2017

Cheatgrass (Bromus tectorum) distribution in the intermountain Western United States and its relationship to fire frequency, seasonality, and ignitions

Bethany A. Bradley  
*University of Massachusetts Amherst*

Caroline A. Curtis  
*University of Massachusetts Amherst*

Emily J. Fusco  
*University of Idaho*

John T. Abatzoglou  
*University of Idaho*

Jennifer K. Balch  
*University of Colorado*

Follow this and additional works at: [https://scholarworks.umass.edu/nrc_faculty_pubs](https://scholarworks.umass.edu/nrc_faculty_pubs)

Part of the [Environmental Monitoring Commons](https://scholarworks.umass.edu/ems_commons) and the [Natural Resources and Conservation Commons](https://scholarworks.umass.edu/nrc_commons)

Recommended Citation

Bradley, Bethany A.; Curtis, Caroline A.; Fusco, Emily J.; Abatzoglou, John T.; Balch, Jennifer K.; Dadashi, Sepideh; and Tuanmu, Mao-Ning, "Cheatgrass (Bromus tectorum) distribution in the intermountain Western United States and its relationship to fire frequency, seasonality, and ignitions" (2017). *Biological Invasions*. 400. 10.1007/s10530-017-1641-8

This Article is brought to you for free and open access by the Environmental Conservation at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Environmental Conservation Faculty Publication Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
Cheatgrass (*Bromus tectorum*) distribution in the intermountain western United States and its relationship to fire frequency, seasonality, and ignitions

Bethany A. Bradley¹, Caroline A. Curtis², Emily J. Fusco², John T. Abatzoglou³, Jennifer K. Balch⁴, Sepideh Dadashi⁴, Mao-Ning Tuanmu⁶

1. Department of Environmental Conservation, University of Massachusetts, Amherst MA 01003
2. Graduate Program in Organismic and Evolutionary Biology, University of Massachusetts, Amherst MA 01003
3. Department of Geography, University of Idaho, Moscow, ID 83844
4. Earth Lab, CIRES, University of Colorado, Boulder, CO 80309
5. Department of Geography, University of Colorado, Boulder, CO 80309
6. Biodiversity Research Center, Academia Sinica, Taipei 11529, Taiwan

*Corresponding Author

Email: bbradley@eco.umass.edu

Phone: 413 545 1764

ORCID: 0000-0003-4912-4971
Abstract

Cheatgrass (*Bromus tectorum*) is an invasive grass pervasive across the Intermountain Western US and linked to major increases in fire frequency. Despite widespread ecological impacts associated with cheatgrass, we lack a spatially extensive model of cheatgrass invasion in the Intermountain West. Here, we leverage satellite phenology predictors and thousands of field surveys of cheatgrass abundance to create regional models of cheatgrass distribution and percent cover. We compare cheatgrass presence to fire probability, fire seasonality and ignition source. Regional models of percent cover had low predictive power (34% of variance explained), but distribution models based on a threshold of 15% cover to differentiate high abundance from low abundance had an overall accuracy of 74%. Cheatgrass achieves ≥15% cover over 210,000 km² (31%) of the Intermountain West. These lands were twice as likely to burn as those with low abundance, and four times more likely to burn multiple times between 2000-2015. Fire probability increased rapidly at low cheatgrass cover (1-5%) but remained similar at higher cover, suggesting that even small amounts of cheatgrass in an ecosystem can increase fire risk. Abundant cheatgrass was also associated with a 10 day earlier fire seasonality and interacted strongly with anthropogenic ignitions. Fire in cheatgrass was particularly associated with human activity, suggesting that increased awareness of fire danger in invaded areas could reduce risk. This study suggests that cheatgrass is much more spatially extensive and abundant than previously documented and that invasion greatly increases fire frequency, even at low percent cover.

**Keywords:** Bromus tectorum; fire regime alteration; grass-fire cycle; invasive grass; invasive plant; Moderate Resolution Imaging Spectroradiometer (MODIS) burned area product
Introduction

Biogeographical analysis of invasive plants aims to use species distributions to assess broad scale macroecological questions such as invasion risk (e.g., Allen and Bradley 2016), niche shifts (e.g., Petitpierre et al. 2012), and invader impacts (e.g., Balch et al. 2013). Much of our understanding of invasion biogeography is dependent on distribution data acquired through museum collections or through contributed databases of management records (e.g., Early Detection and Distribution Mapping System; EDDMapS; Bargeron and Moorhead 2007). However, these databases tend to contain spatial biases (Pyšek et al. 2008; Cross et al. 2017) and over-represent populations with low abundance (Bradley 2013), potentially leading to an inaccurate representation of the invaded range. It may be possible to detect abundant populations for a subset of invasive plants using airborne or satellite remote sensing (Bradley 2014), thereby creating distribution datasets that are well suited for understanding the spatial attributes of invasion and invader impacts. One example of an invasive plant that, under ideal conditions, can be detected remotely is cheatgrass (*Bromus tectorum*).

