

Deep Learning Classification of Cheatgrass Invasion using Biophysical and Remote Sensing Data

Kyle Larson^{*}¹ and Aaron Tuor[†]¹

¹Pacific Northwest National Laboratory

September 28, 2017

Contents

1	Introduction	2
2	Related Work	3
2.1	Cheatgrass Mapping	3
2.2	Machine Learning Methods for Geo-spatial Prediction	3
2.3	PNNL study	3
2.4	Other works	3
3	Data Set	4
3.1	Field Measurements	4
3.2	Bio-physical Data	4
3.2.1	Potential relative radiation	4
3.2.2	DAYMET	5
3.2.3	Growing degree days	5
3.2.4	PRISM	5
3.3	Ecological Data	5
3.3.1	Ecoregions	6
3.3.2	LANDFIRE Landcover	6
3.4	Satellite Imagery	6
3.4.1	Normalized Difference Vegetation Index	7
3.4.2	Enhanced Vegetation Index (EVI)	7
3.4.3	MODIS (annual, seasonal)	7
3.4.4	Landsat (annual, seasonal)	7
3.5	Data Synopsis	7
4	Methods	8
4.1	Baseline Models	8
4.2	Deep Learning Models	8
4.2.1	Deep Neural Network	10
4.2.2	Recurrent Neural Network	11
4.3	Experimental Setup	11

^{*}kyle.larson@pnnl.gov

[†]aaron.tuor@pnnl.gov

5 Results and Analysis	12
5.1 Performance on Field measurement locations	12
5.1.1 Quantitative assessment	12
5.1.2 Training Statistics	12
5.1.3 Qualitative assesment	13
5.2 Performance on Full Study Region	16
5.2.1 Pixel-wise count of cheatgrass class across model predictions	17
5.2.2 Block statistics (count) of cheatgrass class (analogous to block area), then look at range of block counts	17
6 Conclusion and Future Work	17

Abstract

In this study we explore deep learning approaches which incorporate geophysical data and readings from multiple remote sensing platforms for predictive mapping of vegetation over very large, ecologically diverse regions. We focus specifically on mapping an invasive exotic annual grass, cheatgrass (*Bromus tectorum*), that has become a major land management issue in the Western United States. We explore the relative advantages of deep neural networks to integrate imagery from the LandSat7 and MODIS platforms which have complementary advantages in their spectral bandwidth, spatial resolution, and temporal frequency. We show that integrating imagery from multiple platforms is beneficial to the predictive mapping of cheatgrass in the Western United State's historic range of Sage Grouse. Our analysis of prospective machine learning models demonstrates that a deep neural network can more effectively take advantage of the signal from multiple platforms than traditional classification models such as Random Forest, and Linear Discriminant Analysis for the classification of cheatgrass.

1 Introduction

Accurate and up to date maps of vegetation coverage are critical assets for land management agencies tasked with mitigating damage wrought by invasive species. A variety of methods have been developed to map invasive annual grasses from satellite and aerial imagery, each uniquely tailored to the limitations of the imagery. Current practices usually involve choosing one platform best suited for the study, although more than one may potentially be useful, e.g., using finer spatial resolution imagery may improve results, but doing so comes with concomitant loss of temporal resolution which is also important for detecting species like invasive cheatgrass which has maintained a growth cycle distinct from native vegetation.

Introduced in the late 19th-century, cheatgrass is now found in every state in the contiguous U.S.. Nowhere has its invasion been more prevalent than in western states where it has become a dominant component in many shrubland and grassland ecosystems. It is now estimated to dominate at least 40,000 km² in the states of Nevada and Utah alone. Cheatgrass invasion poses a variety of threats to ecosystem function, rangeland health, and human safety. A central thread to many of these threats is the increase in fine fuels associated with cheatgrass, which can lead to increased fire frequency and irreversible loss of native vegetation and wildlife habitat. Following a fire, cheatgrass is able to more effectively compete with native vegetation, giving rise to a positive feedback cycle of further invasion and fire. [Refs: Mack 1981; Pellatt 1996; Bradley & Mustard 2005; USDA Plant Database]

While cheatgrass is considered ubiquitous throughout much of the western U.S., detailed spatial information about its presence and abundance are still lacking for much of its range within this region. Previous efforts to map cheatgrass have focused on core areas of invasion such as the Great Basin and Snake River Plain. These efforts help paint a clearer picture of cheatgrass invasion in the western U.S., but their disparate nature (due to differing extents, methods, data, and focal periods) prevents easily reconciling them to inform range-wide management decisions. Conversely, mapping cheatgrass (or other vegetation) at a range-wide scale is also not straightforward or without its challenges; most notably the need for a greater volume of training and input predictor data to adequately model the phenomenon across many ecological gradients. [Refs:]

The study area encompasses a large part of the Intermountain West as well as portions of Montana, Wyoming, Colorado, and New Mexico east of the Rockies. The areal extent of the area identified as the historic sage-grouse range covers more than 308 million acres and includes portions of southern Alberta and British Columbia, Canada. This study focused on the U.S. portion of the sage-grouse range, which is approximately 288 million acres. Figure 1 shows the region considered and location of field samples.

2 Related Work

2.1 Cheatgrass Mapping

A variety of methods to map cheatgrass are described in the literature, ranging from imagery-driven methods that focus on spectral signatures or phenological indicators of cheatgrass to ecological niche modeling approaches aimed at mapping biophysical and climatic predictors. Much attention has been given to deriving phenological indicators of cheatgrass presence from overhead imagery (typically using spectral indices such as the Normalized Difference Vegetation Index, or NDVI) because its life cycle differs from many of the native plant species in its western range. Cheatgrass is a winter annual that may begin growth in the late fall and senesce in late spring, whereas many native plants begin growing in mid to late spring and continue growth through summer under favorable precipitation conditions. Thus, cheatgrass presence can be identified indirectly by comparing pixel-chronologies of NDVI, particularly in years when winter or early spring precipitation is [Refs: Rice et al. 1992; Loik 2007; Bradley 2009;] Phenological - attempt to exploit seasonal differences in growth of target and non-target species Can be highly variable across years, plant communities, and regions

Often target specific time periods to optimize detection of phenological signatures

- Spectral
- Biophysical
- Studies that attempt to combine these approaches?

2.2 Machine Learning Methods for Geo-spatial Prediction

Machine learning methods have become increasingly popular for remote sensing classification due to their ability to model complex class signatures, accept a variety of input data, and often outperform traditionally used parametric methods. Some of more commonly used machine learning methods in remote sensing include support vector machines (SVMs), Random Forests (RFs), single decision trees (DTs), boosted DTs, k-nearest neighbor (k-NN), and artificial neural networks (ANNs). While there has been considerable use and comparison of these as well as other machine learning methods not mentioned, there appears to be little consensus on whether one is more superior than the others. A number of excellent reviews of these machine learning methods in remote sensing exist, which we defer to for readers seeking more in-depth information and guidance on those specific methods. Here, we focus on illustrating the use and performance of a newer and rapidly growing class of machine learning methods called deep learning for classification in remote sensing, with an emphasis on providing an applied perspective from multiple domains. [Refs: Maxwell et al. 2018; Maxwell et al. 2015; Crisci et al. 2012; Lu & Weng 2007; Mountrakis et al. 2011; Rodriguez-Galiano et al. 2012; Rogan et al. 2008; Shao et al. 2012]

2.3 PNNL study

Impetus for previous study

In 2016, PNNL developed a statistical model to map the occurrence of cheatgrass in the Intermountain West that uses satellite-based Normalized Difference Vegetation Index (NDVI) a widely used spectral index for mapping cheatgrass, vegetation, and other phenomena and biophysical parameters. NDVI was derived from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) which offers low- spatial, high-temporal resolution imagery that is well-suited for assessing land cover over large geographic areas. The model achieved an acceptable level of accuracy for the study area (71%), but accuracy was variable when examined at finer scales.

