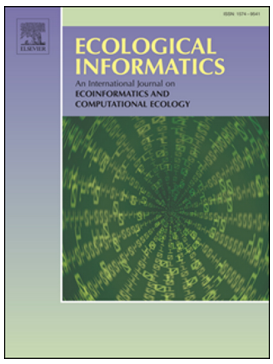

Mapping an invasive grass in the northwestern US with fused satellite time series and biophysical features

Ty C. Nietupski, Hailemariam Temesgen, Becky K. Kerns

PII: S1574-9541(24)00515-6

DOI: https://doi.org/10.1016/j.ecoinf.2024.102973

Reference: ECOINF 102973

To appear in: *Ecological Informatics*

Received date: 20 May 2024

Revised date: 9 November 2024

Accepted date: 18 December 2024

Please cite this article as: T.C. Nietupski, H. Temesgen and B.K. Kerns, Mapping an invasive grass in the northwestern US with fused satellite time series and biophysical features, *Ecological Informatics* (2024), https://doi.org/10.1016/j.ecoinf.2024.102973

Mapping an invasive grass in the northwestern US with fused satellite time series and biophysical features

Short title: Mapping invasive grass with time series

Ty C. Nietupski[1]*, Hailemariam Temesgen[2], Becky K. Kerns[3]

[1] *USDA Forest Service, Geospatial Technology and Applications Center, Fort Collins, CO, USA*

[2] *Department of Forest Engineering, Resources and Management, Oregon State University, Corvallis, Oregon, USA*

[3] *USDA Forest Service, Pacific Northwest Research Station, Corvallis, OR, USA*

Address correspondence to: ty.nietupski@usda.gov;

ORCID IDs: Ty C. Nietupski, 0000-0003-0248-5753; Hailemariam Temesgen, 0000-0002-9604-650X; Becky K. Kerns, 0000-0003-4613-2191

Abstract

The introduction and spread of the invasive annual grass *Ventenata dubia* (ventenata) has incited concern from land managers in the Inland Northwestern United States. Maps describing ventenata's distribution would be a valuable management asset but have not been developed. Techniques using satellite time-series have been used to detect annual grasses' unique phenological qualities (timing of biological events); however, detection is complicated by the co-occurrence of phenologically similar species. This study aimed to examine the capability of land surface phenology derived from fused satellite imagery to map and gain insight into the ventenata invasion. We evaluated the influence of land surface phenology, climate, and topo-edaphic predictors on ventenata classification and examined differences in the distributions predicted from three random forest models: 1) hybrid (phenology, climate, topo-edaphic), 2) bioclimatic (climate, topo-edaphic), and 3) phenology (phenology). The hybrid model indicated that 7.7 % (5,454 km^2) of the Blue Mountains Ecoregion may contain heavy ventenata invasion and that many populations were located in forest/non-forest transition zones and forest openings. The phenology model predicted ventenata populations in regions

occupied by other annual grasses, suggesting that land surface phenology characteristics alone could not differentiate ventenata from other invasive annual grasses. The bioclimatic model identified suitable habitat but overpredicted invasion extent in heavily treed areas. These results suggested that incorporating phenology with climatic predictors effectively differentiates invasive annual grasses when phenological patterns are similar, but habitat requirements differ.

Keywords: land surface phenology; *Ventenata dubia*; invasive annual grass; species distribution; remote sensing

1. Introduction

Numerous nonnative annual grass species have been introduced to the western United States. Some introductions have resulted in extensive changes to ecosystem structure and function (Levine et al. 2003; Keeley et al. 2005; Simberloff et al. 2013), with some of the greatest impacts coming from these species' ability to alter disturbance regimes (D'Antonio and Vitousek 1992; Mack and D'Antonio 1998; Brooks et al. 2004). Winter annual grasses including cheatgrass (*Bromus tectorum*), medusahead (*Taeniatherum caput-medusae*), and red brome (*Bromus rubens*), have long been recognized for their negative effects to native plant communities and fire regimes (Beatley 1966; Young and Evans 1970; Mack 1981; Young 1992; Knapp 1996). Additional annual grass introductions may distribute negative impacts to new habitats and wider geographic domains (Tortorelli et al. 2020).

Addressing the present and future implications annual grass invasion requires an understanding of the spatial and temporal distribution of these species (Forsyth et al. 2012). The development and evaluation of management and policy strategies is hampered by a lack of information about newly introduced species' invasion status and trends, including range and abundance (Jetz et al. 2019; Funk et al. 2020). Therefore, the accuracy of invasive species maps can substantially impact the long-term decision outcomes (Cheney et al. 2018). Species detection and mapping from remotely sensed data at

local and regional levels can assist decisions by revealing locations of problematic populations, the spread rate, and site characteristics that promote spread or persistence (Funk et al. 2020).

Advances in plant species detection using remotely sensed earth observations have created opportunities for rapid and accurate mapping (Huang and Asner 2009; Bradley 2014; Royimani et al. 2019). Winter annual grass species, like cheatgrass, were initially mapped with land surface phenology (LSP; timing of cyclical land surface events like green up) estimated from satellite time series from sensors like the Moderate Resolution Imaging Spectroradiometer (MODIS; Bradley and Mustard 2008; Boyte and Wylie 2016; Bradley et al. 2017; Peterson 2005; West et al. 2017). However, species mixtures with similar phenology may reduce confidence in models based on coarse resolution phenology (i.e., 250+ m) because the phenology data likely represent a combination of introduced grasses (Pastick et al. 2020; Weisberg et al. 2021). Additionally, low-resolution data (e.g., MODIS; 250 m) are often too coarse to detect sparse populations. Earth observation data with high spatial or spectral resolution (e.g., hyperspectral and hyperspatial) can more readily capture smaller populations (Ustin et al. 2002; Noujdina and Ustin 2008; Irisarri et al. 2009; Olsson et al. 2011; Rupasinghe and Chow-Fraser 2021), but data costs, storage, processing, and spatiotemporal coverage across large extents can become limiting factors. To accommodate the spatial and temporal requirements for finer resolution LSP, one can use a combination of sensor data (fused or harmonized data) to characterize the cycle of vegetation development. Spatio-temporal image fusion is one cost-effective approach for creating moderate resolution (i.e., 30 m) time series, which can then be used to estimate LSP (Nietupski et al. 2021), and potentially be applied to map plant species. Harmonized Landsat and Sentinel-2 data have recently been used to map the distribution of cheatgrass in the western US (Pastick et al. 2020), demonstrating the potential of similar methods.

Models based solely on environmental relationships, referred to as habitat suitability models, have also been employed to determine the potential distribution of invasive species (Austin 2007;

Franklin 2010; Elith 2016). Remotely sensed data have occasionally been incorporated in habitat suitability models with the assumption that model predictions represent the potential distribution (He et al. 2015; Ahmed et al. 2020), but by using remotely sensed data, the models likely capture something closer to the actual distribution (Bradley et al. 2012). While a species' potential distribution can be helpful for management, models, and maps of the actual distribution provide an understanding unattainable from habitat suitability alone. Especially in circumstances where multiple species have similar phenological or spectral characteristics, models of the actual distribution may benefit by incorporating multiple sources of biophysical, climatic, and spectral-temporal predictors (Zimmermann et al. 2007). Biophysical variables like those derived from digital elevation models have demonstrated use in previous efforts to map invasive grasses' actual distribution (Peterson 2005; Boyte and Wylie 2016; Bradley et al. 2017; Pastick et al. 2020). In these cases, biophysical information helps to draw relationships between the species presence or abundance and relatively stable features of the physical environment (Franklin 2010). By drawing this relationship, the biophysical data helps to limit the scope of areas where the species could exist while spectral-temporal data helps to identify locations where spectral-temporal patterns align with vegetation characteristics.

The test case and subject of this study, *Ventenata dubia* (wiregrass; henceforth ventenata), is a winter annual grass species from Mediterranean Eurasia and northern Africa (Wallace et al. 2015) that was first recorded the interior Pacific Northwest of the United States in the 1950s. Ventenata's range and abundance did not increase rapidly until more recently (e.g., approximately the early 2000's; personal communication with local land managers). Ecological and economic impacts resulting from the rapid expansion into grasslands and shrublands (Jones et al. 2018; Averett et al. 2020; Endress et al. 2020) and non-forested openings within forested mosaics (Tortorelli et al. 2020) are only beginning to unfold. While other exotic annual grass species, especially cheatgrass, are also present within the Pacific Northwest, these species have been less problematic at higher elevations, a result of reduced habitat

suitability and increased competition from native perennial plants (Chambers et al. 2007, 2014; Roundy et al. 2018).

The existence of substantial ventenata populations in forest openings of the northwestern US appears to be unique compared to other annual grass species and is likely related to ventenata's occupancy of distinctly different niche space (Tortorelli et al. 2020; Applestein and Germino 2021). Some forest openings of the interior Pacific Northwest provide unique habitat for species tolerant of harsh environmental conditions and the low productivity of these openings reduces fire ignition and spread (Agee 1994). Consequently, these forest openings have served as fuel breaks in wildfire control efforts. An accumulation of fine fuel biomass from annual grass invasion in these openings may lead to uncharacteristic fire behavior and harm unique and vulnerable species, like low sagebrush (*Artemisia arbuscula*) and stiff sagebrush (*Artemisia rigida*). In at least one recent fire within the interior Pacific Northwest, ventenata accelerated wildfire spread in areas previously resistant to wildfire (Hallmark and Romero 2015), inciting concern and interest from public and private land managers.

Due to ventenata's potential impact to fire behavior, we investigated the utility of a 30 m LSP product (Nietupski et al. 2021), niche-related biophysical predictors, and a combination of these data to detect and map areas heavily invaded by ventenata (i.e., populations with 20 % or more ventenata cover; hereafter referred to as 'heavily invaded'). Populations with low cover have reduced potential to influence both land surface phenology (Bradley et al. 2017; West et al. 2017) and the local plant community composition, structure, and fire behavior (Bradley et al. 2017; Endress et al. 2020). We focused on the 71,000 km^2 Blue Mountains Ecoregion (BME) of the interior Pacific Northwest, which is also invaded by other nonnative annual grasses (e.g., cheatgrass). We developed three models: 1) a hybrid model incorporating phenology and biophysical predictors, 2) a bioclimatic model using only biophysical predictors, and 3) a phenology model using only phenology predictors. We refer to these as the hybrid, bioclimatic, and phenology models. Our main objectives were to 1) produce a model with

the best possible discrimination power to predict and map the actual distribution of heavily invaded areas, and 2) evaluate and compare the discriminatory capacity, predictor importance and relationships, and spatial prediction patterns from the three models. Lastly, we used our results to gain insights into the invasion by examining spatial patterns and associations with plant communities.

2. Materials and methods

We applied a two-phase framework to investigate the detection and mapping of ventenata from land surface phenology and biophysical features (Figure 1). First, in situ observations of ventenata, imagery, and biophysical data were acquired. In situ observations were aggregated from multiple sources, filtered, and cleaned. A time series of Landsat 8 and MODIS imagery were processed to produce a daily time series, which was used to extract phenological metrics. Gridded climate, soils, and elevation data were used to calculate seasonal climate, soil texture, and terrain variables. Next, three random forest models were developed using the phenology and biophysical predictors. Model evaluation was performed on each model to compare their performance across the region and within elevation bands and vegetation types. Classifications from final models used to investigate associations between heavily invaded areas and vegetation types.

