Final Report

A survey of non-indigenous plant species in the northern range of Yellowstone National Park, 2001-2004.

Lisa J Rew, Bruce D Maxwell, Frank L Dougher, Richard Aspinall,
Tad Weaver and Don Despain, Montana State University,
Bozeman, MT 59717

Acknowledgements

This project was funded by the Yellowstone Inventory and Monitoring program, under the supervision of Cathie Jean. Sincere thanks to Lane Cameron for initiating this project and his continued interest and support, even after leaving for the Channel Islands National Park. Thanks also to Mary Hektner and Roy Rankin for their non-indigenous plant species and logistics insight; to Jennifer Whipple for her help with plant identification; and Ann Rodman and the GIS lab in Yellowstone National Park for sharing their GIS data layers.

Special praise and thanks is reserved for the field crew particularly Jerad Corbin, Mara Johnson and Amanda Morrison who worked on the project for a couple of summers but also to Judit Barroso, Matthew Hulbert, Jeremy Gay, Rebecca Kennedy, Erik Lehnoff, Ben Levy, Nathaniel Ohler, Charles Repath, John Rose and Kit Sawyer. All of them exerted considerable energy and enthusiasm for plant identification, data collection and hiking all sorts of terrain in all sorts of conditions; the project would not have achieved as much without such dedication.

Table of Contents

Executive Summary

An inventory/survey of the occurrence and location of non-indigenous plant species (NIS) is seen as the first of three phases in their management. A survey needs to be designed to obtain an unbiased assessment of NIS extent over the whole area of interest. Data from the inventory or survey can then be used as baseline data to select patches (populations) for monitoring. Monitoring methods should aim to evaluate changes in spatial and temporal extent of NIS, impacts of NIS on the ecosystem, and impacts of management on NIS and surrounding ecosystem. The survey data when developed into probability of occurrence maps can also be used to strategically survey for new populations. Only after following these steps to estimate NIS plant occurrence, extent, dynamics and impacts, can effective management strategies be developed for the currently occurring species as well as new species, that may invade National Parks and other wilderness/natural areas.

The overall aim of this project (2001-2004) has been to survey the occurrence (presence/absence) of NIS within the northern elk winter range of YELL. Covering an area of 152,785 ha, the northern range is too large to look for NIS species over its entirety. In 2001, we focused on identifying which sampling methods provide the most representative sample of species occurrence in the landscape. This objective was achieved through computer simulation and field sampling. As a result we adopted a stratified-random sampling methodology that would maximize our ability to predict occurrence of the NIS. In field seasons 2001 through 2003, our crews sampled 295 transects in the area between Gardiner on the west and Silvergate on the east and from the northern boundary of YELL to as far south as Tower Falls.

Sixty-three NIS are on the YELL target list. Over the four year study we have observed and recorded 29 of them. Only nine of the observed NIS occurred on more than 1% of the surveyed area (in descending order of occurrence): *Phleum pratense*, *Poa pratensis*, *Bromus tectorum*, *Cirsium arvense*, *Bromus inermis*, *Alyssum desertorum*, *Linaria dalmatica, Elymus repens* and *Cardaria chalepensis*. Maps of where these nine species were located on our transects are provided in this report.

While such maps are useful they do not provide any information on areas which were not surveyed. Therefore, the 2001-2003 NIS data were analyzed for correlations with human disturbances (roads and trails) and environmental variables to aid in the prediction of NIS occurrence in areas not sampled. Thus, we have generated maps of probability of occurrence for the nine most frequent species, which are also in this report. The probability of occurrence maps provide insight as to where a particular NIS is most likely to be, based on correlations with variables in the model. However, the species may not be there yet, which makes model evaluation difficult. In 2004, we stratified our sampling transects based upon the results of our probability maps to validate the predictive model results. The results all showed the correct trend, i.e. the target NIS was found increasingly more often as the predicted occurrence also increased. However, the model tended to predict a higher percentage occurrence than we observed in the field. As one would expect some models and observations agreed more (*e.g. Phleum pretense* and *Bromus tectorum*) than others (*e.g. Bromus inermis*). Statistical analysis of the model predictions suggested the model fits were good. The fact we observed less NIS in the field than the model predicted, may suggest that the species has not arrived or reached its full potential in the environment yet. We feel that it may be better to over-

predict then under-predict when the purpose of the study was to identify all locations of each species.

This study was not intended to estimate the extent of populations (density or hectares infested), but rather the relative likelihood of the occurrence of NIS in any given area within the northern range of YELL. Probability of occurrence maps provide the current best prediction of where target species are more of less likely to be. While the accuracy of the predictions were better for some species than others, we believe they provide managers with information on areas which may not have been visited and information on which areas should be targeted to observe new populations when time and money allows. The data also provides information which can be used to select populations for monitoring. Monitoring select populations provides the opportunity to better understand where and under what conditions NIS are spreading and impacting the ecosystem; and, should be the next step taken with this project. The survey and monitoring information could then be used to help target specific species, populations and environments on a larger scale so that limited time and effort could be spent where it will have most effect.

Project description

Introduction

The United States Department of Interior National Park Service is required by law to keep the 34 million hectares designated as National Parks classified as "natural areas". Natural areas must be "unaltered by human activities" as much as possible (U.S. National Park Service, 1996). Maintaining the Parks as "natural areas" includes removal of non-native plant species. The definition of non-native is "any animal or plant species that occurs in a given location as a result of direct, indirect, deliberate or accidental actions by humans" (U.S. National Park Service, 1996). This definition permits the user to recognize and distinguish between changes to animal and plant distributions caused by natural processes and human influences. In reality this statement needs some further clarification. "Human influence" really refers to disturbance by white settlers, more so in the past century and most specifically in the last 50 years.

Many countries have designated specific areas as "wilderness" or "natural ecosystems" and seek to preserve these in their "pristine" state, however pristine is defined. Taking this desire to "protect and retain" such areas, one can argue from the ecological purist point of view, that all non-indigenous species should be removed. However, this is currently impossible from a practical standpoint. In most cases we do not know which non-indigenous species are present within an ecosystem, their frequency or their distribution pattern; how much their distribution is changing and finally what impact they are having on the endemic ecosystem. It is only armed with all of this information that land managers can effectively target and manage non-indigenous species populations.

The language used to describe the presence and impact of non-indigenous plant species (NIS) is often very emotive: "aggressive non-indigenous plants, which spread quickly into natural areas replacing native flora and reducing habitat for native flora and fauna". Often the simple presence of a NIS is stated as proof enough of present or future environmental damage, particularly if it is a highly competitive species and/or if the increase in the non-indigenous species is associated with the decline of native species. However, Weaver *et al.,* (2001) in a study of the northern Rocky Mountains found that of the 29 most commonly found exotic species the majority were intentionally introduced (*e.g. Phleum pratense* and *Poa pratensis*) and none of the most common were generally considered a noxious weed.

A number of studies have shown that when non-indigenous species are introduced to environments and ecosystems different from those in which they evolved, they may disrupt the ecosystem processes and alter biological diversity (*e.g.* Braithwaite & Lonsdale, 1989; Hobbs & Mooney, 1991; see Davis *et al.*, 2000 and Mack *et al.*, 2000 for reviews). Invasion by a new species is influenced by three factors:

1. ecosystem properties, which could be related to the level or frequency of disturbance;
2. number of propagules entering a new environment (propagule pressure); and,
3. the properties of the invading species (Lonsdale, 1999).

Davis *et al.* (2000) and Davis and Pelsor (2001) offer a new theory, that the fluctuation of resource availability is a key factor in controlling invasion. This theory allows for the integration of resource availability with disturbance and fluctuating environmental conditions.

Disturbance is often suggested as a key factor in enhancing the probability of NIS establishment in native plant communities. Natural disturbance has a variety of biotic and geomorphic causes including soil disturbance by fauna, weather related events such as mudflows, floods, wind, fire and geological events such as landslides. Fire is sometimes a quasi-human disturbance if management practices suppress, contain or intentionally ignite them, or if fires are ignited accidentally or intentionally by vandals, whichever way, the natural occurrence of fires has usually been altered. Human disturbance includes construction and use of roads and trails, buildings, utility corridors and campgrounds.

As stated above, the National Park Service has a mandate to preserve the natural systems under their control (National Park Service Organic Act of 1916). There are several phases necessary to achieve this objective:

Phase 1	creating an inventory/survey (documenting occurrence);
Phase 2	monitoring (quantifying changes in distribution, abundance or impact); and,
Phase 3	control or management of non-indigenous species on a large scale.

To a certain extent these phases can be performed concurrently (Fig. 1). The aim of the current project is Phase 1, development of an inventory/survey program.

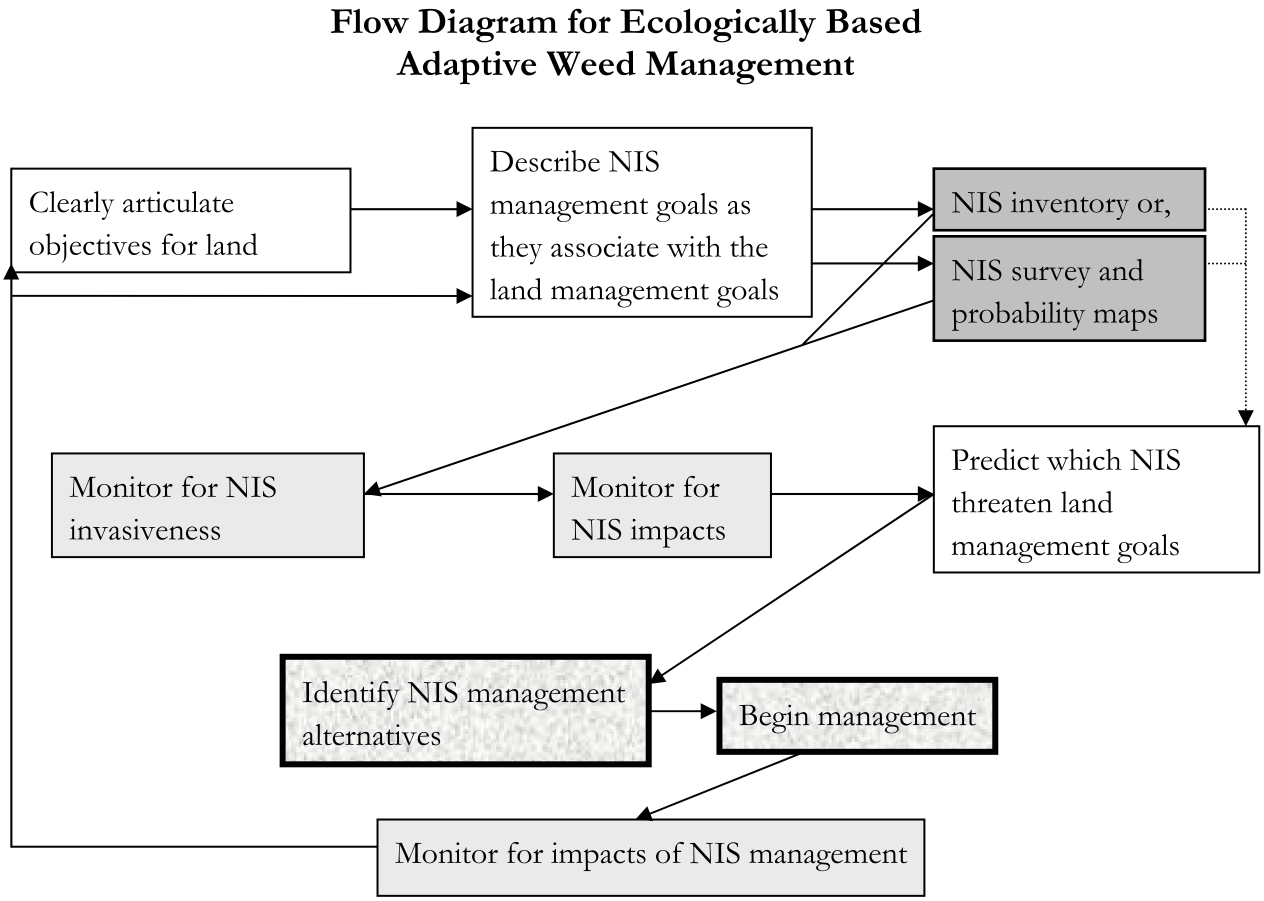


Fig. 1. Flow diagram for ecologically based adaptive weed management.

The problem with developing an inventory/survey

Conducting an inventory/survey of non-indigenous plants in a large region where many of the non-indigenous species have infrequent occurrence is a difficult task. The definition of an inventory is a list of all NIS species and their locations in a delineated management area when the entire area can be observed. A survey is defined as a list of NIS species and their locations in a delineated management area when all of the area cannot be observed. A survey requires careful consideration of sampling methods. As the area of the northern range if so large and we cannot sample the entire area we are by definition completing a survey. This term will be used from now on.

Considering the ultimate use of the survey is essential in the design. In the case of the National Park Service, management of NIS is the objective, but because the NIS are relatively infrequent and spread over large areas, it will never be possible to manage all NIS or all their occurrences. Thus, a survey of the NIS and the subsequent assessment of population and metapopulation dynamics must have the objective of creating an unbiased sample in order to prioritize management of those metapopulations that pose the greatest threat to the ecosystem. An unbiased sample requires locating populations or metapopulations over the extent of the environments where they may exist. Therefore, we are reliant upon a survey that maximizes the probability of finding the non-indigenous invasive species (NIS) and simultaneously builds a data set from which models that predict NIS occurrence can be developed to ensure that we represent, through observation or prediction, all environments where the NIS may be found. It is tempting to combine survey and population assessments. If the survey is strictly a means of finding the NIS so that they can be killed, then an estimate of each metapopulation extent in the survey could serve the purpose of knowing approximately how much herbicide/hand-weeding will be required to control the observed metapopulations. However, if the intent of the survey is to maximize the potential of knowing where all of the NIS are located and subsequently using the survey to select a random sample of metapopulations to monitor for an unbiased determination of population dynamics and prioritization of management, then the survey approach that we are suggesting is most appropriate.

Methods

Study area

Yellowstone National Park (YELL) covers an area of 899,121 ha (2,220,829 ac). Approximately 1265 plant species have been recorded in YELL of which 187 (15%) are non-indigenous plant species (Whipple, 2001). This study concentrates on the area within the northern elk winter range of the Park (152,785 ha, 377,379 ac).

Prior knowledge of non-indigenous species occurrence in the study area

The relative proportional importance of the different forms of disturbance and environmental factors on non-indigenous species establishment and survival has not been quantified. The general perception from the National Park staff involved with NIS surveys and members of this research group was that most infestations occur close to roads, trails and human habitation. From the data collected by YELL park staff in 1998, it was calculated that 278 of 422 (66%) NIS occurrences were less than 100 m from roads or trails, and all observations were made less than 500 m from roads or trails. These data were not collected using a formal sampling strategy and the sites searched were biased by their proximity to roads and trails. Therefore, this information was treated as anecdotal and although considered, the data were not used for any subsequent analysis.

It is assumed that most of the species we are targeting are at a low frequency within the landscape and therefore collecting large numbers of observations is important to provide a reliable estimate of the species occurrence. A large sample combined with an appropriate strategy for estimating geographic distribution is also necessary if the goal is to estimate the distribution of the NIS in the landscape. Survey design is, therefore, a tradeoff between collecting a sufficiently large sample to provide reliable estimates of occurrence, and using a sampling strategy that is efficient for both a) field work and b) estimating the geographic distribution of the species.

To ensure the best use of the limited funds and time available in the field, a desktop study was conducted to develop the most effective sampling regime. This was performed in ESRI ArcView© GIS using routines developed by Aspinall and Dougher. This implemented several different sampling strategies including simple random sample, random walk, random transects, transects normal to specified linear features, stratified random sampling and regular grid sampling. Additionally, different sampling intensities were evaluated for different infestation levels (frequencies) of NIS.

The simulations and sampling strategies implemented within the GIS allowed us to evaluate which sampling strategy provides the highest number of sample points for the shortest time in the field and, also provides geographic coverage necessary for estimating distribution of the NIS. Random points or grid intersections for example, are not as efficient for collecting data as random walks or transects since time used moving from one survey location to another location is not used for data collection. Surveying along transects allows data to be collected continuously and a large sample size be generated. Additionally, surveying along transects allows changes in underlying environmental variables to be recorded. This is important for estimating the geographic distribution of the species from the sample data. This work has subsequently been accepted to the Biological Invasions journal.

If the occurrence of a target species is known to be correlated with an environmental variable, we can stratify the sampling scheme on that variable and improve our probabilities of finding the target (Hirzel and Guisan, 2002). We accepted the assumption that human disturbance in the form of roads and trails increases the chance of finding NIS, and stratified our sampling using this variable. However, to test this hypothesis we also needed to sample away from roads and trails. Therefore, transects established perpendicular to roads and trails were accepted as the most effective sampling methodology. The use of 2000 m transects allows the importance of other factors to be evaluated, since each transect is sufficiently long to cross a number of cover or habitat types and other environmental transitions.

Collection of field data

In 2001, the position of each transect was randomly selected along a road or trail, prior to arrival in the field, and ran perpendicular to roads or trails. This approach needed to be partially modified for the 2002 and 2003 field seasons to ensure a similar number of data points were collected at all distances from roads and trails. The location of transects was still randomly generated but within a set of confines:

- Starting on a road and finishing 2000 m from all roads but at all times the transect runs more than 2000 m from any known trail
- Starting on a trail and finishing 2000 m from all trails but at all times the transect runs more than 2000 m from any known road
- Starting on a road or trail and finishing 2000 m from all roads and trails.

As an additional confine, transect lines were generated in pairs separated no less than 100 m (when possible) and by no more than 500 m. Paired transects were used to maximize surveying time while in the field; crews would survey moving away from roads/trails on one member of a transect pair, and survey moving back towards the road/trail on the second member of the transect pair.

Non-indigenous plant species occurrence data from 2001-2003 were used to generate predictive models for five of the most frequently observed NIS; *Phleum pretense, Bromus tectorum, Cirsium arvense, Bromus inermis,* and *Linaria dalmatica.* In order to acquire a dataset suitable for validating the predictive models, transects for 2004 were stratified not on roads/trails, but on the probability levels generated by these models. Transects were generated separately for areas of higher to lower probability of NIS occurrence. Apart from the different criteria in transect delineation; all survey methods used in the 2004 season were consistent with prior seasons' methods.

Transects were walked and survey observations were made within a 10 m wide swath. Information was gathered when a target NIS was located, the habitat type changed or a disturbance feature was reached. The habitat classifications were based on the classifications devised by D. Despain and incorporated into the YELL GIS layers. For each NIS infestation, width and length along transect were estimated by pacing or visual determination from a central location within a patch when the patch size was small enough. When the length of the patch was too large to visually perceive or pace from a single location, the start and end of the patch length along the transect was recorded with GPS. The total length of these start and end point patches were determined by data analysis in post-processing. Patch widths were estimated up to a maximum width of 64 m.

Transect observations and location data were recorded on to GPS by two-person survey crews. Trimble Pro XR receivers and GeoExplorer® 3 GPS units were used and the data post-processed to improve spatial accuracy. The coordinate system and projection used was Universal Transverse Mercator (UTM) Zone 12N, WGS 1984 Datum. This projection and datum are the same as used for GIS data maintained by YELL Center for Resources, and the Greater Yellowstone Area Spatial Data Clearinghouse managed and maintained by the Geographic Information and Analysis Center (GIAC) at Montana State University.

From 2002 through 2004, all data were collected directly into a data dictionary on a GeoExplorer® 3 unit that contained the same data fields as used in 2001, plus additional information on patch parameters and fields required by North American Weed Mapping Association (NAWMA, 2005). These included the location of target species, with additional information on density (in predefined classes of 0, 0-1, 1-11, 12-32, 33-100, 101-316, 317-1000 and >1000 m^{-2}), percentage cover m^{-2}, length (m) and width (m) of infestation, and spatial pattern type. Percent cover estimates were collected in accordance with NAWMA. Environmental variables included climax habitat type, dominant vegetation cover species (up to twelve species), aspect, topography and disturbance. Additional data fields included NAWMA's "Values at risk" and "Ecological status of site/survey unit" and, time and date. Minor alterations in the structure of the GPS data dictionary were made with each successive field season to improve the efficiency of data collection and processing. Where these alterations resulted in discrepancies with the older data structures, the previous data files were reformatted to bring them into compliance with the newer format. Fields that were not collected but could be added to the database at a later stage include information about the site/region, I&M network, park unit, state, county, ownership, type of survey, and non-indigenous plant species and ITIS code, all of which can be added to the database in the office.

Though the survey method applied in the field was a continuous sampling along each transect, the digital representation of the survey observations was in the form of GPS points marking NIS presences, habitat-type transitions and disturbances. Using custom-made applications for Arcview (Version 3.2) and Microsoft Excel, these data points were transformed into continuous linear transect data. The continuous data were then partitioned into discrete sample points at a regular 10 m sampling interval. These 10 m x 10 m sample points were attributed with the presence or absence of each observed NIS, as well as associated environmental variables (*e.g.* burned or unburned condition, elevation, distance to nearest road, etc.). These 10 m interval data were the final data format for analyses. This included calculating percent occurrence of each species; calculated as the percent of 10 m x 10 m data in which the species was present versus absent for the entire area sampled.

These 10 m data were also analyzed with generalized linear model regressions to derive species/environment relationships. Generalized logistic models (GLM) were used because the dependent variable – the NIS species data - is binary (presence/absence) data. The best model was determined with backward stepwise procedure using Akaike's Information Criterion (Akaike 1977; Burnham and Anderson 1998). In order to generate predictive NIS occurrence maps for the northern range of YELL, we needed to relate the presence/absence data to spatial variable data covering the entire study area. Therefore, we used environmental data derived from digital elevation maps (10 m resolution) and remotely sensed data (30 m resolution). The topographic data included aspect, elevation, slope and solar insolation (the latter using the method of Swift (1976)). Distance from roads and trails were also used in the GLM and calculated from data layers within the GIS database. LANDSAT Enhanced Thematic Mapper (ETM) remote sensing data, acquired July 13th 1999, were included as individual spectral bands and as an unsupervised classification layer (n=128), generated using ISOCLUSTER in ERDAS Imagine. These data had accuracies of between 63 and 100% for individual land cover classes (Legleiter et al. 2003). These remotely sensed data were used to provide some information on the different reflectance of the vegetation; they were used instead of the dominant vegetation GIS layer due to their finer resolution. Other environmental data, such as burned or unburned condition, and presence/absence of trees, were obtained from the Park Service and converted into raster format as needed to work within the framework of the predictive model. GLM Analyses were performed in S-PLUS 2000

Probability of occurrence predictions and maps of the target species were generated using coefficient values from the GLM applied to continuous spatial variables in rasterized format, using an extension we wrote in Arcview. The extension generated the logit of the GLM by summing the product of each variable in the model and its coefficient value, plus the beta intercept value. More explanation of the GLM analysis and generation of prediction maps can be found in Appendix 3.

Results

Survey data

From 2001 through 2003, 295 transects were walked in the northern range with an overall sampled length of 528,960 m x 10 m wide (Fig. 2). Sixty-three species listed on the YELL priority list (Appendix 1) were targeted. Of these 63 species, 29 were observed in the field (Table 1). Nine of these species were observed to occur over greater than 1% of the surveyed area, as analyzed from the 10 m sample points: *Phleum pratense*, *Poa pratensis*, *Bromus tectorum*, *Cirsium arvense*, *Bromus inermis*, *Alyssum desertorum*, *Linaria dalmatica, Elymus repens* and *Cardaria chalepensis*.

In field season 2004, 80 transects were surveyed, covering 81,660 m x 10 m (Fig. 3) Twenty-one of the 63 target species were observed in the field. Since the 2004 transects were stratified on high to lower probability of occurrence of *Phleum pratense, Bromus tectorum, Cirsium arvense, Bromus inermis* and *Linaria dalmatica*; the occurrence rates of the 21 species observed in this season are not considered to be representative of occurrence rates throughout the northern range, and are not included in the species occurrence statistics (Table 1).

NIS Distribution and Probability of Occurrence Maps

The 2001-2003 field data for the seven most extensively occurring NIS: *Phleum pratense* (Fig. 4), *Poa pratensis* (Fig. 5), *Bromus tectorum* (Fig. 6), *Cirsium arvense* (Fig. 7), *Bromus inermis* (Fig. 8), *Alyssum desertorum* (Fig. 9) and *Linaria dalmatica* (Fig. 10) have been used to characterize the distribution of NIS occurrences relative to spatial variables, such as distance to roads (Fig. 11). The analysis of the relationship between NIS occurrence and environmental variables by statistical methods, such as generalized linear regression, allows us to extrapolate those relationships throughout the northern range of YELL through predictive modeling.

Probability of occurrence maps were created for the seven most frequently occurring NIS using the coefficient values provided by the GLM analyses (Table 2): *Phleum pratense* (Fig. 12), *Poa pratensis* (Fig. 13), *Bromus tectorum* (Fig. 14), *Cirsium arvense* (Fig. 15), *Bromus inermis* (Fig. 16), *Alyssum desertorum* (Fig. 17) and *Linaria dalmatica* (Fig. 18). The value of each cell in the output raster, ranging from zero to one, represents the probability that the target species could be present within the area defined by that cell. In this study the raster cell size was 10 m by 10 m.

