THESIS

INCORPORATING THE DISTURBANCE PROCESS OF

FIRE INTO INVASIVE SPECIES HABITAT SUITABILITY

MODELS

Submitted by

Joel Silverman

Department of Forest, Rangeland and Watershed Stewardship

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2008

COLORADO STATE UNIVERSITY

May 14, 2008

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY

JOEL SILVERMAN ENTITLED "INCORPORATING THE DISTURBANCE PROCESS OF FIRE

INTO INVASIVE SPECIES HABITAT SUITABILITY MODELS" BE ACCEPTED AS FULFILLING

IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate work

_____Signature
Dr. Monique Rocca
Adviser

_____Signature
Dr. Melinda Laituri

_____Signature
Dr. Patrick Martin

_____Signature
Department Head/Director

THESIS ABSTRACT

This study is motivated by the difficulties land managers face while attempting to simultaneously maintain the natural role of fire in ecosystems and prevent the spread and proliferation of invasive plants. I developed habitat suitability models to predict the responses of three invasive species to fire and other environmental variables: one species in each of three National Parks. For each species, model comparisons tested whether the inclusion of nationally-available data on burn severity, time since fire, and fire occurrence could improve habitat suitability models relative to non-burn data alone. Each species demonstrated significant responses to fire, although incorporation of fire information into the models improved model performance for some species more than for others. In Yellowstone NP, *Linnaria dalmatica* was more likely to occur in areas of low burn severity, while in Sequoia & Kings Canyon NP, *Cirsium vulgare* was more likely to occur in areas of high severity and areas of rapid vegetation recovery. In Chapter 2, I further examined populations of *Carduus nutans* within Grand Teton NP using detailed vegetation map data and specifying interactions between vegetation class and the burn variables. This technique demonstrated some vegetation types were more susceptible to post-fire invasion than others. Between post-fire years one and twelve, the odds of *C. nutans* presence decrease within lodgepole and hydrophilic vegetation types, but increase within deciduous forests, sagebrush and other vegetation types. Predictive models which incorporate dynamic ecological processes such as fire offer great potential to guide post-fire invasive plant management. Such models can help determine whether an invasive is significantly influenced by fire, and if so, which burn variables are most relevant. When the species is fire-influenced these models can more efficiently and precisely prioritize

areas in need of treatment than either habitat maps that lack burn information or simply

prioritizing using burn severity maps.

Key words:

invasive species, burn severity, modeling, fire, invasibility, habitat suitability

Joel Forest Silverman
FRWS Department
Colorado State University
Fort Collins, CO 80523
Fall 2008

ACKNOWLEDGEMENTS

Chapter 2: A Case-Study of *Carduus nutans*: Evaluating Post-Fire Vulnerabilities Amongst Several Vegetation Types in Grand Teton NP 51

Preface

Although fire is a key processes regulating patterns of succession, heterogeneity, and diversity in many ecosystems (Agee 1993; DeBano 1998; Bond&Keeley 2005), it can also promote invasion by non-native plant species (Davis et al. 2000; Brooks et al. 2004). When non-native plant species, defined here as species absent from North America prior to European contact, are deemed to cause negative impacts on ecosystems, they are termed invasive species. Invasive species often compete with native species, contribute to the loss of wildlife habitat, and disrupt the processes that underlie ecosystem integrity. Predictive maps of suitable habitat can serve as a tool to assist in the management of invasive species. In particular, these maps are useful in identifying areas of suitable habitat during early stages of invasion, so populations can be targeted for management before they get out of control. This study uses a model-based approach to predict invasive species habitat within fire-prone ecosystems. It expands upon previous efforts by specifically incorporating variables representing fire disturbance as predictors of invasive species habitat (Underwood et al. 2004; Floyd et al. 2006).

The thesis is structured into two main chapters (Chapters 1 and 2); each one is a complete article written for peer-reviewed publication. Chapter 1 contains three objectives:

1) Create accurate invasive species habitat suitability maps using national environmental data sources as predictors and invasive species presence / absence field observations as the response variable.

2) Develop a methodology to incorporate burn variables into statistical models that can be used to generate habitat suitability maps for current post-fire conditions.

3) Test whether burn variables improve the model predictions for each of three invasive species, and describe the relationship of each species to the burn variables (if any).

For each species, field data on species observations was partitioned into a training set (80%) and a validation set (20%). Using the training set, a full model with burn variables was compared to a reduced model without burn variables to determine if burn variables improved the explanatory power, model fit, and/or model accuracy of the predictions. The same full vs. reduced model was then compared using the validation set. The models that included burn information demonstrated a range of improvement over models without the burn information, but the nature, strength, and shape of the relationship between invasive species and fire depended on the species. All species showed some relationship to fire occurrence, burn severity, or fire age. Habitat maps generated from the best models showed a great deal of heterogeneity in habitat suitability across each study area, but burn variables did not visually dominate the arrangement of suitable vs. non-suitable habitat across the landscape. This suggests that a combination of burn and non-burn environmental factors together form the pattern seen in habitat maps for the three species modeled.

Chapter 1 demonstrates that existing nationally-available burn data can be used effectively in modeling habitat for some invasive species. First, such burn data can help to determine the importance of fire as a factor promoting or exacerbating invasion of individual species. Secondly, for species strongly associated with fire occurrence, burn

severity, or time-since fire, models that include burn information will generate better predictions of suitable habitat compared to models without burn data. Thus, burn information can be a valuable tool in analyzing invasive species distributions within fire-prone ecosystems. The preferred publication for this chapter is the International Journal of Wildland Fire.

Chapter 2 develops a more complex model analysis for *Carduus nutans* (musk thistle) in Grand Teton National Park. The objectives of Chapter 2 are to:

1) Test whether model performance improves when high-accuracy, fine-scale vegetation data and regional climate data are used in addition to the national data sources used previously.

2) Test whether the odds of *C. nutans* occurrence differs between vegetation classes.

3) Test whether there is an interaction between the burn factor(s) and the vegetation classes.

Chapter 2 uses similar methods to those described in the preceding chapter, however, the focus is on testing *C. nutans* responses across different types of vegetation. Since each vegetation community possesses a different set of functional traits and competitive abilities, some are generally more vulnerable to particular invaders (such as *C. nutans*) than others. In addition, the post-fire vulnerability of a vegetation community differs according to the prevalence of native species' traits such as resprouting ability and seed banking. A better understanding of the distribution of *C. nutans* across vegetation types and the interactions between vegetation and burn variables will contribute to our ecological understanding of *C. nutans* as an invader. Climate variables are also added to

the model to improve overall model explanatory power and to determine whether remotely sensed phenology indices can serve as a surrogate for climatic variables assessed at a coarser spatial resolution.

Together, climate and vegetation variables dramatically improved the percent deviance explained (from 24% to 65%) in the *C. nutans* habitat suitability models. Burn variables were helpful in identifying particular conditions under which the odds of *C. nutans* presence would be raised, however they had a marginal effect on overall model performance.

The results show that *C. nutans* occurrence depends not only on burn factors and vegetation classes independently, but that the response to one burn factor (time-since-fire) depends on the vegetation class. The odds of presence for *C. nutans* decreases with burn age within hydrophilic and lodgepole pine communities, while it increases with burn age within deciduous, mixed conifer and sagebrush communities. Since a combination of fire-related and non-fire-related environmental factors underlie *C. nutans* occurrence, we conclude that models which incorporate both suites of variables are better suited to guide post-fire weed prevention than simply a) treating areas of high burn severity, or b) using models that omit burn factors. The results will be especially useful when managers wish to understand and predict invasions in particular vegetation types. For example, the results may assist managers burning to regenerate aspen stands or those who are in the process of restoring sage grouse habitat after wildfire. Chapter 2 is intended for

submission in the journal Diversity and Distributions because of its emphasis on invasion

ecology and biogeography.

Chapter 1: Evaluating Species Habitat Models That Incorporate the Disturbance Process of Fire within Three Ecosystems

Abstract

A growing body of literature links fire to the rapid colonization and proliferation of particular invasive plants. The likelihood of post-fire invasion and the contributing factors involved remain unexamined for many invasive species at the regional level. In order to assist in the prevention and control of post-fire invasions, we built model-based habitat suitability (HS) maps that incorporate the disturbance process of fire in addition to a standard suite of environmental variables (e.g. topographic data, vegetation indices, and hydrologic data). We tested our model on three invasive species that are currently proliferating within their respective study areas: *Carduus nutans* (musk thistle) in Grand Teton NP, and *Linnaria dalmatica* (Dalmatian toad-flax) in Yellowstone NP, and *Cirsium vulgare* (bull thistle) in Sequoia & Kings Canyon NP. We compare the models with burn information (fire occurrence, time since fire, and burn severity) to reduced models without burn information, demonstrating the additional explanatory power added by burn predictors. Our results indicate all three species are influenced by burn characteristics, although the relative importance of burn variables for improving model prediction varied by species. HS models which incorporate both burn and non-burn variables offer greater potential predictive ability than either non-burn HS maps or burn severity maps (such as those sometimes employed by Burned Area Emergency Response teams). The final HS models are useful for demarcating areas of concern prior to prescribed fire, and for directing post-fire invasive species management. By using input

data that are free and readily-available for the continental US, along with open source

modeling software, this paper demonstrates that HS models can serve as a low-cost,

accessible tool for invasive species research and fire management decision-making.

Introduction

Habitat suitability (HS) models, also known as species distribution models, are statistical

or theoretical models that link species observations at specific locations to environmental

data in order to predict the geographic distribution of one or more species

(Guisan&Zimmermann 2000). HS models were developed based on the assumption that,

at regional to continental scales, species distributions are controlled by environmental

conditions (Lugo et al. 1999; Guisan&Zimmermann 2000). Environmental conditions

are usually approximated from geographic data layers in categories such as topography,

climate, hydrology, soil characteristics, anthropogenic features, and vegetation

information. HS models are useful for predicting the distributions of rare plants, species

distributions under future climates, and mapping vegetation communities (Franklin 1995;

Guisan&Thuiller 2005). Although HS models are most often used to predict distributions

of native species, they are increasingly employed to describe non-native species invading

new territory (Higgins et al. 1999; Guisan&Zimmermann 2000; McNab&Loftis 2002;

Rouget et al. 2004; Underwood et al. 2004; Jones&Halpern 2005; Rew et al. 2005;

Thuiller et al. 2005; Araujo&Guisan 2006; Floyd et al. 2006; Morisette et al. 2006;

Schussman et al. 2006; Barnett et al. 2007).

Here, "non-native species" are defined as plant species not found in the US prior to European contact. "Invasive species" are a subset of non-native species that demonstrate an ability to create a significant direct or indirect negative impact on native plants or animals. Often the impact of invasive species is assessed based on a simple decrease in native plant abundance, richness, or cover, even though the mechanisms underlying this change may be complex or poorly understood.

Predictive maps of suitable habitat are important in assisting invasive plant managers in planning early-detection or control efforts because management areas are often too large and/or remote for complete survey beyond roads and trails. In many cases, species may still be spreading to suitable habitat (as propagules are dispersed or conditions change at a site) and little quantitative information exists regarding the habitat preferences of the invasive species in the recipient community of interest. In this paper, habitat suitability maps were constructed for three species from different ecosystems. In particular, I describe the extent to which the distributions of these species reflect past fire disturbances. Free data sources available across the contiguous US in addition to open source statistical software were utilized to demonstrate the potential for low-cost modeling in national-level management support systems. This work will aid fire and invasive-plant managers in assessing potential risks and making appropriate management responses.

To date, most habitat modeling efforts have limited environmental predictor variables to those that are temporally static - at least at annual to decadal scales - effectively ignoring

the dynamics of ecosystem processes such as fire (Barnett et al. 2007; Crossman&Bass 2008; Nielsen et al. 2008), but see (Gumpertz et al. 2000; Pausas&Verdu 2005; Rew et al. 2005; Lynch et al. 2006; Chabrerie et al. 2008). Because fire and other disturbance processes are considered key factors in invasion biology and HS modeling research (Davis et al. 2000; Guisan&Zimmermann 2000; Brooks et al. 2004; Guisan&Thuiller 2005; Keeley 2006), the inclusion of ecosystem *processes* in addition to ecosystem *structures* in management decisions has become critical to successful prevention and control of invasive species. Due to the complexity of such processes, this paper concentrates solely on fire disturbance. Incorporating ecological processes into HS models has the potential to improve habitat predictions and provide inferential knowledge about the species of interest. The difficulty in obtaining accurate fire disturbance information, and the added complexity of incorporating a time-dependent environmental variable, have previously limited the ability habitat modelers to incorporate fire information into HS models. Omitting *processes* has thus far limited the scope of HS models in fire-prone ecosystems, however current data availability allows for wider applications.

HS models are useful for testing ecological theory as well as for assisting in invasive plant management (Higgins et al. 1999; Zimmermann&Kienast 1999; Rouget et al. 2004; Thuiller et al. 2005; Kneitel&Perrault 2006; Pausas et al. 2006a; Pausas et al. 2006b; Jones&Halpern 2007). The goal of this study is to employ modeling methods to predict the potential habitat for three invasive species, particularly in relation to fire. Development of the HS models was accomplished by linking x/y coordinates of field

observations (presence or absence of each species) to environmental predictor variables using logistic regression. Coefficients from the regression were then used to calculate model predictions across the study area to create an HS map for each species. The modeling process described herein combines several techniques, which together expand upon the work of previous authors by: 1) incorporating both temporally dynamic burn variables and standard environmental variables as predictors; 2) checking the model parameter estimates for the effects of spatial-autocorrelation; and 3) taking advantage of the recent increase in both remote sensing and geographic information system (GIS) data, including a Landsat-derived measure of burn severity and a series of MODIS-derived seasonal vegetation metrics known as TIMESAT.

The Roles of Fire in Ecological Restoration & Invasion Biology

It is now widely acknowledged that fire plays a critical role in maintaining diversity in ecosystems throughout the world and well as influencing canopy structure, spatial heterogeneity, fuel loading, and vegetative regeneration (Baker 1992; Agee; Swetnam 1993; Turner et al. 1999; Dale et al. 2001; Rollins et al. 2001; Schwilk&Ackerly 2001; Pausas et al. 2004; Bond&Keeley 2005; Smithwick et al. 2005). This realization has transformed 20th century fire management policy in the U.S. from one of fire-exclusion to one of fire regime restoration and maintenance (Brown et al. 2004; Dellasala et al. 2004; Stephens&Ruth 2005). Fire managers use both management-ignited prescribed fires and naturally-ignited "wildland fire-use fires" as tools for reintroducing fire to ecosystems adapted to a natural fire regime. Unfortunately, some invasive species can quickly colonize and even dominate burned areas, displacing native vegetation and reducing

biodiversity. This presents a problem for fire managers who are charged with minimizing the adverse effects of fire through planning, monitoring, and mitigating fires under their jurisdiction. Currently, post-fire invasive plant management funding is limited and key decisions must be made with limited time and information. HS models can assist fire managers in understanding *which* invasive species thrive under *what* specific post-fire conditions. Developing habitat suitability maps *prior* to fires can help mangers identify key areas that, if burned, may link known propagule sources with highly suitable habitat. Habitat suitability maps updated soon *after* a fire can determine areas within the burn that are most at risk to non-native plant invasion. We suggest that, by incorporating disturbance information, HS models can 1) quickly summarize existing data, 2) objectively map habitat suitability for a species, and 3) direct resources where they will be most effective in post-fire invasive plant management.

Species of Concern

Invasive species rank amongst the major agents of global change such as climate change and human land use change (Mack et al. 2000). To date, invasive species have caused more extinctions than any other factor except human land use change – a factor with which it strongly interacts (Dantonio&Vitousek 1992). Invasive species take a major toll on agriculture, the integrity of ecosystems, and the services which ecosystems provide to humans (Vitousek et al. 1996; Chapin et al. 2000; Pimentel et al. 2000). Some invasive species are known to affect the disturbance processes which normally maintain ecosystems; in fact, this is one of the strongest impacts an invader can force on an ecosystem (Dantonio&Vitousek 1992; Dantonio 1993; Davis et al. 2000; Brooks et al.

11

2004; Guisan&Thuiller 2005). Building on this notion, the theory of resource fluctuation

suggests that the pulse of resources available following disturbance (in conjunction with

the presence of propagules) is a key factor underlying non-native plant invasion (Davis et

al. 2000). These findings provided the impetus for this study and directed the choice of

species to ones with a hypothesized, yet poorly understood, relationship with fire.

We modeled three invasive species: *Cirsium vulgare*, *Carduus nutans*, *Linnaria*

dalmatica).   Each of these species is considered non-native in the US and has been found

to be an aggressive invader in some portion of the country (USDA 2002). Fire effects are

not as well documented for these species compared to other post-fire invaders such as

Bromus tectorum (cheatgrass)*,* but each of the three species is known to colonize recently

burned sites (one in each study area discussed below).

Methods

Study Areas

Our study areas encompassed three lands managed by the National Park Service (NPS):

Sequoia / Kings Canyon (SEKI), Grand Teton (GRTE), and Yellowstone (YELL). These

areas covered a broad range of elevations, fire regimes, and invasive species. SEKI

covers 349,767 ha, with an elevation gradient ranging from (485 m to 4392 m)

(Caprio&Lineback 2002). Fire regimes range from frequent surface fires in the

Ponderosa pine forests (*Pinus ponderosa* Dougl.) and Sequoia groves (*Sequoiadendron*

giganteum (Lindl.) Buchh.) to infrequent high severity fires of the lodgepole (*Pinus*

contorta Dougl.ex Loud.) and red fir forests (*Abies magnifica* Murr.) (Caprio&Lineback 2002).

Covering 135,073 ha, GRTE ranges in elevation from 1935 m to 4197 m, with broad expanses of sagebrush (*Artemisia tridentada* Nutt.) that have moderate fire frequency & high severity. GRTE also contains mixed-conifer forests that have moderate frequency and mixed severity, aspen groves (*Populus tremuloides* Michx.) that have moderate frequency and high severity, and spruce-fir forests (*Picea engelmannii* Parry ex Engelm and *Abies lasiocarpa* (Hook.) Nutt.) that have low-frequency, high- severity fires.

YELL encompasses a total of 898,317 ha with elevations ranging from 1610 m to 3462 m. Lodgepole pine dominates YELL with accompanying infrequent, high severity canopy fires such as the famous fires of 1988. Other vegetation types of YELL include spruce-fir, aspen, and sagebrush. Vegetation which burns only rarely, or in small extents, such as tundra or high elevation woodlands without fuel connectivity are not covered in this study.

Independent Variables - Environmental Data

Logistic regression was employed to relate observations (presence or absence of each species) to environmental predictors. We temporally aligned species observations with the most recent prior fire, if any existed, using geo-referenced fire history records. A spatial covariate term was added to the logistic regression model using general estimate equations to determine if spatial-autocorrelation affected the predictor coefficients and

standard errors (Zeger&Liang 1986). Final habitat maps were created from the model

coefficients, yielding a HS map for each species.

In order to build models which would be easily updatable by land managers, only free

and accessible data sources were considered as candidate predictor variables. These

varied in spatial resolution (listed below) and several of them represent averages over

time. We termed the non-burn variables "static" because they remain relatively constant,

whereas burn variables are termed "disturbance process" variables because they represent

temporally dynamic ecosystem attributes. Data were selected from several broad

categories: topography (30 m), tree cover (30 m), hydrology (30 m), satellite-derived

phenology (250 m), and fire (30 m).

 "Static" Environmental Variables

A set of remotely sensed indices of seasonal greenness (referred to hereafter as RS

phenology data) were used as predictor variables. These data layers, made available

through NASA, are based on infrared and near infrared bands of MODerate Imaging

Spectrometer (MODIS) satellite images. The images are first calibrated to the

Normalized Difference Vegetation Index (NDVI) algorithm and then fit to a smoothed

sine wave using a time-series function developed by Jonsson and Eklundh (2004). The

layers utilize the high temporal frequency of MODIS images (2-3 day cycle) to quantify

seasonal changes in greenness at a moderate resolution (250 m) across the contiguous

US. The series, known as TIMESAT, includes 14 indices of seasonal change, 11 of

which were used for this study (Jonsson&Eklundh 2004).

Candidate topographic variables included a 30m resolution digital elevation model (DEM) and annual potential solar radiation. The annual potential solar radiation combines slope, aspect, and nearby topography to determine the amount of solar radiation possible for each pixel (ESRI 2006). Each static variable was treated as a linear predictor with the exception of elevation, for which elevation-squared was also considered as a candidate. Euclidean distance to open water was calculated using NPS hydrology geo-databases from each study area (NPS 2008) and output was scaled to a 30m cell resolution.

Disturbance Process Variables

The burn variables in this study describe three spatio-temporal aspects of a disturbance process: the presence or absence of fire (fire occurrence), the time since fire (TSF), and the severity of the burn (burn severity). In order to create these variables, two sources of information were utilized: the NPS fire history GIS databases and a remotely sensed index of burn severity (NPS-USGS 2007; NPS 2008). While fire occurrence and TSF are relatively simple concepts based upon fire dates and locations within a GIS, burn severity requires some further explanation.

Burn severity is the combined change in soils and vegetation caused by fire, measured in this study using the differenced Normalized Burn Ratio (dNBR). DNBR is calculated using a pair of remotely sensed spectral images: one pre-fire and one post-fire. The images are captured using NASA's Landsat sensors, which have a moderate temporal and

15

spatial resolution (~16 day interval and 30 m^2 pixels). The dNBR values correlate to observed burn severity as recorded in field plots, when information on each strata is evaluated for first-order biologic and physiologic responses to the burn (Brewer et al. 2005; Key&Benson 2005). DNBR data are freely available for many fires on public lands through the USGS National Burn Severity Mapping Project and the Monitoring Trends in Burn Severity (MTBS) project (NPS-USGS 2007; USGS 2008). The dNBR algorithm uses bands 4 and 7, which reflect portions of the electromagnetic spectrum that correlate well to green vegetation and bare ground, respectively (Key&Benson 2005). First, the Normalized Burn Ratio (NBR) algorithm is calculated for both a pre- and a post-fire image during the growing season using the equation:

NBR = (Band 4 - Band 7) / (Band 4 + Band 7)

Next, the difference between the pre-fire and post-fire image is calculated by the equation:

dNBR = NBR$_{PRE}$ - NBR$_{POST}$

Values ranging from about -550 to 1350 are considered to represent a true burn response, although extreme values ranging from -2000 to +1500 can occur due to clouds, smoke, and other natural phenomena. Within the valid range, burn severity is classified as follows *in general*:

Enhanced regrowth: -550 to -101

Unburned: -100 to +99

Low: +100 to +269

Moderate: +270 to +439

Moderated-high: +440 to +659

High: +660 to +1300

Further, all the dNBR images used in this study were calibrated using a sampling

technique which controls for bias caused by differences in mean unburned value between

images (Key&Benson 2005). Polygons were digitized in uniform unburned areas outside

burn perimeters and the average value of the pixels in these zones was calculated. This

mean unburned value was then subtracted from each dNBR pixel value within the burn

perimeter.

The burn variables are different than other environmental variables in that they represent

the aftermath of a spatially and temporally discrete set of disturbances. We designed a

procedure to address the spatial and temporal relations among burns, and, specifically, to

relate field observations to the prior fire history at each location. This process is

described in the model array construction section below.

Response Variable

The response variable for this study – the presence or absence of each species – consisted

of an aggregated set of geo-referenced field observations collected from a variety of NPS

monitoring and mapping projects (Appendix C). All plots met several criteria including:

complete plot search for the species of interest, an accurate GPS coordinate (+/- 5 to 10 m) representing the site searched for the species, and search area ranging from 300 to 1000 m^2. Spatial distribution of points in SEKI appear to be biased toward roads, rivers and sequoia groves, leading to local areas of clustered points and other large, vacant zones. Spatial distribution of observations in YELL is concentrated in the northern portion of the park, although transects extended perpendicular to roads for substantial distances. GRTE points were spatially well distributed due to a rich source of vegetation mapping data available for the study area.