Cheatgrass, a non-native annual grass, is one of the most notorious and problematic invasive plants in the United States. The grass primarily colonizes and becomes established in shrublands of the Intermountain West, including the Great Basin, where it is negatively associated with the richness and abundance of small mammals (Freeman et al. 2014; Holbrook et al. 2016) and reduces habitat quality for Greater Sage-Grouse (*Centrocercus urophasianus*; e.g., Crawford et al. 2004; Lockyer et al., 2015), a species of concern for regional management. Cheatgrass also increases the abundance and connectivity of fine fuels, and has altered fire frequency throughout the Intermountain West (D’Antonio and Vitousek 1992; Brooks and Pyke 2001; Pilliod et al. 2017). A regional analysis across the southern portion of the Great Basin concluded that the presence of cheatgrass more than doubles the fire frequency that would be expected in native land cover (Balch et al., 2013). Following fire, cheatgrass becomes dominant and recovery of native shrubs and grasses is unlikely, particularly in relatively hot and dry regions (Davies et al. 2012; Knapp 1996; Taylor et al. 2014). These negative ecological impacts underscore the need to refine our understanding of the distribution of cheatgrass throughout its non-native range and its relationship to fire. To better understand invasion risk and fully characterize the cheatgrass-fire cycle, regional models of cheatgrass distribution and abundance are needed.

Unique phenological traits of cheatgrass, relative to the native plant community, potentially enable detection and mapping with satellite imagery. Non-native invasive plants often outcompete native species because earlier phenology gives them a priority effect (Wolkovich and Cleland 2011). These phenological differences can also facilitate satellite detection of non-native invasive plants (Bradley 2014). For example, cheatgrass becomes productive in spring, earlier than native shrubs and grasses. As cheatgrass becomes more abundant, the timing of peak greenness shifts, and this change is observable with satellite time series (Peterson 2005). Additionally, cheatgrass responds strongly to precipitation (Stewart and Hull 1949), which is highly variable year to year in the Intermountain West. This inter-annual variability in greenness is also detectable with satellite time series and relates to the distribution of cheatgrass (Bradley and Mustard 2005). Although the phenological differences of cheatgrass make its distribution potentially detectable, many other regional variations in the landscape reduce the accuracy of resulting models. For example, spatial variability in soil brightness and temporal
variability in cloud cover can confound vegetation signals, while smaller-scale variations in topography and rainfall can affect phenological response.

Nonetheless, the temporal and phenological signals of cheatgrass relative to those of native vegetation have led to multiple models of cheatgrass’ distribution. These studies leverage several satellite platforms, but each covers only a portion of the Intermountain West where cheatgrass interacts with fire. For example, satellite data from Landsat, Moderate Resolution Imaging Spectroradiometer (MODIS), and Advanced Very High Resolution Radiometer (AVHRR) have all been used to model cheatgrass distribution in subsets of the Great Basin including Nevada and Utah (Peterson et al. 2005, 2006; Boyte et al. 2016), the Snake River Plain of Idaho (Clinton et al. 2010), and the southern Great Basin (Bradley and Mustard 2005, 2008). Only Downs et al. (2016) model the broader extents of the historical range of Greater Sage-Grouse across the Intermountain West, but their analysis does not consider relationships with fire.

Another limitation of existing satellite-derived models of cheatgrass’ distribution is a lack of training data used to train and test spatial models. Vegetation surveys are often conducted at small spatial resolutions (e.g., 1 m² quadrats) or along linear transects, which are not necessarily good proxies for vegetation cover observed by satellites within pixel sizes ranging from hundreds of square meters to hectares (Bradley 2014). As a result, many maps of cheatgrass derived from remote sensing have relied on limited field data collected by the authors, which can result in relatively sparse training and testing data that rarely include information on percent cover. For example, both Bradley and Mustard (2008) and Clinton et al. (2010) defined presence as large, homogeneous patches of abundant cheatgrass; they trained models with only 487 and 92 occurrence (presence or absence) points, respectively. Bradley and Mustard (2005) collected testing data at 659 points, but these data were limited to a portion of the map area and were based on qualitative rather than quantitative estimates of cover (e.g., dense understory). Boyte et al. (2016) used 35,000 training points to model percent cover of cheatgrass in the Snake River Plain, but these data were derived from another remote sensing product that itself was poorly correlated ($R^2 = 0.29$) with observed percent cover (Peterson 2006). Only two studies to date have collected or compiled training data throughout the modeled study region and in sufficient quantity to perform a robust accuracy assessment. Peterson (2006) collected 806 field estimates of percent cover of cheatgrass throughout Nevada to train and test a model of percent cover, and Downs et al. (2016) compiled data from 6650 field surveys throughout the Intermountain West to train and test a model of presence. As a result, additional assessments of cheatgrass presence and percent cover across the Great Basin based on regionally extensive training data are needed.

Improving the accuracy of spatial models of cheatgrass presence and percent cover will lead to better understanding of the relations between cheatgrass and regional fire regimes. While the grass-fire cycle associated with cheatgrass has long been known (Brooks and Pyke 2001; D’Antonio and Vitousek 1992), its regional impact was only recently quantified by combining a remotely sensed map with fire records (Balch et al. 2013). Balch et al. (2013) found that presence of cheatgrass was associated with a doubling of fire frequency and some of the largest fires in the Great Basin. Their analysis also revealed strong inter-annual variability in burned area, which is closely associated with an increase in fuel loads (largely cheatgrass) following wet growing seasons (Balch et al. 2013). Furthermore, the presence of
cheatgrass across parts of the Great Basin has augmented fire activity beyond what pyrogeographic models predict based on climatic factors alone (Parks et al. 2015). However, these assessments were based on a remotely sensed product at 1 km resolution, with an overall accuracy of 61%, which was derived from annual variability observed during the 1990s (Bradley and Mustard, 2008). Quantifying relations between cheatgrass and fire at higher resolution, and developing a contemporary map of cheatgrass, could refine assessments of the response of fire dynamics to cheatgrass.