2.4 Other works

- Land cover classification based on NDVI using LANDSAT8 time series: A case study Tirupati region [20]
- Hyperspectral remote sensing of vegetation [32]
- Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review [1]
- A review of hyperspectral remote sensing and its application in vegetation and water resource studies [17]
- Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem [19]

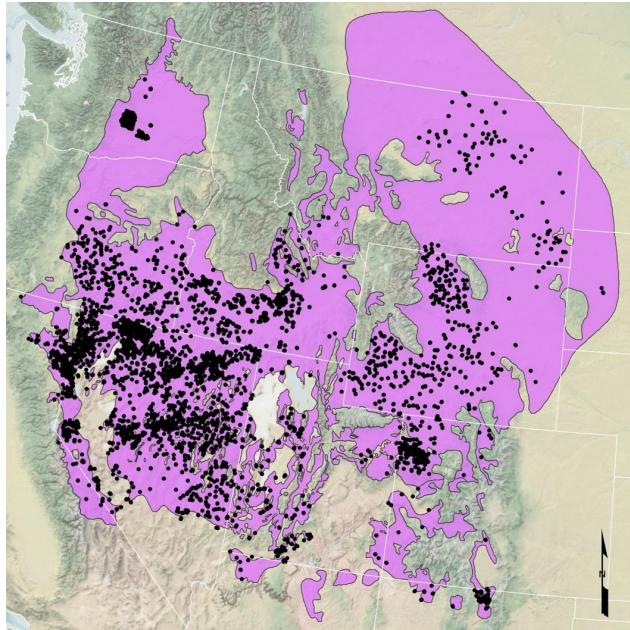


Figure 1: Study Region

- Mapping salt-marsh vegetation by multispectral and hyperspectral remote sensing [3]

3 Data Set

Within the study area boundaries, we solicited more than 24,000 field measurements from various sources to gather information on cheatgrass occurrence. Accompanying these field measurements are bio-physical measurements (weather and soil data), and ecological data, as well as satellite imagery gathered from the MODIS and LandSat platforms.

3.1 Field Measurements

Figure 2 provides the field measurement datasets and sources considered, although not all data sources and data were accepted for use. Data were reviewed for completeness, geographic accuracy, and quality of data source. Review of the field measurement dataset allowed us to identify 6,650 measurement points that could potentially be used for modeling. All field data were collected along transects ranging from 25m to 100m in length. Point intercept data and plot frame (0.25 m to 1 m) data taken along the transects were summarized to calculate percent canopy cover of *Bromus tectorum* and *B. rubens*. The introduced, annual grass, red brome (*B. rubens*) was included because it poses a very similar threat in terms of modifying fire regimes, and its life history characteristics are similar to cheatgrass.

3.2 Bio-physical Data

Bio-physical datasets used for model development included the elevation, potential relative radiation index (PRR) (Pierce et al. 2005), and a growing degree day index. Spatially interpolated climate data (precipitation, temperature) for the study area was acquired from the PRISM (Parameter-elevation Regressions on Independent Slopes Model) Climate Group (Daly et al. 1994; Daly et al. 2008; DiLuzio et al. 2008) and the Daily Surface Weather and Climatological Summaries (DAYMET) program (Thornton et al. 1997; Thornton et al. 2014).

3.2.1 Potential relative radiation

The PRR is a unitless index of available solar radiation for photosynthetic activity at a given location that takes into account the influence of geographic position, seasonal and daily variation in solar inclination, and topography. PRR was calculated by summing digital hillshade interpolations for a given period of interest as described by Pierce et al. (2005). Hourly hillshade interpolations were performed for daylight hours of one day of the month that most closely represents the average solar period for the month (i.e., $PRR = \text{Sum} [\text{Hillshade}_{i-j}, m-n], \text{hours } i-j \text{ for each representative}$

Table 2.1. Data Source and Number of Field Measurements Evaluated.

Description	Number of Field Measurements Reviewed	Number of Field Measurements Accepted
BLM Assessment, Inventory, and Monitoring Program	2043	1927
BLM Landscape Monitoring Framework	2335	2222
Global Invasive Species Information Network	4292	0
Joint Base Lewis-McCord Yakima Training Center	382	375
PNNL – Birds of Prey field campaign	92	81
PNNL – Hanford Vegetation	39	39
PNNL – Owyhee field campaign	30	29
PNNL – Shoshone field campaign	95	89
USGS SAGEMAP GIS database	820	818
Sagebrush Steppe Treatment Evaluation Project	1086	1070
U.S. Forest Service Forest Inventory Analysis	13170	0
Total	24384	6650

Figure 2: Field Data

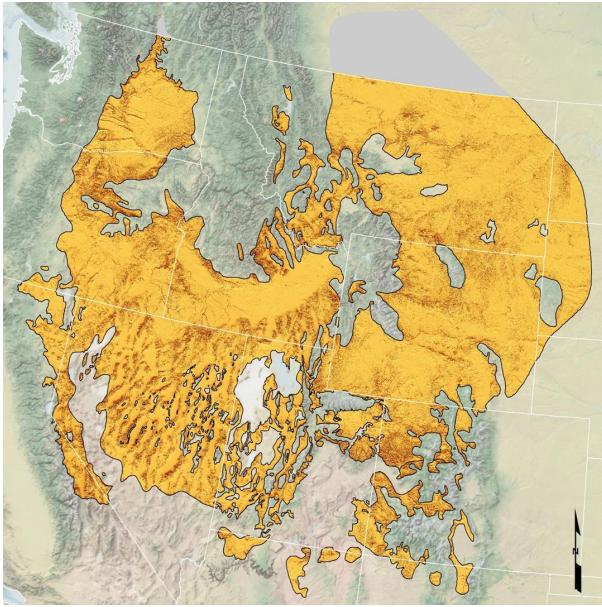


Figure 3: Potential Relative Radiation Index (light orange to dark orange illustrates high to low PRR).

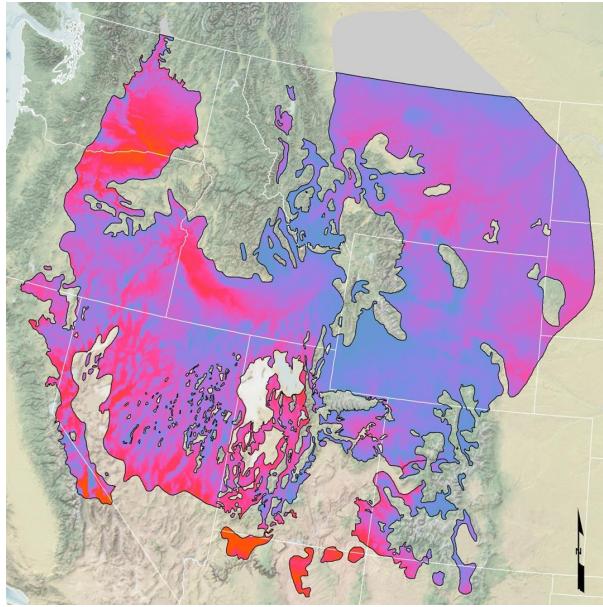


Figure 4: Cumulative Growing Degree Day (blue to red illustrates low to high GDD).

day of months m-n). PRR calculated for use in this study reflects the solar conditions between October and June, which encompasses the bulk of the growing season of cheatgrass across the study area.

3.2.2 DAYMET

DAYMET provides gridded estimates of daily weather parameters for North America at a 1-km resolution, including daily continuous surfaces of minimum and maximum temperature and precipitation. DAYMET daily minimum and maximum temperature data were acquired to calculate growing degree day index for the study area.

3.2.3 Growing degree days

Using the DAYMET daily minimum and maximum temperature data from 1 October 2014 to 30 April 2015, the cumulative growing degree day (GDD) index was calculated to represent the relative period of time when temperatures are suitable for plant growth (Figure 2.3).

The cumulative GDD between October 1 and April 30 was calculated on a daily basis: $[(T_{max} - T_{min})/2] \cdot w$ where w is the minimum temperature for growth of cheatgrass (assumed to be 0C or 32F), and summed for the period. Negative values were set to 0.

3.2.4 PRISM

PRISM uses point measurements of climate data and a digital elevation model of terrain to estimate continuous gridded surfaces of monthly climate elements at a 4-km resolution. PRISM 30-year mean monthly and 30-year mean annual precipitation, minimum temperature, and maximum temperature (totaling 39 separate climate variables) were obtained to explore relationships between cheatgrass occurrence and general climate patterns. We also derived several seasonal cumulative precipitation and average minimum and maximum temperature variables from PRISM data that correspond to important seasonal periods during the life history of cheatgrass.

3.3 Ecological Data

Our data set includes high level ecological data in the form ecoregions, and land cover classification.