Model Array Construction

Aligning the response variable with the associated predictor variables required intersecting the response variable (the species observations) with the environmental predictor layers in a GIS. Each study area was processed separately. To develop the burn variables, a series of grid surfaces were first created by chronologically "stacking" single-year burned area grids, starting with 1988 and continuing until present. For example the "Stack 1999" would include all burns prior to 1999 and, in locations where multiple burns occurred, only the most recent burn value would be retained (Figure 1-A). Once all the stack layers were created (stack 1988 to stack 2006) the layers were intersected with the invasive species observation point file with in a GIS. The point file then contained columns for the year of each species observation as well as columns for each fire year at the given location. Using standard query language logic statements, a new column of years was created which represented the most recent fire *prior to the species observation* at each location. This column of year numbers was subtracted from

the current year to yield the TSF burn variable within each model array. Invasive species

observations with no known fire year were given a TSF value of 100, because most parts

of the park likely burned at some time in the past, but not during the last century.

Observation points associated with sites that only burned prior to 1987 were removed

from the model array, primarily because dNBR data prior to this date were not readily

available. This simplification allowed each observation to include both a TSF value and

a burn severity (BSEV) value which related to the same (most recent) burn. It was also

assumed that recent disturbances (last 20 years) would yield the strongest effects on

invasive species distributions. The other two burn variables, fire occurrence and burn

severity, were also processed using the same temporal rule set as described for TSF.

Analysis

Removing Collinear Predictor Variables

Once the full model array was complete, the predictor variables were analyzed to identify

and remove the most strongly collinear terms. First, we ran a logistic regression

including all candidate predictor variables and computed the variance inflation factor

(VIF) using R Statistical Software (companion to applied regression package)

(Fox&Monette 1992; Fox 1997; R 2007). The variable(s) with the highest VIF are

considered strongly collinear with other variables. The variable with the highest VIF was

removed and the regression was run repeatedly, in an iterative procedure, until no more

collinear variables remained (no VIF > 10). Quadratic terms were collinear with their

linear counterpart, but were added back at the end of this process.

All low VIF terms plus the polynomial terms were then run in a logistic regression again, and the polynomial terms with significant p-values (using the Wald test) were retained. If the polynomial term was significant, the linear term was added back to subsequent models to maintain model hierarchy. For example, dNBR and dNBR^2 were both retained for the *C. vulgare* and *L. dalmatica* model arrays because dNBR^2 was significant.

Modeling Framework

Models arrays for each species were partitioned into a fitting set (80% of the observations) and a validation set (the remaining 20%). We fit a general linear model to the fitting data using logistic regression with a forwards / backwards stepwise model selection process within R statistical software (R 2007). During the stepwise procedure, Akiake's Information Criteria (AIC) was used to select the most parsimonious model (Venables&Ripley 2002; R 2007). The resulting model, which selected from amongst all the non-collinear burn and non-burn predictor variables, is referred to as the *burn model*. A second (reduced) model was fit by logistic regression after removing any burn variables from the burn model. We will refer to the reduced model as the *non-burn model*. An analysis of deviance determined whether the burn variables explained significantly more deviance, after accounting for the additional number of parameters.

Considering Spatial Auto-correlation

To check the model results for potential bias or inflation in the standard errors (and subsequent p-values) due to spatial auto-correlation, a general estimate equation (GEE) was utilized (Zeger&Liang 1986). GEE is a type of marginal model which fits a spatial

covariance term into a generalized linear model, and using an iterative procedure, calculates the mean response to each parameter, as opposed to the joint response to all terms found in conditional models (Miller et al. 2007). As input into the GEE procedure, we fitted a Gaussian semi-variance function to the residuals from our logistic regression. First, the semi-variance of each pair of deviance residuals was plotted against the Euclidean distance between the pair of residuals. The semi-variance (V) equation is:

$$V_{(x,y)} = \left(\frac{1}{2}\right)\sum(z_x - z_y)^2$$

Pairs are plotted in bins, allowing initial estimates of the sill (where variance stabilizes) and range (distance where variance reaches the sill). Using these estimates, the pairs are fit to several distributions using functions in R (Gaussian, spherical, and exponential) (Reich&Davis 2003; R 2007). In each case, the Gaussian distribution fit the binned data best, as measured by corrected AIC. A spatial weight matrix was constructed from the fitted residuals and the GEE quasi-log-likelihood calculated the parameter estimates and "robust" standard errors (Zeger&Liang 1986).

This set of $n * n$ pairs was next fit to a Gaussian distribution and a spatial weights matrix was created where points were weighted inversely to the amount of correlation detected in v. This covariance structure of correlation between observations was then used in fitting the GEE model. GEE and GLM burn model parameter estimates, standard errors, and p-values were compared. If estimates or p-values shifted significantly (e.g. change in $p > 0.05$) between GEE and GLM, the GEE estimates would be used for predictions. If p-values did not shift significantly, GLM estimates would be retained, because GLM models have the advantages of allowing comparisons using AIC, deviance statistics, and

computing conditional response instead of marginal response for parameters (Millar et al. 2007)

Model Assessment

Burn and non-burn models were assessed using several measures of overall performance and classification accuracy. These tests were conducted on both the data used to fit the model as well as the validation data . Models were tested using both threshold and threshold-independent methods (Foody 2002). Threshold-dependent methods require the user to choose a subjective threshold probability at which predictions are classified as either zero (if p < threshold) or one (if p > threshold). The choice of threshold influences model accuracy measures. However, thresholds are sometimes necessary in order to meet research or management objectives (e.g. treating or monitoring a certain number of acres predicted as presence, designating preserves, etc.). For each model we used the threshold that minimized differences between specificity and sensitivity, creating a classifier which considers errors of omission and errors of commission equally important. Hence, in our results, sensitivity and specificity always appear quite similar within each model / dataset combination. Three threshold-dependent measures of model accuracy are reported: sensitivity (percent of true presences predicted), specificity (percent of true absences predicted), and classification accuracy (true presence plus true absence as a percent of total).

One of the common measures of classification accuracy, Cohen's Kappa, is not reported because kappa has been proven to be biased to changes in the ratio of presence to absence

data in the model (prevalence) (Allouche et al. 2006). This presents problems when

working with points which are aggregated from various research (such as those used in

this study) because the prevalence is rarely 1:1. Instead, we report an unbiased measure

adapted from medical science for use in ecology known as the true skill statistic (TSS)

(Allouche et al. 2006). TSS is simply:

Sensitivity + Specificity – 1

Values range from -1 to +1, with scores above 0 indicating a better performance than

chance alone, after accounting for prevalence. These model performance measures

provide a quantitative view which is a necessary compliment to any habitat suitability

map.

The Receiver Operation Curve (ROC) serves as a threshold-independent metric of model

performance. ROC plots true presence versus false presence for each possible threshold

($p = 0.01$ to $p = 0.99$, by 0.01) and calculates the area under the curve (AUC) of the ROC.

Values Range from 0.5 (the model predicts no better than random chance) to 1.0 (perfect

unity) (Fielding&Bell 1997).

The Habitat Suitability Maps

As in the model array construction process, temporal dynamics needed to be considered

for the burn variables. For each burn variable within each study area the most

comprehensive stack layer was utilized (include 2000s and 1990s records; in Yellowstone

1988 is also included). These stacks contained all fire years considered within each study

area, giving precedence to the most recent fires in cases of over-lapping burns (Figure 1-

A). This allows the habitat suitability maps to reflect the current state of fire history for each study area.

Each burn and non-burn logistic regression equation was applied across the entire study area using the ArcGIS Raster Calculator, generating a map of relative habitat suitability. All extents were set to the bounding coordinates of the appropriate study area. In order to match the resolution of other layers in the model, the 250 m resolution TIMESAT grids were re-sampled to 30 m using cubic convolution (ESRI 2006).

Results

Each of the six final models (one full and one reduced model for each species) explained between 15% and 46% of the deviance (Table 1-B). Even when extra model parameters were penalized using AIC or the drop in deviance test, the burn models showed some improvement over the same species modeled without burn data (lower AIC scores and significant differences detected between models using analysis of deviance p-values (Table 1-B).

Collinearity tests showed the three burn variables, used together in any given species model, inflated the variance. However, removing one or two of the burn variables (using the variance inflation factor method described in the methods), the remaining burn variable(s) maintained low variance inflation.   Non-burn variables were collinear amongst themselves and some collinearity existed between non-burn and burn variables.

For each species, the step selection model processes led to models that used combinations of burn and non-burn variables to predict species occurrence. In the two species models where burn-variables ranked highly, *L. dalmatica* and *C. nutans* burn models (see p-values, Table 1-C and 1-E), overall model performance improved in comparison to non-burn versions (see analysis of deviance, Table 1-B).

In the species-specific sections below, the models were assessed and compared in terms of 1) the statistical fit to the observations (AIC and drop in deviance p-values); 2) the classifier accuracy measures at a balanced threshold (classification accuracy, sensitivity, specificity, and TSS); 3) the predictive ability as compared to random chance (TSS); and 4) predictive ability independent of thresholds (area under the curve of the receiver operating curve) (Table 1-B).

C. nutans in Grand Teton NP

As the amount of TSF increased, the probability of presence slightly increased for *C. nutans* (log-odds increases 1.03 for every 10 years TSF, p = 0.021) (Table 1-D).. No other burn variables were retained by the VIF and step selection methods for this species. Presence of *C. nutans* decreased with higher tree cover, greater distance to water, and later spring green-up dates (Table 1-D).

The drop in deviance was significant for the burn vs. non-burn comparison (p = 0.02) but both the AIC and deviance explained showed only slight improvements (Table 1-A). All

measures of classification accuracy and model performance were quite similar between models with and without burn variables for this species (Table 1-A).

***L. dalmatica* in Yellowstone NP**

The *L. dalmatica* burn model also exhibited a decreased likelihood of presence in burned areas, when all other factors were equal (Table 1-C). Yet, as with *C. vulgare,* the odds of presence elevated over a portion of the dNBR range for *L. dalmatica*, reaching a peak at around dNBR = 250 (low severity) (Figure 1-B). At dNBR=250, the odds of presence were 1.4 times the odds at an unburned site (dNBR = 1). The species responded to both linear and squared elevation terms, demonstrating presence is linked to moderate elevations (Table 1-C). Two TIMESAT variables were also highly significant: Green-up and Base Level. Presence decreased with later spring green-up, while presence increased with increased Base Level (growing-season productivity; see Appendix A, section 1-h.)

The deviance explained between burn and non-burn models for *L. dalmatica* were very similar, but the drop in deviance test suggests that the burn model was a significant improvement over a model without burn information (Table 1-A). The threshold-dependent accuracy measures showed a modest increase in sensitivity and specificity within the fitting dataset, but a drop in classification accuracy. Validation data showed the models to be similar across all three measures. However, the TSS score raised by about 16% in the burn model (fitting dataset). The AUC statistic improved by 17% when the fitting set was used to compare burn vs. non-burn, but the AUC remained similar when the same comparison was made using validation data.

C. vulgare in Sequoia & Kings Canyon NP

C. vulgare is less likely to occur in burned areas than unburned areas, averaging over all other factors (odds ratio of 2:100) (Table 1-E). However, when burn severity is considered for *C. vulgare*, the odds of presence increased substantially within certain portions of the full dNBR range (Figure 1-B). The log-odds response is U-shaped, with highest likelihoods at enhanced regrowth (-550 to -100) and high-severity portions of the range. For instance, a high severity burn site (dNBR = 650) is 7 times as likely as an unburned sites (dNBR = 1) to support the presence of the species. *C. vulgare* was also more 9 times more likely to occur in an enhanced regrowth site (dNBR = -175) than in an unburned site (dNBR = 1).

C. vulgare showed a strong response to fire occurrence and burn severity. Together, these burn variables tripled the deviance explained (from 15% to 45%). Similarly, the burn model fit, as measured by AIC and the drop in deviance test, showed strong improvement over the non-burn model (Table 1-A). All threshold-dependent methods of accuracy were raised by about 20% in the burn model, except TSS, which jumped from 0.3086 to 0.7026 (Table 1-A). Overall model performance as measured through the AUC, was also improved in burn model (0.76 to 0.93) (Table 1-A).

Spatial Autocorrelation Tests

For each species modeled, parameter estimates based on the GEE were nearly identical to the logistic regression estimates. All p-values for individual variables remained significant (p< 0.5) in the GEE if they were significant in the GLM. Also, all changes in p-value were less than one order of magnitude (e.g. no changes from 0.001 to 0.01). Because there was so little change and the assessment models are more advanced for GLM (i.e. AIC, etc.), we report only GLM logistic regression results.

Habitat Suitability Maps

Maps of relative probability of occurrence were created for each species in its respective study area. The pattern of probability is heterogeneous near the pixel resolution (30 m), but change more gradually when viewed at a courser scale. HS surfaces rarely change abruptly between burned area and adjacent lands in any fire-related pattern (Figure 1-C; but see Chapter 2 for an exception), nor do dNBR burn severity surfaces show much resemblance to the HS map patterns within burns (Figure 1-D). Within all three study areas, the highest probabilities of presence appear associated with areas of relatively low elevation, close proximity to rivers, and higher levels of solar radiation (Figures 1-E, 1-F, and 1-G)

Discussion

Although "static" environmental (non-burn) variables still accounted for the majority of the deviance explained in the models, burn variables strongly improved model fit, performance and accuracy for *C. vulgare*, and provided inferential ability regarding all three species modeled. Even after removing collinear predictors, one or two burn

28

variables were always retained by the step-wise model selection process, indicating the

fire can partially explain each of the three species' distributions. Burn occurrence

decreased the overall odds of presence in *C. vulgare* and *L. dalmatica* models, but certain

ranges of burn severity were linked to increased odds of presence in the same species. In

the third species modeled, *C. nutans,* likelihood of presence significantly increased with

increasing TSF. Model assessments demonstrated that the three species respond to

specific burn conditions, but that improvement in model prediction *via burn information*

varied between the three species under consideration.

C. vulgare in Sequoia & Kings Canyon NP

The occurrence of *C. vulgare* is related to both burn occurrence and burn severity.

Although the species prefers unburned areas overall, certain burns – those with either

high severity or enhanced post-fire vegetation – increase the odds of presence (Figure 1-

B). It is likely that severely burned sites offer more light and bare soil resources,

especially important to the ruderal strategy of the species. Enhanced post-fire vegetation,

as measured through dNBR, is usually associated with open meadows, which often

experience increased green-up the year following fire (Key&Benson 2005). It is likely

that, like the native meadow flora, *C. vulgare* is also able to quickly capitalize on the

flush of newly available nutrients following fire.

C. vulgare has existed in portions of the park for many years, but has recently been

observed proliferating within recently-burned sequoia groves (Gerlach et al. 2001). Since

the niche that *C. vulgare* inhabits within the park is expanding the habitat suitability map

generated may only stay accurate for a few years. Therefore, we suggest the model

should be re-calculated each year as more presence/absence data become available.

One disadvantage to using dNBR as a variable in SEKI is that the spectral bands utilized

have a poor ability to penetrate the dense vertical canopies of sequoia groves. Therefore,

it is difficult to assess the severity of surface fires in the understory using Landsat

imagery. If managers wished to study *C. vulgare* invasion exclusively within sequoia

groves, fire occurrence, fire frequency, and TSF would be more appropriate candidate

variables to use for modeling. Regardless, inclusion of a vegetation class variable is

recommended for any future modeling within SEKI, as vegetation varies greatly across

the park and may have no surrogate (see Chapter 2). Ranking among variables in the

final burn model as well as comparisons between the burn and non-burn models for this

species (the drop in deviance test and measures of accuracy improved), indicate this

species is substantially influenced by fire (Table 1-B and 1-E). Other non-burn variables

such as elevation, distance-to-water, and seasonal RS phenological amplitude were also

significant in the final burn model (Table 1-E). Overall, we have moderate to high

confidence in the predictions for this species throughout the park, yet predictions in

forests with high canopy cover may be poor. Additional random or random-stratified

sample points across the park should refine predictions further. Future sampling should

assess the range values present across the landscape for each environmental variable and

compare this to the range of values represented with the current sample.

C. nutans in Grand Teton NP

C. nutans demonstrated almost no gain in model performance with the incorporation of burn information (Table 1-B), although one burn variable, TSF was retained as significant explanatory variable. In the burn model for *C. nutans*, odds of occurrence increases by 0.3% with each additional year of TSF (Table 1-D). Since the fire data is mainly limited to fires which occurred from 1 to 12 years prior to weed observation, it is not recommended to infer beyond this range. Almost all non-burn variables ranked more significant in the burn model than TSF. As with many invasive species, *C. nutans* appears to prefer low tree canopy cover, close proximity to water, high solar radiation, areas of high net primary productivity (Base level variable), and areas where early spring growth is possible (Green-up variable) (Table 1-D).

The similarity in fit, accuracy, and performance, between burn and non-burn models seems to indicate fire is of less importance in determining the species distribution compared to the other factors in the model (Table 1-D). However interactions between fire and other variables not tested here may be important. GRTE includes a set of substantially different vegetation types (from sagebrush to spruce). It is possible that *C. nutans* response to fire depends on the vegetation type. We explore this species / study area in detail in Chapter 2.

L. dalmatica in Yellowstone NP

As with the other two species modeled, the odds of *L. dalmatica* occurrence are reduced in burned areas overall, but are elevated in particular portions of the dNBR range.- in this

case moderate-low severity (Figure 1-B). The drop in deviance test was highly significant, showing the burn model was an improvement compared to the non-burn model (Table 1-A). Thresholded accuracy measures suggested a modest improvement between the burn and non-burn model when using the fitting dataset, but the validation set failed to confirm this. This is likely due to extremely low prevalence in the validation set, which is known to affect measures of classification accuracy (Table 1-A) (Fielding&Bell 1997; Foody 2002; Allouche et al. 2006). Low to moderate elevations within the park are most likely to have presence of the species when elevation and elevation squared terms are considered (Table 1-C). In addition, the species was present more often in areas with an early growing season (Green-up variable) and high net primary productivity (Base level variable) (Table 1-C).

L. dalmatica exemplifies a mild response to fire, where model performance improved modestly with the addition of burn variables, and individual burn variables showed mid-level ranking amongst the retained variables in the model (Tables 1-B and 1-C). This species is currently found in the northern range of Yellowstone NP, although isolated pockets of the species were found during the 2007 field season further south crossing into the higher, colder Swan Valley region. *L. dalmatica* is usually associated with open vegetation, which may explain the response to low burn severity (these areas do not tend to burn as severely), although the species may also be damaged directly by higher burn severity levels. In either case, the species is able to take advantage of some burned areas. Although the species has existed in the region since at least the 1960s, it is slowly expanding as mentioned above. Due to this slow spread, predictive maps should reflect

the habitat requirements of the species for three to five years before updating is necessary. Despite this fact, observations remain highly skewed toward the northern region and roadside locations. It is recommended that further sampling be undertaken as soon as possible. We have moderate confidence in these predictions overall, with accuracy probably greatest in the northern region and open vegetation types.

Spatial Auto-correlation

Detecting and mitigating potential sources of error is an important consideration in any bio-geographic study and discussion of such errors are available for many particular fields of interest (Guisan&Zimmermann 2000; Pearce&Ferrier 2000; Foody 2002; Lentile et al. 2006). We were concerned about errors caused by spatial-autocorrelation because geographic data often violate the basic GLM assumption of independence between observations. We used a GEE model to check parameter estimates and p-values computed in the GLM logistic regression, but no estimates demonstrated more than slight changes in these values. The fact the estimates changed so little when the spatial structure of the residuals was considered, shows the original logistic regression was a sufficient technique for the data in this study. The techniques available to account for spatial auto-correlation are limited by the binomial (presence/absence) nature of the response variable used in this study. Other techniques for binomial data are beginning to show promise such as Markov-chain Monte-Carlo (Lynch et al. 2006), but in the meantime it may be simpler to collect continuous response data where more spatial techniques are available (Miller et al. 2007).

Implications for Management

The burn models presented in this paper may assist with fire and invasive plant management in the respective study areas. Further, the methods presented here can be easily applied to other species or ecosystems. For fire management planning, invasive species habitat maps can be used to identify areas at risk to invasion under current or potential burn conditions. In planning prescribed burns, HS maps could be calculated using the anticipated burn conditions and could even be linked to software which spatially predicts fire behavior. Such pre-fire HS maps would help managers assess the degree to which the burn plan objectives may coincide with or differ from weed management goals. In the context of prescribed fire, managers could then manipulate burn frequency and burn severity to reduce the probability of invasion in likely habitat. Using HS maps prior to burns could also help fire managers identify wildland fire trigger points, where invasion by high-impact species such as *Bromus tectorum* (cheatgrass) could be prevented from spreading to newly-burned highly-suitable habitat. In some cases, it may be important to modify the intended fire intensity or frequency to strike a balance between maintaining fire regimes and catalyzing new invasions (Keeley 2006). This is particularly important when considering invasive species which are known to alter fire regimes or other ecosystem processes (Mack&D'Antonio 1998; Brooks et al. 2004).

Using HS maps effectively will also require matching burn variables to the constraints presented by the area of interest. For example, while dNBR is a significant predictor of *C. vulgare* presence across the entire area of SEKI, it is difficult for dNBR to capture burn severity in old-growth sequoia groves. The dense canopies in these groves usually

remain intact following the frequent surface fires, which means little change can be

detected via satellite. Instead, other burn variables including simple burn occurrence,

TSF, and fire frequency may be more useful predictors of habitat in this forest type. We

recommend caution in using the inferences presented here to understand response in these

forests because the measurements presented are averaged across entire regions, which

contain a variety of vegetation types and fire regimes. Chapter two of this thesis adds

new terms and interactions to the model for *C. nutans* in order to identify differences in

the species response to fire within several vegetation types. Similar work should be

undertaken for *C. vulgare* and *L. dalmatica.*

Despite the fact dNBR is a poor index of burn severity in some vegetation types, it

remains a useful variable for use in regional HS modeling, as demonstrated in the cases

of *L. dalmatica* and *C. vulgare*. Inclusion of dNBR in HS models tracks which (if any)

severity levels are most conducive to invasion by the species of interest. DNBR may also

improve model prediction accuracy, as in the above two cases. There are also many

pragmatic reasons why dNBR is an important tool for HS modeling at a national/regional

level. DNBR has proven to be a robust and flexible index of burn severity at moderate

spatial scales (Brewer et al. 2005; Lentile et al. 2006; Zhu et al. 2006) and is freely

available for many burned areas across the US (Key&Benson 2005; NPS-USGS 2007).

In addition, LANDSAT sensors have been capturing images since the 1980s, allowing the

creation of full retrospective "fire atlases" for specific regions such as SEKI (NPS-USGS

2007). Finally, dNBR image processing is being made available for all lands in the

continental US through the Monitoring Trends in Burn Severity Project (MTBS). Burn

severity for all fires greater than 1000 acres in the western US and greater than 500 acres

in the eastern US will become available over the course of the next several years (USGS

2008). For these reasons, we recommend incorporating dNBR into HS models in

addition to other burn (burn occurrence, TSF, and fire frequency) and environmental

variables (topography, phenology, hydrology, geologic) variables.

While burn-related variables show promise for use in mapping species habitat, they are

by no means a replacement for widely-used "static" environmental variables. The HS

maps presented respond most strongly to topographic variables, but it is a combination of

factors that leads to the final landscape pattern (Figures 1-C, 1-D, 1-E, 1-F and 1-G). In

the cases of *C. vulgare* and *L. dalmatica*, post-fire weed management which treated only

high severity areas would omit many areas of highly suitable habitat and select many

areas with low suitability (Figure 1-D).

In order to make the final decisions regarding management actions , the HS maps and

model statistics should be considered in relation to the dispersal mechanism and rate of

spread for the species of interest. The models presented thus far purposely omitted

dispersal mechanism variables such as distance-to-road. This decision was based on the

following premises:

    1) The primary objective was to accurately predict backcountry areas of habitat

    (current weed infestations near roads are relatively well-known and managed).

    2) Including human dispersal routes as an environmental variable confounds the

    species habitat with a vector of dispersal, biasing predictions towards habitat

similar to those near roads and trails. These areas may have more infestations due

to constant immigrating seed, not because of the quality of the habitat.