Here, we present a satellite-based model of cheatgrass distribution for the full extent of the hydrographic Great Basin over the 2001-2014 time period. We compiled extensive data from field surveys for model training and testing and used MODIS observations of both early productivity and annual variability as predictor variables. We use this updated model to reevaluate relations between cheatgrass and fire frequency over a study region larger than that analyzed by Balch et al. (2013). Additionally, we expand previous fire analyses by assessing relations between percent cover of cheatgrass and fire frequency, differences in fire seasonality in regions in which cheatgrass is dominant, and links to anthropogenic ignition sources.
Methods

Cheatgrass percent cover

We compiled data from 15 datasets that contained estimates of percent cover of cheatgrass within the Great Basin ecoregion (Table 1), including points with cheatgrass absent. Where reported in metadata, percent cover estimates were typically derived from either point-line transects or transects of ocular estimates (e.g., Daubenmire frames). Where transect lengths were reported, transects ranged in length from 30 – 150 m. We excluded data from the Southwest Exotic Plant Mapping Program (Thomas and Guertin 2007) because, although spatially extensive, they typically referred to a small infested area and were based on qualitative, visual estimates of percent cover. For datasets that reported a range of cover values (e.g., 1-5%), we used the mean value (3%). Since the mapping focus was on terrestrial native land cover, we excluded all cheatgrass points located on agricultural or developed land based on the Landfire existing vegetation type classification (EVT 130; Rollins 2009). We limited the extents of the analysis to the hydrographic Great Basin (USGS 2017).

We scaled-up all transect and plot-level percent cover measurements by calculating the average percent cover of all co-occurring data within a 250 m grid cell. This spatial resolution was consistent with the finest resolution of the phenology predictor layers derived from MODIS (see below). We chose to use average percent cover assuming that MODIS reflectance and phenology responds to average cover within a pixel. Although none of the datasets were collected at a 250 m resolution, cheatgrass cover is likely to be spatially auto-correlated across relatively large extents (Bradley and Mustard 2005; Peterson 2005). Therefore, we assumed that these data provide a reasonable approximation of cheatgrass cover for the corresponding MODIS pixel.

Predictor layers

Cheatgrass primarily grows in early spring, whereas most native vegetation growth occurs in summer. This difference in timing has been used to differentiate cheatgrass on the basis of remotely sensed estimates of phenology (Boyte et al. 2015; Peterson 2005). Additionally, cheatgrass (an annual grass) has much higher inter-annually variability in biomass than that of native vegetation (primarily perennial species). For example, cheatgrass biomass has been observed to vary by an order of magnitude between wet and dry years (Hull and Pechanec 1947), whereas native shrubs show much lower variability (e.g., Elmore et al. 2003; Renwick et al. 2017). Differences in the Normalized Difference Vegetation Index (NDVI), a metric of photosynthetic greenness, between wet and dry years have also been used previously to identify areas where cheatgrass is present with remotely sensed data (Bradley and Mustard 2005, 2008).

We used the eMODIS phenology product from 2000-2014 (USGS EROS 2016) to create metrics of annual variability, such as the standard deviations of phenology metrics for the time series and differences between wet and dry years (Table 2). We used MODIS NDVI data (Carroll et al. 2004) to measure differences between spring (May-June) and summer (July-August) greenness, focusing specifically on years with above average precipitation across the Great Basin (2005, 2010, 2015) when infested areas are most likely to show strong spring growth. We also included predictor layers such as
average time-integrated NDVI and the MODIS Vegetation Continuous Field measurement of tree cover (DiMiceli et al. 2011) to help to identify other land cover classes distinct from cheatgrass. We used 21 total MODIS-derived layers plus one elevation layer (GMTED 2010) as predictors in a random forest model of cheatgrass percent cover across the Great Basin. All predictor layers and resulting models were standardized to 250 m resolution, which is the native resolution of all but the 1 km resolution elevation dataset.

Spatial models

We used a random forest regression model in R (package randomForest, R version 3.1.2, (Breiman 2001)) to predict cheatgrass percent cover based on the MODIS predictor layers. The random forest algorithm uses an ensemble approach to create a set of decision trees (we used the default n trees = 500) using bootstrap sampling of the percent cover data such that ⅔ of the data are used to train each tree and ⅓ (the ‘out-of-bag’ data) are retained for model testing. Predictor variables are randomly selected in each tree, so variable importance can be estimated based on the variables that contribute the most splits to the overall model and thereby most improve overall accuracy. Because the model is based on decision trees, random forest regression is not strongly affected by auto-correlation amongst predictor variables. Nonetheless, we performed variable selection using random forest (package VSURF, R version 3.1.2) to reduce the number of predictor variables. VSURF first ranks variable importance based on contribution to the random forest model, removes low-importance variables whose addition does not significantly reduce overall model error, and, finally, iterates combinations of the remaining variables to select only those whose inclusion significantly reduces model error. Model validation is performed within the random forest algorithm through cross-validation of the out-of-bag data. The random forest model results reflect the average of individual regression trees, which split the predictor variables based on cheatgrass percent cover.

In addition to predicting cheatgrass continuous cover, we created a regional classification of cheatgrass presence. We created a second random forest model of cheatgrass continuous cover, retaining 30% of data points for testing, and created a post-hoc classification. We used this approach rather than creating a random forest classification directly so that we could test overall classification accuracy based on varying thresholds of observed cover. Lower values of observed cover are increasingly likely to be misclassified as absent because they are not detectable based on the satellite predictors. For visualization, we used the testing dataset to create histograms of predicted cover for different bins of observed cover, which illustrates separability of percent cover classes. We chose a threshold of predicted cover that maximized the overall model accuracy based on Cohen’s kappa (Cohen 1960). As a result, our classification differentiates between cheatgrass present at high abundance (≥15%) and cheatgrass absent or at low abundance (<15%; see testing data and random forest predictions in Appendix 1). Hereafter, we refer to this classification as ‘presence’, however, presence should be interpreted as areas with abundant infestations.