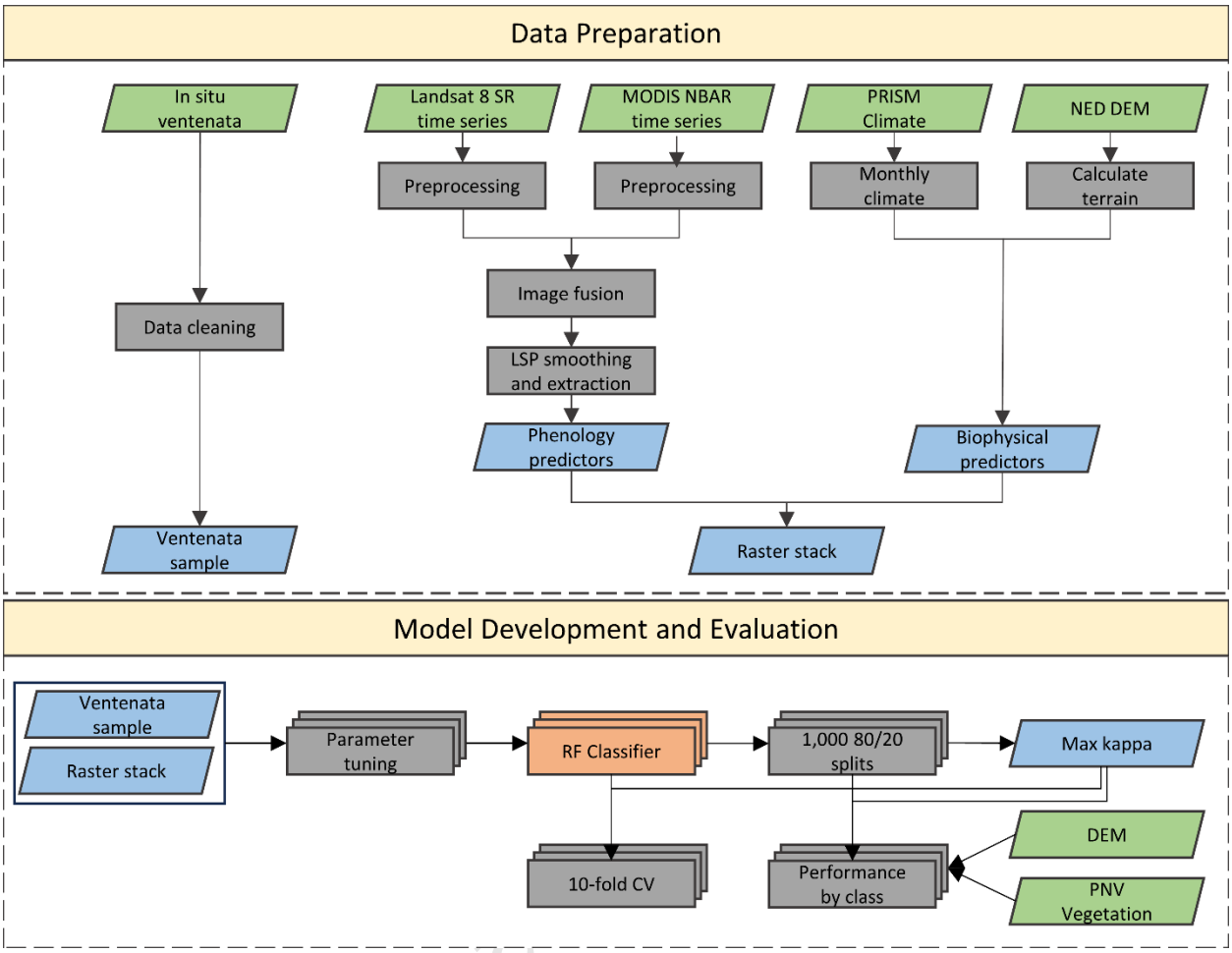

Figure 1. The framework applied to investigate the detection and mapping of ventenata populations.2.1 Study area

The BME (Omernik 1987) is a mosaic of mountains, valleys, and plateaus covering approximately 71,000 km^2 of the interior Pacific Northwest (Figure 2). The BME is bordered to the west by the Cascade Mountains and to the east by the Rocky Mountains. Consequently, the BME flora represents a transition between these major mountain ranges where unique geology drives local community composition. Grasslands, including remnants of Palouse Prairie, are primarily present in the northern portion of the ecoregion, while shrublands, dominated by sagebrush (*Artemisia spp.*) and Western juniper (*Juniperus occidentalis*) woodlands, are more common in the south. A substantial portion of the region is forested,

of which the majority is dominated by ponderosa pine (*Pinus ponderosa*). Still, parts of the region also support dry and moist mixed conifer and subalpine forests. Elevation ranges from 235 m at the Snake River along the Washington-Idaho border to 2,997 m at Sacajawea Peak in the Wallowa Mountains. Annual temperature and precipitation vary widely, but the region's climate is generally characterized by warm, dry summers and cold, wet winters. Mean annual temperatures range from -1 to 13 °C (PRISM Climate Group 2012). Precipitation is primarily received as snow and rain during the winter and spring, with annual totals from 20 to 195 cm (PRISM Climate Group 2012).

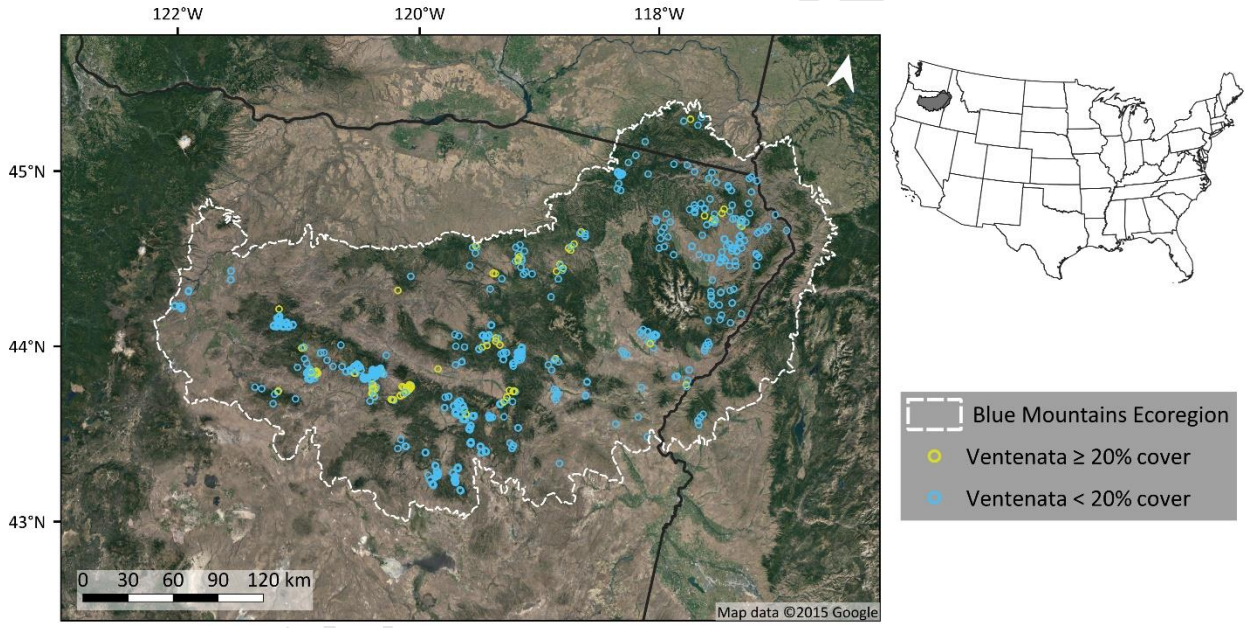

Figure 2. Field-observed populations of ventenata across the Blue Mountains Ecoregion.

2.2 Data

2.2.1 In situ

We compiled records of ventenata absence and abundance from multiple sources. We evaluated these data for spatial accuracy and temporal proximity to the target year (i.e., 2017) where observations recorded within one year of 2017 and with plot sizes close to 30 m were retained. When

multiple observations occurred within 30 m, the observation with the best combination of temporal proximity, spatial accuracy, and plot size was retained. While original data sources were collected using differing protocols (i.e., presence/absence, line-point intercept, ocular cover estimation), no single data source provided a reasonable sample size due to the region's scarcity of ventenata monitoring efforts. Some previous studies have concluded that different cover methods result in similar cover estimates (Stohlgren et al. 1995, Korb et al. 2003, De Stefano et al. 2021), but others have determined that ocular cover estimation can result in systematically higher estimates than line-point intercept estimation (Dethier et al. 1993, Bonham 2013).

To develop a classification model from the data, we converted ventenata cover observations to binary classes where populations with greater than or equal to 20 % cover were assigned to the positive class and all others were assigned to the negative class. This threshold aligns with the minimum cover reported for the detection of cheatgrass populations with land surface phenology (Bradley et al. 2017) and was chosen to represent the potential fuel-changing influence of heavily invaded areas. After applying the cover threshold, 156 of observations included in the negative class had a median venentata cover of 3.7 %. The resulting dataset had 944 observations split between nine sources, including the United States Forest Service (USFS) and Oregon State University (Table 1).

Table 1. Observation sources used to model areas heavily invaded by *Ventenata dubia* in the Blue Mountains Ecoregion.

Source	Heavily invaded (n = 94)	Absence (n = 850)	Year	Sample Radius (m)	Cover Estimation Method	Notes
Dean	3	34	2016	20	Line-point intercept	Prescribed fire study field surveys, unpublished.
Downing	2	139	2017	35	Ocular Quadrat	Fire refugia field surveys (Downing et al. 2019)
Hogrefe	0	92	2017	25	Line-point intercept	Rangeland biomass field surveys, unpublished
Johnston	0	377	2017	18	Presence-absence	Blue Mountains Forest Partners survey data, (Johnston et al. 2021)
Lawrence	0	2	2017	30	Line-point intercept	Sage-grouse field surveys, unpublished
Nietupski	39	129	2017	30	Line-point intercept	Ventenata field surveys, unpublished
Tortorelli	42	62	2018	35	Ocular Quadrat	Ventenata field surveys (Tortorelli et al. 2020)
USFS FIA	3	11	2016	35	Ocular circular subplot	National forest monitoring program administered by the USDA Forest Service. (https://www.fs.usda.gov/pnw/page/pnw-fia-inventory-data)
FSGeodata Clearinghouse	5	4	2017/2018	3 - 23	Ocular within mapped polygon	USDA Forest Service Current Invasive Plants database. (https://data.fs.usda.gov/geodata/edw/datasets.php?xmlKeyword=current+invasive+plants)

2.2.2 Phenology from satellite time series

Image fusion can be used to fill missing data in earth observation datasets like Landsat (Belgiu and Stein 2019), thus allowing for higher resolution estimates of LSP (Gao et al., 2017). A time series of normalized difference vegetation index (NDVI) from Landsat 8 and images synthesized through image fusion of Landsat 8 (30 m resolution) and MODIS (500 m resolution) were used to estimate LSP for 2017 (Nietupski et al. 2021). Below, we briefly summarize the image fusion and LSP methods, but specific details are available in Nietupski et al. (2021). The image fusion process was implemented in Google Earth Engine, and synthesized images were evaluated using true Landsat 8 imagery. Fusion results indicated high correspondence (i.e., high correlation and low error) with Landsat imagery and were comparable to another image fusion algorithm. The Landsat 8 and fused images in combination resulted in a near-daily observation interval at 30 m resolution. The time series was smoothed and annual phenometrics were extracted using a double logistic function. Phenometrics represent various aspects of the annual cycle, including the day-of-year (DOY) associated with transitions to and from a dormant state, the NDVI related to different transition dates, and the rate of change at various transition points. The smoothing and extraction process resulted in 11 phenometrics characterizing the seasonal pattern in reflectance. LSP was evaluated using near-surface cameras and field observed phenology from locations within the BME. Phenometrics showed reasonable performance in detecting phenological phases of shrubland, grassland, and open-forest sites.

Cumulative growing degree days (cGDD; starting from January 1) were also calculated as an alternative to DOY units for five phenological transition dates (start-of-green-up, start-of-season, maturity, end-of-season, dormancy) using the Parameter-elevation Relationships on Independent Slopes Model (PRISM) daily gridded estimates of historical temperature at 4 km resolution (PRISM Climate Group, 2012). These metrics account for temperature-driven plant development and may have lower

spatial and interannual variability than DOY metrics (Russelle et al. 1984). Daily minimum and maximum temperatures were used to calculate GDD following

$$GDD = \begin{cases} (7 - T_{base}), & if\ (T_{max} + T_{min})/2 < 7 \\ (30 - T_{base}), & if\ (T_{max} + T_{min})/2 > 30 \\ \dfrac{(T_{max} + T_{min})}{2} - T_{base}, & otherwise \end{cases} \quad (1)$$

where T_{max} is the maximum daily temperature in Celsius, T_{min} is the minimum daily temperature in Celsius, and T_{base} is the base temperature for ventenata in Celsius. The T_{base} (i.e, minimum temperature threshold) for ventenata is 7°C (Wallace et al., 2015). Mean daily temperatures outside a ventenata-specific range were censored because plant development is unlikely under these conditions (McMaster and Wilhelm 1997). The maximum temperature threshold of cheatgrass (30°C; Thill et al. 1984) was used to approximate ventenata's upper threshold (Wallace et al., 2015).