Instead of including distance to roads and trails in habitat models, we suggest managers

use available information about a species life history and dispersal to select among the

suitable habitat locations those sites where propagule dispersal is most likely. For

example, post-fire monitoring for a slow-spreading species such as *L. dalmatica* could

reasonably be limited to a buffered zone (1-2 km) around the current infestation sites in

the park. In species where seeds are transported moderate distances (5-10 km) via wind,

animal fir, or bird seed caches, such methods will probably be less effective, making the

direct use of HS maps the logical choice.

Conclusion

Fire can be a strong driver of non-native plant invasion and even vegetation type

conversion (Dantonio&Vitousek 1992; Levine et al. 2003; Brooks et al. 2004; Hunter et

al. 2006; Keeley 2006). However, the sheer number of invasive species establishing in

most ecosystems along with the stochastic nature of fire events precludes conducting

mechanistic *in situ* experiments to test the response of each potentially invasive species to

burn characteristics.   For this reason, testing burn variables as predictors of invasion in

HS models makes sense in fire-prone ecosystems. If the species responds to fire, the

response can be quantified and geographically represented. Even if the species modeled

shows little response to fire, that information still increases the body of knowledge

regarding fire effects, with implications for fire and invasive plant management priorities.

The three species tested here showed significant responses to fire effects, although only

in one case was a burn variable top-ranked in importance (Table 1-B). With the current

freely available national burn data, open source statistical software, and the ubiquity of

GIS software, testing disturbance/burn variables in HS models is relatively simple and

inexpensive. We recommend continuing to incorporate the disturbance processes which

shape the ecosystem of study in future HS modeling endeavors.

Tables and Figures (Chapter 1)

Table 1–A Variables Considered in Three Species Model Comparison
The response variable is the presence or absence of the invasive species observed, as collected and aggregated from various NPS field surveys and studies (plot sizes ranged from 315 to 1000 meters^2). NPS Fire History and Hydrology data layers are available online through the NPS data clearinghouse. Burn severity (difference normalized burn ratio) data are available through the online Burn Severity and Mapping Project database. Digital elevation maps and National Land Cover Classification tree canopy cover layers are available through USGS online databases. Solar Radiation is derived from elevation. All remaining layers are remotely sensed phenology indices supplied directly from NASA through the Earth Observation Science program, and are available for land management agencies. An asterisk (*) indicates the layer was calculated in ArcGIS v9.2 using the input data source listed. A double asterisk (**) indicates the layer was first converted from vector type data, then rasterized using Spatial Analyst raster tools in Arc GIS v9.2. Pixel resolution values shown are the length of each side of the pixel.

Type of Variable	Variable	Data Source	Spatial Resolution
Response:			
	Species observation (Presence or Absence)	NPS (See Appendix C)	Varies
Predictors:			
Fire-Related	Fire Occurrence (2 Classes)** Burned / Unburned	NPS Fire History	30 meter
	Time Since Fire**	NPS Fire History	30 meter
	Time Since Fire Squared**	NPS Fire History	30 meter
	Burn Severity (dNBR)**	USGS / NPS	30 meter
	Burn Severity (dNBR) Squared**	USGS / NPS	30 meter
Hydrologic	Euclidean Distance To Water**	NPS Hydrology	30 meter
Topographic	Elevation	USGS	30 meter
	Elevation Squared	USGS	30 meter
	Solar Radiation*	USGS	30 meter
Vegetative	Tree Cover Percent	USGS / NLCC	30 meter
Phenology	Average Base Level	NASA	250 meter
	Large Integral	NASA	250 meter
	Brown Down Rate	NASA	250 meter
	Green Up Rate	NASA	250 meter
	Amplitude	NASA	250 meter
	Max Peak Date	NASA	250 meter
	Peak Date	NASA	250 meter
	Base Level	NASA	250 meter
	Brown Down	NASA	250 meter
	Green Up	NASA	250 meter
	Season Length	NASA	250 meter

Table 1–B Model Performance and Accuracy Measures for Three Invasive Species.

For each species, a model with burn information was compared to a model without burn information. Data was partitioned into a fitting set (80%) and validation set (20%) and performance measures were calculated separately. The area under the curve (AUC) compares model predictions across all thresholds, where values range from -1 to +1 and any value above 0 is considered better than random chance. Prevalence is the ratio of presence to total sample size and is known to influence thresholded measures of accuracy. The "fair cutoff" threshold was chosen to provide a balanced assessment of sensitivity and specificity, since neither was consider more important than the other. The true skill statistic is similar to Cohen's Kappa, but is unaffected by prevalence. It ranges from -1 to +1 and values greater than zero are considered better predictions than by chance alone. Sensitivity is the ratio of true presence to total predicted presence. Specificity is the ratio of true absence to total predicted absence. Akiake's Information Criterion (AIC) allows model fit comparisons and adds a penalty (2k) for each model parameter. Deviance explained is analogous to R^2, and measures the percent of null deviance the model explains. Analysis.of Dev. is an analysis of deviance fit to a Chi-square distribution and tests the null hypothesis the burn model is the same as the no burn model (only tested using the fitting data).

Species & Study Area	Dataset	Model	Sample Size	AUC	Pre-valence	Fair Cutoff	True Skill	Class. Accuracy	Sen-sitivity	Spe-cificity	AIC	Deviance Explained	Analysis of Dev. (p-val)
C. vulgare *SEKI*	Fitting	Burn	338	0.927	0.80	0.770	0.703	0.852	0.852	0.851	203	0.456	<0.0001
		No Burn	338	0.760	0.80	0.840	0.309	0.653	0.652	0.657	297	0.158	
	Validation	Burn	84	0.974	0.84	0.720	0.856	0.928	0.928	0.929	n/a	n/a	
		No Burn	84	0.855	0.84	0.805	0.453	0.735	0.739	0.714	n/a	n/a	
C. nutans *GRTE*	Fitting	Burn	2160	0.819	0.38	0.445	0.514	0.750	0.785	0.729	2219	0.246	0.02
		No Burn	2160	0.817	0.38	0.440	0.497	0.748	0.751	0.747	2222	0.244	
	Validation	Burn	536	0.830	0.39	0.415	0.508	0.754	0.756	0.752	n/a	n/a	
		No Burn	536	0.826	0.39	0.433	0.508	0.764	0.746	0.762	n/a	n/a	
L. dalmatica *YELL*	Fitting	Burn	2031	0.838	0.27	0.595	0.495	0.747	0.747	0.747	2020	0.290	<0.0001
		No Burn	2031	0.669	0.27	0.503	0.338	0.835	0.689	0.649	2044	0.279	
	Validation	Burn	470	0.780	0.04	0.600	0.443	0.721	0.722	0.721	n/a	n/a	
		No Burn	470	0.763	0.04	0.590	0.445	0.723	0.722	0.723	n/a	n/a	

Table 1–C Burn Model Estimates and Odds Ratios for *Linnaria dalmatica* in Yellowstone National Park
Environmental variables (burn and non-burn) are listed in order of decreasing significance. Estimates are given along with the log-transformed odds ratio. Odds ratios are interpreted as a change in odds per unit change in the predictor variable. A cross (†) indicates a burn-related variable. Burn severity is measured through the remotely sensed differenced normalized burn ratio (dNBR). The "Fire Occurrence" (burned / unburned classification) odds ratio is given for burned areas relative to unburned areas. An asterisk (*) indicates a TIMESAT variable.

Variable	Units	Estimate	Odds Ratio	Std. Error	z Value	Pr(>|z|)
Elevation Squared	Meters	-8.56E-06	0.999991	1.67E-06	-5.126	< 0.0001
Green Up*	Days	-1.24E-02	0.987706	2.66E-03	-4.655	< 0.0001
Elevation	Meters	3.10E-02	1.031527	6.75E-03	4.601	< 0.0001
Base Level*	NDVI	2.89E-04	1.000289	6.36E-05	4.543	< 0.0001
(Intercept)	none	-2.81E+01	n/a	6.82E+00	-4.121	< 0.0001
Fire Occurrence †	Burned / Unburned	-5.50E-01	0.577238	1.52E-01	-3.607	0.0003
Burn Severity Squared †	dNBR	-5.81E-06	0.999994	1.67E-06	-3.477	0.0005
Solar Radiation	Watt hours / m^2	1.56E-06	1.000002	4.85E-07	3.215	0.0013
Tree Cover	Percent	-8.28E-03	0.991757	2.94E-03	-2.815	0.0049
Burn Severity †	dNBR	2.86E-03	1.002861	1.02E-03	2.791	0.0053
Max. Peak*	NDVI	-1.27E-04	0.999873	6.66E-05	-1.903	0.0571
Season Length*	Days	1.96E-03	1.001963	1.25E-03	1.575	0.1152

Table 1–D Burn Model Estimates and Odds Ratios for *Carduus nutans* in Grand Teton National Park
Environmental variables (burn and non-burn) are listed in order of decreasing significance. Estimates are given along with the log-transformed odds ratio. Odds ratios are interpreted as a change in odds per unit change in the predictor variable. A cross (†) indicates a burn-related variable such as "time since fire" (TSF). An asterisk (*) indicates the variable is a remotely sensed phenology index computed by NASA – each of which was created from multiple Normalized Differenced Vegetation Index (NDVI) values taken over the course of a year. Elevation has both a linear and first-order polynomial (squared) term. Distance to water represents the Euclidean distance to the closest open water such as a lake or stream.

Variable	Units	Estimate	Odds Ratio	SE	z Value	Pr(>|z|)
Tree Cover	Percent	-1.37E-02	0.986403	2.24E-03	-6.122	< 0.0001
Distance To Water	Meters	-1.74E-03	0.998262	3.15E-04	-5.526	< 0.0001
Green Up*	Days	-2.36E-02	0.976676	4.91E-03	-4.805	< 0.0001
Solar Radiation	Watt Hours / m2	2.60E-06	1.000003	7.80E-07	3.337	0.000847
Elevation	Meters	-2.49E-02	0.975437	8.29E-03	-3.002	0.002681
Peak Date*	Days	1.30E-01	1.138373	4.78E-02	2.714	0.006657
(Intercept)	None	2.40E+01	n/a	9.19E+00	2.606	0.009165
Time Since Fire †	Years	3.10E-03	1.003105	1.35E-03	2.302	0.021338
Elevation Squared	Meters	3.91E-06	1.000004	1.87E-06	2.091	0.036516

Table 1–E Burn Model Estimates and Odds Ratios for *Cirsium vulgare* in Sequoia & Kings Canyon National Parks:
Environmental variables (burn and non-burn) are listed in order of decreasing significance. Estimates are given along with the log-transformed odds ratio. Odds ratios are interpreted as a change in odds per unit change in the predictor variable. A cross (†) indicates a burn variable. "Fire Occurrence" (burned / unburned classification) is the odds of presence in burned areas relative to unburned areas. Burn severity is measured through the remotely-sensed differenced normalized burn ratio (dNBR). An asterisk (*) indicates the variable is a remotely sensed phenology index computed by NASA – each of which was created from multiple Normalized Differenced Vegetation Index (NDVI) values taken over the course of a year. Two variables have both a linear and a first-order polynomial (squared) term. Distance to water represents the Euclidean distance to the closest open water such as a lake or stream.

Variable	Unit	Estimate	Odds Ratio	SE	z Value	Pr(>|z|)
Fire Occurrence †	Burned / Unburned	-3.80E+00	0.022326	6.41E-01	-5.935	< 0.0001
Max. Peak*	NDVI	8.45E-04	1.000845	2.45E-04	3.444	0.000574
Distance to Water	Meters	4.46E-03	1.004471	1.32E-03	3.371	0.000748
Amplitude*	NDVI	6.76E-04	1.000676	2.35E-04	2.882	0.003952
Brown Down Rate*	NDVI Units / day	-1.26E-03	0.998739	6.09E-04	-2.074	0.038111
Burn Severity †	dNBR	-9.45E-03	0.990591	4.66E-03	-2.03	0.042407
Burn Severity Squared †	dNBR	1.92E-05	1.000019	1.15E-05	1.668	0.095302
Elevation Squared	Meters	-1.15E-06	0.999999	7.69E-07	-1.493	0.135549
(Intercept)	None	-4.86E+00	n/a	4.35E+00	-1.118	0.263779
Elevation	Meters	3.02E-03	1.003023	3.46E-03	0.871	0.383504
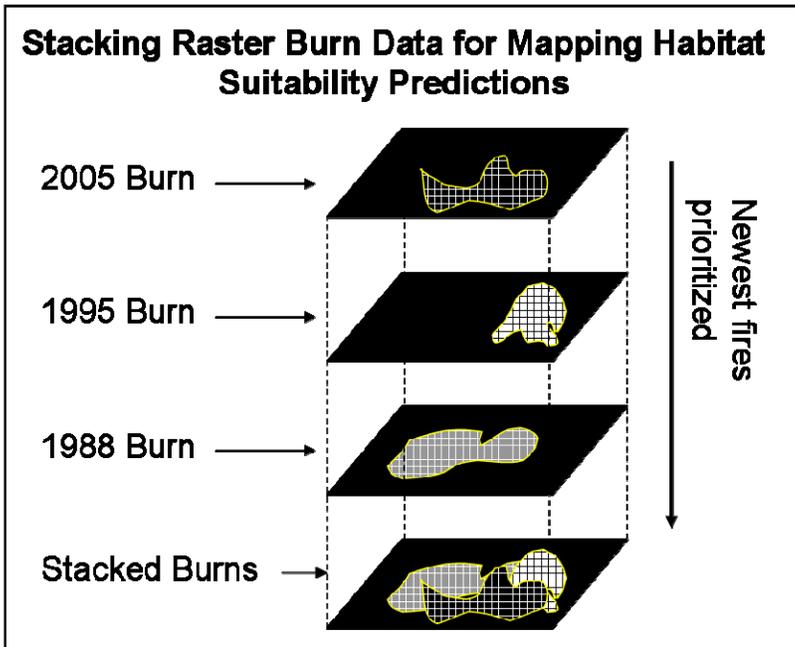
Figure 1-A   Sequencing Burn Data in a GIS for Use in Modeling
Burn GIS layers are "stacked" chronologically, with newest burns taking precedence over older burns in areas of overlap. This type of process allowed the extraction of correct chronologic burn data for model building as well as creating predictions for current post-fire conditions across the study area.
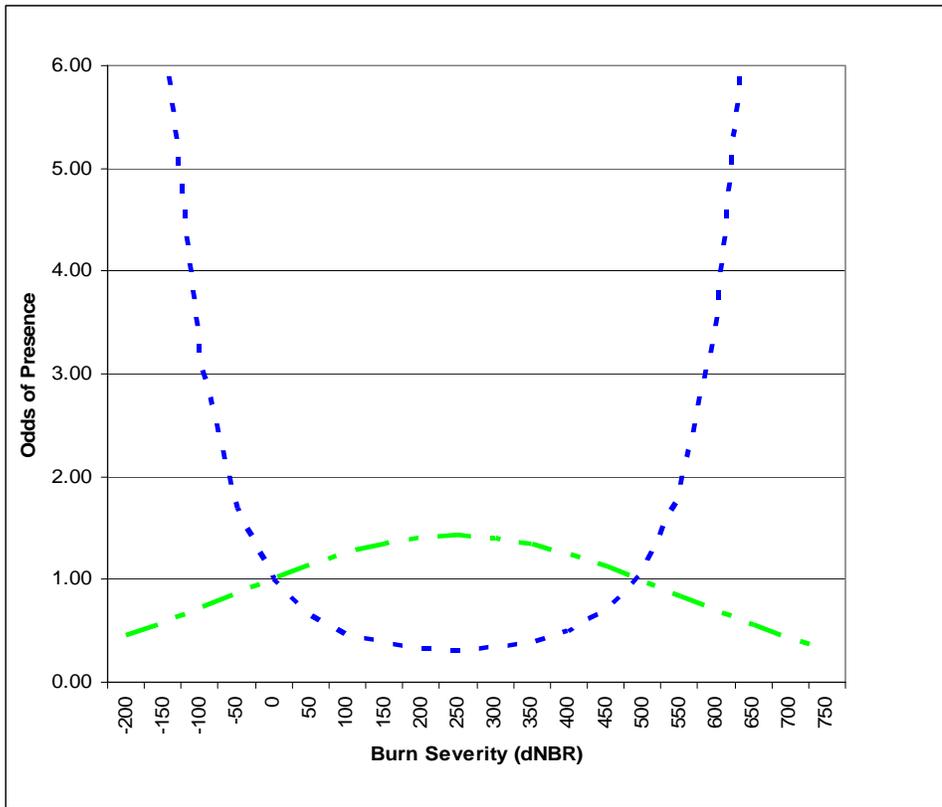
Figure 1-B Odds of Presence in Response to Changes in dNBR for Two Species
C. vulgare (blue line) and *L. dalmatica* (green line) each respond differently to burn severity. *C. vulgare* prefers either end of the burn severity dNBR spectrum: meadows and grasslands with a post-fire "green up" response (less than -100) as well as areas of high burn severity (greater than +500). In contrast, *L. dalmatica* prefers areas of moderate severity.   *C. nutans* is not shown because burn severity was not retained during the best selection process for this model.
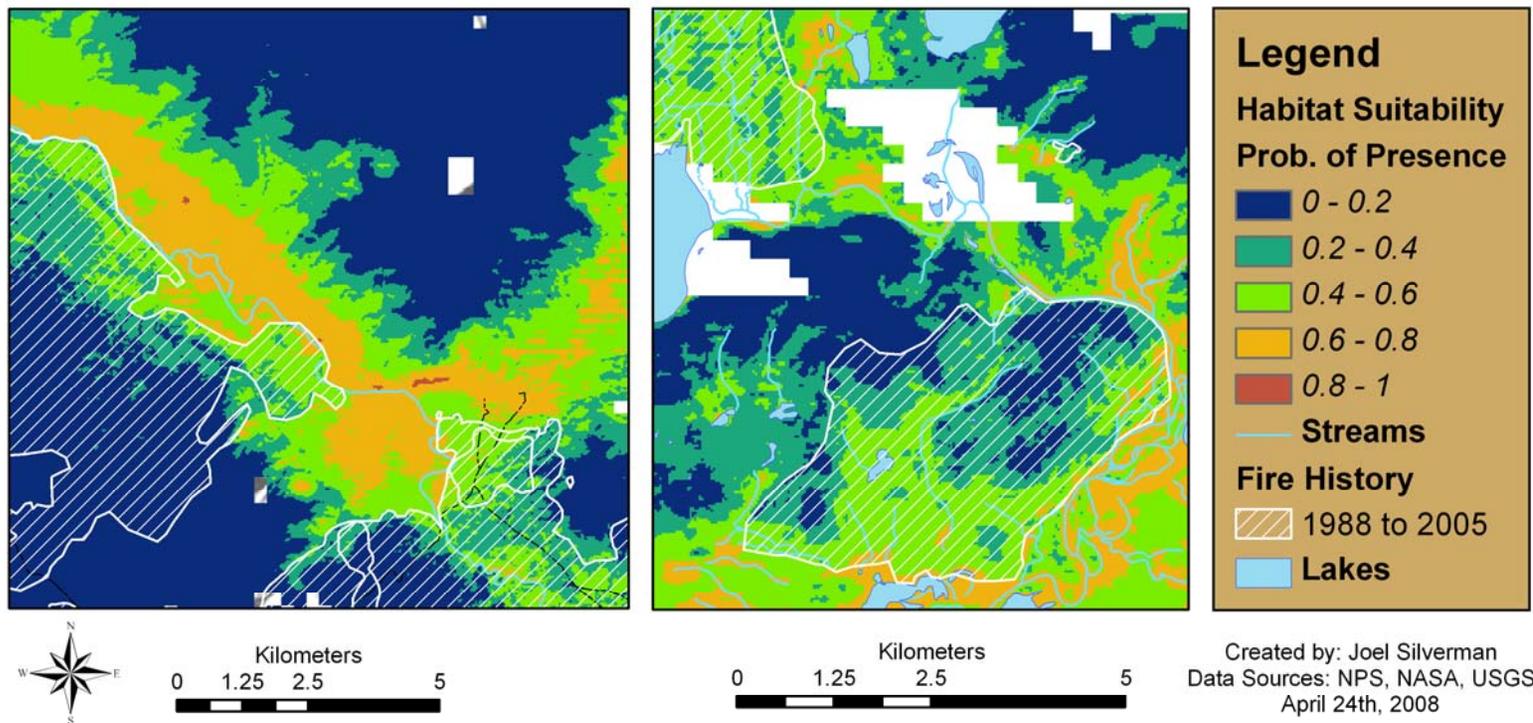
Figure 1-C Comparing Habitat Within Burned vs. Unburned Areas in Grand Teton NP
Habitat suitability for *L. dalmatica* within a portion of Yellowstone NP (left) and *C. nutans* within a portion of Grand Teton NP (right) are depicted below. Habitat suitability varies both within and outside of burns. However this variable alone (burned / unburned status) does not explain habitat suitability. Rather, it is the combination of burn and other environmental factors such as elevation, tree cover, and distance to water which together predict species occurrence in these cases. Areas with white hatch marks have burned since 1987. White block areas are missing data needed for prediction.
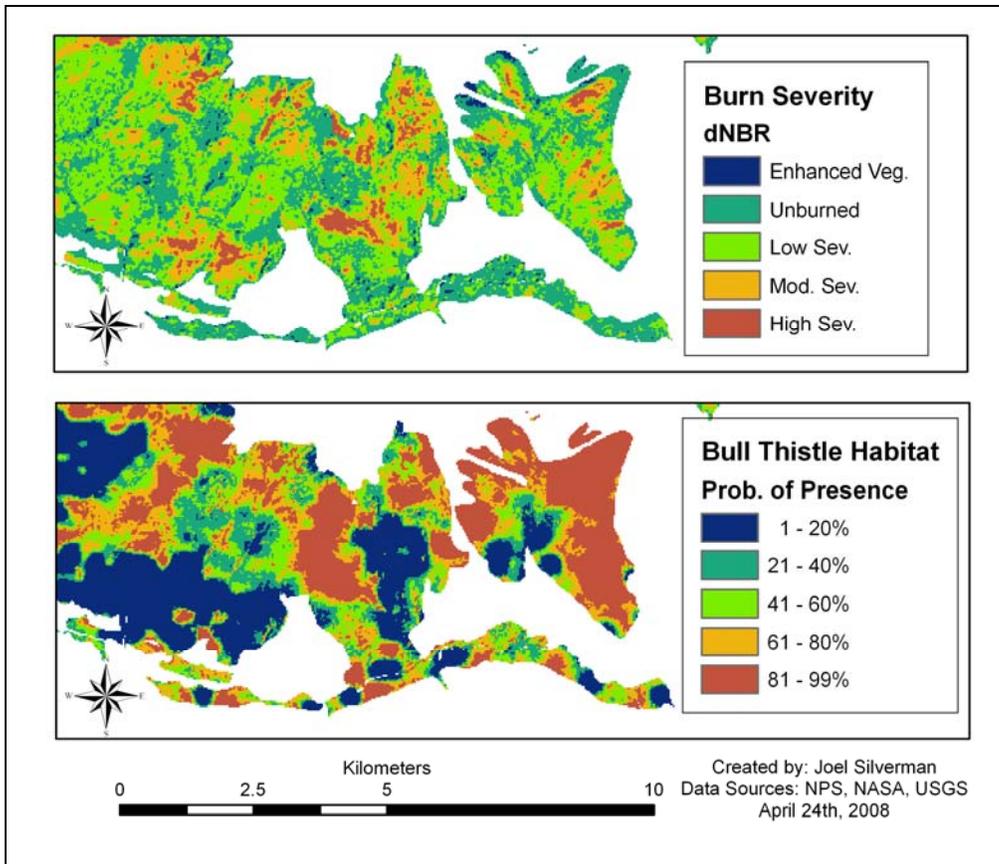
Figure 1-D Comparing Burn Severity and *Cirsium vulgare* Habitat in Sequoia & Kings Canyon NP
Pictured below is a portion of Kings Canyon NP which burned in one or more fires between 1990 and 2004. Since high burn severity (top) does not simply correspond to high habitat suitability (bottom) within the boundaries of these burns, post-fire weed prevention treatments targeting only zones of high burn severity would not be effective.
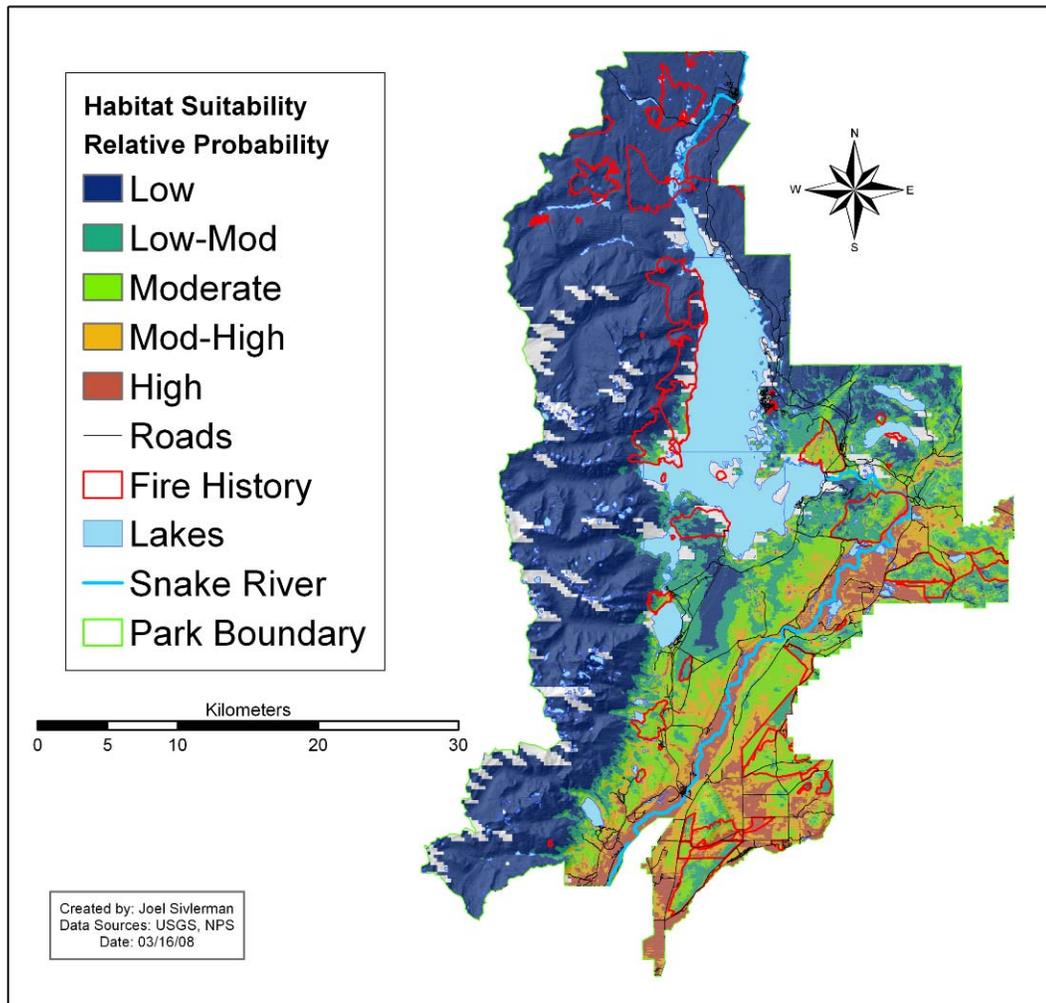
Figure 1-E Habitat Suitability Map for *Carduus nutans* in Grand Teton NP
Areas in red and orange have the highest probability for the presence of *C. nutans*. Topographic relief is shown using shading. High elevation areas (west) show far lower probability of presence while areas in the Snake River flood plain (south east) show high probability for the species. The model is in general agreement with field observations.
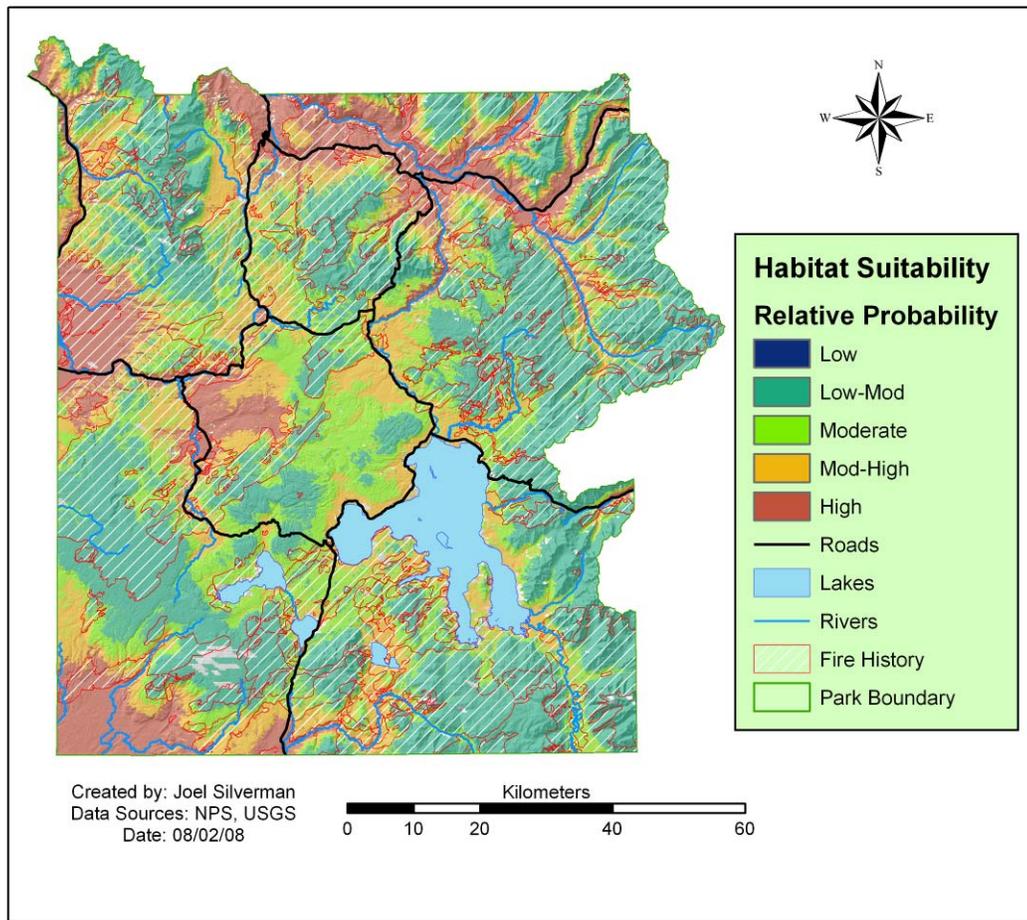
Figure 1-F Habitat Suitability Map for Dalmatian Toad-flax in Yellowstone NP
Areas in red and orange have the highest probability for the presence of Dalmatian toad-flax. Topographic relief is shown using shading. Areas of low elevation and close proximity to rivers have the highest probability of presence for the species. The model is in agreement with observed presence of the species in the northern region, but high probability predictions in the west and south have few field observations off trail for validation. White hatch marks indicate burned areas. Red lines mark the burn perimeters.
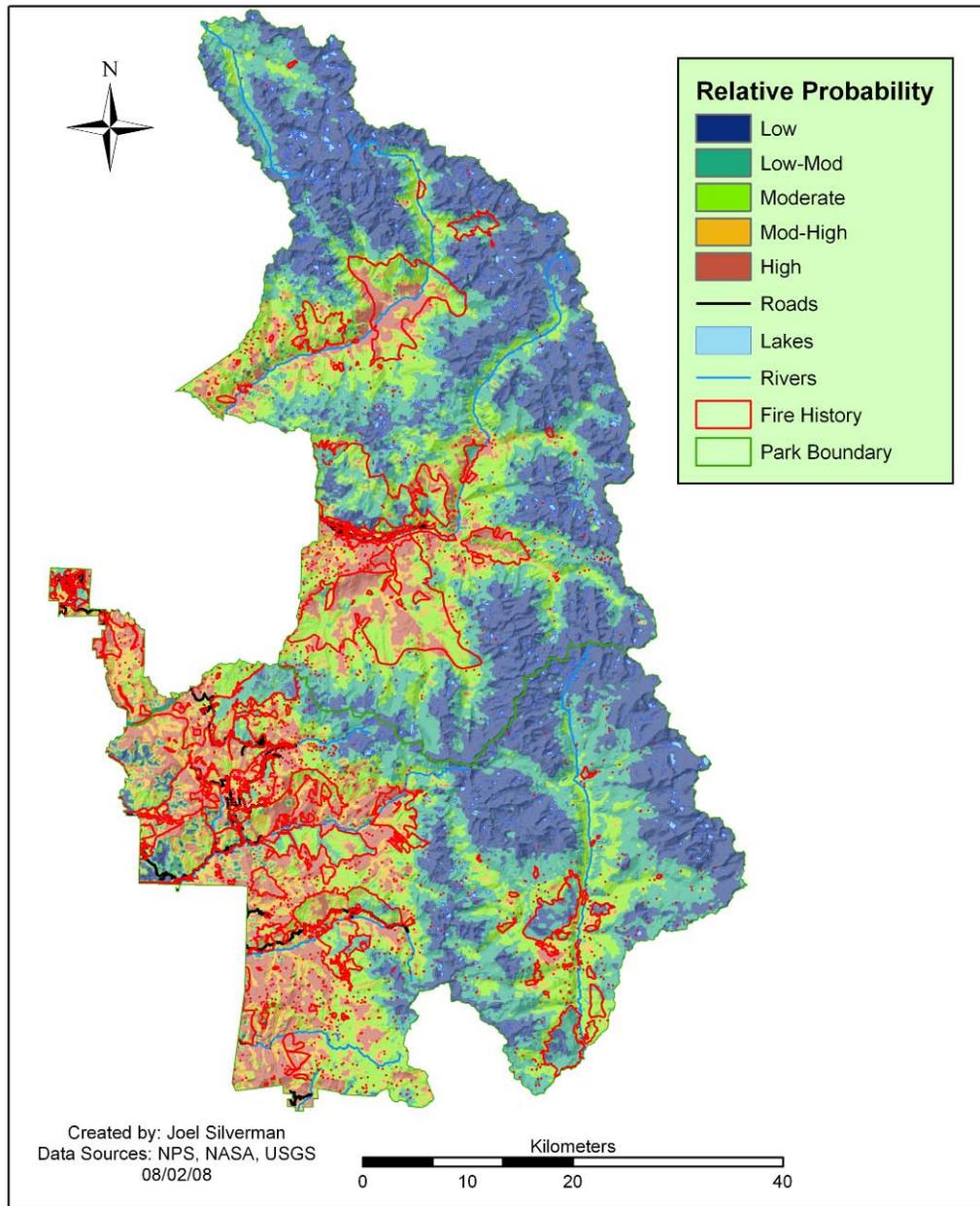
Figure 1-G Habitat Suitability Map for *Cirsium vulgare* in Sequoia and Kings Canyon NP
Areas in red and orange have the highest probability for the presence of *C. vulgare*. Topographic relief is
shown using shading. Areas of low elevation and river corridors have the highest probability of presence
for the species. The model is in general agreement with field observations, but a limited sample of
observations have been collected and were not randomly generated.

