



# Two invasive *Hieracium* species' potential distributions within the Greater Yellowstone Ecosystem were defined using invasion susceptibility models and habitat typing

Christie H. Guetling · Lisa C. Jones ·  
Eva K. Strand · Don W. Morishita ·  
Julia Piaskowski · Timothy S. Prather

Received: 2 August 2022 / Accepted: 27 February 2023 / Published online: 23 March 2023  
© The Author(s) 2023

**Abstract** Invasive *Hieracium* plant species are invading the Greater Yellowstone Ecosystem. The potential distribution of orange hawkweed (*Hieracium aurantiacum*) and meadow hawkweed (*Hieracium caespitosum*) were estimated using habitat susceptibility models to assist land managers' management of these invasive plants. The objectives of this study were to: (1) develop models describing susceptibility of ecosystems to hawkweed invasion, (2) identify indicator species of orange hawkweed and meadow hawkweed, (3) determine habitat types where these invasive hawkweeds might occur, and

(4) create habitat susceptibility maps for management planning and ground surveys. Models were developed using a Mahalanobis distance similarity technique from remotely sensed biotic and abiotic variables, as well as known location data for orange and meadow hawkweed. Ground validation was conducted to assess model weaknesses and subsequent model modification. Indicator plant species were identified as surrogates to determine the likelihood of hawkweed presence during ground survey. Transect data collected from areas susceptible to invasion also were used to determine habitat types where hawkweed might occur. The best model included eight variables: north–south aspect, east–west aspect, slope, NDVI, NDWI, blue spectral band, green spectral band, and precipitation. High susceptibility (65+ % likelihood of suitable habitat) consisted of 66,000 ha for meadow hawkweed and 35,000 ha for orange hawkweed, 5.0% and 2.7% of the study area, respectively. Meadow hawkweed and orange hawkweed had seven and three indicator plant species, respectively. Predicted hawkweed habitat susceptibility encompassed nine habitat types, ranging from xeric sagebrush steppe to wet forests and they overlapped except at the xeric habitat type. Habitat susceptibility models save costs and allow survey prioritization to those areas most susceptible to invasion.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10530-023-03037-z>.

C. H. Guetling  
USDA-Agriculture Research Service, Eastern Oregon  
Agricultural Research Center, Burns, OR 97720-9394,  
USA

L. C. Jones · D. W. Morishita · T. S. Prather (✉)  
Department of Plant Sciences, University of Idaho, 875  
Perimeter Drive MS 2333, Moscow, ID 83844-2333, USA  
e-mail: tprather@uidaho.edu

E. K. Strand  
Department of Forest, Rangeland and Fire Sciences,  
University of Idaho, 875 Perimeter Drive MS 1133,  
Moscow, ID 83844-1133, USA

J. Piaskowski  
Statistical Programs, University of Idaho, 875 Perimeter  
Drive MS 2337, Moscow, ID 83844-2337, USA

**Keywords** Geographic information systems (GIS) · Habitat suitability · Indicator species · Mahalanobis distance · NDVI · NDWI · Remote sensing · Sentinel-2

## Introduction

Invasive plants, defined here as non-indigenous plants that change systems where they occur, are affecting areas of significant importance around the world as anthropogenic forces alter land use and transport seeds to new areas (Shiferaw et al. 2019; Szilassi et al. 2019; Pappgh et al. 2021). Invasive plants can reduce forage production, alter fire return intervals, reduce soil water content, and change native plant composition (Balch et al. 2013; Pfeiffer and Gorchov 2015; Shiferaw et al. 2019; Fenouillas et al. 2021; Papp et al. 2021). Land managers currently lack sufficient resources to reduce the ecological threat of invasive plants so it is imperative to prioritize areas with the greatest risk of invasion to receive focused management efforts.

Invasive plants cost billions of dollars each year from monitoring and control efforts, to higher food costs and reduced incomes on farms (Pyšek and Richardson 2010). Early detection of invasive plants, when infestations are small (<1000 hectares), can mean the difference between efficient eradication and billions of dollars for ongoing control to stem biodiversity loss (Rejmánek and Pitcairn 2002; Shiferaw et al. 2019). Shiferaw et al. (2019) found the invasion of mesquite (*Prosopis* L.) in Ethiopia did not increase linearly when comparing Landsat data over time. They found that early on, the annual rate of spread of mesquite was over four-times slower compared to the spread in later years, highlighting the importance of early detection and early management strategies. Scientists are using remote sensing to better understand the distribution of invasive plants as they expand and move into more remote locations (e.g., Shiferaw et al. 2019; Masemola et al. 2020; Fenouillas et al. 2021; Papp et al. 2021; Pathak et al. 2021; Ghorbanian et al. 2022; Matas-Granados et al. 2022) because remote sensing reduces the cost associated with in situ sampling while providing greater spatial breadth and more frequent repeated measures than ground surveys could possibly achieve (Ahmed

et al. 2020; Papp et al. 2021). Remote sensing has been essential in providing land managers with accurate data to manage invasive plants: Sentinel-2 multi-spectral data have been used to map *Acacia dealbata* (Link) and *A. mearnsii* (De Wild.) spread into grasslands and forests of South Africa (Masemola et al. 2020); Normalized Difference Vegetation Index (NDVI) was used to investigate long-term changes in land cover and identify vulnerability of threatened plants in Spain (Matas-Granados et al. 2022); and slope, aspect and elevation have been used to predict the likelihood of yellow star-thistle (*Centaurea solstitialis* L.) in the United States (Shafii et al. 2004).

Modeling has been used to predict habitat susceptibility to weed infestations using information of existing or previous infestations to prioritize survey areas (Shafii et al. 2003). Prioritizing efforts across expansive landscapes may be aided through use of Habitat Susceptibility Models (HSM) created within geographic information systems (GIS). GIS tools are used to combine spatial environmental data at known infestations and predict areas likely to be invaded based on similar ground conditions (Shafii et al. 2003, 2004; Rew and Maxwell 2006). Areas that are similar to the known infestations of a target species are considered to be highly susceptible to invasion by that species, whereas areas that are different, are considered to have low susceptibility. HSM can efficiently direct ground survey efforts to locate invasive plant populations by identifying highly suitable habitat, saving land managers time and money.

The first step to developing an HSM is to identify abiotic and biotic factors (e.g., slope, aspect, elevation, vegetation indices) linked to the distribution of a target species (Lass et al. 2011; Adhikari et al. 2020). Remotely sensed digital elevation models can be used to interpret abiotic conditions such as aspect and slope, which have a direct impact on solar radiation and soil moisture levels (Holland and Steyn 1975). Some biotic conditions can be assessed using multispectral reflectance data (Lass et al. 2011; John et al. 2018; Ghorbanian et al. 2022). The reflectance bands encompass visible and near-infrared wavelengths; blue, green, red, and near-infrared (NIR) (central wavelengths: 490 nm, 560 nm, 665 nm, and 865 nm, respectively) often are used in ecological studies (Lass et al. 2011). The normalized difference vegetation index (NDVI,

where  $NDVI = \frac{(NIR-Red)}{(NIR+Red)}$  (Tucker 1979) is widely used because it differentiates between dense, living vegetation and sparse or senesced vegetation (e.g., Eitel et al. 2011; Lass et al. 2011; Ghorbanian et al. 2022). Recent studies have also incorporated the short-wave infrared band (SWIR, central wavelength = 1610 nm) to calculate additional indices such as the normalized difference water index (NDWI, where  $NDWI = \frac{(NIR-SWIR)}{(NIR+SWIR)}$ ) which relates to plant water content (John et al. 2018). When models are unavailable, habitat types or indicator species may be used to aid on-the-ground surveys.

Habitat type is a term used to describe ecological characteristics of a location that act as a guide to distinguish between sites that can or cannot support certain plant species (Daubenmire 1984). This can be beneficial during field surveys for ground-truthing models or if maps cannot be loaded due to cellular service limitations or dead batteries in field tablets. Another option for determining habitat suitability when models are unavailable is to use an indicator species, a native or naturalized plant species with similar life-history characteristics of the target species, which may be considered an “indicator” of the potential presence of that target (Fleishman et al. 2005). An indicator species can inform managers on the susceptibility risk when the target species is absent or in low abundance.