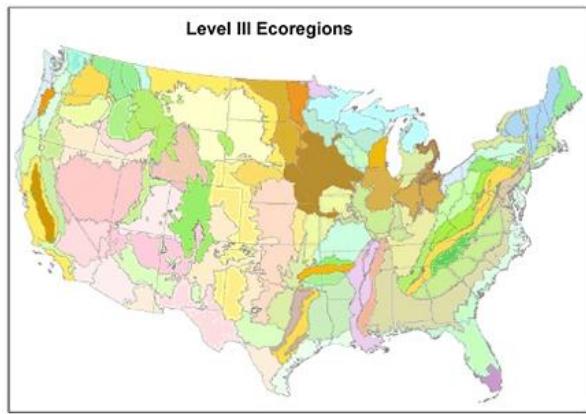


Figure 5: Figure from <https://www.epa.gov/eco-research/ecoregions>

Region Number	%	Region Name
4	0	Cascades
5	0	Sierra Nevada
9	1.9	Eastern Cascades slopes and foothills
10	8.7	Columbia plateau
11	1.7	Blue Mountains
12	9.3	Snake River plain
13	36	Central basin and range
15	0	Northern Rockies
14	0	Mojave basin and range
16	0.2	Idaho batholith
17	1.2	Middle Rockies
18	7.6	Wyoming basin
19	0.3	Wasatch and Uinta Mountains
20	3.7	Colorado plateaus
21	0.02	Southern Rockies
22	1.2	Arizona, New Mexico plateaus
23	0	Arizona, New Mexico Mountains
25	0.02	Wester high plains
41	0	Canadian Rockies
42	0.05	Northwestern glaciated plains
43	2	Northwestern great plains
77	0	North Cascades
80	25.3	Northern basin and range

Table 1: Study area ecoregions

3.3.1 Ecoregions

Ecoregions are areas which contain similar collections of plant and animal life, and similar ecological conditions such as geology, landforms, soils, vegetation, climate, land use, and water resources [27, 30, 8]. The study area spans 23 of the 182 level III ecoregions as defined by the U.S. Environmental Protection Agency (EPA) and the Commission for Environmental Cooperation (CEC). The ecoregion classifications for the study area were retrieved from the EPA hosted GIS data sets¹. Figure 5 shows the distributions of 16 ecoregions contained among the field measurement locations, while Table 1 shows the distribution of ecoregions throughout the study area.

3.3.2 LANDFIRE Landcover

Categorical value for vegetation cover type generalized from LANDFIRE (www.landfire.gov) vegetation data. 15 classes.

3.4 Satellite Imagery

For each location i in the full study region, we have retrieved yearly and seasonal image data from the MODIS and LandSat platform and derived vegetation indexes.

¹<https://www.epa.gov/eco-research/ecoregions>

Column	Band/Index
$M_{:,1}$	NDVI
$M_{:,2}$	EVI
$M_{:,3}$	Blue (B)
$M_{:,4}$	Red (R)
$M_{:,5}$	Near Infra-Red (NIR)
$M_{:,6}$	Mid Infra-Red (MIR)
$M_{:,7}$	dNDVI
$L_{:,1}$	dNDVI
$L_{:,2}$	Blue (B)
$L_{:,3}$	Green (G)
$L_{:,4}$	Red (R)
$L_{:,5}$	Near Infra-Red (NIR)
$L_{:,6}$	Short Wave Infra-Red 1 (SWIR1)
$L_{:,7}$	Thermal, low gain channel (T1)
$L_{:,8}$	Thermal, high gain channel (T2)
$L_{:,9}$	Short Wave Infra-Red 2 (SWIR2)

Table 2: Columns for MODIS spectral bands and vegetation indexes.

Table 3: Columns for LandSat spectral bands and vegetation indexes.

3.4.1 Normalized Difference Vegetation Index

NDVI values provided by the LandSat [34] and MODIS platforms are calculated as a ratio between the red (R) and near infrared (NIR) values as:

$$NDVI = (NIR - R)/(NIR + R) \quad (1)$$

We define dNDVI as the peak NDVI (over a span of a year or a season) minus long-term median annual peak NDVI (across the years we have images for).

3.4.2 Enhanced Vegetation Index (EVI)

EVI incorporates an L value to adjust for canopy background, C values as coefficients for atmospheric resistance, and values from the blue band (B) in order to reduce the background noise, atmospheric noise, and saturation [34].

$$EVI = G * ((NIR - R)/(NIR + C_1 * R - C_2 * B + L)) \quad (2)$$

3.4.3 MODIS (annual, seasonal)

The MODIS annual and seasonal composite imagery data consists of 250m resolution pixel values associated with each study region location i . We have composite annual, spring, and summer pixel values for each location in the study area for each year from 2001-2016. MODIS composite pixels are formed as the band or index value at the time of the peak NDVI value during the respective time range (spring, summer, or entire year). Introducing some notation we define three matrices $M_{\text{annual}}^{(i)}$, $M_{\text{spring}}^{(i)}$, and $M_{\text{summer}}^{(i)}$ for each study location i . Table 2 shows the spectral bands and derived indices that form a each matrix $M \in \mathbb{R}^{16 \times 7}$ of values with rows corresponding to the year, and columns corresponding to the spectral band or derived index.

3.4.4 Landsat (annual, seasonal)

The LandSat imagery differs from the MODIS imagery in resolution, spectral bands, years of operation, and composite product preparation. It consists of 30m resolution pixel values from the years 2003-2012. As in the case of MODIS, for convenience we define three matrices $L_{\text{annual}}^{(i)}$, $L_{\text{spring}}^{(i)}$, and $L_{\text{summer}}^{(i)}$ for each study location i . Table 3 shows the spectral bands and derived indices that form a each matrix $L \in \mathbb{R}^{10 \times 9}$ of values with rows corresponding to the year, and columns corresponding to the spectral band or derived index.

3.5 Data Synopsis

Tables 4, 5, and 6 contain brief descriptions of all variables used in modeling cheatgrass coverage. The complete set of data used in training predictive mapping models consists of 6602 locations with available field measurements due to missing values from 49 data points in the original data set. We define, $w \in \mathbb{N}^{6602 \times 1}$ as the matrix of unique integer ids for each data point, $y \in \mathbb{R}^{6602 \times 1}$, where $y_{i,1} = y_1^{(i)}$ the matrix of coverage values for each data-point, $C \in \mathbb{R}^{6602 \times 3}$,

Variable	Description
y_p	Proportion (0-1) of cheat cheat grass cover.
y_c	Cover class. 0 if $y_r \leq 0.2$, 1 otherwise.

Table 4: Prediction Targets

Variable	Description
c_1	Soil temperature and moisture regime class value.
c_2	Categorical value for vegetation cover type generalized from LANDFIRE (www.landfire.gov) vegetation data. 15 classes.
c_3	EPA Level III Ecoregions.

Table 5: Categorical Variables

where $\mathbf{C}_{i,j} = c_j^{(i)}$ the matrix of categorical variables associated with each data point, and $\mathbf{X} \in \mathbb{R}^{6602 \times 664}$, where $\mathbf{x}_{i,j} = x_j^{(i)}$. These values $\begin{bmatrix} \mathbf{w} & \mathbf{y} & \mathbf{C} & \mathbf{X} \end{bmatrix}$ are stored in comma delimited plain text csv format and may be accessed at <http://tobedetermined.gov>.

4 Methods

In our search for prospective models for producing high quality predictive maps at finest resolution across the full range of our study area, we explored classification models for the presence of cheatgrass. For the categorical mapping of cheatgrass occurrence, we chose two cover classes based on an evident separation in cheatgrass cover in the field data: 2% (low cover or absent) or cover > 2% (high cover or present). The location of field points colored by these classifications is shown in Figure 7. Continuous variables used in model development were standardized by subtracting the mean and then dividing by the standard deviation.

4.1 Baseline Models

For this study we explored two baseline methods for classification, Random Forest (RF), and Linear Discriminant Analysis (LDA). We chose LDA as a baseline model for comparison of the current approach to prior work on mapping cheatgrass with a linear model using satellite imagery. Random Forest was chosen as it has competitive performance with deep learning models in low resource data regimes as is the case of our 6,000 field locations. We used the implementation of Random Forest and Linear Discriminant Analysis provided in the Sci-kit learn python library.