Table 1. Number of observations sampled and average length, width, percentage cover and density of non-indigenous pland species, and the percentage occurrence of each species within the area studied, for 2001-2003 data.

Species	Number of observations	average of observed infestations				% occurrence
		Length (m)	Width (m)	% cover (per m^2)	density (plants/m^2)	
Phleum pratense	822	164.73	38.48	2.26%	4.38	22.875%
Poa pratensis	315	192.08	61.93	2.99%	6.40	10.577%
Bromus tectorum	316	111.65	26.02	9.53%	19.94	6.275%
Cirsium arvense	501	50.02	28.79	5.87%	4.98	4.980%
Bromus inermis	255	90.55	37.38	14.01%	19.27	4.399%
Alyssum desertorum	115	229.23	44.25	9.06%	23.32	3.923%
Linaria dalmatica	296	98.91	16.03	2.54%	2.60	3.439%
Elymus repens	40	572.88	67.63	6.95%	8.30	2.565%
Cardaria chalepensis	48	219.06	16.69	2.98%	23.02	2.000%
Trifolium repens	36	127.07	27.56	13.00%	7.11	0.900%
Cardaria draba	31	141.65	16.55	5.52%	17.10	0.849%
Poa palustris	13	348.92	44.23	1.85%	4.77	0.586%
Cynoglossum officinale	87	31.75	3.36	3.18%	1.26	0.548%
Trifolium hybridum	75	28.78	31.72	18.73%	6.69	0.484%
Poa bulbosa	14	176.96	40.07	5.43%	31.21	0.459%
Melilotus officinale	115	10.15	13.26	1.19%	1.02	0.234%
Cardaria spp.	17	123.76	11.53	5.41%	12.29	0.040%
Verbascum thapsus	7	20.12	20.86	1.86%	4.00	0.036%
Poa annua	8	20.38	21.00	1.38%	3.50	0.032%
Cirsium vulgare	10	4.57	3.80	12.80%	5.40	0.025%
Centaurea maculosa	11	10.00	0.00	0.00%	0.00	0.021%
Chrysanthemum leucanthemum	9	13.33	0.56	0.11%	0.00	0.021%
Potentilla recta	5	4.60	24.40	1.80%	1.00	0.017%
Medicago lupulina	8	2.13	47.88	4.00%	3.00	0.013%
Trifolium aureum	5	1.20	1.60	23.00%	6.00	0.011%
Poa compressa	2	1.50	1.00	10.50%	6.00	0.006%
Carduus nutans	1	6.00	3.00	1.00%	1.00	0.002%
Cardaria pubescens	1	1.00	1.00	1.00%	6.00	0.002%
Hieracium floribundum	1	3.00	30.00	1.00%	6.00	0.002%

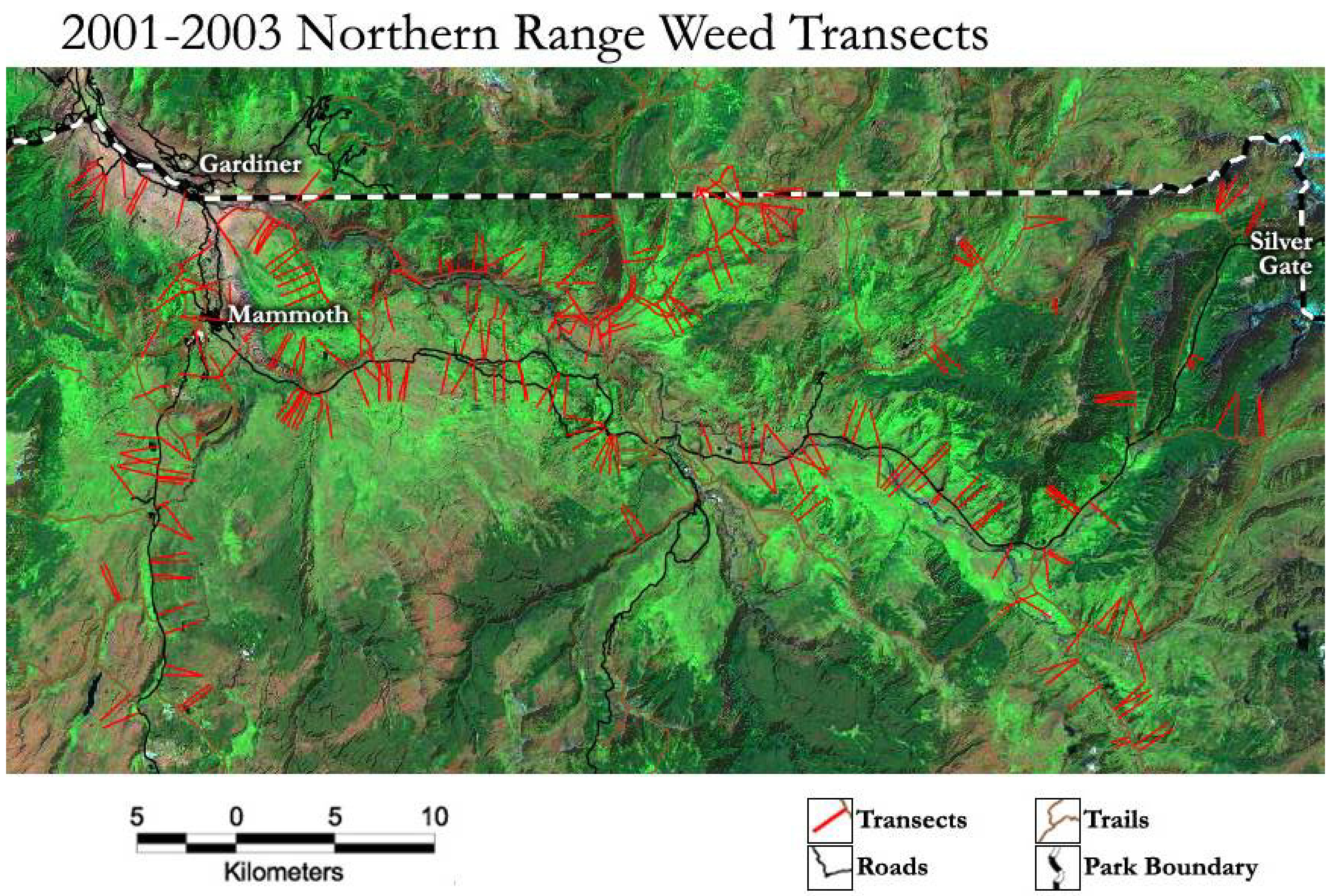


Fig. 2. Locations of all transects walked in the northern range of YELL from 2001 - 2003.

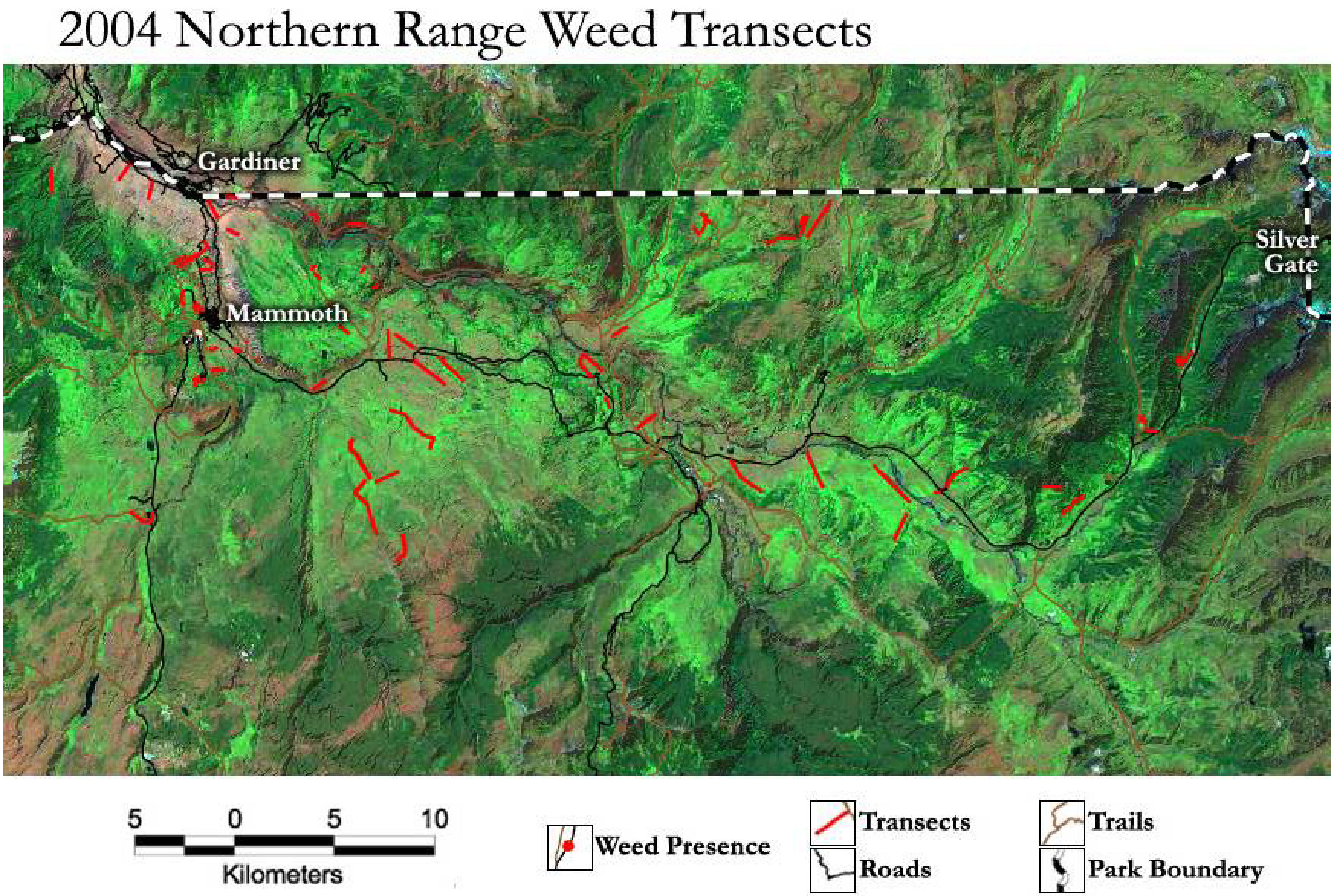


Fig. 3. Locations of all transects walked in the northern range of YELL in 2004.

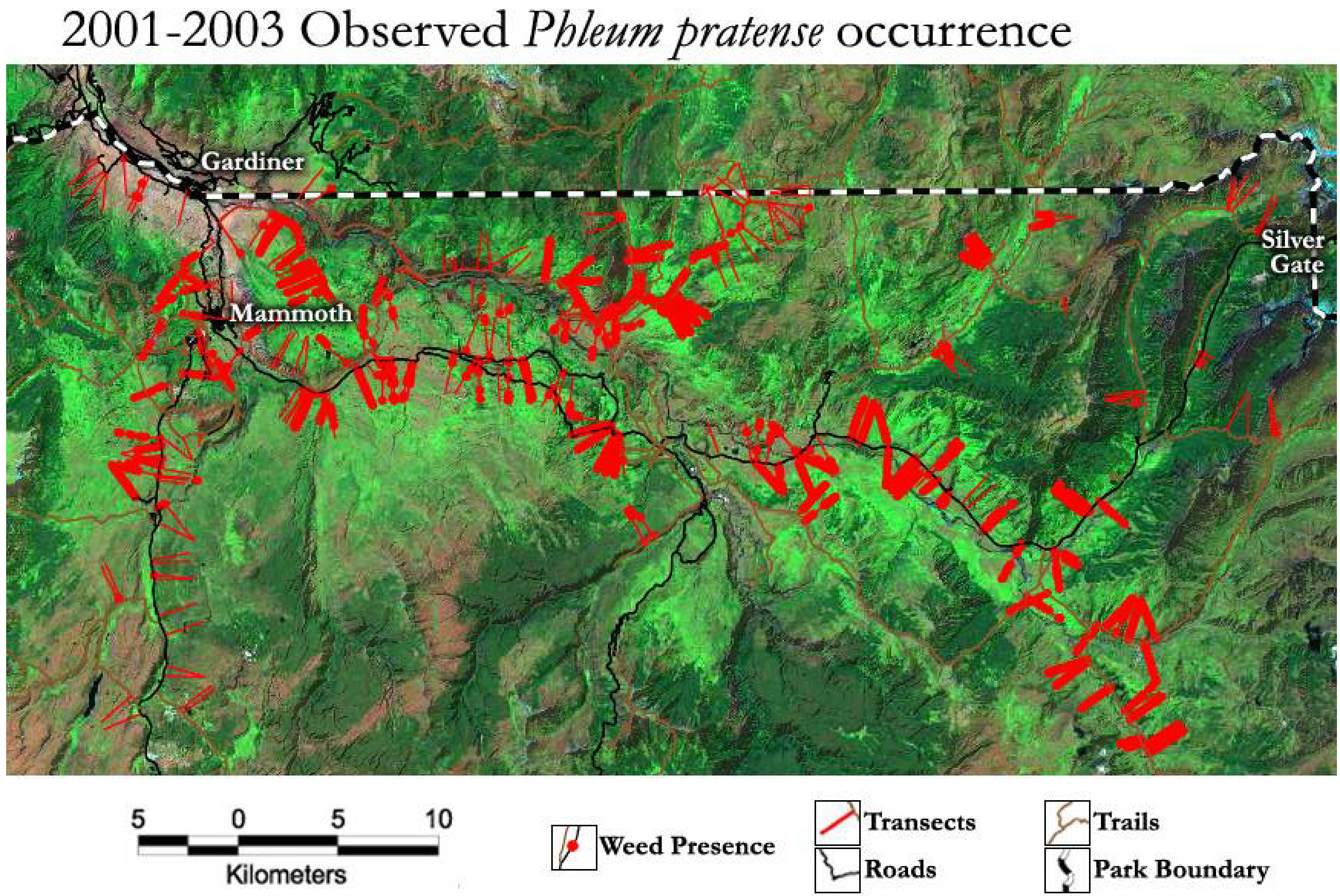


Fig. 4. Observed presence of *Phleum pratense* in the northern range of YELL from 2001-2003.

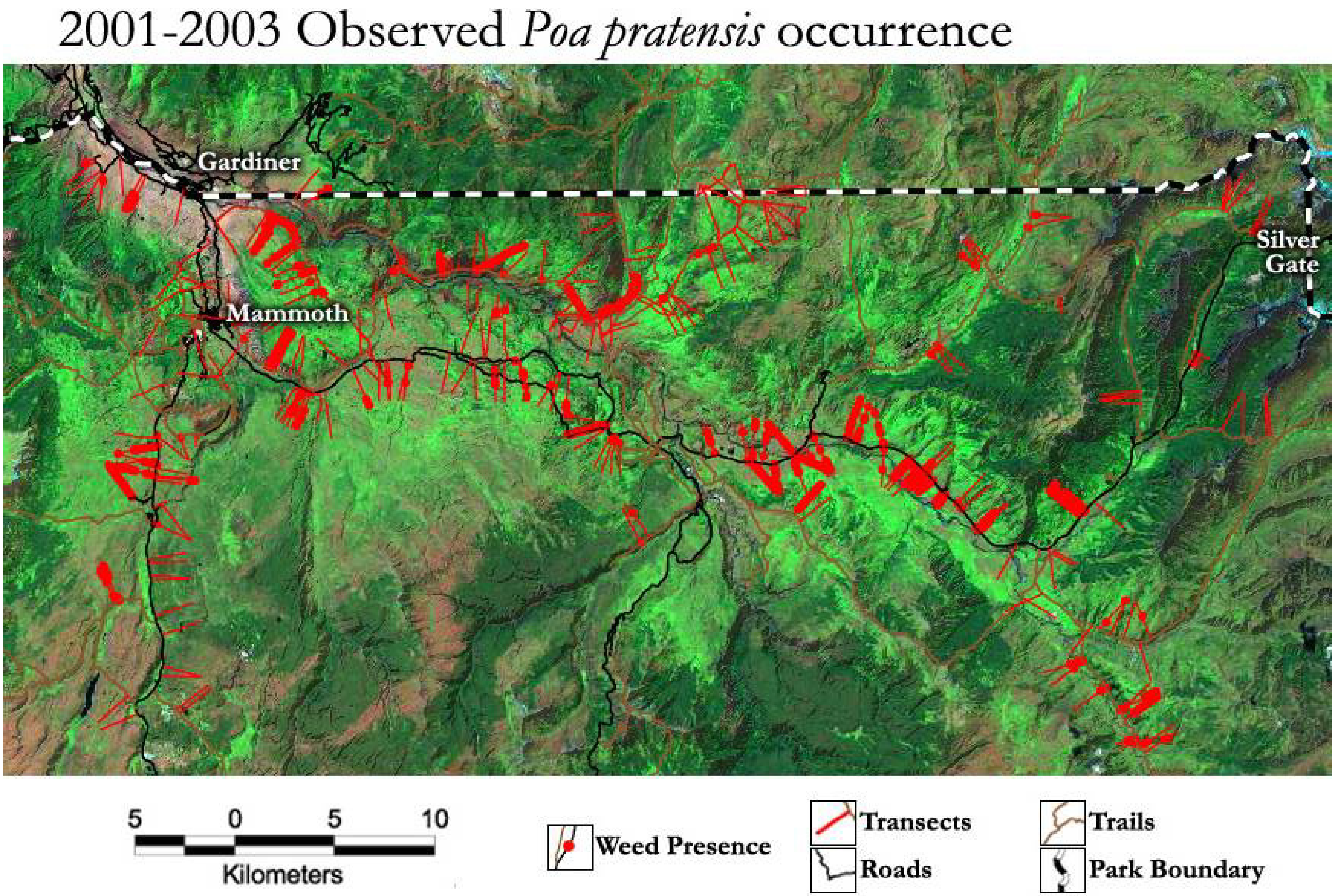


Fig. 5. Observed presence of *Poa pratensis* in the northern range of YELL from 2001 - 2003.

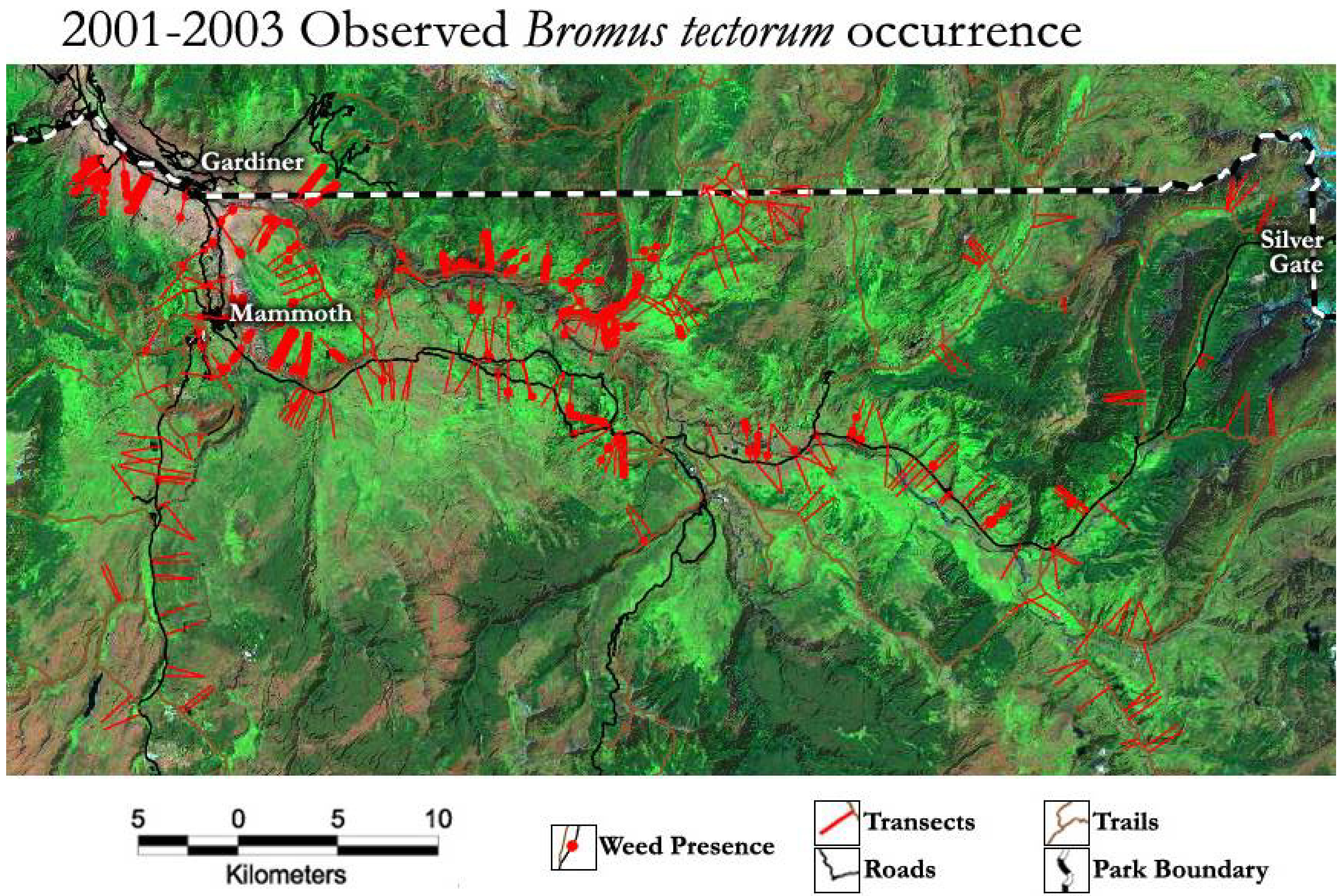


Fig. 6. Observed presence of *Bromus tectorum* in the northern range of YELL from 2001 - 2003.

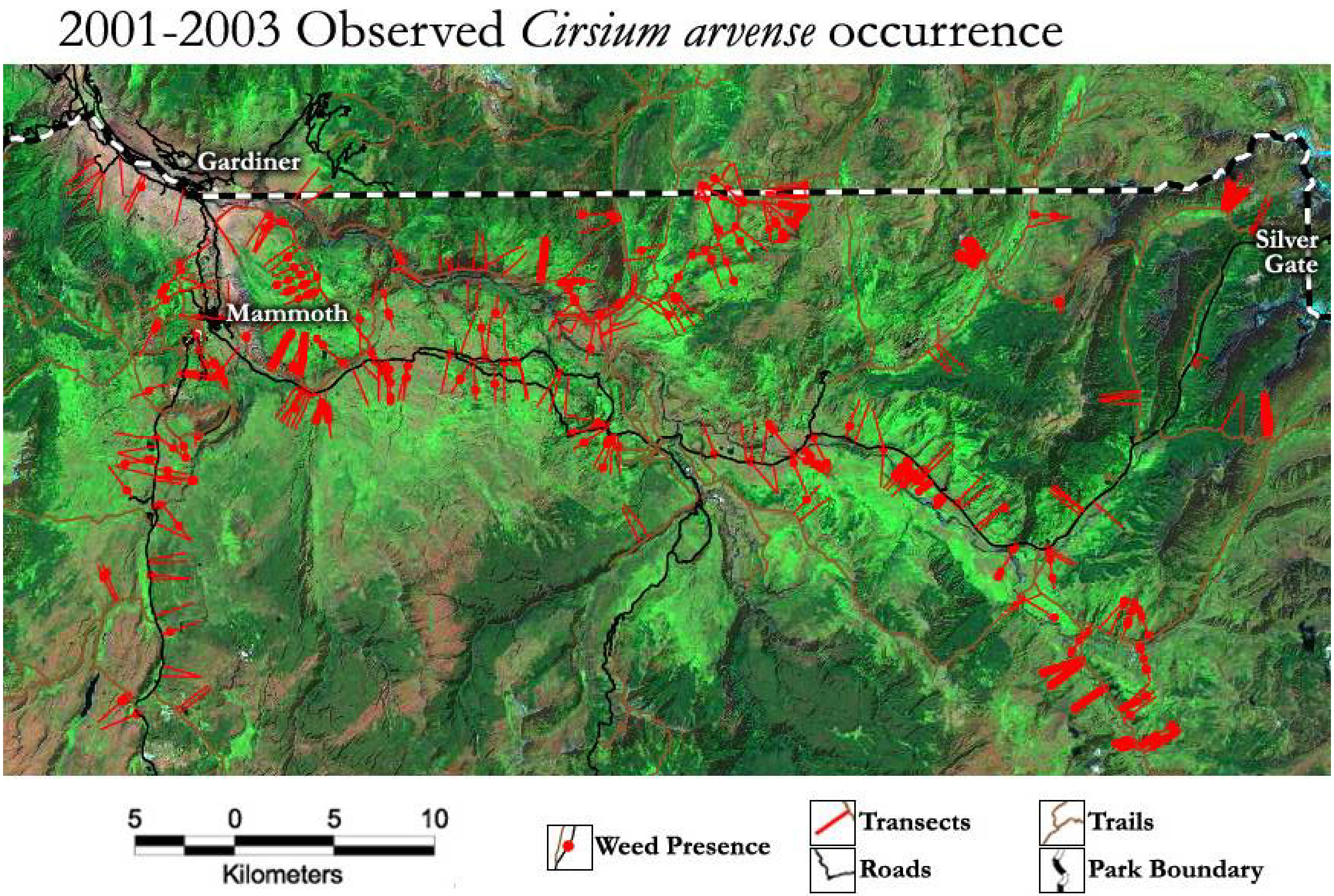


Fig. 7. Observed presence of *Cirsium arvense* in the northern range of YELL from 2001 - 2003.