In the United States as of 2017, 566,600 ha of National Parks were infested with invasive plants, with only 17,400 ha controlled (NPS 2019). Yellowstone National Park (YNP) is part of the Greater Yellowstone Ecosystem (GYE) which spans more than 8-million hectares across portions of Montana, Idaho, and Wyoming, USA. The GYE is a unique and intact ecosystem that is threatened by invasive species. It is characterized by a diverse flora with many vegetation types including: conifer forests, sagebrush steppes, and mountain meadows (Despain 1990). Invasive plants within the GYE challenge conservation goals, particularly meadow hawkweed (*Hieracium caespitosum* Dumort. = *H. pratense* Tausch) and orange hawkweed (*H. aurantiacum* L.). Meadow hawkweed and orange hawkweed were introduced into North America as ornamentals from Europe, as early as the nineteenth century (Wilson 2007). Hawkweed seeds are primarily wind-dispersed but can also be moved by animals. Once established, rhizomes and stolons

allow plants to expand locally, at times excluding other plant species (Wilson and Callihan 1999).

Our study aimed to assist early detection of meadow hawkweed and orange hawkweed in the GYE by providing habitat susceptibility models for risk assessment and directing both plant survey and control activities for hawkweeds. The objectives were to: (1) develop susceptibility models for meadow hawkweed and orange hawkweed using known locations and environmental data within the GYE, (2) identify indicator species of meadow and orange hawkweed then assess the indicator power, (3) determine habitat types where meadow and orange hawkweed occur, and (4) create susceptibility maps that can be served on software platforms to allow invasive plant species mapping on mobile devices.

## Materials and methods

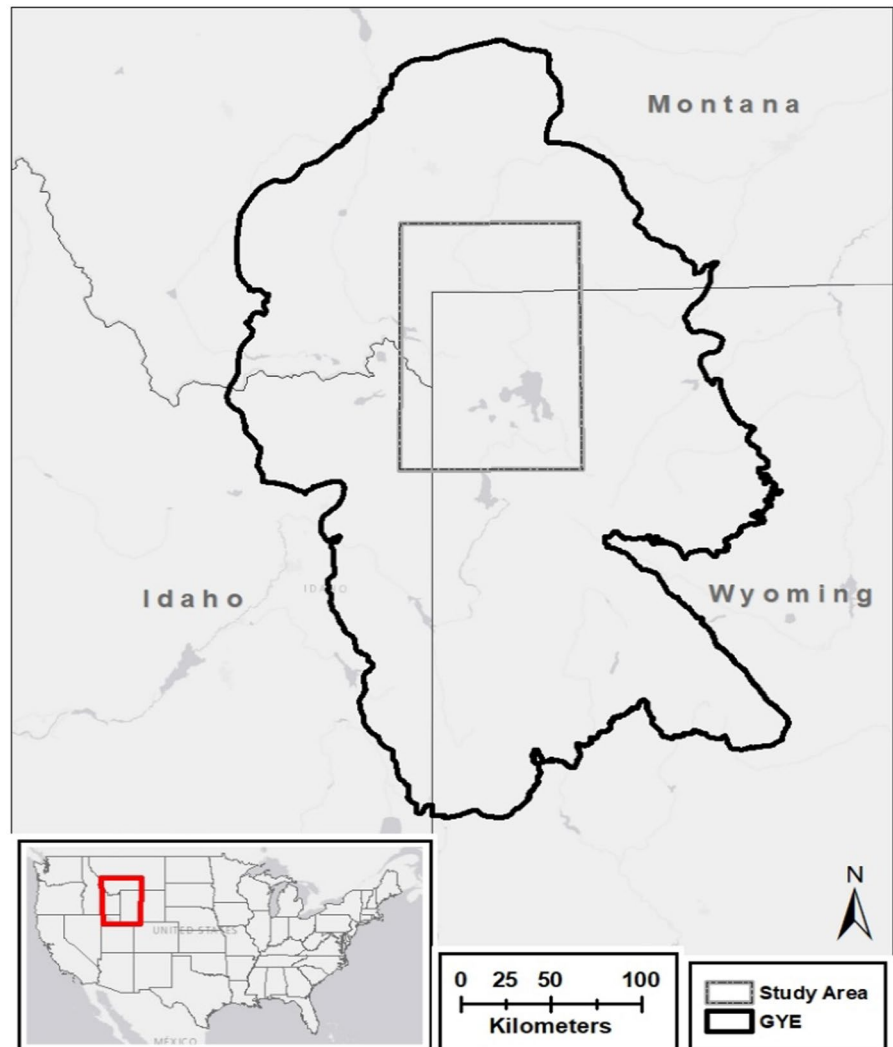
### Study area

The study area (Fig. 1) encompasses YNP and extends 38 km north and 13 km east of YNP (bounding coordinates: 111° 15' 36" W, 110° 0' 36" W, 45° 21' 36" N, 44° 11' 24" N). The exact extent was selected to include essentially all target species occurrence data. The total study area is approximately 1.31 million hectares, representing 16.5% of the entire GYE land area (Yellowstone National Park 2016). The elevation within the study area is between 1550 m (5090 ft) and 3300 m (10,830 ft).

### Spatial data acquisition

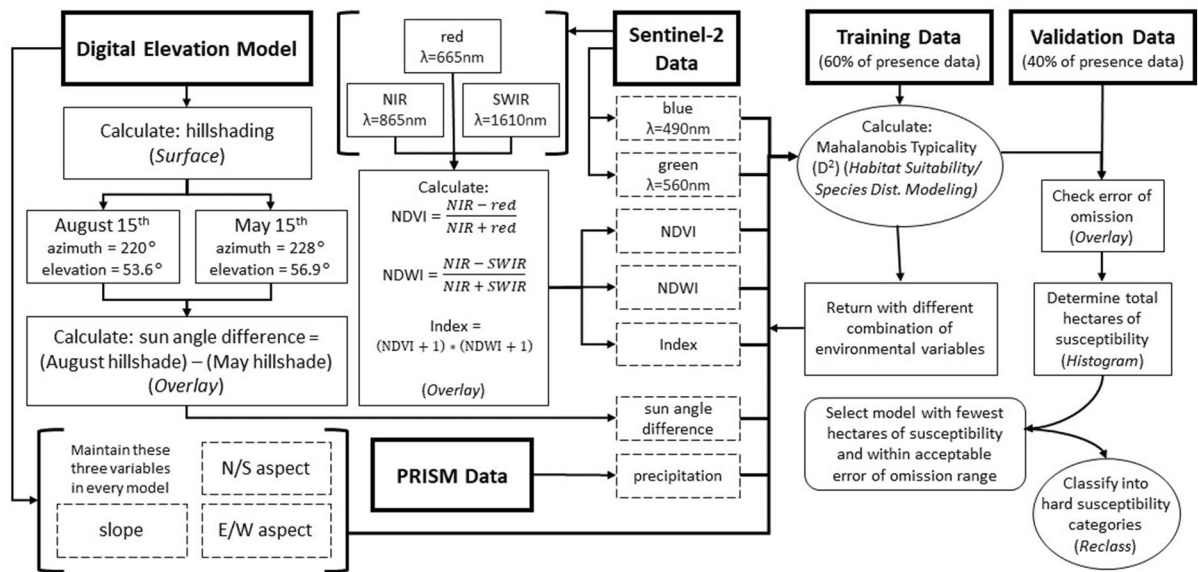
Ten environmental covariates were acquired for modeling: north–south aspect, east–west aspect, slope, hill shade, precipitation, blue and green spectral bands, NDVI, NDWI, and a moisture gradient index  $[(NDWI+1) * (NDVI+1)]$  (Fig. 2). Five spectral bands of Sentinel-2 data were downloaded from Earth Explorer USGS (European Space Agency 2018) (Fig. 2). All spectral data were analyzed at 20-m resolution. The NDVI and NDWI were calculated then combined to create a single variable to represent a moisture gradient of both. Each index was normalized by adding a value of one to remove negative values, then they were multiplied together.