4.2 Deep Learning Models

Deep learning models were chosen as prospective modeling candidates due to their potential for mapping high dimensional input features (such as our 664 continuous variables and 3 categorical values) into a lower dimensional representation which captures the inherent relationships between variables to greatest effect. The final layer of our deep learning models is the same transformation that is learned for logistic regression in the case of classification. A convenient way to view a deep learning model in our context is that it learns a parametrized non-linear mapping from input variables to a latent variable which is then used as input to a linear model. The parameters of the non-linear mapping and ultimate linear model are fit jointly, end to end using the ADAM variant of stochastic gradient descent.

We define the inputs to the neural network as a real valued vector $\hat{\mathbf{x}}^{(i)}$, its predictions $\hat{\mathbf{y}}^{(i)} \in \mathbb{R}^2$, a probability distribution over 2 classes, and the ground truth for a field location as $\mathbf{y}^{(i)}$. Since we are interested in an ablative analysis of the contributions in modeling performance from using data from satellite platforms jointly and independently we

Variable	Description
x_1	Latitude.
x_2	Longitude.
x_3	Elevation above sea level.
x_4	Potential Relative Radiation (available solar radiation for photosynthesis).
x_5	Long-term median cumulative winter precipitation based on PRISM.
x_6	Long-term median cumulative growing degree days (GDD), Oct-Apr.
$x_{7:18}$	Normal 30 year monthly max temperatures. Original study used month 5.
x_{19}	30-year normal annual maximum temperature.
$x_{20:31}$	30-year normal monthly min temperatures. Original study used months 3,11.
x_{32}	30-year normal annual minimum temperature.
$x_{33:44}$	30-year normal monthly precipitation for all months. Original study used months 3,6,7.
x_{45}	30-year normal annual precipitation.
x_{46}	Average of 30-year normal monthly maximum temperature during select winter months (Nov-Feb)
x_{47}	Cumulative precipitation during select winter season months (Dec-Feb).
x_{48}	Cumulative precipitation during select spring season months (Apr-May).
x_{49}	Cumulative precipitation during select summer season months (Jun-Sep).
x_{50}	30-year normal monthly minimum temperature during select winter months (Nov-Dec).
x_{52}	Delta NDVI (Normalized Difference Vegetation Index) from long term median peak NDVI in year of peak cumulative winter precipitation.
x_{53}	Landsat7 long-term median annual peak NDVI.
x_{54}	Landsat7 long-term median spring peak NDVI.
x_{55}	Landsat7 long-term median summer peak NDVI.
$x_{56:325}$	LandSat annual, spring, summer composite pixels.
$x_{326:661}$	MODIS annual, spring, summer composite pixels.
x_{662}	MODIS annual peak NDVI.
x_{663}	MODIS spring peak NDVI
x_{664}	MODIS summer peak NDVI.

Table 6: Continuous Variables. Shaded variables were used in making the original Linear Discriminant Analysis model.

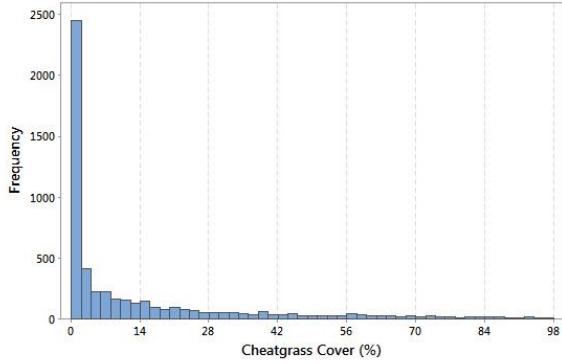


Figure 6: Histogram Showing Distribution of Cheatgrass Cover Values Measured at Field Locations.

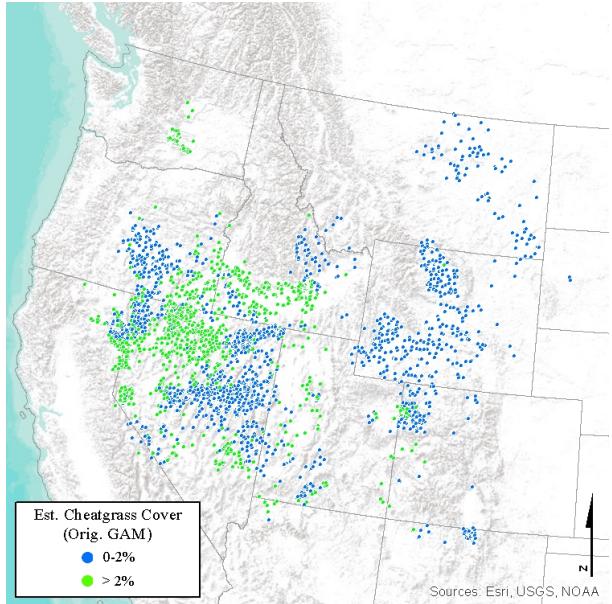


Figure 7: Classification of field samples

define subsets on the full set of continuous variables defined as \mathbf{x} in Table 6:

$$D_1(\mathbf{x}) = \mathbf{x}_{1:50} \quad \text{Biophysical variables} \quad (3)$$

$$D_2(\mathbf{x}) = \mathbf{x}_{1:50,326:664} \quad \text{Biophysical variables + MODIS} \quad (4)$$

$$D_3(\mathbf{x}) = \mathbf{x}_{1:325} \quad \text{Biophysical variables + LandSat} \quad (5)$$

$$D_4(\mathbf{x}) = \mathbf{x} \quad \text{All variables} \quad (6)$$

$$(7)$$

Since our data contains categorical variables derived from ecological analysis which we wish to incorporate into the model we define three learned embedding matrices for the categorical variables, $\mathbf{W}_{\text{cover}} \in \mathbb{R}^{15 \times k}$, $\mathbf{W}_{\text{eco}} \in \mathbb{R}^{23 \times k}$, and $\mathbf{W}_{\text{soil}} \in \mathbb{R}^{9 \times k}$, corresponding to the 15 classes of derived generalized LANDFIRE cover classes, the 23 level III ecoregions, and the 9 soil temperature and moisture regime class values exhibited in the full study region. The rows of these matrices correspond to classes for their respective categories. The number of columns k is the size of the embedding vector for a class; a so called hyperparameter or modeling choice. These embedding vectors are learned jointly with the other parameters of the neural network. Our input to a prospective neural network model is then defined using block matrix notation as:

$$\hat{\mathbf{x}}^{(i)} = \left[\mathbf{c}_1^{(i)} \mathbf{W}_{\text{soil}} \quad \mathbf{c}_2^{(i)} \mathbf{W}_{\text{cover}} \quad \mathbf{c}_3^{(i)} \mathbf{W}_{\text{eco}} \quad D_*(\mathbf{x}^{(i)}) \right] \quad (8)$$

where $\mathbf{c}_1^{(i)}$, $\mathbf{c}_3^{(i)}$, and $\mathbf{c}_2^{(i)}$ are so called one hot representations² of soil, land cover, and ecoregion class respectively, and $D_*(\mathbf{x}^{(i)})$ is one of the four subsets of continuous variables defined in Equations 3–7. The row vector matrix multiplications select the embedding vector associated with a particular class so that it can be concatenated to the rest of the continuous variables in the neural network input. For simplicity of notation the rest of discussion for deep learning model development we will omit subscripts i .

4.2.1 Deep Neural Network

Particular deep neural network (DNN) configurations for classification and regression are depicted in Figure 8. The model consists of L hidden layers \mathbf{h}_l which are recursively defined as:

$$\mathbf{h}_l = \text{RELU}(\mathbf{W}_l \mathbf{h}_{l-1} + \mathbf{b}_l) \quad (9)$$

$$\mathbf{h}_0 = \hat{\mathbf{x}} \quad (10)$$

²A one hot class vector representation has a number of elements equal to the number of classes with a 1 for the class represented and 0 everywhere else

where the nonlinear RELU operation is defined as the elementwise vector valued operation:

$$\text{RELU}(\mathbf{x})_i = \max(\mathbf{x}_i, 0) \quad (11)$$

The output of the DNN maps the final hidden representation \mathbf{h}_L to a probability distribution over cheatgrass occurrence classes of the form:

$$\hat{\mathbf{y}} = \text{softmax}(\mathbf{W}_{L+1}\mathbf{h}_L + \mathbf{b}_{L+1}) \quad (12)$$

where the softmax operation is defined as the elementwise vector valued operation:

$$\text{softmax}(\mathbf{x})_i = \frac{e^{\mathbf{x}_i}}{\sum_j e^{\mathbf{x}_j}} \quad (13)$$

The parameters \mathbf{W} , \mathbf{b} are optimized using stochastic gradient descent with the objective function being the cross-entropy loss between the predicted class distribution $\hat{\mathbf{y}}$ and true class distribution \mathbf{y} . Beyond this basic formulation we also employ dropout for normalization to avoid overfitting and batch normalization for better conditioned gradient updates.