2.2.3 Biophysical

Long-term climatic conditions were determined from the 800 m PRISM 30-year annual and monthly climate normal data (Norm81; PRISM Climate Group 2012). Seasonal climate conditions were calculated from monthly data, where seasonal means were weighted by the days per month. The seasons were defined as spring (April-June), summer (July-September), fall (October-December), and winter (January-March). All gridded climate data, including cGDD, were resampled to 30m using bilinear interpolation for model prediction.

Soil texture in the top 20 cm of the soil profile was derived from the 30 m gridded Soil Survey Geographic Database (Soil Survey Staff 2015). The missing data was filled with Soil Resource Inventory data from the USFS. Only the top layer of the soil profile was included due to its importance for shallow-rooted species. Terrain attributes, including slope, eastness, and northness, were calculated from the 30 m National Elevation Dataset (NED; Gesch et al., 2018).

2.3 Statistical analysis

2.3.1 Model development

Models differed by their inclusion of different predictor groups and were split as follows: 1) hybrid (48 predictors), 2) bioclimatic (32 predictors), and 3) phenology (16 predictors) (Appendix A: Table A1). Supervised classification of heavily invaded areas was performed using the random forest (RF) algorithm, which is known for achieving high prediction accuracy while accommodating missing data and correlated predictors (Cutler et al. 2007; Belgiu and Drăgu 2016). The RF classifier uses an ensemble of decision trees (CART) to form a prediction through a majority voting process (Breiman 2001). We implemented the RF algorithm through the Python package XGBoost (Chen and Guestrin 2016). Tuning parameters were determined through a cross-validated grid search to maximize discriminatory capacity and performance with an unbalanced dataset while reducing the chance of overfitting. A few parameters must be set to fixed values to fit a RF style model in XGBoost. Other parameters may be adjusted to control overfitting or improve model performance for low-prevalence data (Table 2). Parameters were adjusted and determined in sequence and included the total number of trees (number estimators), the maximum depth of the trees (maximum depth), the minimum number of observations per node (minimum child weight), the minimum loss reduction required to make another partition (gamma), and the weight given to classes (scale positive weight).

Table 2. Random forest model parameters determined through cross-validated grid search.

Hyperparameter	**Model**		
	Hybrid	Bioclimatic	Phenology
Number Estimators	100	500	900
Max Depth	7	3	6
Min Child Weight	3	1	4
Gamma	0.1	0.4	0.1
Scale Positive Weight	9	11	13

2.3.2 Model evaluation

Overall model performance was evaluated using 10-fold cross-validation. While model validation using an independent data set would have provide additional insight into model performance, a cross-validation approach was used due to the small sample size of the positive class. Threshold independent and dependent evaluation metrics were selected to characterize the model's discriminatory power (Fielding 2002), including the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity (Cohen 1960; Fielding and Bell 1997). In addition to these metrics, variable importance and partial dependence plots were examined to understand the contribution of each type of predictor to the model and the marginal relationship between the highest ranked predictors and ventenata probability.

While classification model predictions are probabilities, their performance metrics are derived from a confusion matrix that describes the relationship between the observed and predicted classes. Conversion of the probability values to classes for the confusion matrix requires a probability threshold. For threshold-dependent metrics like accuracy (Equation 3), sensitivity (Equation 4), and specificity (Equation 5), the threshold is chosen by the user. For threshold-independent metrics like AUC (Equation 2), many thresholds are used to calculate a curve that relates the true positive rate (TPR) to the false positive rate (FPR).

Observed	**Predicted**	
	Positive	Negative
Positive (P)	True Positive (T_p)	False Negative (F_n)
Negative (N)	False Positive (F_p)	True Negative (T_n)

$$AUC = \sum_{i=1}^{k} \left(\left(TPR_{i-1}(FPR_i - FPR_{i-1}) \right) + \frac{1}{2}(TPR_i - TPR_{i-1})(FPR_i - FPR_{i-1}) \right) \quad (2)$$

, where $TPR = \frac{T_p}{P}$ and $FPR = \frac{F_p}{N}$.

$$Accuracy = \frac{T_p + T_n}{P + N} \quad (3)$$

$$Sensitivity = \frac{T_p}{P} \quad (4)$$

$$Specificity = \frac{T_n}{N} \quad (5)$$

Determination of an appropriate probability threshold to distinguish ventenata populations above (positive class) or below 20 % cover (negative class) was required to assess model performance using threshold-dependent metrics. We performed cross-validation with the parameters determined by the grid search process to calculate the threshold for each model. Data were randomly partitioned into training and testing sets 1,000 times while preserving the prevalence of the original data. For each of these partitions, a model was fit to the training data (80 %) and evaluated with the testing data (20 %). The kappa statistic was calculated for each iteration at 0.01 increments across the probability range, and the median threshold that maximized kappa was retained to distinguish presence and absence (Freeman and Moisen 2008). This process resulted in thresholds of 0.58, 0.73, and 0.7 for the hybrid, bioclimatic, and phenology models, respectively.

Elevation and vegetation classes were used to investigate patterns in the predictive performance that overall performance measures might not capture. The regional elevation range was evenly split into eight classes between 610 and 1,829 m, with one additional class representing

elevations below 610. Vegetation classes were determined from a potential natural vegetation (PNV) map developed by the USFS (Simpson et al. 2019). Sample sizes in each elevation or vegetation class were not adequate to fit separate models for each class, so classes were associated with each record and evaluation metrics were calculated from cross-validation datasets (1,000 random 80-20 partitions). Model performance was evaluated in each class by aggregating metrics from the cross-validation datasets.

3. Results

3.1 Model evaluation

The hybrid model achieved the highest AUC and sensitivity but similar overall accuracy and specificity (Table 3). The models differed most substantially in sensitivity, where omission error was 23 % and 19 % higher in bioclimatic and phenology models. A confusion matrix calculated from the 1000 cross-validation datasets illustrates that the three models perform similarly for a given probability threshold of the positive class (heavily invaded areas) but substantially different for the negative class (Figure 3). The bioclimatic and phenology models required a higher probability threshold to achieve the best balance between performance in predicting ventenata populations above or below 20 % cover (i.e., to maximize kappa).

Figure 3. Confusion matrix showing the distribution of predicted vs. observed positive and negative classes in the three models. The positive class represents ventenata populations of 20 % or greater cover and the negative class represents populations with less than 20 % cover. The upper and lower bounds of each box plot show the 25th and 75th quantiles, the whiskers show the 5th and 95th quantiles, and the mid line shows the median.

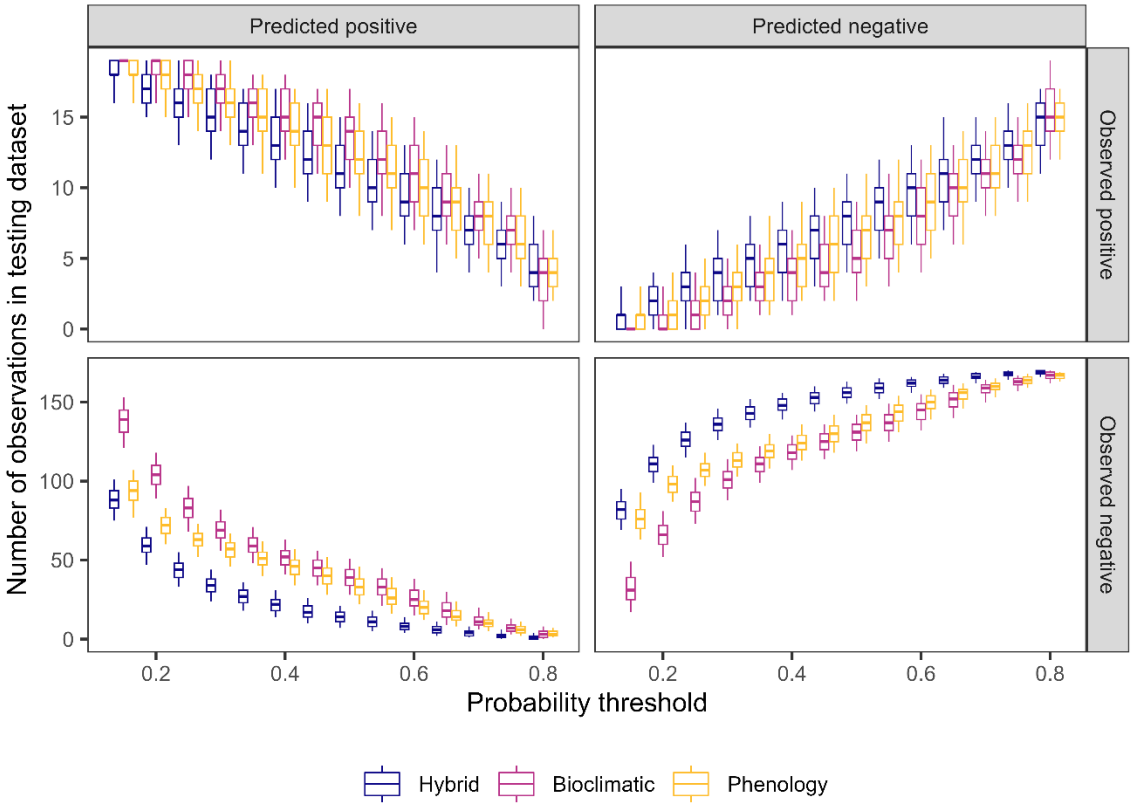

Table 3. Median and interquartile range of evaluation metrics showing the overall model performance for each of the three models. Threshold-dependent metrics (accuracy, sensitivity, specificity) were calculated based on thresholds of 0.58 (hybrid), 0.73 (bioclimatic), and 0.7 (phenology).

Model	AUC	Accuracy	Sensitivity	Specificity
Hybrid	0.88 (0.044)	0.90 (0.021)	0.56 (0.275)	0.95 (0.032)
Bioclimatic	0.85 (0.041)	0.89 (0.038)	0.33 (0.347)	0.96 (0.041)
Phenology	0.84 (0.089)	0.88 (0.047)	0.37 (0.167)	0.94 (0.05)

Evaluation metrics within each elevation and vegetation class showed comparable results to the overall model performance. However, evaluating performance by elevation revealed differences not represented in the overall model performance (Table 4). Notably, while the phenology model's performance is comparable to the hybrid and bioclimatic models in the 914 to 1,676 m elevation range, its performance decreases below the 762 - 914 m elevation class. When evaluated by vegetation class,

the hybrid model generally showed equivalent or better performance than the bioclimatic and phenology models (Appendix A: Table A2). The hybrid model had higher accuracy than the other two models in Scabland Shrub and higher sensitivity across vegetation types, with noteworthy differences occurring in Scabland Shrub, Upland Shrub, and Juniper Woodlands. However, in the Juniper Steppe vegetation class, the hybrid model had lower performance both in terms of sensitivity and specificity.

Table 4. Summary of model performance by elevation class (median and interquartile range). Models are abbreviated as HY (hybrid), BIO (bioclimatic), and PHEN (phenology).