Relationship between cheatgrass and fire

We assessed relations between the area burned, number of fires, and ignition sources with the
regional distribution of cheatgrass. Because the accuracy of our model of continuous cover of cheatgrass was relatively low (see below), our response variables were derived instead from the regional classification of cheatgrass presence or from the field survey measurements of cheatgrass cover.

To test whether cheatgrass presence and percent cover were related to area burned between 2000-2014, we used the MODIS MCD64 Burned Area product (Giglio et al. 2013). We calculated the percentage of cheatgrass presence versus the percentage of cheatgrass absence that burned once or multiple times. We also calculated the percentage of burned area on cheatgrass presence and absence relative to total area burned.

To test whether cheatgrass presence and percent cover were related to the number of fires, we calculated the total number of fires that burned in areas classified as cheatgrass present and cheatgrass absent. Individual fires were identified from the MODIS burned area product by associating pixels that burned within 5 pixels of distance and 9 days with the same fire event. These spatial and temporal distances have been shown to aggregate MODIS burned area pixels into events that closely match fire perimeters derived from the Monitoring Trends in Burn Severity product (MTBS; Eidenshink et al. 2007) across the western United States (Dadashi et al. 2017). For counting numbers of fire events, we associated a given fire perimeter with cheatgrass presence or absence if at least 50% of pixels burned on the first day of the fire (i.e. ignition pixels) were located on that land cover. We focused on interactions between ignitions and land cover for the fire event analysis assuming that numbers of fires would be most influenced by land cover at the ignition point, whereas burned area would be most influenced by land cover throughout the fire perimeter. We compared the percentage of ignitions pixels within areas in which cheatgrass was classified as present or absent. To assess whether the number of fires was related to the percent cover of cheatgrass, we calculated mean cheatgrass cover from field observations within 250 m grid cells and calculated the percentage of pixels that burned within 5% cover bins. We used a Tukey test to compare multiple percent cover bins to evaluate whether burns became significantly more likely as cheatgrass cover increased.

To test whether fires burning where cheatgrass is present have different ignition sources and seasonality than fires burning where cheatgrass is absent, we use the Fire Program Analysis fire occurrence database (Short 2015). The fire occurrence database identifies the point of origin (at least to the nearest Public Land Survey System grid) of fires that were actively managed by municipal, state or federal agencies. This database includes small fires identified as <1 acre and thus reports many more fire events than the MODIS dataset (Short 2015). We used the ‘stated cause’ information from the fire occurrence data from 2001-2013 (the last date of record) to test whether fires that burned on cheatgrass were more likely to be ignited by lightning or human activities. We also explored whether any specific causes of ignition were more likely to cause fires where cheatgrass is present. Because humans start fires throughout the year (Balch et al. 2017), we used a t-test that assumed unequal variance to compare the mean discovery date of human-ignited fires where cheatgrass is present vs. absent.

Finally, of the MODIS fire perimeters (Dadashi et al. 2017) that contained pixels with cheatgrass
presence, we asked whether cheatgrass was more likely to be found at ignition pixels. This pattern could suggest that fires ignite on cheatgrass and spread into surrounding land cover. We identified all pixels burning on the first and last day of each MODIS fire perimeter (Dadashi et al. 2017) for fires that burned over multiple days. If any burned pixels from the first day overlapped cheatgrass presence, we considered the fire to have ignited on cheatgrass. Similarly, if any burned pixels from the last day overlapped cheatgrass presence, we considered the fire to have extinguished on cheatgrass. We used a paired t-test to evaluate whether the likelihood of cheatgrass presence differed between the first and last day of the fire.
Results

Cheatgrass cover data

The data on cheatgrass cover within the Great Basin included 11,307 points (Table 1; Appendix 2). When the data were spatially aggregated to a 250 m resolution, we had data on percent cover of cheatgrass for 8,779 MODIS pixels, which spanned the Great Basin region (Figure 1). Cheatgrass was absent from 3,532 (40%) of these pixels. Among pixels in which cheatgrass was present, the average percent cover was 21 (± 21 St. Dev., Median: 15). A histogram of percent cover values for the MODIS pixel resolution dataset is presented in Figure 2.

Spatial models

Of the initial predictor variables included in the Random Forest model, ten were retained during variable selection and included in the final model (Table 3). The predictors most strongly associated with percent cover of cheatgrass were early season phenology (spring – summer NDVI during 2005, 2010, and 2015, all of which had above-average precipitation) and elevation. Other predictors retained in the model characterized mean and maximum annual NDVI and the difference between the most productive and least productive years.

Thirty four percent of overall variance in percent cover was explained by the random forest regression based on the internal out-of-bag validation. When we created our own secondary testing using a random selection of 2634 points (⅓ of the data), the relationship was similar (R² = 0.32, p<.001). While significant, this relationship was not strong enough to enable a robust projection of percent cover (Figure 3). The most accurate classification of cheatgrass land cover used a threshold of 15% predicted cover to identify cheatgrass present at high abundance (≥15% cover). The resulting model correctly predicts 67% of cheatgrass observed at ≥15% cover and 77% of cheatgrass observed at <15% cover (including absences). The classification had an overall accuracy of 74% (Table 4). The kappa statistic for the model was 0.41, which is considered ‘moderate’ (Landis and Koch 1977). The model estimated that cheatgrass is present at high abundance across nearly one third of the Great Basin (31.4%; 210,000 km²), primarily in northern Nevada, Idaho, Oregon, and Washington (Figure 4).