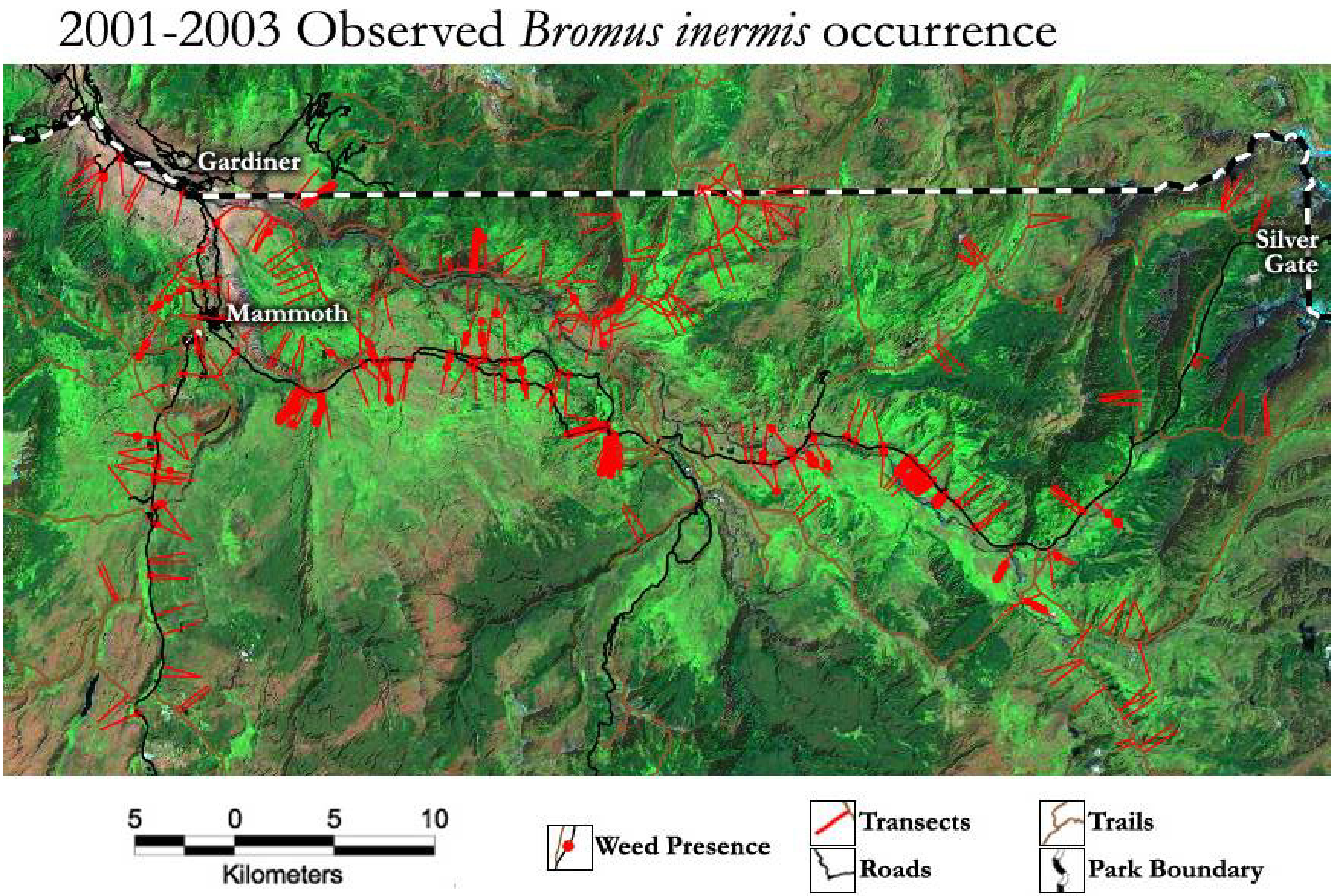


Fig. 8. Observed presence of *Bromus inermis* in the northern range of YELL from 2001 - 2003.

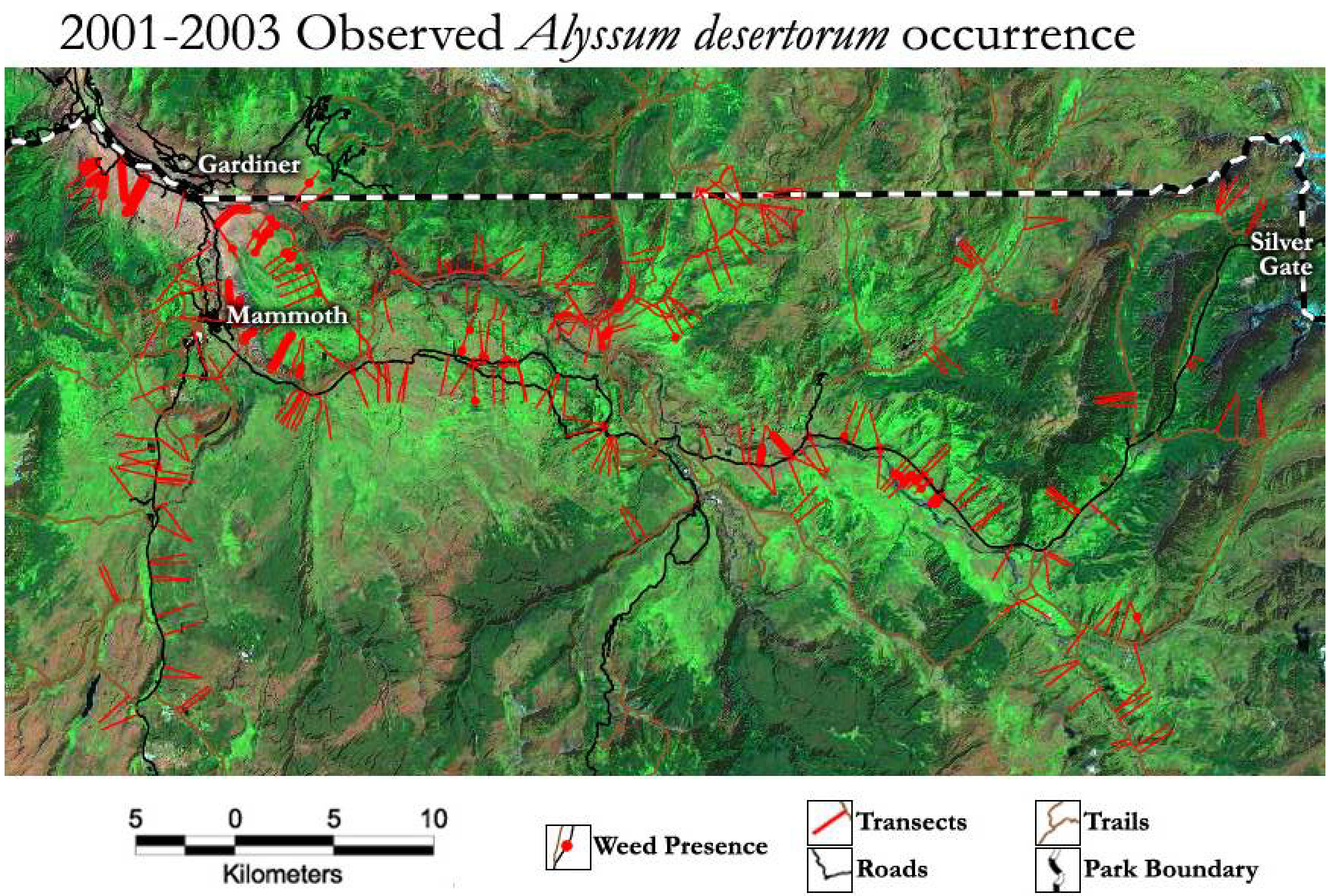


Fig. 9. Observed presence of *Alyssum desertorum* in the northern range of YELL from 2001 - 2003.

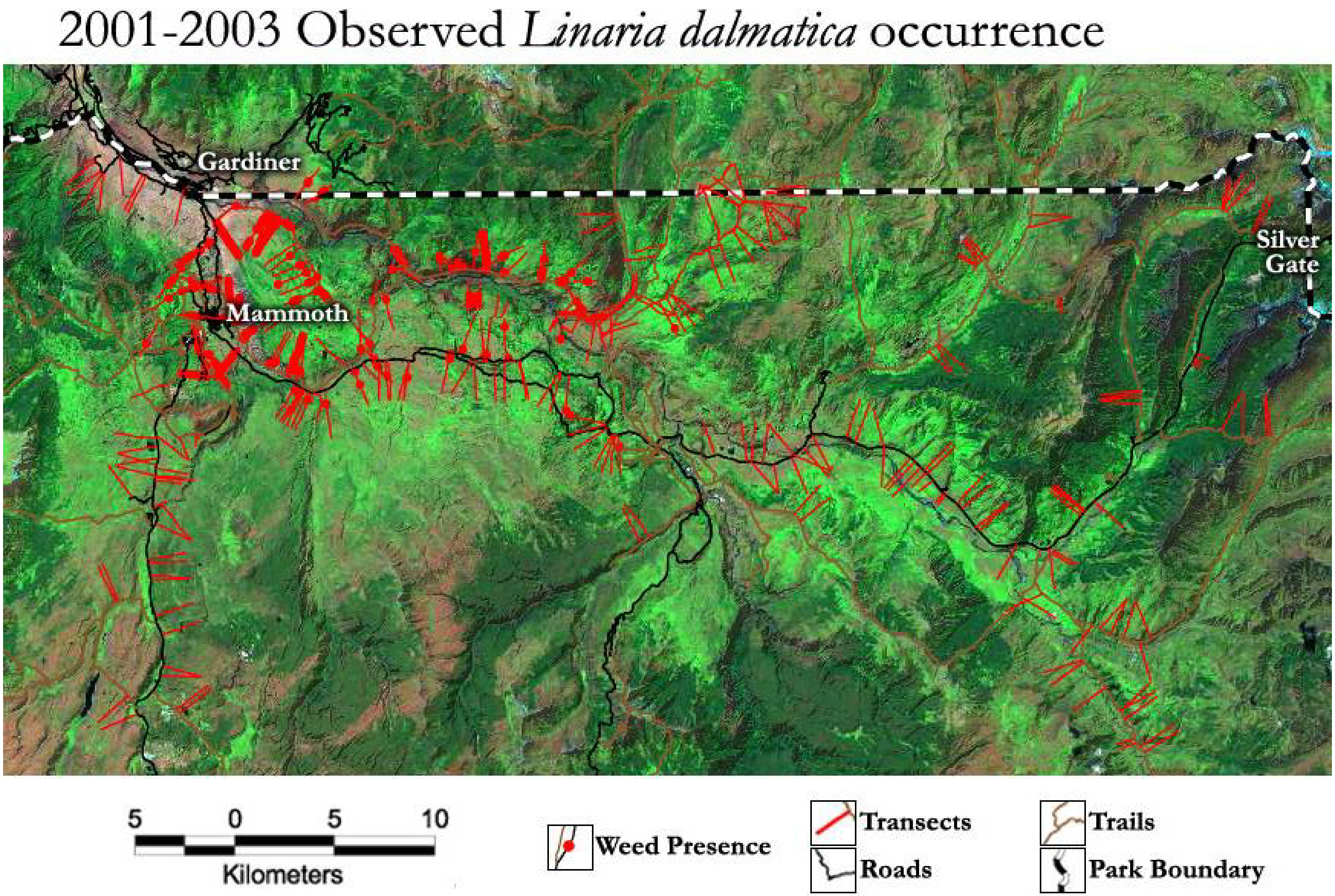


Fig. 10. Observed presence of *Linaria dalmatica* in the northern range of YELL from 2001 - 2003.

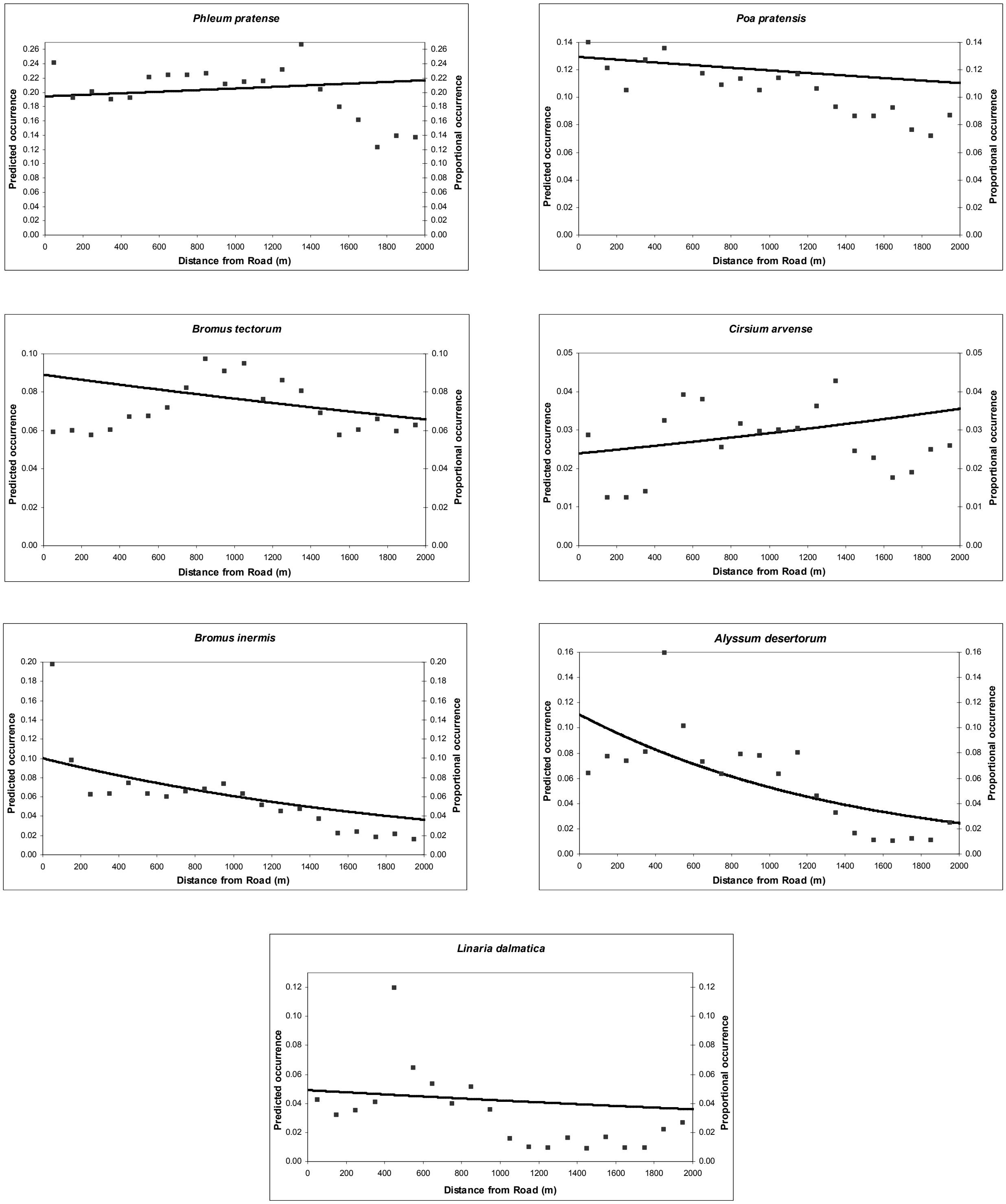


Fig. 11. Proportion of selected species observed within 100 m intervals of distance to roads, and the fitted curve for the logistic regression of NIS occurrence and distance from roads in the northern range of Yellowstone National Park for the 2001-2003 data.

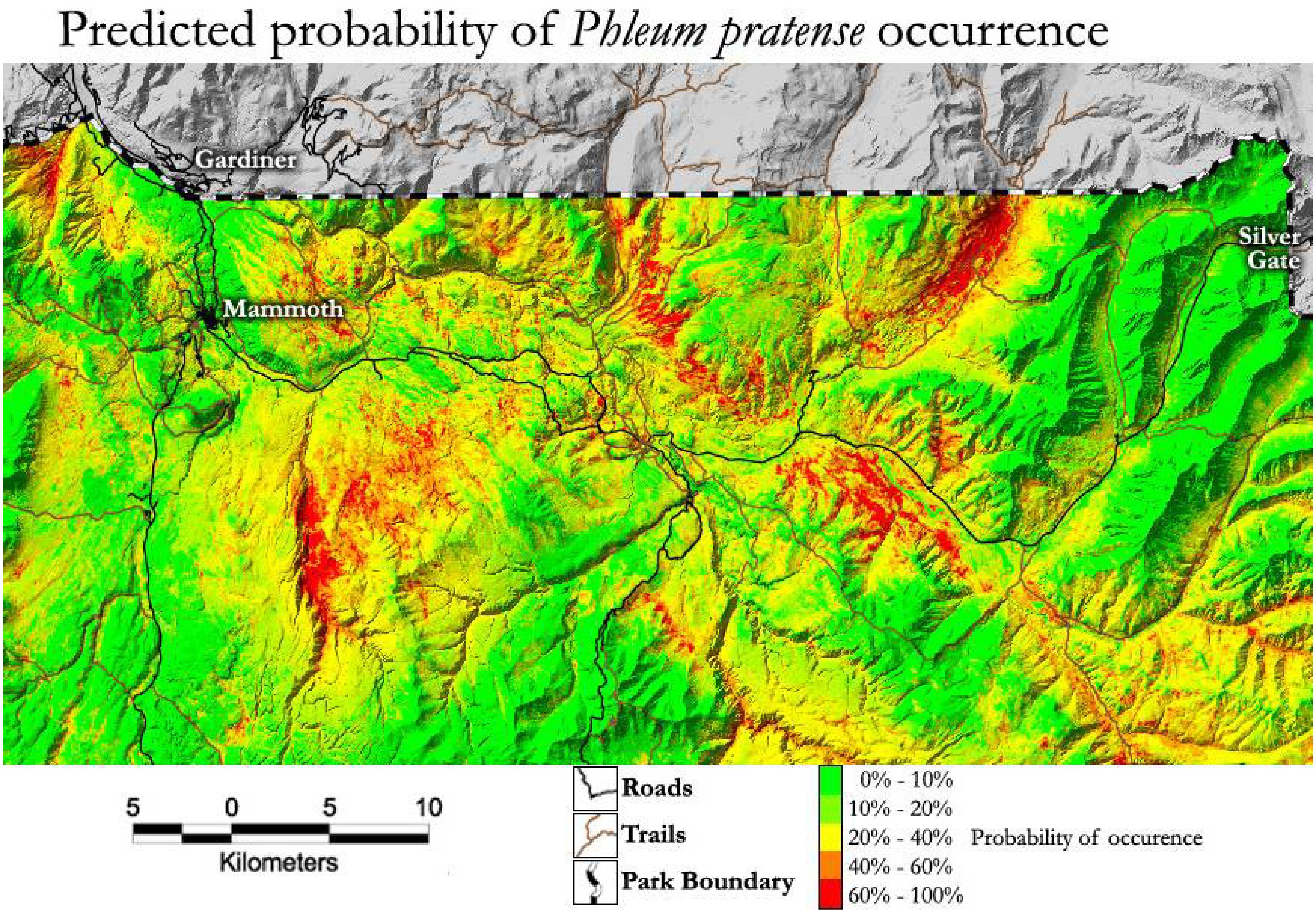


Fig. 12. Predicted occurrence of *Phleum pratense* in the northern range of YELL.

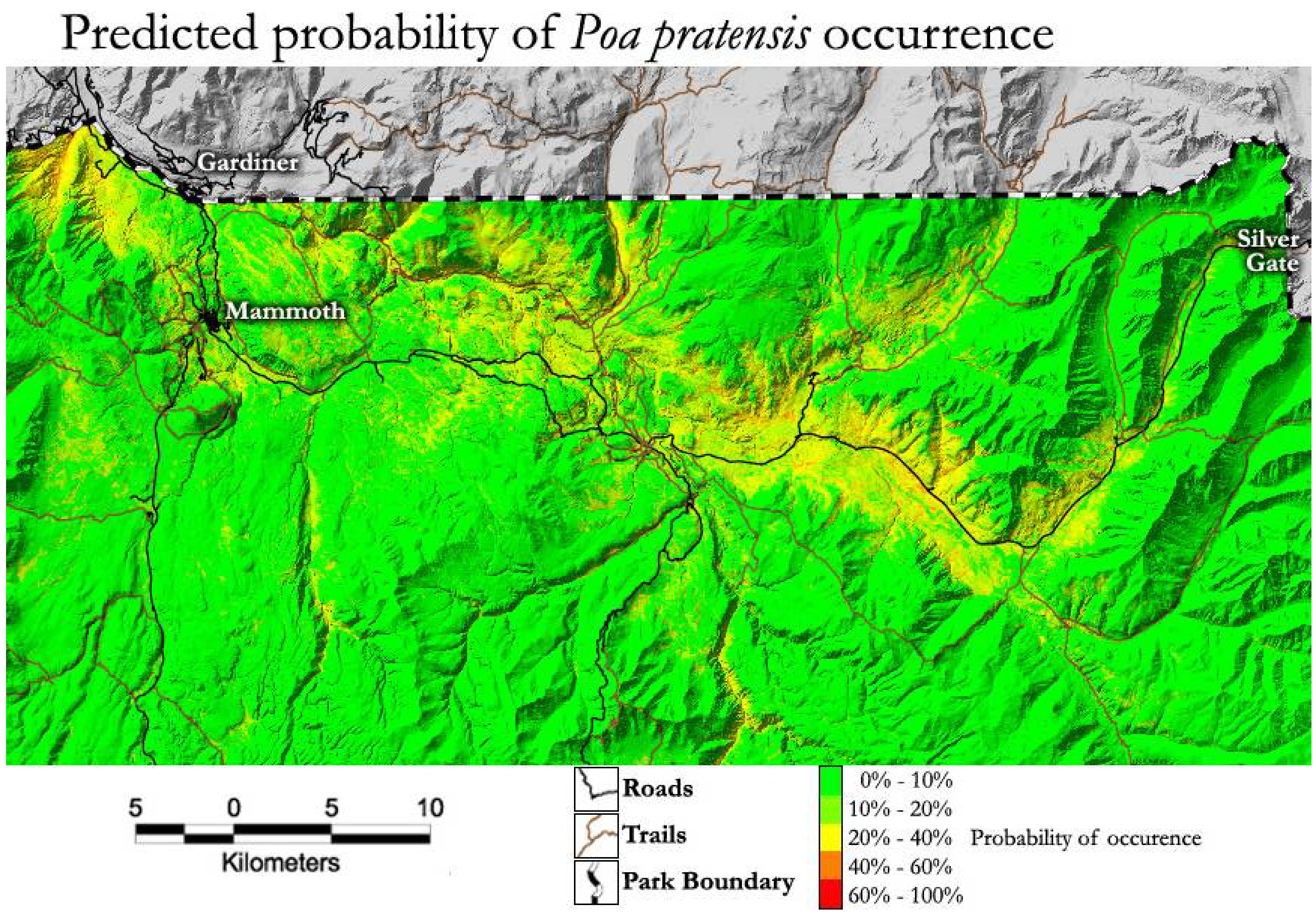


Fig. 13. Predicted occurrence of *Poa pratensis* in the northern range of YELL.