4.2.2 Recurrent Neural Network

A deep neural network as described in 4.2.1 has the potential as a universal function approximator to model cheatgrass distribution given the field data with an optimal Bayes error. In practice however, since the optimal function may be difficult to find during the optimization we developed a joint recurrent neural network (RNN) model intended to take greater advantage of the time series implicit in the satellite imagery data. In this section we detail a model for fusing time series of data from different satellite platforms into a latent representation that can be used for linear classification. The model consists of two separate bidirectional RNNs which model time series of the MODIS and LandSat derived imagery products respectively. A particular instantiation of this modeling approach is depicted in Figure 9. More formally we define two bidirectional Long Short Term Memory networks which are operations on sequences of pixel vectors from the MODIS and LandSat platforms respectively, $\text{LSTM}_{\text{MOD}}(\mathcal{X})$ and $\text{LSTM}_{\text{LAN}}(\mathcal{X})$. Given a sequence of vectors \mathcal{X} each LSTM outputs a condensed vector representation of the time series. The time series for the MODIS and LandSat data consist of sequences of 16, and 10 vectors respectively, one for each year that pixel values were obtained. From the notation developed in 3.4, the t -th vector in the LandSat series is:

$$\mathcal{X}_t^{\text{LAN}} = \left[\mathbf{L}_{t,1}^{\text{annual}} \dots \mathbf{L}_{t,10}^{\text{annual}} \mathbf{L}_{t,1}^{\text{spring}} \dots \mathbf{L}_{t,10}^{\text{spring}} \mathbf{L}_{t,1}^{\text{summer}} \dots \mathbf{L}_{t,10}^{\text{summer}} \right] \quad (14)$$

whereas the t -th vector in the MODIS time series is:

$$\mathcal{X}_t^{\text{MOD}} = \left[\mathbf{M}_{t,1}^{\text{annual}} \dots \mathbf{M}_{t,10}^{\text{annual}} \mathbf{M}_{t,1}^{\text{spring}} \dots \mathbf{M}_{t,10}^{\text{spring}} \mathbf{M}_{t,1}^{\text{summer}} \dots \mathbf{M}_{t,10}^{\text{summer}} \right] \quad (15)$$

These resulting representation vectors are concatenated with the categorical embeddings and the continuous biophysical variables for the input to a deep neural network as described in 4.2.1 where:

$$\hat{\mathbf{x}} = \left[\mathbf{c}_1 \mathbf{W}_{\text{soil}} \mathbf{c}_2 \mathbf{W}_{\text{cover}} \mathbf{c}_3 \mathbf{W}_{\text{eco}} D_1(\mathbf{x}) \text{LSTM}_{\text{MOD}}(\mathcal{X}^{\text{MOD}}) \text{LSTM}_{\text{LAN}}(\mathcal{X}^{\text{LAN}}) \right] \quad (16)$$

4.3 Experimental Setup

In order to gage the effectiveness of incorporating satellite data from multiple platforms we trained each prospective model on each of four subsets of the complete data. For the neural network models we also include the categorical variables for each subset of the data.

- D_1 : Ecological Niche Factors
- D_2 : Ecological Niche Factors + MODIS imagery
- D_3 : Ecological Niche Factors + LandSat imagery
- D_4 : All Variables.

We performed a random hyperparameter search for each prospective model. For each model we chose 200 random configurations of hyperparameters within reasonable ranges for performance and computation time. For each of these runs we tested performance using 5-fold stratified cross-validation where field data was split into each of the 5 folds with equal distributions over the different ecoregions and generalized land cover classifications.

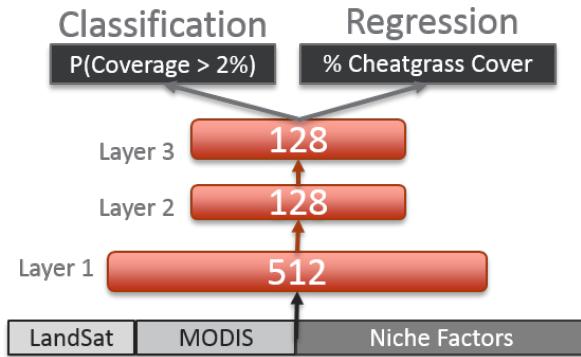


Figure 8: Deep Neural Network for Classification

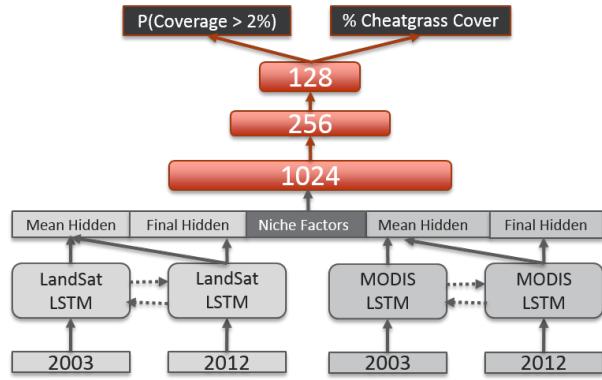


Figure 9: Recurrent Neural Network for Classification

5 Results and Analysis

In this section we discuss the results of experiments and analyze the model predictions for the final mapping product. In section 5.1.1 we discuss performance of all models trained for the field locations in the dataset. Following this discussion we perform a detailed ablation analysis for the described subsets of variables on the best performing model. To this end we first discuss training statistics for models from each variable subset, performing an ablation analysis on key training statistics. We then look more closely at the classification performance to assess potential weaknesses in the prospective model by analysis of predictions versus true class considering high level derived variables such as ecoregion, coverage class and soil regime. Finally, we shift our focus to analysis and assessment of the complete mapping product derived from the highest performing model, with specific focus on the variation of confidence for various points in the full study region that were not present in the training data.

5.1 Performance on Field measurement locations

In this section we report statistics on field measurement locations.

5.1.1 Quantitative assessment

Quantitative results are reported for the average performance across cross-validation folds for each of the 4 subsets of input variables. Table 7 shows the outcomes of experiments in terms of accuracy for classification prediction. The best performing subset of input variables is bolded in each row of the tables, while the best overall performing model, the DNN using all available input data is highlighted in red. We note a few trends. Incorporating satellite data of either form leads to better results. In addition, Random forest and LDA do not gain any advantage from incorporating both MODIS and LandSat imagery whereas the neural network models improve in accuracy from incorporating imagery from both platforms. Finally, we note that the recurrent neural network did not perform as well as the DNN, possibly due to the larger number of parameters to be fit on the same number of data points.

We chose the best performing model in terms of accuracy (DNN with all input variables) to create the final mapping product, and for this model we report several classification statistics for each of the four input variable subsets in Table 10. Here we see the same trend for recall, F-score, and area under the ROC curve in that adding imagery from either satellite platform gives the most significant improvement, with the LandSat imagery demonstrating more improvement. The best results for all these metrics are once again from the inclusion of imagery from both satellite platforms.

5.1.2 Training Statistics

In this section we analyse training statistics for the best performing model. We first look at plots of train versus test performance for the best performing cross-validation fold for each of the input variable subsets. Looking at several different metrics we see training is noisier and has a greater tendency to overfit with inclusion of the modis data. Conversely we see that incorporating the LandSat data into the model has a regularizing effect which limits the amount of overfitting so that test performance converges closer to train performance. We also consider the dropout probability for each data subset (for each model run a random dropout probability was selected to be between 0.5 and 1.0. Table ?? shows the dropout probabilities and hidden layer sizes for the different data subsets. The columns for using MODIS

	D_1	D_2	D_3	D_4
LDA	73.816	78.279	80.653	79.826
RF	80.165	81.677	82.502	82.798
DNN	79.604	82.004	82.456	83.203
RNN	—	81.388	81.820	82.707

Table 7: Classification accuracy of models for subsets of data

	Prec.	Rec.	F-scr	AUC	Acc.	Std.
D_1	—	—	—	—	—	—
D_2	—	—	—	—	—	—
D_3	—	—	—	—	—	—
D_4	0.830	0.840	0.835	0.904	0.827	0.011

Table 9: Performance across a range of metrics for most accurate RNN models for each subset of input variables. The last column reports the standard deviation of the accuracy across different cross-validation folds.

