Early Detection of Invasive Plants—Principles and Practices

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Cover: *Phragmites australis* (Cav.) Trin. Ex Steud. Photograph by Bradley A. Welch (2000).

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
SALLY JEWELL, Secretary

U.S. Geological Survey
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2014

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit http://www.usgs.gov or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit http://www.usgs.gov/pubprod

To order this and other USGS information products, visit http://store.usgs.gov

Suggested citation:
Welch, B.A., Geissler, P.H., and Latham, Penelope, 2014, Early detection of invasive plants—Principles and practices: U.S. Geological Survey Scientific Investigations Report 2012–5162, 193 p., *http://dx.doi.org/10.3133/sir20125162*.

ISSN 2328-031X (print)
ISSN 2328-0328 (online)
ISBN 978-1 4113 3672-8

Foreword

By Penelope Latham, Pacific West Region, Inventory and Monitoring Program, National Park Service

One if by land, and two if by sea…..
Henry Wadsworth Longfellow 1860

In 1860 Henry Wadsworth Longfellow chronicled our Nation's most famous early-detection warning system as he described the hanging of lanterns in the tower of the Old North Church and the subsequent midnight ride of Paul Revere to warn colonists of the impending British invasion. "One if by land, and two if by sea" is the warning that nearly every American schoolchild commits to memory. If only the early-detection of invasive, nonnative plants that threaten the integrity of national park ecosystems were as simple to address.

This document was developed to provide guidance and insight for National Park Service parks and other natural areas engaged in developing early-detection monitoring protocols for invasive plants. It is our hope that this information will demystify and ease the burden for those engaged in early-detection protocol development, whether national parks or other land managers faced with similar problems, rather than to be prescriptive. There are many approaches and often little consensus regarding how to develop early-detection monitoring; however, because of the pervasiveness of this issue and the fact that invasive species are not place bound (they freely cross ecological boundaries, States, nations, and continents), we have tried to provide guidance to encourage a consistent approach to monitoring. It is our intent to facilitate communication and enhance the ability to integrate and interpret results. We hope this will result in an efficient, effective tool for combating the considerable challenge of invasive plants to native ecosystems and biodiversity in national parks. We recognize that the task of detecting invasive species when populations are small is complex and will require a flexible approach to meet varying objectives, financial situations, and ecological conditions of individual parks.

Approaches to large and complex problems are slow in coming and are often developed in a stepwise fashion, each building on and indebted to the progress made previously. The Director of the National Park Service (NPS) distributed the first NPS plan for managing nonnative invasive plants ("A Strategic Plan for Managing Invasive Nonnative Plants on National Park System Lands"). Six key strategies were:

1. Preventing invasions,
2. Increasing public awareness,
3. Conducting inventory and monitoring of nonnative plants,
4. Conducting research and transfer technology,
5. Integrating of nonnative plant management into every aspect of park planning, and
6. Managing (control) of invasive nonnative plants.

Essential components of the proposed recommendations for preventing biological invasions were an early warning system to identify and eradicate new infestations of nonnative plants, the creation and maintenance of park-based lists of plant species that had not yet invaded the park but that were known to occur in the region and were likely to invade, and cooperation with other regulatory agencies. In 2000, a detailed servicewide action plan was drafted. A revision drafted in 2006 reemphasized the function of prevention through early-detection and rapid response (for control and management) and specifically identified Vital Sign monitoring networks in addition to parks as cooperators in the prevention strategy.

The organization of 270 national parks into 32 Vital Sign Inventory and Monitoring (I&M) Networks nationwide is the result of the Natural Resource Challenge National Parks Omnibus Management Act of 1998 (P.L. 105–391) and the subsequent 2001 National Park Service

Management Policies. This legislation instructed the networks to establish baseline information on the condition of park ecosystems through monitoring, to determine long-term trends in park resources, and to detect changes due to human influences.

An Invasive Species Workshop held in Fort Collins, Colorado, in June 2002 began the compilation of existing work related to developing protocols for monitoring invasive species, adopted the North American Weed Management Association minimum mapping standards, and made recommendations regarding adopting standards for field sampling and monitoring methods (for example Elzinga and others, 1998; Hiebert, 2002; Benjamin and Hiebert, 2004). In May 2003, following a survey that determined the majority of networks had identified Invasive Species as one of their top three Vital Signs, the Inventory and Monitoring Advisory Council (IMAC) appointed a working group to evaluate progress and make recommendations for next steps. Upon presentation of these recommendations at the National Inventory and Monitoring Annual Meeting in Lansdowne, Virginia, August 2003, the NPS Inventory and Monitoring networks agreed to adopt a nationwide strategic approach for monitoring invasive plants.

Following this decision, an NPS Invasive Species Coordinator position was established in Fort Collins. Proposals were then solicited from U.S. Geological Survey, Status and Trends scientists to assist NPS Inventory and Monitoring Networks with early-detection. It was agreed that the project would include developing a synthetic approach to early-detection monitoring that would provide guidance, establish standards and methodologies, and incorporate established and state-of-the-art technologies for early-detection protocol development. This document is the result of the multi-year collaborative effort on the part of USGS, university, and NPS scientists to meet this need.

The following chapters include a synthesis of current knowledge regarding invasion theory, identify Steps to Early-detection, and provide advice and decision trees regarding the utility of various approaches for prioritizing species and sites for early-detection monitoring at different scales. To clarify guidance incorporated in this document, examples are provided that cover a range of different options for early-detection in several parks and networks.

The Klamath Network (Pacific West Region) example includes research on the key issues and decisions that land managers may face in developing and implementing invasive species early detection monitoring programs for large and diverse natural areas. Other examples include approaches to early-detection monitoring over large areas by using remote sensing technology (Big Bend National Park), and a small park approach to early-detection monitoring using volunteers (San Francisco Bay Area Network). Finally, the document includes a protocol template with an outline to guide development of protocol sections and standard language included in key areas.

We are deeply indebted to our partners and collaborators for the technical aspects of this document and the persistence of their efforts in helping to achieve a technically useful but broadly applicable document. It is our heartfelt wish that this work will be of benefit to many other parks, I&M networks, and conservation areas and that we will have succeeded in furthering the ability of land managers to combat the insidious challenge of biological invasions to the integrity of native ecosystems.

Acknowledgments

There is always someone behind the scenes of any project of this scope and magnitude that provides the motivation and encouragement—especially when obstacles appear most formidable—to see the project to fruition.

We also benefited greatly from the knowledge and perspective proffered by Daniel Sarr. Along with the staff of the Klamath Network of parks, he has provided excellent insight as to the realistic needs of the National Park Service (NPS) Inventory and Monitoring (I&M) Program and the individual parks. Likewise, Susan O'Neil (Puget Sound Partnership) has provided invaluable understanding of the needs of resource managers. Her common sense, professional criticism, and diligence have kept us on track.

We thank all of the contributing authors for their tireless efforts and patience with this endeavor. The authors have served their turn long after their time because they recognize the importance of these issues to the natural resource management world. We appreciate and acknowledge their sacrifices, both personal and professional.

Several NPS members provided helpful reviews of early drafts of this document, including: Steve Acker, Bruce Bingham, Tamara Naumann, Regina Rochefort, Emily Spencer (now BLM), Andrea Williams, and Craig Young. We also recognize the participants from the NPS–USGS collaborative workshop held in Portland, Oregon, in August 2004, who contributed early in this process to the formation of conceptual models and the development of key components of the framework for early detection of invasive plants.

The U.S. Geological Survey by way of the Ecosystems, Status and Trends Program, National Park Monitoring Project, has provided financial, technical, and professional support for this project, including grants for associated research. Linda Drees, NPS Invasive Species Branch Chief, and Steven Fancy, Monitoring Coordinator for the NPS I&M Program, (now retired) have provided logistical, institutional, and financial support throughout this project.

Finally, we must acknowledge the many invasive-plant management professionals in the U.S. National Parks and from other natural areas who, day after arduous day, devote their lives to protecting our natural and cultural resources from the impacts of invasive plant species. The future of our natural and cultural heritage is in your hands!

Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Contents

Acronyms

Acronym	Definition
AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
AR	Arkansas
ARPO	Arkansas Post National Memorial
AUC	Area Under the Curve
AVHRR	Advanced Very High Resolution Radiometer
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
AZ-WIPWG	Arizona Wildland Invasive Plant Working Group
BAEDN	Bay Area Early Detection Network
BIBE	Big Bend National Park
BRMD	Biologic Resource Management Division
CaEPMT	California Exotic Plant Management Team
Cal-HIP	California Horticultural Invasives Prevention
Cal-IPC	California Invasive Plant Council
CART	Classification and Regression Trees
CDFA	California Department of Food and Agriculture
CD-ROM	Compact Disc-Read Only Memory
CEM	Climatic Envelope Model
CRLA	Crater Lake National Park
CUVA	Cuyahoga Valley National Park
DEM	Digital Elevation Model
DLG	Digital Line Graph
DO	Director's Order
DOI	Department of the Interior
DOQ	Digital Orthophoto Quadrangle
DRG	Digital Raster Graphics
DVD	Digital Versatile Disc
EDRR	Early Detection and Rapid Response (from appendix, not main body of handbook)
EFMO	Effigy Mounds National Monument
EPIgdb	Exotic Plant Information Geodatabase
EPMT	Exotic Plant Management Team
EPPdb	Exotic Plant Prioritization Database
ERD	Entity Relationship Diagram
ESRI	Environmental Systems Research Institute

ETM +	Enhanced Thematic Mapper Plus
EZ	Elevation Zone
FICMNEW	Federal Interagency Committee for the Management of Noxious and Exotic Weeds
FPC	Finite Population Correction
FTP	File Transfer Protocol
FY	Fiscal Year
GAM	Generalized Additive Model
GB	Gigabyte
GIS	Geographic Information Systems
GLM	Generalized Linear Model
GOGA	Golden Gate National Recreation Area
GPRA	Government Performance and Results Act
GPS	Global Positioning System
GRTS	Generalized Random-Tessellation Stratified
GWCA	George Washington Carver National Monument
ha	Hectare
HEAR	Hawaiian Ecosystems at Risk
HEHO	Herbert Hoover National Historic Site
HOCU	Hopewell Culture National Historic Park
HOME	Homestead National Monument of America
HOSP	Hot Springs National Park
HRP	Habitat Restoration Program
HRT	Habitat Restoration Team
HTTP	Hyper Text Transfer Protocol
HUC	Hydrological Unit Code
I&M	Inventory and Monitoring
IA	Iowa
IEP	Invasive Exotic Plant
IL	Illinois
IMAC	Inventory and Monitoring Advisory Council
IN	Indiana
IPAW	Invasive Plants Association of Wisconsin
IRMA	Integrated Resource Management Applications
ISED	Invasive Species Early Detection
ISSG	Invasive Species Specialist Group (part of IUCN)
ITIS	Integrated Taxonomic Information System

IUCN	The World Conservation Union
KS	Kansas
LABE	Lava Beds National Monument
LAVO	Lassen Volcanic National Park
LDA	Linear Discriminant Analysis
LIBO	Lincoln Boyhood National Memorial
m	meter
mm	millimeter
MN	Minnesota
MO	Missouri
MSAVI	Modified Soil Adjusted Vegetation Index
MSDI	Moving Standard Deviation Index
MSS	Multispectral Scanner
MTMF	Mixture Tuned Matched Filtering
NAIP	National Agriculture Imagery Program
NASA	National Aeronautics and Space Administration
NAISMA	North American Invasive Species Management Association
NDVI	Normalized Difference Vegetation Index
NE	Nebraska
NEPA	National Environmental Policy Act
NHD	National Hydrography Datasets
NIR	Near Infrared
NISC	National Invasive Species Council
NISC	National Invasive Species Council
NM	National Monument
NOAA	National Oceanic and Atmospheric Administrations
NPS	National Park Service
NRC	Natural Resource Council
NRCS	Natural Resources Conservation Science
NRDT	Natural Resource Database Template
NR-MAP	Natural Resource Management and Assessment Program (from appendix)
NRPC	Natural Resource Program Center
NRTR	Natural Resource Technical Report
OH	Ohio
OR	Oregon
ORCA	Oregon Caves National Monument

OS	Opportunistic Sampling
OTA	Office of Technology Assessment
PCC	Percent Correct Classification
PDA	Personal Digital Assistants
PERI	Pea Ridge National Military Park
PINN	Pinnacles National Monument
PIPE	Pipestone National Monument
PORE	Point Reyes National Seashore
QA	Quality Assurance
QC	Quality Control
QDA	Quadratic Discriminant Analysis
RAM	Random Access Memory
RAWS	Remote Automated Weather Stations
ROC	Receiver Operator Characteristic
SAVI	Soil Adjusted Vegetation Index
SCA	Soil Conservation Association
SDM	Species Distribution Model
SFAN	San Francisco Bay Area Network
SOP	Standard Operating Procedure
SSP	Site Stewardship Program
SU	Sampling Unit
TAPR	Tallgrass Prairie National Preserve
TM	Thematic Mapper
TNC	The Nature Conservancy
TSN	Taxonomic Serial Number
UNESCO	United Nations Educational, Scientific and Cultural Organization
USDA	United States Department of Agriculture
USFS	United States Forest Service
USGS	United States Geological Survey
USPEO	United States Presidential Executive Order
UTM	Universal Transverse Mercators
WHIS	Whiskeytown National Recreation Area
WHR	Wildlife Habitat Relationship
WICR	Wilson's Creek National Battlefield
WIMS	Weed Information Management System
WMA	Weed Management Area

Invasive Plant Early-Detection Decision Tree

This key walks the reader through 17 questions regarding the planning, setup, and implementation of an invasive plant early-detection program. The key will send the reader to important chapters within this document (as well as other resources) to help make the most of its resources. We recognize that the document is long and may provide more detail than some managers will need in order to set up a program. This decision tree will help expedite the use of this information on the basis of the current resources available to the park or network. Many chapters will also have a decision tree specific to that topic which will facilitate gleaning the necessary information at a finer level. Although the decision tree is designed with the National Park Service Inventory and Monitoring (I&M) networks in mind, other parks and organizations will also benefit from its direction.

1. Do you already have an early-detection program in place and just need to formalize a protocol in the I&M format?

Yes See Chapters 1–3 for background information and then see the examples in Chapters 11–13. Also see the Quick-Start Guide to identify pieces of an early-detection program that you may have overlooked and Chapter 3, figure 3.1, for an overview of successful components of an early-detection program.

No Go to Question 2.

2. Are you a small park with small natural resource staff and little funding for an early-detection program?

Yes Go to Question 3.

No Go to Question 4.

3. Is the I&M Network or another local partnership (for example Weed Management Areas, Invasive Plant Councils) developing an early-detection program?

Yes Talk to your Network I&M Coordinator (or partner) to ensure that you do indeed fit into the “small park/few resources” category — you may actually be part of a larger effort with available resources through your network. You’ll notice throughout the document that many chapters have options for “small parks.” These tend to be less expensive, less intensive options that still allow you to meet your objectives and goals. Then see Question 4.

No Go to Question 4.

4. Do you have a current inventory/survey of invasive plants in your park?

Yes Go to Question 5.

No Conduct an inventory/survey. (See *Inventory and Survey Methods for Nonindigenous Plant Species* by Rew and Pokorny 2006; available at *http://www.weedcenter.org/store/images/books/INVENTORYBOOK.pdf*, accessed March 24, 2014). If you are interested in building a predictive model to help prioritize invasion-prone sites, be sure to consult with an invasive-species ecologist and a statistician before conducting the inventory. In the meantime, compile lists of species from existing park data, regional watch lists, and NPSpecies (NPS certified species database). See Chapter 13 for an example of this approach. Update your list after conducting the inventory. Then go to Question 5. Also see Question 13.

5. Have you identified the scope, goals and objectives for an early detection program? This will include deciding whether to focus on species and (or) sites.

Yes Go to Question 6.

No See Chapter 4.

6. Do you have one or more prioritized lists of species to search for?

Yes Go to Question 7.

No See Chapter 5 and Chapter 10. Research and join formal groups and (or) have informal discussions with neighboring landowners/managers to determine what species might be headed your way. See also Chapter 5 for an approach to amalgamating multiple regional species lists. Then see Question 7.

7. Are you interested in searching vulnerable sites (that is susceptible to invasion) as well as valuable sites (that is uninvaded and (or) sensitive habitat) at your park?

Yes Go to Question 8.

No Go to Question 9.

8. Do you have the expertise to develop models or the funding to get assistance in developing models for predicting sites likely to be invaded by priority species?

Yes See Chapters 6, 7 and 9, then see Question 11.

No Develop your own strategy using conceptual models (written or graphical). All early-detection programs should be based on a model, but it need not be a complex, computer-generated predictive model. Get the practical approach that is in your head (from working with species in your park) down on paper to justify such things as searching along roads and trails, riparian areas, and so forth. Seek review from local experts and revise your model. Be sure that your plan corresponds with your early-detection objectives. Then see Question 11.

9. Is your main objective to keep invasive plants out of valuable sites (“weed-free” and (or) sensitive habitat), either because your available resources for early detection are extremely limited or because the extent of the infestation is large and the number of valuable sites is small?

Yes Consult with park managers to identify and prioritize high value sites. Then see Chapter 8 to develop a sampling plan.

No Just vectors and pathways……….Go to Question 10.

10. If you are only planning on searching along likely vectors and pathways of invasion (e.g., roads, trails, riparian zones), you should recognize that it is very likely that the species you encounter are moving into adjacent areas which may be of management concern. Please consider this decision carefully. If that is still your objective, you can use stretches of roads, trails and riparian zones as your sample units. Then see Chapter 8 to develop a sampling plan.

11. Are you interested in developing a sampling design that incorporates probability so that you can determine trends and report on improvements in park condition?

Yes See Chapters 6, 7 and 8, then go to Question 12.

No Go to Question 13.

12. Are you interested in developing a sampling design that incorporates probability so that you can assess the severity of the infestations and (or) improve your search efficiency?

Yes See Chapter 8, particularly the discussion of cluster sampling and adaptive sampling, then go to Question 14.

No Go to 13.

13. Please consider that even though your program may not currently have the capacity or expertise to develop a statistically valid sampling design, it might still be in your best interest to consult with some other I&M Networks, USGS, local partners, and (or) academics to obtain professional advice on sampling design before you collect data. If data are collected properly (including both presence and absence data), it can serve as the foundation to develop predictive models, and assess invasion patterns of species and can be used in other analyses, which might seem unlikely to you now but could benefit the program in the future. Whether or not you decide to use probability sampling, see Chapters 8 and review Chapters 6, 7, and 9. Then go to Question 14.

14. Does your park or network have a data-management system in place to track and analyze the early detection information (GIS and tabular data)?

Yes See Chapter 10 to be sure that it meets NAWMA standards and is compatible with the upcoming National I&M Invasive Plant database, then see Question 16.

No Go to Question 15.

15. Work with your Network I&M data manager or partners to determine if there are existing databases you can adopt for park use. You can also consider using The Nature Conservancy’s WIMS (Weed Information Management System) as some other parks are doing (available

free for download). See Chapter 10 for more information on Data Management and what data fields are recommended, then see Question 16

16. Are you planning to use volunteers for early-detection monitoring?
    - **Yes** See Chapter 13 as an example of how volunteers are used for early detection at Golden Gate National Recreation Area, then see Question 17.
    - **No** Go to Question 17.
17. Are you ready to start developing your early detection protocol?
    - **Yes** See Chapter 11 for example.
    - **No** See the Quick Start Guide which has a list of questions about starting an invasive plant early-detection program. It will send you to the part of the document that can assist in answering those questions and get you ready to develop your early-detection protocol. Talk to people. Read the Applications and Principles examples (Chapters 11–13) to get ideas!

Quick-Start Guide

Focus area	Reader's question	Where to look
Getting started	Why should I consider starting an early-detection program in my park? It's more than I can do to fight the weeds I know about now.	Foreword, Chapters 1 and 3.
	What is an invasive species?	Chapter 1, Glossary.
	Are there other NPS documents and programs that relate to early-detection?	Chapter 1.
	What are the costs and benefits of an early-detection program?	Chapters 1 and 2. Fig 2.1, 2.2
	How did early-detection become an important part of the I&M program?	Foreword, Chapter 1.
	What are the key components to an early-detection program?	Chapter 3.
	What NPS networks are working on early-detection protocols?	Chapters 11–13. (See also *http://irma.nps.gov/App/ProtocolTracking*, accessed March 24, 2014.)
	I am about to inventory my park for invasive species. Where should I start?	Chapters 4 and 5. See also Rew and Pokorny 2006. (See References Cited for details.)
	I have completed an inventory of invasive plants and I'm ready for an early-detection program. How do I get started?	Start with the Decision Tree then read Chapters 1 and 4.
	I work at a small park with a small staff and an even smaller budget. Where should I start?	Start with the Decision Tree then read Chapters 1, 4, and 11.
	I work at a large park with extensive natural areas or wilderness. Where should I start?	Start with the Decision Tree then read Chapters 1, 4, 6 (especially fig. 6.3), 11, and 12.
What to consider	What do I need to know about how biological invasions occur that will help me to design an early-detection program?	Chapter 2.
	Can I tell if a plant will be invasive from its biology or life-history characteristics?	Chapter 2, Chapter 5.
	Where can I find life-history information on particular species?	Chapters 2 and 5.
	What characteristics contribute to a site being susceptible to invasion?	Chapter 2, Chapter 5.
	Should I consider vectors and pathways?	Chapter 2, Chapter 5.
Defining the scope and objectives	What is the difference between a goal and an objective?	Chapter 4, Glossary.
	How specific does an objective need to be for a monitoring protocol?	Chapter 4.
Data needs	What type of data do I need to predict where species will invade?	Chapter 5, Chapter 7.

Focus area	Reader's question	Where to look
	Should I look for new species or new occurrences of existing species?	Chapters 4, 5, and 10.
	What type of data do I need to prioritize invasive species?	Chapter 5.
	Why should I collect absence data?	Chapters 5, 7, 8, and 9.
	If I don't have absence data for surveys, are my data still useful? What can I do to improve the data?	Chapters 5, 7, 8, and 9.
	Should I follow the NAWMA standards?	Chapter 3, Chapter 10.
	Where can I find climatic and environmental data for my park?	Chapter 5 and associated links.
	How do I prioritize what species to look for and where to look for them at different scales?	Chapter 5.
Predictive models	Does a predictive model have to be computer-based?	Chapter 3, 7 and 8.
	What is species distribution modeling and how does it relate to early-detection?	Chapter 7.
	What are the important considerations for invasive species modeling?	Chapter 7, Box 7.2 and 7.3.
	What are training data for modeling and how do I collect this type of data?	Chapter 7, Box 7.4.
	What is the difference between plot-based and remote-sensing models?	Chapter 6, Chapter 7.
	When should I use remote-sensing for early-detection?	Chapter 6, Figure 6.3.
	What type of imagery should I use?	Chapter 6, Figure 6.3.
	Why would I use a spatially explicit versus a non-spatial model?	Chapter 7.
	Are there examples of predictive models?	Chapter 7; Chapter 6; and Chapters 12.
Sampling design	When should I use probability sampling for early-detection?	Chapter 8.
	What are the sampling design options for small parks with few resources?	Chapter 8.
	What are the steps in developing a sampling design?	Chapter 8.
	Should I use stratified sampling?	Chapter 8.
	Are there benefits to using cluster or adaptive cluster sampling?	Chapter 8.
	Do I analyze the data differently for different sampling designs?	Chapter 8.
	What is GRTS? Should I use it for sample selection?	Chapter 8.
	Are there free software programs to help me design and analyze my data?	Chapter 8.
	What about the difficulty of detecting small infestations or species?	Chapter 8.