Chapter 2: A Case-Study of *Carduus nutans*: Evaluating Post-Fire Vulnerabilities Amongst Several Vegetation Types in Grand Teton NP

Abstract

Fire plays an important role in maintaining biodiversity and landscape heterogeneity, yet this disturbance process can also lead to non-native plant invasion, threatening native plant communities as well as wildlife. In order to stem non-native plant invasion while sustaining natural fire regimes, managers need accurate predictive models, maps, and statistical analysis to effectively target the sites most likely to be invaded. We know that habitat for some invasive species is highly predictable using statistical relationships between species occurrences and nationally-available environmental data layers including fire, yet for other species, such models account for only a small portion of the variance in species distribution (Chapter 1). Here we aim to improve a habitat model for *Carduus nutans* L. (musk thistle) by incorporating detailed, locally-available data sources in addition to nationally-available, moderate resolution data sources on topography, phenology, burn severity, and hydrology among others. Adding a high-resolution vegetation class variable, a fire-vegetation interaction, and several climate variables as predictors of *C. nutans* habitat increased the variance explained by the model while providing valuable insight into vulnerabilities within some vegetation communities. The odds that *C. nutans* would be present at a site increased with increasing burn severity in a uni-modal fashion, while the magnitude and direction of response to time since fire (TSF) was dependent on the vegetation type. *C. nutans* is likely to diminish with TSF in some

vegetation types (lodgepole and hydrophilic), but may continue to increase for multiple years after fire in other types (mixed conifer, deciduous, spruce/fir, and sagebrush). The species generally preferred hydrophilic vegetation when other factors were held constant. A full model, which included burn variables (TSF and burn severity) showed a small but significant gain in explanatory power when compared to a reduced model of only non-burn variables. Both the full model and reduced model in this chapter explained approximately 65% model deviance, whereas previous models, which only included nationally-available moderate-resolution data (Chapter 1), only explained 24% of the deviance.   While burn severity, TSF, and the vegetation*TSF interaction showed only minor overall model improvement, the final model demonstrates the odds of *C. nutans* presence are affected by certain burn conditions. Models which include disturbance-related variables in combination with high-quality vegetation may be valuable when nationally available data sources do not produce adequate habitat maps or when post-fire management objectives differ among vegetation types.

Introduction

Habitat suitability (HS) models, also known as species distribution models, are based on the premise that species distributions are dependent upon the surrounding environmental conditions (Lugo et al. 1999; Guisan&Zimmermann 2000; Miller et al. 2007). In HS models, various statistical methods link geo-referenced field observations to environmental data in order to predict the distribution of a species across a study area of interest (Guisan&Zimmermann 2000). Environmental data commonly used in HS models include information regarding the topography, climate, hydrology, geology, anthropogenic features, and vegetation across the study area. Early HS models predicted distributions of mostly native species, but HS models are increasingly applied to predict invasive species distributions (Higgins et al. 1999; Guisan&Zimmermann 2000; McNab&Loftis 2002; Rouget et al. 2004; Underwood et al. 2004; Jones&Halpern 2005; Rew et al. 2005; Thuiller et al. 2005; Araujo&Guisan 2006; Floyd et al. 2006; Morisette et al. 2006; Schussman et al. 2006; Barnett et al. 2007). The term invasive species is defined here as any non-native species which has a high potential or actual impact on the recipient ecosystem. Invasive species are particularly important to manage because their current unprecedented levels of colonization are negatively impacting native plants, native wildlife, and the underlying processes which maintain ecosystems (Dantonio&Vitousek 1992; Chapin et al. 2000; Davis et al. 2000). Habitat models and their products, maps and statistical inferences, can help managers efficiently monitor the most suitable habitat over areas that are too large or remote for a complete inventory.

When modeling invasive species, it is critical to understand and incorporate relevant ecological theory, and although the field of invasion ecology has made significant gains in understanding the processes involved in invasion (Richardson&Pysek 2006), these concepts have rarely been integrated into invasive species HS models (Guisan&Thuiller 2005). For instance, many theories agree that fluctuating resources due to disturbance provide a window of opportunity for invaders to colonize communities (Davis et al. 2000). Despite this ecological understanding, disturbance processes have rarely been incorporated into empirical HS models (but see Gumpertz et al. 2000; Rew et al. 2005; Lynch et al. 2006; Chabrerie et al. 2008). Even in these few cases, the disturbance modeled is usually a single simplified variable (e.g. binary: disturbed vs. undisturbed). The lack of disturbance process representation in invasive species models may be due to the difficulty of procuring long-term disturbance data or the challenge involved in temporally aligning response data with the given disturbance history. However, the inclusion of ecosystem *process* in addition to ecosystem *structure* is critical to successful restoration and conservation management - especially in light of future conditions imposed by a rapidly changing climate (Harris et al. 2006).

Since disturbance processes, and fire in particular, are often strongly linked to invasion (Dantonio&Vitousek 1992; Levine 2000; Brooks et al. 2004; Keeley 2006), the use of fire-related variables as predictors in HS models should improve habitat maps and/or offer insight into invasion biology. The goals of this study are to employ HS modeling methods in order to predict the potential distribution of one invasive species, *Carduus. nutans* (musk thistle), across Grand Teton National Park (GRTE), and to describe the

relationship between *C. nutans* and fire. Previous work (Chapter 1) identified a negative

relationship between the species and amount of time since the last known fire. However,

the relationship was averaged over diverse vegetation types found across a large

geographic area and wide elevation span. In Mesa Verde NP, *C. nutans* post-fire

colonization and persistence was strongly linked to particular vegetation types such as

pinyon-juniper (Floyd et al. 2006).   Because *C. nutans* is known to prefer some post-fire

vegetation types to others, a high resolution vegetation classification map was

incorporated as a predictor in the model, allowing us to test the post-fire "invasibility" of

six vegetation types found in the park. Field observations (presence or absence of *C.

nutans*) were linked to environmental predictor variables using logistic regression and the

results were used to calculate the probability of presence of *C. nutans* across the study

area (the habitat suitability map). By incorporating vegetation class variables, climate

variables and a "vegetation * burn" interaction we are able to ask:

> 1) How much additional explanatory power can these variables add compared to
>
> the national suite of variables assembled in Chapter 1?
>
> 2) Does the likelihood of *C. nutans* presence vary among different post-fire
>
> vegetation types?

Specifically, the *C. nutans* HS model helps to answer several ecological questions:

> 1) Do burn characteristics (burn severity, time since fire (TSF)) help to predict
>
> the spatial distribution of *C. nutans* in GRTE?
>
> 2) Which post-fire vegetation communities are most easily invaded?
>
> 3) How can these models be applied to fire and invasive plant management?

Assessing the questions and goals outlined above requires understanding the role of fire

in both ecosystem maintenance and frequent promotion of invasion. Fire helps to

maintain biotic structure (fuels), function, and diversity in many ecosystems.(Baker 1992;

Agee; Swetnam 1993; Turner et al. 1999; Dale et al. 2001; Rollins et al. 2001;

Schwilk&Ackerly 2001; Pausas et al. 2004; Bond&Keeley 2005; Smithwick et al. 2005).

Fire helps cycle nutrients, creates a mosaic of seral stages, and creates conditions

necessary for the regeneration of important species such as lodgepole pine and aspen.

For these reasons, fire and conservation management in many parts of the US have

moved from a policy of fire-exclusion to one of fire regime restoration or maintenance.

Although restoring and maintaining fire regimes is a priority, invasive species can

confound fire management efforts because they often colonize recent burns, creating

competition with native plants, lowering diversity, and altering ecosystem processes.

In order to minimize these negative effects, burned areas are sometimes treated to reduce

in invasion. Unfortunately, such post-fire invasive species management is usually limited

by time, funding, and a lack of synthesized information. Managers who are required to

mitigate the risks of fire need tools which can spatially target areas for post-fire invasive

plant management in a cost-effective manner. Creating HS models and maps which

specifically incorporate burn variables and are tailored to specific vegetation types can

help managers understand not only where species are most likely to occur, but also the

species response to specific burn parameters and interactions. This type of information

has value both in planning prescribed burns and mitigation other types of wildland fire

after the actual event. By incorporating burn and vegetation information, HS models can:

summarize responses to fire by vegetation type, create accurate HS maps for the species of interest, and efficiently target areas for weed management.

We chose to model the species *C. nutans*, an invasive plant considered noxious in 22 states and 4 Canadian provinces (USDA 2002). *C. nutans* has long been associated with pasturelands and more recently has invaded natural areas. In our study area, the species is found in most native vegetation types below 9000 feet in elevation and is present in burned and unburned areas. To date, only a few locations are dominated by *C. nutans*. However in other western lands, such as Mesa Verde National Park, *C. nutans* dominates post-fire pinyon-juniper woodlands in burned areas. These infestations are continuing to inhibiting natural post-fire succession (1989 to present), particularly in vegetation communities where few native re-sprouting species exist (Floyd et al. 2006).

Methods

Study Area

Covering 135,073 ha, GRTE ranges in elevation from 1935 m to 4197 m, with broad expanses of sagebrush (*Artemisia tridentada* Nutt.) that have moderate fire frequency & high severity. GRTE also contains mixed-conifer forests that have moderate frequency and mixed severity, aspen groves (*Populus tremuloides* Michx.) that have moderate frequency and high severity, and spruce-fir forests (*Picea engelmannii* Parry ex Engelm and *Abies lasiocarpa* (Hook.) Nutt.) that have low-frequency and high- severity fires.

Logistic regression and general estimate equations (GEE) were used to link field observations of *C. nutans* to the environmental variables. The former technique explains the direction and relative importance of each environmental variable. Once the parameter estimates were computed, a model selection process was used to find the most parsimonious set of non-burn variables (non-burn model). To this set, burn variables and an interaction between TSF and vegetation type were added (burn model). Burn and non-burn models were compared using measures of model fit, accuracy, overall performance, and predictive ability compared to random chance. Odds ratios for specific variables were graphed and a final habitat suitability map was created.

Since regression estimates can be distorted due to spatial-autocorrelation inherent in geographic data (non-independent observations), a second technique was used to detect the degree of distortion caused in the final logistic regression model. This technique – GEE - gives consistent estimates of the marginal response to the environmental variables by accounting for spatial auto-correlation in the residuals (Zeger&Liang 1986; Miller et al. 2007). GEE and Logistic regression estimates and standard errors were compared.

Environmental Data

"Static" Environmental Variables

A number of environmental variables commonly used in HS models were initially chosen for testing. These were categorically different than the burn variables in that they were relatively stable and did not have discrete starting points in time. These are referred to as *static variables* to differentiate them from the disturbance-related variables. In addition,

only free data sources were utilized, in order to build models which would be easily updatable by land managers. These data varied in spatial resolution to some degree and several of them represent temporal averages. Data were selected from several broad categories: climate (1000 m grid resolution), topography (30 m), hydrology (30 m), vegetation (30 m), and satellite-derived phenology (250m) (Table 2-A). Appendix A describes the details of the GIS analysis.

Climate variables were obtained from DayMet climate data (Thornton 2001). DayMet layers contain interpolated climate surfaces from eighteen years of daily meteorological station observations (1980 to 1997) (Thornton et al. 2000). These grids are 1 km resolution and are available for the continental US. Most layers were used directly but several were calculated in ArcGIS v9.2 (ESRI 2006). The calculated layers include summer precipitation, driest month, minimum average monthly temperature, and temperature seasonality (see Appendix A). These methods were adopted from Jones and Halpern (2007).

Candidate topography variables included a 30m resolution digital elevation model (DEM) and a DEM-derived annual potential solar radiation layer calculated using ArcGIS (ESRI 2006). Annual potential solar radiation calculates the potential energy using information from nearby cells, sun angle, slope, and aspect. In order to keep the number of parameters to a reasonable number, only one quadratic term was specified - elevation. The Euclidean distance to open water sources such as lakes, rivers, ditches, and reservoirs

was calculated using the GRTE hydrology geo-database (NPS 2008) and output was scaled to a 30 m resolution.

Six vegetation types were aggregated from the GRTE vegetation map, which followed the National Vegetation Classification (NVC) methodology (Cogan et al. 2005). Several vegetation types: high elevation conifer, tundra, shrub and barren sites were combined into vegetation class 1. Since this class was least likely to burn and would have the harshest physiological conditions, all other classes were compared against this one in model analysis. The remaining classes were as follows: hydrophilic (wetlands, riparian, etc.), lodgepole, mixed forests (lodgepole, subalpine fir, spruce, deciduous mixed, etc.), pure deciduous forest, & sagebrush (no re-sprouting sagebrush species included) (Appendix B).

The phenology layers were based on a series of algorithms developed by NASA, which use MODIS images to quantify seasonal changes in greenness at a continental to regional scale. The indices collectively known as TIMESAT, is a series of National Differenced Vegetation Index (NDVI) images compiled over the course of 1 to 2 years and fit to a smoothed sine wave function that represents one annual cycle of change in greenness (see Appendix A) (Jonsson&Eklundh 2004).