Relationship between cheatgrass and fire

Over 22,000 km², or 10.7% of land with cheatgrass present (≥15% cover) burned between 2000-2014. In contrast, only 5.2% of land with cheatgrass absent (<15% cover) burned during the same time period. Cheatgrass dominated land cover was more likely to burn multiple times. On land with cheatgrass present, 0.9% of pixels burned more than once compared to 0.2% of pixels with cheatgrass absent. Not surprisingly, cheatgrass was overrepresented in total area burned within the Great Basin and was present in 47.9% of burned area.

Cheatgrass was even more strongly overrepresented relative to the total number of MODIS fire perimeters burned. Based on our temporal and spatial aggregation (Dadashi et al. 2017), 4886 fire events occurred in the Great Basin from 2000-2014. A majority of the ignition pixels within 54.9% of
these fire events occurred where cheatgrass was present. Focusing on the 3394 multi-date fires to ask whether fires starting in cheatgrass were likely to spread into native land cover, we found that a majority of ignition pixels occurred on cheatgrass in 56.4% of fires. However, this percentage was not significantly higher than the proportion of multi-date fires with cheatgrass present on the last day (54.8%; t=1.35; p = 0.09).

Based on the observed cover data, the relationship between cheatgrass presence and increased fire frequency was even more pronounced. Of the 5318 MODIS cells in which cheatgrass was present, 13.6% burned from 2000-2014. In contrast, of the 3461 cells in which cheatgrass was absent, only 4.9% burned from 2000-2014. Areas in which observed cover of cheatgrass was <1% were not significantly more likely to burn than areas in which cheatgrass was absent (z=1.56; p=0.61; Figure 5). Fire frequency increased substantially where observed cheatgrass cover was greater than 1% (z>6; p<.001), but did not differ significantly between higher cheatgrass cover values (Figure 5).

Cheatgrass presence showed a marked interaction with fires ignited by human activity (as opposed to lightning). Cheatgrass was present in 44% of fires in the agency firedata (Short 2015). Humans ignited 74.5% of the 19,492 fires in cheatgrass. In contrast, humans ignited only 27.1% of the 24,584 fires in non-cheatgrass land-cover (Short 2015). Of the stated causes of fires ignited by humans in cheatgrass, fireworks were the most disproportionately skewed towards cheatgrass land cover (88% of fireworks ignitions occurred in cheatgrass), whereas campfires were less likely to start fires in cheatgrass (58% of campfire ignitions occurred in cheatgrass). Human-ignited fires within cheatgrass burned significantly earlier in the season (mean burn date July 19) than human ignited fires in other land-cover types (mean burn date July 29) (t=10.62; p<0.001). Fires in cheatgrass were more likely to occur during summer, especially on July 4 (Figure 6).
Discussion

Our regional model of cheatgrass presence suggested that cheatgrass cover is ≥15% across nearly one-third of the Great Basin (210,000 km²), and is particularly prevalent in the Snake River Plain and eastern Washington (Figure 4). This area is about ten times larger than that estimated by Bradley and Mustard (2005). However, that study did not include the northern Great Basin, and focused on identifying the strong temporal variability caused by cheatgrass rather than creating an accurate map of regional presence. Later studies that focused on map accuracy and covered a more extensive region reported much higher estimates land area with cheatgrass present. Bradley and Mustard (2008) estimated the extent of cheatgrass to be 125,000 km² (still excluding eastern Washington), and Downs et al. (2016) estimated that 240,000 km² across the historic range of Greater Sage-Grouse (the Great Basin and northern Great Plains) was dominated by cheatgrass. Collectively, these estimates suggest that cheatgrass dominates a much larger area than previously reported.

The overall accuracy of our model of cheatgrass presence was 74%, which is comparable to previous models. For example, Bradley and Mustard (2008) reported an overall accuracy of 61%, and Downs et al. (2016) reported an overall accuracy of 71%. Clinton et al. (2010) reported an overall accuracy of 90%, but their results were based on 92 testing points, and those testing points were concentrated in areas with abundant cheatgrass. Hence, the estimate of Clinton et al. (2010) is likely inflated relative to estimates based on more extensive testing data that include a range of cheatgrass cover values as presences.

Our percent cover model explained 34% of the variation of cheatgrass percent cover. This accuracy is comparable to previous models of percent cover. For example, Peterson (2005, 2006) reported R² values between modeled and observed cheatgrass cover of 0.5 (based on 75 testing points) and 0.24 (based on 806 testing points). West et al. (2017) reported an R² value of 0.92 for the relationship between predicted probability of cheatgrass and observed cover in Wyoming. However, West et al.’s testing data included almost entirely points with <20% and >80% cover, so this regression estimate is likely inflated. Boyte et al. (2016) also reported a much higher correspondence between modeled and observed percent cover (R² = 0.83). However, their assessment was based on testing data derived from the Peterson (2006) map, which itself had low accuracy (R² = 0.24). Therefore, the accuracy of the model of Boyte et al. (2016) is likely lower than their correlation coefficient suggests. Collectively, these results suggest that modeling the percent cover of cheatgrass remains challenging across large spatial extents.