Focus area	Reader's question	Where to look
	Can I use areas of land (for example, subwatersheds) as sample units?	Chapter 13.
Validation	How do I evaluate my predictive model to know if it will work?	Chapter 9.
	What is the difference between internal and external validation? Do I need to do both?	Chapter 9.
	What metrics are used to assess models?	Chapter 9.
Logistics and finance	How do I manage data once an early-detection program is in place?	Chapter 10.
	Are there existing databases that I can use?	Chapters 10, 13
	How will the data be used by the park and network?	Chapter 10.
	If I work at a small park and we remove each infestation as we find them, how do I document that?	Chapters 10 and 13.
	Can my existing opportunistic data (legacy data) be imported into a new system?	Chapter 10.
	Can my paper maps of weed locations be updated into a new geodatabase system?	Chapter 10.
	How do I decide if I should remove a species or wait for treatment when I find it?	Chapter 10.
	How do I decide if I should focus effort on spread pathways and vectors or sensitive habitat?	Chapters 5 and 10.
	Are there additional resources for funding early-detection or rapid response?	Chapter 10; See also National Invasive Species Web site (*http://www.invasivespeciesinfo.gov/*).
	How do I accomplish early-detection at a park and still consider multiple scales of invasion?	Chapters 4 and 11. See also Chapter 12 for examples.
	Is it worth the time to train and manage volunteers for help with early-detection?	Chapter 13.
	How do I get involved in regional efforts to know what species to look for?	Chapters 5 and 10. See also Chapters 11–13 for examples.
Other information	Where can I find acronyms and a glossary for this document?	Pages vii and 187.
	What if my park is only planning to implement a few components of the handbook?	Chapters 1, 3.
	I would like to collaborate with USGS to develop a predictive model; where should I start?	Chapters 5, 6, and 7.

Extended Abstract

Invasive plants infest an estimated 2.6 million acres of the 83 million acres managed by the National Park Service (NPS) in the United States. The consequences of these invasions present a significant challenge for the NPS to manage the agency's natural resources "unimpaired for the enjoyment of future generations." More NPS lands are infested daily despite diligent efforts to curtail the problem. Impacts from invasive species have been realized in most parks, resulting in an expressed need to control existing infestations and restore affected ecosystems. There is a growing urgency in the NPS and other resource management organizations to be proactive—to protect resources not yet affected by current and future invasive species. Invasive species most certainly will continue to be a management priority for the national parks well into the 21st century as nonnative plants increasingly become established in U.S. ecosystems.

In 1996, the Director of NPS distributed the first NPS plan for managing nonnative plants with six key strategies identified including preventing invasions and conducting inventory and monitoring of nonnative plants. In 2000, a detailed action plan was drafted, with a revision in 2006, which reemphasized the function of prevention and early-detection/rapid response. The Inventory & Monitoring (I&M) Program, which was established in 1998 as part of the Natural Resource Challenge National Parks Omnibus Management Act, organized 270 national parks into 32 networks to inventory resources and conduct long-term monitoring of key indicators, used to monitor park resources. Invasive species, particularly early-detection of invasive plants, has consistently ranked high among I&M networks as a key indicator to be monitored to assess resource condition.

The NPS I&M Program, in collaboration with the U.S. Geological Survey (USGS) Status and Trends Program, compiled this document to provide guidance and insight to parks and other natural areas engaged in developing early-detection monitoring protocols for invasive plants. While several rapid response frameworks exist, there is no consistent or comprehensive guidance informing the active detection of nonnative plants early in the invasion process. There are many approaches and often little consensus regarding how to develop efficient early-detection monitoring; however, because of the pervasiveness of this issue and the fact that invasive species are not place bound (that is they freely cross ecological boundaries, states, nations, and continents), we have tried to provide guidance to encourage a consistent approach to monitoring. We hope this will result in an efficient, effective tool for combating the considerable challenge of invasive plants to native ecosystems and biodiversity in national parks and elsewhere. We recognize that the task of detecting invasive species when populations are small is complex and will require a flexible approach to meet varying objectives, financial situations, and ecological conditions of individual parks and natural areas.

Early-detection was selected as a primary focus for invasive-species monitoring because, along with rapid response, it is a key strategy for successful management of invasive species. Eradication efforts are most successful on small infestations (that is less than 1 hectare) and become less successful as infestation size increases, to the point that eradication is unlikely for large (that is greater than 1,000 hectares) populations of invasive plants. By tracking new species and new infestations, resource managers may begin to understand the invasion patterns affecting their parks and subsequently formulate strategies that will allow for improved management actions. Too often early-detection is conducted passively in parks. Managers rely on erratic reports from visitors, groundskeepers, and backcountry rangers as the source of information to trigger management action. The nature of this approach requires spontaneous decisions to be made about allocation of staff time and other resources, directing them away from current projects toward unexpected and unconfirmed issues. Alternatively, active detection methods may be used such that managers respond rapidly to predictable, confirmed reports in a timely and cost-effective manner. This document provides guidance for natural resource managers wishing to detect invasive plants early through an active, directed monitoring program.

Chapters 1–3 establish the need for early-detection of invasive plants, provide linkages to relevant policies and programs, summarize dominant invasion theories relevant to this text, and outline the key components required to implement a successful early-detection monitoring program. Chapters 4–10 address each of the early-detection steps in detail including setting goals and objectives, prioritizing species and resources for monitoring, identifying remote sensing options, choosing analytical processes, conducting evaluation and assessment, formulating sampling design, managing data, and implementing detection methods. Chapters 11–13 provide examples of applications of early-detection principles, including: insights into early-detection protocol development across multiple parks, an approach to probabilistic predictive modeling of invasive plants with detailed examples, an example using remotely sensed and geographically referenced data for predicting the risk of occurrence for target species, a small park example using volunteers, and a protocol for integrating early-detection and long-term trends monitoring across large and small parks. Subsequent materials include citations, a glossary, detailed reports from participating researchers, and a protocol template that meets NPS I&M program standards. The protocol template will be useful for those who may be required to document their specific early-detection procedures.

This document also has a Quick-Start Guide to direct readers to specific chapters and text relevant to their needs. Each chapter was written by a USGS researcher(s) or NPS manager(s)/researcher(s). Decision trees and flow charts contained within several of the chapters assist the reader in deciding what methods to choose and when to use them. The various steps (each presented as a separate chapter) in Section 2 are meant to follow a conceptual model developed by the I&M program that encapsulates the idealized components of an early-detection program (see Chapter 3). A park or network may decide to implement only a few of the relevant components. This document is written in a modular format to accommodate use of individual chapters. It may also be approached in a linear fashion, as a sequence of steps leading to a comprehensive approach to early-detection.

We have written this document to reach a large audience. Our primary audience comprises resource professionals within the National Park Service (NPS) Inventory and Monitoring (I&M) Program's networks of parks, but we think that the knowledge and experience captured in this document is more broadly applicable to include other natural areas professionals. Even within the NPS I&M Program, there is broad variation in the technical expertise and resources available. Consequently, some readers may find parts of this document trivial or obvious, while others may find it too technical. We have chosen to emphasize the technical side of invasive species early-detection because this is the arena in which most professionals need more guidance. That is to say, we are recommending a comprehensive approach to early-detection of invasive plant species. This approach includes but is not limited to complex techniques that may seem to be just beyond the budgetary and (or) time-bound grasps of some resource professionals. Experience and recent reviews of the state of invasive-species management suggest that more definitive work is required to provide state-of-the-art early-detection strategies and tools of this kind to resource professionals (see Chapter 1). Nonetheless, we have provided low-cost options, examples, and approaches within the central chapters of this document as well as several case studies that are currently exploring early-detection principles in U.S. parks and networks.

This is intended to be a living document that will be updated regularly by way of an electronic Internet copy as new materials become available. In particular, we anticipate specific protocols developed by individual I&M networks will be linked to this document to provide readers with a range of specific implementation examples.

Introduction

By Bradley A. Welch

Chapter 1 of
Early Detection of Invasive Plants—Principles and Practices
Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Insets

Chapter 1
Introduction

By Bradley A. Welch

Globalization of commerce, transportation, human migration, and recreation in recent history have introduced nonnative species (also referred to as exotic, alien, or non-indigenous species; see Glossary section) to new areas at an unprecedented rate. Biogeographical barriers that once restricted the location and expansion of species have been circumvented, leading to increased homogenization of the Earth's biota. Although only 10 percent of introduced species become established and only 1 percent become problematic (Williamson, 1993; Williamson and Fitter, 1996) or invasive, nonnative species have profound impacts worldwide on the environment, economies, and human health. Invasive species have been directly linked to the replacement of dominant native species (Tilman, 1999), the loss of rare species (King, 1985), changes in ecosystem structure, alteration of nutrient cycles and soil chemistry (Ehrenfeld, 2003), shifts in community productivity (Vitousek, 1990), reduced agricultural productivity, and changes in water availability (D'Antonio and Mahall, 1991). Often the damage caused by these species to natural resources is irreparable and our understanding of the consequences incomplete. Invasive species are second only to habitat destruction as a threat to wildland biodiversity (Wilcove and others, 1998). Consequently, the dynamic relations among plants, animals, soil, and water established over many thousands of years are at risk of being destroyed in a relatively brief period.

The consequences of these invasions present a significant challenge for the National Park Service (NPS) to manage the agency's natural resources "unimpaired for the enjoyment of future generations." National Parks, like other land management organizations, are deluged by new nonnative species arriving through predictable, sudden, and unexpected anthropogenic pathways. "Predictable pathways" include roads, trails, and riparian corridors easily accessible to humans. Almost all nonnative plants, arthropods, and pathogens in the United States were transported to the United States by humans. Natural introductions by wind, birds, and so on, have occurred occasionally for pathogens but are rare for other taxa (National Research Council, 2002). Natural introductions occur more infrequently with increasing isolation; for example, a long distance plant dispersal rate of one species every 100,000 years has been estimated for the Hawaiian flora (Fosberg, 1948, and Carlquist, 1974, in National Research Council, 2002). "Sudden pathways" may include long-distance dispersal through cargo containers, air freight, and importation for agriculture and horticulture. "Unexpected introductions" may occur when weed seeds are unexpectedly found in restoration planting mixes or when nonnative species exhibit sudden shifts in host species.

Box 1.1. Nonnative plants are making their presence known across U.S. ecosystems, accounting for an ever-increasing proportion of State flora:

- 43 percent in Hawaii,
- 36 percent in New York,
- 25 percent in Missouri,
- 18 percent in California,
- and 10 percent in Texas

(Rejmanek and Randall, 1994).

At the continental scale, trade liberalization increases the likelihood that some of the hundreds of thousands of species of plant pests not yet found in the United States will someday arrive here (National Research Council, 2002). Nonnative plants claim an estimated 4,600 acres of public lands each year in the United States (Asher and Harmon, 1995), significantly altering local flora and extracting economic consequences estimated at greater than $120 billion dollars per year (Pimentel and others, 2005). For example, the flora of the States of Hawaii and New York comprise approximately 43 percent and 36 percent nonnative plants, respectively (Rejmanek and Randall, 1994). Invasive plants infest an estimated 2.6 million acres of the 83 million acres managed by the NPS.

More NPS lands are infested daily despite diligent efforts to curtail the problem. Impacts from invasive species have been realized in most parks, resulting in an expressed need to control existing infestations and restore affected ecosystems. For example, the invasion of North American rangelands by cheatgrass (*Bromus tectorum*) and Florida wetlands by Melaleuca (*Melaleuca quinquenervia*) has altered the frequency and intensity of fires and wetland hydrology, respectively, in these ecosystems, thus substantially affecting ecosystem services (National Research Council, 2002). There is a growing urgency to be proactive—to protect resources not yet affected by current and potential invasive species (Marler, 1998). Invasive species most certainly will continue to be a management priority for the National Parks well into the 21st century as nonnative plants increasingly become established in U.S. ecosystems.

Listing 1.1. Several sources of guidance that provide assistance to National Park Service managers in formulating a strong and clear policy on managing invasive species in the parks.

International laws and conventions on endangered species, wetlands, and plant protection—among others (see *http://www.invasivespeciesinfo.gov/laws/intl.shtml* for relevant documents) influence management of invasive species within national parks.

Federal laws and regulations influencing invasive plant management in the United States also pertain to such issues as noxious weeds, plant introduction and transport, pesticide use, endangered species, aquatic ecosystems, and agricultural practices (see *http://www.invasivespeciesinfo.gov/laws/federal.shtml* for a list and links to pertinent documents).

Executive Order 13112 (*http://www.invasivespeciesinfo.gov/laws/execorder.shtml*) established the National Invasive Species Council. Currently the Council comprises 13 Departments and Agencies that deal with invasive species issues at the national level.

State and local laws and regulations (see *http://www.invasivespeciesinfo.gov/laws/statelaws.shtml for a state-by-state database*) offer specific legal guidance to parks and their managers pertaining to issues closer to home.

NPS Management Policies 2006 (*http://www.nps.gov/policy/MP2006.pdf*) provide an agency-wide framework for nonnative plant management in the context of other natural and cultural resources.

Natural Resources Management Guidelines (NPS-77) offer more detailed direction on actions required and (or) appropriate with respect to natural resource issues on NPS lands. [N.B.–The nonnative species section of NPS-77 is being revised and is unavailable at this time. Revisions will be included in the Natural Resource Management Reference Manual #77 (*http://www.nature.nps.gov/RM77/*) when they are completed.]

NPS Invasive Species Action Plan Final Draft (March 2006) provides national guidance for the national parks.

A park's individual **Natural Resource Management Plan** details specific goals, objectives, and actions for a specific park.

Many parks may also have a specific **invasive species management plan**.

Invasive species have been consistently ranked as a top vital sign for long-term monitoring as part of the NPS Inventory & Monitoring (I&M) Program. At present, 21 of 32 networks and 205 parks have invasive plants as a top vital sign for monitoring under the broad category of biological integrity. Early detection (and rapid response) is a key component of a long-term monitoring program to protect biodiversity in the national parks and elsewhere. Although this document was originally designed with the NPS I&M networks in mind (see Foreword), the information detailed within this document will have application to any parks or other natural areas engaged in developing long-term early detection monitoring protocols for invasive plants.

Need for Early Detection

Prevention and early detection are the principal strategies to successful invasive plant management. While there is a need for long-term suppression programs to address high impact species (for example, velvet tree (*Miconia calvescens* DC.) in the Pacific Islands or Tamarisk (*Tamarix* species) in the Southwest), eradication efforts are most successful for infestations less than one hectare in size (Rejmanek and Pitcairn, 2002). Eradication of infestations larger than 100 hectares is largely unsuccessful, costly, and unsustainable (Rejmanek and Pitcairn, 2002). Further, in their detailed review of the nonnative species problem in the United States, the U.S. Congress, Office of Technology Assessment (1993), stated that the environmental and economic benefits of supporting prevention and early detection initiatives significantly outweigh any incurred costs, with the median benefit-to-cost ratio being 17:1 in favor of being proactive.

NPS Management Policies (National Park Service, 2006) address both the introduction of nonnative species (Section 4.4.4.1) and the removal of established nonnative species (Section 4.4.4.2), dictating that "(e)xotic species will not be allowed to displace native species if displacement can be prevented" (Listing 1.1).

Box 1.2. Early detection involves active, planned measures and passive, incidental reports used to locate newly introduced nonnative species in a given area <u>before</u> populations become established. Early detection may be used successfully to:

Find nonnative species that are **new or undiscovered** in a defined area.

Find **incipient or new satellite populations** of nonnative species existing elsewhere in a previously invaded area.

Although preventing the introduction of invasive plants is the most successful and preferred strategy for resource managers, the realities of globalization, tight fiscal constraints, and limited staff time guarantee that invaders will get through park borders. Fortunately, invasive plants quite often undergo a lag period between introduction and subsequent colonization of new areas (Smith and others, 1999). Diligent managers, then, can take advantage of early detection monitoring techniques to make certain invasive species are found and successfully eradicated before populations become well established. However, policy and management implications only become clear when common processes and probabilistic transitions during invasion are recognized (Lodge and others, 2006).

This strategy requires the successful manager to (1) detect species early (that is, find a new species or an incipient population of an existing species while the infestation is small [less than 1 hectare]), and (2) respond rapidly (that is, implement appropriate management techniques to eliminate the invasive plant and all of its associated regenerative material).

All too often early detection is conducted passively. Managers rely solely on erratic reports from visitors, groundskeepers, and backcountry rangers as the source of information that triggers management action. The nature of this approach requires spontaneous decisions to be made about allocation of staff time and other resources, directing them away from current projects toward unexpected and unconfirmed issues. Alternatively, active detection methods may be used such that managers respond rapidly to predictable, confirmed reports in a timely and cost-effective manner. Rather than relying solely on serendipity, systematically collected information keeps managers well informed about the status of ecosystems in their parks such that limited resources can be focused on high priority sites and (or) areas highly susceptible to invasion. Lower priority and less susceptible sites also are surveyed, but less intently, according to the designated sampling plan. When conducted properly, active early detection is the most profitable investment for invasive species management (Rejmanek, 2000; Smith and others, 1999).

Handbook Context

In 2002, the National Research Council thoroughly reviewed the state of knowledge on invasions, including predicting introductions and establishment. Their key recommendations for furthering knowledge and preventing invasions of nonnative species include the following:

- "… Careful recording of the circumstances of arrival, persistence, and invasion of non-indigenous species in the United States would substantially improve prediction and risk assessment."
- "Information on the structure and composition of natural ecosystems in North America (and the disturbance regimes within them) should be reinterpreted by the scientific community to analyze these ecosystems' vulnerability to biotic invasion. Attention should be paid to identifying groups of native species that could be vulnerable or could facilitate the establishment of non-indigenous species."
- "A central repository of information relevant to immigrant species would accelerate efforts to strengthen the scientific basis of predicting invasion. Information collected by federal, state, and international agencies, academic researchers, and others should be brought together in a single information facility or service so that it can be evaluated collectively, to permit the construction of needed data sets and the design of appropriate experiments, and to document the circumstances surrounding invasions."

The Ecological Society of America (Lodge and others, 2006) expands on these recommendations in a recent review of United States invasive species policy and management. Specifically, the report states:

- "Without improved strategies based on recent scientific advances and increased investments to counter invasions, harm from invasive species is likely to accelerate."
- "Federal leadership, with the cooperation of state and local governments, is required to increase the effectiveness of prevention of invasions, detect and respond quickly to new potentially harmful invasions, control and slow the spread of existing invasions, and provide a national center to ensure that these efforts are coordinated and cost effective."
- "Specifically, the Ecological Society of America recommends that the Federal government take the following six actions: (1) Use new information and practices to better manage commercial and other pathways to reduce the transport and release of potentially harmful species; (2) Adopt more quantitative procedures for risk analysis and apply them to every species proposed for importation into the country; (3) Use new cost-effective diagnostic technologies to increase active surveillance and sharing of information about invasive species so that responses to new invasions can be more rapid and effective; (4) Create new legal authority and provide emergency funding to support rapid responses to emerging invasions; (5) Provide funding and incentives for cost-effective programs to slow the spread of existing invasive species in order to protect still uninvaded ecosystems, social and industrial infrastructure, and human welfare; and (6) Establish a National Center for Invasive Species Management (under the existing National Invasive Species Council) to coordinate and lead improvements in Federal, State, and international policies on invasive species."

The report emphasizes the need for a broad-based, organized, and coordinated approach to invasive species management to ensure long-term success. Early detection and rapid response strategies are emphasized throughout the report and highlighted explicitly in the details of the third recommendation as key components of successful management plans.

By way of this text and its associated materials, land management organizations will be building on these important recommendations as a contribution to the greater body of knowledge regarding the threat of invasive species in the United States.

This document will assist NPS park managers and Inventory & Monitoring (I&M) networks across the United States in meeting invasive species management directives issued by the NPS as part of the Natural Resource Challenge (National Parks Omnibus Management Act of 1998 [P.I. 105–391]). In 2002 the NPS I&M program held a workshop to recommend guidelines and tools for developing protocols for inventory and monitoring of invasive plants. One of the four adopted goals is to “prevent and detect new alien plant invasions, and eradicate new invasives” (Hiebert, 2002; Benjamin and Hiebert, 2004). The group developed a preliminary flowchart of the components of an effective weed-monitoring program and adopted the North American Weed Management Association standards (Beard and others, 2001). This document meets the goal established in 2002.

Additionally, the NPS Invasive Species Action Plan (National Park Service, 2006) includes specific, recommended actions ranging from leadership and coordination to restoration. This document and associated materials meet or help parks and networks to meet the guidelines and suggestions of the following actions from the plan:

- 1A.2: Develop NPS capability at a regional or multi-park level to help build and coordinate park invasive species programs, maximize existing efforts such as Exotic Plant Management Teams (EPMT), and integrate regional and local NPS efforts with State and local agencies and organizations.
- 1B.1: Expand partnerships to maximize results. The NPS will actively participate in local weed-management areas and regional and State panels on aquatic nuisance species or invasive species to foster coordinated invasive species management and education.
- 1B.2: Establish citizen-steward partnership coordinator programs in parks.
- 1B.4: Enhance national, regional, and State interagency coordination. It is important to seek opportunities to expand NPS involvement on Interagency Panels. Also, it is necessary to increase coordination with State and local invasive species committees such as Weed Management Areas, Exotic Pest Plant Councils, State invasive councils and boards.
- 1B.5: Identify mechanisms to work on land adjacent to a park in discretionary cooperative efforts.
- 1C.3: Rank invasive species for each park unit with significant invasive species concerns. Species will be ranked as to level of threat, invasion potential, and feasibility of control…. Additionally, areas of each park unit likely to be invaded shall be identified.
- 3A.1: Implement a system for reporting and rapidly communicating new occurrences of high-priority or other invasive species.
- 3A.3: Contribute to the development of national standards for all aspects of invasive species management.
- 6A.2: Improve the quality of the invasive species data in NPSpecies.
- 6A.3: Improve the quality of the invasive species data in NR–MAP.
- 8A.5: Work with nursery and pet trade industry to promote responsible ownership and marketing of noninvasive species.

The action items that apply to a specific natural area will be dependent upon the explicit management goals and monitoring objectives detailed by a park, network, or other land-management organization.

Handbook Purpose

This document provides guidance for natural resource managers wishing to detect invasive plants early through an active, directed monitoring program. Currently several regional electronic mail notifications track new species occurrences, and conceptual models and limited documentation provide managers with broad-scale frameworks for rapid response systems triggered after a species has been detected. The Invasive Plant Atlas of New England (IPANE) provides an example of a database and clearinghouse for occurrence data from the Northeastern United States (*http://www.eddmaps.org/ipane/earlydetection/species_scientific.htm*). As with other regional efforts, interested parties may choose to receive invasive plant alerts through the database system or visit the site periodically. Both the National Invasive Species Council (*http://www.invasivespecies.gov/global/EDRR/EDRR_index.html*) and the Federal Interagency Committee on the Management of Noxious and Exotic Weeds (*http://www.fws.gov/ficmnew/FICMNEW_EDRR_FINAL.pdf*) provide excellent frameworks for rapid response systems once a species has been detected. No literature, however, details the steps required to successfully implement an early detection protocol (that is, how one actively searches for new species or new populations of existing species given typical constraints).

We recognize that locally specific detection strategies depend upon management priorities, monitoring objectives, available data sets, and available resources. Therefore, we do not attempt to prescribe a single methodology to fit all eventualities. Rather, we recognize a series of choices that collectively need to be considered when developing a comprehensive early detection program. We provide information pertinent to each of these decision points to aid managers in making informed choices tailored to their individual circumstances. How monitoring objectives and available data sets influence the choice of analytical tools is articulated in this book. Details on how to use a particular statistical tool, however, can be found elsewhere, and the reader is directed toward relevant citations and resources for further consultation. We also recognize that some managers will require guidance only on specific steps within the early detection process. Consequently, we have designed the document such that each chapter may stand alone or be integrated with the other chapters to form a comprehensive approach to early detection of invasive plants. Chapters that present options or that require decisions to be made with respect to available data and resources (for example, choosing among predictive modeling approaches) include a decision tree diagram to assist users in making appropriate choices given their relevant circumstances.