**Fig. 1** Study area within the Greater Yellowstone Ecosystem



Average annual precipitation data were acquired from Parameter-elevation Regressions on Independent Slopes Model (PRISM 2020) for 2012 through 2018 at 4-km spatial resolution. Data from each year were projected to 20-m resolution and then resampled using focal statistics with a 300 by 300-cell window to smooth the edges of 4-km data (TerrSet v. 18.31, Clark Labs, Worcester, MA; ArcMap v. 10.6.1, Esri, Redlands, CA). Average annual precipitation across the study area was 36 cm to 147 cm. A 30-m resolution digital elevation model (DEM) was acquired from Earth Explorer USGS (2018) and resampled to 20-m resolution using the nearest neighbor technique in ArcMap (Esri, Redlands, CA). Aspect, slope, and hillshade were derived from the DEM. In the northern

hemisphere, south facing slopes receive more direct sunlight than north facing slopes and are typically drier. Because aspect is a circular variable, it was split into north–south ( $\cos R$ ) and east–west components ( $\sin R$ ) (Woodcock et al. 2008) where  $R$  = aspect in radians. Species presence data were obtained from the Greater Yellowstone Coordinating Committee (GYCC), Northern Rocky Mountain Exotic Plant Management Team, YNP Exotic Plant Management Team, and the US Forest Service. Locations were represented by polygons surrounding infestations. Ground surveys were conducted July 10–13, 2018 and July 22–26, 2019 to refine polygons and digitize new polygons using Collector (Esri, Redlands, CA). Location data were later converted from vectors to



**Fig. 2** Steps for building a susceptibility model. Bold rectangles represent raw data and dashed rectangles represent environmental data layers created from the raw data. Ovals indicate soft or hard classification steps. Words in *(italics)* indicate the TerrSet tool used. Different combinations of environmental

layers were tested with the species training data to find the best model. The final model was chosen to minimize susceptible area while maintaining a 9% error of omission for validation data. Refer to Table 1 for specific environmental variables used in each model iteration

raster files at 20-m resolution (TerrSet v. 18.31, Clark Labs, Worcester, MA) then approximately 60% of the known occurrence pixels were used as training data to develop the model (hereafter, training data) and 40% were used to validate the model (hereafter, validation data).

### Model building

Susceptibility models were developed using Mahalanobis distance ( $D^2$ ) among landscape variables using *Mahalanobis Typicality* in the *Habitat and Biodiversity Modeler: Habitat Suitability/Species Distribution Modeling* tool (Clark Labs 2017). Mahalanobis Typicality is a semi-supervised classifier that assigns probabilities to pixels indicating their similarity to the training data pixels. The  $D^2$  statistic is the standardized squared distance between habitat covariates for a given sample of training data and the mean vector of the covariates (Griffin et al. 2010). A  $D^2$  value was calculated for each pixel in the study area based on the value of the covariates within that pixel, relative to the covariate averages in the training data:

$$D^2 = (\hat{\mu} - \underline{x})' \hat{\Sigma}^{-1} (\hat{\mu} - \underline{x})$$

where  $\hat{\mu}$  is the vector of the mean values for habitat covariates and  $\hat{\Sigma}$  is the variance–covariance matrix for habitat covariates at training data locations.

Seventeen different combinations of the environmental variables described above were tested in the *Habitat and Biodiversity Modeler: Habitat Suitability/Species Distribution Modeling* tool in TerrSet for predictions of meadow hawkweed and orange hawkweed susceptibility (Table 1). Each model included north–south aspect, east–west aspect, and slope, along with one to five additional variables. Models were created at 20-m resolution and assessed based on a 9% error of omission rate for the training data. The error rate was achieved by selecting a threshold for  $D^2$  values that maintained 91% of training pixels in the “susceptible” category. A conservative threshold was chosen in hopes of finding invasions that may occur towards the ecological limits of the hawkweeds’ ability to persist.  $D^2$  values were set to create high, moderate, and low susceptibility categories (Table 2).

**Table 1** Variables for meadow hawkweed (MeH) and orange hawkweed (OrH) susceptibility model selection

	NDVI	Sun angle diff	NDWI	(NDVI+1) (NDWI+1)	Blue band (W sr <sup>-1</sup> m <sup>-2</sup> )	Green band (W sr <sup>-1</sup> m <sup>-2</sup> )	Precip. (cm)	Validation error		Area susceptible		D <sup>2</sup> cutoff	
								MeH (%)	OrH (%)	MeH (ha)	OrH (ha)	MeH	OrH
Model 1	x							5.1	12.7	987,000	815,000	0.100	0.160
Model 2	x	x						7.5	15.7	952,000	797,000	0.090	0.140
Model 3			x					4.1	15.3	964,000	781,000	0.110	0.150
Model 4	x		x					6.0	14.7	951,000	718,000	0.110	0.185
Model 5				x				5.0	14.7	960,000	813,000	0.110	0.145
Model 6	x				x			6.0	13.3	896,000	663,000	0.115	0.210
Model 7	x				x		x	3.8	13.7	659,000	330,000	0.100	0.220
Model 8	x				x	x	x	5.0	14.0	615,000	319,000	0.100	0.225
Model 9			x		x			4.1	14.3	882,000	633,000	0.115	0.195
Model 10			x				x	2.6	16.3	701,000	417,000	0.085	0.155
Model 11							x	0.6	10.0	776,000	474,000	0.070	0.170
Model 12	x						x	0.4	11.0	710,000	418,000	0.080	0.170
Model 13					x		x	2.3	9.3	714,000	397,000	0.078	0.210
Model 14	x		x		x		x	4.7	12.7	658,000	335,000	0.101	0.210
Model 15			x		x		x	2.6	15.0	659,000	332,000	0.098	0.210
Model 16			x		x	x	x	4.3	16.0	649,000	303,000	0.093	0.220
Model 17	x		x		x	x	x	6.4	13.7	614,000	314,000	0.103	0.220

All models include two aspect variables (north–south, east–west) and slope plus the variables indicated by x. Sun angle diff. represents the difference in hillshading in August subtracted from hillshading in May. Blue and green bands are Sentinel-2 spectral bands. Precipitation represents the 7-year average from 2012 to 2018. Validation error represents the portions of meadow hawkweed or orange hawkweed validation pixels that were classified as “not susceptible” based on the Mahalanobis cutoff (D<sup>2</sup>) values (0.070–0.115 for MeH models; 0.140–0.220 for OrH models) that allowed for 8.5–9.0% training error (percent of training pixels categorized as “not susceptible”)

### Reducing model error in lodgepole pine stands

Positive model predictions in unsuitable habitat were compared to determine possible areas for model improvement. Although no meadow or orange hawkweed populations were documented or observed within dense lodgepole pine (*Pinus contorta* Douglas ex Loudon) stands, the hawkweed susceptibility models predicted varying levels of susceptibility in those areas. Therefore, we developed a method to identify and remove some of these lodgepole pine locations.

Fourteen polygons of dense lodgepole pine were digitized in summer 2019 and determined to be within fire perimeters that occurred in 1988. Fire polygons from 1982 to 1994 were extracted from the dataset and considered as potential dense lodgepole pine. Fire polygons from 1995 to 2016 (Gibson 2005 [years 1995–2002]; LANDFIRE 2019 [years 2003–2016]) were removed from the potential dense lodgepole pine perimeters to ensure recently burned locations were not considered as they could still be suitable to hawkweed invasion.

To predict locations of dense lodgepole pine, *Habitat and Biodiversity Modeler: Habitat Suitability/Species Distribution Modeling* (Clark Labs 2017) was used across the study area, following methods described above for hawkweed models. Sixty percent of dense lodgepole pine pixels were used to train five models using different combinations of blue and green spectral bands, NDVI, NDWI, and precipitation (Table 3). The output was a map that assigned each

pixel a probability from 0 to 1 based on similarity to the training data. The previously described fire region was then selected from the full study area model. Model accuracy was determined by the proportion of dense lodgepole pine validation pixels properly classified in “predicted dense lodgepole pine”, and the proportion of open forest, meadow hawkweed occurrence pixels, and orange hawkweed occurrence pixels classified as “not dense lodgepole pine”. The dense lodgepole pine model was based on Mahalanobis Typicality values.

### Final model processing steps

A hydrologic features map was digitized in ArcGIS 10.6.1 (Esri, Redlands, CA) then imported into TerrSet as vector data and converted to a raster at 40-m resolution to remove water bodies. Because most paved roads in the study area were less than 15-m across, it did not seem appropriate to convert roads into 40-m pixels and lose susceptibility predictions adjacent to roads.