Multiple factors likely contribute to the poor performance of the regional cheatgrass percent cover model. When cheatgrass cover is low, reflectance signals from soils and native vegetation dominate the phenological patterns observed by MODIS. Thus, there is a broad and heterogeneous range of responses possible from non-cheatgrass, which likely reduces the ability of the model to correctly classify low cheatgrass percent cover (Figure 3). Additionally, inter-annual variation in the biomass and percent cover of cheatgrass is high (Bradley and Mustard 2005; Hull and Pechanec 1947). Because the field data included in our analyses were collected over multiple years (most of which were dry years), it is likely that the training and testing data do not capture maximum cheatgrass cover from 2000-2015. Less than 7% of the training data were collected in wetter years during the study period (2005, 2010, and 2016).
In contrast, the MODIS predictor variables (Table 2) were selected to detect the maximum anomalies in early season growth and inter-annual variability associated with cheatgrass cover during the wettest years. During dry years, cheatgrass has lower percent cover and therefore its unique phenological characteristics are unlikely to be readily detectable with MODIS. This temporal mismatch of observations in dry years vs. prediction focused on wet years likely accounts for some of the model overestimates of percent cover (Figure 3). Moreover, much of the training data were collected during dry years in the mid-2000s. Ongoing spread of cheatgrass (including in conjunction with fire) adds error to the training dataset. In order to improve regional models of percent cover, more current training data with a focus on wetter years when cheatgrass is most productive are needed.

Additionally, many of the compiled estimates of percent cover were based on linear transects 30-150 m in length, which capture a small fraction of a 250 x 250 m MODIS pixel. This spatial mismatch between field survey data and satellite pixel size is a common problem for validation of remote sensing products, including maps of invasive plants (Bradley 2014), and likely contributes to the low model performance. Field survey data at larger spatial scales are rare, so studies such as this one using coarse resolution sensors must assume that smaller-scale field data are representative of the larger landscape (e.g., the locally measured percent cover estimates reflect the regional environmental conditions). Cases where local measurements are not representative of the larger landscape add uncertainty to model projections. Another approach would be to use a higher spatial resolution sensor such as Landsat, which has a pixel size of 30 x 30 m. However, the use of Landsat currently requires a considerable amount of data storage and processing (e.g., there are more than 50 Landsat scenes encompassing the Great Basin) and there are not yet Landsat-based products of seasonal or inter-annual phenology (although the technology to do those analyses exists; e.g., Pasquarella et al. 2016). As a result, using higher resolution satellite imagery to model cheatgrass percent cover could soon be feasible across broad spatial areas.

Relationship between cheatgrass and fire

We found that cheatgrass was twice as likely as other land-cover types to burn from 2000-2014. These estimates are similar to those of Balch et al. (2013), who estimated that the percentages of cheatgrass and other land-cover types that burned were 12.8% and 3.4%, respectively. This slight discrepancy could be due to a difference in study area extents. Here, the study area extends further north into eastern Washington than did that of Balch et al. (2013). Cheatgrass is abundant in eastern Washington, but relatively few large fires have occurred in the region (possibly because extensive agriculture in the region reduces fire spread and motivates fire-control efforts). This study area also excluded much of the Mojave, which has a lower fire frequency than the Great Basin (Balch et al. 2013).

Although fires in the Great Basin are primarily ignited by lightning (Balch et al. 2017), the vast majority (74.5%) of fires in cheatgrass were ignited by humans. The strong association between human ignitions and fires in cheatgrass could be due to cheatgrass’ association with human-dominated landscapes. For example, cheatgrass occurrence increases near roads (Bradley and Mustard 2006; Gelbard and Belnap 2003), railroads (Mack 1981), and powerlines (Bradley and Mustard 2006). These features likely serve both as vectors for the transport of invasive plant propagules and sources of
disturbance for cheatgrass establishment. In the Great Basin, roads, railroads, and powerlines also were the covariates most strongly associated with human ignitions of fire (Fusco et al. 2016). Thus, the close association between humans and cheatgrass, coupled with the close association between humans and fire, may further increase fire likelihood in the Great Basin.

Fireworks resulted in the strongest interaction between cheatgrass and a specific cause of human ignition (Figure 6). Although fireworks are the stated cause of 5% of fires in the Great Basin, 88% of those fires occurred in cheatgrass. There is a clear spike in the number of fires in cheatgrass associated with fireworks on Independence Day (July 4th), whereas the number of fires in other land-cover types changes little on Independence Day (Figure 6). Balch et al. (2017) noted an increase in the number of fires ignited by fireworks around July 4th, but previous work has not detected an association between fireworks ignitions and a particular land cover type. The higher likelihood fireworks-ignited fires in cheatgrass may reflect the proximity of cheatgrass to human settlement, its high fine-fuel biomass (Brooks and Pyke 2001; Pilliod et al. 2017), and its early senescence. Early senescence of cheatgrass relative to native vegetation may also explain the 10-day difference in average fire date in cheatgrass and other land-cover types.

Cheatgrass is thought to increase fire frequency by increasing fine fuel biomass and horizontal continuity of fine fuels (Brooks and Pyke 2001; Pilliod et al. 2017). Our comparison between observed percent cover and fire frequency (Figure 5) suggests that even low cover of cheatgrass is associated with an increased likelihood of fire. The proportion of burned grid cells in which cheatgrass was absent or at <1% cover was similar (5.2% burned). However, 13% of cells with 1-5% observed cheatgrass cover burned, and this percentage remained statistically similar as cheatgrass cover increased. It is possible that inter-annual variability in measured cheatgrass cover adds uncertainty to this analysis, such that areas observed to have low cover in dry years later had higher cover and promoted fire following wet years. Nonetheless, these results suggest that even low cheatgrass cover increases fire risk in the Great Basin.