The reader may choose to approach the development of early detection procedures in a linear, comprehensive fashion. If this is the case, we suggest following this document in sequential order and following the logical progression of steps involved in the early detection process (see Chapter 3). It may be that the reader requires guidance only on specific steps in the early detection process, though. In this case, the reader may wish to go directly to the chapters highlighting the relevant details. In either case, we suggest that the reader review the early detection conceptual model presented in Chapter 3 and then scan the decision trees and overviews associated with each chapter to gain context and a better perspective of the entire process. Particular attention should be given to Chapter 4, which discusses formulating management goals and monitoring objectives. If goals and objectives have not been well defined, it is likely that any monitoring program will provide less than desirable results in the long term. A review of the case studies will add a practical perspective to the development of early detection procedures.

Chapters 1–3 introduce the text, summarize dominant invasion theories relevant to this text, and outline the key components required to implement a successful early detection monitoring program. Chapters 4–10 address each of the early detection steps in detail, including setting goals and objectives, acquiring appropriate information, choosing analytical processes, conducting evaluation and assessment, and implementing detection methods. Chapters 11 and 13 provide applications of early detection principles within the context of the Klamath and San Francisco NPS Inventory and Monitoring (I&M) Networks. One chapter presents information from the Klamath Network: Chapter 11 shares insights with respect to early detection protocol development across multiple parks. The San Francisco Network presents a small park example using volunteers (Chapter 13). An example using remotely sensed and geographically referenced data for predicting the risk of occurrence for target species in Big Bend National Park is presented in Chapter 12. Subsequent materials include citations and a glossary. Cited materials and sources of additional information appear in the "References Cited" chapter and are grouped by chapter.

This text is intended to be a living document on the internet that will be updated regularly as new materials become available. In particular, we anticipate specific protocols developed by individual I&M networks will be linked to this document to provide readers with a range of specific implementation examples. Although this document has been designed with the NPS I&M networks and parks in mind, it undoubtedly will have broad application for natural resource professionals everywhere who want to improve invasive plant-management strategies.

Definitions

Executive Order 13112 though there are others—(*http://www.invasivespeciesinfo.gov/laws/execorder.shtml*) established the National Invasive Species Council (NISC) and defined the term "invasive species" for all Federal agencies. Accordingly, an "invasive species" is an alien species whose introduction causes or is likely to cause economic or environmental harm or harm to human health (U.S. Presidential Executive Order, 1999). Executive Order 13112 further defines an "alien species" as, with respect to a particular ecosystem, any species, including its seeds, eggs, spores, or other biological material capable of propagating that species, that is not native to that ecosystem (U.S. Presidential Executive Order, 1999).

Determining what constitutes "harm" or which species are likely to cause harm is a difficult undertaking and can be subjective. No measure of harm is available for most nonnative and (or) invasive species. Where measures of harm (or impact) exist, there are no corresponding measures of benefit for appropriate evaluation. Nevertheless, we have chosen a definition for invasive species that includes a reference to harm. We have done so because Federal policy directs the National Park Service and other agencies to adopt the National Invasive Species Council definition under Executive Order 13112 and because many different terms related to invasive species are currently used interchangeably and, based on the literature, truly are not equivalent (see Glossary). Terms associated with invasive species ecology and management, as with any ambiguous terms, have multiple pathways for interpretation depending on the audience being addressed and, therefore, are not interchangeable. The definitions adopted by an agency or organization influence communication with various segments of the public who provide the political will that fuels policy, management, and funding sources. For those with the freedom to do so, we advise that invasion terminology be defined clearly, with various audiences in mind, and in accordance with the legislation, mission statement, scientific theory, and (or) management practice associated with one's agency or organization.

Unless otherwise noted, we will define invasive species in accordance with Executive Order 13112. We are substituting the term "nonnative" for "alien" in an attempt to provide a more neutral contrast to the term "native" and to limit negative connotations associated with the term “alien.” We define nonnative species as those that are not naturally occurring in an ecosystem but are capable of living and reproducing in the wild without continued human agency, including invasive species and genetically modified native species (Odell, 2004).

Please note that we consider the early detection of invasive plant species to be an iterative process that is updated and improved after each iteration through updated information and revised objectives (see fig. 3.1). In this sense, we consider early detection to be a long-term monitoring process (that is, “…a collection and analysis of repeated observations or measurements to evaluate changes in condition and progress toward meeting a management objective….” Elzinga and others, 1998). Alternatively, one can view early detection as an inventory or survey technique. Each visit to a site is considered a distinct event, especially if rapid response procedures have removed identified populations of target species. We take a broader, longer term perspective. We believe that the success of any monitoring program is contingent upon first completing a rigorous inventory or survey of the appropriate resources to establish a baseline from which to compare future monitoring and management activities. We also believe that successful implementation of an early detection monitoring program will lead to fewer established populations of target species in selected management units or sites over time. Just as important, information gained from early detection monitoring will allow managers to evaluate the success of their overall management strategies, build a better understanding of the primary vectors and pathways leading to invasion, and provide one measure of ecological integrity.

Recommended Reading

- A report by the Ecological Society of America outlines the state of invasive species management and research across the United States in the context of Federal guidance and policy. Recommendations are made to improve the current system. See Lodge and others (2006).
- In 1993, the U.S. Office of Technological Assessment conducted a thorough review of harmful non-indigenous species in the United States. The review covers research, management, policy, economics, ecological impacts, and so forth. Although it is quite lengthy, it contains much useful information, facts, figures, diagrams, and summary material. See Office of Technology Assessment (1993). The document can be downloaded at *http://www.princeton.edu/~ota/disk1/1993/9325/9325.PDF*.
- The National Park Service Management Policies are useful for putting invasive species management into the greater context of natural and cultural resource management across the Park Service. Subsections covering invasive species refer to subsections covering restoration, integrated pest management, and enhancement of native species and ecosystems. See National Park Service (2006) (accessed January 5, 2007 at *http://www.nps.gov/policy/MP2006.pdf*).
- The USDA National Invasive Species Information Center, accessed January 5, at *http://www.invasivespeciesinfo.gov/laws/main.shtml* has a good directory with links to relevant laws and policies that apply to invasive plants (and other taxa).

References Cited

Asher, J.A., and Harmon, D.W., 1995, Invasive exotic plants are destroying the naturalness of U.S. wilderness areas: International Journal of Wilderness, v. 1, p.35–37.

Beard, R., Searle, C., Bruno, D., Rife, J., Cline, D., and Mullen, B., 2001, Weed mapping standards: Granby, Colo., North American Weed Management Association, accessed April 13, 2014, at *http://www.naisma.org*.

Benjamin, P., and Hiebert, R., 2004, Assessing the invasive plant issue: Park Science, v. 22, p. 27–31.

Carlquist, S., 1974, Island Biology, New York: Columbia University Press, 660 p.

D’Antonio, C.M., and Mahall, B.E., 1991, Root profiles and competition between the invasive, exotic perennial, *Carpobrotus edulis*, and two native shrub species in California coastal scrub: American Journal of Botany, v. 78, p. 885–894.

Ehrenfeld, J.G., 2003, The effects of exotic plant invasions on soil nutrient cycling processes: Ecosystems, v. 6, p. 503–523.

Elzinga, Caryl L., Salzer, D.W., and Willoughby, J.W., 1998, Measuring and monitoring plant populations: Bureau of Land Management Technical Reference 1730–1, BLM/RS/ST–98/005+1730.

Fosberg, F.R. 1948, Derivation of the flora of the Hawaiian Islands, in Zimmerman, E.C., ed., Insects of Hawaii: Honolulu, University of Hawaii Press, v. 1, p. 107–119.

Hiebert, R., ed., 2002, Guidelines for inventory and monitoring of invasive plants—Report of the workshop on invasive plant inventory and monitoring protocols: Fort Collins, Colo., National Park Service, 59 p.

King, W.B., 1985, Island birds——Will the future repeat the past?, *in* Moors, P.J., ed., Conservation of island birds: Cambridge, United Kingdom, International Council for Bird Preservation, Cambridge University Press, p. 3–15.

Lodge, D.M., Williams, S., Macisaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, B., Smith, M., Andow, D.A., Carlton, J., and McMichael, A., 2006, ESA Report—Biological invasions, recommendations for U.S. policy and management, Ecological Applications, v. 16, p. 2035–2054.

Marler, M., 1998, Exotic plant invasions of federal wilderness areas—Current status and future directions: Missoula, Mont., The Aldo Leopold Wilderness Research Institute, Rocky Mountain Research Station, 9 p.

National Park Service, 2006, Management policies 2006: Washington, D.C., U.S. Department of the Interior, accessed January 5, 2007, at *http://www.nps.gov/policy/MP2006.pdf*.

National Park Service, 2006, Invasive species action plan—Final draft: Fort Collins, Colo., National Park Service.

National Research Council, 2002, Predicting invasions of nonindigenous plants and plant pests: Washington, D.C., National Academy of Sciences, 198 p.

Odell, T., 2004, Invasive species control—A comprehensive model state law: Washington, D.C., Environmental Law Institute.

Office of Technology Assessment, 1993, Harmful non-indigenous species in the United States–OTA–F–565: Washington, D.C., U.S. Congress, Government Printing Office.

Pimentel, D., Zuniga, R., and Morrison, D., 2005, Update on the environmental and economic costs associated with alien-invasive species in the United States: Ecological Economics, v. 52, p. 273–251.

Rejmanek, M., 2000, Invasive plants—Approaches and predictions: Australian Journal of Ecology 25, p. 497–506.

Rejmanek, M., and Pitcairn, M.J., 2002, When is eradication of exotic pest plants a realistic goal?, *in* Veitch, C.R., and Clout, M.N., eds., Turning the tide—The eradication of invasive species: Gland, Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, Species Survival Commission Invasive, Species Specialist Group, p. 249–253.

Rejmanek, M., and Randall, J.M., 1994, Invasive alien plants in California—1993 summary and comparison with other areas in North America: Madrono, v. 41, p. 161–177.

Smith, H.A., Johnson, W.S., Shonkwiler, J.S., and Swanson, S.R., 1999, The implications of variable or constant expansion rates in invasive weed infestations: Weed Science, v. 47, p. 62–66.

Tilman, D., 1999, The ecological consequences of changes in biodiversity—A search for general principles: Ecology, v. 80, p. 1455–1474.

U.S. Presidential Executive Order, 1999, Executive Order 13112 of February 3, 1999, Federal Register: v. 64, no. 25.

Vitousek, P.M., 1990, Biological invasions and ecosystem processes—Towards an integration of population biology and ecosystem studies: Oikos, v. 57, p. 7–13.

Wilcove, D.S., Rothstein, D., Dubow, J., Phillips, A., and Losos, E., 1998, Quantifying threats to imperiled species in the United States: Bioscience, v. 48, p. 607–615.

Williamson, M., 1993, Invaders, weeds and risk from genetically modified organisms: Experientia, v. 49, p. 219–224.

Williamson, M., and Fitter, A., 1996, The varying success of invaders: Ecology, v. 77, p. 1661–1666.

Plant Invasion Process—Implications for Land Managers

By Bradley A. Welch

Chapter 2 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figures

Insets

Chapter 2.

Plant Invasion Process—Implications for Land Managers

By Bradley A. Welch

Box 2.1. "Control of biotic invasions is most effective when it employs a long-term, ecosystem-wide strategy rather than a tactical approach focused on battling individual invaders" (Mack and others 2000).

Overview

Understanding the invasion process and basic attributes correlated with invasibility can inform the strategic planning process. This knowledge is invaluable for setting management goals and monitoring objectives, prioritizing target species and search areas, and selecting and implementing search strategies for early detection. From a practical standpoint, for example, we understand that costs of managing invasive plants increase exponentially as invaders become more entrenched (fig. 2.1). These costs increase to the point where management actions become cost-prohibitive once populations have fully invaded an ecosystem. Costs—measured in dollars or degree of impact—to ecosystem components and processes resulting from invasion also increase substantially over time, making ecosystem restoration improbable in the latter stages of invasion. Hence, experienced natural-areas managers and scientists have stressed the need to detect nonnative plants early and to respond rapidly before populations become invasive. Additionally, early detection efforts are most successful if they are integrated into a long-term, ecosystemwide strategy rather than a tactical approach focused on managing individual species (Mack and others, 2000).

Defining management goals, deciding where to look, deciding how to look, choosing which existing and potential threats should be managed, and addressing each species effectively are daunting tasks to undertake. This chapter outlines the key stages in the invasion process and summarizes the basic attributes associated with invasibility. Both sets of principles are linked to management implications where relevant.

Stages of Invasion

Like native species, invasive species must grow, reproduce, disperse propagules, and colonize new territory to perpetuate themselves. Invasive species must overcome geographic, resource, environmental, and dispersal barriers before they truly become problematic (fig. 2.2). Consequently, only about 10 percent of nonnative species arriving in natural areas become established (Williamson and Fitter, 1996) and about 10 percent of established nonnative species disperse widely enough to survive stochastic environmental events. Of these, approximately 10 percent significantly affect natural areas (Kowarik, 1995). The invasion process, then, is best understood by correlating barriers to invasion with plant population biology and dividing these relations into stages: introduction, establishment, lag time, spread, and invasion (Groves, 1986; Cousens and Mortimer, 1995; Rejmanek, 2000; Richardson and others, 2000). We include establishment and spread as separate stages from invasion even though there can be overlap among these stages. We can be more complete in our coverage and articulate the relevance of these stages to management options by doing so. Similarly, we discuss lag time (that is, the time between initial establishment and initial spread) separately because this latent period provides a window in which nonnative populations can be detected and can still be cost-effective to manage (Lodge and others, 2006).

There is much debate over the accepted nomenclature and partitions that segment the invasion process, but scientists generally recognize similar barriers at each stage which must be overcome for a species to move on to the next level of biological success (fig. 2.2). Barriers are reversible: climate change may alter the ability of a species to compete, native species may recover, or new resources may become available. Likewise, the invasion process is not necessarily smooth. Richardson and others (2000) and Williamson (1996) offer the reviews and illustrations of the invasion process, each offering a slightly different perspective.

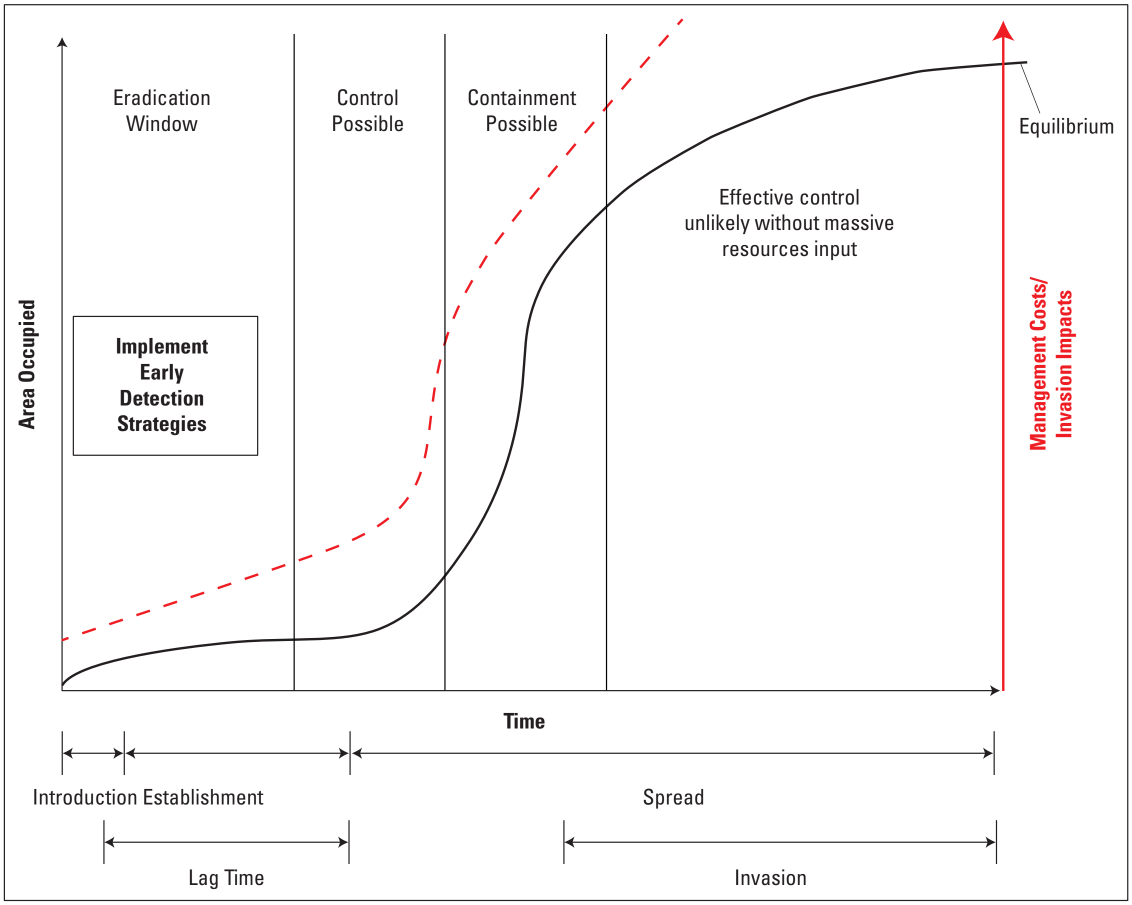

Figure 2.1. Relation between biological stages of invasion and associated management costs and (or) invasion impacts. Appropriate management strategies are shown to illustrate the relevance of the invasion process to management options. As invasive species become more entrenched, the costs of managing them become significantly greater, to the point of being cost-prohibitive. Eradication = total elimination of all plants and associated plant material, control = partial elimination of new and existing invasive plants and associated material, and containment = limiting the spread of invasive plants to existing populations (also see Glossary; after Chippendale, 1991; Naylor, 2000; and McNeely, 2001).

Introduction

Plants first have to overcome geographic barriers before becoming successful invaders (fig. 2.2). Both intentional (for example, horticulture, agriculture, erosion control) and inadvertent means (for example, ballast water, seed mixes, freight) facilitate transport of nonnative plant material across obstacles such as mountains, oceans, or significant distances in a short period of time. Consequently, introductions (that is, the transport and subsequent arrival of nonnative plant material) are often to accessible habitats by transportable species (Williamson, 1996). Most arrivals to new areas result from human importation, though movement of propagules by way of natural vectors and pathways (for example, wind, water, frugivorous birds, mammals) contributes to the introduction of incipient populations within a given region or locality (Office of Technical Assessment, 1993; Williamson, 1996). Unfortunately, new introductions occur continuously, making invasion imminent to any given conservation area. Propagule pressure (that is, the number of individuals of a species that is released) improves the likelihood that propagative material will be introduced to natural areas downstream or downwind from existing source populations. Propagule pressure is closely correlated with settlements and roads, further emphasizing the effect of human activity in introducing nonnative plant materials to new localities even along seemingly natural pathways (Timmins and Williams, 1991). Anthropogenic and natural modes of introduction, including propagule pressure, should be evaluated in and around natural areas as part of any early detection program.

Establishment

Once a species has arrived, it must encounter favorable environmental factors (for example, available propagation sites, adequate resources, and limited predation) to establish self-perpetuating populations (fig. 2.2). Stochastic events such as flooding, drought, herbivory, and frost, which prove

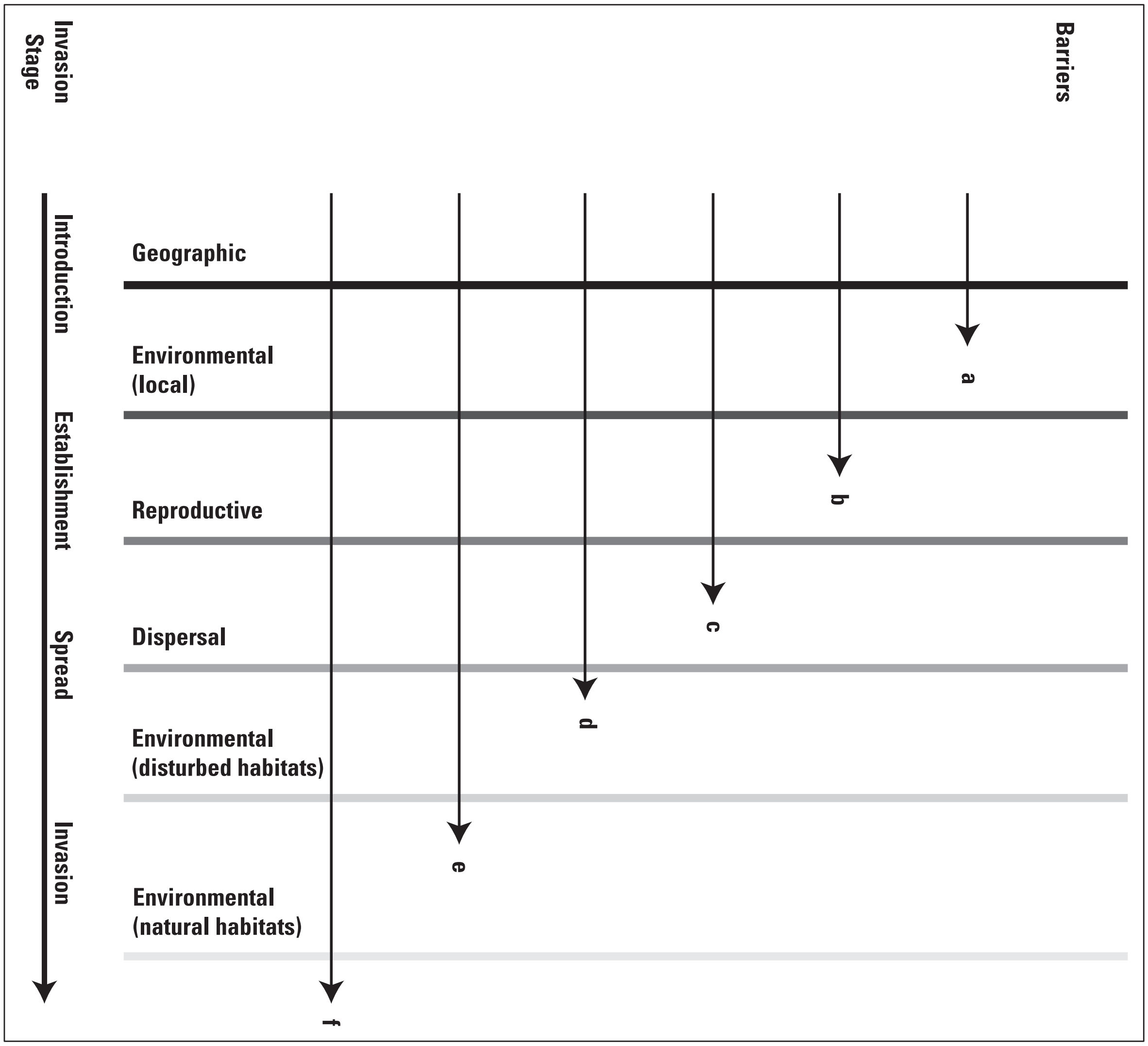

Figure 2.2. Major barriers limiting the introduction, establishment, and spread of nonnative plants. The barriers are: a. Major intercontinental and (or) intracontinental geographical barriers; approximate scale is greater than 100 kilometers; b. Abiotic and biotic environmental barriers at the site of introduction; c. Reproduction barriers that prevent consistent and long-term vegetative and (or) generative production of offspring; d. Local/ regional dispersal barriers; e. Environmental barriers in human-modified or alien-dominated vegetation; and f. Environmental barriers in natural or seminatural vegetation. Arrows a through f indicate the paths followed by taxa to reach different states from introduced to invasive in natural vegetation. Crossing of the barriers is reversible. For example, climatic fluctuations can either pose new barriers (which could drive alien taxa to extinction at local and (or) regional scales), or enable the taxon to survive or spread. (Adapted from Richardson and others, 2000)

favorable to some nonnative plants later in the invasion process, often hamper the establishment of most introduced species (Hobbs and Huenneke, 1992; Mack, 1995 and 2003). Small population size may lead to a genetic bottleneck, which also squelches establishment (Williamson, 1993). Despite these hurdles to establishment, the continuous barrage of different nonnative species ensures that at least one species eventually will encounter appropriate conditions for establishment in a given area. Similarly, broad-scale and (or) repeated introductions of the same species increase its chances of establishment (Perrings and others, 2002). Marketing of ornamental species through nursery sales, the aquarium trade, or use for erosion control, for example, may lead to broad public acceptance of nonnative species. The popularity of the species increases its chances of establishment in residential areas and its subsequent establishment in conservation areas (Ensernick, 1999). Marketing pathways should be considered in any early detection protocol. Note that Richardson and others (2000) refer to species that require repeated introductions to sustain populations as "casual" species, whereas others (Williamson, 1996; Sharma and others, 2005) include establishment as part of the introduction phase.