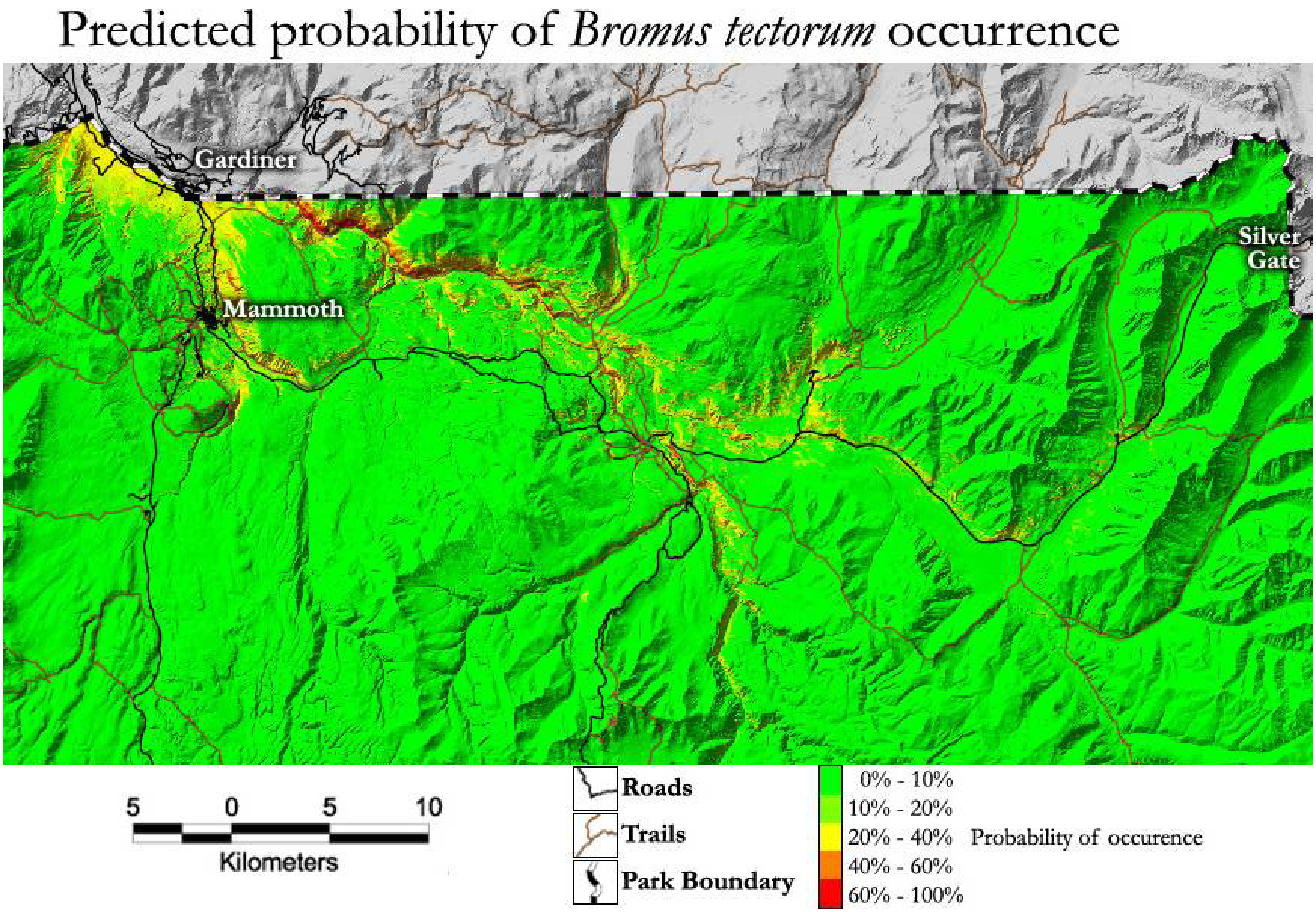


Fig. 14. **Predicted occurrence of *Bromus tectorum* in the northern range of YELL.**

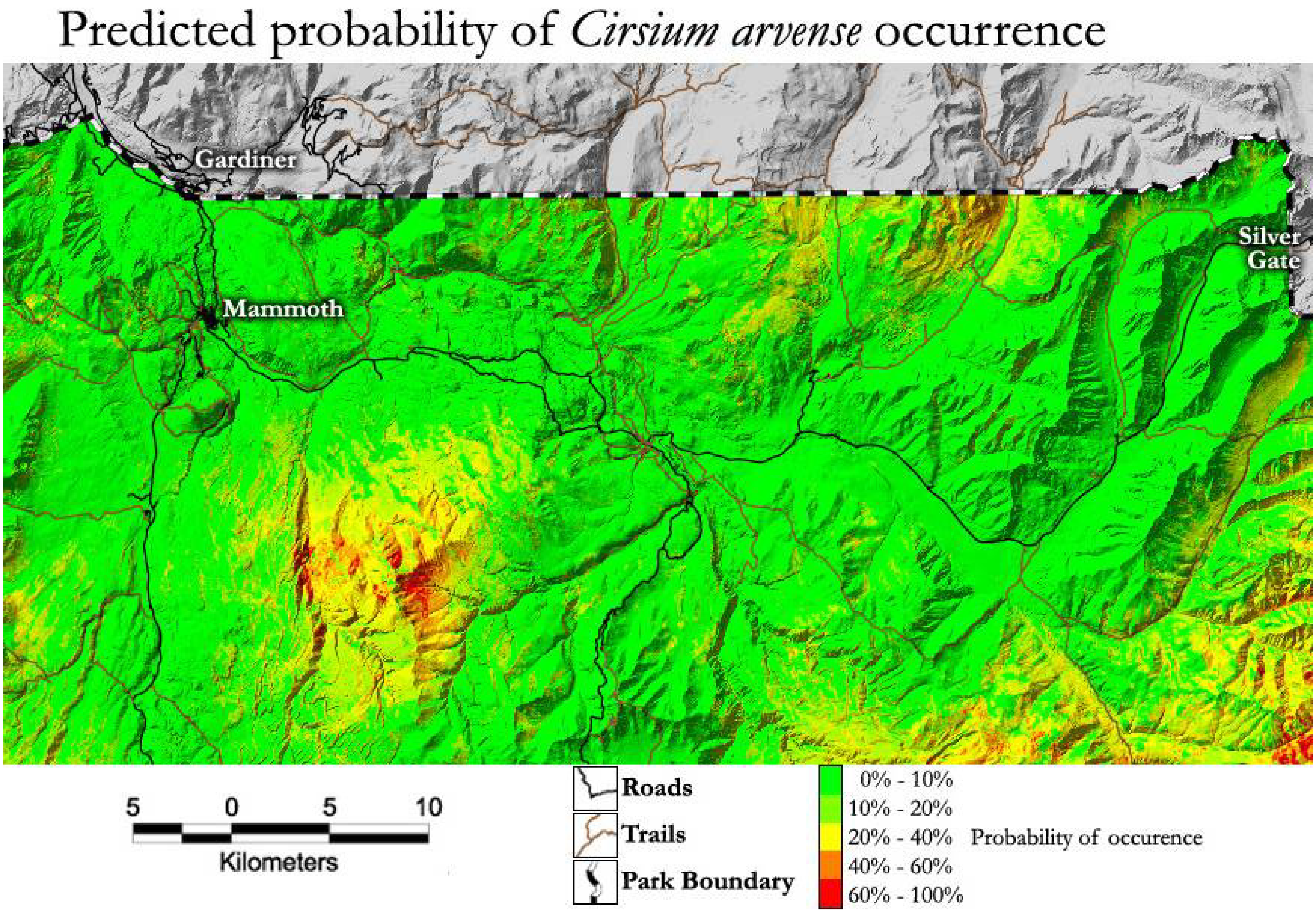


Fig. 15. Predicted occurrence of *Cirsium arvense* in the northern range of YELL.

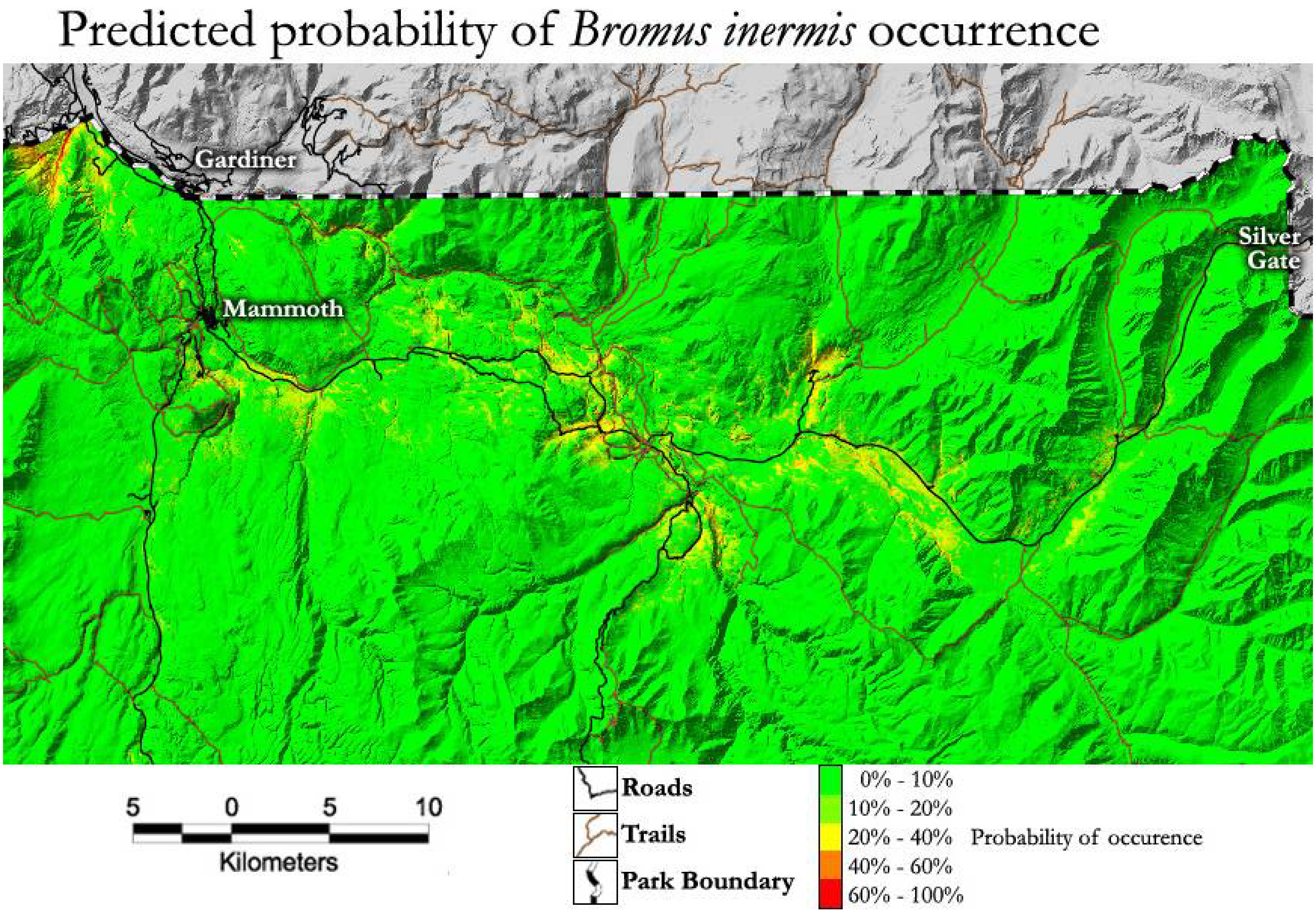
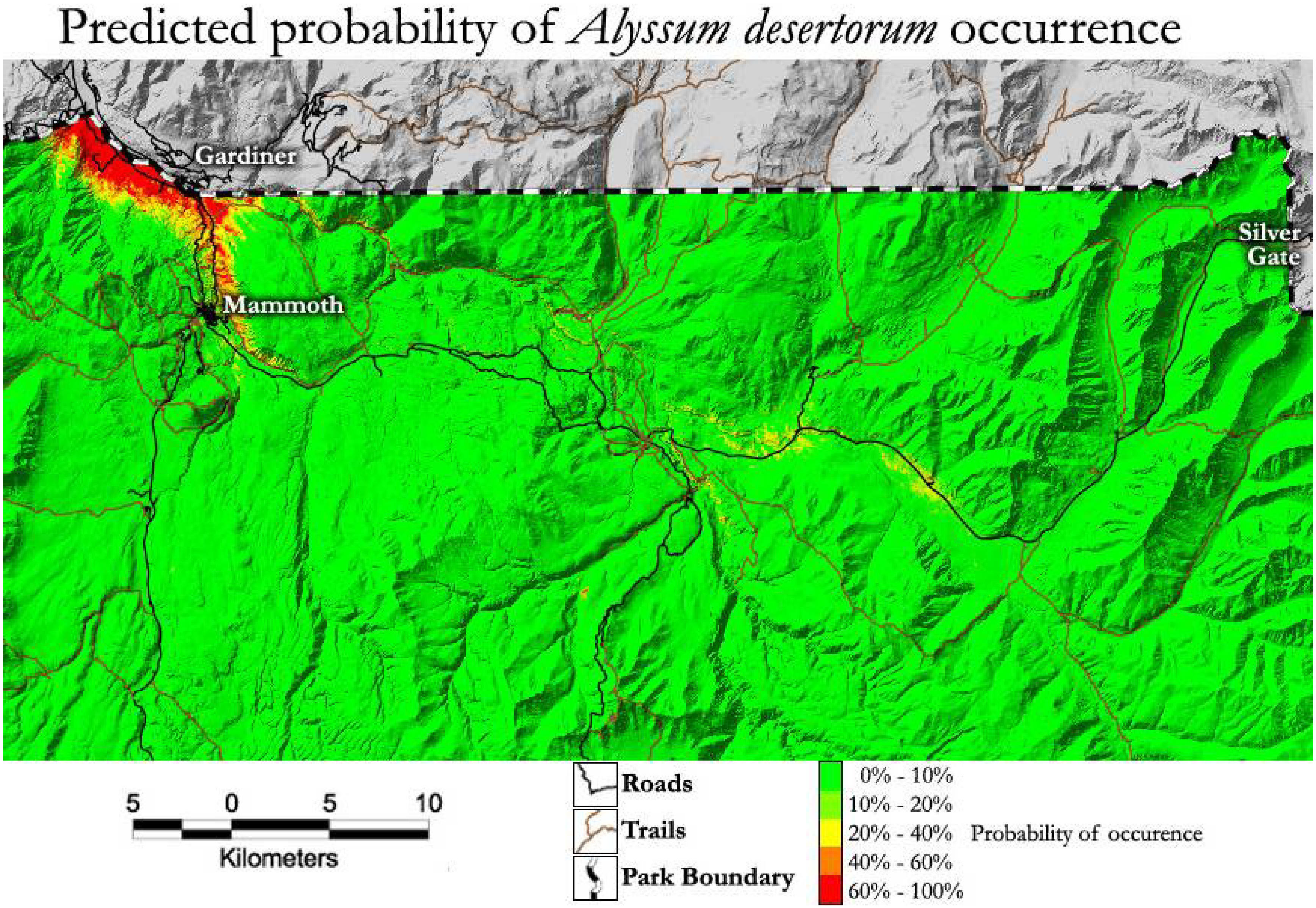

Predicted probability of *Bromus inermis* occurrence

Fig. 16. Predicted occurrence of *Broumus inermis* in the northern range of YELL.



Predicted probability of *Alyssum desertorum* occurrence

Fig. 17. Predicted occurrence of *Alyssum desertorum* in the northern range of YELL.

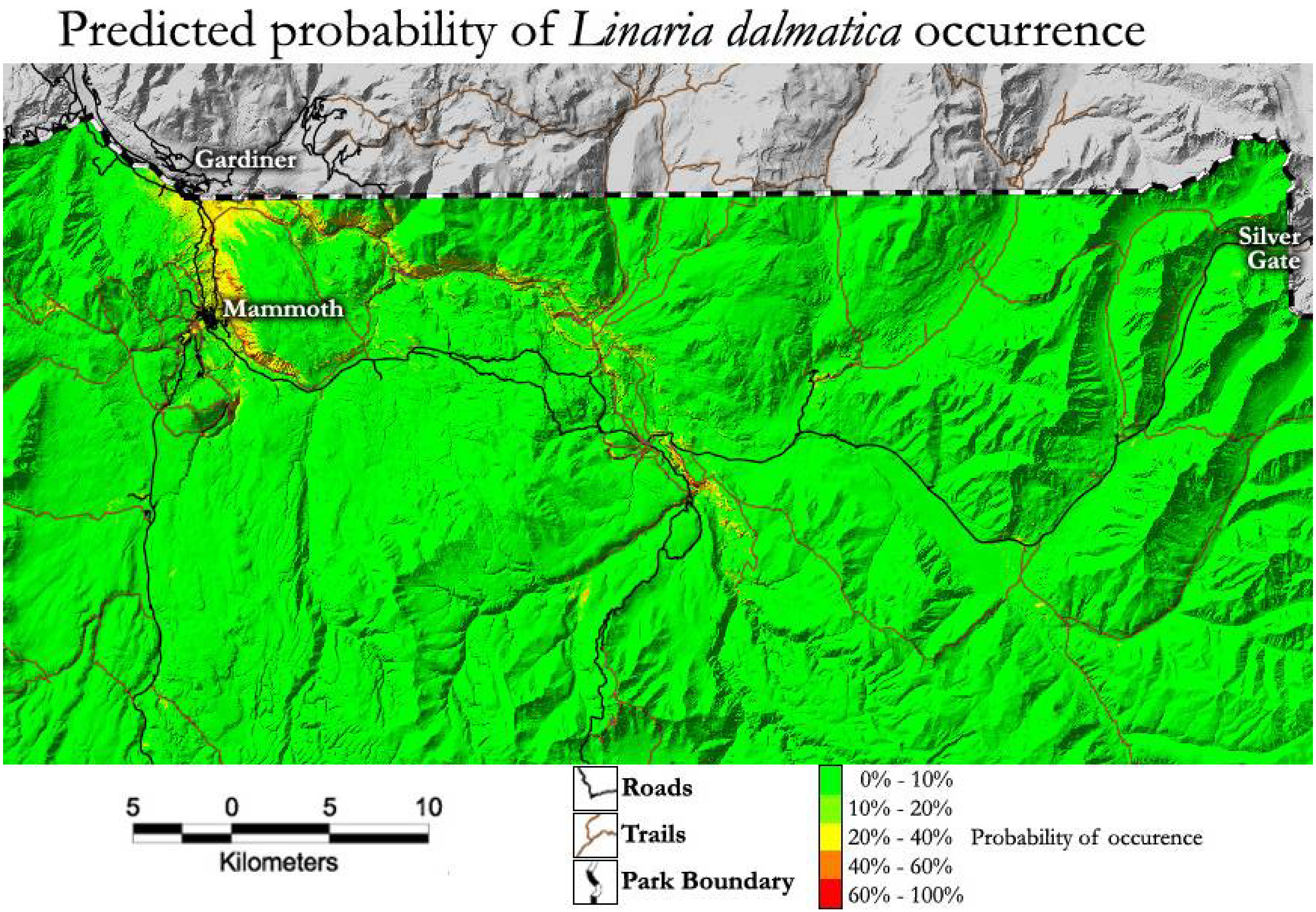


Fig. 18. Predicted occurrence of *Linaria dalmatica* in the northern range of YELL.

Table 2. Coefficient values from the generalized linear models for the seven most prevalent non-indigenous species in the northern range of YELL, using 2001-2003 data.

	Alyssum desertorum	*Bromus inermis*	*Bromus tectorum*	*Cirsium arvense*	*Linaria dalmatica*	*Poa pratense*	*Phleum pratense*
(Intercept)	13.34898	8.82634	4.93962	-7.79169	2.84698	2.66630	1.71355
Wildfire in 1988	-1.80501	0.94902	-0.93381	0.60346	-1.45364	-0.71334	0.61913
Cosine of aspect	-0.51170	-0.21317	-0.73911	-0.10163	-0.28439	-0.38956	-0.14390
Sine of aspect	-	-	-0.20856	0.52720	-0.25857		0.15676
Elevation (m)	-0.00864	-0.00566	-0.00509	-0.00018	-0.00268	-0.00307	-0.00305
Distance from roads	-0.00061	-0.00055	-0.00020	0.00020	-0.00018	-0.00002	0.00017
Slope (°)	-0.87419	-0.00809	0.04356	0.01126	0.03158	-0.00633	0.00519
Solar insulation	0.00885	0.00013	0.00014	0.00023	-0.00004	0.00004	0.00002
Distance from trails	0.00032	0.00034	0.00005	0.00019	-0.00065	0.00017	0.00024
Trees or shrub/grasslands	0.62229	1.42114	1.13287	1.46224	0.40682	0.19354	-0.08311
ETM01 - Landsat 07/99	-0.03105	-0.07464	0.02204	-	-0.03905	0.01807	0.03314
ETM02 - Landsat 07/99	-0.05800	0.13947	-0.05236	0.04523	0.03584	0.02178	0.04354
ETM03 - Landsat 07/99	0.06337	-0.11417	0.01124	-0.02616	0.01502	-0.04077	-0.08752
ETM04 - Landsat 07/99	0.02998	-	0.01832	-0.01317	0.00716	0.01341	0.03027
ETM05 - Landsat 07/99	-0.05223	-	-0.01537	0.02149	-0.03227	0.01112	0.00716
ETM07 - Landsat 07/99	0.07337	0.01772	0.02598	-0.04727	0.01937	-0.02324	-0.01022
ISO128 - Vegetation classification	-	0.00674	0.00439	0.02147	0.01593	0.00637	0.00402
Number of parameters	15	14	17	16	17	16	17
Akaike Information Criterion	8449.5	15289.9	18484.1	17949.3	11928.9	32067.4	48452.6
Degrees of freedom	52896.0	52896.0	52896.0	52896.0	52896.0	52896.0	52896.0
Total residuals	52881.0	52882.0	52879.0	52880.0	52879.0	52880.0	52879.0
Residual deviance	8419.5	15261.9	18450.1	17917.3	11894.9	32035.4	48418.6

The 2004 validation data points, which were not used in the GLM, were overlaid on the appropriate probability maps in the GIS. At each validation data point, the predicted probability value was recorded and collated into six probability classes (0, >0-20, >20-40, >40-60, >60-80, >80-100). The aim of the 2004 sampling was to sample equally in each of the percentile bins. However, an additional variable (trees or shrub/grassland) was added to the predictive models this winter (2005) which improved the model performance and therefore not all bins were sampled equally. When the number of observations (presence and absence) fell below 200, no data were displayed in Fig. 2. *P. pratense* was observed in the same percentage as predicted until the 20-40% bin. In all other cases the observed occurrence in the field was lower than the predicted, although all species generally observed the correct positive trend. *P. pratense* and *B. tectorum* field and model data agreed best, *B. inermis* and *C. arvense* the worst (Fig 19).

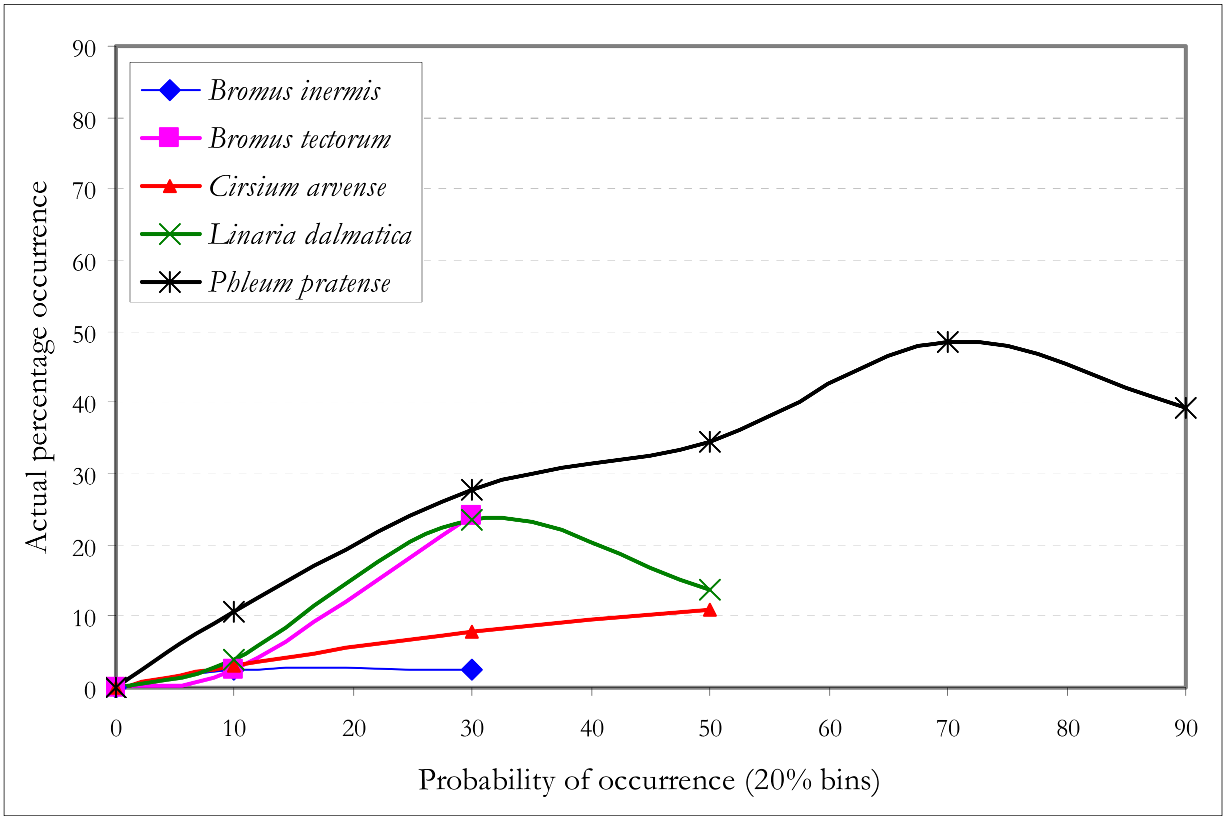


Fig. 19. The percentage of presences observed in the field during 2004 in each of the six predetermined probability of occurrence (presence) bins generated from the model, for five of the most prevalent species on the landscape.

Conclusions

The aim of this project was to determine the best sampling approach for a NIS inventory/survey in YELL, and complete the field survey. We have achieved this, and a bit more. The best sampling method was evaluated using a computer simulation model and some field sampling in 2001. The most efficient and accurate method to represent a species' frequency over the landscape as a whole was used. This method entailed randomly-stratified transects on roads and trails, i.e. transects started on roads or trails and finished 2000 m from either or both of them. By using this unbiased sampling approach it was possible to develop probability of occurrence models for individual species which provide information on the entire area of interest. We believe such probability of occurrence maps will be useful to managers, we hope we are right.