	Prec.	Rec.	F-scr	AUC	Acc.	Std.
D_1	0.795	0.820	0.808	0.866	0.796	0.025
D_2	0.844	0.803	0.823	0.893	0.820	0.014
D_3	0.831	0.833	0.832	0.900	0.825	0.015
D_4	0.837	0.842	0.839	0.902	0.832	0.012

Table 8: Performance across a range of metrics for most accurate DNN models for each subset of input variables. The last column reports the standard deviation of the accuracy across different cross-validation folds.

	Prec.	Rec.	F-scr	AUC	Acc.	Std.
D_1	0.807	0.815	0.811	0.883	0.802	0.024
D_2	0.825	0.823	0.824	0.891	0.816	0.024
D_3	0.827	0.838	0.833	0.901	0.825	0.021
D_4	0.835	0.835	0.835	0.904	0.828	0.025

Table 10: Performance across a range of metrics for most accurate RF models for each subset of input variables. The last column reports the standard deviation of the accuracy across different cross-validation folds.

only or LandSat show that these models have comparable dropout probabilities and network size. The MODIS only input model does have the largest network which greatens the chance of overfitting but a large dropout probability which should have more of a regularizing effect. A more careful analysis with identical sized networks and dropout probabilities would confirm what the present analysis suggests (that the MODIS data has some properties which encourage overfitting). With careful use of early stopping we can see in figure 53 that the best loss can in fact be achieved by incorporating all the data.

5.1.3 Qualitative assessment

In this section we characterize the overall performance of the best performing DNN models. Figure 71 shows a confusion matrix for classification for the best performing DNN model using all the input variables. The types of mistakes the model makes in classification are evenly balanced in terms of false positives and false negatives, however since a smaller proportion of the data contains negative examples it in fact slightly favors false positives. We can see this more clearly in figure 63 which plots accuracy in terms of actual cheat grass coverage values. The worst performance is in the 0-2 percent range which contains only negative examples, so the only mistake that can be made is a false positive.

Looking closer at prediction performance we next analyze the model in terms of accuracy across individual ranges of cheatgrass coverage, ecoregions, land cover classes, and soil and climactic conditions. Figures 60-61 show that the models incorporating the LandSat data attain 100 % accuracy for actual coverage ranging from 70-100 % cheat grass coverage and accuracy above 90 % up down to 40 % coverage. This is despite the fact that there are fewer examples in

	D_1	D_2	D_3	D_4
Dropout Probability	0.275	0.350	0.30	.5
Layer 1	256	256	256	256
Layer 2	256	256	64	32
Layer 3	64	128	256	256

Table 11: Hyperparameters for most accurate DNN models for each subset of input variables.

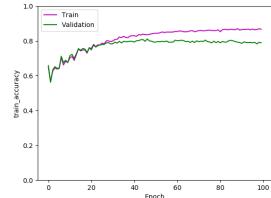


Figure 10: D_1

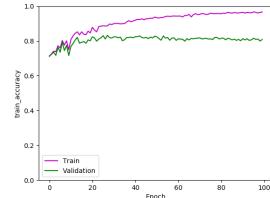


Figure 11: D_2

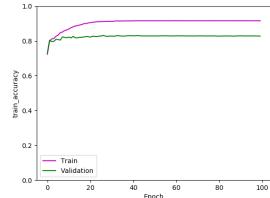


Figure 12: D_3

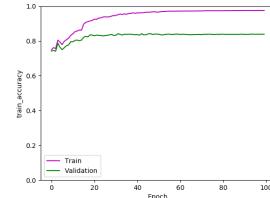


Figure 13: D_4

Figure 14: Train versus test accuracy for the best performing cross-validation fold for input variable subsets

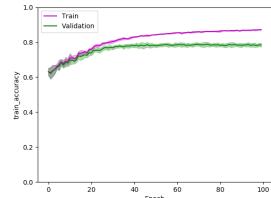


Figure 15: D_1

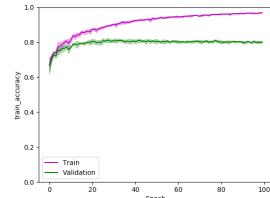


Figure 16: D_2

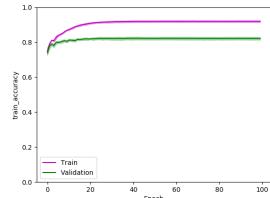


Figure 17: D_3

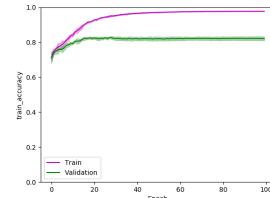


Figure 18: D_4

Figure 19: Train versus test accuracy for the average of cross-validation folds for input variable subsets

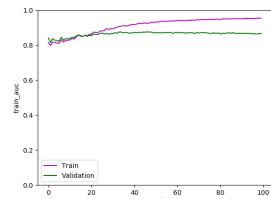


Figure 20: D_1

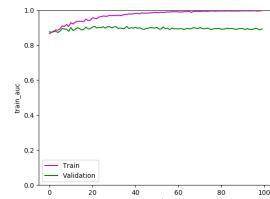


Figure 21: D_2

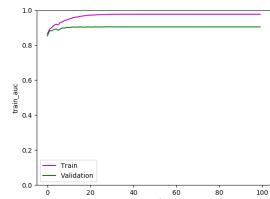


Figure 22: D_3

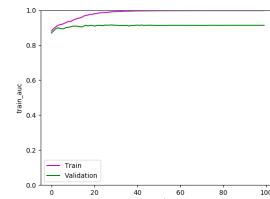


Figure 23: D_4

Figure 24: Train versus test area under the ROC curve for best performing cross-validation fold for input variable subsets

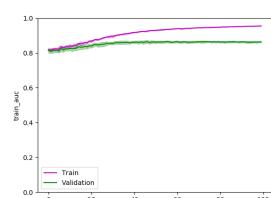


Figure 25: D_1

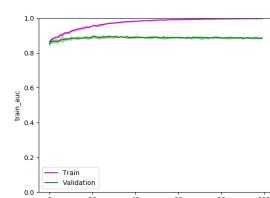


Figure 26: D_2

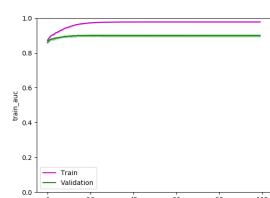


Figure 27: D_3

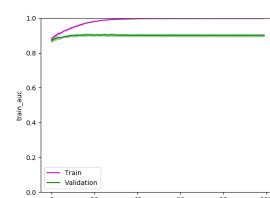


Figure 28: D_4

Figure 29: Train versus test area under the ROC curve for average of cross-validation folds for input variable subsets

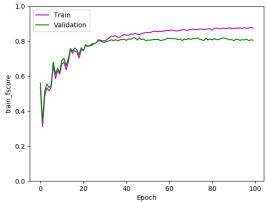


Figure 30: D_1

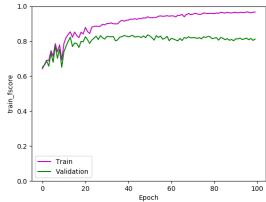


Figure 31: D_2

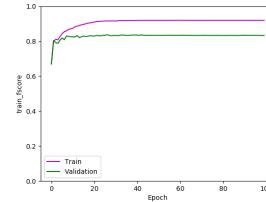


Figure 32: D_3

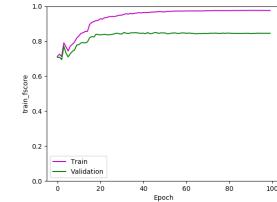


Figure 33: D_4

Figure 34: Train versus test F-score of best performing cross-validation fold for input variable subsets