Established species may exist in small populations for long periods, making detection difficult. It is at this stage, though, that eradication is most probable and cost-effective (fig. 2.1). Predicting the location of these populations can be accomplished through climatic and habitat matching because the primary factors driving success at this stage are local environmental conditions (Williamson, 1996; Richardson and others, 2000). Habitat matching, of course, is reliant upon intimate knowledge of species' habitat requirements, which for newly introduced species may not be fully understood. Sifting through a plethora of research and management reports to find key life-history traits for each target species is an overwhelming task for anyone. For this reason, several agencies and organizations have summarized relevant information for an increasing number of nonnative species invading a variety of ecosystems. Collections of species' profiles can assist managers in quickly pinpointing critical life-history characteristics of target species.

Sources of species profiles include:

- The Nature Conservancy—*http://www.imapinvasives.org/GIST/ESA/index.html/*.
- USGS Invasive Species Program—*http://www.usgs.gov/ecosystems/invasive_species/*.
- Global Invasive Species Programme—*http://www.issg.org/database/welcome/*.
- Hawaiian Ecosystems at Risk (HEAR)—*http://www.hear.org/*.
- Center for Aquatic and Invasive Plants (University of Florida)—*http://plants.ifas.ufl.edu/*.

Lag Time

Contrary to popular perception, most nonnative plant species do not arrive in a new area and immediately blanket it as an impenetrable mass. Rather, many nonnative species experience a time lag between introduction and rapid expansion (Binggeli, 2001). An understanding of lag time is relevant to early detection of new invasions and for planning management strategies primarily because it underscores the need to closely monitor potential invaders that have not yet become problematic. Selecting and prioritizing sites for monitoring will depend on an understanding of the variation between a potential invader's new and native habitat, its history of immigration, and the anthropogenic and abiotic disturbances that may release it from its lag phase. Much of this information is contained in species' profiles found at the electronic links previously listed.

Lag times have been observed across the globe in a variety of ecosystems and range from 3 years for woody tropical shrubs to over 170 years for temperate trees (Binggeli, 2001). Hobbs and Humphries (1995) suggested that lag times are often associated with a sudden or rapid spread of nonnative species rather than being associated with a gradual increase in population size over time. This "sudden release," whether real or perceived, results because (1) genetic adaptations allow the species to overcome environmental barriers, (2) environmental barriers to invasion are removed by disturbance, stress, or biotic interaction with other species, or (3) observation and monitoring techniques fail to detect population growth at an early invasion stage.

Box 2.2. A **lag time** is often observed between initial introduction of a nonnative species and subsequent rapid spread. The rapid expansion of a nonnative species after a perceived lag period may result from:

- genetic adaptations,
- a combination of environmental factors including disturbance, or
- poor observation and monitoring techniques on the part of the observer (Hobbs and Humphries, 1995).

Himalayan balsam (*Impatiens glandulifera*), for example, has become a successful invader in Europe and, more recently, in the United States as a result of genetic adaptation following a lag period (Kollmann and Bañuelos, 2004). Strong evidence supports the hypothesis that disturbance and other environmental factors release nonnative species from time lags, too. Flooding, hurricanes, fire, deforestation, and drought have all led to large gap formations in regions where nonnative species existed but did not spread until after the disturbance. Note that stochastic events such as these are likely to hinder establishment of nonnative species early in the invasion process but can promote spread in postestablishment stages of invasion (see Invasion and fig. 2.2). Such was the case for beach sheoak (*Casuarina equisetifolia*) in Florida following two hurricanes in the 1980s (Binggeli, 2001). Similarly, significant changes in grazing regimes in Hawaii led to rapid expansion of gorse (*Ulex europaeus*) (Binggeli, 2001).

Quite often, "time lags" are perceived to have occurred because observations were missed or populations of species introduced for agricultural, horticultural, or other economic reasons were insufficiently monitored. For example, velvet tree (*Miconia calvescens*) was detected in Tahiti 30 years after it was introduced and had already become widespread. It was detected in remote areas of Hawaii in unexpectedly large populations only after officials had received warnings from French Polynesia (Conant and others, 1997). If for no other reason, an understanding of the factors contributing to lag time is necessary in that it reinforces the need to be ever-vigilant in the early detection and monitoring of potential invasive plant species even if they appear benign.

Spread

Box 2.3. Management thresholds can be used to screen high priority invasive plant species from lower priority nonnative species. As a general rule, nonnative species can be considered a priority if their dispersal distance per unit time exceeds:

- 100 meters in less than 50 years for seed dispersed species, or
- 6 meters per year for plants spreading vegetatively

(Richardson and others, 2000).

A population of nonnative plants that sustains itself over several life cycles without human intervention is considered established or, in some cases, naturalized (Sharma and others, 2005). The likelihood of eradication of the population as a result of an environmental stochastic event is low at this point, and dispersal of propagules away from the parent population begins (fig. 2.2; Richardson and others, 2000). The principal barriers, if any, to the spread or colonization phase are typically dispersal limitations. Limitations are determined in part by the life-history traits of a given species as applied to population dynamics (for example, vegetative reproduction, bird dispersal, dormancy requirements), though spread can be in any direction and at any speed (see Williamson, 1996). Note that the rate of spread has been positively correlated with the abundance of a species in its native range, its global distribution, degree of taxonomic isolation, degree of habitat matching, availability of human-made structures facilitating spread (for example, roads, bridges, canals), the presence of other nonnative species (that is, facilitation, as discussed by Connell and Slatyer 1977), and the degree of habitat alteration. Approximating the rate of spread through these surrogates can assist managers in separating high priority nonnative species from lower priority species.

At this stage, management strategies should be used to stifle the dispersal of nonnative species to new areas, targeting vectors and pathways—especially anthropogenic mechanisms—that are under some degree of management control. A “slow-the-spread” strategy is a rational management choice to augment local control efforts, particularly when the environmental or economic costs of allowing an invader to proceed unmanaged are likely to outstrip management costs. For each unit of time during which an invader is prevented from occupying new range, a benefit accrues (for example, current efforts to stop the spread of zebra mussel (*Dreissena polymorpha*) and other freshwater invasive species to water bodies of the Western United States) (Lodge and others, 2006). While control programs for widespread species are expensive, they may still be cost-effective when important natural resources are at risk.

Invasion

Ecologically speaking, an established nonnative species that founds new self-perpetuating populations, undergoes widespread dispersal, and becomes incorporated within the resident flora is considered invasive (Richardson and others, 2000). Richardson and others (2000) also suggest that nonnative species should not be considered invasive unless propagules—particularly in large numbers—have been dispersed from the parent plant a distance of more than 100 meters in less than 50 years for seed dispersed species and more than 6 meters per 3 years for plants spreading vegetatively. Note that these guidelines may be useful for separating invasive nonnative species from noninvasive nonnative species, but they are not intended to separate nonnative species from native species. Many native species spread faster than these rates. Likewise, it is the characteristic formation of new self-perpetuating populations and widespread dispersal that set the invasion stage apart from the establishment and spread stages. While these figures are somewhat arbitrary, they appear to have practical utility, at the least, as a management threshold to aid screening and prioritization procedures.

In addition to overcoming dispersal barriers, Richardson and others (2000) suggest that for established and spreading species to be considered invasive, they must adjust to a broad range of biotic and abiotic pressures allowing them to invade both disturbed and successionally mature communities (fig. 2.2). Some evidence suggests successionally mature, undisturbed communities, especially productive areas, are as susceptible to invasion as disturbed, early successional communities (Williamson, 1996; Stohlgren and others, 1999). Hence, the last barrier in the Richardson and others (2000) model may not distinguish well between invasive and noninvasive nonnative species. In either case, eradication at this stage with limited resources is virtually impossible. Management strategies should focus primarily on limiting spread and monitoring and containing known populations.

Equilibrium

Eventually, the rate of expansion of a new species into suitable habitat slows as the species reaches a state of dynamic equilibrium with its invaded environment (Williamson, 1996; fig. 2.1). Invasive species reach equilibrium because they occupy all suitable habitat, native plant communities recover and are competing effectively, biotic and abiotic pressures are limiting growth, management and control strategies are effective, long-term environmental conditions (for example, climate change) are less suitable for the invader, or a combination of these factors is limiting the rate of increase. Many common plant species in habitats of the Northeastern United States have long since reached equilibrium (for example, dandelion [*Taraxacum officianale*], annual bluegrass [*Poa annua*], and ground ivy [*Glechoma hederacea*]). Nonetheless, the long-term costs and effects incurred from invasive species by the time they reach equilibrium far exceed the benefits of full-scale control programs except in areas where they have been excluded.

Key Predictors of Invasiveness

The study of invasion biology has its relevance to management because it informs the risk-assessment process used to prioritize among the myriad potential and existing invaders threatening park resources. While some natural resource professionals prioritize threats based on perceived risk, most attempt to use documented predictors of invasiveness to inform their assessment methods. Unfortunately, no single rule or set of rules can be applied uniformly to characterize accurately all successful invasions. The range and variation in life-history strategies associated with species that have successfully invaded different ecosystems under varying conditions are too vast given our current understanding of plant invasions.

Nevertheless, invasion biologists have defined a number of attributes linked to successful invasions that can be used individually or in combination as predictors of invasiveness. Both species-based (biological) attributes and site-based (environmental) attributes may confer invasiveness. Natural resource professionals can use a combination of species attributes and ecosystem traits and an understanding of local stochastic events to formulate a risk-assessment tool to help filter likely invaders from unlikely invaders (Davis and others, 2000; Mack, 2003; Chapman and others, 2001). Similarly, environmental conditions may be defined under which invasions by a subset of species are likely to occur (for example, guilds and genera). These predictors should be applied to the risk assessment (see Chapters 6 and 7) and modeling processes (see Chapter 9) with specific species and specific ecosystems in mind. Generalized rules applied to nonspecific circumstances will not separate high-risk species or sites from low-risk species or sites to any appreciable extent (see previous paragraph). The following information summarizes dominant invasion hypotheses and associated predictors that may be used to define a local recipe for invasion.

Species-Based Attributes

Along with initial population size and number of introduction attempts (Rejmanek, 2000), researchers suggest that invading species possess a number of attributes conferring invasion success. Predominant factors are described below.

Invasive History

The simplest and most pragmatic generalization drawn from invasion biology is that a species is likely to become invasive in a new habitat if it has a prior history of invasion elsewhere (Rejmanek, 2000; National Research Council, 2002). This rule can be applied to national and regional risk assessment models as a broad-spectrum filter. It also can be used locally. Nonnative plants that have successfully invaded similar ecosystems in surrounding landscapes are likely to find suitable habitats within the target management area. For example, the invasive history and devastating impact of velvet tree on the flora of Tahiti triggered early detection reconnaissance in Hawaii (Conant and others, 1997).

Fitness Homeostasis and Geographic Range

Species that maintain relatively constant fitness over a range of environments (fitness homeostasis) are likely to overcome environmental constraints and (or) climatic changes in the introduced range relatively quickly (Rejmanek, 2000). Similarly, species with a wide distribution and abundance across habitat types in their native range are more likely to find suitable environmental conditions in the introduced range than species with a narrower niche breadth (Williamson, 1996; Rejmanek, 2000). A broad native range also improves the likelihood of propagule distribution to new territories. Fitness homeostasis is difficult to quantify, though, and some species occupy broader habitat ranges in introduced regions than in their native range (Rejmanek, 2000). Common reed (*Phragmites australis*) occupies brackish and freshwater wetland ecosystems on every continent except Antarctica, ranging from the Tibetan plateau to the Nile River (Haslam, 1971). Its ability to invade new wetlands across the globe is purely a matter of transport.

Phenotypic Plasticity

Many organisms have the ability to alter their physical features, change growth rates, and produce chemical compounds in response to environmental change, both biotic and abiotic. Polygonum species, for example, alter their morphology in response to water depth, producing a floating-leaf morphology under submerged conditions and mud-flat morphology under drier conditions (Grime, 2001).

High Intrinsic Rate of Natural Increase

Plant species that are able to grow, reproduce, establish, and spread quickly are often successful invaders. Traits such as small seed size, prolific propagule production, short juvenile period, rapid growth rate, and high transpiration rates allow species such as purple loosestrife (*Lythrum salicaria*) and common mullein (*Verbascum thapsus*) to quickly dominate extant and potential native vegetation (Rejmanek and Richardson, 1996; Grime, 2001).

Competitive Ability

Some plant species inherently use resources better than other species. Tamarisk (*Tamarix* species) holds a competitive advantage over native shrubs of the Southwestern United States because tamarisk roots grow deeper and faster than those of native species, allowing the invader to draw the water table down below the roots of native shrubs (Kennedy and Hobbie, 2004). Other invaders successfully transform environmental conditions in the introduced habitat. Tamarisk also competes with native plant species through allelopathy, depositing excessive salts in upper soil layers and effectively limiting the growth and germination of native propagules (Kennedy and Hobbie, 2004). Cheatgrass (*Bromus tectorum*) has thrived throughout the Western United States, in part, because it has altered the preexisting fire regime to one more suitable to its life cycle (Brooks and others, 2004). Yet, other invasive species rely on traits such as high leaf area index (Grime, 2001), nonspecific mutualism (Rejmanek, 2000), physiological integration (that is, the ability to translocate nutrients and other resources from ramets living under ideal conditions to ramets living under suboptimal conditions by rhizomes and other connective plant tissue; Hara and others, 1993), and tolerance to low resource levels or high disturbance levels (Grime, 2001) to outcompete native plants.

Reproductive Strategies

Despite the insurance afforded by multiple reproductive strategies, vegetative reproduction tends to increase habitat compatibility and increase invasive success with latitude and in aquatic systems (Rejmanek, 2000).

Effective Dispersal Mechanisms

In the Eastern United States and on the Pacific Islands, animals, in particular frugivorous birds, serve as the predominant dispersers of nonnative plant propagules (Rejmanek, 2000). In the Western United States, however, dispersal of successful invasive plants tends to be by way of physical mechanisms (for example wind and water).

Many other traits and mechanisms have been correlated with invasion potential. Rejmanek (2000), Grime (2001), National Research Council (2002), Mack (2003), and Sharma and others (2005) offer more complete reviews of these attributes.

Community- or Site-Based Attributes

While all communities are invasible, some may be more invasible than others or may become more susceptible to invasion under a given set of conditions (Williamson, 1996; Davis and others, 2000; National Research Council, 2002). Certain environmental factors increase the likelihood that an invasion will be successful. In particular, human modifications of the environment (for example, soil disturbance, habitat fragmentation, climate change, atmospheric nitrogen deposition, increased carbon dioxide) undeniably exacerbate invasion susceptibility. The following attributes are frequently associated with sites susceptible to invasion.

Novel Life Forms and Vacant Niches

No plant community possesses all the possible plant taxa, dispersal mechanisms, or life forms existing in the plant kingdom (Bell, 1991). Geographic isolation, environmental forces, and evolution over geologic time have restricted floristic diversity in any given area. Novel species now overcome these barriers relatively quickly with human assistance. Unique tree species have invaded undisturbed grasslands. Forests have been invaded by novel herbaceous and shrub species. For example, pricklypear (*Opuntia monacantha*) was introduced by humans to Yunnan, China, where its life-history traits are well suited to the environment and no succulent species existed previously (Mack, 2003). It now dominates the native flora with which it has not co-evolved. Many other examples support this hypothesis as well (see Mack, 2003).

Nonnative species are problematic on oceanic islands and in other isolated ecosystems where the native floral communities are often considered vulnerable to invasion (Elton, 1958). The vacant niche hypothesis suggests that the reason for this vulnerability is a lack of biological resistance on the part of the native plant community (Simberloff, 1986, 1995). Put differently, communities are most resistant to invasion when all possible mechanisms for resource uptake are being used by native species. Little research exists to support this hypothesis (see Simberloff, 1995, and Williamson, 1996).

Release from Biotic Constraint

Plants and plant propagules frequently arrive in an introduced range unaccompanied by specialist diseases, herbivores, or parasites from the native range. As a result, local barriers to establishment, reproduction, and subsequent spread are less restrictive to such nonnative plants allowing the invader to flourish. Scotch broom (*Cytisus scoparius*), which originates from the British Isles and Europe, lacks specialist herbivores in the Pacific Northwest though the number of generalist herbivores is comparable between the two geographic regions (Memmott and others, 2000). It may take some time for specialist species in the introduced range to "recognize" the new species as a potential host or food source by which time the nonnative species has successfully invaded. This hypothesis provides the basis for biological control efforts and the choice of species for agricultural production (Mack and others, 2000).

Disturbance and Stress

Human-induced disturbance (for example, fire, substrate removal, grazing) and stress (for example, flooding, thermal effluent, salinity, atmospheric nitrogen deposition) have been implicated as primary pathways for invasion across plant community types (Harper, 1965; Pickett and White, 1985; Mack, 1989; Hobbs and Huenneke, 1992; Office of Technical Assessment, 1993; Grime, 2001). Native species are adapted to fluctuations in the magnitude, frequency, and duration of local disturbance and stress regimes. Humans often alter disturbance cycles substantially or introduce new disturbances or stresses. Consequently, introduced species preadapted to these new conditions may thrive while native species may be geographically limited, weakened, or eliminated. Disturbance also may create gaps in the vegetation, allowing introduced species to establish founder populations from which they can spread through novel or superior life-history traits.

Resource Availability

A variant on the disturbance and stress hypothesis is the hypothesis of resource availability. Grime (1979, 2001) asserted that increased availability of resources (for example, light, water, nutrients) favors competitively dominant species such as invasive species—species that normally are kept in check by moderate disturbance or stress. Similarly, Davis and others (2000) suggested that all plant communities are susceptible to invasion but are most susceptible when they experience a period of resource fluctuation. The fluctuation may be an increase in available resources favoring dominant or invasive species, or resources may decline, limiting productivity and competitive ability of the native species. Stohlgren and others (1999) found a significant positive correlation between native community species richness and invasive species richness, suggesting that richer communities have more resources available for introduced species to exploit. Impoverished communities have fewer available resources. Introduced species must be present in the extant vegetation or as a propagule source for resource availability to be a viable invasion mechanism (Davis and others, 2000).

Diversity

Based on the food web theory, co-evolution theory, and vacant niche hypothesis, some ecologists postulate that communities with low species richness are more susceptible to invasion than species-rich communities (Elton, 1958; Tilman, 1997). Low species richness implies that niches are available for occupation. As mentioned previously, an alternative view considers species-rich communities to be more susceptible to invasion because they have more potential and available resources to be exploited (Stohlgren and others, 1999). Others suggest that high species diversity (species richness and abundance; Sala and others, 1996), complex structural diversity (Tilman and others, 1997), or high functional diversity (Mack, 2003) may be better predictors of community resistance to invasion. Relations between diversity and invasion also may be scale-dependent. In general, small-scale studies show a negative exponential relation between species richness and invasion, whereas large-scale field studies indicate that species-rich habitats are more susceptible to invasion (Sharma and others, 2005). It is also possible that no relation exists between diversity and invasion potential.

There is much debate associated with this hypothesis. Managers should, therefore, be cautious when employing this hypothesis for management purposes. It would be unwise to assume that species-rich communities or diverse ecosystems are immune to invasion.

Invasional Meltdown

Connell and Slatyer (1977) proposed three mechanisms of change in plant communities, one of which was facilitation. Facilitation involves species arriving in a new environment, modifying conditions, and making environmental conditions more suitable for successive colonizers. Simberloff and Von Holle (1999) suggested that in an invasive species context, invaders beget invaders. In some instances, facilitation may lead to a snowball effect, or invasional meltdown, which becomes a self-perpetuating cycle of invasions until native species are an insignificant proportion of the local flora. Species that transform nutrient cycles and alter disturbance regimes, such as the nitrogen-fixing firetree (*Myrica faya*) in Hawaii (Vitousek and Walker, 1989), and cheatgrass (Brooks and others, 2004), which alters desert fire regimes, are likely candidates for this hypothesis.

Vector and Pathway Analysis

Many new nonnative species tend to be introduced into communities through similar vectors (that is, the mechanisms of plant introduction) or along the same pathways (that is, the routes taken) as previous introductions (Ruiz and Carlton, 2003). Identifying the sources of repeated introductions, evaluating the risk associated with each, and interrupting the dominant pathways to invasion are the foundations of vector and pathway analysis in an applied setting. Vector analysis integrates concepts from other invasion hypotheses into a more applied, strategic framework. Vector analyses use several criteria—species-specific information (for example, life history characteristics), degree of site-susceptibility, potential invasion consequences, pathway magnitude (that is, degree to which a mode of transport contributes to the invasion potential of a target area), and probable frequency of pathway use—to prioritize among the predominant sources of current and future invasions (see Box 2.4 and Ruiz and Carlton, 2003). In this approach, management actions are directed toward turning off the faucet, so to speak, rather than trying to catch each drip. Limited resources are then maximized for the long term. Volunteer and partnership approaches currently are being used in and around parks in the Pacific Islands, for example. The goal is to survey increasingly distant areas from parks to stop invasive species before they establish anywhere on the island. Monitoring conducted inside the parks may be complemented by surveys performed outside park boundaries. Surveys start with an evaluation of the nursery trade to identify target species and proceed along roads and trails that are the significant modes of transmission for plant material. Subsequent surveys will work away from the parks along vectors and pathways in order of priority.

Box 2.4. Vector and pathway analyses use invasive species traits, community susceptibility, and evaluation of dominant vectors to prioritize management goals and objectives. Management action is directed toward the source of the invasions rather than the sink. Pathway analysis criteria (probability criteria) include:

- Pathway magnitude,
- Viability of organisms during transit,
- Likelihood of transmitting species difficult to detect/manage,
- Comparability of destination and original habitats,
- Ease of spread,
- Difficulty of control if established, and
- Consequence criterion—level of potential damage.

Invasion Theory and Management Applications

Why is invasion theory important to resource managers? Invasion theory points managers toward the types of information they need to collect for screening, prioritizing, modeling, searching for, and, ultimately, treating plant invasions. Despite the differences among theoretical camps, research helps define:

- what species to look for,
- where to look for them,
- how to look most efficiently, and
- what to do about them once they have been found.

Information needs vary with the goals, objectives, and priorities set by managers (see Chapter 4). The approach managers take to setting priorities may be based on species, sites, or both (fig. 2.3; see also Chapter 3). This approach assists in formulation of management goals and later in determining where and how management resources will be allocated. Once articulated, management goals determine which management strategy will be required and what monitoring procedures (articulated as monitoring objectives) best support the chosen goals and strategies. If the goal is to find small populations of invasive plants before they become established such that eradication is still feasible, early detection monitoring and rapid response strategies will be required as eradication is not feasible later in the invasion process. An understanding of the processes most relevant to each invasion stage of a targeted population aids in strategy selection (refer to fig. 2.1 and Chapter 5). We elaborate on the connection between monitoring objectives and management strategies in relation to the invasion process in Chapter 4.

The formulation of management priorities and goals, monitoring objectives, and management strategies from theoretical principles is not a static process or as linear as it appears (fig. 2.3). Priorities shift. Goals and objectives change in response to new priorities or as a result of new information obtained from monitoring and management activities. Information needed to set goals and objectives does not demand the level of detail required for modeling species or designing sampling regimes. Therefore, the entire process becomes iterative, improving with each repetition. This is particularly true in the context of early detection efforts that use predictive models to aid search strategies. Predictive models improve with the quality of available data.