The conventional way to determine the accuracy of predictive maps is to generate an omission/commission table, but that requires determining an arbitrary threshold and the proportion of presences/absences at that given threshold. The resolution provided by this approach is much coarser than our model predicts and we did not consider this approach appropriate. Evaluating the observed percentage of presences in different percentile bins is more appropriate and our 2004 data did follow a positive trend. However, we generally observed less of our target NIS in the field than we expected from the model. While this may mean that the model is poor it can equally mean that the target species has not arrived and established at a given location yet. This is because the probability of occurrence maps are essentially providing information on where the conditions, i.e. the environmental variables put into the model, are more or less suitable for the given species. It does not mean that seeds have arrived at that spot, and the model does not currently include any dispersal from known populations. We believe the validation results are encouraging and are currently investigating other methods to evaluate model accuracy.

With reference to what the data can be used for in the future. The probability of occurrence maps provide information on areas which should be searched first when looking for new populations; and, they provide data on populations which could be used for the next phase of NIS management, monitoring. As with the inventory/survey phase it is not possible to monitor all populations. We do not believe that all populations of a particular NIS are invasive, increasing in spatial extent and density, and having a negative impact, in all the environments in which they occur. Therefore, we need a better understanding of where particular species' populations are increasing/having and impact and where they are not having such a negative effect. Thus, by sampling a number of populations from different environments we can get a better and more informative understanding of this issue. This should be the next phase of this project.

The current project has received a great deal of interest in the scientific and management arena. Over the four year period we have set up similar sampling systems in other areas including Bighorn National Recreation Area, Gallatin National Forest and Kootenai National Forest. Plus, we have one scientific manuscript in press and another published, and have presented different aspects of the project at several conferences – some of which are detailed below.

Publications and Presentations

Rew LJ, Maxwell BD, Aspinall R (2005) Predicting the occurrence of nonindigenous species using environmental and remotely sensed data. *Weed Science* 53, 236-241.

Rew LJ, Maxwell BD, Aspinall RJ and Dougher FL (in press) Searching for a needle in a haystack: evaluating survey methods for sessile species. *Biological Invasions. (Accepted 12/17/04)*

Bruce Maxwell, Richard Aspinall and Lisa Rew have been invited to present this work at several meetings including the IPINAMS conference held in Florida in 2003; WSSA meeting in Kansas 2004. They have also presented aspects of the project at several conferences or meetings. Some of these meetings are listed below:

Dougher FL, Rew LJ and Maxwell BD (2005) Scale effects in the evaluation of the spatial distribution of non-native species in wildland ecosystems. Western Science Weed Society - Vancouver (abstracts not yet in print).

Rew, LJ, Maxwell BD, Taper MD and Aspinall R (2005) Environmental suitability pattern and scale effects on non-indigenous species dispersion. Weed Science Society of America Abstracts – Hawaii, p37.

Rew, LJ and BD Maxwell. (2004) Site-specific management of species with invasive potential. Weed Science Society of America Abstracts – Kansas City, MO. Feb. 2004. (Invited) p. 70.

Rew, LJ and BD Maxwell (2004) Sampling to understand non-indigenous plant species occurrence and develop probability maps of occurrence. Montana Academy of Sciences: Invasive Species Symposium April 16, 2004 Billings, MT.

Rew, LJ and BD Maxwell (2004) Sampling to understand non-indigenous plant species occurrence and develop probability maps of occurrence. The Ecological Society of America Conference August 1-6, 2004 Portland, OR. p25.

Maxwell BD, Rew LJ and Aspinall R (2003) Exotic Plant Survey and Monitoring: Methods to Discover Distribution with Low Frequency Occurrence. In: Tom Philippi and Robert Doren (eds), *Proceedings of Detecting Invasive Exotic Plants, Workshop and Conference*, Florida International University, Miami, FL, Feb. 2003.

Rew LJ, Maxwell BD, Aspinall RJ, and Dougher FL (2003) Sampling to understand non-indigenous plant species occurrence and assist with management objectives. *7th International Confernce on the Ecology and Management of Alien Plant Invasions,* Florida. 90. (invited)

Aspinall R J, Rew LJ, Maxwell BD (2003) Models for predicting the distribution of rare species: non-indigenous plant species in Yellowstone National Park. *7th International Conference on the Ecology and Management of Alien Plant Invasions,* Florida. 6. (invited)

Maxwell BD, Aspinall RJ and Rew LJ (2003) Statistical and simulation model approaches to understand and manage non-indigenous plant species: case studies from the Greater Yellowstone Ecosystem. *7th International Conference on the Ecology and Management of Alien Plant Invasions,* Florida. 57. (invited)

Rew LJ (2002) The threat of exotic vegetation in the Greater Yellowstone Area. *Greater Yellowstone Coordination Committee*, Jackson, WY, May 8th 2002.

Rew LJ (2002) Developing a predictive weed survey methodology for exotic weeds in northern range of Yellowstone National Park. *Rocky Mountain Summit*, Whitefish, MT, Sept 18th 2002.

References

Akaike H (1977) Likelihood of a model and information criteria. *J. Econometrics*. 16:3-14.

Braithwaite RW, Lonsdale WM & Estbergs JA (1989) Alien vegetation and native biota in tropical Australia: the spread and impact of *Mimosa pigra*. *Biol Cons.*, 48, 189-210.

Burnham KP and Anderson DR (1998) Model Selection and Inference: A Practical Information-Theoretic Approach. Springer-Verlag, New York.

Davis MA and Pelsor M (2001) Experimental support for a resource-based mechanistic model of invasibility. *Ecology Letters,* 4, 421-428.

Davis MA, Grime JP & Thompson K (2000) Fluctuating resources in plant communities: a general theory of invisibility. *J. of Ecology*, 88, 528-534.

Hirzel and Guisan (2002). Which is the optimal sampling strategy for habitat suitability modeling. *Ecological Modeling.*

Hobbs RJ and Mooney HA (1991) Effects of rainfall variability and gopher disturbance on serpentine annual grassland dynamics. *Ecology*, 72, 478-491.

Legleiter C J, Lawrence RL, Fonstad MA, Marcus WA and Aspinall RJ (2003) Fluvial response a decade after wildfire in the northern Yellowstone ecosystem: a spatially explicit analysis. Geomorp. 54:119-136.

Lonsdale WM (1999) Global patterns of plant invasions and the concept of invisibility. *Ecology*, 80, 1522-1536.

Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M and Bazzaz FA (2000) Biotic invasions: causes, epidemiology, global consequences, and control. *Ecol. Appl.*, 10, 689-710.

National Park Service 1996 (Online) Preserving our natural heritage – a strategic plan for managing invasive non-indigenous plants on national park system lands. http://www.nature.bps.gov/wv/strat_pl.htm

NAWMA (2005) North American Weed Mapping Association. http://www.nawma.org

Swift LW (1976) Algorithm for solar radiation on mountain slopes. Water Res. 12:108-112.U.S.

U.S. National Park Service Organic Act of 1916 (Online) (Exerts from) http://www.nps.gov/legacy/nomenclature.html

Weaver T, Gustafson D and Lichthardt J (2001) Exotic plants in early and late seral vegetation of fifteen Northern Rocky environments. *West. North Am. Nat.*, 61, 417-427.

Whipple JJ (2001) Annotated checklist of exotic vascular plants in Yellowstone National Park. *West. North Am. Nat.*, 61, 336-346.

Project Timetable

Project timetable	Jan	Feb	Mar	Apr	May	June	Jul	Aug	Sept	Oct	Nov	Dec
Project implementation	**2002**											
Advertise positions			🟨	🟨								
Reserve accommodation			🟨									
Purchase GPS, rent vehicles etc.			🟨	🟨								
Apply for Research Permit				🟨								
Program GPS				🟨								
Phase I - Non-indigenous survey												
Field Assistants commence					🟩							
GPS and botanical initiation (1 week)					🟩							
Data collection						🟩	🟩	🟩				
Data collation and analysis												
Data collated								🟦				
Data analysis & report									🟦	🟦	🟦	
Project implementation	**2003**											
Advertise positions			🟨	🟨								
Reserve accommodation			🟨									
Apply for Research Permit				🟨								
Program GPS				🟨								
Phase I - Non-indigenous survey												
Field Assistants commence					🟩							
GPS and botanical initiation (1 week)					🟩							
Data collection						🟩	🟩	🟩				
Data collation and analysis												
Data collated								🟦				
Data analysis & report									🟦	🟦	🟦	🟦
Project implementation	**2004-2005**											
Advertise positions			🟨	🟨								
Reserve accommodation			🟨									
Apply for Research Permit				🟨								
Program GPS				🟨								
Phase I - Non-indigenous survey												
Field Assistants commence					🟩							
GPS and botanical initiation (1 week)					🟩							
Data collection						🟩	🟩					
Data collation and analysis												
Data collated								🟦				
Final report due May 2005									🟦	🟦	🟦	🟦

Appendix 1: Non-indigenous NIS of interest for Yellowstone National Park

Watch List: Exotic species not documented/established in the park. The goal is to prevent establishment through staff education, early detection, and eradication. Those species noted with an asterisk (*) have been found in the park, but were removed prior to seed dispersal.

1.	*Arctium lappa**	(great burdock)
2.	*Arctium minus**[1]	(common burdock)
3.	*Centaurea pratensis**	(meadow knapweed)
4.	*Centaurea solstitialis*	(yellow starthistle)
5.	*Chondrilla juncea*	(rush skeletonweed)
6.	*Crupina vulgaris*	(common crupina)
7.	*Isatis tinctoria**	(dyer's woad)
8.	*Lepidium latifolium*	(perennial peppergrass)
9.	*Lythrum salicaria*	(purple loosestrife)
10.	*Onopardum acanthium**	(scotch thistle)
11.	*Senecio jacobaea**	(tansy ragwort)

Priority 1: Species that have produced seed in the park, but populations are small and limited in number. These species have a high probability for eradication with continued annual monitoring and treatment. They are also the most cost effective species to control (<1 acre infestation).

1.	*Astragalus cicer*	(chick-pea milkvetch)
2.	*Carduus acanthoides*	(plumeless thistle)
3.	*Centaurea diffusa*	(diffuse knapweed)
4.	*Centaurea repens*	(Russian knapweed)
5.	*Chorispora tenella*	(blue mustard)
6.	*Conium maculatum*	(poison hemlock)
7.	*Dianthus armeria*	(grass pink)
8.	*Euphorbia esula*	(leafy spurge)
9.	*Hyoscyamus niger*	(black henbane)
10.	*Potentilla recta*	(sulfur cinquefoil)
11.	*Ranunculus acris*	(tall buttercup)
12.	*Tamarix chinensis*	(tamarisk)
13.	*Tanacetum vulgare*	(tansy aster)
14.	*Trifolium aureum*	(yellow clover)

Priority II: Aggressive invaders, some of which are well established in some localities making eradication impractical (identified by •), but most are confined to relatively small areas at specific locations. Containment will be the primary goal for these species in established infestations, and as funding permits as a secondary goal, annual control to reduce seed production with possible future eradication. Individual plants or small infestations away from core infestation areas will be a high priority for aggressive control. Control efforts have a high probability of successfully limiting the spread, and will be undertaken. Monitoring of and for these species should by frequent and regular.

---

[1] Only basal rosettes have been found, so identification to species is uncertain

1.	*Berteroa incana*•	(berteroa)
2.	*Cardaria* spp.[2]	(whitetop)
3.	*Carduus nutans*	(musk thistle)
4.	*Centaurea maculosa*•	(spotted knapweed)
5.	*Chrysanthemum leucanthemum*•	(oxeye daisy)
6.	*Cirsium vulgare*	(bull thistle)
7.	*Convolvulus arvensis*	(field bindweed)
8.	*Cynoglossum officinale*•	(houndstongue)
9.	*Hieracium auranticum*	(orange hawkweed)
10.	*Hieracium caespitosum*	(yellow king devil)
11.	*Hieracium floribundum*	(glaucous king devil)
12.	*Hieracium flagellare*	(whiplash hawkweed)
13.	*Hypericum perforatum*	(St. Johnswort)
14.	*Linaria dalmatica*•	(Dalmatian toadflax)
15.	*Linaria vulgaris*•	(yellow toadflax, butter and eggs)
16.	*Melilotus albus*	(white sweet clover)
17.	*Melilotus officinalis*•	(yellow sweet clover)
18.	*Silene vulgaris*	(bladder campion)
19.	*Sonchus arvensis*	(perennial sow-thistle)
20.	*Verbascum thapsus*	(wooly mullein)
21	*Veronica biloba*	(bilobed speedwell)

<u>Priority III</u>: Aggressive exotics, which are dispersed over large areas of Yellowstone and have deleterious effects on the park ecosystem. Control efforts are likely to be ineffective and costly. However, work may be done to confine the spread of these plants in sensitive areas. Monitoring would be beneficial, but will come after Priorities I & II.

1.	*Alyssum desertorum*	(desert elyssum)
2.	*Bromus inermis*	(smooth brome)
3.	*Bromus tectorum*	(cheatgrass, downy chess)
4.	*Cirsium arvense*	(Canada thistle)
5.	*Elymus repens*	(quackgrass)
6.	*Medicago lupulina*	(black medic)
7.	*Phleum pratense*	(common timothy)
8.	*Poa* spp.[3]	(bluegrass)
9.	*Trifolium hybridum*	(alsike clover)
10.	*Trifolium repens*	(white clover)

<u>Priority IV</u>: Exotics, for which little or no control efforts are foreseen. Even though many of these plants displace native vegetation, control of high priority species takes precedence. Limited monitoring actions may be undertaken. Approximately 134 species fall into this category. None of the plants in this category are listed noxious by the surrounding states.

---

[2] *Cardaria chalepensis, Cardaria draba,* and *Cardaria pubescens*

[3] *Poa annua, Poa bulbosa, Poa compressa, Poa palustris,* and *Poa pratensis*

Weed Science, 53:236–241. 2005

Symposium

Predicting the occurrence of nonindigenous species using environmental and remotely sensed data

Lisa J. Rew
Corresponding author. Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT 59717;
lrew@montana.edu

Bruce D. Maxwell
Department of Land Resources and Environmental Sciences, Montana State University,
Bozeman, MT 59717

Richard Aspinall
Geographic Information and Analysis Center,
Montana State University, Bozeman, MT 59717

To manage or control nonindigenous species (NIS), we need to know where they are located in the landscape. However, many natural areas are large, making it unfeasible to inventory the entire area and necessitating surveys to be performed on smaller areas. Provided appropriate survey methods are used, probability of occurrence predictions and maps can be generated for the species and area of interest. The probability maps can then be used to direct further sampling for new populations or patches and to select populations to monitor for the degree of invasiveness and effect of management. NIS occurrence (presence or absence) data were collected during 2001 to 2003 using transects stratified by proximity to rights-of-way in the northern range of Yellowstone National Park. In this study, we evaluate the use of environmental and remotely sensed (LANDSAT Enhanced Thematic Mapper +) data, separately and combined, for developing probability maps of three target NIS occurrence. Canada thistle, dalmation toadflax, and timothy were chosen for this study because of their different dispersal mechanisms and frequencies, 5, 3, and 23%, respectively, in the surveyed area. Data were analyzed using generalized linear regression with logit link, and the best models were selected using Akaike's Information Criterion. Probability of occurrence maps were generated for each target species, and the accuracies of the predictions were assessed with validation data excluded from the model fitting. Frequencies of occurrence of the validation data were calculated and compared with predicted probabilities. Agreement between the observed and predicted probabilities was reasonably accurate and consistent for timothy and dalmation toadflax but less so for Canada thistle.

Nomenclature: Canada thistle, *Cirsium arvense* L. CIRAR; dalmation toadflax, *Linaria dalmatica* (L.) P. Mill. LINDA; timothy, *Phleum pratense* L. PHLPR.

Key words: Generalized linear model, invasive species, logistic regression, nonnative species, predictive mapping, survey, stratified sampling.

Considerable resources are directed toward the management of nonindigenous species (NIS), and obtaining information on their location is important. However, if NIS are relatively infrequent and spread over large areas, financial and logistical constraints will make it impossible to locate and manage all populations. Thus, small portions of the total management area are generally sampled (surveyed). If such data are collected using an unbiased survey design, in which data on NIS occurrence and possibly associated variables are recorded, the data can be used to produce probability maps of species occurrence for the areas which were not surveyed (Franklin 1995; Guisan and Zimmerman 2000; Shafii et al. 2003). Such probability of occurrence maps would help land managers to decide where to send crews to search for additional NIS populations. The survey data also could be used to select populations or patches to monitor from the range of environments within which the target species exists. The relative invasiveness and the potential effects of populations in the different environments can then be evaluated. These monitoring results would serve to prioritize management of populations in the environments that pose the greatest threat to the ecosystem.

Many countries have designated specific areas to be maintained as "wilderness" or "natural areas" for recreational or wildlife benefit, or both. Exactly how management of these wildlands is defined obviously varies, but in many cases it is linked to maintaining flora and fauna at a level observed before settlement by Europeans or at least the early 1900s. For example, the National Park Service has a mandate to maintain natural areas under their jurisdiction as unaltered by human activities as possible (National Park Service 1996). Thus, considerable effort is extended to the management of NIS, particularly plant species.

Disturbance is often suggested as a key factor enhancing the probability of nonindigenous plant establishment in plant communities (Grime 1979). Natural disturbance has a variety of biotic and geomorphic causes including soil disturbance by fauna, weather-related events such as drought, floods, wind, fire, and geological events such as landslides. In most areas of the United States, the natural fire regime has been altered; so, fire should be considered a quasihuman disturbance. Human disturbance also includes construction and use of roads and trails, buildings, utility corridors, and campgrounds. Anthropogenic disturbances such as roads or trails (Gelbard and Belnap 2003; Parendes and Jones 2000; Trombulak and Frissell 2000; Tyser and Worley 1992; Watkins et al. 2003), cultivation, grazing, trampling, and domestic ungulates (Mack and Thompson 1982; Tyser and

Table 1. Coefficient values for the best fit combined variable model for Canada thistle, dalmation toadflax, and timothy for Data Subset 1 ($n = 42{,}317$).

Coefficients	Canada thistle	Dalmation toadflax	Timothy
Intercept	−7.92551	3.12522	2.00790
Proximity to road (m)	0.00019	−0.00017	0.00016
Proximity to trail (m)	0.00010	−0.00065	0.00025
Elevation (m)	0.00033	−0.00249	−0.00308
Cosine of aspect (°)	−0.18760	−0.39421	−0.22335
Sine of aspect (°)	0.66578	−0.30232	0.17488
Presence of wildfire (binary)	0.67134	−1.32851	0.65564
Slope (°)	0.02320	0.03210	0.00336
Solar insolation (Wh m^{-2})	0.00022	−0.00007	—
LANDSAT ETM[a] Band 1	0.03656	−0.04306	0.02932
LANDSAT ETM Band 2	0.05022	0.05954	0.04364
LANDSAT ETM Band 3	−0.07421	—	−0.08256
LANDSAT ETM Band 4	−0.01688	—	0.02936
LANDSAT ETM Band 5	0.00764	−0.03073	0.00831
LANDSAT ETM Band 7	−0.01699	0.01859	−0.01336
Isocluster class	0.01768	0.01423	0.00388

[a] Abbreviation: LANDSAT ETM+ LANDSAT Enhanced Thematic Mapper+.

Key 1988; Young et al. 1972) are often considered to have more effect on the occurrence of NIS than natural disturbances.

If the occurrence of a target species is known to be correlated with a particular variable, one could stratify the sampling scheme on that variable and improve the probability of finding the target (Hirzel and Guisan 2002). In this study of NIS in the northern range of Yellowstone National Park, we accepted the assumption that human disturbance in the form of rights-of-way (ROW) increases the chance of finding NIS and stratified our sampling using this variable, but sampled away from this disturbance to generate an unbiased data set. The aim of this study was to generate predictive maps of target NIS occurrence using generalized linear models for the entire area of interest—the northern range of Yellowstone National Park. To generate a predictive map requires that the independent variable data are available for the entire area of interest. To achieve this, we used environmental data obtained from digital elevation maps and reflectance data from LANDSAT Enhanced Thematic Mapper (ETM)+ imagery. The influence of environmental and reflectance data on the occurrence of target NIS was assessed, and the benefit of using the environmental and reflectance data, independently or combined, to improve model fit was evaluated. The accuracy of the resultant probability of occurrence predictions and maps was evaluated for three target species.

Materials and Methods

Yellowstone National Park covers an area of 899,121 ha predominantly in Wyoming, United States. A total of 187 nonindigenous plant species have been recorded within the Park, which comprises 15% of the total plant species (Whipple 2001). This study concentrates on the area within the northern elk winter range of the Park (152,785 ha). Sixty-two NIS were targeted by this study, but we are only reporting on three of those species here.

A stratified sampling approach was used to collect field data. Transects were stratified on ROW, which include roads and trails in this instance. Field sampling was performed from early June to late August in 2001 to 2003. During the 3 yr, a total of 305 transects were completed, most of which were 2,000 m in length, although some were shorter if the terrain proved impassable. All transects were 10 m wide. The total area surveyed was 53 ha, representing 0.035% of the study area.

Transect start locations were randomly allocated on ROW in a geographical information system (GIS), before commencing field work. In 2001, the start position of each transect was randomly located on a ROW but ran 2,000 m perpendicular to ROW from that point. This approach needed to be partially modified for subsequent years to provide a more similar number of data points at all distances from ROW. In 2002 and 2003, the start locations of transects were still randomly generated but fit the following set

Table 2. Best model fits for Canada thistle, dalmation toadflax, and timothy using seven remotely sensed (LANDSAT ETM+)[a] and eight environmental data variables, combined and independently, for Data Subset 1 ($n = 42{,}317$). Akaike's Information Criterion values of the best fit models are provided with number of variables retained in the best model in parentheses.

Target species	All variables (15)	LANDSAT ETM+ variables (7)	Environmental variables (8)
Canada thistle	14,657.42 (15)	16,066.92 (7)	14,768.24 (7)
Dalmation toadflax	9,513.46 (13)	11,293.14 (6)	9,789.77 (7)
Timothy	38,388.81 (14)	40,702.72 (7)	42,956.91 (7)

[a] Abbreviation: LANDSAT ETM+ LANDSAT Enhanced Thematic Mapper+.

TABLE 3. Best model fits for Canada thistle, dalmation toadflax, and timothy using all 15 independent variable data (reflectance and environmental data variables) for Data Subsets 1 to 3 (n = 42,317). Akaike's Information Criterion values of the best fit models are provided with number of variables retained in the best model in parentheses.

Target species	Subset 1	Subset 2	Subset 3
Canada thistle	14,657.42 (15)	14,763.96 (15)	14,914.79 (15)
Dalmation toadflax	9,513.46 (13)	9,575.50 (13)	9,575.65 (15)
Timothy	38,388.81 (14)	38,528.05 (13)	38,799.90 (14)

of confines: starting on a road and finishing 2,000 m from all roads but at all times traversing more than 2,000 m from any known trail; starting on a trail and finishing 2,000 m from all trails but at all times traversing more than 2,000 m from any known road; and starting on a road or trail and finishing 2,000 m from all ROW.

Transects were walked and location and other data recorded with a Global Positioning System (GPS) by two-person teams. Trimble Pro XR and GeoExplorer3® GPS receivers[1] were used, and the data were differentially postprocessed to improve positional accuracy (mean horizontal precision was 1.5 m). The coordinate system and projection used was Universal Transverse Mercator (UTM) Zone 12N, WGS 1984 Datum. Along each transect when a target NIS was intersected, the length of the patch was recorded in the GPS data dictionary. Additional location data were also recorded along each transect. All these data were used to generate continuous NIS data using extensions we created in Arcview[2] (Version 3.2) and an Excel[3] macro. The continuous data were generated at 10- by 10-m resolution.

Environmental and remote sensing data were used as independent variables. To generate predictive NIS maps of the entire area of interest, we need to have variable information of the entire area. Therefore, we used the environmental data from digital elevation maps (10-m resolution) and remote sensing data (30-m resolution). The environmental data including aspect, elevation, slope, and solar insolation were calculated from 10-m resolution digital elevation map; distance from roads and trails were calculated from data layers within the GIS database. Solar insolation was calculated for the summer months using only Swift's method (Swift 1976). LANDSAT ETM+ remote sensing data, acquired July 13, 1999, were included as individual spectral bands and as an unsupervised classification layer. The unsupervised classification layer was generated using ISOCLUSTER in ERDAS Imagine,[4] and 128 classes were identified. These classes were used by Legleiter et al. (2003) to develop a land-cover map of the Yellowstone watershed with accuracies of between 63 and 100% for individual land-cover classes. The 128 individual ISOCLUSTER classes were used in this analysis. The 30-m resolution Bands 1 to 5 and 7 of the LANDSAT ETM+ data were pan-sharpened to 15-m resolution with the panchromatic data from LANDSAT ETM+ Band 8 and resampled to 10-m resolution using nearest neighbor resampling so that the resolution of the LANDSAT ETM+ data matched the resolution of the digital elevation model available for the study area.

All these data layers were queried at 10-m intervals along the continuous sampling transects, and the sample values stored in the transect attribute database in Arcview. Thus, the final data set contained presence and absence points for 28 NIS, eight environmental variables (aspect was transformed into cosine and sine of aspect), six LANDSAT ETM+ bands, and one unsupervised classification layer, at 52,896 locations. Twenty percent of the data (n = 10,579) were randomly selected from the main data set and set aside to validate the accuracy of the probability models and maps. This random selection was performed thrice to produce Data Subsets 1 to 3, which contained the majority of the data (n = 42,317).