### Plant community surveys

Plant community surveys were conducted July 10–13, 2018 and July 22–26, 2019. Due to a late spring in 2019, seasonal variation in sampling times was minimal. Surveys were less than 1.2 km from a road, though usually within 0.5 km. Each transect was 20-m long with five quadrats (i.e., plots), sized

**Table 2** The D<sup>2</sup> values for categories of meadow and orange hawkweed susceptibility. Percent of training data, validation data, and study area in each category

Susceptibility categories	D <sup>2</sup> cutoff	Percent of training data	Percent of validation data	Percent of study area	Total hectares
<i>Meadow hawkweed</i>					
Not susceptible	$0.000 < x \leq 0.103$	9	6	53	694,000
Low	$0.103 < x \leq 0.35$	21	19	27	356,000
Moderate	$0.35 < x \leq 0.65$	29	27	15	192,000
High	$0.65 < x \leq 1.0$	41	47	5	66,000
<i>Orange hawkweed</i>					
Not susceptible	$0.00 < x \leq 0.22$	9	14	76	994,000
Low	$0.22 < x \leq 0.50$	20	26	15	200,000
Moderate	$0.50 < x \leq 0.70$	30	28	6	79,000
High	$0.70 < x \leq 1.0$	41	33	3	35,000

The D<sup>2</sup> values were set to ensure high, moderate, low and not susceptible categories contained approximately 40%, 30%, 20%, and 9% of the training data, respectively

**Table 3** Variables for dense lodgepole pine (DLP) model selection

	NDWI	NDVI	Precipitation	Training error (%)	Validation error (%)	D <sup>2</sup> cutoff	OrH error (%)	MeH error (%)	OrH and MeH error (%)	Open forest error (%)
DLP Model 1	x			15	27	0.35	2	2	15	9
DLP Model 2		x		15	13	0.33	1	1	14	14
DLP Model 3	x		x	15	40	0.30	1	0	9	0
DLP Model 4		x	x	10	40	0.30	0	0	0	8
DLP Model 5			x	15	40	0.30	1	0	7	0

All models include the blue and green spectral bands plus variables indicated with an x

0.25 m by 0.5 m, placed along the transect at 5-m intervals starting at 0-m. Within plots, the percent cover of bare ground, rock, and each plant species were estimated visually and assigned a ranked value based on estimated cover classes of: 0%, 1 to 5%, 5 to 12.5%, 12.5 to 25%, 25 to 50%, 50 to 75%, 75 to 95%, and 95 to 100%. Habitat types were identified using plant cover data from transects. Each transect was oriented to stay within a single community type. In 2018, 15 transects were within known locations of historic meadow hawkweed and orange hawkweed infestations. In 2019, survey locations were within several contiguous pixels of a susceptibility category (high, moderate, or low) based on preliminary model predictions. A total of 45 transects were surveyed within meadow hawkweed or orange hawkweed susceptibility at elevations between 1750 m (5740 ft) and 2430 m (7970 ft). Transect locations were categorized by susceptibility category for meadow and orange hawkweed based on the final model.

#### Indicator species analysis

An indicator species analysis was conducted using a chi-squared test based on the historical occurrence of meadow hawkweed or orange hawkweed against each species found along each transect using the *vegan* package in R v. 3.5.1 (R Core Team 2018; Jones et al. 2018). Species found in fewer than three transects were removed from the statistical analysis, but were considered for habitat typing, leaving 66 species for statistical testing. A Benjamini and Hochberg (1995) *p* value adjustment was used to control the false discovery rate.

A quantitative measure of the indicator power (IP) of an indicator species for the target invasive species was calculated from a presence-absence matrix on

species with unadjusted chi-squared *p* values  $\leq 0.05$ . A species is considered a strong indicator when it frequently co-occurs with the target invasive species while also infrequently occurring in the absence of the target. The IP equation:

$$IP_I = \sqrt{\{[S \div O_I][1 - (O_T - S) \div (N - O_I)]\}}$$

where  $O_I$  is the frequency of the indicator species (*I*) occurrence,  $O_T$  is the frequency of the target invasive species (*T*) occurrence, *S* is the frequency of shared occurrences of the indicator and the target species, and *N* is the total number of transects surveyed (Halme et al. 2009).

#### Habitat typing

A habitat vector layer for ArcMap and vegetation guide, both created by Despain (1990), were used to identify habitat types at transect locations. Due to the large spatial scale of Despain's (1990) habitat types, habitat types were further refined using Steele et al. (1983) forest descriptions and Mattson (1984) meadow or open riparian descriptions. Transect photographs and species lists were compared to descriptions to confirm habitat types.

## Results

#### Susceptibility models

Model 17 containing 8 variables: north–south aspect, east–west aspect, slope, NDVI, NDWI, blue spectral band, green spectral band, and precipitation had low error rates for both hawkweed species validation data while minimizing the total area



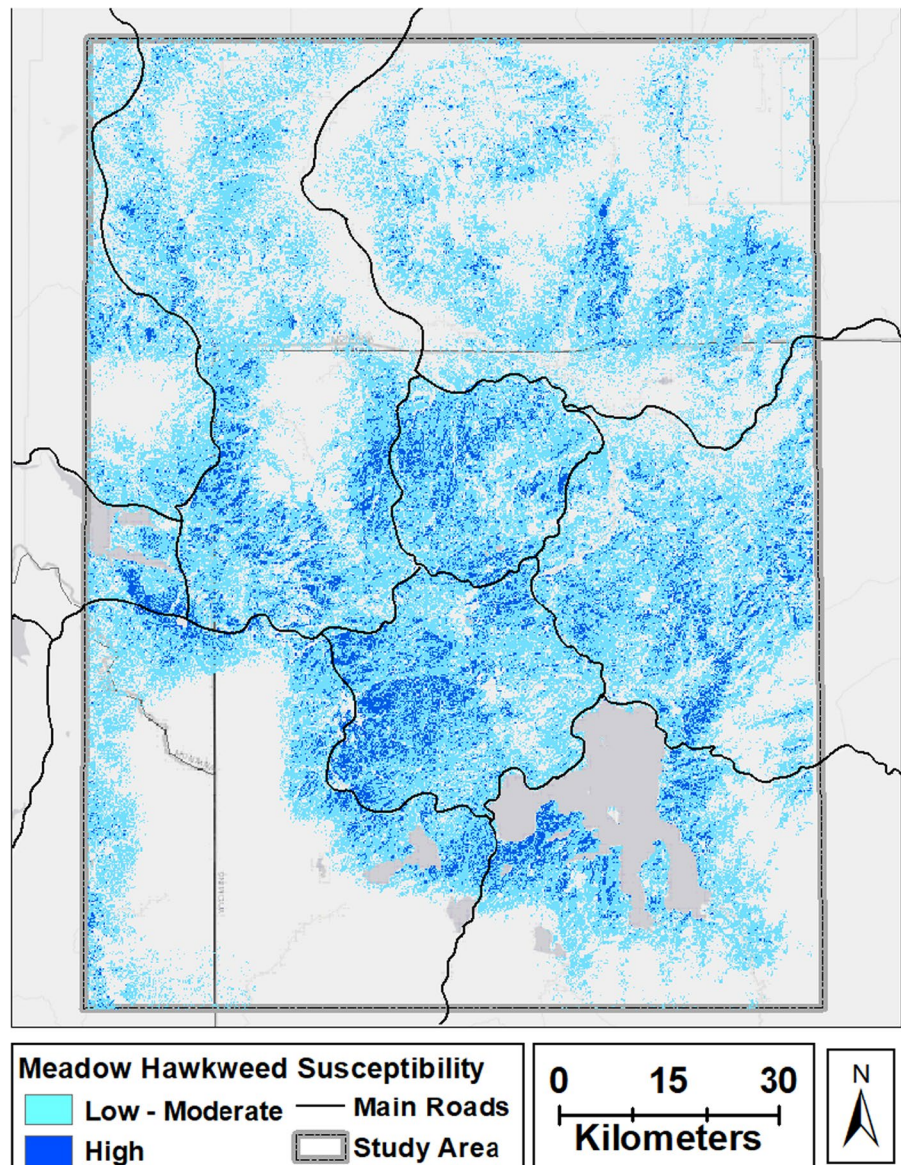
predicted to be susceptible to better focus ground survey efforts (Table 1; Figs. 3, 4). Generally, areas that received less than 40 cm (e.g., along the Yellowstone River north of YNP) or more than 95 cm (e.g. south-western portion of YNP and high elevations) of annual precipitation were not susceptible to meadow or orange hawkweed. Incorporating both blue and green spectral bands in model 17 greatly reduced the area predicted to be susceptible, compared to other models (Table 1).