Because early detection targets species before they become established, the important factors to consider are those that allow the species to be introduced to the area of concern and enable plants to then germinate and survive. Consequently, habitat matching and vector analysis are useful tools in the planning stages. These tools, in turn, require life-history information for each target species, including information on the native habitats of potentially invasive species, and environmental data pertaining to the targeted sites. Because the information needs associated with the list of potential species and sites can become overwhelming, management goals should specify whether priorities are driven by target species, areas of management concern, or both. This decision will make objectives realistic and achievable.

As a general rule for screening purposes, species that are invasive in one part of the world are likely to become invasive elsewhere. This criterion works well at multiple scales (for example, international, regional, local) as a crude risk assessment filter. It has limited utility as a means of prioritizing among the extensive list of species that will satisfy this criterion. Neither will it indicate where these species are likely to occur in a natural area. Unfortunately, there is no single, ubiquitous life-history trait that confers invasiveness for all species. Certain life-history traits allow species to dominate sites under some circumstances. Thus, life-history characteristics need to be compared among species during the prioritization process because some nonnative species have the potential to be more problematic than others. By knowing which characteristics are well suited to local environmental attributes, natural resource professionals can focus efforts on a subset of the list of all possible invaders.

All sites are susceptible to invasion, but some are more susceptible than others. Any level of site susceptibility requires propagule pressure from extant vegetation or seed banks to fuel the invasion. Even if plant material reaches a susceptible site, there is no guarantee that invasion will occur. Local, stochastic circumstances may preclude subsequent germination and establishment of viable populations. Nevertheless, managers must be aware of activities and disturbances that facilitate invasion since successful invasions occur when species attributes, environmental context, and interactive stochastic events (that is, disturbances or stresses that make resources available to invading plants that would otherwise be

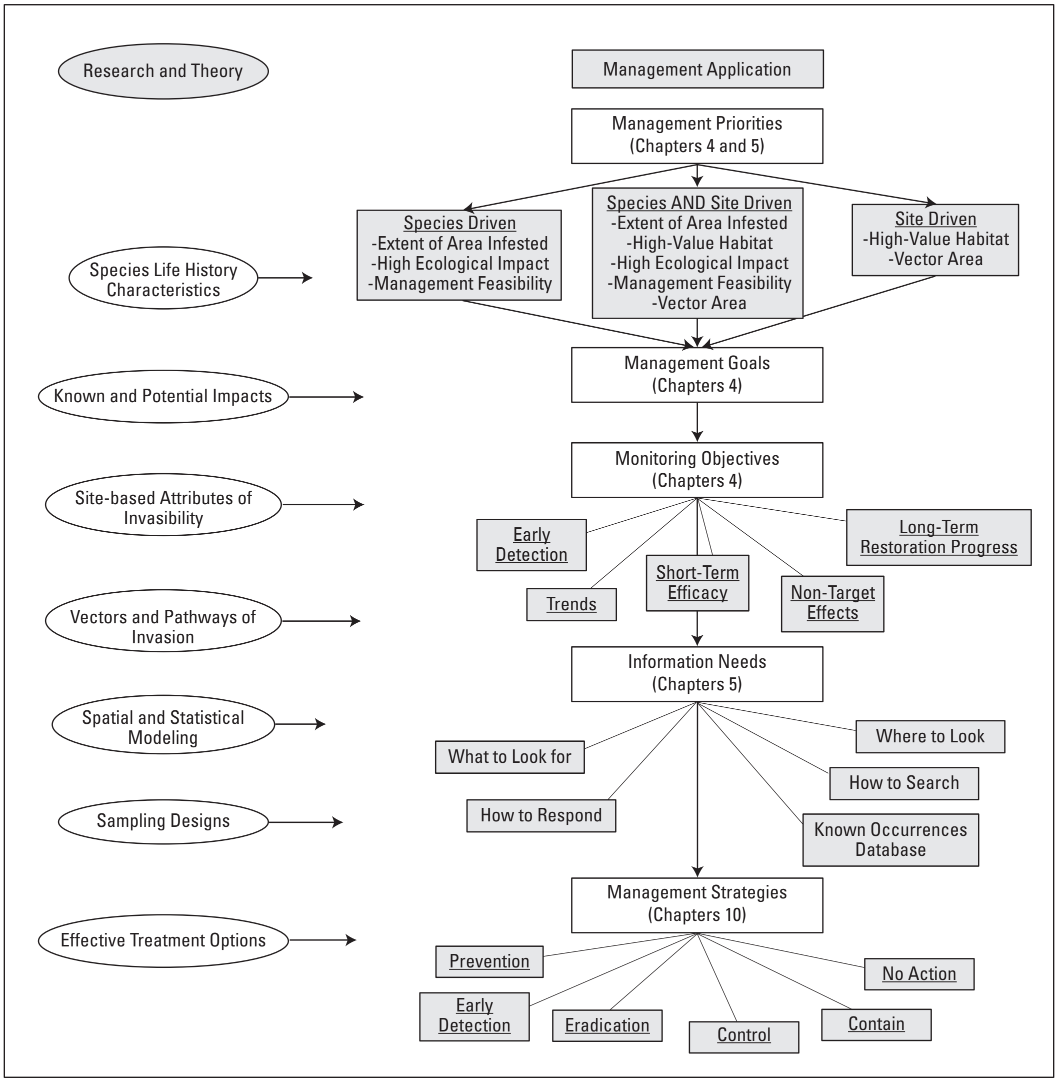

Figure 2.3. Conceptual linkage between invasive species theoretical research (ovals) and management application (boxes). Information in this figure also creates a link to the conceptual framework presented in Chapter 3 and figure 3.2.

used by resident flora) are complementary (Smith and others, 1999; Davis and others, 2000; Mack, 2003). Today's digital tools such as geographic information systems (GIS) offer managers a means to evaluate a variety of site characteristics simultaneously. Coupled with local knowledge, these tools can assist managers in prioritizing sites on the basis of susceptibility to invasion. Vector and pathway analysis may be used as a proactive means of utilizing these sets of information to improve early detection and prevention strategies.

Until recently the primary focus of research, assessment tools, and management practices has been species-driven. Management products derived from a species-driven approach may prove useful for predicting and detecting the occurrence of well-established invaders, but they are not effective tools for long-term, strategic management of new invasive plants (Smith and others, 1999). Strategic management requires nonnative species to be detected early before they are well established and problematic. To be successful as a management strategy, therefore, early detection of invasive plants demands consideration of all three elements of invasion—species attributes, community susceptibility, and interactive stochastic circumstances. The effective application of such knowledge needs to be applied in a systematic, iterative, and thorough procedure to optimize available resources.

This document is designed to put invasion theory into practice for managers requiring invasive plant early detection methods. Chapter 3 describes conceptually the steps involved in the early detection process. These steps are built on the aforementioned theoretical underpinnings and are described in more detail in subsequent chapters. Examples are given where appropriate, and practical applications are included in the Applications and Principles chapters (Chapters 11–13).

Recommended Reading

For a brief synopsis of invasion theory, see Sharma and others (2005). Undoubtedly, managing vectors and pathways to invasion is the most effective means of minimizing impacts from invasive species. It can also be perceived as complex and difficult to achieve. Ruiz and Carlton (2003) include information on a variety of vectors and pathways, risk assessment models, policy initiatives, and conceptual frameworks. Spatial and temporal patterns in terrestrial, marine, and freshwater ecosystems are discussed, including different taxonomic groups. See Ruiz and Carlton (2003).

Richardson and others (2000) include a concise discussion of one set of theoretical concepts behind plant invasion. The terminology they use is consistent with terms used in other parts of the world, but may differ slightly from that used in North America. Steps have been taken herein to rectify discrepancies

For a good overview of all aspects (for example, management, research, policy, economics, social) of biological invasions across taxonomic groups, see Mack and others (2000).

Predicting plant invasions is a complex task with conflicting opinions on the best methods for achieving practical results. The National Research Council (2002) publication reviews the state of the science and management applications relative to prediction.

Some ecosystems seem to be more susceptible to invasion than others or may be more susceptible at different times. This may be a result of the fluctuation in available resources that occurs in an ecosystem from natural and anthropogenic causes. See Davis and others (2000).

A recent publication of Montana State University addresses inventory and survey methods for nonnative species which are vital as a comparative baseline for any invasive plant monitoring work (Rew and Pokorny, 2006).

Box 2.5. Characteristics of nonnative species and susceptible sites that could lead to successful biological invasions are reasonably well known; however, the types of data needed to quantify these characteristics are less well understood. It is also poorly known how to assign parameters to the characteristics and how the parameters would be used in algorithms to determine the likelihood of each stage of the invasion process (National Research Council, 2002).

Klinger and Brooks (2007) formulated a quantitative stage-based prioritization using the Analytical Hierarchy Process (AHP) to prioritize species for early detection surveys at Whiskeytown National Recreation Area. In this method, a final score for a species is the sum of all of the products of several weighted criteria multiplied by the abundance value for the species from plot data (see Chapter 5 for more information).

Quantitative methods such as AHP potentially have utility, but existing data collected in a manner suitable for this analysis were extremely limited. In addition, to use this approach for early detection, information would need to be available for species that might be introduced and would not, therefore, be available at the management unit of concern.

References Cited

Bell, A.D., 1991, Plant form—An illustrated guide to flowering plant morphology: Oxford University Press, Oxford, United Kingdom, 432 p.

Binggeli, P., 2001, Time-lags between introduction, establishment and rapid spread of introduced environmental weeds: Proceedings of the Third International Weed Science Congress no. 238, Oxford, Miss., International Weed Science Society.

Brooks, M.L., D'Antonio, C.M., Richardson, D.M., Grace, J.B., Keeley, J.E., Ditomaso, J.M., Hobbs, R.J., Pellant, M., and Pyke, D., 2004, Effects of invasive alien plants on fire regimes: BioScience, v. 54, p. 677–688.

Chapman, R.A., Le Maitre, D.C., and Richardson, D.M., 2001, Scenario planning—understanding and managing biological invasions in South Africa, *in* McNeely, J.A., ed., The great reshuffling—Human dimensions of invasive alien species: Gland, Switzerland, and Cambridge, United Kingdom, International Union for Conservation of Nature, p. 195–208.

Chippendale, J.F., 1991, The potential returns to research into rubber vine (*Cryptostegia grandiflora*) in north Queensland: Master of Agriculture Studies thesis, University of Queensland, Australia, 173 p.

Conant, P., Medeiros, A.C., and Loope, L.L., 1997, A multiagency containment program for miconia (*Miconia calvescens*), an invasive tree in Hawaiian rain forests, *in* United Kingdomen, J.O., and Thieret, J.W., eds., Assessment and management of plant invasions: New York, Springer, p. 249–254.

Connell, J.H., and Slatyer, R.O., 1977, Mechanisms of succession in natural communities and their role in community stability and organization: American Naturalist, v. 111, p. 1119–1144.

Cousens, R., and Mortimer, M., 1995, Dynamics of weed populations: Cambridge University Press, Cambridge, United Kingdom, 348 p.

Davis, M., Grime, J.P., and Thompson, K., 2000, Fluctuating resources in plant communities—A general theory of invasibility: Journal of Ecology, v. 88, p. 528–534.

Elton, C.S., 1958, The ecology of invasions by animals and plants: London, United Kingdom, Methuen, 196 p.

Ensernick, M., 1999, Biological invaders sweep in: Science, v. 285, p. 1834–1836.

Grime, J.P., 1979, Plant strategies and vegetation processes: Chichester, United Kingdom, John Wiley & Sons, Inc., 222 p.

Grime, J.P., 2001, Plant strategies and vegetation processes (2d ed.): Chichester, United Kingdom, John Wiley & Sons, Inc., 456 p.

Groves, R.H., 1986, Plant invasions of Australia—An overview, *in* Groves, R.H., and Burdon, J., eds., Ecology of biological Invasions—An Australian perspective: Canberra, Australia, Australian Academy of Science, 137–149 p.

Hara, T., van der Toorn, J., and Mook, J.H., 1993, *Phragmites australis*, a clonal plant: Journal of Ecology, v. 81, p. 47–60.

Harper, J., 1965, Establishment, aggression, and cohabitation in weedy species, *in* Baker, H., and Stebbins, G., eds., The genetics of colonizing species: New York, Academic Press, p. 243–265.

Haslam, S.M., 1971, Community regulation in *Phragmites communis* Trin. I. monodominant stands: Journal of Ecology, v. 59, p. 65–73.

Hobbs, R.J., and Huenneke, L.F., 1992, Disturbance, diversity, and invasion—Implications for conservation: Conservation Biology, v. 6, p. 324–337.

Hobbs, R.J., and Humphries, S.E., 1995, An integrated approach to the ecology and management of plant invasions: Conservation Biology, v. 9, p. 761–770.

Kennedy, T.A., and Hobbie, S.E., 2004, Saltcedar (*Tamarix ramosissima*) invasion alters organic matter dynamics in a desert stream: Freshwater Biology, v. 49, p. 65–76.

Klinger, R., and Brooks, M.L., 2007, Incorporating weighted hierarchical criteria and uncertainty into invasive plant prioritization schemes—A case study from the National Park Service Klamath Network, *in* Holloran, P., ed., Measuring performance of invasive plant management efforts: Rohnert Park, Calif., Proceedings of the California Invasive Plant Council Symposium, v. 10, p. 7.

Kollmann, J., and Bañuelos, M.J., 2004, Latitudinal trends in growth and phenology of the invasive alien plant *Impatiens glandulifera* (Balsaminaceae): Diversity and Distributions, v. 10, p. 377–385.

Kowarik, I., 1995, Time lags in biological invasions with regard to the success and failure of alien species, *in* Pysek, P., Prach, K., Rejmánek, M., and Wade, P.M., eds., Plant invasions—general aspects and special problems: Amsterdam, The Netherlands, SPB Academic Publishing, p. 15–38.

Lodge, D.M., Williams, S., MacIsaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, P.B., Smith, M., Andow, D.A., Carlton, J.T., and McMichael, A., 2006, Ecological Society of America report—Biological invasions: Recommendations for U.S. policy and management: Ecological Applications, v. 16, p. 2035–2054.

Mack, R.N., 1989, Temperate grasslands vulnerable to plant invasion—Characteristics and consequences, *in* Drake, J.A., Mooney, H.A., di Castri, F., Groves, R.H., Kruger, F.J., Rejmanek, M., and Williamson, M., eds., Biological invasions—A global perspective: Chichester, United Kingdom, Wiley & Sons, p. 155–179.

Mack, R.N., 1995, Understanding the processes of weed invasions—The influence of environmental stochasticity, *in* Stirton, C., ed., Weeds in a changing world: Brighton, United Kingdom, British Crop Protection Council, Symposium Proceedings number 64, p. 65–74.

Mack, R.N., 2003, Phylogenetic constraint, absent life forms, and preadapted alien plants—A prescription for biological invasions: International Journal of Plant Science, v. 164, p. S185–S196.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M., and Bazzaz, F.A., 2000, Biotic invasions—Causes, epidemiology, global consequences and control: Ecological Applications, v. 10, p. 689–710.

McNeely, J.A., 2001, An introduction to human dimensions of invasive alien species, *in* McNeely, J.A., ed., The great reshuffling—Human dimensions of invasive alien species: Gland, Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, p. 5–20.

Memmott, J., Fowler, S., Paynter, Q., Sheppard, A., and Syrett, P., 2000, The invertebrate fauna on broom, *Cystisus scoparius*, in two native and two exotic habitats: Acta Oecologia, v. 21, p. 213–222.

National Research Council, 2002, Predicting invasions of nonindigenous plants and plant pests: Washington, D.C., National Academy of Sciences, 198 p.

Naylor, R.L., 2000, The economics of alien species invasions, *in* Mooney, H.A., and Hobbs, R.J., eds., Invasive species in a changing world: Washington, D.C., Island Press, p. 241–259.

Office of Technology Assessment, 1993, Harmful non-indigenous species in the United States: OTA–F–565: Washington, D.C., U.S. Congress, Government Printing Office.

Perrings, C., Williamson, M., Barbier, E.B., Delfino, D., Dalmazzone, S., Shogren, J., Simmons, P., and Watkinson, A., 2002, *in* Biological invasion risks and the public good—An economic perspective: Conservation Ecology, v. 6, p. 1.

Pickett, S.T.A., and White, P.S., 1985, The ecology of natural disturbance and patch dynamics: New York, Academic Press, 472 p.

Rejmanek, M., 2000, Invasive plants—Approaches and predictions: Austral Ecology, v. 25, p. 497–506.

Rejmanek, M., and Richardson, D.M., 1996, What attributes make some plant species more invasive?: Ecology, v. 77, p. 1655–1660.

Rew, L.J., and Pokorny, M.L., eds., 2006, Inventory and survey methods for nonindigenous plant species: Bozeman, Montana State University Extension, 80 p.

Richardson, D.M., Pysek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D., and West, C.J., 2000, Naturalization and invasion of alien plants—Concepts and definitions: Diversity and Distributions, v. 6, p. 93–107.

Ruiz, G.M., and Carlton, J.T., eds., 2003, Invasive species—Vectors and management strategies: Washington, D.C., Island Press, 484 p.

Sala, O.E., Lauenroth, W.K., McNaugton, S.J., Rusch, G., and Zhang, X., 1996, Biodiversity and ecosystem functioning in grasslands, *in* Mooney, H.A., Cushman, J.H., Medina, E., Sala, O.E., and Schulze, E.D., eds., Functional role of biodiversity—A global perspective: Chichester, United Kingdom, John Wiley & Sons, Inc., p. 129–149.

Sharma, G.P., Singh, J.S., and Raghubanshi, A.S., 2005, Plant invasions—Emerging trends and future implications: Current Science, v. 88, p. 726–734.

Simberloff, D., 1986, Introduced insects—A biogeographic and systematic perspective, *in* Mooney, H.A., and Drake, J.A., eds., Ecology of biological invasions of North America and Hawaii: New York, Wiley, p. 3–26.

Simberloff, D., 1995, Why do introduced species appear to devastate islands more than mainland areas?: Pacific Science, v. 49, p. 87–97.

Simberloff, D., and Von Holle, B., 1999, Positive interactions of nonindigenous species—Invasional meltdown?: Biological Invasions, v. 1, p. 21–32.

Smith, C.S., Lonsdale, W.M., and Fortune, J., 1999, When to ignore advice—Invasion predictions and decision theory: Biological Invasions, v.1, p. 89–96.

Stohlgren, T.J., Binkley, D., Chong, G.W., Kalkhan, M.A., Schell, L.D., Bull, K.A., Otsuki, Newman, Y.G., Bashkin, M., and Son, Y., 1999, Exotic plant species invade hot spots of native plant diversity: Ecological Monographs, v. 69, p. 24–46.

Tilman, D., 1997, Community invasibility, recruitment limitation, and grassland biodiversity: Ecology, v. 78, p. 81–92.

Tilman, D., Knopps, J., Wedin, D., Reich, P., Ritchie, M., and Siemann, E., 1997, The influence of functional diversity and composition on ecosystem processes: Science, v. 277, p. 300–1305.

Timmins, S.M., and Williams, P.A., 1991, Weed numbers in New Zealand’s forest and scrub reserves: New Zealand Journal of Ecology, v. 15, p. 153–162.

Vitousek, P.M., and Walker, L.R., 1989, Biological invasion by *Myrica faya* in Hawaii—Plant demography, nitrogen fixation, ecosystem effects: Ecological Monographs, v. 59, p. 247–265.

Williamson, M., 1993, Invaders, weeds and the risk from genetically manipulated organisms: Experientia, v. 49, p. 219–224.

Williamson, M., 1996, Biological invasions: London, United Kingdom, Chapman & Hall, 256 p.

Williamson, M., and Fitter, A., 1996, The varying success of invaders: Ecology, v. 77, p. 1661–1666.

Strategic Approach to Early Detection

By Bradley A. Welch

Chapter 3 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figures

Chapter 3. Strategic Approach to Early Detection

By Bradley A. Welch

Overview

In their review of the status of biotic invasions, Mack and others (2000) stated in no uncertain terms that "control of biotic invasions is most effective when it employs a long-term, ecosystem-wide strategy rather than a tactical approach focused on battling individual invaders." Effective management of invasive plants, then, requires managers to match strategies with the task at hand. In other words, early-detection methods should be used to find and subsequently eradicate incipient populations of invasive plants while these measures are still feasible and cost effective (see Chapter 2, particularly fig. 2.1). Control strategies should be applied to existing populations that require a more substantial resource investment to mitigate associated impacts. Containment measures should be used when populations have become so entrenched that removal and effective control are impractical.

The planning and direction (that is, strategy) for accomplishing any one of the component approaches to invasive plant management can be complex and multifaceted. Insufficient attention to detail or incomplete implementation of a given strategy can yield less than desirable results. Early-detection methods are susceptible to incomplete execution because data management and reporting are often sporadic or because search efforts commonly are inadequate and uncoordinated (National Research Council, 2002; Rejmanek and Pitcairn, 2002; DeAngelis and others, 2003). The complexities inherent to early detection should come as no surprise. Given the nature of this management strategy, one is literally searching for the proverbial needle in a haystack. However, the rewards one reaps by implementing a comprehensive early-detection strategy far exceed investment costs (Perrings and others, 2002). A little planning can go a long way.

This document provides guidance for designing and implementing a comprehensive, invasive plant early-detection monitoring strategy. Our intent is to provide a solid framework for those managers just starting out. We also hope to improve the likelihood of success for those managers currently using early-detection methods but who desire a more effective protocol or want to explore alternatives. This chapter provides an overview of the key components contributing to a successful early-detection and rapid-response program (fig. 3.1). We then outline the steps required for a comprehensive invasive plant early-detection monitoring program (fig. 3.2) because other sources provide examples of rapid response frameworks (Federal Interagency Committee for the Management of Noxious and Exotic Weeds, 2003; National Invasive Species Council, 2003). Chapters 4–10 address in more detail each step in the early-detection process. Information in this chapter is based on conceptual models generated by researchers and natural resource managers who met in Portland, Oregon, in August 2004, and on subsequent input from U.S. Geological Survey and National Park Service staff.

Early Detection and Rapid Response

Early detection and rapid response go hand in hand. The effort invested in early detection has limited value if management action is not taken to eliminate targeted populations before they spread. Similarly, the value of rapid response is realized only if populations are identified when they are small and manageable. From a pragmatic standpoint, early-detection and rapid-response methods encompass more than just finding and killing invasive plants. Rapid-assessment techniques as well as data management and reporting are also essential to an effective program (fig. 3.1).