The three subsets of data were analyzed with generalized linear regression models, with binomial distribution and logit link in S-PLUS 2000.[5] Generalized logistic models (GLM) were used because the dependent variable—the NIS species data—is binary (presence and absence) data. The best model was determined with backward stepwise procedure using Akaike's Information Criterion (AIC), where the change in AIC value between models is used to define the "best" model, with the lowest AIC value representing the best model fit (Akaike 1977; Burnham and Anderson 1998). In our analysis, we determined three best models for each of the data subsets, using the AIC value for model selection. The three best models were selected using the reflectance data and environmental data variables, separately and combined. This was to determine if only one type of data were available—i.e., environmental or reflectance, which would make a better model, and how do those models compare with models from the combined data. All the analyses were performed in S-PLUS 2000.

Probability of occurrence predictions and maps of the target species were generated using coefficient values from the GLM applied to continuous spatial variables in rasterized format, using an extension we wrote in Arcview. The extension generated the logit of the GLM by summing the product of each variable in the model and its coefficient value, plus the beta-intercept value. The value of each cell in the output raster, ranging from zero to one, represents the probability that the target species could be present within the area defined by that cell. In this study, the raster cell size was 10 by 10 m. The validation data points, which were not used in the GLM, were overlaid on the appropriate probability maps in the GIS. At each validation data point, the predicted probability value was recorded and collated into 10 probability classes. The frequencies of occurrence were then calculated for the associated validation data, for each target species.

Three target species were chosen for the analysis with GLM and development of probability of occurrence maps. These were: Canada thistle, a wind-dispersed species with rhizomatous growth; dalmation toadflax, a non–wind dispersed rhizomatous species; and timothy, a non–wind dispersed nonrhizomatous species.

Results and Discussion

The frequency of Canada thistle was 5%, dalmation toadflax 3%, and timothy 23% within the area surveyed. Al-

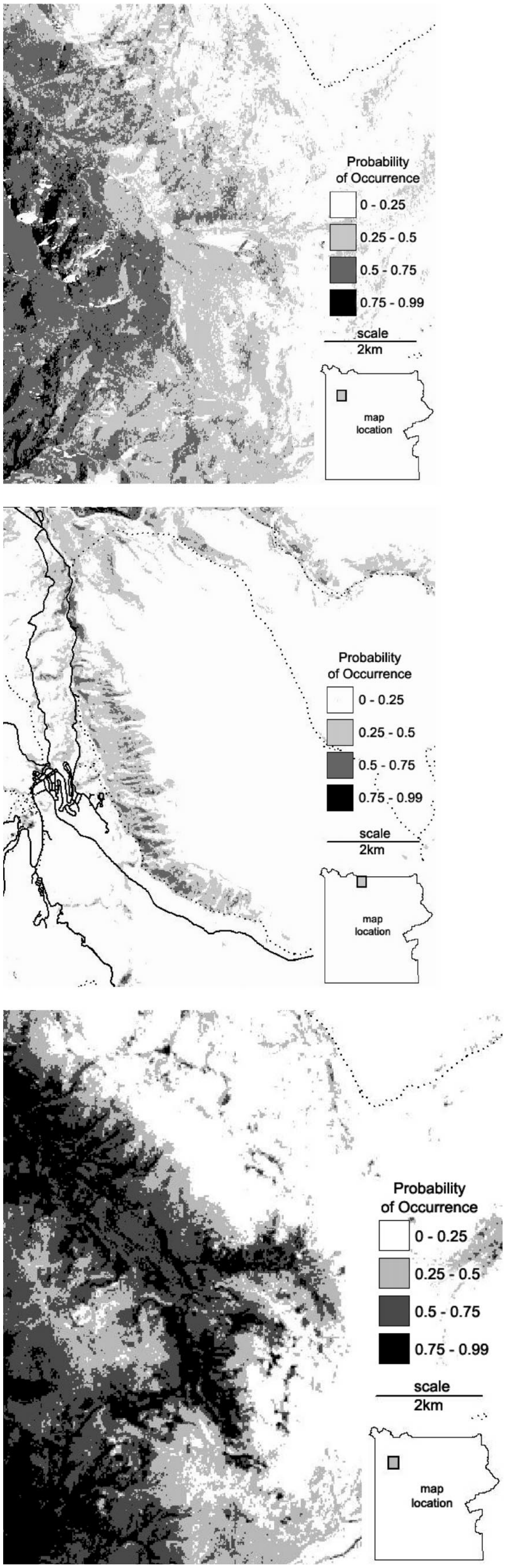


FIGURE 1. Predicted probability of occurrence maps for (a) Canada thistle, (b) dalmation toadflax, and (c) timothy for selected areas of the northern range of Yellowstone National Park. Solid lines represent roads; dashed lines represent trails.

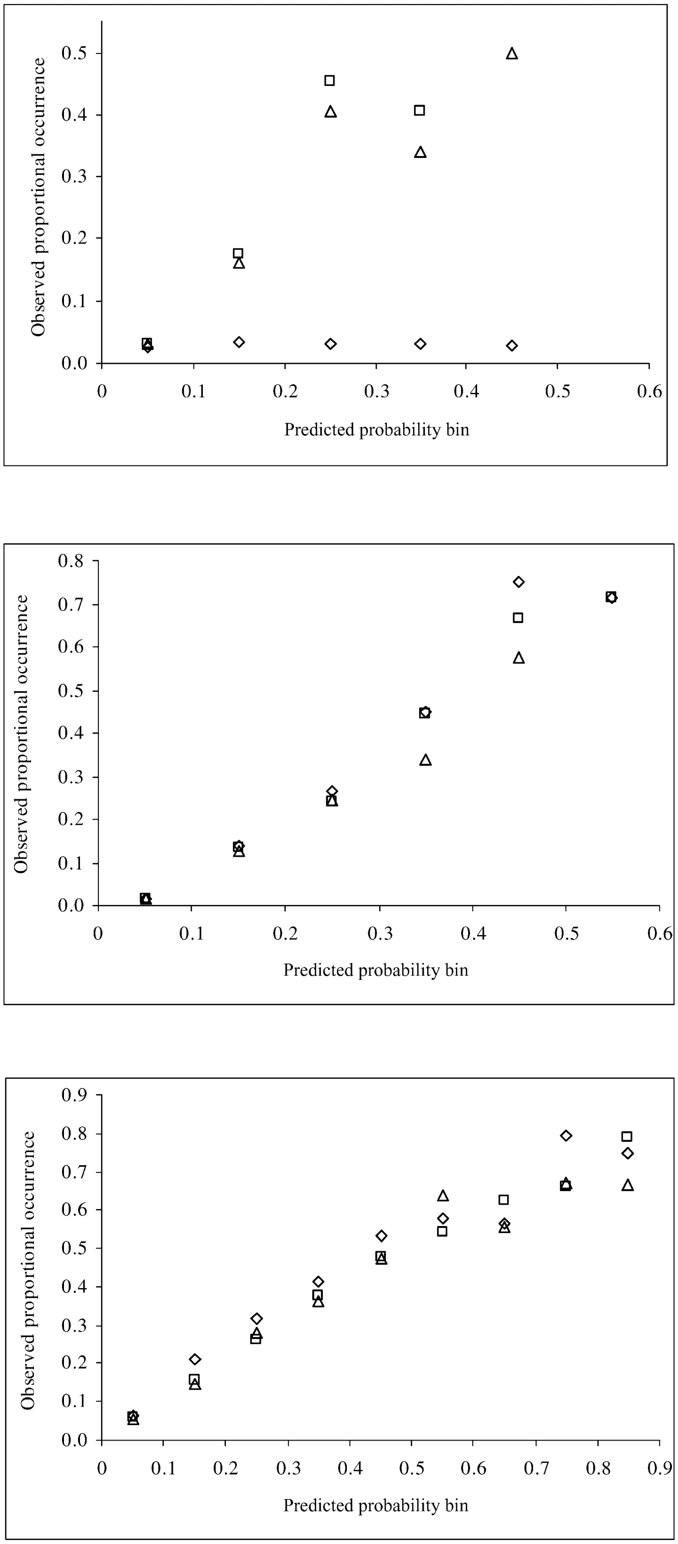


FIGURE 2. Observed frequency of occurrence of the validation data plotted against the predicted probability values, collated into 10 classes for (a) Canada thistle, (b) dalmation toadflax, and (c) timothy. ◇, □, and △ represent validation Data Sets 1, 2, and 3, respectively.

though these values are of interest, they provide no information to improve our understanding of where the species occurred on the landscape. Analyzing the binary NIS data using GLM provides some indication of the environmental variables that are associated with the occurrence of target NIS.

The occurrence of Canada thistle, dalmation toadflax, and timothy was correlated with most of the environmental variables and the reflectance measurements of the remote sensing bands, but the importance of the independent variables differed for the three target species (Table 1). This demonstrates that the occurrence of the target species is driven by numerous environmental parameters, but none of the target species had very specific associations with any one of the variables measured.

If only one type of predictor variable, either the environmental or remotely sensed data, were fit with a GLM, the environmental data produced a better model for Canada thistle and dalmation toadflax occurrence than the LANDSAT ETM+ data, whereas the converse was true for timothy (Table 2). Selecting the best model fit from all the available variables always provided a better model than environmental or reflectance variables separately (Table 2—only results from Subset 1 shown). However, the number of variables retained in the best model differed according to the subset of data used, and this was reflected in the different AIC values (Table 3).

Predictive maps of Canada thistle, dalmation toadflax, and timothy were generated from the best models for each data set, which happened to be Subset 1. Examples of approximately 10- by 10-km areas are provided for display purposes (Figures 1a–c); these smaller areas provide better observation of the probability maps than those of the entire area. The validation data sets were then used to evaluate the agreement between the predicted probabilities and the observed frequencies of occurrence. For example, if 200 validation points were recorded in probability class 0.1 to 0.2, we would expect on average 30 presences and 170 absences; in the probability class 0.6 to 0.7 we would expect 130 presences and 70 absences, etc. Agreement between the validation data and the probability predictions was better at the lower than at the higher occurrence probabilities for each of the target species (Figure 2) because too few of the validation data were located within the higher probability classes. And, because the model is predicting locations where the target species is more or less likely to establish and survive, although it may not have arrived there yet. Agreement between the observed and predicted data was good for timothy, particularly in the first six classes, with more variability in the agreement for the next three probability classes (0.6 to 0.7, 0.7 to 0.8, and 0.8 to 0.9); insufficient validation data were recorded in the 0.9 to 1 classes for comparison. Observed vs. predicted agreement of the dalmation toadflax data was good; there was more variation between the validation data sets for the 0.3 to 0.4 and 0.4 to 0.5 classes (Figure 2). Insufficient occurrence data were located in the higher probability classes (more than 0.6). Variation between the validation data sets was greatest for Canada thistle, with Validation Sets 2 and 3 providing similar results to each other but different to Validation Set 1 (Figure 2). Canada thistle model performance was poor for the lower probability classes, and insufficient validation data were available for probabilities greater than 0.5 (Figures 2a–c). This was expected on the basis of the high-residual sum of squares for Canada thistle compared with the other two species.

Conclusions

Many management areas are too large to sample entirely, so developing predictive maps of species occurrence provides information on conditions that are conducive for that species. However, in order for such predictions to be accurate, it is important that the survey methods used to collect the data on which the probability maps are based are unbiased and sample the environmental conditions present in the study area (Hirzel and Guisan 2002). Sampling may be randomly stratified on variables or gradients that are believed to be associated with a species distribution (Hirzel and Guisan 2002). However, stratifying on a number of variables becomes more complex as the number of target species increases because each species may have a different response to individual and multiple variables (Maggini et al. 2002). Therefore, Hirzel and Guisan (2002) suggest that unless correlations between the target species and variables are well known, sampling equally, not proportionally, within the multiple variables would probably be most effective. Because relationships between NIS occurrence and environmental variables are poorly understood, and the knowledge we do have suggests that species respond differently, we stratified on the one variable which is known to be important, proximity to ROW, and extended transects 2,000 m from ROW to provide the best possibility of sampling all environments equally. Analysis of the data using GLM with logit link provided information on target species correlations with environmental and reflectance data variables. Output from these models produced good predictions, particularly for dalmation toadflax and timothy, and we believe that the approach shows potential and will be validated further. Accurate probability maps of species occurrence could be used by land managers to prioritize where to spend the limited resources available for managing NIS in wildland and rangeland areas.

Sources of Materials

[1] Trimble Pro XR and GeoExplorer3® GPS receivers, Trimble Navigation Limited, 749 North Mary Avenue, Sunnyvale, CA 94085.

[2] Arcview (Version 3.2) ESRI Inc., 380 New York Street, Redlands, CA 92373-8100.

[3] Excel (Microsoft Excel 2002®), Microsoft Corporation, 1 Microsoft Way, Redmond, WA 98052-6399.

[4] ERDAS Imagine, Leica Geosystems GIS & Mapping, LLC, Worldwide Headquarters, 2801 Buford Highway, N.E., Atlanta, GA 30329-2137.

[5] S-PLUS 2000: Mathsoft Inc., 1700 Westlake Avenue North, Suite 500, Seattle, WA 98109.

Acknowledgments

This project was partially funded by the Yellowstone Inventory and Monitoring program. We thank Frank Dougher for writing Arcview extensions and collating the field data; the field crew including Judit Barroso, Jerad Corbin, Matthew Hulbert, Jeremy Gay, Mara Johnson, Rebecca Kennedy, Erik Lehnoff, Ben Levy, Amanda Morrison, Nathaniel Ohler, Charles Repath, John Rose, and Kit Sawyer; and Ann Rodman and the GIS lab in Yellowstone National Park for sharing their GIS data layers.

Literature Cited

Akaike, H. 1977. Likelihood of a model and information criteria. J. Econom. 16:3–14.

Burnham, K. P. and D. R. Anderson. 1998. Model Selection and Inference: A Practical Information-Theoretic Approach. New York: Springer-Verlag. 353 p.

Franklin, J. 1995. Predictive vegetation mapping: geographical modelling of biospatial patterns in relation to environmental gradients. Prog. Phys. Geogr. 19:474–499.

Gelbard, J. L. and J. Belnap. 2003. Roads as conduits for exotic plant invasions in a semiarid landscape. Conserv. Biol. 17:420–432.

Grime, J. P. 1979. Plant Strategies and Vegetation Processes. Chichester, UK: J. Wiley. 222 p.

Guisan, A. and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135:147–186.

Hirzel, A. and A. Guisan. 2002. Which is the optimal sampling strategy for habitat suitability modeling? Ecol. Model. 157:331–341.

Hitchcock, C. L. and A. Cronquist. 2001. Flora of the Pacific Northwest. 12th ed. Seattle, WA: University of Washington Press. 730 p.

Legleiter, C. J., R. L. Lawrence, M. A. Fonstad, W. A. Marcus, and R. J. Aspinall. 2003. Fluvial response a decade after wildfire in the northern Yellowstone ecosystem: a spatially explicit analysis. Geomorphology 54:119–136.

Mack, R. N. and J. N. Thompson. 1982. Evolution in steppe with few, large, hooved mammals. Am. Nat. 119:757–773.

Maggini, R., A. Guisan, and D. Cherix. 2002. A stratified approach for modelling the distribution of a threatened ant species in the Swiss National Park. Biodivers. Conserv. 11:2117–2141.

National Park Service. 1996. Preserving our Natural Heritage—A Strategic Plan for Managing Invasive Non-Indigenous Plants on National Park System Lands. www.nature.nps.gov/biology/invasivespecies/strat_pl.htm.

Parendes, L. A. and J. A. Jones. 2000. Role of light availability and dispersal in exotic plant invasion along roads and streams in the H. J. Andrews Experimental Forest, Oregon. Conserv. Biol. 14:64–75.

Shafii, B., W. J. Price, T. S. Prather, L. W. Lass, and D. C. Thill. 2003. Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics. Weed Sci. 52:748–751.

Swift, L. W. 1976. Algorithm for solar radiation on mountain slopes. Water Res. 12:108–112.

Trombulak, S. C. and C. A. Frissell. 2000. Review of ecological effects of roads on terrestrial and aquatic communities. Conserv. Biol. 14:18–30.

Tyser, R. W. and C. H. Key. 1988. Spotted knapweed in natural area fescue grasslands: an ecological assessment. Northwest Sci. 62:151–160.

Tyser, R. W. and C. A. Worley. 1992. Alien flora in grasslands adjacent to road and trail corridors in Glacier National Park, Montana (U.S.A.). Conserv. Biol. 6:253–262.

Watkins, R. Z., J. Chen, J. Pickens, and K. D. Brosofske. 2003. Effects of forest roads on understory plants in a managed hardwood landscape. Conserv. Biol. 17:411–419.

Whipple, J. J. 2001. Annotated checklist of exotic vascular plants in Yellowstone National Park. West. N. Am. Nat. 61:336–346.

Young, J. A., R. A. Evan, and J. Major. 1972. Alien plants in the Great Basin. J. Range Manag. 25:194–201.

Received May 20, 2004, and approved November 16, 2004.

Biological Invasions Special Issue, 2005 In press

Searching for a needle in a haystack: evaluating survey methods for non-indigenous plant species.

Lisa J. Rew[1*], Bruce D. Maxwell[1], Frank L. Dougher[1] & Richard Aspinall[2]

[1] Land Resources and Environmental Sciences Department and [2] Geographic Information and Analysis Center, Montana State University, Bozeman, MT 59718.

[*] Author for correspondence (email: lrew@montana.edu; fax: +1 406 994 7966

Keywords: survey, inventory, sampling method, stratified sampling, non-native species, invasive species, weeds, simulation model, weed management.

Abstract

The control and management of non-indigenous plant species (NIS) can be conceptually defined into three phases: inventory/survey, monitoring and management. Here we focus on phase one, determining which species are present and where they are located within the environment. Sampling for NIS is inherently time consuming and thus costly. Many management areas are large and therefore can only be surveyed (partial observation of the total area by sampling) and not inventoried (total observation of area). Survey data should reflect the spatial distribution of the target species populations over the landscape. Such data can then be used in combination with environmental data, to create probability maps of target species occurrence for the entire area of interest. We used a GIS model to evaluate seven different survey methods for consistency and reliability of intersecting NIS species' patches and producing samples which reflected the spatial distribution of the population, and which can be performed in a cost and time efficient manner. The GIS model was developed to create NIS populations which were then sampled using the different survey methods, and the results recorded. To improve the applicability of the model, four patch sizes and levels of occurrence were used, along with random and weighted distribution patterns in relation to patch proximity to roads and trails. Grid and random points, and targeted (stratified continuous) transects (starting on a road or trail (RoW) and finishing 2 km from any RoW) methods provided the most consistent samples of the population. Logistically, point methods required an unrealistic distance and time commitment in comparison with transect methods. The importance of collecting information on the size of NIS patches was demonstrated as more small patches were intersected than larger ones when the area infested was held constant. Thus, if frequency of patches is used to explain the results of a survey then comparisons between species and methods are difficult to interpret thus leading to erroneous conclusions. However, using percentage of area infested estimates provides for easier comparison between species and sample methods. The targeted transect method provided the most reliable, efficient and consistent sample with the expected spatial distribution.

Abbreviations: NIS - non-indigenous plant species; RoW - right of way e.g. roads and trails; SaD - Seek and destroy survey method; perp.RoW - perpendicular to right of way transect survey method; YNP - Yellowstone National Park.

Introduction

Invasion of natural communities by non-indigenous plant species is a threat to native biodiversity and is currently rated as one of the most important global-scale environmental problems (Vitousek et al. 1996). Significant effort is expended to manage or eliminate non-indigenous plant populations in natural and semi-natural areas.

In the USA, the National Park Service is required by law to inventory the 34 million ha of Parks classified as "natural areas" under their jurisdiction and to maintain these in a state as "unaltered by human activities as possible" (U.S. National Park Service 1996). Stohlgren et al. (1995) completed a report on the status of the inventory lists for flora and fauna of National Parks; they found that less than 80% of inventories were complete in taxonomic, geographical and ecological coverage. The sampling procedures varied between Parks and often provided just a list of species present in that Park. More recently there have been advances to ensure that data attributes of different studies are comparable (Cooksey & Sheley 1997; NAWMA 2002; Stohlgren et al. 2003).

There has been little work addressing the advantages and disadvantages of different sampling methods for large natural areas. Stolgren et al. (1995) stated that there is "no off the shelf sampling design and technique for optimizing geographical, ecological and taxonomic completeness in biotic inventories. Trade-offs are inevitable." In this paper we evaluate seven sampling/survey methods on their relative ability to detect non-indigenous species (NIS). Much of the terminology surrounding NIS has emotive connotations and is loosely defined and confusing (Rejmánek et al. 2002). Terms such as 'invasive', 'weed', and 'exotic' are often used interchangeably; we use the term 'non-indigenous species' to represent all of these terms (see Maxwell et al. 2003 for definitions). Non-indigenous species are those that are not originally from the area of interest, they could be from a different continent or a different area of the same continent.

Management of NIS should be regarded as a three phase process: inventory/survey, monitoring and management. Although all three phases can be performed simultaneously, considering them as different phases is important for conceptual, theoretical, methodological, logistical and practical reasons. The first phase, inventory/survey, determines which species are present and their distribution within the environment. The second phase, monitoring, provides information on how patches are changing with time, their impact on the ecosystem and the impact of management on the patch. Monitoring objectives define and constrain the objectives and methods for inventory/survey because the require locating sample populations across the widest possible set of environments which define the habitat for a species. The final phase, management, is important for control of non-indigenous patches and populations, by reducing their distribution and impact. This paper focuses on phase one, inventory/survey of NIS.

An inventory/survey is required to provide baseline information regarding which species are present and where, within an area. The "where" should include information on both NIS presence and absence. For many natural areas it will not be possible to inventory the entire area. Therefore, a survey that obtains a representative sample of the population is required so that the frequency of the species in the area of interest can be determined. In addition, the survey protocol should collect information on environmental variables associated with both presence and absence of NIS, plus an approximation of patch size. Environmental data help to improve our understanding of factors affecting NIS occurrence (Stohlgren et al. 2003). Correlations between the environmental variables and NIS occurrence data can then be used to produce probability of occurrence maps of target species (Franklin 1995; Guisan and Zimmerman 2000).

The design of a sampling method for a survey depends on the objectives of the study. For example, if the objective is to "seek and destroy" NIS, an approach that is biased towards areas perceived to have higher occurrences of such species would generally be chosen. If the objective is to evaluate the frequency and distribution of a species in the landscape, an unbiased approach should be used. In addition, if there is prior knowledge about a species distribution, or how the distribution is associated with particular variables that are known across the study area a stratified approach can be used (e.g. Fortin et al. 1989, Thompson 2002). If the intention is to use the survey data to develop probability of occurrence maps for the entire area of interest a stratified approach linked to empirical modelling can be employed (see Guisan & Zimmerman 2000, for review). However, even when NIS are widespread and occur locally at high abundance, they frequently are scarce when considering a whole landscape. In this case, when NIS have very low frequency in the environment, an adaptive cluster sampling method may be appropriate (Thompson 2002). In this paper we consider the question of sampling for NIS in a natural ecosystem; the same approach could be used for agricultural and rangeland situations or, for native plants that occur at low frequencies or with low density distributions.

Specifically, we evaluate sample methods for use when the study/management area is large and cannot be inventoried. To achieve this, a sampling methodology is required that is based on an understanding of plant distribution. The appropriate theories to guide sample design are drawn from an understanding of the geographical structure and ecological organization of plant communities. There are two general categories of theory. The more prevalent category is based on the view that plant communities are groups of interacting species whose presence, absence, or even relative abundance can be deduced from "assembly rules" which are based on the functional roles of each species (inherently defined ecological niches) (Levins 1970; MacArthur 1972; Diamond 1975; Weiher & Keddy 1999). According to this theory, species coexist in interactive equilibrium with the other species in the community. The stability of the community and its resistance to invasion are derived from the adaptive equilibrium of member species, each of which has evolved to be the best competitor in its own ecological niche (i.e. niche separation). The theory assumes that the species which co-occur within a community are determined by interspecific competition for limited resources and other biotic interactions. That is, species distributions follow Gaussian curves but the curves overlap in environmental space. The other general theory that has emerged to explain community structure is based on an assumption that communities are open, nonequilibrium assemblages of species largely thrown together by chance, history including genetic drift, and random dispersal (Hubbell 2001). Thus, this theory assumes that species come and go, their presence or absence is dictated by random dispersal and stochastic local establishment and extinction in environmental space. Hubbell (2001) called this the "Unified Neutral Theory" and contrasted it with the most prevalent theory described above, which he called the "Niche-Assembly Theory".

If one considers the unified neutral theory as a null hypothesis then a first order determinant of NIS would be associated with dispersal. The second order determinant of NIS would be environmental variables, i.e. niche assembly theory. Both first and second order determinants would be subject to random factors. The mix of theories associated with plant distribution could be used to establish stratification of sampling.