Disturbance Process Variables

The two primary sources used to create the disturbance variables were fire history polygons and a widely-available index of burn severity – differenced Normalized Burn

Ratio (dNBR) raster data. The fire history polygons provided information on each fire perimeter and the year of the fire. These were subsequently converted to raster data for analysis and were also used for extracting the dNBR values within the burn perimeter. The second source, dNBR, detects change between a pre- and a post-fire image using infrared and near-infrared spectral bands. These bands (4 and 7 on the Landsat Thematic Mapper platform) correlate to change in greenness and soil exposure. The Normalized Burn Ratio (NBR) algorithm is given as:

NBR = (Band 4 - Band 7) / (Band 4 + Band 7)

NBR is calculated for an image prior to the fire (usually the same year as the fire or the previous year) and a second image one year post-fire. The two images are usually taken during the growing season and during the same time of year (e.g. June 2002 and June 2003) to ensure a match in the phenology between images. The difference between the two images (dNBR) is then found by the following:

dNBR = NBR$_{PRE}$ - NBR$_{POST}$

The dNBR index corresponds to biotic and physiologic changes due to fire such as tree scorch, char, litter consumption, and the consumption of other live fuels (Key&Benson 2005). The values of dNBR range from -1000 to 1500 but only values within -550 to +1350 are usually attributed to fire. Within this range the basic thresholds are as follows:

Enhanced Vegetation: -550 to -100

Unburned: -100 to +90

Low Severity: +91 to +200

Moderate Severity: +201 to +350

Moderate-High Severity: +351 to +600

High Severity: +601 to +1350

Extreme values may represent cloud, smoke, snow or other obstructions (Key&Benson 2005). The images were also calibrated to correct for differences in mean image value found between images using techniques described by Key & Benson (2005). Calibrated values were extracted from within burn perimeters and all surrounding areas were given a value of zero (no change).

Since the burn variables represent the aftermath of a spatially and temporally discrete set of disturbances, it was necessary to develop rules for incorporating them into the models. Burn variables needed to be arranged for two separate steps in the model protocol. First, invasive species observations and burn variables had to be correctly fit into a matrix for analysis. Burn values had to be the most recent values preceding the weed observations. No burn values after the weed observation would be used. Second, burn values needed to be chronologically "stacked" (figure 4) for the entire set of years analyzed (1988 to 2006). These were used in calculating the final predictive surface (HS map) for the study area, thus making the predictions relative to current post-fire conditions. This required some simplifications described previously (see Chapter 1 and Appendix A).

Response Variable

The geo-referenced observations of *C. nutans* (presence or absence) used as the response

variable in this study were compiled from several NPS monitoring and mapping project

databases (provided courtesy of GRTE). All plots met several criteria including: a

complete plot search for *C. nutans*, accurate GPS coordinates (+/- 5 to 10 m), and search

area ranging from 300 to 1000 m^2 (if marked absent). See Appendix C for details on

each data source.

Using simple queries in a GIS, we determined few of the existing field observations were

located *spatially within* burn perimeters and *temporally after* fires. To correct this lack of

data, field sampling was conducted in the summer of 2007 using a random stratified

sampling design. Sample locations were selected within burn perimeters using ESRI Arc

GIS v9.2 and Hawth's Ecological Tools (Beyer 2004; ESRI 2006). The stratification

combined burn severity class (low, med, high), and forested / non-forested classification

based on the GRTE vegetation map. Points were selected within recent burns (1990 to

2006) that were greater than 40 acres in size. No additional points were collected within

the 1988 burn perimeters because these were already well represented from vegetation

mapping plots. In addition, a few "opportunistic" presence points were collected en route

to the random sample points because the species occurs relatively infrequently, making

collecting a sufficient number presence points difficult. At each opportunistic point a

plot was installed and geo-referenced using GPS. Our circular field plots measured 314

m^2 and represent a compromise between the spatial resolution of the predictor grids (30

m x 30 m) and the time required to confidently search a plot for individual rosettes of *C. nutans*. Plots were designed to be similar in size and shape to the standard dNBR field calibration plots (CBI plots) and other standard plot protocols used by the National Park Service (Key&Benson 2005).

Analysis

Model Array Construction

The process of forming a data matrix listing each observation and the corresponding "static" environmental data was a simple GIS operation, however, adding the relevant burn information to the array required additional steps. A series of stacked burn layers were intersected with the point observations and conditional statements were used to select the most recent burn information prior to the given species observation (See Appendix A).

The burn variable, TSF, was calculated for the data matrix as the number of years between the *C. nutans* field observation and the most recent previous burn. If no known fire date existed for a given observation, it was given a TSF value of 100 years. Observations which had only burned between 1931 (first year of spatial fire history) and 1987 were removed from the model array. Thus, only burn history information from 1988 to 2006 was utilized in the model array. There were several reasons for this simplification: 1) *C. nutans* was not found in the park prior to the mid-1990s which means fires prior to 1988 probably had little to no propagule pressure from the species, 2) forest succession reverted canopy cover in most old burns back to a similar level as the

"unburned" vegetation leaving poor light conditions for *C. nutans* (Zouhar 2002), 3) the combined acres burned between 1931 and 1987 only comprise a small percent of the total burned area (1931 to 2006) or the total area within GRTE, and 4) this simplification allows each observation to include both a TSF value and a corresponding burn severity value (available for 1988 to 2006). A total of 91 observations within burns (burned between 1931 to 1987 *and* observed after the fire date) were removed from the model for this reason, leaving 2161 observations for the model fitting process.

Predictor Variable Collinearity

Collinearity can lead to inflated significance levels in regression. To prevent this, we used a variance inflation factor (VIF) package available for R Statistical Software (Fox&Monette 1992; R 2007). Variable(s) with a VIF > 10 are considered strongly collinear with other variables (Chatterjee 2000). The variable with the highest VIF was removed and the regression was run again in an iterative fashion until no more variables were collinear (no VIF > 10). Quadratic terms were collinear with their linear counterpart, but were added back at the end of this process because they were deemed likely to give certain predictors a more realistic response curve. In addition, non-significant quadratic terms were likely to be dropped during the following model procedure.

Modeling Framework

A partitioned portion of the data (80%) was used for fitting the model. Logistic regression was run on just the non-burn variables using the Step function. This function

iteratively run regressions using both backward and forward selection to determine the

most parsimonious model as measured by Akiake's Information Criteria (AIC). This is

referred to as the *non-burn model*.

Next, the burn variables and one interaction term were added to the non-burn terms

chosen by the step-selection process outlined above and a regular logistic regression with

out step selection was computed. This model is referred to as the *burn model*. Note that

the non-burn model is a reduced version of the burn model, allowing an analysis of

deviance test to be computed.

Considering Spatial Auto-correlation

The basic assumption of independence between observations is inherently violated in

regression models which use geographic data such as elevation. This is because

geographic observations close together are generally more similar than would be

expected by random chance. This lack of independence can lead to biased estimates

and/or inflated standard errors known as spatial auto-correlation (Zeger&Liang 1986). In

order to check for spatial auto-correlation in the logistic regression estimates, a general

estimate equation (GEE) was also run on the step-selected variable set. GEE is a type of

marginal model, which is resistant to the distortions caused by dependence in the

response variable (see Chapter 1). The GEE and GLM burn model parameter estimates,

standard errors, and p-values were compared. If estimates or p-values shifted > 0.05 in

GEE, the GEE estimates would be used for predictions. If p-values did not shift by this

amount, GLM estimates would be retained because GLM models have the advantages of

allowing comparisons using AIC, deviance statistics, and provide conditional response

instead of marginal response (Millar et al. 2007).

Model Assessment

The burn and non-burn models were assessed using several measures of overall

performance and classification accuracy (Table 2-C). Data were tested on both the data

used to fit the model (80%) as well as a validation set (20%). Because modelers disagree

on the best measures of accuracy a number of techniques were used, including both

threshold-dependent and threshold-independent methods (Foody 2002). Thresholds

require the user to *choose* a threshold probability at which predictions are classified as

either zero (if p < threshold) or one (if p > threshold), but the choice of threshold can

influence the measures, making comparison between studies difficult. However,

thresholds allow users to choose how much type I or type II error they are willing to

accept when creating a predictive map for treatment or monitoring (Fielding&Bell 1997).

Choosing a threshold is also necessary when delineating a set number of acres for survey,

treatment, protection, etc. We assess sensitivity (percent total presences predicted),

specificity (percent total absences predicted), classification accuracy, and the true skill

statistic (TSS). TSS measures how much better than chance a model classifies data, and,

unlike Kappa, it is unbiased by prevalence (Allouche et al. 2006).

Mapping Predictions – The Habitat Suitability Map

The coefficients for each parameter estimate in the most explanatory model were used to create a final habitat suitability map. The analysis was computed in ArcGIS using the raster grids representing each variable. The equation is the basic logistic form:

$$y_i = \left(\frac{\exp^{\beta_0 + \beta_1 x_k}}{1 + \exp^{\beta_0 + \beta_1 x_k}} \right)$$

All input layer extents were set to the bounding coordinates of GRTE and all input grid resolutions were resampled to 30m (if not already in this resolution) and smoothed using cubic convolution (ESRI 2006).

Results

Using climate variables, high quality vegetation class variables and a vegetation * TSF interaction, provided ecological inference about post-fire conditions as well as increased deviance explained (from 24% within the burn model of Chapter 1 to 67% within the burn model in Table 2-C). Despite this fact, a best-selection set of variables *which included burn variables* performed quite similarly to a subset *which omitted burn variables* (deviance explained only increased by 1.8% between the burn and non-burn models within the fitting dataset) (Table 2-C). Thus while certain burn characteristics may raise the odds of presence (e.g. in certain vegetation types or in certain burn severity levels) they only marginally improved the predictive map for *C. nutans* across the study area.

Burn vs. Non-Burn Model

Using collinearity diagnostics and an AIC-based step-selection method, a best subset of potential "static" and disturbance process variables were chosen. This model was then compared using various measures of accuracy and overall model performance. For example, sensitivity and specificity were very similar between burn and non-burn models (~ 77% accuracy). The TSS scores were both better than prediction by chance, but also similar between burn and non-burn models (Table 2-C). Finally, the deviance explained was near 66% in both cases and both AUC of ROC values approached 86%. A drop in deviance test between models however did show a small but significant (p = 0.02) drop in AIC when using the burn model.

Environmental Variables Selected

The collinearity diagnostic and step-wise selection processes identified a parsimonious subset of variables needed to predict habitat suitability for *C. nutans*. This subset included "static" and disturbance process variables from all categories: climatic, hydrologic, phenologic, topographic and burn-related (Table 2-B). The model selection process retained both GIS and remote sensing imagery variables, demonstrating that a diverse pool of data sources is important in HS modeling. The final model variables are listed in descending order of importance in Table 2-B.

Climate variables were well-represented in the final model despite their coarse resolution (1 Km) and the possibility that they could have been collinear with phenologic variables.

Probability of presence for the species decreased with an increase in precipitation frequency, an increase in annual humidity, increasing summer precipitation, and lower monthly minimum temperature. Vegetation information was also included in the final model. Class 1 High elevation/sparse vegetation was used as the baseline of comparison (treated as the intercept) and was significant at $p = 0.0043$. In comparison to Class 1, sagebrush showed a decreased odds of presence while hydrophilic vegetation odds of presence increased five- fold (Figure 2-A). Odds of presence also increased with canopy cover level. Increased distance to water sources decreased the odds of presence. Topographic influences such as solar radiation and elevation were important as well. Presence increased directly with increased radiation, while the highest odds of presence in relation to elevation were at moderate levels (peak near 1600 m).

NASA's TIMESAT Phenologic variables were also important in the final model. While areas of low productivity - as measured by "average base" and "base level" increased the odds of presence, early "peak date" greenness decreased the odds of presence. Fall seasonal "brown down" rate was only marginally significant ($p=0.15$) and probably not reliable for inference (see Appendix A for phenologic definitions).

Two burn variables, burn severity and TSF were both retained in the final model. Burn severity retained both a linear and quadratic term which together, increase the odds of presence as the severity score rises toward 725. Relatively little area in GRTE burned at scores higher than 500, making it difficult to tell if high severity fires continue to increase the odds of presence, or if the odds begin to level off at moderate-high severity (Figure 2-

B). At a dNBR score of 500, the odds ratio for *C. nutans* presence rose to its maximum

of 2.5 times that of unburned sites (dNBR of 0). The second burn variable, TSF,

indicated a small increase in odds of presence as burn age increases – at least within the

range of most of the data (1 to 12 years post-fire). For instance, the odds of presence in a

twelve year old burn was 1.10 times that of a one year old burn site (Figure 2-C).

Finally, the interaction between vegetation class and TSF showed several vegetation

types responded significantly different to TSF than Class 1 (high elevation/sparse)

(Figure 2-C). As stated above, *C. nutans* presence was positively correlated with

increasing TSF in general. Mixed conifer and deciduous interactions with TSF were not

significantly different from this relationship. In sagebrush vegetation an increase in TSF

also increased the odds of presence, but at a significantly different rate than the former

types. Lastly, lodgepole forests and hydrophilic vegetation both showed decreasing odds

of presence as TSF increased; each at a different rate (Figure 2-C).

Spatial Auto-Correlation

Results from the GLM (Model 2) were compared to those computed by the GEE to

determine if spatial auto-correlation had inflated the estimates and standard errors. These

values were very similar, with no significant term (terms with p-value less than 0.10)

showing a change in p-value of more than 0.05. This indicated there was little spatial

dependence left within the GLM model residuals and that GLM results were statistically

valid. Only GLM statistics are reported.

Mapped Predictions

The habitat suitability map shows likelihood of *C. nutans* presence is highest in the southwestern portion of the study area (Figure 2-E). This portion of the park is characterized by low elevations, a substantial river corridor (the Snake River), higher road density, higher bison and elk densities, and river recreation. It also encompasses the greatest proportion of non-forest and open woodland vegetation types. Many of these trends match the significant correlations between the response and the predictor variables such as elevation, water resources, climate, and vegetation (Table 2-B). In some cases, habitat suitability within burns is visibly different than adjacent lands (Figure 2-F).

Discussion

Our methods found variables from a variety of categories to be significant predictors of *C. nutans* occurrence, including the disturbance process of fire, phenology, topography, climate, and vegetation. This breadth of sources led to models with higher explanatory power than the *C. nutans* models in Chapter 1. In the case of *C. nutans*, using only "static" variables (the non-burn model) yielded similar results to a model that included disturbance process variables as well (the burn model). Model predictive ability was quite similar between the two (Table 2-C), yet the inclusion of disturbance process data offered valuable inference regarding the effects of a fire on the distribution of *C. nutans*. For instance, there is ecological value in knowing that a dNBR value of 450 doubles the odds of presence compared to an unburned area (Figure 2). Further, presence of *C. nutans* appears to respond to burn age; and the direction and magnitude depends on the type of vegetation (Figure 3). Describing these differences in post-fire vegetation type

invasibility is an important contribution to invasion ecology. Disturbance processes are often mechanisms underlying invasion because temporary pulses of resources coupled with increased propagule pressure frequently co-occur following disturbance (Davis et al. 2000; Levine et al. 2003; Pausas et al. 2006b). For this reason, the combination of process variables and "static" variables are important to future invasive species HS modeling.

Burn Severity

Perhaps one of the most important discoveries in this study was the positive relationship between the presence of *C. nutans* and burn severity. Increasing burn severity correlates to lower canopy cover, higher tree & understory mortality, and an increase in bare mineral soil (Key&Benson 2005; Lentile et al. 2006; Abella et al. 2007). These changes provide a temporary lack of competition and an influx of available resources such as light, water, and nitrogen that are critical to invasion success (Davis et al. 2000). Negative dNBR values associated with post-fire enhanced regrowth (Key&Benson 2005) did not show an increase in odds of *C. nutans* presence. However, since only 1.6% of our total number of burn observations occurred in this range of dNBR (-550 to -100), our sample was likely ineffective in measuring any effect in this range. In addition, much of the open vegetation in GRTE is sagebrush, a vegetation type which has shown poor correlation to burn severity as measure in the field (Zhu et al. 2006). If dNBR does not correlate to actual burn severity on the ground in sagebrush, it will likely be a less effective predictor of *C. nutans* occurrence in post-fire sagebrush. Despite this, dNBR is

a good indicator of burn severity for most vegetation types in the study area and was important to understanding the effects of fire on the species.

Incorporating Temporal Patterns in HS models

Adding burn variables to the model required consideration of the temporal dimensions of the invasion. This was important to analyze because some invasive species form only a very temporary post-fire dominance, while others persist indefinitely. Most of the observations occurred between 1 and 12 years post-fire; hence predicting beyond this range is unadvised. While the model array construction disentangled the invasive species observation dates from the fire dates, the predictive map reflects the current temporal and spatial burn pattern. Because of this, it is difficult to visually assess model accuracy by simply overlaying field presence and absence points over the predictive surface. In other words, *C. nutans* observations and fires are stacked layers in time – you cannot view them in two dimensions. It is more meaningful to use the performance measures such as ROC of AUC because they compare the observation conditions to the predictions *as they were at the time of observation.* As with all invasive species, they are not at equilibrium with their environment and may be quickly adapting to new ecotones. Thus, collecting additional observations and making new maps every few years is necessary to reflect increases in the species potential range.

Long-term, repeated-measurements of invasive species would allow for greater confidence in elucidating cause and effect between disturbance processes and invasive plant responses. Vegetation mapping observations collected in the early 2000s could be

re-sampled as more fires (wild or otherwise) occur in future years. Long-term monitoring would also aid in modeling dispersal spread-rates for the species.

Invasibility of Vegetation Types – General vs. Post-fire Vulnerabilities

Simply looking at *C. nutans* general response between vegetation types, hydrophilic is the most likely to be invaded and sagebrush is the least likely. The interaction between vegetation type and burn age suggests that some vegetation types remain easily invaded after fire while others return to unsuitable habitat relatively quickly. The relative odds of *C. nutans* occurrence in lodgepole and hydrophilic vegetation decreases as TSF increases. Conversely, deciduous, mixed conifer, sagebrush and vegetation class 1 may exhibit prolonged increases in odds of presence after fire.

In Mesa Verde National Park (MEVE), vegetation communities lacking post-fire resprouting species were clearly linked to post-fire invasion by *C. nutans*, but resprouting shrublands remained resistant (Floyd et al. 2006). Further, *C. nutans* persisted in non-resprouting vegetation for over 13 years after fire. In GRTE, we could not demonstrate vegetation functional traits (such as resprouting) were linked to the responses of *C. nutans* in GRTE. However, it is possible that sagebrush and deciduous communities remain open for a longer window of time following fire because such sites offer greater light resources to invading forbs. In contrast, lodgepole and hydrophilic communities may be quicker to close off light resources. More research is necessary to determine which functional traits are responsible for the patterns we found (Figure 3).

Other Environmental Variables

It is also worth noting the strong positive correlation between *C. nutans* presence and hydrophilic vegetation (also sites closer in proximity to open water), while conversely *C. nutans* often also occupies sites associated with climatic aridity (see climate variables, Table 1). This may be an indicator that the species prefers sites with high variation in seasonal water availability or it may simply mean the species can tolerate a wide range of moisture conditions. The relatively low probability of presence in sagebrush (not post-fire sagebrush) habitat (Figure 1) seems to suggest the former, but there are likely other unknown explanatory factors such as soil type as well. Some, but not all of the densest infestations in the park occur in areas near a reservoir (Jackson Lake) or the Snake River. Both experience human-caused seasonal water level fluctuations due to downstream agricultural water rights. Thus, another disturbance process, human manipulation of water resources, may also be driving the invasion of *C. nutans* in GRTE. Conversely, neither area receives much direct impact from park visitor, thus the species cannot be assumed to simply follow roads, trails, and infrastructure.

One final disturbance in the park worth considering is the pressure exerted by large populations of herbivores (elk, bison, and antelope). In GRTE, some of the highest densities of *C. nutans* observed are in areas where the disturbances of heavy elk browsing, fire, and/or seasonal inundation occur, suggesting complex interactions may exist between multiple disturbances and *C. nutans*.

Data Resolution & Interpretation

Other important considerations in HS models include the resolution of the input layers,

the plot size, and the scale of interest. It is almost impossible to survey a complete 250

meter or 1 kilometer area and hence, absences scaled up from the plot size will likely fail

to record all presences within the pixel. Even if plot and layer resolution match well,

they may both fail to address the scale of interest, that is, the scale at which managers or

researchers wish to understand patterns, apply treatments, etc.. In this study we used

variables with pixel resolutions ranging from 1000 m to 30 m. We considered the

implications of mixing variables of differing resolutions and chose to include several

coarse-scale data layers: climate, phenology & tree cover layers (1000 m, 250m, and

250m respectively). For the final HS map, smoothed versions (cubic convolution) of the

coarse-resolution layers were used, however, because of the use of these coarse layers,

the scale of our prediction surface becomes somewhat difficult to interpret. A very

conservative user may wish to only consider results at the 1000 m or 250 m resolution,

while others will undoubtedly wish to harness the information provided by the 30 m

resolution data sources. When interpreting odds ratios it is important to remember the

scale of the variable of interest. We concluded the use of the coarse-resolution

environmental layers was ecologically appropriate because they tended to influence

vegetation responses (such as invasion) at coarse spatial scales (e.g. climate affects

species distributions at a coarser scale than topography).

Mapping Habitat vs. Mapping Risk: Incorporating Dispersal

We purposely omitted "distance to road" and "distance to trail" variables from our model to focus on modeling species habitat somewhat independent from the issue of dispersal. Dispersal is a complex spatio-temporal process which requires a mechanistic analysis that to our knowledge has not been accomplished for *C. nutans*. A simple overlay operation in a GIS or a more advanced cost-path analysis could be combined with the habitat map to directly incorporate information regarding known infestations and vectors of dispersal. However, without quantitative studies on the mechanisms of dispersal, such exercises are somewhat subjective. In GRTE, *C. nutans* is usually low-density but widely-scattered. The few sites where *C. nutans* is known to form dense, dominant infestations are not simply near roads and trails. This indicates the species is 1) already well-dispersed and 2) uses other vectors besides humans traveling along roads and trails. Marking buffer-zones around known infestation sites could help focus limited resources on the nearby habitat at greatest risk to high propagule pressure, but this still assumes all infestation sites are already known. Until dispersal vectors are well-understood for *C. nutans* in the study area, GIS buffer operations combined with habitat suitability may be the best maps of overall invasion risk.

The final habitat maps, odds ratio table, and model assessment indicate diversifying environmental data sources can improve predictions. The odds ratio results and model coefficient p-values indicate fire has a moderate influence on the distribution of *C. nutans*. Collecting additional data within certain infrequently-occurring portions of the dNBR spectrum, across rare vegetation types, and across a longer time span (greater than

twelve years post-fire) would help to increase confidence in some responses of the species. These caveats aside, these maps provide a clearer picture of the species distribution than was previously known either through expert knowledge or more generic environmental data modeling (Chapter 1).

Management Implications & Conclusions

HS models with burn information can be of benefit in determining and quantifying fire effects (if any) on a given invasive species – allowing for a number of applications. Simply knowing the response to burn severity may be important for some fire management plans, while spatially quantifying potential habitat may be important in other cases. If the species responds strongly to fire, highly suitable habitat in units slated for prescribed burning should be pre-treated, given special prescription, or excluded. Known infestations near or within burned areas should be prioritized for treatment depending on the post-fire invasibility of the vegetation type. Burn severity inferences could help determine the optimum fire behavior conditions needed to meet invasive plant cover-reduction targets. For GRTE, we recommend monitoring: 1) burns 2-3 years post-fire, 2) sites with dNBR values of moderate to high burn severity, and 3) prioritizing vegetation types which remain vulnerable after fire (sagebrush, mixed conifer, deciduous, & class 1).

HS models can quantify acres needing post-fire treatment depending on constraints of budgets. If limited reseeding or herbicide application is possible, our HS model map could guide area selection by setting the probability threshold to a point which allows a

feasible number of treatable acres.   This method would be cost-effective, reasonably

objective, and spatially explicit.   In contrast, basing post-fire invasive plant treatment

decisions simply in areas of high burn severity or within close proximity to roads may be

ineffective. Model-based maps offer statistically sound predictions which incorporate

both disturbance process and environmental structure.

In order to advance HS models predictive and inference abilities, researchers should

consider incorporating not only the "static" ecological characteristics of the study area,

but the ecological processes that shape it. This may include disturbances, interactions

across trophic hierarchies, and their inherent temporal and spatial components (Levin

1992; Bailey&Whitham 2002; Pausas et al. 2006b). Analysis and applications should be

developed with the specific life history, dispersal vectors, and distribution characteristics

of the specific species of concern. Developing GIS layers which map the extent, severity,

and temporal dimensions of disturbances should be a priority for future habitat suitability

modeling.