Conclusions

These results show that cheatgrass invasion is widespread across one-third of the Great Basin desert, a figure much more extensive than previously reported. Cheatgrass invasion is associated with a doubling of regional fire frequency and fire risk is elevated even at low levels of cheatgrass abundance. Our results also highlight a strong association between human ignitions and fire on cheatgrass. Collectively, these results suggest that reducing the grass-fire cycle in the Great Basin should focus on management and control of invaded ecosystems regardless of cheatgrass cover and in areas proximal to human habitation and land use. Additional education and outreach aimed at reducing human ignitions due to fireworks and other anthropogenic causes that interact with cheatgrass could also reduce fire risk.

Appendices

Appendix 1. Histograms of predicted percent cover for different bins of observed percent cover and contingency table calculator for different mapped percent cover thresholds.
Appendix 2. Cheatgrass percent cover data, locations and sources compiled for this analysis.

Acknowledgments

We thank E. Fleishman, J. Finn, N. Horning, A. Mahood, N. Mietkiewicz, and R.C. Nagy for valuable discussion. E. Fleishman, S. Hanser, M. Holton, M. Jesus, M. Lavin, A. Mahood, B. Rau, and L. Turner greatly assisted this project by providing percent cover data. This research was supported by the National Aeronautics and Space Administration Terrestrial Ecology Program under Award NNX14AJ14G and the Joint Fire Sciences Program 15-2-03-6.
Table 1. Source data for field collections of *B. tectorum* cover within the Great Basin used as training data. Note that these data were later aggregated to average percent cover at a 250 m scale consistent with MODIS predictor layers.

<table>
<thead>
<tr>
<th>Source</th>
<th>N pts</th>
<th>Notes/Contact</th>
<th>Spatial extents</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLM AIM</td>
<td>3489</td>
<td>Bureau of Land Management (BLM AIM, 2016)</td>
<td>3 x 50 m* transect</td>
</tr>
<tr>
<td>NDOW</td>
<td>2309</td>
<td>Nevada Department of Wildlife surveys 2011-2015</td>
<td>3 x 50 m* transects</td>
</tr>
<tr>
<td>Fleishman</td>
<td>1092</td>
<td>Field Surveys – unburned plots (Fleishman, 2015a, 2015b)</td>
<td>3 x 50 m* transects</td>
</tr>
<tr>
<td>USGS 2010-2011</td>
<td>1058</td>
<td>USGS/BLM sagebrush regeneration plots (Knutson et al., 2014)</td>
<td>Unknown</td>
</tr>
<tr>
<td>NW_ReGAP</td>
<td>831</td>
<td>Validation plots (Hanser, 2011)</td>
<td>Unknown</td>
</tr>
<tr>
<td>USGS shrubmap</td>
<td>588</td>
<td>Oregon sagebrush regeneration plots 2003 (Hanser, 2011)</td>
<td>Unknown</td>
</tr>
<tr>
<td>Brummer</td>
<td>373</td>
<td>Field surveys (Brummer et al., 2016)</td>
<td>20 20x50 cm frames in 1 ha</td>
</tr>
<tr>
<td>CalFlora</td>
<td>341</td>
<td><a href="http://www.calflora.org/">http://www.calflora.org/</a></td>
<td>Unknown</td>
</tr>
<tr>
<td>Rau</td>
<td>339</td>
<td>Sage Step control plot field surveys (McIver et al., 2014)</td>
<td>3 x 30 m transects</td>
</tr>
<tr>
<td>Peterson</td>
<td>337</td>
<td>Field surveys (Peterson, 2005)</td>
<td>0.1 ha</td>
</tr>
<tr>
<td>Fleishman_Birds</td>
<td>230</td>
<td>Vegetation measurements in bird habitat, Erica Fleishman unpublished</td>
<td>3 x 50 m* transects</td>
</tr>
<tr>
<td>Fleishman_JFSP</td>
<td>101</td>
<td>Field Surveys, Erica Fleishman unpublished</td>
<td>3 x 50 m* transects</td>
</tr>
<tr>
<td>Pseudo-absence</td>
<td>97</td>
<td>Pseudo-absence points on salt pans, water and alpine non-vegetated areas—digitized from aerial photos</td>
<td>N/A</td>
</tr>
<tr>
<td>EDDMaps</td>
<td>50</td>
<td>Verified contributed records with percent cover (Bargeron and Moorhead, 2007)</td>
<td>Unknown</td>
</tr>
<tr>
<td>Bradley</td>
<td>52</td>
<td>Field surveys (Bradley and Mustard, 2005)</td>
<td>6 x 30 m* transects</td>
</tr>
<tr>
<td>Mahood</td>
<td>20</td>
<td>Field surveys, Adam Mahood unpublished</td>
<td>6 x 50 m transects</td>
</tr>
<tr>
<td><strong>TOTAL POINTS</strong></td>
<td><strong>11307</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Measurements radiate out from a center point.
Table 2. List of predictor layers tested in the Random Forest model of cheatgrass percent cover