Elevation Class (m)	n	Prevalence	Accuracy			Sensitivity			Specificity		
			HY	BIO	PHEN	HY	BIO	PHEN	HY	BIO	PHEN
< 610	2	0	1 (0)	1 (0)	0 (1)	-	-	-	1 (0)	1 (0)	0 (1)
610-762	7	0	1 (0)	1 (0)	1 (0.5)	-	-	-	1 (0)	1 (0)	1 (0.5)
762-914	22	0	1 (0)	1 (0)	1 (0)	-	-	-	1 (0)	1 (0)	1 (0)
914-1067	36	0.19	0.71 (0.2)	0.71 (0.23)	0.71 (0.23)	0 (0)	0 (0.33)	0 (0)	0.89 (0.25)	0.86 (0.25)	0.9 (0.2)
1067-1219	45	0.09	0.9 (0.18)	0.78 (0.21)	0.83 (0.16)	1 (1)	0.5 (1)	0 (0.58)	1 (0.14)	0.82 (0.2)	0.89 (0.18)
1219-1371	131	0.15	0.84 (0.09)	0.77 (0.1)	0.82 (0.09)	0.67 (0.3)	0.5 (0.33)	0.5 (0.33)	0.88 (0.1)	0.82 (0.11)	0.89 (0.09)
1372-1524	238	0.22	0.83 (0.07)	0.83 (0.06)	0.8 (0.07)	0.58 (0.2)	0.45 (0.19)	0.43 (0.21)	0.9 (0.08)	0.94 (0.06)	0.91 (0.06)
1524-1676	307	0.03	0.96 (0.03)	0.97 (0.03)	0.97 (0.03)	0 (0.33)	0 (0)	0.4 (0.5)	0.98 (0.03)	1 (0)	0.98 (0.03)
1676-1829	156	0	1 (0)	1 (0)	1 (0)	-	-	-	1 (0)	1 (0)	1 (0)

Predictors shared among the three models had similar importance (Appendix A: Figure A1-A3). The top ten predictors from the hybrid model were split between phenology and climate. The top few climate predictors were similar between the hybrid and bioclimatic models, but the soil predictors were ranked higher in the bioclimatic model. The highest-ranking phenology predictors in the hybrid and phenology models were minimum NDVI (NDVI Min), DOY at the start-of-green-up (SOG), cumulative growing degree days at the end-of-season (cGDD at EOS), and DOY at the end-of-season (EOS). The highest-ranking climate predictors in the hybrid and bioclimatic models included average winter maximum temperature (Wint Tmax), average fall maximum temperature (Fall Tmax), average total winter precipitation (Wint Ppt), and average winter maximum vapor pressure deficit (Wint MaxVPD). The hybrid model indicates that heavily invaded areas were associated with lower baseline NDVI values, earlier start-of-green-up, and lower cumulative growing degree days at the end-of-season (Figure 4). Heavily invaded areas were also associated with higher fall and winter maximum temperatures and relatively moderate winter precipitation.

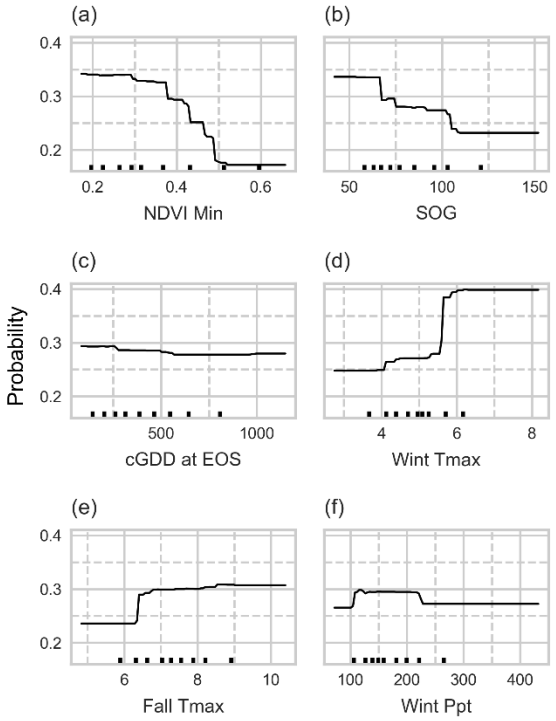

Figure 4. The relationship between the probability of heavily invaded areas and the top predictors from the hybrid model. Tick marks in the rug plot along the x-axis indicate deciles. Predictors include a) minimum NDVI (NDVI), b) start-of-green-up (day-of-year), c) cumulative growing degree days at the end-of-season (°C), d) average winter maximum temperature (°C), e) average fall maximum temperature (°C), and f) average total winter precipitation (mm).

3.2 Ventenata's distribution

The hybrid model predicts heavy invasion is present in 5,454 km^2 of the BME (7.7 %) while the bioclimatic and phenology models predict 5,986 km^2 (8.4 %) and 7,441 km^2 (10.5 %), respectively (Figure 5a-c). There are a few notable spatial differences in the predictions among the three models, especially in the areas above the calculated probability threshold (Figure 5, areas shown in yellow to red). A comparison between the hybrid and bioclimatic models shows that the hybrid model frequently predicted heavily invaded areas in a subset of the area predicted by the bioclimatic model. However,

some hybrid model predictions occurred outside the range of the bioclimatic model (e.g., heavily invaded areas in the northwest but not the northeast). To a lesser extent, the hybrid model also predicted the presence in a subset of the area predicted by the phenology model. In this case, the phenology model predicts presence broadly, including lower elevations and substantial parts of the southern and western parts of the BME.

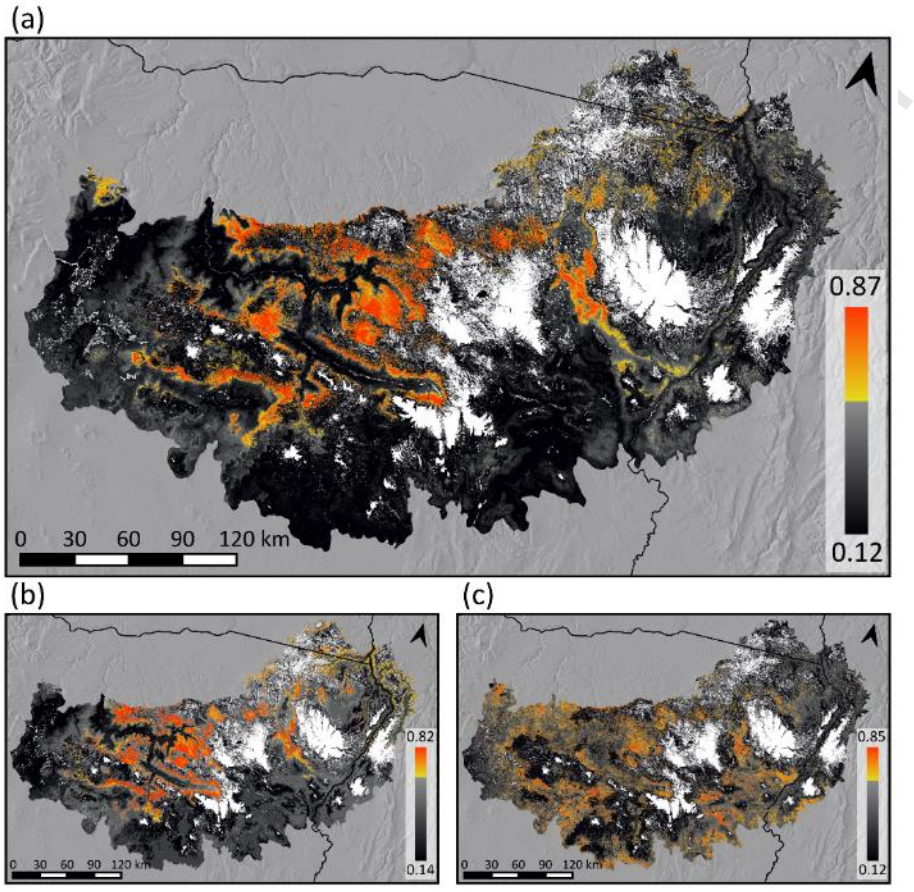

Figure 5. Predictions from a) the hybrid model, b) the bioclimatic model, and c) the phenology model. The color ramp is split at the threshold that maximizes kappa (hybrid=0.58, bioclimatic=0.73, phenology=0.7), where areas above the threshold (heavily invaded) are yellow to red and areas below the threshold are grey to black. Masked areas (white) are presumed to be currently uninvadable due to high elevation, high forest canopy cover, or perennial water.

Heavily invaded areas predicted by the hybrid model primarily occurred within 946 and 1,451 m (90 % interval). While the elevational range for the bioclimatic model was comparable to the hybrid model, the lower bound of the phenology model was lower in elevation (753 to 1,495 m). The hybrid model predicted that approximately 73 % of heavily invaded areas were found within Upland Shrub, Juniper Woodland, Xeric Pine, and Dry Ponderosa Pine vegetation types (Figure 6).

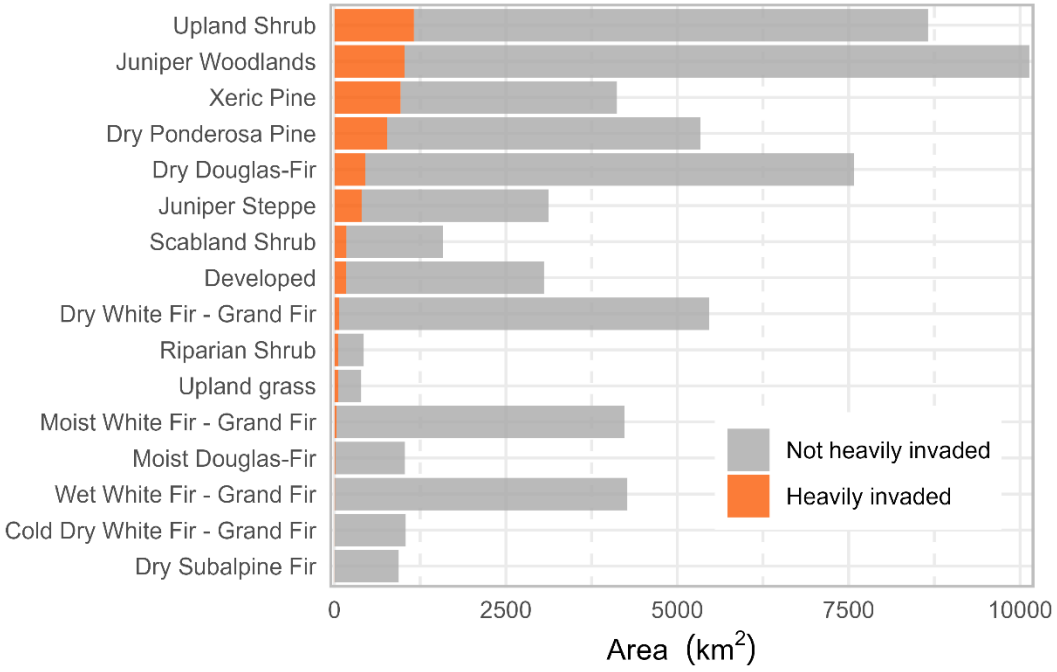

Figure 6. Potential natural vegetation classes that cover greater than 0.5 % of the Blue Mountains Ecoregion and have greater than 0.1 % heavy ventenata invasion. The area heavily invaded (orange) and not heavily invaded area (grey) by ventenata with the percent of each vegetation type heavily invaded as an inset. Vegetation types are presented in descending order by the area heavily invaded.

The probability difference between the hybrid and the other two models also helped illustrate the effect of including both predictor sets (Figure 7a). The bioclimatic model generally predicted higher probability across the BME, especially in areas with high conifer canopy cover (see Figure 2 for forested areas). In contrast, some forest openings had higher probability in the hybrid model. In comparison to

the phenology model, the hybrid model predicted higher probability across the higher elevations and lower probability across lower elevations. The greatest difference between the hybrid and phenology models occurred along the southern and northwestern parts of the BME. The spatial patterns of negative difference (phenology greater than hybrid) between these models closely align with the area predicted as cheatgrass by another recent phenology modeling effort (Figure 7b).