Rapid Assessment

Rapid-assessment procedures follow the detection or reporting of a suspected invasive plant. The plant's identity must be validated and verified by a reputable source. This task prevents the unwanted destruction of desirable native species that look similar to undesirable nonnative species. A standard risk assessment must be conducted so that the plant can be ranked and prioritized along with other new and existing invasive species. Chapter 5 discusses prioritization in further detail and presents a methodology. Other prioritization and risk-assessment procedures such as the Alien Plant Ranking System (*http://www.npwrc.usgs.gov/resource/literatr/aprs/*) and the Invasive Species Assessment Protocol (*http://www.natureserve.org/library/invasiveSpeciesAssessmentProtocol.pdf*) are also available and are summarized in Chapter 5. Forms with standardized fields for data entry improve the

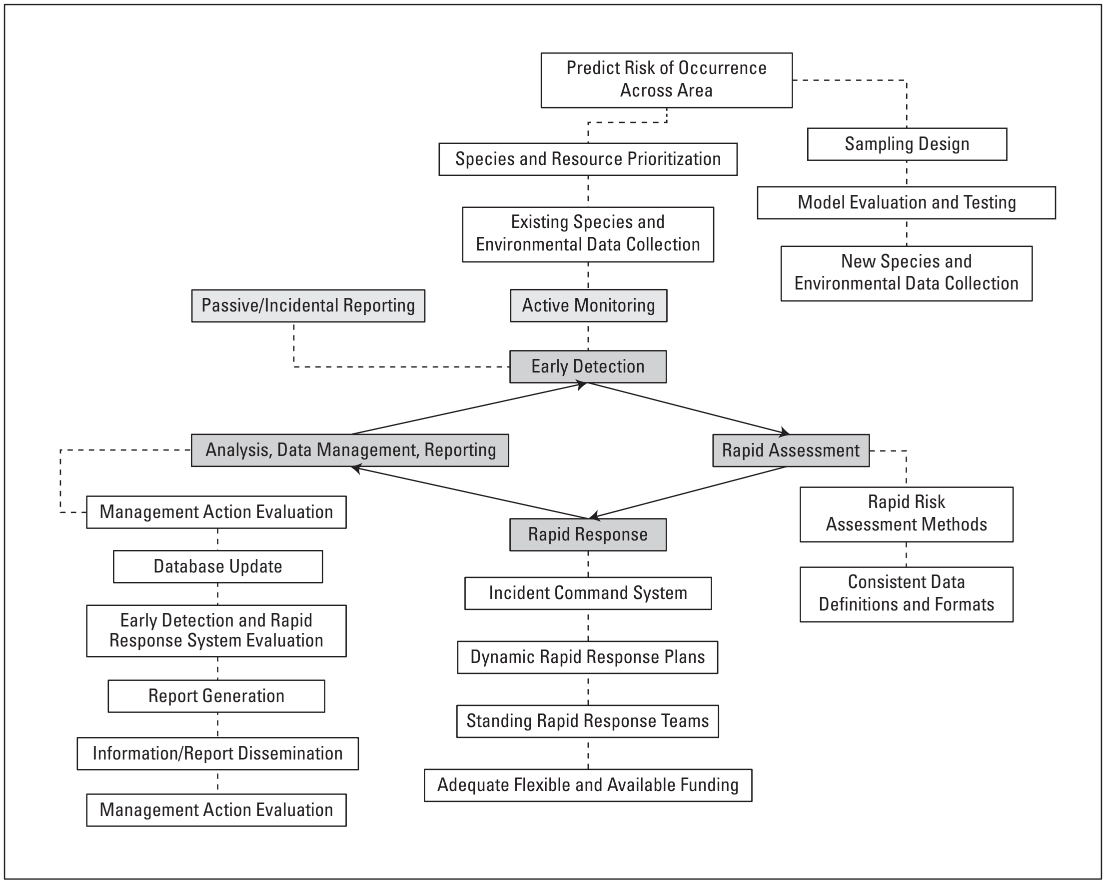

Figure 3.1. Key stages (dark-grey boxes) and associated subtasks (white boxes) of an invasive plant early-detection and rapid-response system. Note that the early-detection stage may involve active and (or) passive monitoring (light-grey boxes). Also note that early detection is a cyclic process, though it is not illustrated as such here.

efficiency of the assessment process and the comparability of data sets (see Chapter 10 and examples above). If conducted systematically, the rapid-assessment procedure improves the efficiency of the rapid-response procedure.

Rapid Response

Rapid response can also be improved to varying degrees. An incident command system improves communication and response time by having established points of contact and standard response procedures already in place. Response efficiency can also be improved by having designated response personnel. Response personnel may be limited to a single person or may involve response teams depending on the resources available and the potential need. Of course, no response plan can be implemented without adequate and flexible funding. Detections of plant invasions do not always occur at predictable times, making fiscal planning difficult, and many resource managers do not have the budgetary flexibility to set aside funds for potential invasions. However, strategic early-detection protocols can alleviate fiscal uncertainty for managers by incorporating periodic, planned search and response procedures into standard management plans. See the National Invasive Species Council (accessed November 12, 2010 at *http://www.invasivespecies.gov/global/EDRR/EDRR_index.html*) for guiding principles to early detection and rapid-response, and the Federal Interagency Committee for the Management of Noxious and Exotic Weeds (*http://www.fws.gov/ficmnew/FICMNEW_EDRR_FINAL.pdf*) for a rapid-response framework.

Analysis, Data Management, and Reporting

Early detection and rapid response programs quite often end after management actions have been implemented. This prevents managers from evaluating the success of detection, assessment, and response procedures and, consequently, precludes managers from improving the process for the future. Post response evaluation can pinpoint hurdles to effective management as well as highlight fruitful activities. Input and analysis of standardized data fields can also help identify species that have been introduced repeatedly, areas that have been invaded repeatedly, and vectors and pathways that have contributed to multiple invasions. This information improves predictive capabilities which, in turn, improves early-detection procedures and may elicit a change in park management to halt repeated invasions where possible. Data management can also be used to inform neighbors of potential invasive plant threats and can assist staff with public outreach campaigns aimed at reducing future sources of invasion. Report generation, of course, can serve as a communication tool as well as a record of events for current and future staff (see Chapter 10).

Steps to Early Detection

Detection methods can be active (that is, reliant upon planned search strategies), passive (that is, reliant upon incidental reporting from staff and visitors), or both. While incidental information is helpful, active detection methods remove much of the uncertainty about the state of the target resource, providing better knowledge of where invasive plants do and do not exist.

Active detection methods require an upfront investment in planning and decision making. Management goals and program scope (including ecological scale and context) must be defined from the beginning to ensure that program success can be obtained and evaluated (see fig. 3.2, Step 1, and Chapter 4). Lists of known and potential target species need to be developed (Step 2). The lists should be evaluated and updated periodically. Species and environmental data need to be collected and evaluated for each species on the target list and for all areas of management interest (Step 3). Much of this information may be obtained from literature reviews and from local knowledge. Assembled information includes existing life-history information, spatial habitat data (for example, soils, elevation, vegetation type, distance to rivers or streams), and potential vectors and pathways (see Chapter 5). Ideally, this information would be stored in a geographic information system (GIS) or be compatible with GIS for ease of use with modeling and analysis later in the process. An occurrence database should be constructed with information on presence and absence data for target species inside and outside the park where possible (Step 4). Given this assemblage of information, resource staff need to clearly define the monitoring objectives that will drive the approach to risk assessment and search strategies (Step 5; see Chapter 4). Note that most of the first five steps require significant input and data mining on the part of the parks and their partners (fig. 3.2, grey boxes). Consultants and contractors can provide assistance with many of these steps, but ultimately the responsibility lies with the resource management staff, particularly for goal and objective setting. Well-framed goals and objectives will assist staff, consultants, and contractors in producing an early-detection protocol that best serves management needs (see Chapter 4).

Consultants and contractors can provide significant assistance with many of the other, more complex steps in the early-detection process (fig. 3.2, dotted boxes). Priority assessment is one such step (Step 6). Managers need to choose whether to take a species-driven (that is, target species are prioritized for detection), a site-driven (that is, sites, ecosystems, or management units are prioritized for their conservation value and (or) management significance), or a combined approach to setting priorities (fig. 3.2, dashed white boxes). Many tools that exist to assist with this task emphasize different aspects of invasion biology and impact, so one should choose a tool that is well suited to the management and conservation priorities of the particular natural area (Chapter 5). Whatever tool is selected, we suggest using peer-reviewed risk-assessment methods to improve consistency and repeatability of priority assessments.

For each species and (or) site on the target list, some form of prediction will need to be made regarding the risk of occurrence across the park or areas of interest (Step 7; see Chapters 6 and 7). This prediction may be as simple as identifying dominant invasion pathways such as roads and trails for subsequent searches (a simple mental model). Or, risk of occurrence predictions may be applied across the entire management unit, mapping the probability of occurrence for target invasive species (complex models requiring a computer algorithm and GIS to generate risk probabilities). Procedures should be well documented so they can be repeated and implemented in other parks no matter what predictive tool is used.

Once predictions are made of the occurrence of invasive plants in areas of concern, an optimal search strategy (sampling or survey design) must be developed to efficiently cover the priority species and areas defined in Steps 1–6 (Step 8; see Chapter 8). High risk areas identified in Step 7 may be sampled more intensely to increase the probability of finding invasive plants, but other areas should also be sampled to improve future predictive capabilities and in case the predictions are inaccurate. Areas of high conservation value such as rare species habitat, wetlands, and special ecological areas may be sampled more intensely because of their importance. Allowance should be made for differing costs (for example, time) and constraints (for example, safety) of travel to a sampling unit. Again, procedures and design considerations should be well documented so they can be implemented in other parks.

Before implementing search strategies and the early-detection protocol, predictive models and sampling designs should be evaluated and field tested (Step 9; see Chapter 9). The evaluation step will ensure that predictive models and sampling designs are efficient and appropriately match monitoring objectives. Kettenring and others (2006) provide guidance to model developers and users for documenting and evaluating simple and complex models. Managers should recognize the limits of models relative to their objectives, identify limitations and constraints, evaluate options for parks of varying sizes, determine the probability that a method may fail to detect species occurrence, and ensure that sample sizes are adequate for critical habitats. Local staff and volunteer efforts may contribute significantly to the actual search strategies and field testing steps, so allowance should be made for these options during the planning stages (see Chapter 13 for an example from the NPS San Francisco Bay Area Network that describes the use of volunteers in the early-detection strategy).

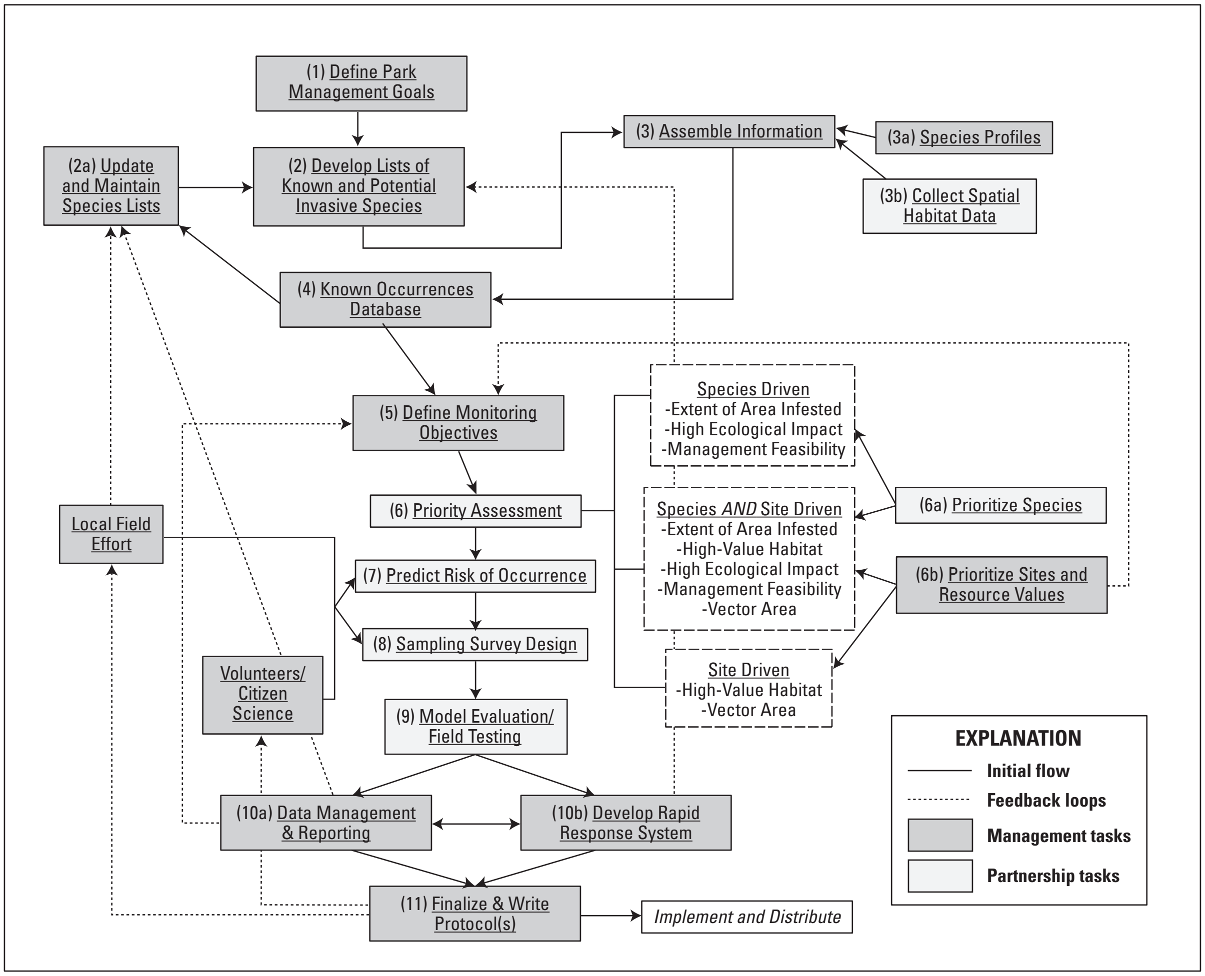

Figure 3.2. Principal steps and feedback loops involved in a comprehensive, invasive plant early-detection monitoring program. Some steps require significant input and planning on the part of resource managers and staff (dark-grey boxes). Note that defining monitoring objectives (Step 5) is a vital step to ensuring the success of early-detection monitoring. Step 6 requires managers to decide whether priorities will be assessed based on species, sites, or both (white boxes). Steps 10a, 10b, and 11 may involve input and guidance from a regional, national, or centralized office. Other steps may require assistance from partners, consultants, and contractors in order to complete (light-grey boxes).

Plans for data-management and rapid-response procedures (Steps 10a and 10b) should be made once predictive models and sampling designs have been revised. The database should record all search efforts, including all occurrences and absences of target species resulting from the formal search effort, and any incidental observations of target species obtained from other activities or reports. Procedures should allow continual feedback and improvement to life-history information, GIS layers, and other information to improve predictive risk models. The list of target species will be periodically updated as well. Data should be analyzed and results reported at appropriate intervals. Data should be shared and efforts coordinated with other parks and partners to foster regional control efforts. If a target population is found but cannot be controlled immediately, arrangements should be made for future control efforts. Input and guidance from NPS regional or national offices may be required for some data-management, reporting, and rapid-response procedures to ensure comparability across parks and monitoring networks (see Chapter 10).

Ultimately, all procedures must be well documented so that early-detection strategies can be implemented consistently in the future by new staff and revised as needed (Step 11; see Chapter 10). The National Park Service and the U.S. Geological Survey have adopted formatting and content standards for monitoring protocols that should be followed by those agencies (accessed January 5, 2007 at *http://science.nature.nps.gov/im/monitor/protocols/ProtocolGuidelines.pdf*). Chapter 11 discusses protocol development.

Once the protocol and procedures have been finalized, resource managers can implement the procedures, using the predictive search model to direct search effort by staff, backcountry rangers, volunteers, researchers, and others. If one of the target species is found, the park will want to map the infestation, using North American Invasive Species Management Association (NAISMA) standards (accessed April 13, 2014 at (*http://www.naisma.org/*), and eradicate the invader as quickly as possible. This may involve acting immediately or reporting the infestation to others for future removal depending on the size of the infestation and management tools required. In any event, this finding should also trigger more thorough searches in the area, using perhaps an adaptive sampling design or similar sampling strategy to locate other satellite populations nearby (see Chapter 8 for sampling options).

Most steps in the early-detection process feed information back into previous steps (fig. 3.2, dashed lines) with a view to improving the efficiency and effectiveness of the entire process over the long term. Ideally, all stages are interactive and iterative, improving the early-detection and rapid-response processes with successive iterations.

Recommended Reading

Consult the National Invasive Species Council's Guidelines (*http://www.invasivespecies.gov/global/EDRR/EDRR_index.html*) for a review of the general elements required for an early detection and rapid response system. Alternatively, see National Invasive Species Council (2003)

The Global Invasive Species Programme's Web site (*http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/ENVIRONMENT/EXTBIODIVERSITY/0,,contentMDK%3A20473193~menuPK%3A1170331~pagePK%3A148956~piPK%3A216618~theSitePK%3A400953,00.html*) contains general guidelines for international efforts.

See the Federal Interagency Committee for the Management of Noxious and Exotic Weeds (*http://www.fws.gov/ficmnew/FICMNEW_EDRR_FINAL.pdf*) documents for more details regarding the integration of early detection and rapid response. Alternatively, see Federal Interagency Committee for the Management of Noxious and Exotic Weeds (FICMNEW) (2003).

See the Ecological Society of America's review of U.S. invasive species management and policy with national-level recommendations for early detection. See Lodge and others (2006).

A recent publication by Montana State University addresses inventory and survey methods for nonnative species which are vital as a comparative baseline for any invasive plant monitoring work (Rew and Pokorny, 2006).

References Cited

DeAngelis, D.L., Gross, L.J., Comiskey, E.L., Mooij, W.M., Nott, M.P., and Bellmund, S., 2003, The use of models for a multiscaled ecological monitoring system, *in* Busch, D.E., and Trexler, J.C., eds., Monitoring ecosystems—Interdisciplinary approaches for evaluating ecoregional initiatives: Washington, D.C., Island Press, p. 167–188.

Federal Interagency Committee for the Management of Noxious and Exotic Weeds (FICMNEW), 2003, A national early detection and rapid response system for invasive plants in the United States—Conceptual design: Washington, D.C., 27 p.

Kettenring, K., Martinez, B.T., Starfield, A.M., and Getz, W.M., 2006, Good practices for sharing—Ecological models: BioScience, v. 56, p. 59–64.

Lodge, D.M., Williams, S., Macisaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, P.B., Smith, M., Andow, D.A., Carlton, J.T., and McMichael, A., 2006, ESA report, Biological invasions—Recommendations for U.S. policy and management: Ecological Applications, v. 16 p. 2035–2054.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M., and Bazzaz, F.A., 2000, Biotic invasions—Causes, epidemiology, global consequences and control: Ecological Applications, v. 10, p. 689–710.

National Invasive Species Council, 2003, General guidelines for the establishment and evaluation of invasive species early detection and rapid response systems, version 1, 16 p.

National Research Council, 2002, Predicting invasions of nonindigenous plants and plant pests: National Academy of Sciences, Washington, D.C., 198 p.

Perrings, C., Williamson, M., Barbier, E.B., Delfino, D., Dalmazzone, S., Shogren, J., Simmons P., and Watkinson, A., 2002, Biological invasion risks and the public good—An economic perspective: Conservation Ecology, v. 6, p. 1.

Rejmanek, M., and Pitcairn, M.J., 2002, When is eradication of exotic pest plants a realistic goal?, *in* Veitch, C.R. and Clout, M.N., eds., Turning the tide—The eradication of invasive species: Gland, Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, Species Survival Commission Invasive, Species Specialist Group, p. 249–253.

Rew, L.J., and Pokorny, M.L., eds., 2006, Inventory and survey methods for nonindigenous plant species: Bozeman, Montana State University Extension, 80 p.

Early Detection Strategy—Scope, Goals, and Objectives

By Susan O'Neil

Chapter 4 of
Early Detection of Invasive Plants—Principles and Practices
Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Inset

Chapter 4.
Early Detection Strategy—Scope, Goals, and Objectives

By Susan O'Neil[1]

Overview

The concerns raised by invasive species are so extensive that implementation of a comprehensive, long-term invasive plant monitoring plan could easily exceed a manager's entire monitoring budget. This is particularly true because managers must weigh early detection, trend detection, removal, efficacy monitoring, monitoring of secondary effects of management, and restoration/recovery monitoring when designing an invasive species management program. These options must all be considered within the bounds of a severely limited budget, facilities, and staff time. Further, early detection must compete with other monitoring priorities and a host of other natural resource management issues. For these reasons, it is critical that scientists and managers develop clear scope, goals, and objectives, consistent with the operational realities of their program. Herein, we outline suggestions for the development of scope, goals, and objectives for an invasive species early-detection (ISED) monitoring program.

It is crucial that all ISED programs articulate their scope, goals, and objectives. Documenting this information at the onset will facilitate agreement among participants and stakeholders, facilitate the review and refinement of the program, and avoid disruptive misunderstandings and failure to achieve critical objectives. Changes in conditions, resources, and personnel are inevitable, and well-documented scope, goals, and objectives will allow an ISED program to adapt to these changes. Some constraints and realities will become apparent only after a program has been implemented. This documentation is as essential for a small historical site with one volunteer looking for new infestations as it is for a large park with an extensive invasive species program and dedicated year-round staff.

Scope

With invasive species, it is important to consider geographic scope in every step in the planning process. The scope of the plan could be defined by a portion of the park, the entire park, or the park plus surrounding environs (that is, a park buffer area). Monitoring a buffer area around the park provides an excellent opportunity to detect invasive species before they reach the park. Monitoring a buffer area also encourages collaboration and partnerships with adjacent landowners.

To determine the appropriate geographic scope, managers need to consider the primary sources of invasive plants in or around the park and the primary vectors or potential entry points for species. These ecological issues frequently transcend ownership boundaries. Consequently, a more functional approach to defining scope employs ecological units to set boundaries for invasive plant species. Appropriate ecological units include ecoregions with distinctive physical and biological features, particular soil formations, vegetation types, or land-use categories. Although funding is rarely sufficient to allow full inventory and monitoring of even a modest sized park, knowledge of invasive species that exist outside park boundaries is often critical to planning an early-detection and rapid-response program.

Fiscal realities obviously also affect the scope of a monitoring program. Small parks may have the ability to completely census the area within the boundary and environs using staff and volunteers. Large parks may need to focus on certain high priority areas and search a random sample of other areas. Therefore, the goals should aim to specify a geographic extent that is ecologically sound at a level of detail that is feasible. Another approach could be to develop the early-detection program in phases. For example, to develop computerized predictive models to assess invasion risk that will require an investment in preliminary inventories and model design, a large park may wish to adopt a phased approach to early detection. Phase 1 might consist of developing an early-detection strategy for accessible park areas such as areas adjacent to roads, trails, and other corridors while model development is underway for relatively inaccessible areas. During Phase 2, early-detection monitoring for inaccessible areas is implemented using a portion or all of the funds previously designated for model development. Scope will continue to be an important consideration when selecting sampling design, field methodology, modeling approaches, and species prioritization as discussed in later chapters.

[1]Puget Sound Partnership, Tacoma, Washington (United States).

As an example, planning and research conducted in the Klamath Network (Chapter 11) reveals many invasive species in the lower elevation parks that are close to population centers such as Redwood and Whiskeytown National Parks. The most effective strategy for these parks is likely to be intensive sampling along vectors and invasive-species corridors such as roads and trails. Their geographic scope should include buffer zones outside of the park boundaries to find invasive plants before they reach the parks. Where funding is limited and invasives are not as prevalent, as is the case for a high-elevation park like Crater Lake, it may be appropriate for the scope to include only high priority sites rather than the entire park. Of course, the scope of the project always needs to be reflected in the objectives.

The Pacific Island Network presents a different context for defining the scope of an early-detection program. Oceanic islands are extremely vulnerable to invasion by nonnative plant species from continents. The catastrophic consequences of these plant invasions for native biodiversity and ecosystem processes in island ecosystems have been documented in the ecological literature more frequently in recent years. Because islands are surrounded by water, they can be protected from conspicuous terrestrial plant invasions by prompt detection and eradication, if there is the political will and the public support to do so (for example, Timmins and Braithwaite, 2002; Loope and Reeser, 2002). Under these circumstances, it may be appropriate to define the scope of the program from an islandwide perspective. This approach would incorporate buffer zones outside the parks for early-detection monitoring and would rely heavily on cooperators and other resource management agencies and organizations for implementation of early detection across a broader area than most NPS units would need to consider.

Goals

The goals of an ISED program should include a broad vision of the purpose of the program. The goals should concisely state what the manager hopes to achieve, in a broad sense, by implementing the program. According to Hendee and Dawson (2002) goals are general portraits of ideal ends or effects. They provide direction and purpose to potential objectives, which are attainable in the short term and more specific than goals. Goals are often lofty statements of intent (Hendee and Dawson, 2002), while objectives are statements of specific conditions to be achieved. When objectives are achieved, it shows progress in the direction of the established goals (Hendee and Dawson, 2002).

Examples of goals for an ISED program are to:

- Prevent new infestations from establishing within park boundaries,
- Ensure that high priority sites within the park are not invaded by new species or new infestations of existing species, and
- Develop partnerships with neighbors to detect new invasions in buffer areas surrounding parks.