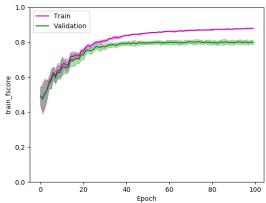


Figure 35: D_1

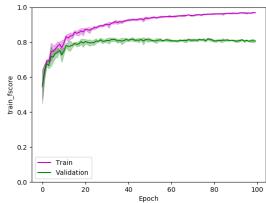


Figure 36: D_2

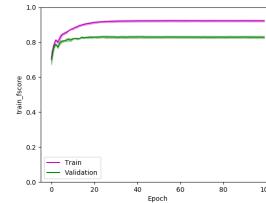


Figure 37: D_3

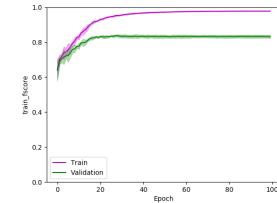


Figure 38: D_4

Figure 39: Train versus test F-score of average cross-validation score for input variable subsets

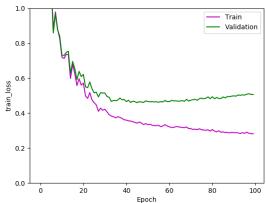


Figure 40: D_1

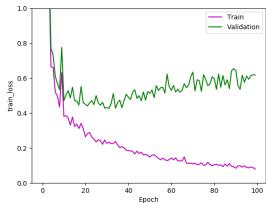


Figure 41: D_2

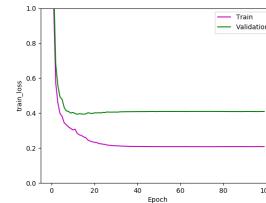


Figure 42: D_3

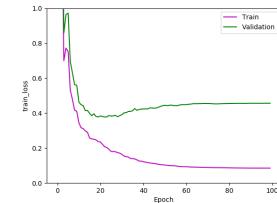


Figure 43: D_4

Figure 44: Train versus test cross entropy loss for best performing cross-validation fold for input variable subsets

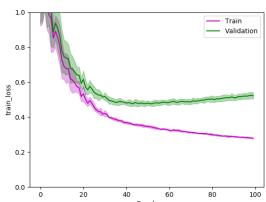


Figure 45: D_1

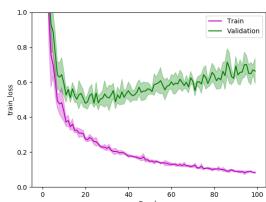


Figure 46: D_2

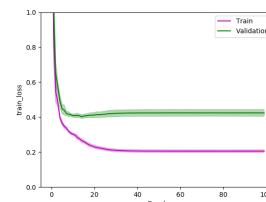


Figure 47: D_3

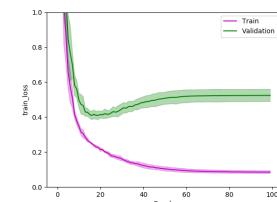


Figure 48: D_4

Figure 49: Train versus test cross entropy loss for average of cross-validation folds for input variable subsets

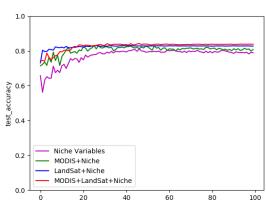


Figure 50: Accuracy

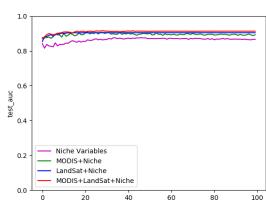


Figure 51: Area under ROC

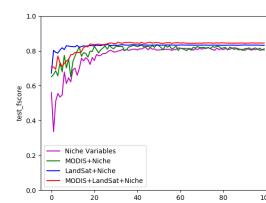


Figure 52: F-score

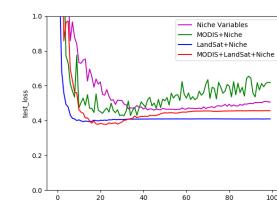


Figure 53: Cross-entropy

Figure 54: Test metrics plotted for input variable subsets

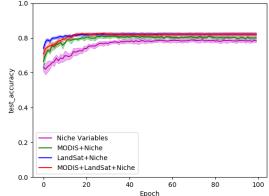


Figure 55: Accuracy

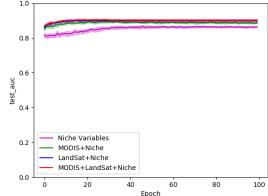


Figure 56: Area under ROC

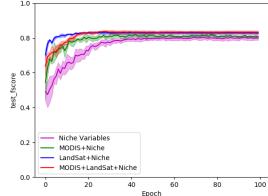


Figure 57: F-score

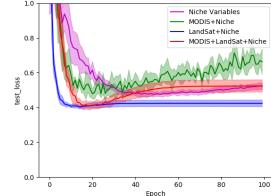


Figure 58: Cross-entropy

Figure 59: Test metrics plotted for input variable subsets

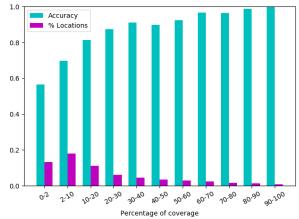


Figure 60: D_1

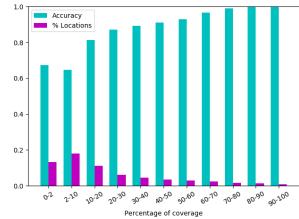


Figure 61: D_2

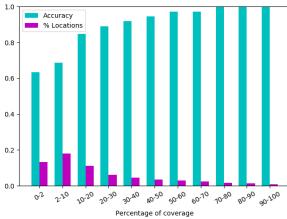


Figure 62: D_3

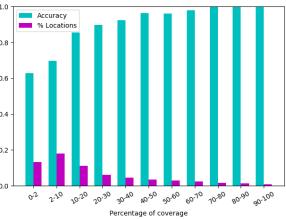


Figure 63: D_4

Figure 64: Accuracy for ranges of cheat grass coverage.

the data set for these high coverage locations. Since the overwhelming majority of false positives and false negatives fall within the 0-40 % range, this suggests that changing the class boundary from 2% to 40% may lead to more accurate classification at the expense of the overall utility of a final mapping product. We pay particular interest in judging if the model performs better across different ecoregions due to an unbalanced representation in the data set or something intrinsic in the environment itself. Across the other ways of splitting the data into ecoregion, soil moisture temperature regime, and land cover class we see that accuracy is not correlated with frequency of occurrence. Although there does seem to be a minimum threshold for number of examples for an ecoregion to attain acceptable accuracy, considering the low performance of ecoregions 25 and 16 and corresponding small number of examples.

- Map of field data symbolized by classification error

5.2 Performance on Full Study Region

In this section we analyze the geographic variability in cross-validation model predictions among other things.

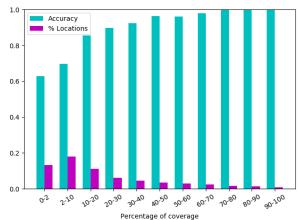


Figure 65: D_1

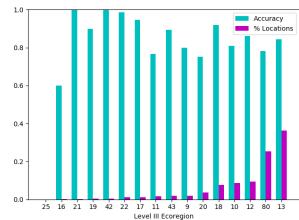


Figure 66: D_2

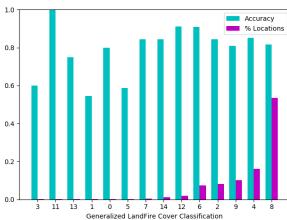


Figure 67: D_3

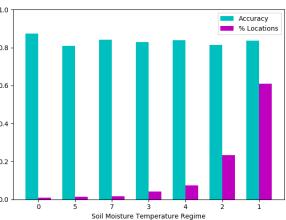


Figure 68: D_4

Figure 69: Accuracy for different subsets of the field locations.