Disturbance is often suggested as a key factor in enhancing the probability of non-indigenous plant establishment in plant communities (Grime 1979). The role of anthropogenic disturbances such as cultivation, trampling, domestic ungulates (Young et al. 1972; Mack & Thompson 1982; Tyser & Key 1988) and roads/trails (Tyser & Worley 1992; Parendes & Jones 2000; Trombulak & Frissell 2000; Gelbard & Belnap 2003; Watkins et al. 2003; Pauchard & Alaback 2004) and fire management practices (Godwin et al. 2002) are often seen as more influential than natural disturbances. Natural disturbances include soil disturbance by fauna, weather related events including floods, wind damage and fires, and geological events such as landslides, mudflows and floods. Both anthropogenic and natural disturbance, depending on the intensity and extent, can destabilize the community equilibrium. Or, put another way they add noise to the metrics used to measure community stability such as relative abundance and frequency.

There is sufficient evidence to suggest that anthropogenic disturbances, particularly rights of way (e.g. roads and trails) affect the distribution of NIS (Tyser & Worley 1992; Spellerberg 1998; Parendes & Jones 2000; Gelbard & Belnap 2003; Watkins et al. 2003; Pauchard & Alaback 2004). Therefore, in this study we assume that the distribution of NIS patches is weighted towards roads or trails, such that the abundance of patches increases as proximity to rights of way (RoW) increases. The weighted distributions were compared with random distributions. Although, under both scenarios spatial dependence will be displayed at some spatial scale because of the geographic pattern of environmental factors or other spatial processes such as plant dispersal.

We evaluate one new design and six established survey methods for detecting NIS in natural ecosystems. The established methods are currently employed by land managers and scientific researchers. The methods were evaluated to determine the accuracy of the different approaches in terms of intersecting (detecting) target patches, how representative these data were of the population and, to compare the distance and time required to complete sampling in the field.

Materials and methods

A model was developed to evaluate seven different sampling methodologies using standard and adaptive sampling designs, on random and weighted distributions of four virtual NIS. The model has a simulation component and a sampling component. The simulation component creates distribution and abundance of the NIS. The sampling component implements the seven methods with standard and adaptive designs, to survey the NIS distribution and abundance. Evaluating these survey methods as they would be applied in the field is termed the standard sampling approach. The adaptive cluster sampling design was "added-on" to the standard sampling approach. For clarity we will refer to the seven different sampling methods as survey methods; population will only be used in reference to the simulated "true" population; a sample (of the population) will be used to refer to data produced by an iteration of the simulation model using one of the survey methods; and a patch will be used to refer to a circular group of plants.

The simulation model was developed in ArcView (Version 3.2, ESRI, Inc.). The area modelled was restricted to a randomly selected 10 000 ha area (10 km by 10 km) for computational speed. The area selected for the simulation was in the northern range of Yellowstone National Park (YNP), Wyoming, USA, since the most effective and efficient sampling method was to be used for an operational survey of this area. Environmental data, including the distribution of rights RoW, were available as data layers within the geographical information system (GIS) database in ArcView. Rights of way in the study area included about 12 km of primary road and 39 km of trails. No adjustment or allowance was made for elevation in the generation of the sampling points and transects.

Non-indigenous species abundance and spatial distribution

The four NIS distributions simulated provided a range of distribution and abundance patterns with regard to proximity to RoW, frequency in the study area and mean size of patch where present. To ensure that the simulated populations were representative of "real" situations their parameters were obtained from field data collected in YNP during 2001 and 2002 (Rew unpublished). The four species chosen to provide parameters for the virtual species simulations represent different NIS types and will be referred to as Species A, B, C and D.

Parameters used in the model included patch extent (mean area), percentage infestation (calculated as area covered by patches / the total area surveyed) (Table 1) and a distribution weighting. The distribution weighting was created by fitting the patch-occurrence probability along a distribution curve relative to distance from RoW. The parameters used to generate the distribution curve were estimated from field data describing the distribution of NIS patches relative to RoW. The resultant simulated distributions are more densely populated in the areas proximal to RoW, and more sparsely populated in areas distant from RoW (Figure 1). This is consistent with using the dispersal vector as a first order process assumption (unified null hypothesis). The simulation model was run 100 times for each species, generating 50 random and 50 weighted population distributions.

Survey methods

Seven survey methods were used; one biased and six unbiased (Figure 2). The biased survey method reflects a method used by many land managers who use their knowledge of NIS distributions to develop a search pattern along RoW, particularly roads. We term this the "seek and destroy" (SaD) approach. In the model this approach is performed along RoW, using transects. Of the six unbiased methods, one was systematic, two were totally random and the remaining three were random stratified on RoW or an environmental feature (Figure 2). The systematic method used a rectangular grid and sampling was performed at the grid intersections. The five random, unsystematic methods included two totally random methods; random points and random walk. No limitations were placed on the pattern that the random walk transect could take, other than those required to keep the walking path within the study area. The three random stratified methods included two transect methods which were stratified on RoW. That is, transects started on a RoW; but the selection of the transect start points on the RoW was random. For the first of these methods the transect travelled perpendicular to the randomly located start position. This survey method was termed "perp.RoW". The start points for the second random-stratified method were established in the same way but it was ensured that transects ended 2 km from any RoW and the start location, these are termed "targeted" transects. The third random-stratified method used transects stratified on elevation contours, a technique used by some groups to reduce loss and gain of elevation when completing transects (S. Dewey, pers. comm.), and these were termed "contour" transects.

The total area surveyed within the simulation boundary was held constant at 20 ha for all methods. This represents 0.2% of the total area. The simulation employed a search radius of 5 m at grid intersections and for the random point method. A total of 2546 sample points were required to survey 20 ha. All of the other survey methods used transects which were 10 m wide and 2000 m long. For transects along rights of way (SaD) and elevation contours (contour) minor changes in the length of individual transects had to be permitted due to insufficient length of RoW or contours in some situations. The total area sampled was maintained at 20 ha for all transect methods. Overlap in the area surveyed was not permitted for point methods. No overlap was permitted for a distance of more than 10 m for transect methods. The 10 m overlap was permitted as some transect methods, particularly the random walk, were allowed to cross over previous survey lines. The positions of survey points for the grid method, contour and SaD had to be held constant between iterations as the small study area limited the number of configurations possible. The location of sample points/lines was random and changed for each iteration for all other methods.

The adaptive cluster design (Thompson 2002) was added to the standard survey methods, with the intent of evaluating the increased number of patches intersected when using such an approach. Assuming that presence of NIS patches is spatially autocorrelated, this strategy will gather more data at locations of NIS presence by expanding the search parameters in these areas. The adaptive procedure in the model applies an increased search radius of 50 m (in consecutive 10 m wide circles) around each NIS patch detected by the standard survey methods (Figure 3). The application of a 50 m search radius is iteratively applied to any subsequently detected NIS patches until no further patches are intersected within the search radius. The results of using the adaptive cluster design, in addition to the standard sampling design, were recorded as an independent dataset for each of the seven survey methods. The results of the standard and adaptive sampling design were evaluated independently.

Additional simulations were performed to demonstrate the effect of the four different patch radii on the number of patches intersected, when the overall area covered by the four species' populations was held stable. The populations were then sampled using grid and targeted transect survey methods. These additional simulations were only performed on random distributions and 50 iterations were run for each species.

Model output & analysis

The simulation model and all sampling methods were developed as a module (SampleMod) in ArcView 3.2 using the Avenue programming language. For each model iteration SampleMod generated a new random and weighted NIS population distribution, as well as new sampling locations for each of the survey methods where possible.

For every iteration of the model each of the seven survey methods was applied and the number of target patches intersected by each method recorded. We could have designed the model only to count patches where all or the majority of the patch was within the sampled area (transect or sampling point). However, this is not how people sample in the field. Thus, patches were tallied in the results if any part of them was intersected by the survey transect or point. The simulated population mean (μ) and standard deviation (σ) of distance to RoW were calculated for each iteration, as well as the mean ($\bar{x}$) and standard deviation (s) of the patches intersected by each survey method. These statistics were recorded in separate tables for weighted and random NIS distributions using standard and adaptive designs, with separate records for each iteration of the model. Thus, survey methods which produced a $\bar{x}$ close to μ, and an s close to σ were considered superior to methods which obtained samples with greater differences between the population and sample statistics.

With each iteration of the model the number of patches intersected, the mean and standard deviation of patch distance from RoW was recorded for each survey method. Distance to RoW was chosen as the characteristic measure for these comparisons, primarily because RoW was the principal disturbance factor upon which we were basing the distributions of our populations, and would thus yield the most reliable information about the similarity of the populations and samples.

Results from sampling with the different survey methods were evaluated in several ways. The number of patches intersected by each survey method was used to calculate a percentage frequency (using the number of sample points in grid and random point surveys); the percentage of the sample area infested by the different number of patches was calculated using an average value for the proportion of each patch being intersected (this was derived using geometry). The distribution of the sample means with proximity to RoW was depicted graphically with box and whisker plots. A paired *t*-test was used to determine if the distribution of patches in proximity to RoW intersected with the targeted method differed from samples generated using the other methods. Plus, the minimum distance to be travelled to reach all the survey sites for each method was measured using the survey locations depicted in Figure 2.

The least-cost paths to sample each of the survey points or transects was calculated for each of the survey methods. The least-cost/minimum distance path to connect all 2546 points/ 10 transects was calculated by creating polylines to connect each of the points/transects in ArcView GIS. The time taken to cover the least-cost paths was calculated assuming a walking speed of 1 km/hr when sampling and 3 km/hr when walking between sampling locations. All least-cost path estimates assume: completing the entire survey in one period, rather than taking into account end-of-day return trips; and zero-gradient terrain, rather than taking topography and land cover into account. Thus, time and distance estimates are accepted to be underestimates of the true costs of applying these methods in the field, but apply equally to each method.

Analysis and display of data generated in SampleMod were performed in Arcview 3.2, S-Plus 2000 (Mathsoft Inc.), SigmaPlot 2001 (SPSS, Inc.) and Excel 2002(Microsoft®).

Results

The point survey methods detected considerably more patches than the transect methods. As expected adding the adaptive sampling design to the standard sampling generally increased the number of patches intersected for all survey methods (Table 2). Additionally, when a species distribution is correlated more definitively with RoW (i.e. weighted distribution) the adaptive design combined with the SaD method would be expected to show a dramatic increase in the number of patches intersected. This was true for the Species A – C when using the SaD method, resulting in more patch intersections for those species compared with other transect or point methods. Similarly more interactions were made when using the SaD method plus adaptive sampling for Species B and C. Species D patches showed less correlation with RoW (Figure 1) and, as expected, the proportion of patches intersected did not show much, if any, increase when sampled with the SaD method or any method plus the adaptive design (Table 2).

Distribution of the samples

Sample data collected in a survey should reflect the distribution of the population, which in this model scenario is measured as proximity to RoW. When choosing a survey method one needs to know how well and how consistently sample data reflect the population distribution. In the simulation environment we know the μ and σ of the population and the $\bar{X}$ and s of each sample. There were 50 samples (i.e. model iterations) for each survey method, and box and whisker plots were used to graphically depict these data and aid comparison between sample and population statistics (Figure 4).

Grid and point methods provided samples with similar mean and median patch proximity to RoW (i.e. narrow boxes and whiskers) as populations, for all species and both standard and weighted distributions (Figure 4). Samples generated using the targeted method also produced narrow boxes (representing 50% of the data) and whiskers (10^{th} and 90^{th} percentiles), and the outliers (all remaining data) were generally not spread far on the *y*-axis (proximity to RoW) (Figure 4). Samples from the random walk method provided a mean and median value similar to that of the population for all 50 iterations/samples but the length of the whiskers and position of the outliers demonstrate the variation between the samples for that method (Figure 4). This shows that although the random walk method can provide some samples with similar mean distance to RoW distributions as the population there is a lot of variability. This makes intuitive sense, some random walks will traverse close to RoW, others in the backcountry, while many will traverse the full extent. Samples from the contour methods obtained mean and median values further from RoW than those of the population, and there was high sample variation. In contrast, samples collected using the perp.RoW transects systematically underestimated the median distance to RoW. The SaD method samples provided the narrowest box and whiskers of any method but the distribution of samples in relation to proximity to RoW did not reflect the population for any of the species. Thus, samples generated with the grid, point and targeted methods provided the least sample variation of all methods, with means and medians most similar to those of the populations of each of the virtual NIS.

Distance and time taken to complete surveys

The results above have shown that point and grid methods intersected more patches than transect methods, and that point, grid and targeted methods provided samples that most consistently represented the distribution of the population relative to RoW. However, there could be some logistical advantages of particular methods. Twenty hectares were sampled by each of the survey methods. The additional distance to be traversed to link all the sample points or transects together was calculated assuming that each method was walked in one period and that the most efficient (minimum distance) route was taken. The distances were calculated for the sample locations shown in Figure 2. Approximately 480 km would have to be walked to reach all grid intersections and random points; in comparison, 34 – 48 km would have to be traversed to reach all transects (Table 3). A speed of 3 km/hr was assumed when traversing between sampling areas and 1 km/hr when sampling and recording data. The time taken to complete the transect survey methods was between 31.3 and 36 hours (Table 3). The grid and point methods were estimated to take 180 and approximately 178 hours, respectively. Therefore, for logistical purposes of time and distance traversed in the field, the transect survey methods proved to be most cost-effective.

Comparison between survey methods

From the above evaluations of the model output we can determine that transect methods provide the most effective logistical and financial approach. Targeted transects provided the second shortest and quickest method to implement in the field, and the most representative and consistent sample of a population. Therefore, the targeted method was used as the benchmark against which the other six survey methods were compared. Paired *t*-tests were performed between each of the methods and the targeted approach using proximity to RoW sample data. The proportion of pairs for which the method of interest showed no significant difference from the targeted method are provided in Table 4. Grid, point, walk and contour samples were not significantly different from the target approach samples for the majority of pairs. Agreement of less than 0.7 was more frequent when using the adaptive design, particularly for Species A and B. As expected, perp.RoW samples generally (2 exceptions) showed poor agreement with mean proximity to RoW samples of the targeted methods. No samples generated using the SaD methods had spatial distributions similar to samples generated using the targeted method.

Effect of patch size

The initial model simulation demonstrated that point sampling methods, both grid and random point, intersected many more patches than the transect methods for both the random and weighted NIS distributions. This result was not intuitive as the area sampled equalled 20 ha (20 0000 m^2) for both point and transect methods. The perimeter of the transects equalled 40200 m but the circumference of all the survey points totalled 79985 m, which is almost twice that of the transects. This partially explains why more patches were intersected using the point rather than transect approach. In addition, the total circumference provided by the different patch sizes was positively related to the chance of a patch being intersected (data not shown).

To investigate this issue in more detail we re-ran the model, retaining the four different patch radii (Species A < Species B < Species C < Species D) but ensured that each of the virtual species covered the same area (1.3% of the total area, 1278884 m^2). This was only performed with random distributions, and grid and targeted survey methods. The area surveyed was the same (20 ha) and the results were recorded as before.

When the area covered by the four species was held constant more small radii patches (Species A) were intersected than larger patches (consecutively Species B, C, D) for each of the survey methods (Table 5). As before, more patches were intersected with the grid than targeted method. The different patch numbers intersected by the two survey methods has important implications for estimates of species frequency (number patches/number of observation points). The percentage frequency, sometimes referred to as occurrence, of the smaller radii species is higher. This would lead to incorrect estimates of NIS frequency in the environment. Using information on patch size the percentage, or proportion, of the survey area infested can be calculated. The actual area of each patch intersected was recorded with an additional extension to the model. However, this type of information is generally not recorded in the field so we also calculated the percentage area infested using an average value for the proportion of a patch intersected by each transect or grid method. The resulting infestation estimates are reasonably accurate and would provide a more accurate reflection of the importance of the different species than frequency calculations (Table 5).

The percentage frequency and percentage of the area infested were also calculated for the original data set (Table 2). As expected the percentage frequency values provide little useful information when used to compare across survey methods. Point methods for the same species and distribution provided similar frequency estimates. Transect methods generally provided similar results for the same species with random distributions, but showed less agreement between methods for the weighted distributions (Table 2). This suggests that comparing percentage frequency values between methods or studies would not provide valuable or accurate insight into the data. Relative species frequencies were not calculated for this study.

Calculations of the percentage of the area infested are easier to interpret than frequency estimates. To estimate the percentage of the area infested one has to determine the mean proportion of each patch intersected, which can be achieved through probability and geometry. Percentage area infested calculations demonstrated that the point and target transects methods intersected the appropriate number of patches given the infestation level in the total area. The other transect methods over- or under- sampled the population. Percentage area infested values provided an additional approach for comparing between methods and species occurrence rate in the environment.

Discussion

Prior to developing the model we hypothesized that several of the survey methods would produce samples which would accurately reflect the population infestation extent and spatial distribution in the landscape. Grid and random point sampling methods intersected significantly more patches than the transect methods for both the random and weighted distributions, with the same sampled area. These results have important implications for comparing between studies where different survey methods were employed. If only percentage frequency calculations are performed, surveys conducted using point methods will be seen to have higher populations of NIS than surveys conducted with transect methods (SaD excepted) in the same area. In addition, patch size plays an important role. Populations with small patch size, here Species A, will be estimated to have a higher frequency than species with larger patches, here Species D. Limitations of frequency estimates are well known (e.g. Barbour et al. 1980; Crawley 1997; Thompson 2002) due to size and shape of plots and plants, and different environmental gradients etc. However, they are still used regularly in reports and are thus included in our evaluation. Calculating the percentage of the area infested provides for more reliable comparison between survey methods when the objectives are to estimate species occurrence as well as spatial distribution.

Trombulak and Frissel (2000) reviewed the ecological effects of roads and generally supported the view that roads are associated with negative biological impacts. Several studies have demonstrated that the abundance of NIS increases in closer proximity to RoW though there are differences in systems (Tyser & Worley 1992; Marcus et al. 1998; Parendes & Jones 2000; Gelbard & Belnap 2003; Maxwell et al. 2003; Watkins et al. 2003; Pauchard & Alaback 2004). Road density is high throughout continental USA (Forman 2000). For example, in the continental USA 20% of land is within 127 m of a road, 83% is within 1061 m and only 3% of land was more than 5176 m from a road (Riitters & Wickham 2003). Thus, RoW may provide an effective dispersal vector for species (e.g. Pauchard et al. 2003; Pauchard & Alaback 2004) aiding their dispersal into new environments. Consequently we chose proximity to RoW as an important influence on NIS distributions, thus allowing us to compare a biologically relevant distribution with a random distribution. The results show that the grid, point and targeted survey methods provided samples similar to the population for both random and weighted (proximity to RoW) distributions of NIS.

Most NIS data will be collected as presence/absence data but rarely are other potentially correlated variables recorded to aid prediction of the target species occurrence. Evaluating sampling methods for predictive purposes has recently received more attention. Hirzel & Guisan (2002) evaluated four different sampling methods to predict habitat types. They evaluated random, regular, proportionally-stratified and equal-stratification methods in different habitat types at four sample sizes. They found that the most important factor was number of samples taken, though systematic sampling provided better results than random and, including environmental information in the sampling design also improved the subsequent predictive models. They concluded that choosing the right sampling method can improve the results by a few percent and reduce the risk of making poor predictions; this was particularly apparent for presence/absence predictions. In their model stratified sampling did not improve the predictive models greatly. However, in our model stratifying on RoW (e.g. targeted) proved to be preferable to grid and random point due to practicalities of time in the field.

The Gradsect design is a variation of a random stratified design, which samples along environmental gradients. It represents a compromise between randomized sampling of stratified multiple variables in order to minimize travel and cost (Austin & Heylingers 1989). Maggini et al. (2002) used a random stratified sampling approach to collect data for predicting ant distribution. They used an equal number of replicates per variable (e.g. vegetation, slope, aspect, etc.). To use random stratified methods for multiple variables it is necessary to have access to information on the effect of each of the different variables on the distribution of the response variable. Maggini et al. (2002) stated that stratifying on a number of variables becomes more complex as the number of target species increases because each species may have a different response to individual and multiple variables. Hirzel & Guisan (2002) suggest that unless such correlations are well known sampling equally, not proportionally, within the multiple variables would probably be most effective but this needs to be evaluated further. As relationships between NIS occurrence and environmental variables are poorly understood, and the knowledge we do have suggests species respond differently, we stratified on the one variable which is known to be important, proximity to RoW. We do not intend to suggest that environmental variables do not influence species distribution, neither that the effects of such variables could not be greater than RoW nor, that when such information is available it should not be used to help stratify the sampling method. Our objective here is to evaluate different sampling methods when the number of NIS is not necessarily known, their relation with environmental variables is different or unknown, and when the sampling area is large. Because these are the situations that many land managers face when make sampling decisions. Under these situations it may only be practical to stratify on one variable - RoW.

Many survey studies only collect data on presence/absence of a species and make no record of patch size. Sampling populations which cover the same overall area but have different patch sizes demonstrated clearly that patch size should be included as part of the surveying protocol. That is patches have unequal probability of being sampled and the importance of considering this is explained by Thompson (2002). More smaller patches were intersected than larger ones when the area of NIS was held constant. Evaluating these data purely on numbers of presence/absence and not including data on the estimated area infested will lead to erroneous conclusions of species frequency and consequently their management. When data on the size of individual patches is not recorded as part of a survey protocol, the mean size can be determined from a sub-sample of patches of an individual species, and an estimate of the percentage or proportion of the area infested can be calculated as demonstrated in Table 2 and 5. Calculations of infested area and frequency will also be easier if transect data are collected continuously.

The four species modelled had different distributions relative to RoW, which was highlighted as more patches of three of the species (Species A-C) were intersected using the SaD method with weighted distributions than random. The SaD method was incorporated into the model as some managers assume that sampling along the main vectors of NIS spread will result in a large proportion of the patches being intersected for the same unit of time/effort. There is however, no evidence that this is correct since areas away from RoW are not searched. The simulation model demonstrated that while this method does result in a higher number of patches being intersected than the other transect methods, patches away from RoW are missed, even when using the adaptive design in conjunction with SaD sampling. It is often perceived that it is best to stratify management on new infestations and these may be ones at greater distances from RoW and thus those that would not be detected by the SaD method. Furthermore, the SaD method is essentially a Phase three operation, where the goal is not inventory/survey, but management. Data provided by this method of sampling can be misleading and lead to incorrect conclusions about the distribution and abundance of NIS. Additionally, since the data collected by SaD sampling are biased, they are also not suited to predicting and monitoring NIS distributions although these are important components of phases 1 and 2 of NIS management.

Our evaluation of the different survey methods suggests that none of the methods provide a perfect representation of the population but the targeted method is the most reliable and time efficient. Given these limitations, and if it is not possible to sample an area entirely, then it is prudent to collect additional information regarding environments and conditions where a species is *and* is not located. Such information can be used to increase our understanding of NIS distributions and can be updated periodically to sample unrepresented areas or as a result of geological, weather or anthropological disturbance. Pauchard et al (2003) studied the occurrence of *Linaria vulgaris* at three different scales around West Yellowstone, Wyoming, USA, and the results provided a better understanding of the pattern of *L. vulgaris* invasion and helped to prioritize management of this species. Using the targeted survey method provides information over an extensive area and at a range of scales, depending on how the data are collected in the field. This type of data collection provides information on rare NIS species and new populations which should be targeted for management. The data can be used to develop predictive models (see Rew et al. in press) of target species occurrence at appropriate scales (10-30 m) to help managers target their time managing NIS most effectively. Considering how well the survey method represents the population, and time in the field we feel that the targeted approach is most reliable.

Acknowledgements

This project was partially funded by the Yellowstone Inventory and Monitoring program. Thanks to Ann Rodman and the GIS lab in Yellowstone National Park for sharing their roads and trails data etc., and Nathaniel Ohler for considerable assistance with the geometric calculations.

References

Austin MP and Heyligers PC (1989) Vegetation survey design for conservation: gradsect sampling of forests in north-eastern NSW. *Biological Conservation* 50: 13-32.

Barbour MG, Burk JH and Pitts WD (1980) Terrestrial Plant Ecology. The Benjamin/Cummings Publishing Company, Inc., California, USA, 605 pp.

Cooksey D and Sheley R (1997) Montana noxious weed survey and mapping system. Montana State University Extension Service, MT, 9613, 20 pp.

Crawley MJ (1997) Plant Ecology. 2nd edition. Blackwell Science Ltd, Oxford, Britain, 717pp.

Diamond JM (1975) Assembly of species communities. In: Cody ML and Diamond JM (eds) Ecology and Evolution of Communities, pp 342-444. Belknap Press of Harvard University Press, Cambridge, MA, USA.

Franklin J (1995) Predictive vegetation mapping: geographical modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*. 19: 474-499.

Forman RTT (2000) Estimate of the area affected ecologically by the road system in the United States. *Conservation Biology*: 14: 31-35.

Fortin MJ, Drapeau P and Legendre P (1989) Spatial autocorrelation and sampling design in plant ecology. *Vegetatio* 83: 209-222.

Gelbard JL & Belnap J (2003) Roads as conduits for exotic plant invasions in a semiarid landscape. *Conservation Biology* 17: 420-432.

Godwin K, Sheley R and Clark J (2002) Integrated noxious weed management after wildfires. Montana State University Extension Service, MT, 9613, 46 pp.

Grime JP (1979) Plant Strategies and Vegetation Processes. John Wiley, Chichester, Britain, 222 pp.

Guisan A and Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147-186.