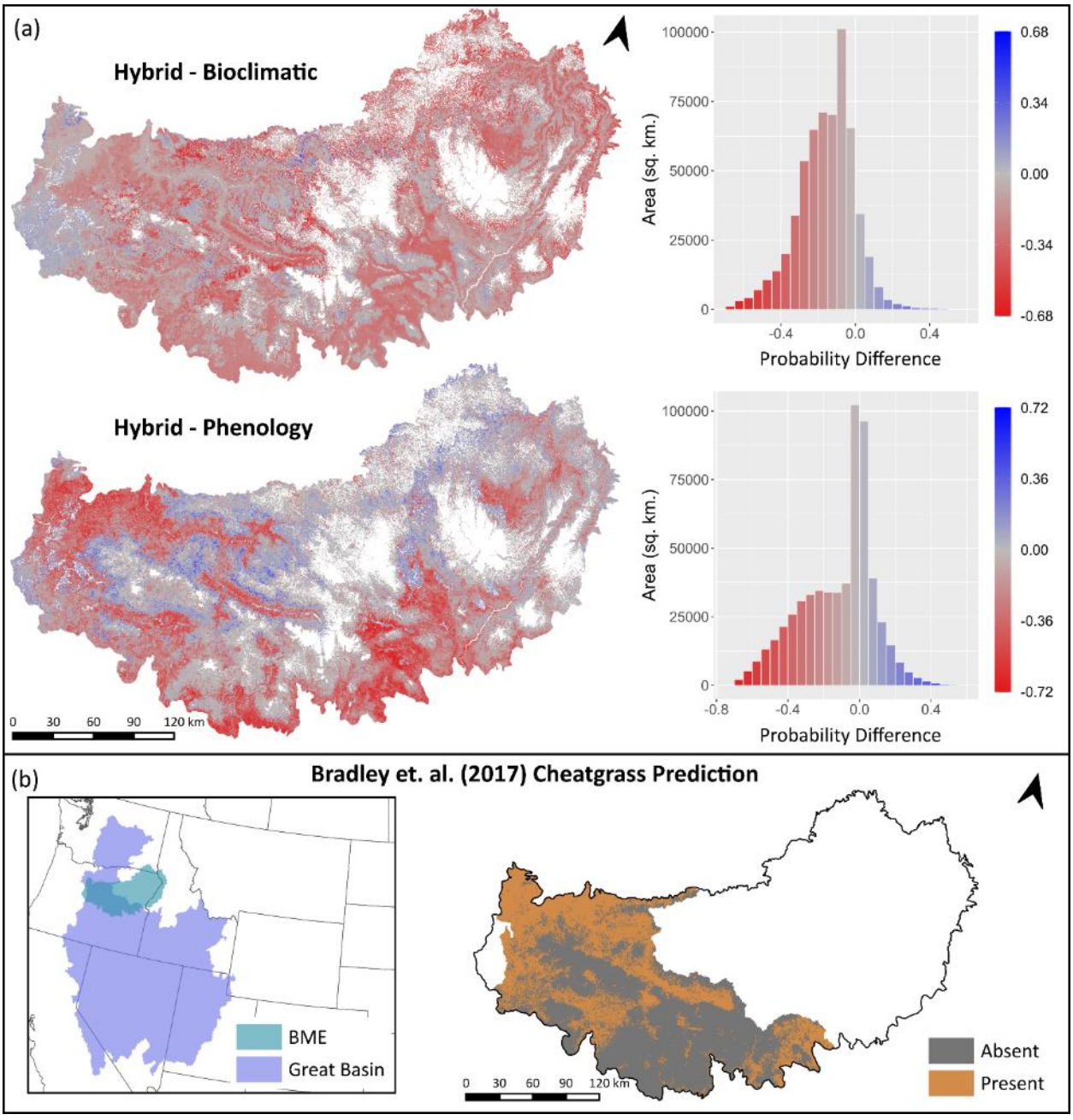

Figure 7. Ventenata model differences and cheatgrass predictions where a) shows the difference in ventenata model predictions and b) shows predicted as cheatgrass (> 15 % cover) within a subset of the BME (Bradley et al. 2017). In panel a, blue indicates that the hybrid model predicted a higher probability, while red indicates that the bioclimatic or phenology model predicted a higher probability.

4. Discussion

We demonstrate that including biophysical predictors can help to distinguish populations of some annual grass species when land surface phenology (LSP) cannot resolve species-level differences. Species-level detection using remotely sensed data is advancing along with sensor technologies and image processing techniques that allow for greater fidelity in land surface observations (Pastick et al. 2020; Weisberg et al. 2021). However, mapping annual grasses over large regions presents a challenge when using coarse-resolution satellite imagery, especially as new invaders with similar phenological characteristics are introduced. Yet, information describing the characteristics of individual species geographic distribution is critical for the management and policy decision-making process (Funk et al. 2020). To span the current knowledge gap for a new invader, we produced the first ecoregional estimate of heavily invaded areas by leveraging both phenology and biophysical predictors in a single hybrid model.

Our use of a 30 m LSP product (Nietupski et al. 2021) allowed for ventenata population detection in the small forest openings (>= 30 m) but came with complications. Increased LSP spatial resolution can be advantageous because the dominant vegetation across large areas can drive LSP in lower-resolution imagery (e.g., MODIS). In such cases, increased resolution may help detect smaller populations, which we observed in dry forests of the BME. However, higher resolution LSP can negatively impact vegetation class separation because LSP variation increases with decreasing spatial grain (Weisberg et al. 2021). Phenological variation can also arise from temperature, moisture, edaphic,

or genetic factors (Piao et al. 2006; Luquez et al. 2008; Andrew and Ustin 2009). Therefore, we attempted to account for environmentally derived variation by calculating phenometrics by accumulated heat (i.e., cGDD) instead of calendar days. Temperature-driven plant development is a better predictor of growth than calendar days (Cudney et al. 1989; Ball et al. 2004) but we found that only one of these metrics (cGDD at EOS) was important to the model. Limitations to the spatial grain of the temperature data may partially explain the limited explanatory power of the cGDD predictors. However, cGDD at EOS indicated that lower accumulated heat units (i.e., earlier in the season) were associated with ventenata's senescence, a characteristic of winter annual grasses including ventenata (Wallace et al. 2015).

The use of phenology and niche-related predictors together increased model performance and, specifically, the hybrid model's ability to classify locations without heavy ventenata invasion. At lower elevations, environmental predictors reduced probability where phenology attributes matched ventenata's, but the habitat was unsuitable. In mid-elevations, phenology predictors decreased probability where forest canopy cover was high and increased probability where the start-of-season was early, a key indicator of winter annual grasses (Bradley et al. 2012). Thus, it is not surprising these issues were evident in the phenology and bioclimatic model outputs. The phenology model had reduced performance at lower elevations where the co-occurrence of other invasive annual grasses is common (Mack 1981; Young 1992), and the bioclimatic model had reduced performance in mid-elevations (e.g., 1067 – 1371 m) where the vegetation varied by tree canopy cover.

Lower performance and higher predictions in low and mid-elevations for the phenology and bioclimatic models substantially impacted the predicted distribution. Phenology and bioclimatic model limitations in classifying heavily invaded areas resulted in less conservative predictions of the total area invaded (i.e., 7.7% vs. 8.4% and 10.5% of the BME). In comparison to the hybrid model, the phenology model predicted more populations at lower elevations, while the bioclimatic model predicted more

populations in mid-elevations. Consequently, the phenology and bioclimatic models are likely overpredicting the distribution of heavily invaded areas and overprediction can lead to poor management resource allocation (Guisan et al. 2013). Overprediction from the bioclimatic model is not entirely negative. Biophysical predictors provided useful information about ventenata's niche and the location of suitable habitat. Erroneous predictions in locations with high tree cover or other attributes that prevent invasion may indicate habitat conditions that are susceptible to future invasion. However, to better understand the true range of habitat suitable to ventenata, the data used for the model would need to include all true presence and absence data, as opposed to observations above or below a 20 % cover threshold.

The performance characteristics of the hybrid model make it better suited to some applications than others. Invasive species presence maps can be used in early detection and eradication efforts to refine the search area for new populations (Freeman and Moisen 2008). For this type of application, a model with high sensitivity would be preferred because missing true populations may have a negative impact on control efforts. All three models in this study generally performed better in specificity than sensitivity. The high specificity of the three models is partially related to the small sample size of heavily invaded areas and is not uncommon when modeling rare, low prevalence species (Cheney et al. 2018). While future iterations of this model would benefit from larger sample sizes of ventenata, the hybrid model's high specificity does benefit certain management applications. A map with high specificity gives the user high confidence in areas where the species is not present. In the context of fire management, it might be equally important to identify locations where fine fuel conditions have not changed as it is to identify locations where fuel conditions have changed. Considering that the hybrid model has the best sensitivity for a given level of specificity, we can have the greatest confidence in the output produced by this model. Thus, for the fire management challenges in the BME, the hybrid model provides better

information about where fuel conditions may have changed because of the grass invasion, knowing that other locations are fairly certain to not have heavy ventenata invasion.

The niche of heavily invaded areas was most closely associated with winter and fall climate conditions, as expected of a winter annual grass. Fall and winter temperature and precipitation are crucial for winter annual grasses because germination and initial leaf elongation occur during this period (Meyer et al. 1997). While maximum winter temperature is also an important predictor of cheatgrass, winter precipitation has been one of the least important cheatgrass predictors in other studies (Bradley 2009). The importance of winter precipitation in our models may indicate that ventenata have higher moisture requirements than cheatgrass during the germination and initial growth period. Ventenata's preference for wetter conditions has also been documented in the field (Jones et al. 2018) and experimentally, where surface moisture from thatch increased local soil moisture, enhancing seedling emergence (Wallace et al. 2015).

Our comparative illustration of ventenata and cheatgrass predictions (Figure 7a-b) also suggests that the coarse-resolution remotely sensed phenology attributes used in this study cannot differentiate heavy ventenata and cheatgrass invasion . The greatest difference between the hybrid and phenology models overlapped with much of the area mapped as cheatgrass. The patterns observed in this comparison suggest that including climatic predictors helped differentiate location with heavy ventenata and cheatgrass invasion. Reliance on satellite phenology in previous studies may mean that parts of the predicted cheatgrass range are occupied by ventenata or other species (Figure 7b). Interestingly, we found areas where the phenology model predicted high probability, but neither the hybrid nor the cheatgrass model predicted grass populations. This could indicate that the phenology model detects missed cheatgrass populations or the presence of other invasive annual grasses such as medusahead. Furthermore, cheatgrass, ventenata, and other nonnative annual grasses exist side-by-side in some locations, but it is pertinent to consider that the cheatgrass and ventenata models compared here were

developed with relatively high cover thresholds (i.e., 15 % cover of cheatgrass and 20 % cover of ventenata) and the extent to which these species coexist at high cover is not well established.

The hybrid model indicates that 8 % of the Blue Mountains Ecoregion (BME) may already be heavily invaded by ventenata. While this is a relatively small fraction of the ecoregion, heavily invaded areas are disproportionately represented in certain plant communities. Heavily invaded areas were most prevalent in Upland Shrub and Juniper Woodland vegetation types. These vegetation types have been reported as being actively invaded (Jones et al. 2018; Nicolli et al. 2020) and are closer to the agricultural lands where the first documented impacts of ventenata were reported (Fountain 2011; Wallace et al. 2015). Within these shrubland and woodland vegetation types, ventenata populations were located close to ecotones between shrub- and forest-dominated plant communities. Additionally, ventenata occupied a relatively high portion of the Dry Ponderosa Pine, Upland Grassland, Scabland Shrub, and Riparian Shrub vegetation types. Ventenata has the greatest impact within the Xeric Pine type, with almost a quarter of this type occupied by ventenata. Within forested areas, heavily invaded areas were located within many openings between patches of higher forest canopy cover. These observations coincide with the growing body of literature that indicates ventenata's ability to occupy higher elevations than cheatgrass and medusahead (Tortorelli et al. 2020; Applestein and Germino 2021).

Annual grass invasion in forested ecosystems has recently drawn greater attention (Flory et al. 2015; Kerns et al. 2020; Reilly et al. 2020). Invasive grass influence on fire regimes in shrublands of the western US is well documented (Brooks et al. 2004; Balch et al. 2013), but the modeled distribution of heavy ventenata invasion in and around forested areas highlights the potentially transformative impact of this grass invasion on forests. While many native plant communities found at higher elevation can respond quickly to fire disturbances and are thought to be more resistant to annual grass invasion, research now suggests that these plant communities do not have the same resistance to ventenata invasion (Tortorelli et al. 2020; Applestein and Germino 2021). Further, fire simulation modeling

suggests that ventenata may influence fire behavior within invaded patches and the transmission of fire to surrounding forests (Tortorelli 2023). If a fire does occur in and transmit through invaded areas, it may further facilitate the spread of ventenata as canopy cover reduction from tree mortality can increase available habitat. The ventenata invasion and related fire effects could potentially influence future ecosystem composition and structure, compounding the forecasted impacts of climate change.