Tables and Figures (Chapter 2)

Table 2-A Variables Considered for Modeling Habitat within Grand Teton
The response variable is the presence or absence of the invasive species observed, collected and aggregated from various NPS field surveys and studies (plot sizes ranged from 315 to 1000 meters^2). Climate variables (DayMet) are interpolated surfaces created at the University of Montana using daily weather station observations compiled from 1980 to 1997 (available online). NPS fire history and hydrology data layers are available online through the NPS data clearinghouse. Burn severity (dNBR) data are available through the online Burn Severity and Mapping Project database. Digital elevation maps and National Land Cover Classification (NLCC) tree canopy cover layers are available through USGS online databases. Solar radiation is derived from the elevation layer using ESRI ArcGIS v9.2 software. All remaining layers are remotely sensed phenology indices supplied directly from NASA through the Earth Observation Science program, and are available for land management agencies. * Indicates the layer was calculated in ArcGIS v9.2 using the input data source listed. ** Indicates the layer was first converted from polygon (vector) type data, and then calculated using various Spatial Analyst raster tools in Arc GIS v9.2. Pixel resolutions are given as the length of each side of a square pixel.

Type of Variable	Name	Data Source	Pixel Resolution
Response:			
	Species observation (Presence or Absence)	NPS (See Appendix C)	Varies
Burn-related Predictors:			
	Fire Occurrence (2 Classes) Burned / Unburned	NPS (Fire History)**	30 meter
	Time Since Fire	NPS (Fire History)**	30 meter
	Time Since Fire Squared	NPS (Fire History)**	30 meter
	Burn Severity (dNBR)	USGS / NPS**	30 meter
	Burn Severity (dNBR) Squared	USGS / NPS**	30 meter
Non-burn Related Predictors:			
Climatic	Annual Humidity	DayMet	1000 meter
	Annual Precipitation	DayMet	1000 meter
	Driest Month for Each Pixel	DayMet*	1000 meter
	Frost Degree Days	DayMet	1000 meter
	Precipitation Frequency	DayMet*	1000 meter
	Season Length	DayMet	1000 meter
	Summer Precipitation	DayMet*	1000 meter
	Temperature Minimum	DayMet*	1000 meter
Vegetative	Tree Cover Percent	USGS / NLCC	30 meter
	Vegetation Type (6 Classes)	NPS (Veg. Map)*	30 meter
	-High Elevation / Sparse	NPS (Veg. Map)*	30 meter
	-Hydrophilic	NPS (Veg. Map)*	30 meter
	-Lodgepole	NPS (Veg. Map)*	30 meter
	-Mixed Conifer	NPS (Veg. Map)*	30 meter

	-Deciduous	NPS (Veg. Map)*	30 meter
	-Sagebrush	NPS (Veg. Map)*	30 meter
Hydrologic	Euclidean Distance to Water	NPS (Hydrology)**	30 meter
Phenologic	Average Base Level	NASA	250 meter
	Large Integral	NASA	250 meter
	Brown Down Rate	NASA	250 meter
	Green Up Rate	NASA	250 meter
	Amplitude	NASA	250 meter
	Max Peak Date	NASA	250 meter
	Peak Date	NASA	250 meter
	Base Level	NASA	250 meter
	Brown Down	NASA	250 meter
	Green Up	NASA	250 meter
	Season Length	NASA	250 meter
Topographic	Elevation	USGS	30 meter
	Elevation Squared	USGS	30 meter
	Area Solar Radiation	USGS*	30 meter

**Table 2-B Final Burn Model Coefficients And Odds Ratios**
Environmental variables (burn and non-burn) are listed in order of decreasing significance. Estimates are given along with the log-transformed odds ratio. In the case of continuous variables, odds ratios are interpreted as a change in odds per unit change in the predictor variable. In the case of the vegetation (veg.) class variable, the odds of *C. nutans* presence is in comparison to the high elevation (elev.) veg. class. A cross (†) indicates a burn-related variable and the "time since fire" term is abbreviated (TSF). Burn severity is measured through the remotely-sensed differenced Normalized Burn Ratio (dNBR). An asterisk (*) indicates the variable is a remotely-sensed phenology index computed by NASA – each of which was created from multiple Normalized Differenced Vegetation Index (NDVI) values taken over the course of a year. A double asterisk (**) indicates the variable is an interaction term. Since the polynomial term "elevation squared" was significant (p< 0.10) when simple "elevation" was omitted from the model, both terms were retained, although p-values now appear non-significant. The "distance to water" variable represents the Euclidean distance to the closest open water such as a lake or stream. Non-significant vegetation classes and interaction terms were omitted from the final habitat suitability map calculations. Terms with p > 0.10 were used in map calculations but are not reliable for odds ratio interpretations.

Variable	Units	Estimate	Odds Ratio	SE	z Value	Pr(>|z|)
Precipitation Frequency	Days / month	-6.59E+01	2.52E-29	9.52E+00	-6.920	< 0.0001
Veg. Class: Sagebrush	Class	-1.56E+00	0.210767	2.92E-01	-5.330	< 0.0001
Distance to Water	Meters	-1.42E-03	0.998586	3.61E-04	-3.920	0.0001
Solar Radiation	Watt hours / m2	3.34E-06	1.000003	8.56E-07	3.900	0.0001
Peak Date*	Day of year	2.03E-01	1.224460	5.56E-02	3.650	0.0003
Veg. Class: Hydrophilic	Class	1.64E+00	5.165490	4.60E-01	3.570	0.0004
Annual Humidity	Pascal	-7.65E-02	0.926325	2.24E-02	-3.420	0.0006
TSF x Hydrophilic	None	-1.67E-02	0.983439	5.03E-03	-3.320	0.0009
Base Level*	NDVI	-4.08E-04	0.999592	1.27E-04	-3.210	0.0013
Average Base*	NDVI	-1.14E-03	0.998861	3.56E-04	-3.210	0.0013
TSF †	Years	7.65E-03	1.007675	2.50E-03	3.060	0.0022
Burn Severity †	dNBR	3.05E-03	1.003053	1.02E-03	3.000	0.0027
Veg. Class: 1 (Intercept)	None	6.00E+01	na	2.10E+01	2.860	0.0043
TSF x Lodgepole**	None	-1.33E-02	0.986818	5.36E-03	-2.480	0.0133
Summer Precipitation	Millimeters	-1.16E+00	0.315058	5.20E-01	-2.220	0.0263
TSF x Sagebrush**	None	7.96E-03	1.007994	3.59E-03	2.220	0.0267
Temperature Minimum	Celsius	1.05E+00	2.849091	5.92E-01	1.770	0.0767
Burn Severity Squared †	dNBR	-2.19E-06	0.999998	1.47E-06	-1.480	0.1379
Brown Down Rate	NDVI Units / Day	2.95E-04	1.000295	2.06E-04	1.430	0.1519
Tree Canopy Cover	Percent	3.35E-03	1.003352	2.48E-03	1.350	0.1780
Elevation Squared	Meters	-4.78E-06	0.999995	4.33E-06	-1.100	0.2699
TSF x Deciduous**	None	4.28E-03	1.004286	4.79E-03	0.892	0.3722
Elevation	Meters	1.47E-02	1.014798	1.92E-02	0.766	0.4437
Veg. Class: Deciduous	Class	-2.94E-01	0.744978	3.85E-01	-0.764	0.4447
Veg. Class: Mixed Conifer	Class	-4.48E-01	0.638905	6.68E-01	-0.671	0.5024
Veg. Class: Lodgepole	Class	2.59E-01	1.296023	4.59E-01	0.565	0.5724
TSF x Mixed Conifer**	Class	1.67E-03	1.001666	7.32E-03	0.228	0.8200

**Table 2-C Model Performance and Accuracy Measures for *Carduus nutans***

A model with burn information was compared to a model without burn information using *C. nutans* as the response and a wide range of predictor variables. Data was partitioned into a fitting set (80%) and validation set (20%) and each was tested separately. The area under the curve (AUC) of the Receiver Operating Curve (ROC) compares model predictions across all thresholds, where values range from -1 to +1 and any value above 0 is considered better than random chance. Prevalence is the ratio of presence to total sample size and is known to influence thresholded measures of accuracy. The "fair cutoff" threshold was chosen to provide a balanced assessment of sensitivity and specificity, since neither was consider more important than the other. The true skill statistic is similar to Cohen's Kappa, but is unaffected by prevalence. It ranges from -1 to +1 and values greater than zero are considered better predictions than by chance alone. Sensitivity is the ratio of true presence to total predicted presence. Specificity is the ratio of true absence to total predicted absence. Akiake's Information Criterion (AIC) allows model fit comparisons and adds a penalty ($2k$) for each model parameter. Deviance explained is analogous to R^2, and measures the percent of null deviance the model explains   The ANOVA is an analysis of deviance fit to a Chi-square distribution and tests the null hypothesis the burn model is the same as the no burn model (only tested using the fitting data).

| Species & Study Area	Dataset	Model	Sample Size	AUC Of ROC	Pre-valence	Fair Cutoff	True Skill	Class. Accuracy	Sen-sitivity	Spe-cificity	AIC	Deviance Explained	ANOVA: (P-value)
C. nutans in GRTE	Fitting	Burn	2160	0.864	0.38	0.430	0.559	0.779	0.780	0.779	1939	0.676	9.58E-09
		No Burn	2160	0.854	0.38	0.440	0.527	0.763	0.763	0.764	1976	0.658	
	Validation	Burn	536	0.853	0.41	0.423	0.545	0.772	0.775	0.770	n/a	n/a	
		No Burn	536	0.843	0.41	0.425	0.517	0.759	0.757	0.760	n/a	n/a	
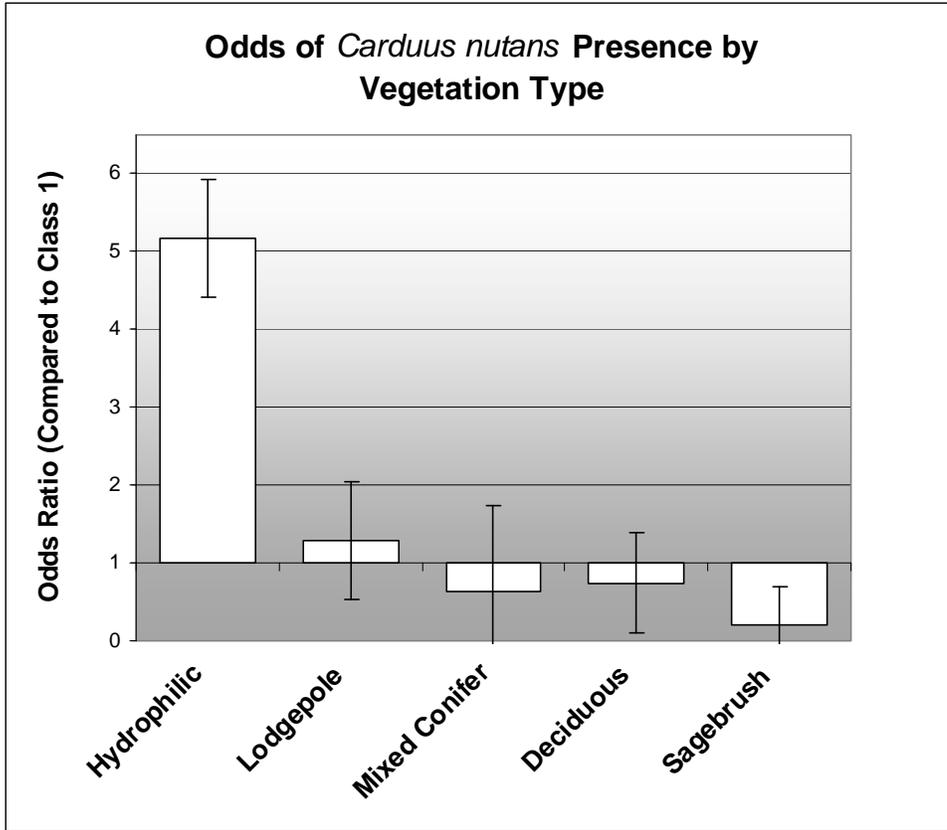
Figure 2-A Odds of Carduus nutans Presence by Vegetation Type
The odds of presence varies across different vegetation class in comparison to class 1 (an aggregate of high elevation and sparse vegetation types). The model holds all other variables constant, in order to compare the general invasibility of each vegetation class to the invader, *C. nutans*. Only hydrophilic vegetation & sagebrush are significantly different from class 1.
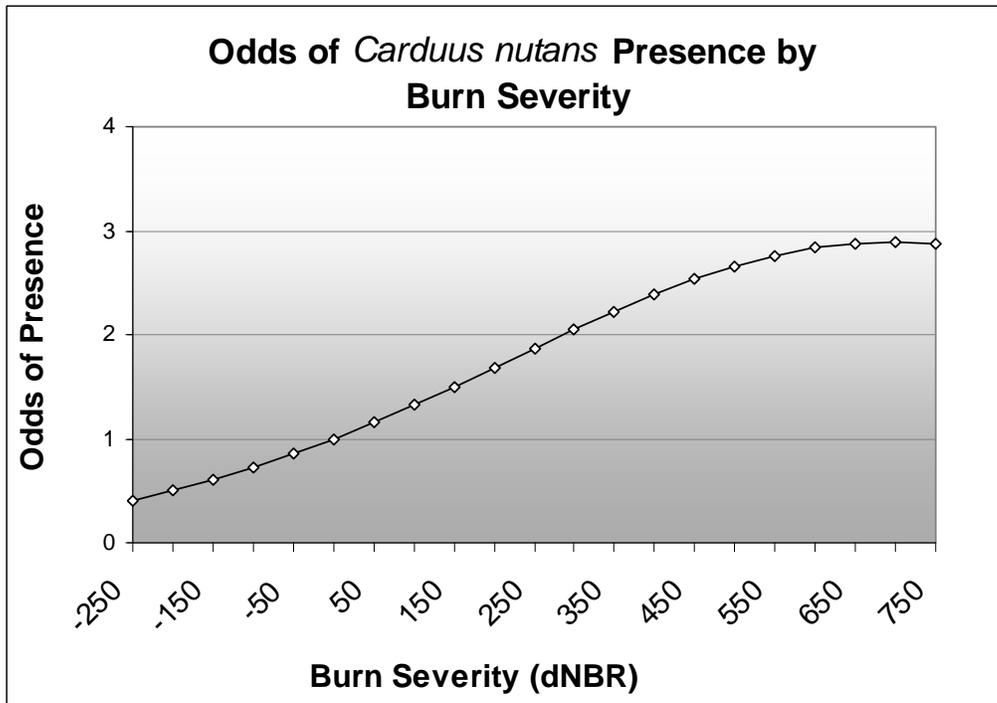
Figure 2-B Odds of Carduus nutans Presence by Burn Severity
The relative odds of *C. nutans* presence increases with burn severity (as measured through the remote sensing index, dNBR) when all other predictors are held constant. The commonly occurring range of dNBR across vegetation classes in GRTE is 100 to 500. Values from -250 to -100 indicate a post-fire "green-up" response, where vegetation becomes greener the year following a fire. Higher burn severity values (above 500) were rare across fire in GRTE and require further sampling for greater confidence.
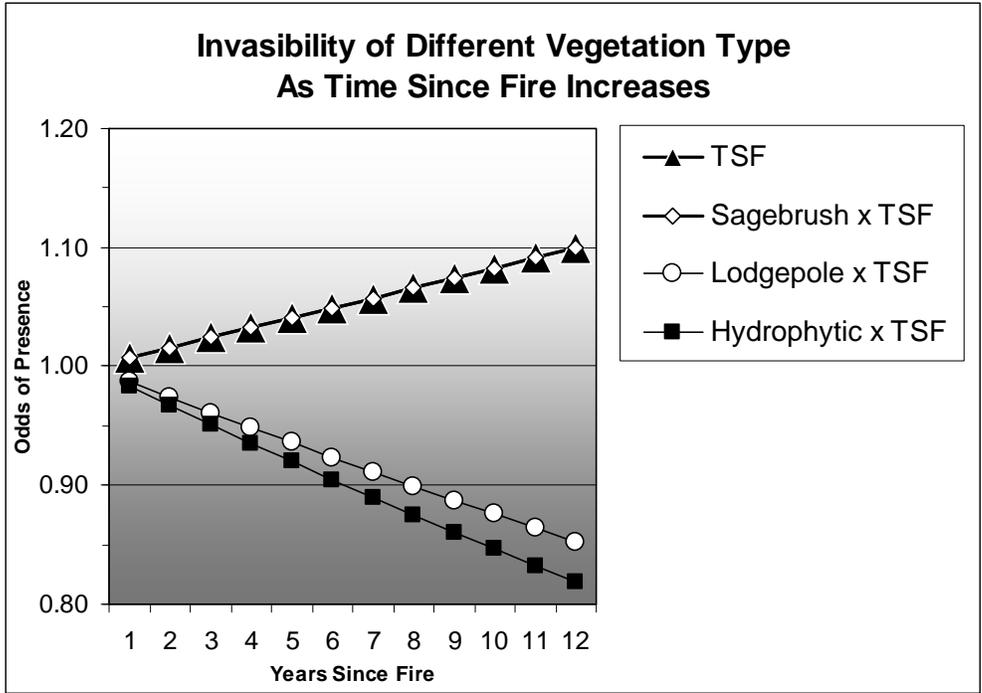
Figure 2-C of Carduus nutans Presence in Response to Vegetation & Time Since Fire
As time since fire (TSF) increases, the odds of *C. nutans* presence increases or decreases, depending on the vegetation class (interaction). The interaction coefficients for sagebrush, lodgepole and hydrophilic vegetation class are significant. The former shows an increased odds of presence over time, where as the latter two show a decreased odds of presence over time. Vegetation class 1, mixed conifer, and deciduous vegetation show an increasing odds of presence with time, as represented by the TSF line.
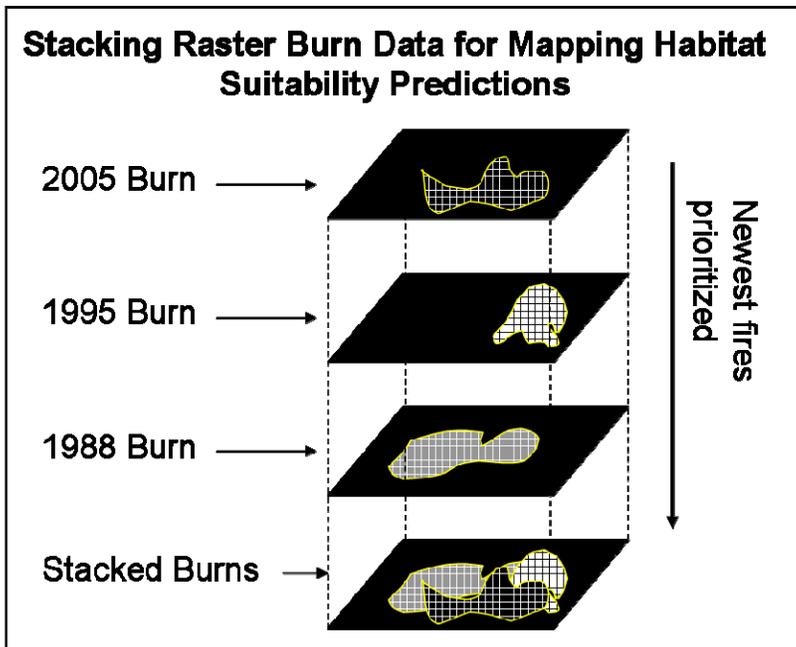
Figure 2-D   Sequencing Burn Data in a GIS for Use in Modeling
Burn GIS layers are "stacked" chronologically, with newest burns taking precedence over older burns in areas of overlap. This type of process allowed the extraction of correct chronologic burn data for model building as well as creating predictions for current post-fire conditions across the study area.
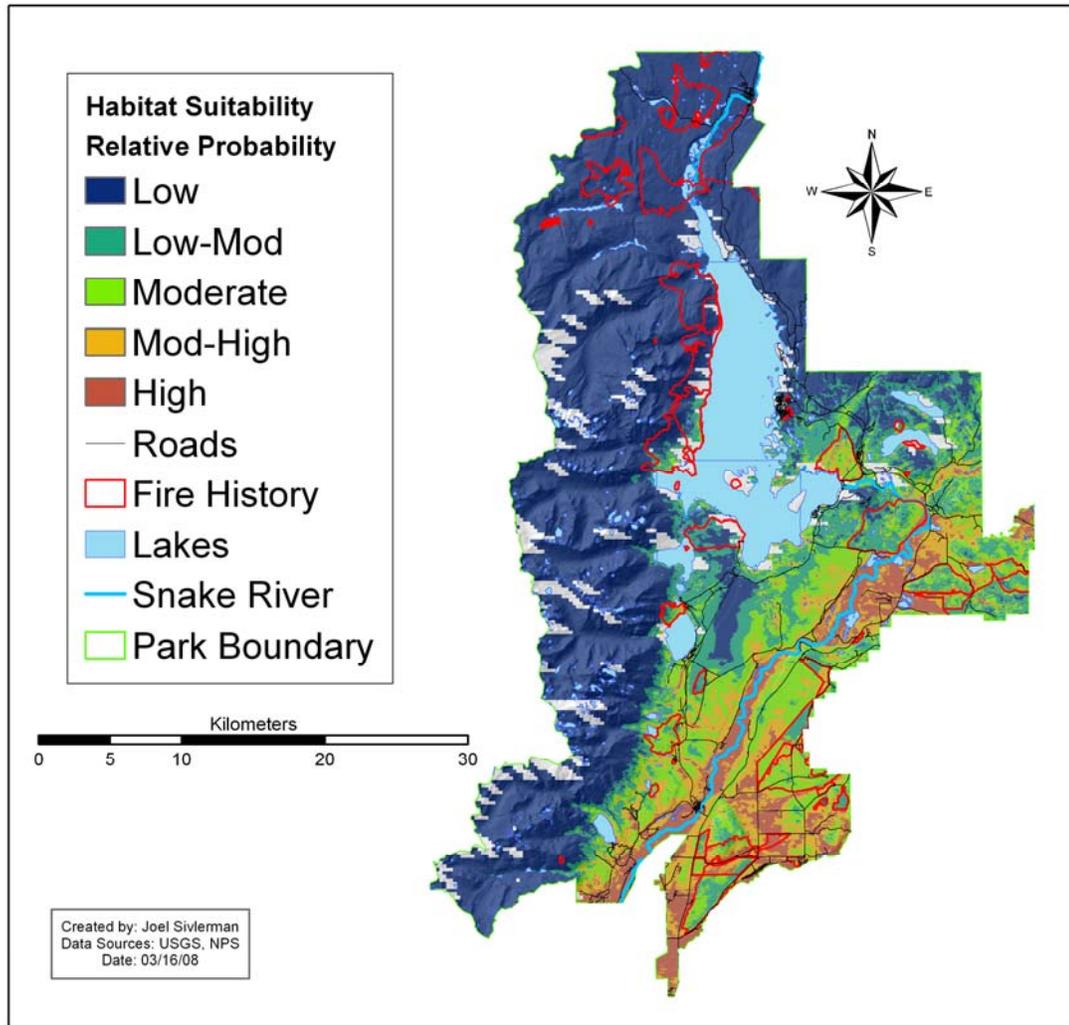
Figure 2-E Suitability Map for *Carduus nutans* in Grand Teton NP
The map of *C. nutans* habitat suitability primarily shows trends along the elevation gradient. The
mountains in the west have the lowest habitat suitability and the lowest areas (near the Snake River) have
the highest habitat suitability.
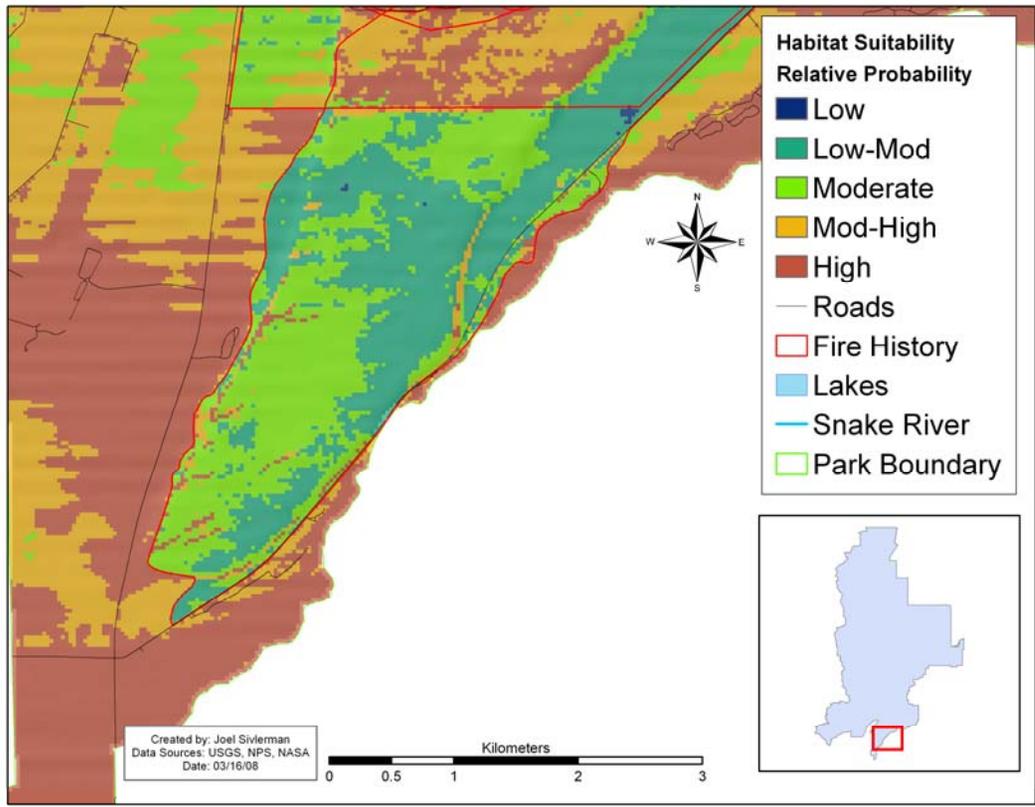
Figure 2-F Decreased Odds of *Carduus nutans* Presence within Antelope Flats Burn
The odds of presence noticeably decreased within the boundary of a moderately age burn in grass and sagebrush vegetation, when compared to its unburned surroundings.