When considering the goals of an early-detection program, it is important to realistically consider the institutional commitment and the scope of what can feasibly be detected. Generally, goals should intersect with major planning documents for the park or program to ensure that park managers consider the monitoring program to be relevant to their needs. If they do not, it is unlikely that the program will be supported or that results will be used to inform management.

It is also important that the goals of the early-detection program fit within the park's or network's larger invasive species programmatic goals. Particularly for larger parks or those with natural resource staff, early-detection monitoring will be one component of a larger invasive species management program. For more information on setting goals for initial survey and program work, please see Pokorny and others (2006).

Objectives

Once broad goals have been outlined, objectives can provide the operational details of the program. Monitoring objectives (referred to as management objectives by Elzinga and others [1998]) serve several purposes that are particularly valuable during program development:

- They focus and sharpen the thinking about the desired state or condition of the resource.
- They describe to others the desired condition of the resource.
- They determine the management that will be implemented and set the stage for alternative management if the objectives are not met.
- They provide direction to determine the appropriate type of monitoring.
- They specify a measurement and threshold for success (Elzinga and others, 1998).
- And, perhaps most importantly, they involve managers at the beginning.

The objectives are often the point where the realities and constraints of an invasive species program must be considered and where the potential risks can be assessed.

An **objective** is a specific statement that provides focus about the purpose or desired outcome of a particular monitoring program. One of the first steps in developing a long-term invasive species monitoring program is to articulate clear monitoring objectives. To be effective, monitoring objectives should be realistic, specific, unambiguous, and measurable. Once formulated, monitoring objectives assist in defining the protocol and sampling design needs (that is, sampling methods, variables to measure, number of replicates, number of plots or transects, plot size and spacing, and frequency and timing of monitoring) and help determine how the information will be used and evaluated after it has been collected (Box 4.1). Together with a well-written justification statement, monitoring objectives should define why monitoring is required, what will be monitored or measured, where the monitoring will take place, and when the monitoring will occur (DeAngelis and others, 2003; Gerlitzlehner, 2003).

Multiple objectives can be written within the scope of the overall program goals. For instance, if the goal mentions the entire park, the objectives should match that goal rather than focus on a specific high priority site(s), unless means have been made to make statistical inferences to the rest of the park. Monitoring objectives should, of course, remain within the fiscal and other resource constraints set by the goals. The important thing is that the scope has been thoughtfully considered and the objectives are not incorporating a different sized area, qualitatively different habitat, or management designation than encompassed in the overall program goals.

Assembly of existing documents, such as General Management Plans, Fire Management Plans, Restoration Action Plans, and so forth, will help to ensure that monitoring is consistent with higher level strategic planning. Such linkages will increase the likelihood of securing funding and broad-based support from park managers. Meeting with adjacent landowners and nearby agencies at this preliminary stage will provide information on existing programs, goals, and objectives. It also will reduce duplication and promote cooperation. These interactions will provide essential information about the pool of existing and encroaching species in the region. Partnerships with other resource-management agencies and organizations can assist in the development or adoption of common objectives.

Many factors can make setting objectives a daunting task. Often sufficient information is available about the ecology of an individual invasive plant species to make a fully informed decision (see Chapter 5). However, it is difficult to detect multiple species with different life-history traits, phenology, and rates and patterns of spread by using a single search strategy. Detection of species with differing life-history strategies complicates the methodology, timing, and locations of monitoring. This may result in the need for an iterative process to ultimately determine monitoring objectives. Many early-detection programs will also have the complicating factor of trying to locate new species in both disturbed areas associated with typical vectors and pathways (for example, roads, trails, campgrounds, riparian areas) as well as in high quality sites (for example, wetlands, backcountry sites, rare species habitat). These are referred to as vulnerable sites and valuable sites by Harris and others (2001), and both are important to consider when setting objectives. Managers are likely to have different objectives for the valuable sites and vulnerable sites. These considerations are often reflected in a stratified sampling design.

Simple conceptual or narrative summaries such as **conceptual ecological models** help describe important ecological processes and components that can assist in formulating good monitoring objectives (see fig. 2.1 for an example of an ecological model of invasive species). Invasive species ecological models can be as simple or as complex as desired. They can focus on individual target species, plant communities, focal ecosystems, or areas of the park. Obtaining information on related species can be helpful when considering certain invasive plants for targets of an early-detection program. Historical information on the park can help provide summaries of past disturbances or land use that is important for setting management objectives for invasive species. Seeking assistance from experts is a great way to learn about new information sources and may provide managers with potential reviewers of their ecological model or objective statements (Elzinga and others, 1998).

Box 4.1. The six components to include as part of a **good management/monitoring objective** as defined by Elzinga and others (1998) are:

- *Species or community*. This defines what will be monitored.
- *Specific place*. This is a geographic area that is often jurisdictional such as a park unit or subunit.
- *Attribute*. This is the aspect of the species or indicator that is measured, such as presence, percent cover, area, or density. For early detection of invasive species, presence will be the attribute measured.
- *Desired trend*. The verb of the objective (increase, decrease, maintain, prevent). This will apply to the species or community being monitored. For an early-detection program, the focus of the management objective can be either the community with invasive species or the invasive species themselves.
- *Amount of change*. This is a measurable state or degree of change for the attribute. This is typically the most difficult part of setting a management objective. It should be biologically possible and feasible to detect. Remember that the monitoring objectives for the early-detection program follow from this.
- *Timeframe*. The time needed for management to prove itself effective. This should be reasonable but short enough to elicit changes in management direction if needed.

Because a manager cannot typically afford to accomplish all of the desired components of a comprehensive invasive species management or monitoring program, a defensible prioritization process is a critical step to planning. When prioritizing tasks and developing associated objectives, a park manager, network, or region should consider several factors, including:

- The highest priority resources to monitor,
- The highest priority concerns for the resources,
- The elements with highest prospects for success,
- Logistical and safety concerns,
- Partnering opportunities,
- Land-use designations such as wilderness areas, and
- The costs and benefits of each component with respect to overall project goals.

The last factor may involve using certain cost-effective field methods such as using volunteers or students instead of paid professionals. Certain field methods should be selected over others if they are found to be more time efficient to provide a cost savings when using paid staff. Consideration should also be given to the time and cost of getting to remote field sites. If the early-detection program is working within the framework of a broad-based NPS Monitoring Program with other vital signs being monitored and measured at the park, consider collocating sampling sites where possible and (or) sharing staff. Rotating panel designs for more remote sites may also reduce the costs associated with field surveys. While budget considerations are often mentioned at the end of monitoring articles or protocols, we advise that they be considered as early as during the development of objectives. The budget for one year should not constrain the overall goals or long-term vision of a program, but the objectives should be considered in the context of what can be feasibly accomplished. It is also important to consider what level of time and effort will be meaningful for managing invasive species and ensuring that the program meets these targets.

The NPS I&M Program guidance focuses primarily on monitoring objectives (accessed April 13, 2014, at *http://science.nature.nps.gov/im/monitor/docs/NPS_Monitoring_Design_Guidance.pdf*). Listed below are examples of monitoring objectives, both simple and complex, that are based on this guidance and might have direct application depending on the size of the park and availability of resources:

Examples of monitoring objectives for small parks with small budgets:

- Walk half of the roads and trails at John Muir National Historic Site every year to search for occurrences of the top 10 potential invasive plants based on the Weed Management Area (WMA) list, so that all roads and trails are surveyed every 2 years.
- Detect new infestations of the top 10 ranked invasive plants at Sites X and Y, which contain endangered species habitat, by using trained volunteers to search random belt transects perpendicular to roads and trails.
- Maintain and update every 3 years a list of the worst existing and potential species at George Washington Carver National Memorial.
- Detect all new infestations of the top five prioritized species in three designated wetland sites within Moores Creek National Battlefield over a 5-year period by censusing the areas and adjacent roads and trails for plants each season.

Examples of monitoring objectives for large parks with larger budgets and more field staff:

- Systematically monitor each public trail in Colorado National Monument at least once per year for the next 10 years using visual assessment and GPS technology to detect and accurately map incipient populations of the top 10 plant species on the weed watch list. Monitoring and mapping will be conducted by trained volunteers and interns.
- Evaluate all invasive plant monitoring and mapping data collected along riparian systems, trails, and roads in Santa Monica Mountains National Recreation Area every 3 years to determine the primary pathways leading to new invasions in the park.
- Monitor each of 10 sites of ecological significance in Grand Teton National Park that are currently weed-free every year for the next 15 years to detect spotted knapweed.
- Detect biennially the presence of any new exotic plant either on the Shenandoah National Park watch list of exotic species or deemed by the Commonwealth of Virginia or surrounding States to be highly invasive within specific ecological zones (to be identified) found within the park. Monitoring efforts will be reevaluated every 7–10 years.
- Plan and implement a probability survey that will have at least a 70 percent chance of detecting garlic mustard before it has become established in more than 10 percent of the area of designated valuable sites at Great Smoky Mountains National Park.
- Detect new species and new infestations using a two-tiered system along invasive plant vectors in all Klamath Network parks. Tier 1 will use volunteers and park staff to adopt randomly selected sections of roads, trails, and riparian areas to search for infestations throughout the growing season. Tier 2 will use seasonal I&M crews to search randomly placed, stratified transects perpendicular to roads and trails using a rotating panel design throughout.

All monitoring objectives should answer the following questions (Guidance for Designing an Integrated Monitoring Program accessed April 3, 2014, at *http://science.nature.nps.gov/im/monitor/docs/NPS_Monitoring_Design_Guidance.pdf*):

- Is each of the objectives measurable?
- Are they achievable?
- Is the location or spatial bounds of the monitoring specified?
- Is the species or attribute being monitoring specified?
- Will the reader be able to anticipate what the data will look like?

In summary, it is extremely important to take the time and energy to develop good goals and achievable and measurable objectives within a scope that meshes with park management and that are ecologically meaningful. While these may seem like simple items to put together at the beginning of a program, they are can be quite challenging to do well. These items will set the tone and the path for the rest of the protocol, so it is recommended that they are reviewed regularly and revised as needed.

Recommended Reading

An excellent discussion of plant-based monitoring techniques is covered by Elzinga and others (1998). It includes information on setting management and sampling objectives that reinforce the information included here.

For more specific coverage of setting goals and objectives for invasive plant species inventory and survey projects, see Pokorny and others (2006).

The NPS I&M Web site (*http://science.nature.nps.gov/im/monitor/docs/NPS_Monitoring_Design_Guidance.pdf*) offers NPS guidance specific to the Inventory and Monitoring Program on formulating goals and objectives.

References Cited

DeAngelis, D.L., Gross, L.J., Comiskey, E.L., Mooij, W.M., Nott, M.P., and Bellmund, S., 2003, The use of models for a multiscaled ecological monitoring system, *in* Busch, D.E., and Trexler, J.C., eds., Monitoring ecosystems—Interdisciplinary approaches for evaluating ecoregional initiatives. Washington, D.C., Island Press, p. 167–188.

Elzinga, C.L., Salzer, D.W., and Willoughby, J.W., 1998, Measuring and monitoring plant populations: U.S. Department of the Interior, Bureau of Land Management, Technical Reference 1730–1, Denver, Colo., 496 p.

Gerlitzlehner, J., 2003, The standard operating procedure for monitoring weed control (QD Code: NH 1234): Department of Conservation, Wellington, New Zealand.

Harris, S., Brown, J. and Timmins, S., 2001, Weed surveillance—How often to search: Science for Conservation 175, Department of Conservation, Wellington, New Zealand, 27 p.

Hendee, J.C., and Dawson, C.P., 2002, Wilderness management—Stewardship and protection of resources and values (3d ed.): Golden, Colo., Fulcrum Publishing, 640 p.

Loope, L.L., and Reeser, D.W., 2002, Crossing boundaries at Haleakala—Addressing invasive species through partnerships, *in* Harmon, D., ed. Crossing boundaries in park management: Proceedings of the 11th Conference on Research and Resource Management in Parks and on Public Lands, Denver, Colo., April 2001.

Pokorny, M.L., Dewey, S.A., and Radosevich, S.R., 2006, Getting started—Fundamentals of nonindigenous plant species inventory/survey, *in* Rew, L.J., and Pokorny, M.L., eds., Inventory and survey methods for nonindigenous plant species: Bozeman, Montana State University Extension, p. 8–16.

Timmins, S.M., and Braithwaite, H., 2002, Early detection of invasive weeds on islands, *in* Veitch, C.R., and Clout, M.N., eds., Turning the tide—the eradication of invasive weeds, International Union for Conservation of Nature, Species Survival Commission, Invasive Species Specialist Group: Gland, Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, p. 311–318.

Prioritizing Species and Sites for Early-Detection Programs

By Matthew L. Brooks and Robert Klinger

Chapter 5 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figure

Table

Insets

Chapter 5. Prioritizing Species and Sites for Early-Detection Programs

By Matthew L. Brooks and Robert Klinger

Overview

An initial step in developing early-detection programs is to prioritize species and sites to identify which ones are most important to monitor. This process focuses on identifying species or life forms that are most likely to pose the greatest management challenge and sites that are most likely to be invaded or are most important to protect from invasion. Specific guidelines may vary depending on the stage of invasion that nonnative species may be in, within the area of interest. In this chapter we explain how species and sites can be prioritized to improve the efficiency of early-detection efforts.

Prioritizing Species for Early Detection

Species Lists

The principal raw materials for the prioritization process are species lists. Even programs designed to survey sites (as opposed to searching for particular species) benefit tremendously if species lists are used in the program design. Species lists, particularly nonnative species lists, vary in their usefulness depending on the detail and relevance of their geographic scope, ancillary information, and age.

Box 5.1. Lists that are compiled to document the status of nonnative plants are preferable over lists that are compiled for other purposes such as general botanical surveys or to validate vegetation maps.

Invasive plant lists have been developed for many States and multistate geographic regions within the United States. Examples from the Western United States include lists for Arizona (AZ-WIPWG, 2005), California (Cal-IPC, 2006), and Oregon and Washington (Reichard and others, 1997). Other regions with State lists include Connecticut (Mehrhoff and others, 2003), Florida (Florida Exotic Pest Plant Council Plant List Committee, 2005), Illinois (Schwegman, 1994), Rhode Island (Gould and Stuckey, 1992), and Virginia (Virginia Department of Conservation and Recreation and Virginia Native Plant Society, 2003; Heffernan and others, 2001).

Species lists can also be derived from coarse-scale regional surveys, or from finer scale local surveys. Regional lists are generally less useful than site-specific lists for programs designed at local scales. However, these two types of lists can be effectively used together. For example, a site-specific list can be used to target management actions for nonnative species already occurring within a management unit, and a regional list can be used to design local programs focused on detecting the initial establishment of species occurring elsewhere in the region.

Lists that are compiled to document the status of nonnative plants are preferable to lists that are compiled for other purposes such as general botanical surveys or to validate vegetation maps. No single list can serve all possible purposes as the quality of data on which the list is based will vary. Data characteristics depend on the sampling design used to collect the data (see Chapter 8), and the sampling design is ultimately determined by its intended purpose. Therefore, there is no single optimal sampling design for all applications. Consequently, data will vary in level of specificity, accuracy, and scope. For example, vegetation maps commonly are concerned with plant associations, noting only dominant species and other species of interest. Rare occurrences (which are the primary targets for early detection) may be left off intentionally or missed entirely. Accordingly, surveys that are not designed to specifically inventory nonnative plants most likely will underreport the actual number of nonnative species present in the sampling area.

Although it may at first sound counter-intuitive, it is often useful to also have data defining environmental conditions where target species do not occur (that is, absence data). Absence data can assist managers in defining areas where target species are unlikely to occur, thereby narrowing the geographic focus of the areas to be prioritized for searches. In addition, life-history information (for example, perennial relative to annual, presence of rhizomatous roots, seed mass, and so forth), tendency to be invasive in other geographic regions, known ecological effects and feasibility of control are highly desirable data sets that can help in the prioritization process. Older species lists (for example, more than 20–30 years) can be useful in documenting occurrence of a species in an area, although data on environmental conditions associated with these lists may be obsolete.

Prioritization Process

If species lists exist and resources to evaluate them are available, then the suite of species that early-detection monitoring should most optimally focus upon can be identified using a process known as prioritization. In the past, prioritization has been almost exclusively applied to narrow the number of species targeted for active management, but it can similarly be used to narrow the number of species targeted for early-detection monitoring. In both cases, prioritization addresses the desire to focus management efforts, whether for control or early detection, on a reduced subset of the total species pool where these efforts will be most effective.

Randall and others (2008) recently reviewed 17 examples of systems used to help place nonnative plants into categories to facilitate their management and compared them to a system that they developed themselves (Morse and others, 2004). Twelve of these systems are designed to prioritize management actions for nonnative species that are already established within a management unit: two prioritize among invaded sites (Timmins and Owens, 2001; Wainger and King, 2001) and 10 prioritize among species within sites, States, or nations (Orr and others, 1993; Weiss and McLaren,1999; Thorp and Lynch, 2000; Champion and Clayton, 2001; Fox and others, 2001; Heffernan and others, 2001; Virtue and others, 2001; Hiebert and Stubbendieck, 1993; Warner and others, 2003; Morse and others, 2004). Only two (Warner and others, 2003; Morse and others, 2004) focus heavily on species' effects on biodiversity, whereas the rest focus mostly on feasibility of control or potential effects on agricultural, horticultural, or other economic factors. Six other systems reviewed by Randall and others (2008) do not focus specifically on prioritization, but rather on prediction of the invasion potential for species that are not yet present within a given area (Rejmanek and Richardson, 1996; Reichard and Hamilton, 1997; Pheloung and others, 1999; USDA Natural Resources Conservation Service, 2000; Williams and others, 2001; USDA Animal and Plant Health Inspection Service, 2004). These predictive systems apply at the state or national scale, and focus almost exclusively on the potential for species to become established and spread within a biogeographic region.

Prioritization decisions are typically made on the basis of some combination of the following four factors:

1. The relative ecological and (or) economic concerns that the nonnative species pose,
2. their potential to spread and establish populations quickly (that is, their "weediness"),
3. their potential geographic and (or) ecological ranges, and
4. the feasibility in which they can be controlled (Timmins and Williams, 1987; Hiebert and Stubbendieck, 1993; Weiss and McLaren, 1999; Fox and others, 2000; Warner and others, 2003; Morse and others, 2004) (table 5.1).

The scoring systems for these prioritization efforts generally emphasize the threat potential and spread potential more than the other two factors, with the weighted sum of the ranks for all four resulting in the net priority assessment (table 5.1).

While the large number of systems may appear bewildering at first, many can be directly applied to a wide variety of areas and situations. Using an existing system will reduce the cost of developing a new system and provide managers with choices and flexibility. However, it is important to stress the necessity of selecting the system that is most appropriate for a given situation (Randall and others, 2008), and not just using one that the individuals involved may be familiar with.

Box 5.2. Although prioritization for control efforts has been most common, prioritization for early-detection monitoring is based on the same basic premises.

Prioritization is generally done for species that are known to be invasive or for sites that have high conservation value but may be susceptible to invasion. In some instances species

Table 5.1. Relations between the scoring of evaluation factors and the net assessments for prioritization efforts designed for control compared to those designed for early detection.

Evaluation factors	Prioritization for control		Prioritization for early-detection monitoring	
Threat potential	High	Low	High	Low
Spread potential	High	Low	High	Low
Range potential	High	Low	High	Low
Control feasibility	High	Low	Low	High
Net priority assessment	High	Low	High	Low

and sites both can be prioritized for management actions (Timmins and Owens, 2001), and if adequate resources are available this can be an extremely useful strategy. Prioritization is most often based on a synthesis of preexisting studies, expert opinion, or both (Randall and others, 2001; Hiebert and Stubbendieck, 1993; Timmins and Owens, 2001).

Although prioritization for control efforts has been most common, prioritization for early-detection monitoring is based on the same basic premises. There is no compelling reason that systems developed to inform control efforts could not be used (with minor modifications) to help inform early-detection monitoring efforts. Both strategies rely on information related to threat potential, spread potential, range of potential geographic/ecological sites, and feasibility of control (table 5.1). The one primary difference is that a species that has low feasibility of control should raise its priority level in terms of early-detection monitoring but lower its priority level in terms of control. Basically, invasive species that are more difficult to control should have higher priority in situations where early-detection monitoring is used to identify new populations and keep them from establishing.

Information Needed for Prioritizing Species

> **Box 5.3.** When prioritizing species, careful attention needs to be given to the phase of the invasion process the rankings are meant to address.

Relatively few life-history characteristics have been found to be consistently good predictors of invasiveness (Kolar and Lodge, 2001). Thus, efforts that spend inordinate amounts of time collecting as much information as possible on a large number of species attributes (the "shotgun" approach) may not yield the results that a more logical and focused approach will. For example, when prioritizing species, careful attention needs to be given to the phase of the invasion process the rankings are meant to address (see Chapter 2). Management objectives differ among the phases, as does the relative importance of species attributes in prioritizing them for early detection.

The management objective for species in the colonization phase of invasion is solely to prevent their introduction. Developing a list of species with the greatest potential for being introduced into the area of interest is a critical step in any effort to prevent such introductions. In most cases, developing a prioritized list for this stage will be the most difficult because the pool of potential species will likely be quite large. Once a list of candidate species is developed, useful information for prioritizing target species includes potential for invasiveness, biogeographical data, land-cover types where invasive, and potential effects.

When species are introduced and become established, their populations commonly are small and limited in geographical extent; therefore, a feasible management objective may be eradication (Rejmanek and Pitcairn, 2002). If eradication is not feasible, then control of populations (that is, reducing abundance and (or) dispersal pathways and vectors) within the boundaries of local infestations may be. However, it is important to recognize that even if eradication or control is successful, species could re-invade previously managed areas. High-priority species in this stage would be those that tend to fit the definition of a "transformer species," a species that causes significant changes in community and ecosystem characteristics (Richardson and others, 2000) and has ecological and life-history characteristics suggesting that the species could spread rapidly. Therefore, the primary focus should be on collecting information on effects, distribution and abundance, and life-history characteristics, as well as management feasibility and regional range.

Information needed to develop prioritized lists of invasive species in different stages of the invasion process. The categories within each level (that is A, B, C) are ranked in general order of importance.

A. Colonization phase

   1. Potential for invasiveness
      Tendency to be invasive elsewhere

   2. Biogeographical
      Natural ("native") range
      Nonnative ("invasive") range

   3. Land-cover types where invasive
      Vegetation type
      Ecosystems

   4. Potential impacts

B. Establishment phase

   1. Impacts
      Ecosystems
      Structure
      Species composition

   2. Ecological patterns
      Distribution in target sites
      Distribution in adjacent sites
      Abundance in adjacent sites

   3. Life-history characteristics
      Dispersal
      Reproduction

   4. Biogeographical
      Regional range

   5. Management feasibility
      Availability of control methods

C. Spread and equilibrium phases

   1. Management feasibility
      Availability of control methods
      Size of infestation
      Accessibility to infestations

2. Ecological patterns
   Trend in target sites
   Distribution in target sites
   Abundance in target sites

3. Life-history characteristics

4. Impacts
   Ecosystems
   Structure
   Species composition

Species that are more widely distributed or abundant are at the latter, or spread and equilibrium, stages of the invasion process. Eradication of these species is unlikely (Rejmanek and Pitcairn, 2002), so containment of existing populations is probably the most reasonable management objective. The likelihood of success for limiting further spread and reducing existing populations will depend on the size of existing populations and availability and effectiveness of containment methods. Data on trends in abundance and distribution and dispersal capability can help distinguish species that are spreading rapidly from those with slower rates of spread. While the general categories of information on species in the spread stage are the same as those in the establishment stage, the specific information that is of most use is generally different. Information on impacts can still be useful for prioritizing species in the spread stage, but information on management feasibility and population dynamics is more desirable. Biogeographical information is not especially helpful at this stage.