Model output was divided into four susceptibility categories based on training data occurrence rates.

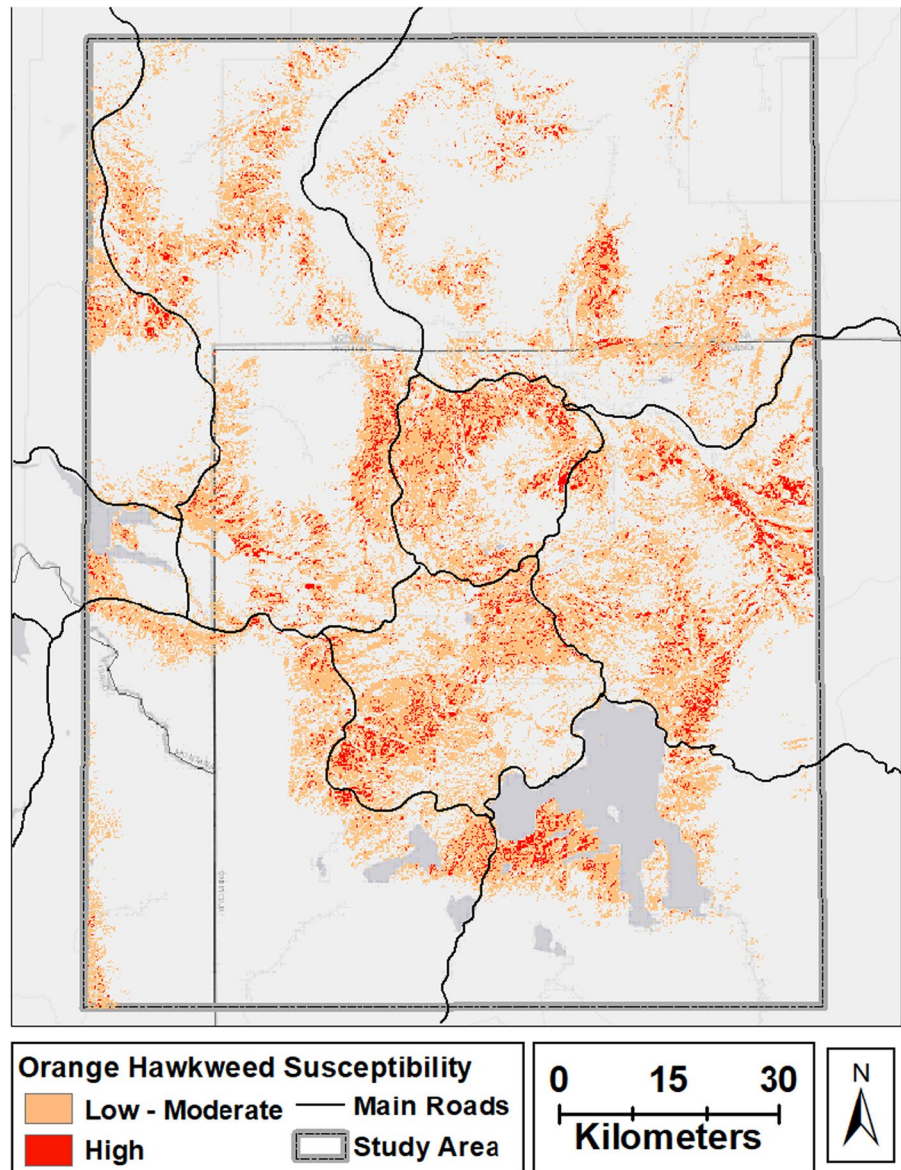
The  $D^2$  values were set to correspond with training data occurrence rates of approximately 9%, 20%, 30%, and 40% for the zero, low, moderate, and high categories, respectively (Table 2).

Approximately 614,000 ha (47%) of the study area were predicted to be susceptible to meadow hawkweed. Of the susceptible area, 66,000 ha were in the high susceptibility category; 192,000 ha in moderate susceptibility; and 356,000 ha in low susceptibility (Table 2). Approximately 314,000 ha (24%) of the study area were predicted to be susceptible to orange hawkweed. Of the area susceptible to

**Fig. 3** Meadow hawkweed susceptibility model



**Fig. 4** Orange hawkweed susceptibility



orange hawkweed, 35,000 ha were in high susceptibility; 79,000 ha were in moderate susceptibility; and 200,000 ha were in low susceptibility (Table 2).

Overlaying susceptibility models indicated 670,400 ha (51%) of the study area were not susceptible to either hawkweed (Table 4). Overlapping areas of low susceptibility encompassed 80,900 ha (6%), overlapping moderate susceptibility encompassed 36,500 ha (3%), and overlapping high susceptibility encompassed 20,800 ha (2%) of the study area (Table 4). Approximately 53% (324,000 ha) of area

susceptible to meadow hawkweed was not susceptible to orange hawkweed and 7% (20,600 ha) of area susceptible to orange hawkweed was not susceptible to meadow hawkweed.

Approximately 413,000 ha (32%) of the study area were considered dense lodgepole pine based on wildfires between 1982 and 1994. The best model for dense lodgepole pine was DLP Model 4, which used blue and green spectral bands, NDVI, and precipitation (Table 3). The area selected for removal (considered to be dense lodgepole pine) had a  $D^2$  value

**Table 4** Meadow hawkweed (MeH) and orange hawkweed (OrH) susceptibility model overlap in hectares and the percent of the study area susceptible

MeH categories (across)	Zero susceptibility		Low susceptibility		Moderate susceptibility		High susceptibility		OrH totals	
	ha	%	ha	%	ha	%	ha	%	ha	%
<i>OrH categories (down)</i>										
Zero susceptibility	670,400	51	255,400	20	63,600	5	5000	< 1	994,400	76
Low susceptibility	20,100	2	80,900	6	81,600	6	17,400	1	200,000	15
Moderate susceptibility	2900	< 1	16,100	1	36,500	3	23,300	2	78,800	6
High susceptibility	500	< 1	3400	< 1	10,800	1	20,800	2	35,500	3
MeH totals	693,900	53	355,800	27	192,500	15	66,500	5		

The first value, 670,400 ha and 51% indicate the portion of the study area in the zero category for both (i.e. not susceptible)

of  $\geq 0.30$  to keep the training error at or below 15% and minimize error for meadow hawkweed, orange hawkweed, and open forest (Table 3). DLP Model 4 was selected because it had  $D^2 \leq 0.30$ , no error for meadow hawkweed or orange hawkweed location data, and only 8% error for open forest. Dense lodgepole pine exclusion was 9170 ha and 8640 ha for the meadow and orange hawkweed models, respectively.

#### Indicator species from plant community surveys

Forty-five transects were surveyed for plant cover in habitats predicted to be susceptible to meadow and orange hawkweeds based on model predictions (Table 5). The meadow hawkweed model predicted 13, 10, 22, and zero transects were in the high, moderate, low, and no susceptibility categories, respectively. The orange hawkweed model predicted 11, 8, 17, and 9 transects were in the high, moderate, low, and no susceptibility categories, respectively.

A total of 209 plant species were recorded within the 45 transects, of which we identified 143 (Appendix A). Three genera were identified to genus only: strawberry (*Fragaria* L.), sedge (*Carex* L.), and horsetail (*Equisetum* L.). The chi-square contingency test revealed seven species were indicator species ( $p < 0.05$ ) for meadow hawkweed and three species were indicator species for orange hawkweed (Table 6). The unadjusted  $p$  values were used for selecting species for the indicator power analysis.