Appendix A - GIS Analysis Documentation

Overview of Methods and Software

Preparation of data, stratified sampling design, map algebra, and data exploration were carried out using ArcGIS/ArcInfo version 9.2 (ESRI 2006). Many of the Spatial Analyst tools were utilized as well as Data Management tools. The Raster Calculator was used extensively because of the variety of operands available. One important macro from Hawth's Tools was used repeatedly to capture values from polygon and raster data in point file attribute fields (Beyer 2004). Some spatial-statistics analysis was performed in R statistical software (R 2007). The processes performed are described in this index through descriptions, diagrams, and flow charts. Although not fully automated, these processes should be repeatable using this appendix as a guide.

1. Ingestion From Original GIS Data Sources

Almost all data required some level of data management to be incorporated into this project. All data were set to the CONUS NAD 83 projection, and used 11 N or 12 N zones. All files were set to the extent of the study area of interest (Grand Teton NP, Sequoia & Kings Canyon NP, or Yellowstone NP). Some data remains in WGS 84, as supplied by NASA in order to preserve original files, but they were transformed to NAD 83, if used.

a. Burn Data Conversion, Calibration, Extraction, and "Stacking"

The Land-sat derived dNBR images were primarily downloaded from the USGS-

NPS Burn Severity Mapping Project (NPS-USGS 2007) website

http://burnseverity.cr.usgs.gov/ although many additional dNBR scenes were

made available through the help of the fire ecologists at Sequoia & Kings Canyon

and Grand Teton NP. First, images were checked for projection and transformed

to NAD 83 if necessary.

      i.   Calibration**:**

Calibration here refers to removing the background grey tone bias for

comparison between different dNBR images. This was calculated using

methods described by Key and Benson (2005). Uniform "grey pixels"

near the burn are sampled and a mean value is calculated for the sample.

To sample the grey areas, polygon shapefiles were created in Arc Catalog.

These were added to a layout in ArcMap, the editor was turned on, and the

shape of each polygon was drawn using the mouse. When at least 5000

pixels had been sampled, the Spatial Statistics Zonal Statistics wizard was

used from the drop down menu. This menu allows you calculate mean

dNBR value contained within the same x/y space as the sample polygons

just created. The mean was always between -75 and 75, a relatively small

adjustment considering the overall scale of values (approximately -1000 to

1350).
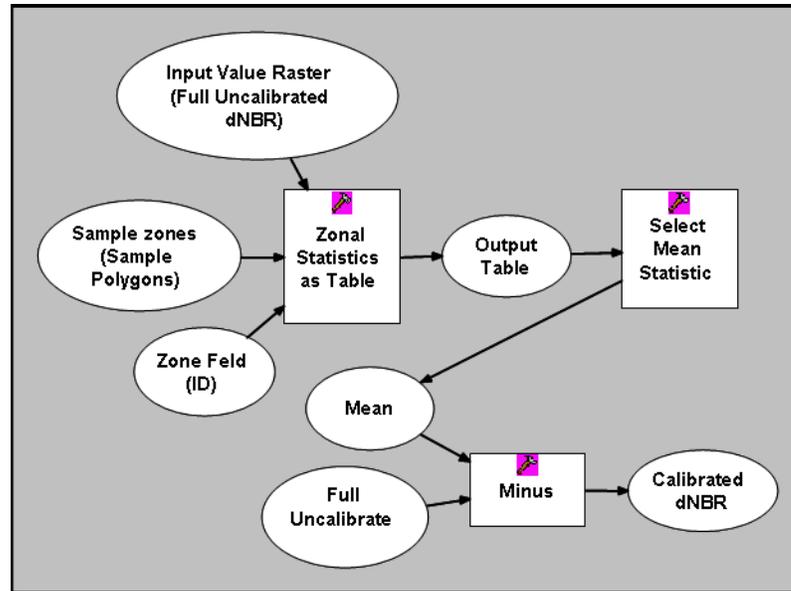
Figure 3-A Calibration for dNBR
The mean unburned value near the fire perimeter is subtracted from the burn values
within the fire perimeter to create equal background values between images. This allows
for comparison among dNBR images. The initial step of creating and drawing sampling
polygons is not included.

ii.   Extraction:

For each burn, the DNBR scene (a large rectangular area of dNBR values

downloaded) was extracted by the fire perimeter using the Spatial Analyst

tool "Extract by Mask". The mean found in the calibration above was

then subtracted from un-calibrated extracted values.   The un-calibrated,

clipped images were not generally saved because it was an intermediate

step, although the zonal statistics and sample polygons were saved. The

abbreviation of "cal" was used in each file name to designate it as a final

dNBR product for each fire.
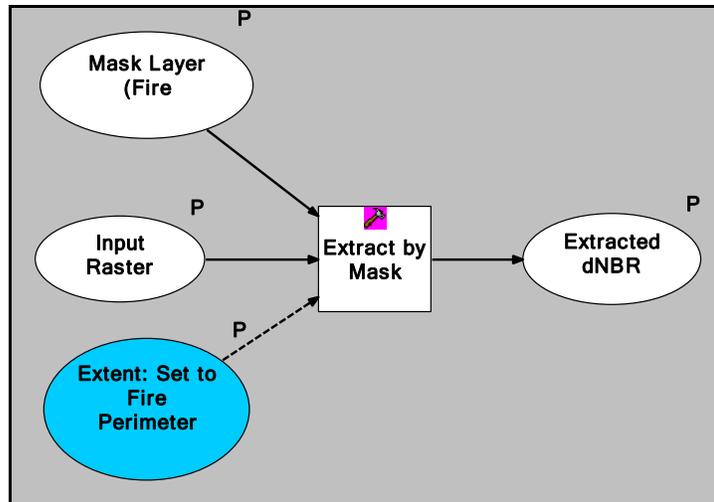
Figure 3-B Extract by Mask Analysis
The "Extract by Mask" tool retains only dNBR pixels within the fire perimeter. All area within the study extent which was not burned is given a dNBR value of 0 (no change).

iii.   Conversion of Fire History Polygons

Fire history shapefiles spatially describe all fires for most of the 20th century in each park (first year varies by park) and are relationally joined to a database of fire information. For most recent fires, perimeters are mapped a combination of GPS ground assessment and dNBR images. Older fires were digitized from physical maps. Vector perimeter data and the associated year of burn were transformed to 30 m resolution raster files. These were set to the extent of the burn and "snapped" to the study area extent to minimize registration error. Individual fires were merged first by year of fire (all 1990 fires, all 1991 fires, etc.).

iv.   Fire Inclusion Criteria and "Stacked" Data

Individual fires were named as mentioned above. Next the fires were merged by year, such that a dNBR layer of all fires for a given year exists.

94

No year layer was created if there was not at least one fire >20 acres in the park (1988 to most recent year available).

These dNBR year layers were then merged or "stacked," such that recent fires values were retained in cases of multi-year overlap. These stacked layers started with dNBR for 1988 and move forward to 2006. Finished "stacked" layers are named either starting with a "d" indicating dNBR (burn severity) or "b" indicating burn year. Each name includes the letters "pt" for "prior to" and is followed by a year number. For example, [dpt03] indicates the dNBR "stacked" for years *prior to* 2003 (1988 – 2002). This file would contain a prioritized merge of dNBR for 2002, 2001, 2000…1988). Note that [dpt03] *doesn't* contain dNBR for 2003. Fires less than 20 acres were not calibrated to conserve time, along with a few fires for which no appropriate dNBR was available. These individual stacked layers ([dpt03], [dpt04], etc.) were used in the model array construction to align of response observation with the appropriate burn history information (values from most recent prior burn).

b. Vegetation Class Coverage

The Grand Teton vegetation map is a relatively high accuracy coverage of vegetation types, with a number of hierarchical vegetation classes associated with the actual individual polygons. Many of these classifications were quite complex. In order to create a reasonable number of vegetation classes for statistical

analysis, the National Vegetation Class Associations were aggregated into six classes: high elevation/sparse, hydrophilic, lodgepole pine, mixed conifer, deciduous, and sagebrush (see Appendix B). The low number of classes allowed for testing first order interactions with burn variables in the statistical models. All polygons were converted to 30 m pixel resolution using the "center setting" for value choice. This was necessary for computational efficiency.

c. Distance to Lakes and Streams

The Euclidean (straight line) distance to each pixel was calculated based on hydrologic geo-databases for each park. All streams features were used to create a raster grid (30 m resolution) of distance values. The same tool was used on the lakes features for each park. Stream and lake distance raster files were combined using the minimum operand in the Raster Calculator. This operand takes input raster files set to the same spatial extent and finds the minimum value for each x/y location. These minimum values are the output raster. Thus, the closest open water is used in the final distance to water raster.
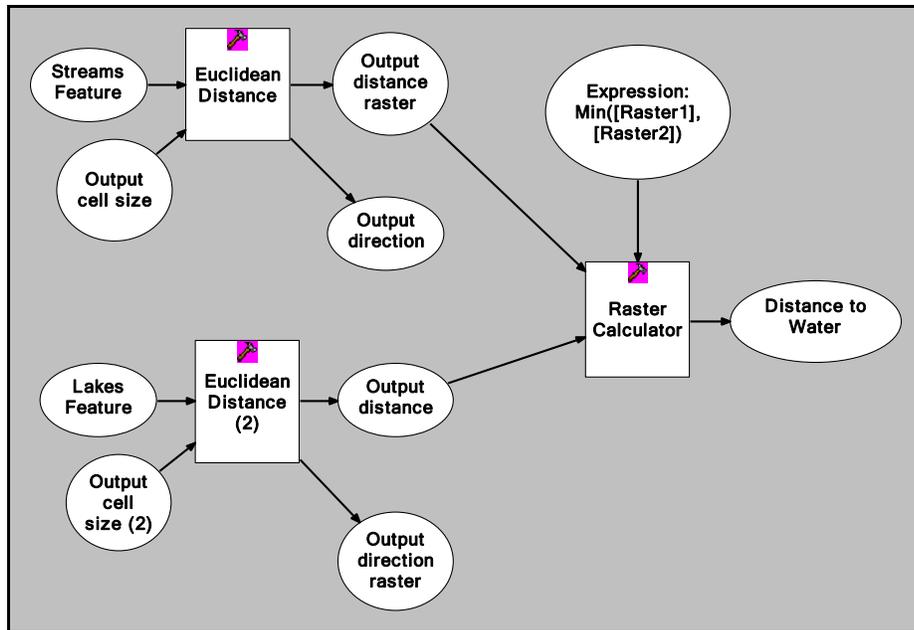
Figure 3-C Distance to Water Calculation
The Eulidean distance to water was calculated for each raster cell using a vector (polygon and polyline) input lake and river layers provided by NPS.

d. DayMet Meteorological Data

Available at: http://daymet.ntsg.umt.edu/data/RecordSum.htm

DayMet is an online data source for 1 km resolution climate surfaces generated from twenty years of observations collected and interpolated from weather stations (Thornton&Running 1999). These data layers were reprojected into NAD 83 and resampled to 30 m resolution for analysis. Some layers were converted using the Raster Calculator and local map algebra operands such as "min," "max," and "minus".

Annual Humidity: Unchanged from DayMet; average annual humidity for each cell.

Annual Precipitation: Unchanged from DayMet: Total annual precipitation for each cell.

97

Driest Month: June through August monthly total precipitation grids were analyzed using the "min" operand in Raster Calculator to determine the minimum value for each x/y coordinate for the output grid.

Frost Degree Days: Unchanged from DayMet; sum of the mean temperatures for all days where the average temperature was 32 degrees or colder.

Precipitation Frequency: Precipitation frequency is measured as the fraction of days where precipitation = 0. This layer was calculated using the "min" operand on precipitation frequency data for the months of June through August.

Temperature Minimum: This layer gives the average monthly temperature of the coldest month within each pixel, using the "min" operand.

Temperature Seasonality: This layer represents the difference in temperature between the warmest month (mean temperature) and coldest (mean temperature) month within each pixel. First a maximum average monthly temperature layer was created along with a minimum average monthly temperature layer. These two layers were subtracted to find the final value within each pixel.

e. Digital Elevation Model (DEM)

Available at: http://seamless.usgs.gov/ or http://www.nps.gov/gis/data_info/

A 30 m resolution DEM file for each park was downloaded from NPS data clearinghouse.

f. Incident Solar Radiation (ISR)

ISR was calculated using the solar radiation tool in Arc GIS 9.2 Spatial Analyst tool box. Total annual radiation was calculated using 16 aspects, a uniform clear sky, and 14 day sun angle changes. See Arc Help files or the process metadata attached to this layer for more details (ESRI 2006).

g. The NLCD Forest Percent Canopy Cover

The National Land Cover Data was used in both Chapters. In GRTE, 250m resolution data was utilized (delivered from NASA partners). In the three-park comparison, a 30 m resolution version was downloaded directly from USGS data center at: http://seamless.usgs.gov/.

h. NASA Remote Sensing Phenological Indices

The NASA Phenological Indices are described in both chapters, and the following figure and extended legend was modified from Jonsson & Eklundh (2004) and provided courtesy of our NASA-NPS project (Benson et al. 2005) partners at Goddard Space Flight Center.
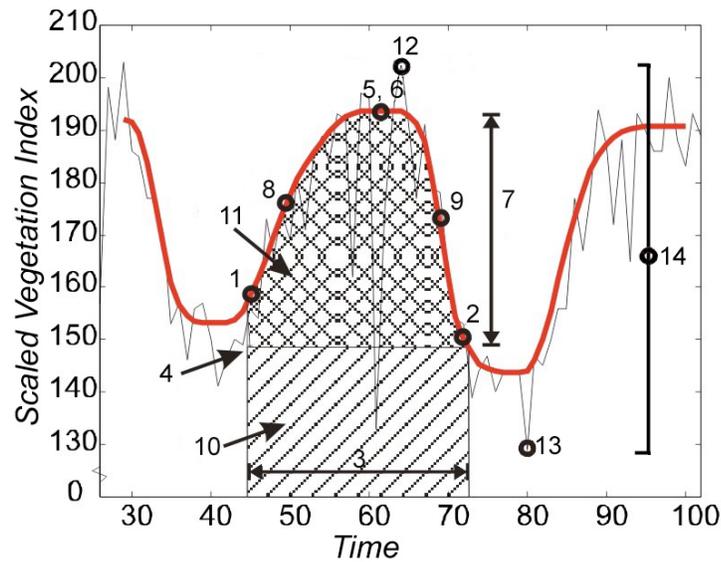
Figure 3-D The TIMESAT Indices
TIMESAT is a remotely-sensed index of seasonal change in greenness fit to a sign wave. Each number corresponds to a different measure of change.

The original TIMESAT algorithm provides 11 phenological parameters for each successfully fitted pixel:

> 1. **Green Up (time for the start of the season)**; time for which the left edge has increased to a user defined level (10% of the seasonal amplitude used here) measured from the left minimum level.
>
> 2. **Brown Down (time for the end of the season)**; time for which the right edge has decreased to a user defined level measured from the right minimum level.
>
> 3. **Season Length;** time from the start to the end of the season.
>
> 4. **Average base level**; given as the average of the left and right minimum values.

5. **Peak Date;** computed as the mean value of the times for which, respectively, the left edge has increased to the 90 % level and the right edge has decreased to the 90 % level.

6. **Maximum Peak;** largest data value for the fitted function during the season;

7. **Amplitude (seasonal);** difference between the maximal value and the base level.

8. **Green Up Rate;** rate of increase at the beginning of the season; calculated as the ratio between the values evaluated at the season start and at the left 90 % level divided by the corresponding time difference.

9. **Brown Down Rate;** rate of decrease at the end of the season; calculated as the ratio between the values evaluated at the season end and at the right 90 % level divided by the corresponding time difference.

10. **Large seasonal integral**; integral of the function describing the season from the season start to the season end.

11. **Base level (small seasonal integral)**; integral of the difference between the function describing the season and the base level from season start to season end.

2. **Chronologically Merging Burn Information Raster Files**

Burn Severity and TSF raster layers were merged in two ways, each for a separate purpose. The first goal was to temporally align species observation year with the last know prior fire data. In some cases, the most recent fire occurred after the species observation. Thus a *series* of "stacked" layers was created for this purpose (e.g. all dNBR prior to 2003, all dNBR prior to 2004, etc.). The values from each yearly "stack" was added as a field in the attribute field of the observation point files described in the array construction topic below

The second goal for merged files was to create a single layer for use in calculating the final habitat suitability map using the most recent burn values at each location within the study area (one dNBR and one TSF layer per study area). Hence, these merged burn layers describe yearly burn information in a chronological manner, omitting older values in cases of spatial overlap. Unburned cells within the study areas were given a placeholder value of 0 for dNBR (no change due to fire) and 100 for TSF (assumes at least 100 years have past since the last fire).

3. The Response Variable

The response variable for each species is a geo-referenced observation of presence or absence for the given species. In general, observations needed to meet several criteria: 1) all absence observations need to make a complete, consistent, reasonable search of an actual "plot" area, 2) all "plot" areas needed to be at least 300 m^2 in area, 3) observations needed to be geo-referenced to an acceptable accuracy (+/-) 15 m (Horizontal Error), 4) all observations needed to include the year of observation, 5) observations needed to be

about 30 m apart. Plots smaller than 300m^2 this would not accurately represent the resolution of the environmental layers. Plots much larger than this would require too much time to survey, ultimately leading to very small sample sizes. Data sources varied and are described in more detail below, as well as in the ArcGIS point shapefiles: [yell_lida.shp], [grte_canu.shp], and [seki_civu.shp].

4. Response Variable Data Sources

The data sources varied by park, but mostly represent park inventories of different sorts: fire effects, vegetation mapping, and invasive plant management weed mapping (presence only). In 2007, we conducted additional field sampling in Grand Teton. All observation data were aggregated by species using the vector "merge" tool in the Data Management toolbox.

In YELL, most observations were collected from an earlier study of several invasive species in the northern range of the park (Rew et al. 2005). These data are a public data source made available through Ann Rodman, an NPS manager at YELL. Unlike many sources, these data were collected along transects where the observer stopped every ten meters and recorded an observation. In discussion with park managers, we accepted this as an implied plot area of 314 m^2 (radius 10 m). Because these observations were closer together than even our finest-scale predictor variables (30 m^2), we needed a method to aggregate observations. This was also necessary because of the overwhelming total number of observations. First, a unique address raster was created for the study area extent (each pixel has a unique number) using a method described by (Theobald 2005).

Hawth's Tool "Intersect Point" tool was used to extract the unique pixel address for all

observations in the point shapefile of observations. Points were then aggregated using

the "Dissolve" tool in the Arc Data Management Toolbox. If any observation in the pixel

was a presence, then all observations in the pixel were considered presence (using the

statistics maximum option). Observations were further aggregated within pixel to a

single center point using the "Mean Center" tool in the Arc Spatial Analyst Toolbox.

This created one point per pixel. A "join" operation was then employed using the unique

address as the join field in both the "dissolve" observation shapefile and the new

aggregated set. This method allowed the transfer of the presence/absence information to

the final aggregated observation point shapefile (see Figure 1).
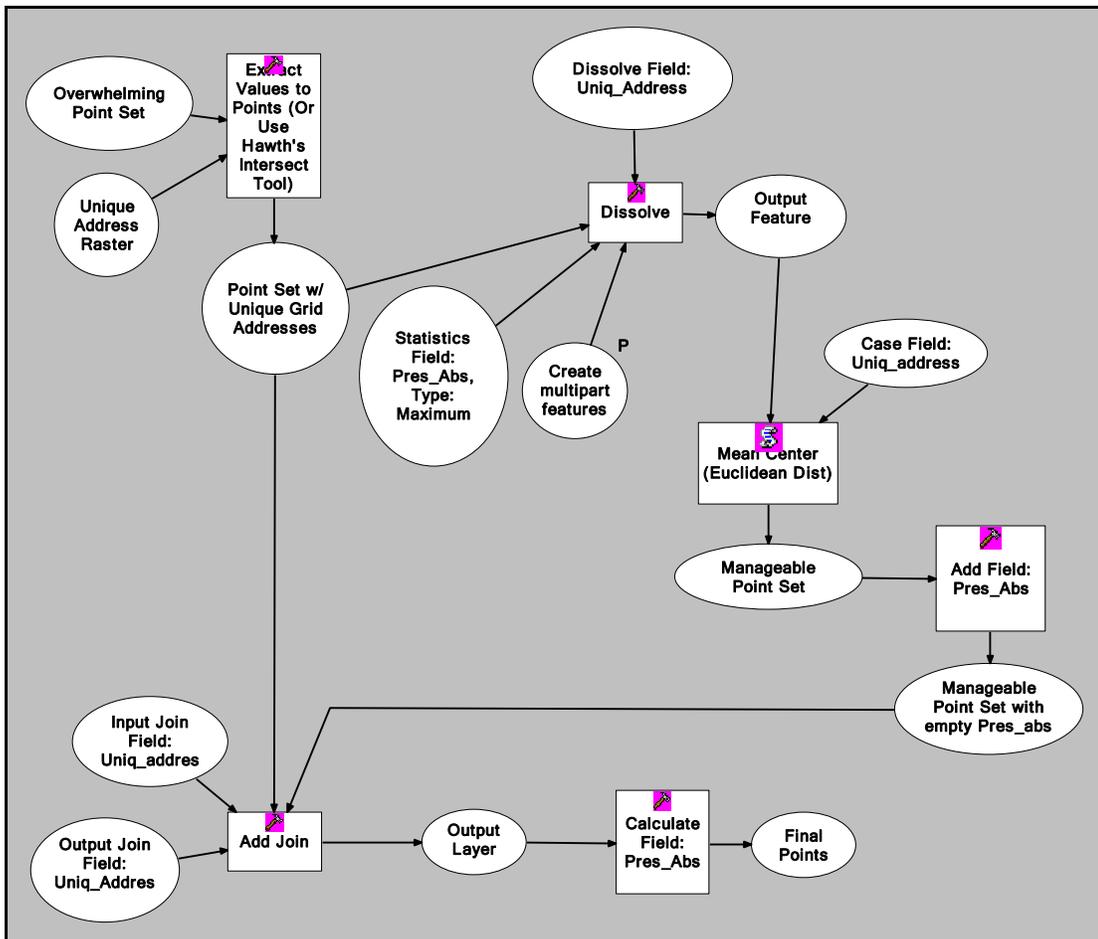
Figure 3-E Yellowstone observation dissolve and aggregation operations

104

5. Array Construction

Stacked dNBR and TSF values were extracted to the point observation files using Hawth's point intersect tool (Beyer 2004). Next a field called most recent burn year was created and calculated using the Field Calculator (advanced Visual Basic setting). This is the difference in years between the species observation and the most recent prior burn year. An example of the conditional statement used is given below:

Most recent burn year for *C. nutans*:

IF ([obs_date] > [bpt99]) AND ([bpt99] <> 0) THEN
 [mrb_yr] = [bpt99]
END IF

IF ([obs_date] > [bpt00]) AND ([bpt00] <> 0) THEN
 [mrb_yr] = [bpt00]
END IF

IF ([obs_date] > [bpt01]) AND ([bpt01] <> 0) THEN
 [mrb_yr] = [bpt01]
END IF

IF ([obs_date] > [bpt02]) AND ([bpt02] <> 0) THEN
 [mrb_yr] = [bpt02]
END IF

IF ([obs_date] > [bpt03]) AND ([bpt03] <> 0) THEN
 [mrb_yr] = [bpt03]
END IF

IF ([obs_date] > [bpt06]) AND ([bpt06] <> 0) THEN
 [mrb_yr] = [bpt06]
END IF

IF ([obs_date] > [bpt07]) AND ([bpt07] <> 0) THEN
 [mrb_yr] = [bpt07]
END IF

Once "most recent burn year" was calculated, TSF could be calculated a simple subtraction from the year of species observation. The most recent dNBR prior to observation was calculated in much the same way. Once burn variables were included in the observation point file, other variables were added using Hawth's intersect tool. The shapefile was then converted to a simple comma separated values (CSV) file and imported into R statistical software for regression and spatial auto-correlation analysis.