Results from this study inform and build upon our understanding of invasive annual grass mapping and the ecological understanding that resulting maps can provide. Satellite image time series helped to describe the characteristics of the annual grass development cycle, while image fusion techniques improved the observation frequency and consistency at 30 m. As a result, LSP derived from the fused image data in combination with biophysical features allowed for the detection of ventenata in relatively small forest openings. Model output illuminated important aspects of ventenata's invasion, including some environmental requirements, associations with vegetation types, and the location of potentially problematic populations. This information can be used to help locate previously unknown heavily invaded areas, inform the placement of fuel breaks in wildfire control efforts, identify potential propagule sources for the future spread of ventenata, direct the application of control methods in high value areas (Denchev and Denchev 2018; Davies and Hamerlynck 2019; Koby et al. 2019; Lamm et al. 2021), or target field monitoring for at-risk sites and vegetation types, particularly in association with proposed vegetation management.

Data availability

The climate data used in this study are available at https://prism.oregonstate.edu/normals. The gridded soil data are available at https://data.nal.usda.gov/dataset/gridded-soil-survey-geographic-database-gssurgo. The elevation data are available at https://www.usgs.gov/the-national-map-data-delivery. Predictions of ventenata's distribution are available in the USDA Forest Service Research Data Archive at

https://doi.org/10.2737/RDS-2022-0011. All other data generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

CRediT author statement

Ty Nietupski: Conceptualization, Methodology, Sofware, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing – Original Draft, Visualization, Funding Acquisition. **Becky Kerns:** Conceptualization, Resources, Writing – Review & Editing, Supervision, Project Administration, Funding Acquisition. **Temesgen Hailemariam:** Writing – Review & Editing, Supervision, Funding Acquisition.

Acknowledgments

This research was largely supported by the Joint Fire Science Program (Proposal ID: 16-1-596 01-21) and a Google Cloud Platform Research Credits grant. This research was also supported in part by an appointment to the United States Forest Service Research Participation Program administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Department of Agriculture. We would like to thank those who contributed data to our ventenata database, including numerous members of the Oregon State University community and employees of the USDA Forest Service. We also thank Robert Kennedy for his contributions to the conceptualization and review of this research. We would like to express gratitude to Bridgett Naylor, Kent Coe, and Michelle Day for field data collection and management support. We would like to thank Claire Tortorelli her review of an early version of this manuscript. We also received helpful feedback during review that strengthened the manuscript and would like to thank the anonymous reviewers for their time and thoughtful suggestions.

References

Agee JK (1994) Fire and weather disturbances in terrestrial ecosystems of the eastern Cascades. Gen. Tech. Rep. PNW-GTR-320. USDA-Forest Service, Pacific Northwest Research Station, Portland, OR

Ahmed N, Atzberger C, Zewdie W (2020) Integration of remote sensing and bioclimatic data for prediction of invasive species distribution in data-poor regions: a review on challenges and opportunities. Environ Syst Res 9:. https://doi.org/10.1186/s40068-020-00195-0

Andrew ME, Ustin SL (2009) Effects of microtopography and hydrology on phenology of an invasive herb. Ecography 32:860–870. https://www.jstor.org/stable/20696295

Applestein C, Germino MJ (2021) Patterns of post-fire invasion of semiarid shrub-steppe reveals a diversity of invasion niches within an exotic annual grass community. Biol Invasions. https://doi.org/10.1007/s10530-021-02669-3

Austin M (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. Ecol Modell 200:1–19. https://doi.org/10.1016/j.ecolmodel.2006.07.005

Averett JP, Morris LR, Naylor BJ, et al (2020) Vegetation change over seven years in the largest protected Pacific Northwest Bunchgrass Prairie remnant. PLoS One 15:1. https://doi.org/10.1371/journal.pone.0227337

Balch JK, Bradley BA, D'Antonio CM, Gómez-Dans J (2013) Introduced annual grass increases regional fire activity across the arid western USA (1980-2009). Glob Chang Biol 19:173–183. https://doi.org/10.1111/gcb.12046

Ball DA, Frost SM, Gitelman AI (2004) Predicting timing of downy brome (*Bromus tectorum*) seed production using growing degree days. Weed Sci 52:518–524. https://doi.org/10.1614/WS-03-067

Beatley JC (1966) Ecological status of introduced brome grasses (*Bromus Spp.*) in desert vegetation of

Southern Nevada. Ecology 47:548–554. https://doi.org/10.2307/1933931

Belgiu M, Drăgu L (2016) Random forest in remote sensing: A review of applications and future directions. ISPRS J Photogramm Remote Sens 114:24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Belgiu M, Stein A (2019) Spatiotemporal image fusion in remote sensing. Remote Sens 11:818. https://doi.org/10.3390/rs11070818

Boyte SP, Wylie BK (2016) Near-Real-Time Cheatgrass Percent Cover in the Northern Great Basin, USA, 2015. Rangelands 38:278–284. https://doi.org/10.1016/j.rala.2016.08.002

Bonham CD (2013) Measurements for terrestrial vegetation. Hoboken, NJ: Wiley 264 p. https://doi.org/10.1002/9781118534540

Bradley BA (2014) Remote detection of invasive plants: A review of spectral, textural and phenological approaches. Biol Invasions 16:1411–1425. https://doi.org/10.1007/s10530-013-0578-9

Bradley BA (2009) Regional analysis of the impacts of climate change on cheatgrass invasion shows potential risk and opportunity. Glob Chang Biol 15:196–208. https://doi.org/10.1111/j.1365-2486.2008.01709.x

Bradley BA, Curtis CA, Fusco EJ, et al (2017) Cheatgrass (*Bromus tectorum*) distribution in the intermountain Western United States and its relationship to fire frequency, seasonality, and ignitions. Biol Invasions 20:1493–1506. https://doi.org/10.1007/s10530-017-1641-8

Bradley BA, Mustard JF (2008) Comparison of phenology trends by land cover class: a case study in the Great Basin , USA. Glob Chang Biol 14:334–346. https://doi.org/10.1111/j.1365-2486.2007.01479.x

Bradley BA, Olsson AD, Wang O, et al (2012) Species detection vs. habitat suitability: Are we biasing

habitat suitability models with remotely sensed data? Ecol Modell 244:57–64. https://doi.org/10.1016/j.ecolmodel.2012.06.019

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Brooks ML, D'Antonio CM, Richardson DM, et al (2004) Effects of invasive alien plants on fire regimes. Bioscience 54:677–688. https://doi.org/10.1641/0006-3568(2004)054[0677:eoiapo]2.0.co;2

Chambers JC, Miller RF, Board DI, et al (2014) Resilience and resistance of sagebrush ecosystems: Implications for state and transition models and management treatments. Rangel Ecol Manag 67:440–454. https://doi.org/10.2111/REM-D-13-00074.1

Chambers JC, Roundy BA, Blank RR, et al (2007) What makes Great Basin sagebrush ecosystems invasible by *Bromus tectorum*? Ecol Monogr 77:117–145. https://doi.org/10.1890/05-1991

Chen T, Guestrin C (2016) XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, pp 785–794. https://doi.org/10.1145/2939672.2939785

Cheney C, Esler KJ, Foxcroft LC, et al (2018) The impact of data precision on the effectiveness of alien plant control programmes: a case study from a protected area. Biol Invasions 20:3227–3243. https://doi.org/10.1007/s10530-018-1770-8

Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas 37–46. https://doi.org/10.1177/001316446002000104

Cudney DW, Jordan LS, Corbett CJ, Bendixen WE (1989) Developmental rates of wild oats (*Avena fatua*) and wheat (*Triticum aestivum*). Weed Sci 37:521–524. https://doi.org/10.1017/S0043174500072349

Cutler DR, Edwards TC, Beard KH, et al (2007) Random forests for classification in ecology. Ecology 88:2783–2792. https://doi.org/10.1890/07-0539.1

D'Antonio C, Vitousek P (1992) Biological invasions by exotic grasses, the grass/fire cycle, and global change. Annu Rev Ecol Syst 23:63–87. https://doi.org/10.1146/annurev.es.23.110192.000431.

Davies KW, Hamerlynck E (2019) Ventenata and other coexisting exotic annual grass control and plant community response to increasing Imazapic application rates. Rangel Ecol Manag 72:700–705. https://doi.org/10.1016/j.rama.2019.02.010

Denchev TT, Denchev CM (2018) Two new smut fungi on Ventenata (Poaceae): *Tilletia elizabethae* from Slovakia and *T. ventenatae* from Turkey. Willdenowia 48:177–183. https://doi.org/10.3372/wi.48.48201

De Stephano A, Fowers B, Mealor BA (2021) Comparison of visual estimation and line-point intercept vegetation survey methods on annual grass–invaded rangelands of Wyoming. Inv Plant Sci and Manag 14:240-252. https://doi.org/10.1017/inp.2021.36

Dethier MN, Graham ES, Cohen S, et al (1993) Visual versus random-point percent cover estimations:'objective'is not always better. Mar Ecol Prog Ser 96:93-100. https://www.jstor.org/stable/24833571

Downing WM, Krawchuk MA, Meigs GW, et al (2019) Influence of fire refugia spatial pattern on post-fire forest recovery in Oregon's Blue Mountains. Landsc Ecol 34:771–792. https://doi.org/10.1007/s10980-019-00802-1

Elith J (2016) Predicting distributions of invasive species. In: Robinson AP, Walshe T, Burgman MA, Nunn M (ed) Invasive species: Risk assessment and management. Cambridge University Press, Cambridge, UK, pp 93–129

Endress BA, Averett JP, Naylor BJ, et al (2020) Non-native species threaten the biotic integrity of the largest remnant Pacific Northwest Bunchgrass prairie in the United States. Appl Veg Sci 23:53–68. https://doi.org/10.1111/avsc.12464

Fielding AH (2002) What are the appropriate characteristics of an accuracy measure? In: Scott JM, Heglund PJ, Morrison ML, et al. (eds) Predicting Species Occurences: Issues of Accuracy and Scale. Island Press, Covelo, CA, pp 271–280

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24:38–49. https://doi.org/10.1017/S0376892997000088

Flory SL, Clay K, Emery SM, et al (2015) Fire and non-native grass invasion interact to suppress tree regeneration in temperate deciduous forests. J Appl Ecol 52:992–1000. https://doi.org/10.1111/1365-2664.12437

Forsyth GG, Le Maitre DC, O'Farrell PJ, van Wilgen BW (2012) The prioritisation of invasive alien plant control projects using a multi-criteria decision model informed by stakeholder input and spatial data. J Environ Manage 103:51–57. https://doi.org/10.1016/j.jenvman.2012.01.034

Fountain B (2011) Producing timothy hay and managing for the impacts for Ventana. In: Proceedings of the Western Society of Weed Science. 64:107–108

Franklin J (2010) Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press, Cambridge, UK

Freeman EA, Moisen GG (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. Ecol Modell 217:48–58. https://doi.org/10.1016/j.ecolmodel.2008.05.015

Funk JL, Parker IM, Matzek V, et al (2020) Keys to enhancing the value of invasion ecology research for management. Biol Invasions 22:2431–2445. https://doi.org/10.1007/s10530-020-02267-9

Gao F, Anderson MC, Zhang X, et al (2017) Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. Remote Sens Environ 188:9–25. https://doi.org/10.1016/j.rse.2016.11.004

Gesch DB, Evans GA, Oimoen MJ, Arundel S (2002) The national elevation dataset. Photogramm Eng Remote Sens 68:5–11

Guisan A, Tingley R, Baumgartner JB, et al (2013) Predicting species distributions for conservation decisions. Ecol Lett 16:1424–1435. https://doi.org/10.1111/ele.12189

Hallmark B, Romero F (2015) Fuel treatment effectiveness on the Corner Creek Fire, Oregon. Boise, ID

He KS, Bradley BA, Cord AF, et al (2015) Will remote sensing shape the next generation of species distribution models? Remote Sens Ecol Conserv 1:4–18. https://doi.org/10.1002/rse2.7

Huang C, Asner GP (2009) Applications of remote sensing to alien invasive plant studies. Sensors 9:4869–4889. https://doi.org/10.3390/s90604869

Irisarri JGN, Oesterheld M, Verón SR, Paruelo JM (2009) Grass species differentiation through canopy hyperspectral reflectance. Int J Remote Sens 30:5959–5975. https://doi.org/10.1080/01431160902791895

