environment 168, 286–300.