Hirzel A and Guisan A (2002) Which is the optimal sampling strategy for habitat suitability modeling. *Ecological Modeling* 157: 331-341.

Hubbell SP (2001) The Unified Neutral Theory Of Biodiversity And Biogeography. Princeton University Press. New Jersey, USA, 448 pp.

Levins R (1970) Extinction. In: Gerstenhaber M (ed) Some mathematical problems in biology, pp 75-107. American Mathematical Society, Providence, RI, USA.

Mack RN and Thompson JN (1982) Evolution in steppe with few, large, hooved mammals. *American Naturalist* 119: 757-773.

Maggini R, Guisan A and Cherix D (2002) A stratified approach for modelling the distribution of a threatened ant species in the Swiss National Park. *Biodiversity and Conservation* 11: 2117-2141.

Marcus WA, Milner G and Maxwell BD (1998) Spotted knapweed distribution in stock camps and trails of the Selway-Bitteroot wilderness. *Great Basin Naturalist* 58: 156-166.

NAWMA (2002) North American invasive plant mapping standards. Retrieved from http://www.nawma.org/documents/Mapping%20Standards/Invasive%20Plant%20Mapping%20Standards.pdf

MacArthur RH (1972) Geographical Ecology. Harper and Row, New York, USA. 269 pp.

Maxwell BD, Rew LJ and Aspinall R (2003) Exotic Plant Survey and Monitoring: Methods to Discover Distribution with Low Frequency Occurrence. In: Philippi T and Doren R (eds), Proceedings of Detecting Invasive Exotic Plants, Workshop and Conference, Florida International University, Miami, FL, Feb. 2003.

Parendes LA and Jones JA (2000) Role of light availability and dispersal in exotic plant invasion along roads and streams in the H.J. Andrews Experimental Forest, Oregon. *Conservation Biology* 14: 64-75.

Pauchard A, Alaback PB and Edlund EG (2003) Plant invasions in protected areas at multiple scales: *Linaria vulgaris* (Scophulariaceae) in the West Yellowstone area. *Western North American Naturalist* 63: 416-428.

Pauchard A and Alaback PB (2004) Influence of elevation, land use, and landscape context on patterns of alien plant invasions along roadsides in protected areas of south-central Chile. *Conservation Biology* 18: 238-251.

U.S. National Park Service (1996) Preserving our natural heritage – a strategic plan for managing invasive non-indigenous plants on national park system lands. Retrieved from http://www.nature.nps.gov/biology/invasivespecies/strat_pl.htm

Rejmánek M, Richardson DM, Barbour MG, Crawley MJ, Hrusa GF, Moyle PB, Randall JM, Simberloff D and Williamson M (2002) Biological invasions: Politics and the discontinuity of ecological terminology. *Bulletin of the Ecological Society of America* 83: 131-133.

Rew LJ, Maxwell DB and Aspinall RJ (*in press*) Predicting the occurrence of non-indigenous species using environmental and remotely sensed data. *Weed Science.*

Riitters KH and Wickham JD (2003) How far to the nearest road? *Frontier in Ecology and the Environment* 1:125-129.

Spellerberg IF (1998) Ecological effects of roads and traffic: a literature review. *Global Ecology and Biogeography Letters* 7: 317-333.

Stohlgren TJ, Quinn JF, Ruggiero M and Waggoner GS (1995) Status of biotic inventories in US National Parks. *Biological Conservation* 71: 97-106.

Stolgren TJ, Barnett TD and Simonson SS (2003) Beyond North American weed management standards. Retrieved from http://science.nature.nps.gov/im/monitor/meetings/FtCollins_02/StohlgrenBeyondNAWMA.pdf

Thompson SK (2002) Sampling. John Wiley and Sons, Inc., New York, USA, 343 pp.

Trombulak SC & Frissell CA (2000) Review of ecological effects of roads on terrestrial and aquatic communities. *Conservation Biology* 14: 18-30.

Tyser RW & Key CH (1988) Spotted knapweed in natural area fescue grasslands: an ecological assessment. *Northwest Science* 62: 151-160.

Tyser RW & Worley CA (1992) Alien flora in grasslands adjacent to road and trail corridors in Glacier National Park, Montana (U.S.A.). *Conservation Biology* 6: 253-262

Vitousek PM, D'Antonio CM, Loope LL, Westbrooks R (1996) Biological invasions as global environmental change. *American Scientist* 84: 468-478

Watkins RZ, Chen J, Pickens J and Brosofske KD (2003) Effects of forest roads on understory plants in a managed hardwood landscape. *Conservation Biology* 17: 411-419.

Weiher E and Keddy PA (1999) Assembly rules as trait-based constraints on community composition. In: Weiher E and Keddy P (eds.) Ecological Assembly Rules, pp 251-271. Cambridge University Press, USA.

Young JA, Evan RA and Major J (1972) Alien plants in the Great Basin. *Journal of Range Management* 25: 194-201.

Table 1. Number of patches and patch radii used in the model for Species A-D used in the model, and the patch area, percentage area infested and circumference this achieved over the 10 000 ha simulation area.

Species	Total number of patches	Patch radius (m)	Area of one patch (m^2)	Area of all patches (ha)	Area infested (%)	Circumference of all patches (km)
A	15055	5.2	85.0	127.9	1.3	491.9
B	5033	10.2	326.9	164.5	1.6	322.6
C	1739	15.3	735.4	127.9	1.3	167.2
D	1220	41.1	5306.8	647.4	6.5	315.1

Table 2. Mean and standard deviation of number of patches intersected using the seven different survey methods, the percentage frequency given the number of observations (n=2546),and percentage of the area infested given the number and proportion of patches intersected, for Species A-D. Results are provided for both random and weighted distributions using both the standard and adaptive sampling designs. See Figure 2 for explanation of survey methods and Figure 3 for explanation of adaptive design.

	Standard sampling				**Adaptive sampling**			
	Mean number of patches intersected	Standard deviation of mean	% frequency of species	% of area infested	Mean number of patches intersected	Standard deviation	% frequency of species	% of area infested
Species A (Population = 15055, Area infested = 1.3 %)								
RANDOM								
Grid	155	13.1	6.1	1.0	659	60.2	25.9	4.5
Point	159	10.7	6.3	1.1	641	60.4	25.2	4.3
Walk	52	7.4	2.0	1.0	132	25.0	5.2	2.7
Perp.RoW	64	9.3	2.5	1.3	191	36.7	7.5	3.8
Targeted	62	6.4	2.5	1.3	200	32.5	7.9	4.0
SaD	63	6.3	2.5	1.3	197	27.6	7.7	4.0
Contour	64	7.5	2.5	1.3	189	22.3	7.4	3.8
WEIGHTED								
Grid	157	15.5	6.2	1.1	5394	215.7	211.9	36.4
Point	160	9.1	6.3	1.1	5291	248.7	207.8	35.8
Walk	48	19.9	1.9	1.0	1313	1000.4	51.6	26.4
Perp.RoW	85	9.3	3.4	1.7	3291	451.8	129.3	66.2
Targeted	52	9.8	2.0	1.0	1143	407.5	44.9	23.0
SaD	487	96.8	19.1	9.8	4731	203.4	185.8	95.2
Contour	40	7.1	1.6	0.8	444	326.0	17.4	8.9
Species B (Population = 5033, Area infested = 1.6%)								
RANDOM								
Grid	116	12.0	4.6	1.8	183	21.1	7.2	2.9
Point	118	11.1	4.6	1.9	184	18.6	7.2	2.9
Walk	25	6.0	1.0	2.0	31	4.8	1.2	2.5
Perp.RoW	32	5.6	1.3	2.6	44	9.7	1.7	3.5
Targeted	31	5.3	1.2	2.5	44	11.5	1.7	3.5
SaD	31	5.5	1.2	2.5	45	9.0	1.8	3.6
Contour	32	5.9	1.2	2.5	43	8.7	1.7	3.5
WEIGHTED								
Grid	116	9.7	4.6	1.8	390	65.4	15.3	6.1
Point	114	9.8	4.5	1.8	391	59.7	15.3	6.1
Walk	24	8.4	1.0	1.9	47	19.6	1.9	3.8
Perp.RoW	45	8.5	1.8	3.6	120	37.3	4.7	9.6
Targeted	27	5.7	1.1	2.2	59	20.5	2.3	4.7
SaD	236	17.4	9.3	18.9	621	44.3	24.4	49.9
Contour	20	3.3	0.8	1.6	26	5.1	1.0	2.1

	Standard sampling				**Adaptive sampling**			
	Mean number of patches intersected	Standard deviation of mean	% frequency of species	% of area infested	Mean number of patches intersected	Standard deviation	% frequency of species	% of area infested
Species C (Population = 1739, Area infested = 1.3%)								
RANDOM								
Grid	74	8.0	2.9	1.3	84	9.6	3.3	1.4
Point	74	8.5	2.9	1.3	81	9.5	3.2	1.4
Walk	12	3.6	0.5	1.5	11	3.8	0.4	1.4
Perp.RoW	14	3.4	0.6	1.8	16	5.0	0.6	2.0
Targeted	13	4.0	0.5	1.7	15	4.3	0.6	1.9
SaD	15	3.9	0.6	1.9	15	4.4	0.6	2.0
Contour	14	3.4	0.5	1.8	15	4.1	0.6	2.0
WEIGHTED								
Grid	71	8.0	2.8	1.2	112	14.1	4.4	1.9
Point	71	8.2	2.8	1.2	115	15.4	4.5	2.0
Walk	14	5.6	0.5	1.8	11	6.8	0.4	1.4
Perp.RoW	27	6.8	1.1	3.4	33	10.3	1.3	4.2
Targeted	15	3.6	0.6	1.9	16	4.5	0.6	2.1
SaD	155	11.3	6.1	19.8	193	17.7	7.6	24.7
Contour	9	3.8	0.4	1.1	8	3.9	0.3	1.0
Species D (Population = 1220, area infested = 6.5%)								
RANDOM								
Grid	248	16.1	9.7	7.7	275	18.4	10.8	8.6
Point	254	13.7	10.0	7.9	277	16.1	10.9	8.6
Walk	14	3.3	0.5	4.8	15	4.6	0.6	5.3
Perp.RoW	23	5.1	0.9	8.1	23	4.9	0.9	8.2
Targeted	22	4.4	0.9	7.8	23	5.3	0.9	8.0
SaD	23	4.9	0.9	8.0	24	5.2	1.0	8.6
Contour	23	4.4	0.9	8.1	24	4.8	0.9	8.5
WEIGHTED								
Grid	265	13.9	10.4	8.2	281	15.3	11.0	8.8
Point	265	11.8	10.4	8.3	274	16.6	10.8	8.5
Walk	15	3.9	0.6	5.3	15	4.5	0.6	5.5
Perp.RoW	24	5.5	0.9	8.4	25	5.6	1.0	8.7
Targeted	23	4.9	0.9	8.0	24	4.4	0.9	8.4
SaD	23	4.9	0.9	8.0	24	5.1	0.9	8.3
Contour	22	5.2	0.9	7.8	23	5.3	0.9	8.1

Table 3. Distance and time travelled between survey areas and total time necessary to complete the seven different survey methods, using the point or transect locations depicted in Figure 2. A speed of 3 km/hr was assumed when travelling between sample areas and 1 km/hr when sampling.

Survey method	Distance travelled (km) to reach survey areas	Time (hr) to travel between survey areas	Total time (hr) to complete survey
Grid	480.0	160.0	180.0
Point	475.0 (± 8)	~158.3	~178.3
Walk	48.0	16.0	36.0
Perp.RoW	33.8	11.3	31.3
Targeted	34.5	11.5	31.5
SaD	37.0	12.3	32.3
Contour	45.5	15.2	35.2

Table 4. Proportion of paired t-tests which showed no significance difference between samples from the targeted method and those from each of the other survey methods using both random and weighted distributions of *Species A-D* and a) standard and b) adaptive sampling designs. Comparisons based on mean distance of patches to rights of way. 95% significance value was used.

	Species A		Species B		Species C		Species D	
	Random	Weighted	Random	Weighted	Random	Weighted	Random	Weighted
a) Standard								
Grid	0.90	0.85	0.98	0.94	0.96	0.94	1.00	0.96
Point	0.94	0.83	0.96	0.96	0.96	0.92	0.98	0.94
Walk	0.81	0.71	0.81	0.81	0.83	0.83	0.83	0.75
Perp.RoW	0.33	0.40	0.17	0.17	0.54	0.71	0.54	0.21
SaD	0	0	0	0	0	0.04	0	0
Contour	0.77	0.69	0.88	0.88	0.92	0.81	0.90	0.88
b) Adaptive								
Grid	0.98	0.60	0.85	0.56	0.92	0.75	0.94	0.96
Point	0.96	0.69	0.94	0.58	0.96	0.88	0.98	0.94
Walk	0.87	0.76	0.56	0.67	0.88	0.83	0.77	0.75
Perp.RoW	0.13	0.36	0.06	0.31	0.44	0.73	0.23	0.29
SaD	0	0	0	0	0	0.04	0	0
Contour	0.87	0.56	0.81	0.40	0.92	0.71	0.90	0.90

Table 5. Number of Species A-D patches intersected when holding the area covered by each non-indigenous species (NIS) constant and sampling with a grid or targeted survey method. Percentage frequency, actual (from the model) and estimated (using geometric calculations) of area infested.

Area covered by NIS held constant (1278887 m2; 1.3%)								
	Mean number of patches intersected		% frequency (n=2546)		% area infested (actual)		% area infested (estimated)	
Survey method	Grid	Targeted	Grid	Targeted	Grid	Targeted	Grid	Targeted
Species A	120	61	4.7	2.4	1.2	1.3	1.2	1.2
Species B	71	24	2.8	0.9	1.3	1.3	1.3	1.9
Species C	55	14	2.2	0.6	1.2	1.3	1.2	1.8
Species D	40	5	1.6	0.2	1.2	1.4	1.2	1.7

Figure 1. Example of the random and weighted distributions of *Species A-D* used in the geographical information system model. Weighted distributions reflect species correlation with proximity to rights of way (shown by double solid lines). Thumbnails to the right hand side of the graphic depict the relative size of the different species' patches/populations and the different occurrence levels.

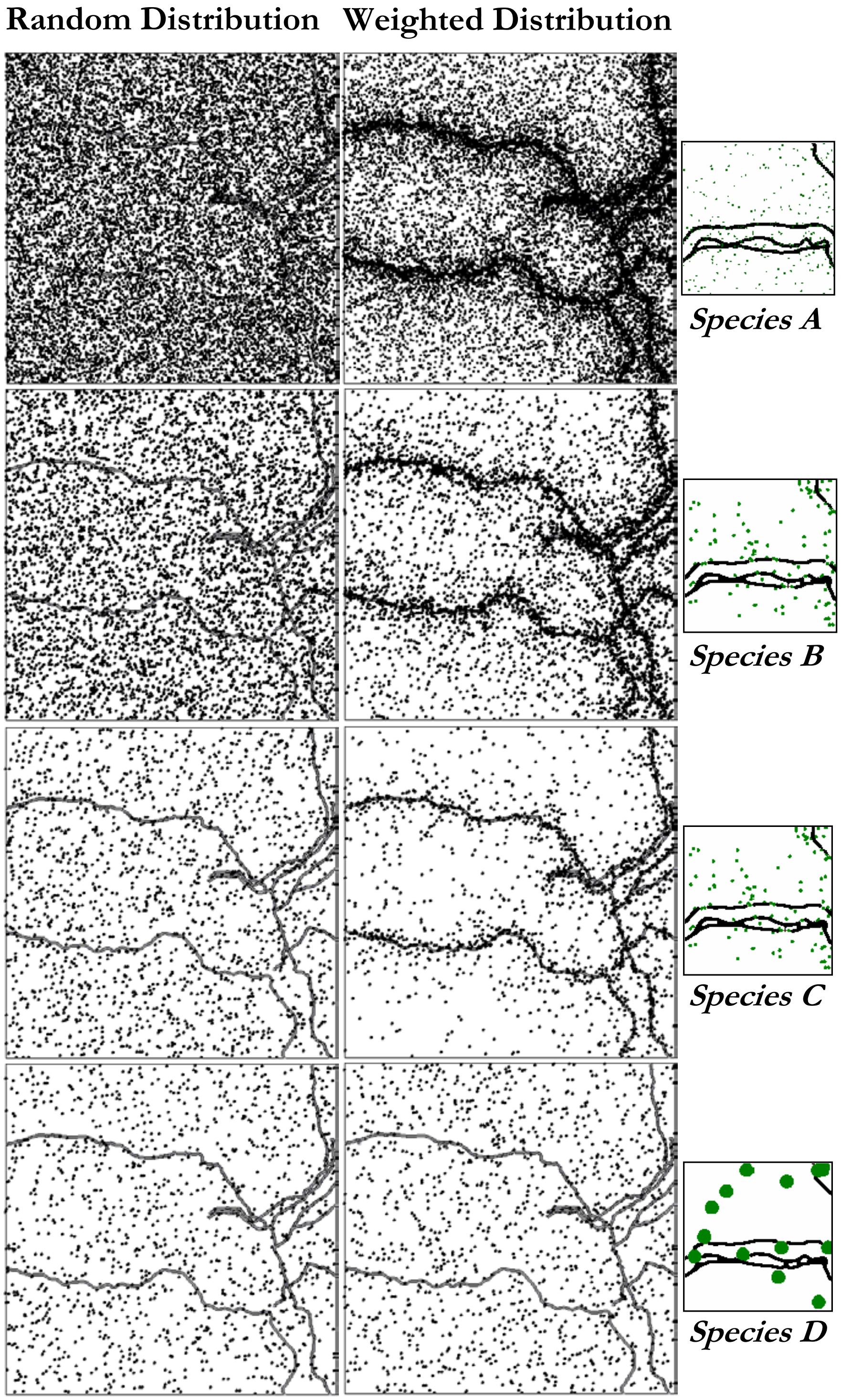
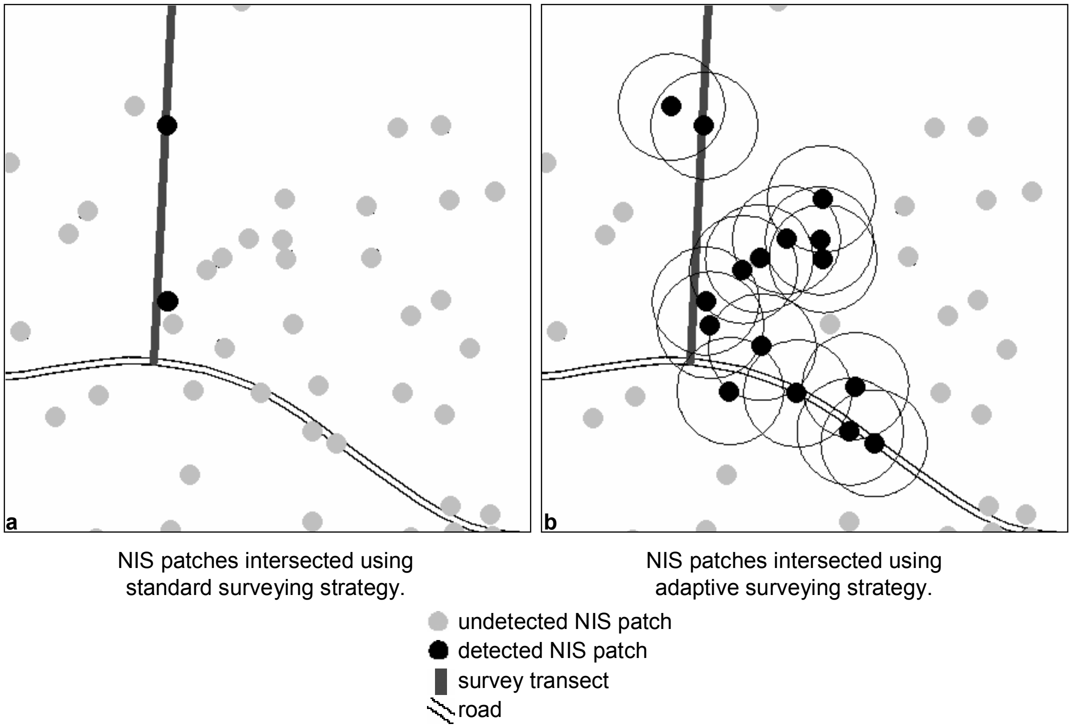


Figure 3. Example of the standard and adaptive sampling design employed in the simulation model for transect survey methods. The same approach was used for point survey methods (not shown). Non-indigenous species (NIS) are drawn to a significantly larger scale for visual clarity. The open circles represent 50 m radius adaptive search area.



NIS patches intersected using standard surveying strategy.

NIS patches intersected using adaptive surveying strategy.

Figure 4. Box and whisker plots showing variation around the median distance from roads and trails for the populations and samples of the population using seven different survey methods. Fifty populations were simulated for both random and weighted distributions of Species A-D. Solid black line within grey box represents the median; dashed white line the mean; grey box represents 50% of the data; whiskers represent the 10th and 90th percentiles; and dots represent outliers. For explanation of different survey methods see Figure 2. Abbreviations include: Pop'n = population p.ROW = Perp.ROW, Tar'd = targeted transects.

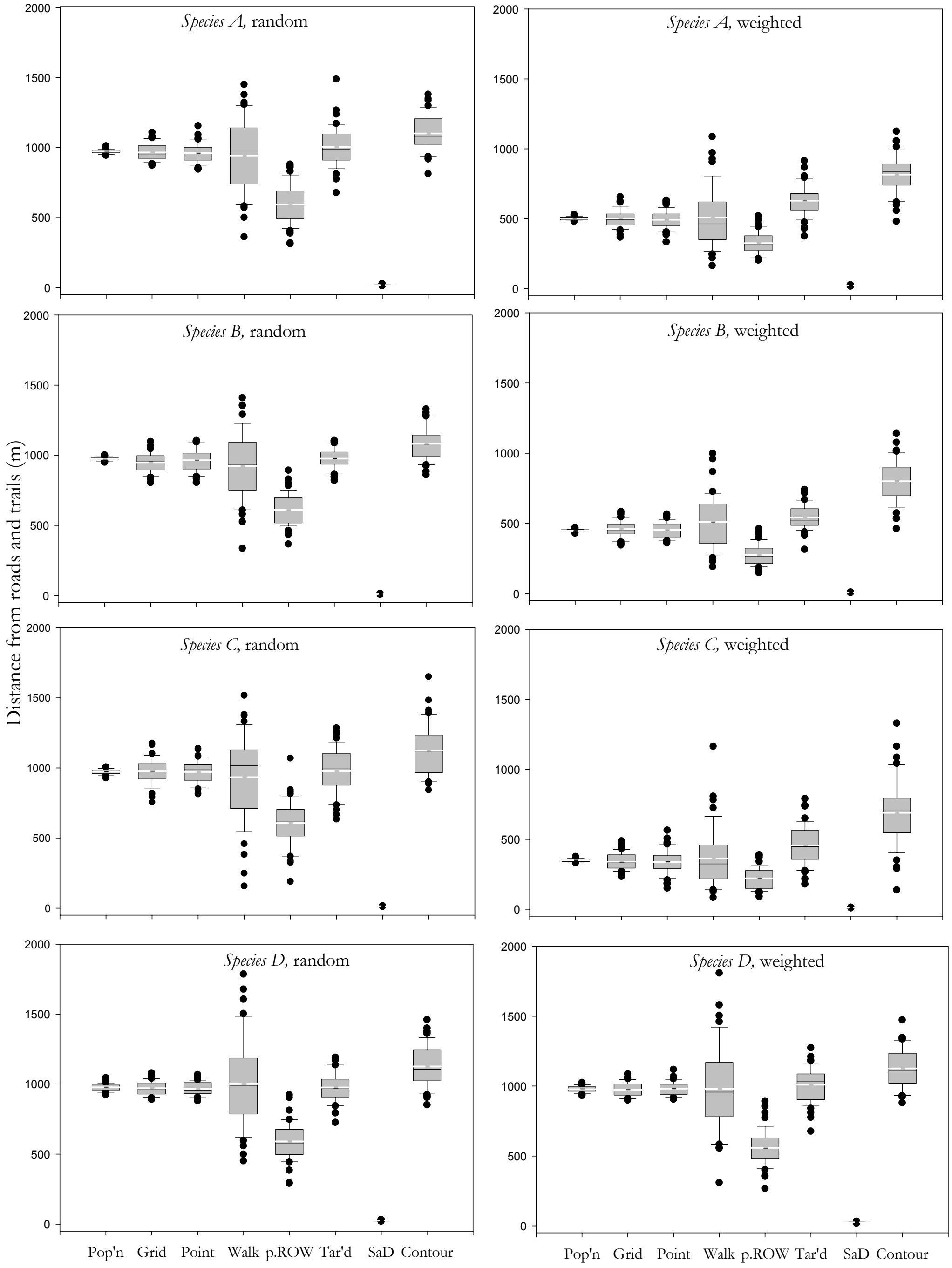


Figure 2. Hierarchical classification tree and diagrams of the seven survey methods evaluated for detection of non-indigenous species' patches using a geographical information system model. Titles given at the fifth level are the abbreviations used for the survey methods throughout the text. A 10 km x 10 km (10 000 ha) area was used in the model and rights of way (RoW) are represented by the light grey double solid lines. Dots represent the point sample locations; black lines the transects.