Jetz W, McGeoch MA, Guralnick R, et al (2019) Essential biodiversity variables for mapping and monitoring species populations. Nat Ecol Evol 3:539–551. https://doi.org/10.1038/s41559-019-0826-1

Johnston JD, Greenler SM, Miller BA, et al (2021) Diameter limits impede restoration of historical

conditions in dry mixed-conifer forests of eastern Oregon, USA. Ecosphere 12:. https://doi.org/10.1002/ecs2.3394

Jones LC, Norton N, Prather TS (2018) Indicators of ventenata (*Ventenata dubia*) invasion in sagebrush steppe rangelands. Invasive Plant Sci Manag 11:1–9. https://doi.org/10.1017/inp.2018.7

Keeley J, Baer-Keeley M, Fotheringham C (2005) Alien plant dynamics following fire in Mediterranean-climate California shrublands. Ecol Appl 15:2109–2125. https://doi.org/10.1890/04-1222

Kerns BK, Tortorelli C, Day MA, et al (2020) Invasive grasses: A new perfect storm for forested ecosystems? For Ecol Manage 463:117985. https://doi.org/10.1016/j.foreco.2020.117985

Knapp PA (1996) Cheatgrass (*Bromus tectorum* L.) dominance in the Great Basin Desert: History, persistence, and influences to human activities. Glob Environ Chang 6:37–52. https://doi.org/10.1016/0959-3780(95)00112-3

Koby LE, Beuschlein J, Prather TS, et al (2019) Management of *Ventenata dubia* in the inland Pacific Northwest with Indaziflam. Invasive Plant Sci Manag 12:223–228. https://doi.org/10.1017/inp.2019.26

Korb JE, Covington WW, Fule PZ. Sampling techniques influence understory plant trajectories after restoration: an example from ponderosa pine restoration. Rest Ecol 11:504-515. https://doi.org/10.1046/j.1526-100X.2003.rec0170.x

Lamm J, Bastow J, Brown R, et al (2021) Short-term nutrient reduction reduces cover of an invasive winter annual grass without negatively impacting the soil microbial community. Restor Ecol 1–11. https://doi.org/10.1111/rec.13469

Levine JM, Vilà M, D'Antonio CM, et al (2003) Mechanisms underlying the impacts of exotic plant invasions. Proc R Soc B Biol Sci 270:775–781. https://doi.org/10.1098/rspb.2003.2327

Luquez V, Hall D, Albrectsen BR, et al (2008) Natural phenological variation in aspen (*Populus tremula*): The SwAsp collection. Tree Genet Genomes 4:279–292. https://doi.org/10.1007/s11295-007-0108-y

Mack MC, D'Antonio CM (1998) Impacts of biological invasions on disturbance regimes. Trends Ecol Evol 13:195–198. https://doi.org/10.1016/S0169-5347(97)01286-X

Mack RN (1981) Invasion of *Bromus tectorum* L. into Western North America: An ecological chronicle. Agro-Ecosystems 7:145–165. https://doi.org/https://doi.org/10.1016/0304-3746(81)90027-5

McMaster GS, Wilhelm WW (1997) Growing degree-days: one equation, two interpretations. Agric For Meteorol 87:291–300. https://doi.org/10.1016/S0168-1923(97)00027-0

Meyer SE, Allen PS, Beckstead J (1997) Seed germination regulation in *Bromus tectorum* (Poaceae) and its ecological significance. Oikos 78:475–485. https://www.jstor.org/stable/3545609

Nicolli M, Rodhouse TJ, Stucki DS, Shinderman M (2020) Rapid invasion by the annual grass *Ventenata dubia* into protected-area, low-elevation sagebrush steppe. West North Am Nat 80:243–252. https://doi.org/10.3398/064.080.0212

Nietupski TC, Kennedy RE, Temesgen H, Kerns BK (2021) Spatiotemporal image fusion in Google Earth Engine for annual estimates of land surface phenology in a heterogenous landscape. Int J Appl Earth Obs Geoinf 99:102323. https://doi.org/10.1016/j.jag.2021.102323

Noujdina NV, Ustin SL (2008) Mapping downy brome (*Bromus tectorum*) using multidate AVIRIS data. Weed Sci 56:173–179. https://doi.org/10.1614/WS-07-009.1

Olsson AD, van Leeuwen WJD, Marsh SE (2011) Feasibility of invasive grass detection in a desertscrub community using hyperspectral field measurements and Landsat TM imagery. Remote Sens 3:2283–2304. https://doi.org/10.3390/rs3102283

Omernik JM (1987) Ecoregions of the conterminous United States. Ann Assoc Am Geogr 77:118–125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x

Pastick NJ, Dahal D, Wylie BK, et al (2020) Characterizing land surface phenology and exotic annual grasses in dryland ecosystems using Landsat and Sentinel-2 data in harmony. Remote Sens 12:1–17. https://doi.org/10.3390/rs12040725

Peterson, EB (2005) Estimating cover of an invasive grass (*Bromus tectorum*) using tobit regression and phenology derived from two dates of Landsat ETM+ data. Int J Remote Sens 26:2491-2507. https://doi.org/10.1080/01431160500127815

Piao S, Fang J, Zhou L, et al (2006) Variations in satellite-derived phenology in China's temperate vegetation. Glob Chang Biol 12:672–685. https://doi.org/10.1111/j.1365-2486.2006.01123.x

PRISM (2012) PRISM Climate Group. In: Oregon State Univ. http://prism.oregonstate.edu. Accessed June 1 2019

Reilly MJ, McCord MG, Brandt SM, et al (2020) Repeated, high-severity wildfire catalyzes invasion of non-native plant species in forests of the Klamath Mountains, northern California, USA. Biol Invasions 22:1821–1828. https://doi.org/10.1007/s10530-020-02227-3

Roundy BA, Chambers JC, Pyke DA, et al (2018) Resilience and resistance in sagebrush ecosystems are associated with seasonal soil temperature and water availability. Ecosphere 9:. https://doi.org/10.1002/ecs2.2417

Royimani L, Mutanga O, Odindi J, et al (2019) Advancements in satellite remote sensing for mapping and monitoring of alien invasive plant species (AIPs). Phys Chem Earth 112:237–245. https://doi.org/10.1016/j.pce.2018.12.004

Rupasinghe PA, Chow-Fraser P (2021) Mapping Phragmites cover using WorldView 2/3 and Sentinel 2

images at Lake Erie Wetlands, Canada. Biol Invasions 23:1231-1247. https://doi.org/10.1007/s10530-020-02432-0

Russelle MP, Wilhelm WW, Olson RA, Power JF (1984) Growth Analysis Based on Degree Days. Crop Sci 24:28–32. https://doi.org/10.2135/cropsci1984.0011183x002400010007x

Simberloff D, Martin JL, Genovesi P, et al (2013) Impacts of biological invasions: What's what and the way forward. Trends Ecol Evol 28:58–66. https://doi.org/10.1016/j.tree.2012.07.013

Simpson MJ, Kertis J, Acker S, et al (2019) Potential vegetation types of California, Oregon, and Washington at the vegzone and subzone scale.

Stohlgren TJ, Falkner MB, Schell LD (1995) A modified-Whittaker nested vegetation sampling method. Vegetatio 117:113-121. https://doi.org/10.1007/BF00045503

Thill DC, Beck KG, Callihan RH (1984) The biology of downy brome (*Bromus tectorum*). Weed Sci 32:7–12. https://doi.org/10.1017/S0043174500060185

Tortorelli CM, Kim JB, Vaillant NM, et al (2023) Feeding the Fire: Annual Grass Invasion Facilitates Modeled Fire Spread across Inland Northwest Forest-Mosaic Landscapes. Ecosphere 14:e4413. https://doi.org/10.1002/ecs2.4413

Tortorelli CM, Krawchuk MA, Kerns BK (2020) Expanding the invasion footprint: *Ventenata dubia* and relationships to wildfire, environment, and plant communities in the Blue Mountains of the Inland Northwest, USA. Appl Veg Sci 1–13. https://doi.org/10.1111/avsc.12511

Ustin SL, Dipietro D, Olmstead K, et al (2002) Hyperspectral remote sensing for invasive species detection and mapping band ratios. IEEE 3:1658-1660. https://doi.org/10.1109/igarss.2002.1026212

Wallace JM, Pavek PLS, Prather TS (2015) Ecological characteristics of *Ventenata dubia* in the Intermountain Pacific Northwest. Invasive Plant Sci Manag 8:57–71. https://doi.org/10.1614/IPSM-D-14-00034.1

Weisberg PJ, Dilts TE, Greenberg JA, et al (2021) Phenology-based classification of invasive annual grasses to the species level. Remote Sens Environ 263:112568. https://doi.org/10.1016/j.rse.2021.112568

West AM, Evangelista PH, Jarnevich CS, et al (2017) 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. Int J Appl Earth Obs Geoinf 59:135–146. https://doi.org/10.1016/j.jag.2017.03.009

Young J (1992) Ecology and management of medusahead (*Taeniatherum caput-medusae* ssp*. asperum* Melderis). Gt Basin Nat 52:245-252. https://www.jstor.org/stable/41712724

Young JA, Evans RA (1970) Invasion of medusahead into the Great Basin. Weed Sci 18:89–97. https://doi.org/10.1017/S0043174500077419

Zimmermann NE, Edwards TC, Moisen GG, et al (2007) Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. J Appl Ecol 44:1057–1067. https://doi.org/10.1111/j.1365-2664.2007.01348.x

Appendix A

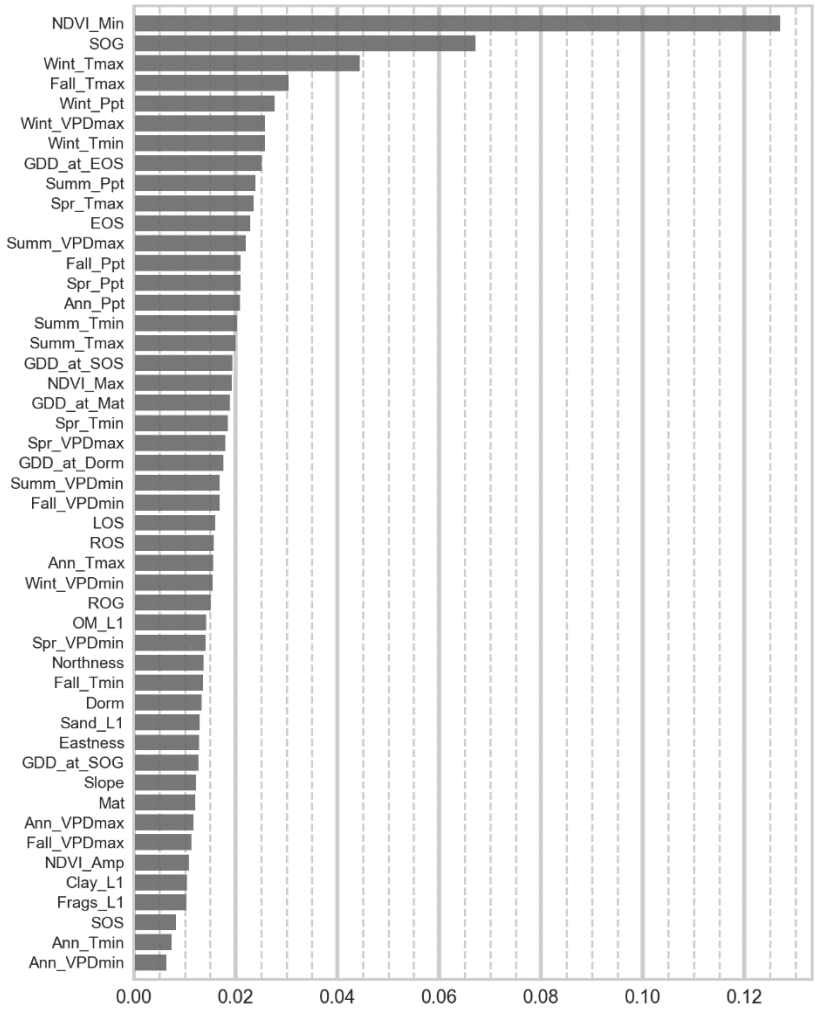

Figure A1. Importance and rank of the predictors included in the hybrid model.