Two of the species significantly associated with orange hawkweed were also significantly associated with meadow hawkweed: Richard's geranium (*Geranium richardsonii* Fisch. & Trautv.) and fringed

**Table 5** Number of transects in each susceptibility category for meadow hawkweed and orange hawkweed

	Meadow hawkweed				Orange hawkweed
	High	Mod	Low	Zero	Sum
<i>Orange hawkweed</i>					
High	9	1	1	0	11
Mod	3	3	2	0	8
Low	1	6	10	0	17
Zero	0	0	9	0	9
<i>Meadow hawkweed</i>					
Sum	13	10	22	0	

willowherb (*Epilobium ciliatum* Raf.) (Table 6). Species significantly associated only with meadow hawkweed were common selfheal (*Prunella vulgaris* L.), arrowleaf ragwort (*Senecio triangularis* Hook.), falsegold groundsel (*Packera pseud aurea* var. *pseud aurea* (Rydb.) W.A. Weber & Á. Löve), starry false lily of the valley (*Maianthemum stellatum* (L.) Link), horsetail (*Equisetum* spp.), sedge (*Carex* spp.) and ballhead ragwort (*Senecio sphaerocephalus* Greene). Alpine timothy (*Phleum alpinum* L.) was only significantly associated with orange hawkweed. Seven species significantly associated with meadow hawkweed had an IP of 0.998, indicating they frequently occurred with meadow hawkweed while also infrequently occurring in the absence meadow hawkweed (Table 6). The strongest indicators of orange hawkweed were fringed willowherb and Richard's geranium (Table 6). Alpine timothy had moderate indicator power for orange hawkweed with an IP of 0.706.

**Table 6** Indicator species results for meadow hawkweed and orange hawkweed based on chi-square values, *p* values, adjusted *p* values, and indicator power (IP)

Species	Meadow hawkweed				Orange hawkweed			
	Chi-sq	<i>p</i> value	Adj <i>p</i> value	IP	Chi-sq	<i>p</i> value	Adj <i>p</i> value	IP
Richard's geranium	16.09	< 0.001*	0.064	0.998	5.21	0.039*	0.531	0.815
Horsetail	13.06	0.002*	0.075	0.889	2.90	0.116	0.544	0.678
Arrowleaf ragwort	10.18	0.003*	0.075	0.998	1.07	0.561	0.898	0.705
Common selfheal	10.18	0.008*	0.128	0.998	4.86	0.060	0.531	0.864
Falsegold groundsel	7.44	0.016*	0.211	0.998	2.41	0.185	0.544	0.814
Starry false lily of the valley	7.44	0.021*	0.224	0.998	0.05	1.000	1.000	0.576
Ballhead ragwort	7.44	0.027*	0.240	0.998	2.41	0.173	0.544	0.814
Fringed willowherb	7.44	0.030*	0.240	0.998	8.33	0.015*	0.531	0.998
Sedge	5.37	0.045*	0.323	0.625	1.94	0.195	0.544	0.548
Alpine timothy	3.85	0.073	0.426	0.706	5.04	0.032*	0.531	0.706

Asterisk (\*) indicates significant *p* value

Species that were important for habitat typing but either occurred in fewer than three transects or were not significant in the chi-square analysis, are listed in Table 7.

#### Habitat types

Eight habitat types were identified within predicted orange hawkweed-susceptible habitat: big sagebrush/Idaho fescue, sticky purple geranium phase; big sagebrush/bluebunch wheatgrass; Idaho fescue/bearded wheatgrass, sticky purple geranium phase; alpine timothy/water sedge; lodgepole pine/Ross' sedge; subalpine fir/grouse whortleberry, grouse whortleberry phase; subalpine fir/bluejoint reedgrass; and subalpine fir/western meadow-rue. All eight habitat types, plus big sagebrush/Idaho fescue, were also identified within predicted meadow hawkweed-susceptible habitat (See Table 8

for habitat type descriptions; see Fig. 5 for habitat type examples). Habitat types ranged from xeric shrubland/grassland to wet forest. Historic infestations of orange and meadow hawkweed were found in fewer habitat types. Historic orange hawkweed infestations encompassed only four habitat types: Idaho Fescue/bearded wheatgrass, sticky purple geranium phase; subalpine fir/bluejoint reedgrass; alpine timothy/water sedge; and Lodgepole pine/Ross' sedge. Historic infestations of meadow hawkweed were found in the same four habitat types as orange hawkweed infestations, plus two additional habitat types: big sagebrush/Idaho fescue and subalpine fir/grouse whortleberry.

**Table 7** Important species for habitat typing that occurred in fewer than three transects and/or were not significant in the chi-square analysis

Common name	Scientific name
Bluejoint reedgrass	<i>Calamagrostis canadensis</i> (Michx.) P. Beauv
Idaho fescue	<i>Festuca idahoensis</i> Elmer
Sticky purple geranium	<i>Geranium viscosissimum</i> Fisch. & C.A. Mey. Ex C.A. Mey
Common yarrow	<i>Achillea millefolium</i> L
Lodgepole pine	<i>Pinus contorta</i> Douglas ex Loudon
Basin big sagebrush	<i>Artemisia tridentata</i> Nutt. ssp. <i>tridentata</i>
Wyoming big sagebrush	<i>Artemisia tridentata</i> Nutt. ssp. <i>wyomingensis</i> Beetle & Young
Huckleberry	<i>Vaccinium</i> sp. L

**Table 8** Habitat type descriptions at transects

Vegetation type	Habitat type	Elevation range (m)	Habitat description	Number of transects
Shrubland/grassland	Big Sagebrush/Idaho fescue	2100–2900	Moist shrubland of big sagebrush and Idaho fescue, occasionally with bluebunch wheatgrass. Old man's whiskers ( <i>Geum triflorum</i> Pursh), prairie sagewort ( <i>Artemisia frigida</i> Willd.), rabbit brush ( <i>Chrysothamnus</i> Nutt.) and Junegrass ( <i>Koeleria</i> Pers.) are common	5
	Big Sagebrush/Idaho fescue Phase: sticky purple geranium	2100–2900	This phase is more moist, denser ground cover of grasses and forbs; sticky purple geranium, California brome ( <i>Bromus carinatus</i> Hook. & Arn.), sulphur-flower buckwheat ( <i>Eriogonum umbellatum</i> Torr.), graceful cinquefoil ( <i>Potentilla gracilis</i> Douglas ex Hook. var. <i>gracilis</i> ), bearded wheatgrass ( <i>Elymus caninus</i> (L.) L.), Reynolds' sedge ( <i>Carex raynoldsii</i> Dewey)	5
	Big Sagebrush/Bluebunch wheatgrass	1800+	Dry shrubland with big sagebrush interspersed with bluebunch wheatgrass. Other common plants include Junegrass, Sandberg bluegrass ( <i>Poa secunda</i> J. Presl), needle and thread ( <i>Hesperostipa comata</i> (Trin. & Rupr.) Barkworth ssp. <i>comata</i> )	2
Grassland/meadow	Idaho fescue/bearded wheatgrass Phase: sticky purple geranium	2300–3000	High moisture; pocket-gopher activity, Idaho fescue nearly absent taller forbs common sticky purple geranium and graceful cinquefoil are indicators. Other common species include Gardner's yampah ( <i>Perideridia gairdneri</i> (Hook. & Arn.) Mathias), Missouri goldenrod ( <i>Solidago missouriensis</i> Nutt.), California brome and elkweed ( <i>Frasera speciosa</i> Douglas ex Griseb.), timber oatgrass ( <i>Danthonia intermedia</i> Vasey)	6
	Alpine timothy/water sedge	(From Mattson 1984, no elevation given)	Alpine timothy, water sedge, flat-top pussytoes ( <i>Antennaria corymbosa</i> E.E. Nelson), bluejoint reedgrass, smallwing sedge ( <i>Carex microptera</i> Mack.), ballhead ragwort, largeleaf avens ( <i>Geum macrophyllum</i> Willd.), Virginia strawberry ( <i>Fragaria virginiana</i> Duchesne), hookedspur violet ( <i>Viola adunca</i> Sm.), varileaf cinquefoil ( <i>Potentilla diversifolia</i> Lehm.). Moderate number of species (19–27 sp/50 m <sup>2</sup> ), moist meadow	3