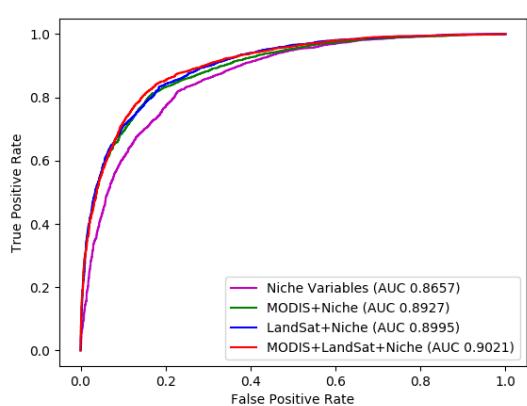


Figure 70: ROC curves for best performing DNN models on different data subsets

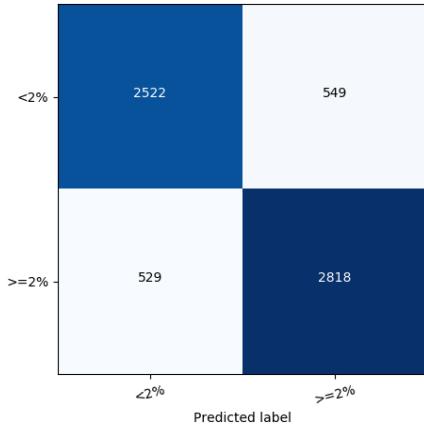


Figure 71: Confusion matrix for best performing DNN model.

5.2.1 Pixel-wise count of cheatgrass class across model predictions

5.2.2 Block statistics (count) of cheatgrass class (analogous to block area), then look at range of block counts

6 Conclusion and Future Work

References

- [1] ADAM, E., MUTANGA, O., AND RUGEGE, D. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management* 18, 3 (2010), 281–296.
- [2] ALLOUCHE, O., TSOAR, A., AND KADMON, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (tss). *Journal of applied ecology* 43, 6 (2006), 1223–1232.
- [3] BELLUCO, E., CAMUFFO, M., FERRARI, S., MODENESE, L., SILVESTRI, S., MARANI, A., AND MARANI, M. Mapping salt-marsh vegetation by multispectral and hyperspectral remote sensing. *Remote sensing of environment* 105, 1 (2006), 54–67.
- [4] BRADLEY, B. A. Regional analysis of the impacts of climate change on cheatgrass invasion shows potential risk and opportunity. *Global Change Biology* 15, 1 (2009), 196–208.
- [5] BRADLEY, B. A. Assessing ecosystem threats from global and regional change: hierarchical modeling of risk to sagebrush ecosystems from climate change, land use and invasive species in nevada, usa. *Ecography* 33, 1 (2010), 198–208.
- [6] BRADLEY, B. A. Remote detection of invasive plants: a review of spectral, textural and phenological approaches. *Biological invasions* 16, 7 (2014), 1411–1425.
- [7] BRADLEY, B. A., AND MUSTARD, J. F. Identifying land cover variability distinct from land cover change: cheatgrass in the great basin. *Remote Sensing of Environment* 94, 2 (2005), 204–213.
- [8] CE, C. Ecological regions of north america: toward a common perspective.
- [9] CHEN, Y., LIN, Z., ZHAO, X., WANG, G., AND GU, Y. Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected topics in applied earth observations and remote sensing* 7, 6 (2014), 2094–2107.
- [10] CRISCI, C., GHATTAS, B., AND PERERA, G. A review of supervised machine learning algorithms and their applications to ecological data. *Ecological Modelling* 240 (2012), 113–122.
- [11] CUTLER, D. R., EDWARDS JR, T. C., BEARD, K. H., CUTLER, A., HESS, K. T., GIBSON, J., AND LAWLER, J. J. Random forests for classification in ecology. *Ecology* 88, 11 (2007), 2783–2792.

- [12] DALY, C., HALBLEIB, M., SMITH, J. I., GIBSON, W. P., DOGGETT, M. K., TAYLOR, G. H., CURTIS, J., AND PASTERIS, P. P. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous united states. *International Journal of Climatology: a Journal of the Royal Meteorological Society* 28, 15 (2008), 2031–2064.
- [13] DALY, C., NEILSON, R. P., AND PHILLIPS, D. L. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of applied meteorology* 33, 2 (1994), 140–158.
- [14] DI LUZIO, M., JOHNSON, G. L., DALY, C., EISCHEID, J. K., AND ARNOLD, J. G. Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous united states. *Journal of Applied Meteorology and Climatology* 47, 2 (2008), 475–497.
- [15] ELITH, J., AND GRAHAM, C. H. Do they? how do they? why do they differ? on finding reasons for differing performances of species distribution models. *Ecography* 32, 1 (2009), 66–77.
- [16] GONG, B., IM, J., AND MOUNTRAKIS, G. An artificial immune network approach to multi-sensor land use/land cover classification. *Remote Sensing of Environment* 115, 2 (2011), 600–614.
- [17] GOVENDER, M., CHETTY, K., AND BULCOCK, H. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water Sa* 33, 2 (2007), 145–151.
- [18] GUIRADO, E., TABIK, S., ALCARAZ-SEGURA, D., CABELO, J., AND HERRERA, F. Deep-learning versus obia for scattered shrub detection with google earth imagery: *Ziziphus lotus* as case study. *Remote Sensing* 9, 12 (2017), 1220.
- [19] HESTIR, E. L., KHANNA, S., ANDREW, M. E., SANTOS, M. J., VIERS, J. H., GREENBERG, J. A., RAJAPAKSE, S. S., AND USTIN, S. L. Identification of invasive vegetation using hyperspectral remote sensing in the california delta ecosystem. *Remote Sensing of Environment* 112, 11 (2008), 4034–4047.
- [20] JEEVALAKSHMI, D., REDDY, S. N., AND MANIKIAM, B. Land cover classification based on ndvi using landsat8 time series: A case study tirupati region. In *Communication and Signal Processing (ICCSP), 2016 International Conference on* (2016), IEEE, pp. 1332–1335.
- [21] KUSSLU, N., LAVRENIUK, M., SKAKUN, S., AND SHELESTOV, A. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters* 14, 5 (2017), 778–782.
- [22] LAWRENCE, R. L., WOOD, S. D., AND SHELEY, R. L. Mapping invasive plants using hyperspectral imagery and breiman cutler classifications (randomforest). *Remote Sensing of Environment* 100, 3 (2006), 356–362.
- [23] LI, L., ZHAO, Y., FU, Y., PAN, Y., YU, L., AND XIN, Q. High resolution mapping of cropping cycles by fusion of landsat and modis data. *Remote Sensing* 9, 12 (2017), 1232.
- [24] LU, D., AND WENG, Q. A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing* 28, 5 (2007), 823–870.
- [25] MAXWELL, A. E., WARNER, T. A., AND FANG, F. Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing* 39, 9 (2018), 2784–2817.
- [26] MOUNTRAKIS, G., IM, J., AND OGOLE, C. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 3 (2011), 247–259.
- [27] OMERNIK, J. M., AND GRIFFITH, G. E. Ecoregions of the conterminous united states: evolution of a hierarchical spatial framework. *Environmental management* 54, 6 (2014), 1249–1266.
- [28] RODRIGUEZ-GALIANO, V. F., GHIMIRE, B., ROGAN, J., CHICA-OLMO, M., AND RIGOL-SANCHEZ, J. P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 67 (2012), 93–104.
- [29] ROGAN, J., FRANKLIN, J., STOW, D., MILLER, J., WOODCOCK, C., AND ROBERTS, D. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sensing of Environment* 112, 5 (2008), 2272–2283.

- [30] SALEQUZZAMAN, M. Ecoregions: A spatial framework for environmental management. *Water Encyclopedia* 2 (2005), 667–668.
- [31] SHAO, Y., AND LUNETTA, R. S. Comparison of support vector machine, neural network, and cart algorithms for the land-cover classification using limited training data points. *ISPRS Journal of Photogrammetry and Remote Sensing* 70 (2012), 78–87.
- [32] THENKABAIL, P. S., AND LYON, J. G. *Hyperspectral remote sensing of vegetation*. CRC Press, 2016.
- [33] THORNTON, P. E., RUNNING, S. W., AND WHITE, M. A. Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology* 190, 3-4 (1997), 214–251.
- [34] USGS, D. Landsat surface reflectance-derived spectral indices, 2014.
- [35] WEST, A. M., EVANGELISTA, P. H., JARNEVICH, C. S., KUMAR, S., SWALLOW, A., LUIZZA, M. W., AND CHIGNELL, S. M. Using multi-date satellite imagery to monitor invasive grass species distribution in post-wildfire landscapes: An iterative, adaptable approach that employs open-source data and software. *International Journal of Applied Earth Observation and Geoinformation* 59 (2017), 135–146.
- [36] ZHANG, L., ZHANG, L., AND DU, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine* 4, 2 (2016), 22–40.