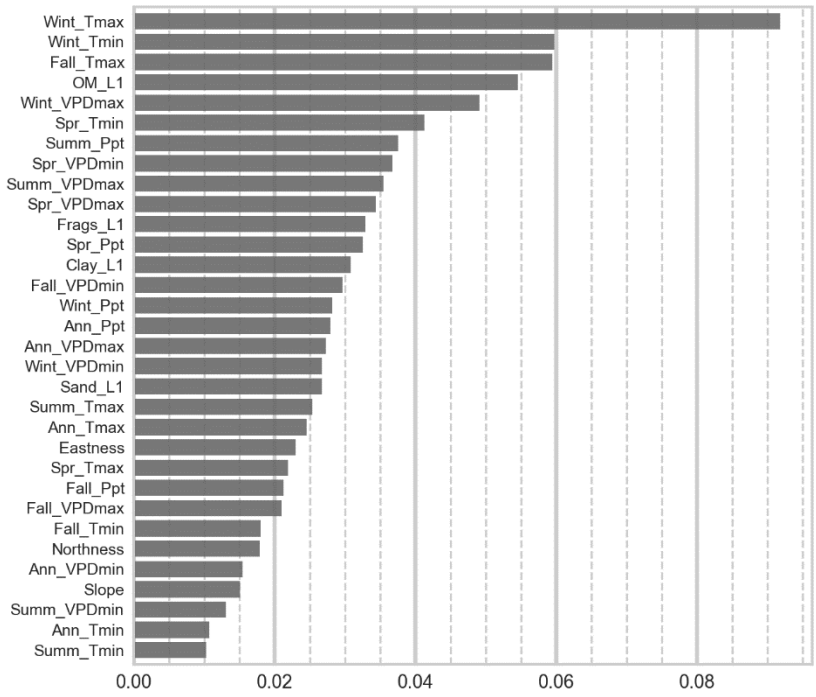

Figure A2. Variable importance and rank of the 32 predictors in the bioclimatic model.

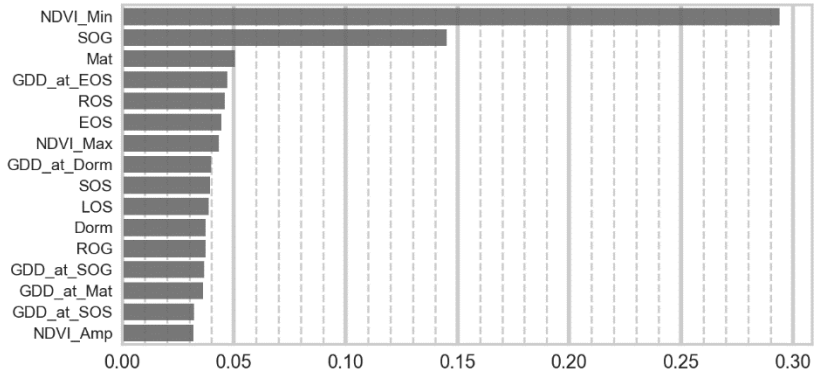

Figure A3. Variable importance and rank of the 16 predictors in the phenology model.

Table A1. Complete list of predictors used to model the presence and absence of ventenata.

Group	Acronym	Units	Description
Phenometrics	Dorm	Day of Year	Inflection point when greenness stops decreasing.
	EOS	Day of Year	Day in which decrease in greenness is steepest.
	LOS	Day of Year	Start and end of season difference.
	Mat	Day of Year	Day of maximum greenness based on function.

	SOS	Day of Year	Day in which increase in greenness is steepest.
	SOG	Day of Year	Inflection point when greenness starts rising.
	ROG	NDVI/Day	Steepness of increase in greenness at the SOS.
	ROS	NDVI/Day	Steepness of decrease in greenness at EOS.
	NDVI_Amp	NDVI	Baseline and maximum NDVI difference.
	NDVI_Max	NDVI	Maximum NDVI observed in 2017.
	NDVI_Min	NDVI	Median NDVI from September 2016 (baseline).
	GDD_at_Dorm	°C	Cumulative GDD at dormancy.
	GDD_at_EOS	°C	Cumulative GDD at EOS.
	GDD_at_Mat	°C	Cumulative GDD at Maturity.
	GDD_at_SOG	°C	Cumulative GDD at SOG.
	GDD_at_SOS	°C	Cumulative GDD at SOS.
Climate	Ann_Ppt	mm	Total annual precipitation.
	Fall_Ppt	mm	Total fall precipitation.
	Spr_Ppt	mm	Total spring precipitation.
	Summ_Ppt	mm	Total summer precipitation.
	Wint_Ppt	mm	Total winter precipitation.
	Ann_Tmin	°C	Mean annual monthly minimum temperature.
	Fall_Tmin	°C	Mean fall monthly minimum temperature.
	Spr_Tmin	°C	Mean spring monthly minimum temperature.
	Summ_Tmin	°C	Mean summer monthly minimum temperature.
	Wint_Tmin	°C	Mean winter monthly minimum temperature.
	Ann_Tmax	°C	Mean annual monthly maximum temperature.
	Fall_Tmax	°C	Mean fall monthly maximum temperature.
	Spr_Tmax	°C	Mean spring monthly maximum temperature.
	Summ_Tmax	°C	Mean summer monthly maximum temperature.
	Wint_Tmax	°C	Mean winter monthly maximum temperature.
	Ann_VPDmin	hPa	Mean annual monthly minimum vapor pressure deficit.
	Fall_VPDmin	hPa	Mean fall monthly minimum vapor pressure deficit.
	Spr_VPDmin	hPa	Mean spring monthly minimum vapor pressure deficit.
	Summ_VPDmin	hPa	Mean summer monthly minimum vapor pressure deficit.
	Wint_VPDmin	hPa	Mean winter monthly minimum vapor pressure deficit.
	Ann_VPDmax	hPa	Mean annual monthly maximum vapor pressure deficit.

	Fall_VPDmax	hPa	Mean fall monthly maximum vapor pressure deficit.
	Spr_VPDmax	hPa	Mean spring monthly maximum vapor pressure deficit.
	Summ_VPDmax	hPa	Mean summer monthly maximum vapor pressure deficit.
	Wint_VPDmax	hPa	Mean winter monthly maximum vapor pressure deficit.
Soils	Clay_L1	%	Clay content in the top 20 cm of the soil.
	Frags_L1	%	Rock content in the top 20 cm of the soil.
	OM_L1	%	Organic matter content in the top 20 cm of the soil.
	Sand_L1	%	Sand content in the top 20 cm of the soil.
Topography	Slope	%	Percent slope
	Eastness	Sine of Aspect	East to west gradient (1 to -1)
	Northness	Cosine of Aspect	North to south gradient (1 to -1)

Table A2. Model performance per potential natural vegetation class. Some metrics could not be calculated if there was a lack of either presence or absence in the class. HY = hybrid, BIO = bioclimatic, PHEN = phenology.

PNV	**Accuracy**			**Sensitivity**			**Specificity**		
	HY	**BIO**	**PHEN**	**HY**	**BIO**	**PHEN**	**HY**	**BIO**	**PHEN**
Developed	1 (0)	1 (0)	0 (1)	-	-	-	1 (0)	1 (0)	0 (1)
Upland grass	1 (0)	1 (0)	1 (0.5)	-	-	-	1 (0)	1 (0)	1 (0.5)
Moist Meadow	1 (0)	1 (0)	1 (0)	-	-	-	1 (0)	1 (0)	1 (0)
Wet Meadow	0.71 (0.2)	0.71 (0.23)	0.71 (0.23)	0 (0)	0 (0.33)	0 (0)	0.89 (0.25)	0.86 (0.25)	0.9 (0.2)
Scabland Shrub	0.9 (0.18)	0.78 (0.21)	0.83 (0.16)	1 (1)	0.5 (1)	0 (0.58)	1 (0.14)	0.82 (0.2)	0.89 (0.18)
Upland Shrub	0.84 (0.09)	0.77 (0.1)	0.82 (0.09)	0.67 (0.3)	0.5 (0.33)	0.5 (0.33)	0.88 (0.1)	0.82 (0.11)	0.89 (0.09)
Riparian Shrub	0.83 (0.07)	0.83 (0.06)	0.8 (0.07)	0.58 (0.2)	0.45 (0.19)	0.43 (0.21)	0.9 (0.08)	0.94 (0.06)	0.91 (0.06)
Juniper Steppe	0.96 (0.03)	0.97 (0.03)	0.97 (0.03)	0 (0.33)	0 (0)	0.4 (0.5)	0.98 (0.03)	1 (0)	0.98 (0.03)
Juniper Woodlands	1 (0)	1 (0)	1 (0)	-	-	-	1 (0)	1 (0)	1 (0)
Ponderosa Pine-Lodgepole Pine	1 (0)	1 (0)	1 (0)	-	-	-	1 (0)	1 (0)	1 (0)
Dry Ponderosa Pine	1 (0.06)	1 (0)	1 (0)	-	-	-	1 (0.06)	1 (0)	1 (0)
Moist Ponderosa Pine	0.86 (0.24)	0.88 (0.2)	0.78 (0.21)	0 (0.5)	0 (0.33)	0 (0)	1 (0.14)	1 (0)	0.88 (0.2)
Xeric Pine	0.89 (0.11)	0.86 (0.12)	0.8 (0.16)	0.67 (0.5)	0.5 (0.33)	0.33 (0.5)	1 (0.1)	1 (0)	1 (0.11)

Dry Douglas-Fir	0.73 (0.15)	0.75 (0.15)	0.73 (0.14)	0.6 (0.35)	0.5 (0.42)	0.67 (0.5)	0.77 (0.15)	0.85 (0.16)	0.75 (0.17)
Moist Douglas-Fir	0.82 (0.09)	0.81 (0.09)	0.79 (0.09)	0.5 (0.27)	0.33 (0.23)	0.33 (0.3)	0.9 (0.09)	0.94 (0.06)	0.9 (0.09)
Dry White Fir - Grand Fir	1 (0)	1 (0)	1 (0)	-	-	-	1 (0)	1 (0)	1 (0)
Moist White Fir - Grand Fir	0.75 (0.5)	1 (0.25)	1 (0.25)	1 (1)	1 (0)	1 (1)	0.75 (0.5)	1 (0.25)	1 (0)
Wet White Fir - Grand Fir	1 (0)	1 (0)	1 (0.11)	-	-	-	1 (0)	1 (0)	1 (0.11)
Cold Dry White Fir - Grand Fir	1 (0)	1 (0)	1 (0)	-	-	-	1 (0)	1 (0)	1 (0)
Moist Subalpine Fir	1 (0)	0.96 (0.05)	1 (0)	0 (0)	0 (0)	0 (0)	1 (0)	1 (0.04)	1 (0)
Wet Subalpine Fir	1 (0)	1 (0)	1 (0.25)	1 (0)	1 (0)	1 (0)	1 (0)	1 (0)	1 (0.25)

Table A3. Composition of the BME by PNV class and their representation in the field data set.

PNV	Percent of BME	n	Prevalence
Developed	4	1	0
Upland grass	11	93	0.19
Moist Meadow	1	4	0.5
Wet Meadow	0	1	0
Scabland Shrub	4	32	0.41
Upland Shrub	20	53	0.15
Riparian Shrub	0	1	0
Juniper Steppe	0	1	0
Juniper Woodlands	5	15	0.2
Ponderosa Pine-Lodgepole Pine	0	1	1
Dry Ponderosa Pine	1	16	0.13
Moist Ponderosa Pine	1	22	0.05
Xeric Pine	5	101	0.25
Dry Douglas-Fir	2	45	0.18
Moist Douglas-Fir	6	74	0.05
Dry White Fir - Grand Fir	2	69	0.04
Moist White Fir - Grand Fir	19	270	0.01
Wet White Fir - Grand Fir	3	45	0
Cold Dry White Fir - Grand Fir	2	88	0.03
Moist Subalpine Fir	1	3	0
Wet Subalpine Fir	2	5	0

Highlights

- Land surface phenology and biophysical predictors resulted in improved maps of an invasive grass
- 30-meter resolution phenology enabled detection of smaller patches of an invasive annual grass
- Notable populations of an invasive annual grass are identified in forest openings and ecotones