<table>
<thead>
<tr>
<th>Name</th>
<th>MODIS predictor variable descriptions</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean AMP</td>
<td>Mean NDVI Amplitude (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Mean MAXN</td>
<td>Mean Annual Max NDVI (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Mean TIN</td>
<td>Mean time-integrated NDVI (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Min AMP</td>
<td>Min NDVI Amplitude (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Min MAXN</td>
<td>Min Annual Max NDVI (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Min TIN</td>
<td>Min time-integrated NDVI (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Pct90 AMP</td>
<td>90th percentile of NDVI Amplitude (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Pct90 MAXN</td>
<td>90th percentile of Annual Max NDVI (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Pct90 TIN</td>
<td>90th percentile of the time-integrated NDVI (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Δ AMP</td>
<td>90th percentile minus min NDVI Amplitude (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Δ MAXN</td>
<td>90th percentile minus min Annual Max NDVI (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Δ TIN</td>
<td>90th percentile minus min time-integrated NDVI (2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>wetdryMAXN</td>
<td>Max MAXN 2005 &amp; 2010 (wet) minus mean MAXN of all other years (dry; 2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>wetdryTIN</td>
<td>Max TIN 2005 &amp; 2010 (wet) minus mean TIN of all other years (dry; 2001-2014)</td>
<td>(USGS EROS, 2016)</td>
</tr>
<tr>
<td>Spr-Sum2016</td>
<td>Mean of cloud/snow-free NDVI from May/June (2016) – Mean of cloud-free NDVI from July/Aug (2016).</td>
<td>(Carroll et al., 2004)</td>
</tr>
<tr>
<td>Spr-Sum2010</td>
<td>Mean of cloud/snow-free NDVI from May/June (2010) – Mean of cloud-free NDVI from July/Aug (2010).</td>
<td>(Carroll et al., 2004)</td>
</tr>
<tr>
<td>Spr-Sum2005</td>
<td>Mean of cloud/snow-free NDVI from May/June (2005) – Mean of cloud-free NDVI from July/Aug (2005).</td>
<td>(Carroll et al., 2004)</td>
</tr>
<tr>
<td>MODIS_VCF</td>
<td>Mean MODIS vegetation continuous field (2001-2005)</td>
<td>(DiMiceli et al., 2011)</td>
</tr>
<tr>
<td>GMTED2010</td>
<td>USGS estimated elevation at ~1 km resolution (2010)</td>
<td>(GMTED, 2010)</td>
</tr>
</tbody>
</table>

Note: We excluded maximum AMP, TIN, MAXN because these layers are susceptible to anomalously high pixel values. We used the 90th percentile temporal values rather than maximum values to reduce these errors.
Table 3. Predictor layers included in the final random forest model. Numbers indicate the order of importance of the layer in the random forest regression model (number 1 has the highest contribution).

<table>
<thead>
<tr>
<th>Predictor Layer</th>
<th>Variable Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMTED2010</td>
<td>1</td>
</tr>
<tr>
<td>Spr-Sum2010</td>
<td>2</td>
</tr>
<tr>
<td>Spr-Sum2005</td>
<td>3</td>
</tr>
<tr>
<td>Δ MAXN</td>
<td>4</td>
</tr>
<tr>
<td>Spr-Sum2016</td>
<td>5</td>
</tr>
<tr>
<td>Mean TIN</td>
<td>6</td>
</tr>
<tr>
<td>Pct90 MAXN</td>
<td>7</td>
</tr>
<tr>
<td>Mean MAXN</td>
<td>8</td>
</tr>
<tr>
<td>Mean AMP</td>
<td>9</td>
</tr>
<tr>
<td>Pct90 TIN</td>
<td>10</td>
</tr>
</tbody>
</table>
**Table 4.** Contingency table for classification of cheatgrass presence/absence using observed points with ≥15% cover as presence and <15% cover as absence. Overall map accuracy is 0.74.

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present (≥15%)</td>
<td>Absent (&lt;15%)</td>
<td>Total</td>
</tr>
<tr>
<td>Present</td>
<td>506</td>
<td>437</td>
<td>943</td>
</tr>
<tr>
<td>Absent</td>
<td>245</td>
<td>1446</td>
<td>1691</td>
</tr>
<tr>
<td>Total</td>
<td>751</td>
<td>1883</td>
<td>2634</td>
</tr>
<tr>
<td>Reliability</td>
<td>0.67</td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Study area map including cheatgrass (*B. tectorum*) percent cover data used to train the classification.
Figure 2. Histogram of number of grid cells associated with bins of average cheatgrass percent cover. Total grid cells = 8779.
Figure 3. Scatter plot of predicted vs. observed percent cover for the testing dataset. Darker points indicate overlap of data. A threshold of 15% predicted cover was used to classify cheatgrass present at high abundance.
Figure 4. Distribution map of abundant cheatgrass presence (≥15% cover; gray areas) within the Great Basin ecoregion based on MODIS predictors.
Figure 5. Grid cells with even small amounts of cheatgrass cover are more likely to have burned between 2000-2014 based on the MODIS MCD64 burned area product. Data are mean +/- stdev and letters indicate a significant difference (p<0.05) based on a Tukey test of multiple percent cover bins.
Figure 6. Human ignited fires in cheatgrass grassland (red bars) are most likely to burn during summer months and on average occur ten days earlier than human ignited fires where cheatgrass is absent (gray bars).
Citations


DiMiceli, C., Carroll, M., Sohlberg, R., Huang, C., Hansen, M., Townshend, J., 2011. Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000 - 2010, Collection 5 Percent Tree Cover, University of Maryland, College Park, MD, USA.


Varies with Climate Conditions. Ecosystems 17, 960–973. doi:10.1007/s10021-014-9771-7