**Table 8** (continued)

Vegetation type	Habitat type	Elevation range (m)	Habitat description	Number of transects
Wet Forest	Lodgepole pine/Ross' sedge	2100–2400	Lodgepole pine, silvery lupine ( <i>Lupinus argenteus</i> Pursh), Rocky Mountain goldenrod ( <i>Solidago multiradiata</i> Aiton), Ross' sedge ( <i>Carex rossii</i> Boott), Wheeler's bluegrass ( <i>Poa wheeleri</i> Vasey). (Although Despain [1990] says it is typically on steep, south-facing slopes, Steele et al. [1983] notes within YNP this habitat type is on gentle terrain at mid-elevations)	10
Open Forests	Subalpine fir/Grouse Whortleberry Phase: grouse whortleberry	2000–2400	Lodgepole pine, Engelmann spruce ( <i>Picea engelmannii</i> Parry ex Engelm.), subalpine fir, Douglas-fir, whortleberry, heart leaf arnica, elk sedge ( <i>Carex garberi</i> Fernald), mosses and lichens	7
	Subalpine fir/bluejoint reedgrass	2000–2100	Lodgepole pine, Engelmann spruce, subalpine fir, bluejoint reedgrass. Typically along stream or ponds. These transects were identified in this habitat type based on characteristics in Mattson (1984) of bluejoint reedgrass/bluejoint reedgrass habitat type which are typically surrounded by shading forest, high species diversity (23–42 sp/50m <sup>2</sup> ). Additional species present include arrowleaf ragwort, yarrow, <i>Fragaria</i> sp. L	5
	Subalpine fir/western meadow-rue	2300–2700	Lodgepole pine, subalpine fir, sweetcicely ( <i>Osmorhiza berteroi</i> DC.), pinegrass ( <i>Calamagrostis rubescens</i> Buckley), heartleaf arnica ( <i>Arnica cordifolia</i> Hook.), fireweed ( <i>Chamerion angustifolium</i> (L.) Holub ssp. <i>angustifolium</i> ), western meadow-rue, gooseberry currant ( <i>Ribes montigenum</i> McClatchie)	2

Habitat type descriptions from Despain (1990). Additional descriptions were included from Steele et al. (1983) and Mattson (1984). The elevation ranges described by Despain (1990) are included however, his GIS layer had some discrepancies. For example, the book listed big sagebrush/Idaho fescue occurring at 6800–9500 ft but a portion of the map labeled as big sagebrush/Idaho fescue was at 5400 ft

## Discussion

Remote sensing provides a cost-efficient method to monitor vegetation cover across a broad spatial extent and through time, something in situ sampling simply cannot achieve (Ahmed et al. 2020; Papp et al. 2021). Predictive modeling can help direct ground surveys to areas that are most likely to be invaded by a target species so new infestations can be treated before

becoming large and unmanageable. Weed infestations can be a multi-billion dollar expense annually for large countries (Pyšek and Richardson 2010) so early detection and swift action are key.

In this study, we developed habitat susceptibility models for meadow hawkweed and orange hawkweed using the D<sup>2</sup> method with known locations and environmental data, determined indicator species, and classified plant communities at transect

locations into habitat types to assist land managers in their efforts to monitor and control these invasive plants in the GYE. Our models were based on presence-only data to avoid assumptions that a species' absence from an area indicated no susceptibility. A challenge of using the  $D^2$  method to develop predictive habitat susceptibility models is the uncertainty if the invader is absent because the habitat is not suitable or simply because propagules have not yet arrived. It takes time for weed propagules to establish and occur in great enough abundance to be detected in a new location; the absence of the target species is not enough to confirm the location is not susceptible to invasion. Another limitation of the  $D^2$  statistic is the lack of a significance test or established method to ascertain importance of the model variables (Griffin et al. 2010), in our case, which is why several models were compared. Once the best model was selected, the model prediction error rate was improved by excluding areas where neither hawkweed species were found within a habitat type. Dense lodgepole pine stands are not suitable habitat for meadow hawkweed or orange hawkweed, so refining models without dense lodgepole pine significantly reduced the amount of land where monitoring might otherwise have been considered. Even with these limitations,  $D^2$  was the best method for presence-only data available for this study (Phillips et al. 2006).

Our meadow hawkweed model indicated 20% of the study area was in the combined moderate and high categories and 27% was in the low category. Combined moderate and high categories for orange hawkweed susceptibility represent only 9% of the study area and the low category encompassed 15%. Prediction of habitat susceptibility for meadow hawkweed was nearly twice that of orange hawkweed which, is similar to the ratio of current mapped infestations of both species. The overlap of meadow and orange hawkweed susceptibility models suggests these species can co-occur across a wide range of habitats yet, portions of their ranges are unique. Survey crews should prioritize efforts along propagule corridors in areas predicted to be highly susceptible to both species and near known populations. Roads, streams, and trails allow propagule movement into new areas so, areas along these corridors should be scoured to maximize survey efficiency.

Outside of the region where the predictive model was developed, surrogate-based approaches may be used to estimate habitat susceptibility based on site conditions that can be evaluated on the ground, such as current habitat conditions or indicator species presence (Halme et al. 2009; Thuiller et al. 2012; Jones et al. 2018). Predicted meadow hawkweed susceptibility encompassed nine habitat types, and predicted orange hawkweed susceptibility encompassed eight of the same habitat types, excluding big sagebrush/Idaho fescue (Table 8). However, historic hawkweed infestations were only recorded in six of these habitat types. The three newly identified potential habitat types: big sagebrush/Idaho fescue, sticky purple geranium phase; big sagebrush/bluebunch wheatgrass; and subalpine fir/western meadow-rue suggest more of GYE is susceptible to meadow and orange hawkweed than previously thought. These habitat types should be prioritized for future monitoring. Invasive hawkweeds are not typically found in dry shrub-steppe grasslands (Wilson 2007); however, meadow hawkweed was observed in a big sagebrush/Idaho fescue habitat type. Orange hawkweed may be found in sites with higher moisture than described in this study as it is known to infest bogs and cranberry fields (Jorgensen and Nauman 1994). Wetter communities like willow and sedge dominated marshes were not surveyed in this study.

Indicator species are another surrogate-based approach we addressed in this study as training new field crews to identify key species may be more realistic than training them to identify broad habitat types. Alpine timothy was the only unique indicator of orange hawkweed. Thus, survey crews could mark locations of this plant to identify likely areas for invasion. Richard's geranium and fringed willowherb were strong indicators for both hawkweeds. If crews encounter those species they should methodically search the area.

Meadow hawkweed and orange hawkweed have the potential to impact a large portion of the GYE across a range of moisture conditions and habitat types. Early detection is vital because patch size of invaded areas grow at faster rates later in their establishment, reducing the likelihood of control (Rejmánek and Pitcairn 2002; Shiferaw et al. 2019). Prioritizing monitoring in habitats suitable for an invasive species can reduce the overall area that is surveyed and ultimately save land managers time and money on ground

**Fig. 5** Meadow and orange hawkweed habitat types. **a** basin big sagebrush/Idaho fescue, **b** basin big sagebrush/Idaho fescue, sticky purple geranium phase, **c** Lodgepole pine/sedges, **d** subalpine fir/bluejoint grass, **a** is only susceptible to meadow hawkweed, **b-d** are susceptible to meadow hawkweed and orange hawkweed



surveys and management efforts, while preserving the GYE in its most natural state. Using the susceptibility models to direct ground surveys within the predicted habitat types that have not yet had documented invasions could be instrumental in locating new infestations, avoiding further expansion, and refining models for these two invasive hawkweed species. Susceptibility models should be updated when new locations of invasive plants are discovered and when updated GIS data are available. When models are not available for an area, understanding the habitat types or indicator species associated with the species of interest can guide survey efforts. While our study focused on plant invasion susceptibility, the same process can be used for any species of concern.

**Acknowledgements** We would like to thank the Greater Yellowstone Coalition and the USDA-Forest Service Special Technology Development Program who jointly funded this project. We are also grateful for the support and contributions of Steven Bekedam with USDI Park Service, Brianna Slothower, former undergraduate student at University of Idaho and Larry W. Lass, retired research scientist at University of Idaho.

**Funding** Funding for the research came from the Greater Yellowstone Coalition and the USDA-Forest Service Special Technology Development Program 15-CA-11010000-015. The authors have no financial or non-financial interests to disclose. All authors contributed to study conception and design, and writing the manuscript, in addition, Christie H. Guetling and Julia Piaskowski conducted data analysis.