6. Habitat Suitability Maps

The final habitat suitability maps were calculated using layers described above and the coefficients computed by the regression in R. The calculation followed the logistic equation:

$$p_i = \frac{\exp(\beta_0 + \beta x_{ki})}{1 + \exp(\beta_0 + \beta x_{ki})}$$

This was performed in the Raster Calculator, as in this example for *L. dalmatica*:

numerator = exp((-28.1000000000) +~
[burned_0_1] * (-0.5495000000) +~
[mr_dnbr0] * (0.0028570000) +~
[mr_dnbr0] * [mr_dnbr0] * (-0.0000058100) +~
[gya_dem30m_masked.img] * (0.0310400000) +~
[gya_dem30m_masked.img] * [gya_dem30m_masked.img] * (-0.0000085600) +~
[area_solar_radiation_yell.img] * (0.0000015590) +~
[tree_cov30m_83_mask] * (-0.0082770000) +~
[n_max_peak_r.img] * (-0.0001266000) +~
[n_an_base_lv_r.img] * (0.0002891000) +~
[n_ssn_lgnth_r.img] * (0.0019610000) +~
[n_grnup.img] * (-0.0123700000))

hs_map = [numerator] / (1 + [numerator])

7. Additional Metadata

Metadata is developed and attached (.xml files) to finished GIS products such as habitat maps and observation point shapefiles. The entire set of GIS data will be made available for NPS, NASA, and other project partners.

Appendix B - Vegetation Classes Aggregation

Class 1 – High elevation, Sparse, Other

Subalpine Fir - Engelmann Spruce Forest

Alder Shrubland

Alpine Herbaceous Vegetation

Alpine Mesic Meadows

Rock Outcrop / Cliff

Ceanothus Shrubland

Cliff and Talus Sparse Vegetation

Exposed Hillside Sparse Vegetation

Herbaceous Aquatics

Rocky Mountain Juniper Woodland Stand

Krummholtz Woodland

Limestone Pavement Sparse Vegetation

Mixed Subalpine Fir - Engelmann Spruce Woodland - Deciduous Shrubland Regeneration

Mixed Deciduous Shrubland

Mixed Grassland Herbaceous Vegetation

Mixed Tall Deciduous Shrubland

Residential and Facilities

Montane Mesic Forb Herbaceous Vegetation

Montane Xeric Forb Herbaceous Vegetation

Glacier / Snow

Whitebark Pine Forest

Limber Pine Forest

Douglas-fir Forest

Bracken Fern Herbaceous Vegetation

Recently Burned Sparse Vegetation

Artic Willow Dwarf Shrubland

Strip Mines, Quarries, and Gravel Pits

Subalpine Mixed Herbaceous Vegetation

Roads and Trails / Trailheads

Class 2 – Hydrophilic

Irrigation Canals

Irrigated Fields

Exposed Lake Shoreline - Stream Deposit Sparse Vegetation

Flooded Wet Meadow Herbaceous Vegetation

Lakes and Reservoirs

Blue Spruce Riparian Forest

Willow Shrubland

Non-vegetated Sand Bars

Streams

Class 3 – Lodgepole Pine

Lodgepole Pine - Ceanothus Woodland Regeneration

Lodgepole Pine Forest

Lodgepole Pine Woodland Regeneration

Class 4 – Mixed Conifer

Mixed Conifer - Cottonwood Riparian Forest

Mixed Conifer Forest

Mixed Conifer Woodland Regeneration

Mixed Evergreen - Poplar Forest

Class 5 - Apen

Cottonwood Riparian Forest

Aspen Forest

Aspen Woodland Regeneration

Class 6 - Sagebrush

Low Sagebrush Dwarf Shrubland

Sagebrush - Antelope bitterbrush Mixed Shrubland

Sagebrush / Shrubby Cinquefoil Mesic Shrubland

Sagebrush Dry Shrubland

Appendix C - Dependent Variable Data Sources

Sequoia and Kings Canyon NP

1. Source: Kern Valley NPS non-native plant survey (McGinnis 2004)

    Type: presence and absence; random and non-random points

    Plot size: Varied; polygons converted to centroids; minimum absence 314 m^2.

2. Source: NPS non-native plant database

    Type: presence and absence; non-random points

    Plot size: Varied; polygons converted to centroids; minimum absence 314 m^2.

3. Source: USGS Survey (Keeley&McGinnis 2005)

    Type: presence; random points

    Plot size: 25 m^2

Grand Teton NP

1. Source: GRTE Vegetation Mapping

    Type: presence and absence; random stratified and non-random points

    Plot size: Varied; minimum absence 314 m^2

2. Source: GRTE Fire Effects and Composite Burn Index Plots

    Type: absence; random stratified points

    Plot size: 706 or 1000 m^2

3. Source: Thesis field work

    Type: presence and absence; random stratified and non-random points

    Plot size: 314 m^2

Yellowstone NP

1. Source: Rew and Maxwell Research

   Type: presence and absence; random stratified transects (starting from roads)

   Plot size: 400 m^2

2. Source: K. Myer, Bechler, G. Crain, McClure, & others

   Type: presence and absence; non-random points

   Plot size: minimum absence 400 m^2

3. Source: YELL Fire Effects Plots

   Type: absence; random stratified points

   Plot size: 1000 m^2

Bibliography

Abella, S. R., W. W. Covington, P. Z. Fule, L. B. Lentile, A. J. S. Meador & P. Morgan (2007). Past, present, and future old growth in frequent-fire conifer forests of the Western United States. *Ecology And Society* 12.

Agee, James K. (1993). *Fire ecology of Pacific Northwest forests,* Washington, D.C., Island Press.

Allouche, O., A. Tsoar & R. Kadmon (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal Of Applied Ecology* 43, 1223-1232.

Araujo, MB & A Guisan (2006). Five (or so) challenges for species distribution modelling. *Journal Of Biogeography* 33, 1677-1688.

Bailey, J. K. & T. G. Whitham (2002). Interactions among fire, aspen, and elk affect insect diversity: Reversal of a community response. *Ecology* 83, 1701-1712.

Baker, W. L. (1992). Effects Of Settlement And Fire Suppression On Landscape Structure. *Ecology* 73, 1879-1887.

Barnett, D. T., T. J. Stohlgren, C. S. Jarnevich, G. W. Chong, J. A. Ericson, T. R. Davern & S. E. Simonson (2007). The art and science of weed mapping. *Environmental Monitoring And Assessment* 132, 235-252.

Benson, Nate, Jeff Morisette & Brad Welch (2005). Using NASA's Invasive Species Forecasting System to support National Park Service decisions on fire management activities and invasive plant species control).

Beyer, H L (2004). Hawth's Analysis Tools for ArcGIS, version 3.26, [Online]. Available: http://www.spatialecology.com/htools. Accessed: June 2007.

Bond, W. J. & J. E. Keeley (2005). Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems. *Trends In Ecology & Evolution* 20, 387-394.

Brewer, C. K., J. C. Winne, R. L. Redmond, D. W. Opitz & M. V. Mangrich (2005). Classifying and mapping wildfire severity: A comparison of methods. *Photogrammetric Engineering And Remote Sensing* 71, 1311-1320.

Brooks, M. L., C. M. D'Antonio, D. M. Richardson, J. B. Grace, J. E. Keeley, J. M. DiTomaso, R. J. Hobbs, M. Pellant & D. Pyke (2004). Effects of invasive alien plants on fire regimes. *Bioscience* 54, 677-688.

Brown, R. T., J. K. Agee & J. F. Franklin (2004). Forest restoration and fire: Principles in the context of place. *Conservation Biology* 18, 903-912.

Caprio, Anthony C & Pat Lineback (2002). Pre-Twentieth Century Fire History of Sequoia and Kings Canyon National Parks: A Review and Evaluation of Our Knowledge. *Association for Fire Ecology Misc. Pub. No. 1*, 180-199.

Chabrerie, O., K. Verheyen, R. Saguez & G. Decocq (2008). Disentangling relationships between habitat conditions, disturbance history, plant diversity, and American black cherry (Prunus serotina Ehrh.) invasion in a European temperate forest. *Diversity And Distributions* 14, 204-212.

Chapin, F. S., E. S. Zavaleta, V. T. Eviner, R. L. Naylor, P. M. Vitousek, H. L. Reynolds, D. U. Hooper, S. Lavorel, O. E. Sala, S. E. Hobbie, M. C. Mack & S. Diaz (2000). Consequences of changing biodiversity. *Nature* 405, 234-242.

Chatterjee, Samprit (2000). *Regression analysis by example,* New York, Wiley.

Cogan, Dan, Klara Varga, Gwen Kittel, Kelly McCloskey, Jennifer Gremer, Diane Abendroth & Cory Bolen (2005). Grand Teton National Park Vegetation Mapping Project Final Report). Bureau of Reclamation.

Crossman, N. D. & D. A. Bass (2008). Application of common predictive habitat techniques for post-border weed risk management. *Diversity And Distributions* 14, 213-224.

Dale, V. H., L. A. Joyce, S. McNulty, R. P. Neilson, M. P. Ayres, M. D. Flannigan, P. J. Hanson, L. C. Irland, A. E. Lugo, C. J. Peterson, D. Simberloff, F. J. Swanson, B. J. Stocks & B. M. Wotton (2001). Climate change and forest disturbances. *Bioscience* 51, 723-734.

Dantonio, C. M. (1993). Mechanisms Controlling Invasion Of Coastal Plant-Communities By The Alien Succulent Carpobrotus-Edulis. *Ecology* 74, 83-95.

Dantonio, C. M. & P. M. Vitousek (1992). Biological Invasions By Exotic Grasses, The Grass Fire Cycle, And Global Change. *Annual Review Of Ecology And Systematics* 23, 63-87.

Davis, M. A., J. P. Grime & K. Thompson (2000). Fluctuating resources in plant communities: a general theory of invasibility. *Journal Of Ecology* 88, 528-534.

DeBano, Leonard F. (1998). *Fire's effects on ecosystems,* New York, J. Wiley.

Dellasala, D. A., J. E. Williams, C. D. Williams & J. E. Franklin (2004). Beyond smoke and mirrors: a synthesis of fire policy and science. *Conservation Biology* 18, 976-986.

ESRI (2006) *ArcGIS 9.2*. Redlands, California.

Fielding, A. H. & J. F. Bell (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24, 38-49.

Floyd, M. L., D. Hanna, W. H. Romme & T. E. Crews (2006). Predicting and mitigating weed invasions to restore natural post-fire succession in Mesa Verde National Park, Colorado, USA. *International Journal Of Wildland Fire* 15, 247-259.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing Of Environment* 80, 185-201.

Fox, J. & G. Monette (1992). Generalized Collinearity Diagnostics. *Journal Of The American Statistical Association* 87, 178-183.

Fox, John (1997). *Applied regression analysis, linear models, and related methods / John Fox,* Thousand Oaks, CA, Sage Publications.

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

Gerlach, John D, Peggy E Moore, Brent Johnson, D. Graham Roy, Patrick Whitmarsh, Daniel Lubin, David Graber, Sylvia Haultain, Anne Pfaff & Jon Keeley (2001). Exotic species threat assessment in Sequoia, Kings Canyon, and Yosemite national parks. *Crossing Boundaries in Park Management: Proceedings of the 11th Conference on Research and Resource Management in*

Parks and on Public Lands. (ed. by D. Harmon), pp. 96-103. The George Wright Society, Inc., Hancock, Michigan.

Guisan, A. & W. Thuiller (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8, 993-1009.

Guisan, A. & N. E. Zimmermann (2000). Predictive habitat distribution models in ecology. *Ecological Modelling* 135, 147-186.

Gumpertz, M. L., C. Wu & J. M. Pye (2000). Logistic regression for southern pine beetle outbreaks with spatial and temporal autocorrelation. *Forest Science* 46, 95-107.

Harris, J. A., R. J. Hobbs, E. Higgs & J. Aronson (2006). Ecological restoration and global climate change. *Restoration Ecology* 14, 170-176.

Higgins, S. I., D. M. Richardson, R. M. Cowling & T. H. Trinder-Smith (1999). Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. *Conservation Biology* 13, 303-313.

Hunter, M. E., P. N. Omi, E. J. Martinson & G. W. Chong (2006). Establishment of non-native plant species after wildfires: effects of fuel treatments, abiotic and biotic factors, and post-fire grass seeding treatments. *International Journal Of Wildland Fire* 15, 271-281.

Jones, Chad & Charles Halpern (2005). Modeling Distribution of High-Priority Exotic Plant Species on the Olympic Peninsula. (ed. by N. P. S. I. a. M. Program). NPS.

Jones, Chad & Charles Halpern (2007). Modeling Distribution of High-Priority Exotic Plant Species on the Olympic Peninsula). National Park Service, Department of Interior.

Jonsson, P. & L. Eklundh (2004). TIMESAT - a program for analyzing time-series of satellite sensor data. *Computers & Geosciences* 30, 833-845.

Keeley, J. E. (2006). Fire management impacts on invasive plants in the western United States. *Conservation Biology* 20, 375-384.

Keeley, J. E. & Thomas W McGinnis (2005). Impact Of Fire And Other Factors On Cheatgrass

Persistence In A Sierra Nevada Ponderosa Pine

Forest. (ed. by USGS), p. 104. Western Ecological Research Center.

Key, Carl H & Nathan C Benson (2005). Landscape Assessment: Sampling and Analysis Methods. (ed. by USDA), p. 55. USDA Forest Service.

Kneitel, J. M. & D. Perrault (2006). Disturbance-induced changes in community composition increase species invasion success. *Community Ecology* 7, 245-252.

Lentile, L. B., Z. A. Holden, A. M. S. Smith, M. J. Falkowski, A. T. Hudak, P. Morgan, S. A. Lewis, P. E. Gessler & N. C. Benson (2006). Remote sensing techniques to assess active fire characteristics and post-fire effects. *International Journal Of Wildland Fire* 15, 319-345.

Levin, S. A. (1992). The Problem Of Pattern And Scale In Ecology. *Ecology* 73, 1943-1967.

Levine, J. M. (2000). Species diversity and biological invasions: Relating local process to community pattern. *Science* 288, 852-854.

Levine, J. M., M. Vila, C. M. D'Antonio, J. S. Dukes, K. Grigulis & S. Lavorel (2003). Mechanisms underlying the impacts of exotic plant invasions. *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 270, 775-781.

Lugo, A. E., S. L. Brown, R. Dodson, T. S. Smith & H. H. Shugart (1999). The Holdridge life zones of the conterminous United States in relation to ecosystem mapping. *Journal Of Biogeography* 26, 1025-1038.

Lynch, H. J., R. A. Renkin, R. L. Crabtree & P. R. Moorcroft (2006). The influence of previous mountain pine beetle (Dendroctonus ponderosae) activity on the 1988 yellowstone fires. *Ecosystems* 9, 1318-1327.

Mack, M. C. & C. M. D'Antonio (1998). Impacts of biological invasions on disturbance regimes. *Trends In Ecology & Evolution* 13, 195-198.

Mack, R. N., D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout & F. A. Bazzaz (2000). Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications* 10, 689-710.

McGinnis, Thomas W (2004). Kern Watershed Weed Inventory of 2004. (ed. by USGS), p. 12.

McNab, W. H. & D. L. Loftis (2002). Probability of occurrence and habitat features for oriental bittersweet in an oak forest in the southern Appalachian mountains, USA. *Forest Ecology And Management* 155, 45-54.

Millar, C. I., N. L. Stephenson & S. L. Stephens (2007). Climate change and forests of the future: Managing in the face of uncertainty. *Ecological Applications* 17, 2145-2151.

Miller, Jennifer, Janet Franklin & Richard Aspinall (2007). Incorporating Spatial Dependance in Predictive Vegetation Models. *Ecological Modelling* 202, 225-242.

Morisette, J. T., C. S. Jarnevich, A. Ullah, W. J. Cai, J. A. Pedelty, J. E. Gentle, T. J. Stohlgren & J. L. Schnase (2006). A tamarisk habitat suitability map for the continental United States. *Frontiers In Ecology And The Environment* 4, 11-17.

Nielsen, C., P. Hartvig & J. Kollmann (2008). Predicting the distribution of the invasive alien Heracleum mantegazzianum at two different spatial scales. *Diversity And Distributions* 14, 307-317.

NPS-USGS (2007). National Burn Severity Mapping Project, [Online]. USGS EROS Data Center. Available: http://burnseverity.cr.usgs.gov/fire_main.php. Accessed: 2007.

NPS (2008). Data Clearinghouse, [Online]. NPS. Available: http://www.nps.gov/gis/data_info/. Accessed: 2007.

Pausas, J. G., R. A. Bradstock, D. A. Keith & J. E. Keeley (2004). Plant functional traits in relation to fire in crown-fire ecosystems. *Ecology* 85, 1085-1100.

Pausas, J. G., J. E. Keeley & M. Verdu (2006a). Inferring differential evolutionary processes of plant persistence traits in Northern Hemisphere Mediterranean fire-prone ecosystems. *Journal Of Ecology* 94, 31-39.

Pausas, J. G., F. Lloret & M. Vila (2006b). Simulating the effects of different disturbance regimes on Cortaderia selloana invasion. *Biological Conservation* 128, 128-135.

Pausas, J. G. & M. Verdu (2005). Plant persistence traits in fire-prone ecosystems of the Mediterranean basin: a phylogenetic approach. *Oikos* 109, 196-202.

Pearce, J. & S. Ferrier (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133, 225-245.

Pimentel, D., L. Lach, R. Zuniga & D. Morrison (2000). Environmental and economic costs of nonindigenous species in the United States. *Bioscience* 50, 53-65.

R, Development Core Team (2007). R: a language and environment for statistical computing, version 2.5.1, [Online]. R Foundation for Statistical Computing, Vienna. Available: http://www.r-project.org/. Accessed: June, 2007.

Reich, Robin & Richard Davis (2003). *Spatial Statistical Analysis for Natural Resources,* Fort Collins, CO, Colorado State University.

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

Richardson, D. M. & P. Pysek (2006). Plant invasions: merging the concepts of species invasiveness and community invasibility. *Progress In Physical Geography* 30, 409-431.

Rollins, M. G., T. W. Swetnam & P. Morgan (2001). Evaluating a century of fire patterns in two Rocky Mountain wilderness areas using digital fire atlases. *Canadian Journal Of Forest Research-Revue Canadienne De Recherche Forestiere* 31, 2107-2123.

Rouget, M., D. M. Richardson, J. L. Nel, D. C. Le Maitre, B. Egoh & T. Mgidi (2004). Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity And Distributions* 10, 475-484.

Schussman, H., E. Geiger, T. Mau-Crimmins & J. Ward (2006). Spread and current potential distribution of an alien grass, Eragrostis lehmanniana Nees, in the southwestern USA: comparing historical data and ecological niche models. *Diversity And Distributions* 12, 582-592.

Schwilk, D. W. & D. D. Ackerly (2001). Flammability and serotiny as strategies: correlated evolution in pines. *Oikos* 94, 326-336.

Smithwick, E. A. H., M. G. Turner, M. C. Mack & F. S. Chapin (2005). Postfire soil N cycling in northern conifer forests affected by severe, stand-replacing wildfires. *Ecosystems* 8, 163-181.

Stephens, S. L. & L. W. Ruth (2005). Federal forest-fire policy in the United States. *Ecological Applications* 15, 532-542.

Swetnam, T. W. (1993). Fire History And Climate-Change In Giant Sequoia Groves. *Science* 262, 885-889.

Theobald, David M. (2005). *GIS concepts and ArcGIS methods,* Fort Collins, Colo., Conservation Planning Technologies.

Thornton, P. E. (2001). DayMet Daily Surface Weather and Cimatalogical Summaries, [Online]. Numerical Terradynamic Simulation Group. Available: http://daymet.ntsg.umt.edu/DaymetPointSum.htm. Accessed: 2007.

Thornton, P. E., H. Hasenauer & M. A. White (2000). Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. *Agricultural And Forest Meteorology* 104, 255-271.

Thornton, P. E. & S. W. Running (1999). An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agricultural And Forest Meteorology* 93, 211-228.

Thuiller, W., D. M. Richardson, P. Pysek, G. F. Midgley, G. O. Hughes & M. Rouget (2005). Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology* 11, 2234-2250.

Turner, M. G., W. H. Romme & R. H. Gardner (1999). Prefire heterogeneity, fire severity, and early postfire plant reestablishment in subalpine forests of Yellowstone National Park, Wyoming. *International Journal Of Wildland Fire* 9, 21-36.

Underwood, E. C., R. Klinger & P. E. Moore (2004). Predicting patterns of non-native plant invasions in Yosemite National Park, California, USA. *Diversity And Distributions* 10, 447-459.

USDA (2002). National Resource Conservation Service PLANTS database, [Online]. Available: http://plants.usda.gov/. Accessed: 2008.

USGS (2008). MTBS, [Online]. USGS EROS Data Center. Available: http://fsgeodata.fs.fed.us/mtbs/index.html. Accessed: 2008.

Venables, W. N. & Brian D. Ripley (2002). *Modern applied statistics with S,* New York, Springer.

Vitousek, P. M., C. M. Dantonio, L. L. Loope & R. Westbrooks (1996). Biological invasions as global environmental change. *American Scientist* 84, 468-478.

Zeger, S. L. & K. Y. Liang (1986). Longitudinal Data-Analysis For Discrete And Continuous Outcomes. *Biometrics* 42, 121-130.

Zhu, Zhiliang, Carl Key, Donald Ohlen & Nate Benson (2006). Evaluate Sensitivities of Burn-Severity Mapping Algorithms for Different Ecosystems and Fire Histories in the United States. (ed. by NPS and USGS), p. 36. Joint Fire Sciences Program.

Zimmermann, N. E. & F. Kienast (1999). Predictive mapping of alpine grasslands in Switzerland: Species versus community approach. *Journal Of Vegetation Science* 10, 469-482.

Zouhar, Kris (2002). Fire Effects Information System, [Online]. USDA. Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory (Producer). Available: http://www.fs.fed.us/database/feis/about.html. Accessed: January, 2008.