## Declarations

**Competing interests** The authors have not disclosed any competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- 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* 9:32
- Adhikari A, Rew LJ, Mainali KP, Adhikari S, Maxwell BD (2020) Future distribution of invasive weed species across the major road network in the state of Montana, USA. *Reg Environ Change* 20:60
- Balch JK, Bradley BA, D'Antonio CM, Gomez-Dans J (2013) Introduced annual grass increases regional fire activity



- across the arid western USA (1980–2009). *Glob Change Biol* 19:173–183
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 57:289–300
- Clark Labs (2017) IDRISI GIS software, TerrSet version 18.31. Clark Labs, Clark University, Worcester, Massachusetts, USA
- Daubenmire R (1984) Viewpoint: ecological site/range site/habitat type. *Rangelands* 6(6):263–264
- Despain DG (1990) Yellowstone vegetation: consequences of environment and history in a natural setting. Roberts Rinehart, Boulder
- Eitel JUH, Vierling LA, Litvak ME, Long DS, Schulthess U, Ager AA, Krofcheck DJ, Stoeckel S (2011) Broadband, red-edge information from satellites improves early stress detection in a New Mexico Conifer woodland. *Remote Sens Environ* 115:3640–3646
- ESRI Software. 2018. ArcMap Version 10.6.1. Redlands, California, USA.
- European Space Agency (2018) Sentinel-2—T12TWQ. U.S. Geological Survey, Sioux Falls, SD USA. <https://earthexplorer.usgs.gov/>
- Fenouillas P, Ah-Peng C, Amy E, Bracco I, Dafreville S, Gosset M, Ingrassia F, Lavergne C, Lequette B, Notter J, Pausé J, Payet G, Payet N, Picot F, Pougavanon N, Strasberg D, Thomas H, Triolo J, Turquet V, Rouget M (2021) Quantifying invasion degree by alien plant species in Reunion Island. *Austral Ecol* 46:1125–1137
- Fleishman DE, Thomas JR, Mac Nally R (2005) Using indicator species to predict species richness of multiple taxonomic groups. *Soc Conserv Biol* 19:1125–1137
- Ghorbanian A, Mohammadzadeh A, Jamali S (2022) Linear and non-linear vegetation trend analysis throughout Iran using two decades of MODIS NDVI imagery. *Remote Sens* 14:3683
- Gibson CE (2005) Fire history polygons for northern Rockies 1889–2003. U.S. Department of Agriculture: Geospatial Data. [https://www.fs.usda.gov/detailfull/r1/landmanagement/gis/?cid=fsp5\\_030970](https://www.fs.usda.gov/detailfull/r1/landmanagement/gis/?cid=fsp5_030970)
- Griffin SC, Taper ML, Hoffman R, Mills LS (2010) Ranking Mahalanobis distance models for predictions of occupancy from presence-only data. *J Wildl Manag* 74:1112–1121
- Halme P, Mönkkönen M, Kotiaho JS, Ylisirniö A, Markkanen A (2009) Quantifying the indicator power of an indicator species. *Conserv Biol* 23:1008–1016
- Holland PG, Steyn DG (1975) Vegetational responses to latitudinal variations in slope angle and aspect. *J Biogeogr* 2:179–183
- John R, Chen J, Giannico V, Park H, Xiao J, Shirkey G, Ouyang Z, Shao C, Laforteza R, Qi J (2018) Grassland canopy cover and aboveground biomass in Mongolia and inner Mongolia: spatiotemporal estimates and controlling factors. *Remote Sens Environ* 213:34–38
- Jones L, Norton N, Prather T (2018) Indicators of ventenata (*Ventenata dubia*) invasion in sagebrush steppe rangelands. *Invasive Plant Sci Manag* 11:1–9
- Jorgensen EE, Nauman LE (1994) Disturbance in wetlands associated with commercial cranberry (*Vaccinium macrocarpon*) production. *Am Midl Nat* 132:152–158
- LANDFIRE: LANDFIRE Disturbance Layer (2019, Feb—last updated) U.S. Department of Interior, Geological Survey. [Online]. Available: <http://landfire.cr.usgs.gov/viewer/> [2019, Nov 6]
- Lass LW, Prather TS, Shafii B, Price WJ (2011) Tracking invasive weed species in rangeland using probability functions to identify site-specific boundaries: a case study using yellow starthistle. In: Clay SA (ed) GIS applications in agriculture volume three: invasive species. CRC Press, Boca Raton, pp 277–299
- Masemola C, Cho MA, Ramoelo A (2020) Sentinel-2 time series based optimal features and time window for mapping invasive Australian native *Acacia* species in Kwa-Zulu Natal, South Africa. *Int J Appl Earth Obs Geoinf* 93:102207
- Matas-Granados L, Pizarro M, Cayuela L, Domingo D, Gómez D, García MB (2022) Long-term monitoring of NDVI changes by remote sensing to assess the vulnerability of threatened plants. *Biol Cons* 265:109428
- Mattson DJ (1984) Classification and environmental relationships of wetland vegetation in central Yellowstone National Park, Wyoming. Master's thesis, University of Idaho
- National Park Service (2019) Invasive and non-native species. <https://www.nps.gov/subjects/invasive/plants.htm> Accessed 23 Aug 2019
- Papp L, van Leeuwen B, Szilassi P, Tobak Z, Szatmári J, Árvai M, Mészáros J, Pásztor L (2021) Monitoring invasive plant species using hyperspectral remote sensing data. *Land* 10:29
- Pyšek P, Richardson DM (2010) Invasive species, environmental change and management, and health. *Annu Rev Environ Resour* 35:25–55
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* 190:231–259
- PRISM Climate Group (2020) Oregon State University <http://www.prism.oregonstate.edu/normals/> created 2020
- R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Rejmánek M, Pitcairn MJ (2002) When is eradication of exotic pest plants a realistic goal? In: Veitch CR, Clout MN (eds) Turning the tide: the eradication of invasive species. International Union for Conservation of Nature, IUCN Species Survival Commission, Gland
- Rew LJ, Maxwell BD (2006) Stratified random sampling method. In: Rew LJ, Pokorny ML (eds) Inventory and survey methods for nonindigenous plant species. Bozeman, MT Montana State University Extension, pp 49–55
- Shafii B, Price WJ, Prather TS, Lass LW, Thill DC (2003) Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics. *Weed Sci* 51:748–751
- Shafii B, Price WJ, Prather TS, Lass LW, Thill DC (2004) Using landscape characteristics as prior information for Bayesian classification of yellow starthistle. *Weed Sci* 52:948–953
- Shiferaw H, Bewket W, Alamirew T, Zeleke G, Teketay D, Bekele K, Schaffner U, Eckert S (2019) Implications of land use/land cover dynamics and *Prosopis* invasion on

- ecosystem service values in Afar Region, Ethiopia. *Sci Total Environ* 675:354–366
- Steele R, Cooper SV, Ondov DM, Roberts DW, Pfister RD (1983) Forest habitat types of eastern Idaho-western Wyoming. [https://www.fs.fed.us/rm/pubs\\_int/int\\_gtr144.pdf](https://www.fs.fed.us/rm/pubs_int/int_gtr144.pdf)
- Thuiller W, Gassó N, Pino J, Vilà M (2012) Ecological niche and species traits: key drivers of regional plant invader assemblages. *Biol Invasions* 14:1963–1980
- Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens Environ* 8:127–150
- Wilson LM (2007) Key to identification of invasive and native hawkweeds (*Hieracium* spp.) in the Pacific Northwest revised March 2007. BC Min For Range, For Prac Br, Kamloops, B.C. [https://www2.gov.bc.ca/assets/gov/environment/plants-animals-and-ecosystems/invasive-species/publications/key\\_to\\_identification\\_of\\_invasive\\_and\\_native\\_hawkweeds\\_in\\_the\\_pacific\\_northwest.pdf](https://www2.gov.bc.ca/assets/gov/environment/plants-animals-and-ecosystems/invasive-species/publications/key_to_identification_of_invasive_and_native_hawkweeds_in_the_pacific_northwest.pdf)
- Wilson LM, Callihan RH (1999) Meadow and orange hawkweed. In: Sheley RL, Petroff JK (eds) *Biology and management of noxious rangeland weeds*. Oregon State University Press, pp 238–248
- Woodcock T, Evans C, Laxson C, Tucker R, Allen J, Mihuc J, Allen E, Mihuc T (2008) Land use designation and vegetation community structure in the Adirondack uplands (New York, USA). *Appl Veg Sci* 11:509–520
- Yellowstone National Park (2016) *Yellowstone Resources and issues handbook*. Yellowstone National Park. [https://www.nps.gov/yell/planyourvisit/upload/RI\\_2016\\_FINAL\\_Intro\\_web.pdf](https://www.nps.gov/yell/planyourvisit/upload/RI_2016_FINAL_Intro_web.pdf) Accessed 20 Aug 2019

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.