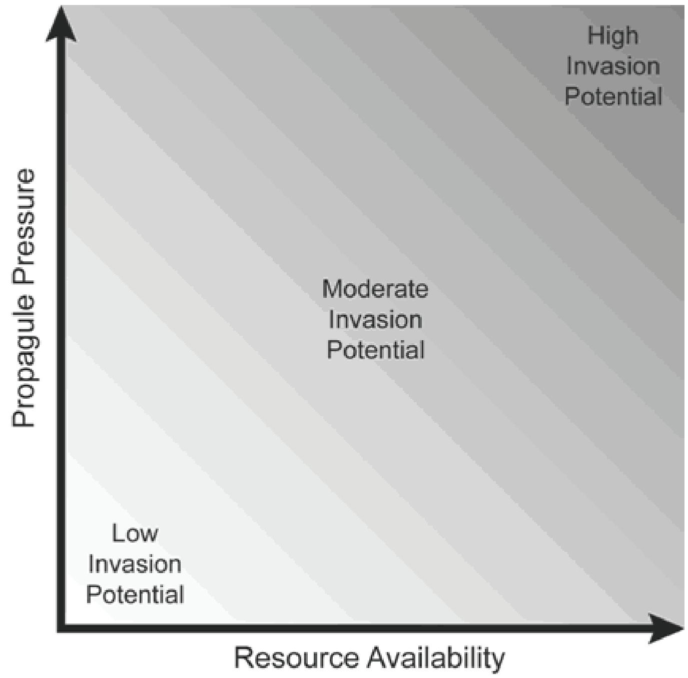

Figure 5.1. Summary of general invasion theory indicating that sites with the highest invasion potential tend to have high resource availability and are subjected to high propagule pressure. Reprinted from Brooks (2007).

Prioritizing Sites for Early Detection

Susceptibility of Sites to Invasion

Numerous interacting factors influence the susceptibility of sites to invasion, and their relative effects have been widely discussed and debated (Hobbs and Huenneke, 1992; Lonsdale, 1999; Williamson, 1999; Davis and others, 2000; Rejmanek and others, 2005). Two factors, however, appear particularly important: propagule pressure and resource availability (Brooks 2007; fig. 5.1).

Propagule pressure is related to the number of disseminules (for example, seeds, rhizomes) introduced into an area per unit time and the types of species they represent. Dispersal rates are positively associated with pathways such as roads and trails, vectors such as livestock, land-use practices such as seeding burned areas, and the extent of area open to invasion. The specific rates of dispersal necessary for a species to become established can vary with the species pool in the region, which is based on their relative abundances and range of life-history characteristics they represent. The species pool is the number of nonnative species in a region; the larger that pool, the greater the likelihood that at least one or several species will invade a local area contained within the region. In regions where there is a substantial pool of potential invaders with varying life-history characteristics, the amount of area to survey can be quite large. This is because there are many species available to potentially exploit a wide range of resource gradients.

Resource availability is a function of the supply of light, water, and mineral nutrients, and the proportion of these resources that are unused by existing vegetation. Resource availability can increase due to direct additions (for example, atmospheric nitrogen deposition), increased rates of production (for example, nutrient cycling rates), or by reduced rates of uptake following declines in plant abundance after they are thinned or removed. Resource requirements and characteristics allowing species to exploit resources differ markedly among annual and perennial herbaceous species, shrubs, and trees. Shade tolerance, root systems (for example, diffuse compared to tap root), metabolic pathways (for example, C3 or C4 photosynthetic pathway), and tolerance to herbivory are just a few examples.

Areas of high resource availability in many cases are disturbed sites. Fire, landslides, and floods not only increase the pool of available resources but may also reduce abundance of native species that would otherwise compete with invading species or reduce invasion rates by consuming potential colonizers (Marty, 2005). The function of disturbance in facilitating many invasions is well established (Lonsdale, 1999; Mack and D'Antonio, 1998; Mack and others, 2000), so disturbed areas need to be given high priority. Areas susceptible to high potential propagule pressure include areas of natural or anthropogenic disturbance that act as pathways for invasion, such as roadways, trails, and pastures where livestock are grazed. Collectively, these factors can be used to develop a basic program for sampling specific sites with a focus on particular species.

> **Box 5.4.** Prioritization of sites that are most susceptible to invasion should focus on areas with high **propagule pressure** and high **resource availability**.

Prioritization of sites that are most susceptible to invasion should focus on areas with both high propagule pressure and high resource availability (fig. 5.1). If resource availability is high but there are few or no vectors and pathways to the site, then the potential for invasive plants becoming established is relatively low. Similarly, if the potential for propagule pressure is high but resource availability is low, then the potential for plant invasion is also low. However, where propagule pressure is potentially high, there is always a chance that invasive plants can establish following a surge in resource availability, which most commonly occurs following a major disturbance and reduces resource uptake (for example, removal of vegetation by fire, flood, or other agents).

Conservation Value of Sites

The definition of conservation value can vary widely, depending on the specific priorities of the management unit. For example, where natural resources are a priority, conservation value may be defined by the presence of robust populations of sensitive species, high biodiversity, or ecosystems characterized by desirable structure and function (such as in a “natural” condition). Where cultural resources are a priority, conservation value is typically defined by the presence of prehistorical or historical artifacts, structures, or landscapes in their desired condition, which are generally defined to be as close to their original condition as possible. Other resources such as viewsheds or recreational opportunities may also be top priorities for management units.

Plant invasions are widely known for the challenges they pose to natural resources. What are less appreciated are the challenges they can pose to other resources of value. Invasive plants can generate concern for prehistorical or historical artifacts and structures by increasing the probability of fires, which can degrade these resources. Fires can obviously consume wooden structures, but the heat from fire can also fracture clay (for example, pottery) or rock (for example, pictographs, obsidian, chert) artifacts. Plant invasions may also alter historical landscapes by type-converting vegetation stands, especially when they shift from one major formation to another. An example is the challenge that invasive annual grasses and altered fire regimes pose to native sagebrush-steppe vegetation at Golden Spike National Historic Site, where a management priority is to promote the native vegetation conditions present at the time the transcontinental railroad was completed at this site in 1869. In cases where viewsheds are a resource, the loss of distinctive vistas dominated by Joshua trees (*Yucca brevifolia*) as a result of frequent fires fueled by invasive annual grasses at Joshua Tree National Monument may be a specific concern. Accessibility to recreational areas may also be degraded by invasive plants. Such is the case with saltcedar (*Tamarix* spp.) in riparian zones within Lake Mead National Recreation Area.

Because these site-based resource priorities vary widely, specific suggestions are difficult to develop that apply to all cases. Accordingly, decisions about prioritizing the conservation value of sites must be made on a case-by-case basis.

Information for Prioritizing Sites

There are six main categories of information to collect when prioritizing sites. Of these, the two most important categories are susceptibility to invasion and the conservation value of the site.

Information needed to develop prioritized lists of sites to protect from invasion by nonnative species. The categories within each level are ranked in general order of importance from highest to lowest.

A. Susceptibility to invasion

1. Intrinsic (site-specific)
   - Species richness of nonnative species
   - Distribution of nonnative species
   - Spatial
   - Vegetation community
   - Abundance of nonnative species
   - Land use
   - Disturbance
   - Historic
   - Contemporary

2. Extrinsic (off-site)
   - Vectors and pathways
   - Roads
   - Trails
   - Watercourses
   - Neighbor perimeter
   - Neighbor area
   - Land use
   - Disturbance
   - Contemporary

3. Invasion rates
   - Temporal trend in nonnative species accumulation

B. Conservation value

1. Biological resources/Biodiversity hotspots
   - Endemics
   - Threatened and endangered species
   - Rare community types

2. Cultural resources
   Prehistoric
   Historic

C. Landscape heterogeneity

   1. Patchiness

D. Connectivity

   1. Corridors

E. Management feasibility

F. Threat of nonnative species relative to other issues

Besides basic ecological information on nonnative species and land use within the area of interest (intrinsic factors), landscape configuration and characteristics are also important (extrinsic factors). This is because invasive species may initially spread from neighboring lands. Conservation value includes information on local hotspots of native diversity, endemism, and threatened and endangered species. Attributes at the landscape scale should also be considered when prioritizing sites, especially patchiness of vegetation communities (some communities are more prone to invasion caused by edge effects; for example, grasslands) and corridors connecting vegetation types to particular sites.

Final Recommendations

Box 5.5. The most effective early-detection monitoring programs include prioritization both of **species** and **sites**.

The most effective early-detection monitoring programs include prioritization both of species and sites. This approach allows monitoring crews to focus on the life forms and species that have the greatest potential to colonize, establish, spread, and ultimately pose significant threats to valued resources within a management unit. At the same time, this combined approach allows crews to focus on areas within the management unit where invasions are most likely to occur or where resources of greatest value for protection are located.

Before beginning any effort to develop an early-detection program, the resources available to implement the program must be evaluated (Brooks and Klinger 2009, also Chapters 10 and 11). Time spent compiling vast amounts of information to develop an early-detection plan is wasted if there is little hope of supporting the efforts needed to synthesize the information into an implementation plan or to implement the plan itself. Time and money are obvious limitations, but so too are institutional support and the personal commitment of staff. Turnover rates of personnel can also be a hindrance, since extensive training can be required to develop effective early-detection teams. Unfortunately, there are many examples of resources expended to develop elaborate sets of management recommendations that have little chance of being implemented because they require funding that is unlikely to be available. Although these efforts may provide important insights, in terms of management on the ground, these are resources largely wasted.

Information collected during the course of any monitoring program should also be used to evaluate and adjust sampling plans as needed (Holling, 1978). Early-detection programs are no exception. One thing that is universal in the management of nonnative plants is that there never seems to be enough information available on species and sites to develop management plans that are truly satisfying. Early-detection programs need to include plans for reprioritizing species, and possibly sites, periodically (perhaps at 5–10 year intervals) to maximize their effectiveness.

Recommended Reading

For an overview of plant community susceptibility to invasion, see Davis and others, 2000.

For an overview of various prioritization and predictions tools used in invasive species management, see Randall and others, 2008.

For a summary of the practical considerations that come into play when designing an early detection program, see Brooks and Klinger, 2009.

References Cited

AZ-WIPWG, 2005, Invasive nonnative plants that threaten wildlands in Arizona: Arizona Wildland Invasive Plant Working Group, accessed June 30, 2006, at *http://www.swvma.org/invasivenonnativeplantsthatthreatenwildlandsinarizona.html.*

Brooks, M.L., 2007, Effects of land management practices on plant invasions in wildland areas, *in* Nentwig, W., ed., Biological invasions—Ecological studies: Heidelberg, Germany, Springer v. 193, p. 147–162.

Brooks, M.L., and Klinger, R.C., 2009, Practical considerations for detecting and monitoring plant invasions, *in* Inderjit, ed., Management of non-native invasive plant species: Heidelberg, Germany, Springer, p. 168–195.

Cal-IPC, 2006, California invasive plant inventory: Cal-IPC Publication 2006–02: Berkeley, Calif., California Invasive Plant Council, accessed June 30, 2006, at *http://www.cal-ipc.org/ip/inventory/index.php.*

Champion, P.D., and Clayton, J.S., 2001, A weed risk assessment model for aquatic weeds in New Zealand, *in* Groves, R.H., Panetta, F.D., and Virtue, J.G., eds., Weed risk assessment: Collingwood, Australia, CSIRO Publishing, p. 194–202.

Davis, M.A., Grime, P.J., Thompson, K., 2000, Fluctuating resources in plant communities—A general theory of invisibility: Berkeley, Calif., Journal of Ecology, v. 88, p. 528–534.

Florida Exotic Pest Plant Council Plant List Committee, 2005, List of Florida's invasive species: Florida Exotic Pest Plant Council, accessed June 30, 2006, at *http://www.fleppc.org/list/05List.htm.*

Fox, A.M., Gordon, D.R., Dusky, J.A., Tyson, L., and Stocker, R.K., 2001, IFAS assessment of nonnative plants in Florida's natural areas: Agronomy Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida, SS-AGR-79.

Gould, L.L., and Stuckey, I.H., 1992, Plants invasive in Rhode Island: Rhode Island Wild Plant Society Newsletter, v. 6, no. 2, p. 1–7.

Heffernan, K.E., Coulling, P.P., Townsend, J.F., and Hutto, C.J., 2001, Ranking invasive exotic plant species in Virginia: Natural Heritage Technical Report 01–13, Richmond, Virginia, Department of Conservation and Recreation, Division of Natural Heritage, 27 p.

Hiebert, R.D., and Stubbendieck, J., 1993, Handbook for ranking exotic plants for management and control: Denver, Colo., U.S. Department of the Interior, Natural Resources Report NPS/NRMWRO/NRR–93/08, National Park Service, Natural Resources Publication Office.

Hobbs, R.J., and Huenneke L.F., 1992, Disturbance, diversity and invasion—Implications for conservation: Conservation Biology, v. 6, p. 324–337.

Holling, C.S., 1978, Adaptive environmental management and assessment: New York, John Wiley & Sons, Inc., 377 p.

Kolar, C.S., and Lodge, D.M., 2001, Progress in invasion biology—Predicting invaders: Trends in Ecology and Evolution, v. 16, p. 199–204.

Lonsdale, W.M., 1999, Global patterns of plant invasions and the concept of invisibility: Ecology, v. 80, p. 1522–1536.

Mack, M.C., and D'Antonio, C.M., 1998, Impacts of biological invasions on disturbance regimes: Trends in Ecology and Evolution, v. 13, p. 195–198.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M., and Bazzaz, F.A., 2000, Biotic invasions—Causes, epidemiology, global consequences, and control: Ecological Applications, v. 10, p. 689–710.

Marty, J.T., 2005, Effects of cattle grazing on diversity in ephemeral wetlands: Conservation Biology, v. 19, p. 1626–1632.

Mehrhoff, L.J., Metzler, K.J., and Corrigan, E.E., 2003, Nonnative and potentially invasive vascular plants in Connecticut: Storrs, Center for Conservation and Biodiversity, University of Connecticut, accessed June 30, 2006, at *http://www.hort.uconn.edu/cipwg/.*

Morse, L.E., Randall, J.R., Benton, N., Hiebert, R., and Lu, S., 2004, An invasive species assessment protocol—Evaluating nonnative plants for their impact on biodiversity, version 1: Arlington, Va., NatureServe, accessed April 3, 2014, at *http://www.natureserve.org/library/invasiveSpeciesAssessmentProtocol.pdf.*

Orr, R.L., Cohen, S.D., and Griffin, R.L., 1993, Generic nonindigenous pest risk assessment process: Washington, D.C., U.S. Department of Agriculture, 40 p.

Pheloung, P.C., Williams, P.A., and Halloy, S.R., 1999, A weed risk assessment model for use as a biosecurity tool evaluating plant introductions: Journal of Environmental Management, v. 57, p. 239–251.

Randall, J.M., Benton, N., and Morse, L.E., 2001, Categorizing invasive weeds—The challenge of rating the weeds already in California, *in* Groves, R.H., Panetta, F.D., and Virtue, J.G., eds., Weed risk assessment: Collingwood, Australia, CSIRO Publishing, p. 203–216.

Randall, J.M., Morse, L.E., Benton, N., Hiebert, R., Lu, S. Killeffer, T., 2008, The invasive species assessment protocol—A new tool for creating regional and national lists of invasive non-native plants that negatively impact biodiversity: Invasive Plant Science and Management, v. 1, p. 36–49.

Reichard, S.H., and Hamilton, C.W., 1997, Predicting invasions of woody plants introduced into North America: Conservation Biology, v. 11, p. 193–203.

Reichard, S., Schuller, R., Isaacson, D., Whiteaker, L., Kruckeburg, A., Kagan, J., Chappell, C., and Old, R., 1997, Nonnative pest plants of greatest concern in Oregon and Washington as of August 1997: Redmond, Oregon, Pacific Northwest Exotic Pest Plant Council, 7 p.

Rejmanek, M., and Pitcairn, M.J., 2002, When is eradication of exotic pest plants a realistic goal, *in* Veitch, C.R., and Clout, M.N., eds., Turning the tide—The eradication of invasive species: Auckland, New Zealand, International Union for Conservation of Nature, Invasive Species Specialist Group, p. 249–253.

Rejmanek, M., and Richardson, D.M., 1996, What attributes make some plant species more invasive?: Ecology, v. 77, p. 1655–1660.

Rejmanek, M., Richardson, D.M., Higgins, S.I., Pitcairn, M.J., and Grotkopp, E., 2005, Ecology of invasive plants—State of the art, *in* Mooney, H.A., Mack, R.N., McNeely, J.A., Neville, L.E., Schei, P.J., and Waage, J.K., eds., Invasive alien species: A new synthesis—SCOPE 63: Washington, D.C., Island Press, p. 104–161.

Richardson, D.M., Pysek, P., Rejmanek, M., Barbour, M.G., Pannetta, F.D., and West, C.J., 2000, Naturalization and invasion of alien plants—Concepts and definitions: Diversity and Distributions, v. 6, p. 93–107.

Schwegman, J., 1994, Exotics of Illinois forests: Erigenia, v. 13, p. 65–67.

Thorp, J.R., and Lynch, R., 2000, The determination of weeds of national significance: Launceston, Australia, National Weeds Strategy Executive Committee, accessed April 3, 2014, at *http://www.weeds.org.au/docs/WoNS/.*

Timmins, S.M., and Owens, S.J., 2001, Scary species, superlative sites—Assessing weed risk in New Zealand's protected natural areas, *in* Groves, R.H., Panetta, F.D., and Virtue, J.G., eds., Weed risk assessment: Collingwood, Australia, CSIRO Publishing, 217–230 p.

Timmins, S.M., and Williams, P.A., 1987, Characteristics of problem weeds in New Zealand's protected natural areas, *in* Saunders, D.A., Arnold, G.W., Burridge, A.A. and Hopkins, A.J.M., eds., Nature conservation and the role of native vegetation: Chipping Norton, Australia, Surrey Beatty and Sons, p. 241–247.

U.S. Department of Agriculture, Animal and Plant Health Inspection Service, 2004, Weed-initiated pest risk assessment: Guidelines for qualitative assessments, accessed June 30, 2006, at *http://www.aphis.usda.gov/plant_health/plant_pest_info/weeds/downloads/wra.pdf.*

U.S. Department of Agriculture, Natural Resources Conservation Service, 2000, National plant materials manual (3d ed.): Title 190, Washington, D.C., U.S. Government Printing Office, p. 540–143—540–151.

Virginia Department of Conservation and Recreation and Virginia Native Plant Society, 2003, Invasive alien plant species of Virginia: Richmond, Virginia Department of Conservation and Recreation, accessed June 30, 2006, at *http://www.dcr.virginia.gov/natural_heritage/invspinfo.shtml.*

Virtue, J.G., Groves, R.H., and Panetta, F.D., 2001, Towards a system to determine the national significance of weeds in Australia, *in* Grove, R.H., Panetta, F.D., and Virtue, J.G., eds., Weed risk assessment: Collingwood, Australia, CSIRO Publishing, p. 124–152.

Wainger, L.A., and King, D.M., 2001, Priorities for weed risk assessment—Using a landscape context to assess indicators of functions, services, and values: *in* Grove, R.H., Panetta, F.D., and Virtue, J.G., eds., Weed risk assessment: Collingwood, Australia, CSIRO Publishing, p. 34–51.

Warner, P.J., Bossard, C.C., Brooks, M.L., DiTomaso, J.M., Hall, J.A., Howald, A., Johnson, D.W., Randall, J.M., Roye, C.L., Ryan, M.M., and Stanton, A.E., 2003, Criteria for categorizing invasive nonnative plants that threaten wildlands: California Exotic Pest Plant Council and Southwest Vegetation Management Association, 24 p., accessed June 30, 2006, at *http://www.cal-ipc.org/ip/inventory/pdf/Criteria.pdf*.

Weiss, J., and McLaren, D., 1999, Invasive assessment of Victoria's State prohibited, priority and regional priority weeds: Frankston, Australia, Keith Turnbull Research Institute, Agriculture Victoria, 16 p.

Williams, P.A., Nicol, E., and Newfield, M., 2001, Assessing the risk to indigenous biota of plant taxa new to New Zealand, *in* Grove, R.H., Panetta, F.D., and Virtue, J.G., eds., Weed risk assessment: Collingwood, Australia, CSIRO Publishing, 100–116 p.

Williamson, M., 1999, Invasions: Ecography, v. 22, p. 5–12.

Predicting Risk of Invasive Species Occurrence—Remote-Sensing Strategies

By Kendal E. Young and T. Scott Schrader

Chapter 6 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figures

Table

Insets

Chapter 6. Predicting Risk of Invasive Species Occurrence—Remote-Sensing Strategies

By Kendal E. Young[1] and T. Scott Schrader[2]

Why Use Remotely Sensed Data?

Knowledge of invasive species occurrence, distribution, and potential invasion pathways is important in developing appropriate long-term monitoring protocols. Costs associated with ground-based visits, however, preclude the National Park Service from inventorying all associated park lands to determine invasive species presence. One potentially cost-effective approach in identifying potential occurrences of invasive species is to predict their distributions by using remotely sensed data and knowledge of species ecology and environmental tolerances (Everitt and others, 1996; Goslee and others, 2003; Osborn and others, 2002; Parker-Williams and Hunt, 2004).Once potential areas of invasive species occurrences are predicted, ground reconnaissance can be more effectively used and applied to an early-detection context for monitoring, verification, and control.

Remote sensing is a means to describe characteristics of an area without physically sampling the area. Remote sensors can be mounted on a satellite, plane, or other airborne structure. Remotely sensed data allow for landscape perspectives on management issues. Sensors measure the electromagnetic energy reflected from an object or area on the Earth's surface (fig. 6.1). These sensors measure energy at wavelengths that are beyond the range of human vision. The guiding principal is that different objects (for example, soils, plants, buildings, water) reflect and absorb light differently at varying wavelengths. Graphically plotting the amount of radiation reflected at a given wavelength provides a unique signature for an object, especially if there is sufficient spectral resolution to distinguish its spectrum from those of other objects (fig. 6.2). Reflectance of clear water is typically low, with initial higher reflectance values in the blue end of the spectrum, which decreases as wavelength increases. Vegetation reflectance is typically low in both the blue and red regions of the spectrum due to absorption by chlorophyll. Because reflectance values peak at the green region, vegetation appears green. In the near infrared (NIR) region, reflectance is much higher than that in the visible bands due to leaf cellular structure. Therefore, vegetation can be identified by the high NIR but generally low visible reflectance. Spectral reflectance curves can be used to discriminate between vegetation types or plant species.

Many Geographic Information System (GIS) data sets are created from remotely sensed data. For example, digital elevation models (DEMs) are derived from space or aircraft-borne sensors. Most GIS software packages can calculate slope and aspect from a DEM. Remotely sensed data can also be used to augment GIS data sets by allowing visual interpretation of images for roads, waterways, fence lines, buildings, and other features. These features can be readily digitized and placed into a GIS environment.

Box 6.1. Use of remotely sensed data. Remotely sensed data are most appropriate for species whose biological characteristics, habitat composition and structure, and landscape context combine to offer a data quality and logistical advantage to ground-based methods (Landenberger and others, 2003).

Remote sensing and GIS technologies were initially developed for different purposes. Traditionally, the two disciplines worked independently, developing new uses for spectral or spatial data (Atkinson and Tate, 1999). Recently, there has been a merging of these two disciplines, as scientists learn the benefits of integrating remote sensing with GIS (Hinton, 1999). Current computer software and hardware facilitate the easy integration of these data sources. Most GIS software packages allow remotely sensed data to be analyzed, or at least viewed. This ability allows the analyst to overlay remote sensing data layers with other spatial data layers. Both spectral and spatial data can provide information about the invasive species occurrences within national parks. As such, integrating remotely sensed data with GIS data shows promise in modeling invasive plant habitats in national parks (Anderson and others, 1996).

Over the last decade, the number of publications pertaining to modeling invasive species by using remotely sensed

[1]U.S. Forest Service, Hathaway Pines, California.

[2]Agricultural Research Service, Jornada Experimental Range, Las Cruces New Mexico.

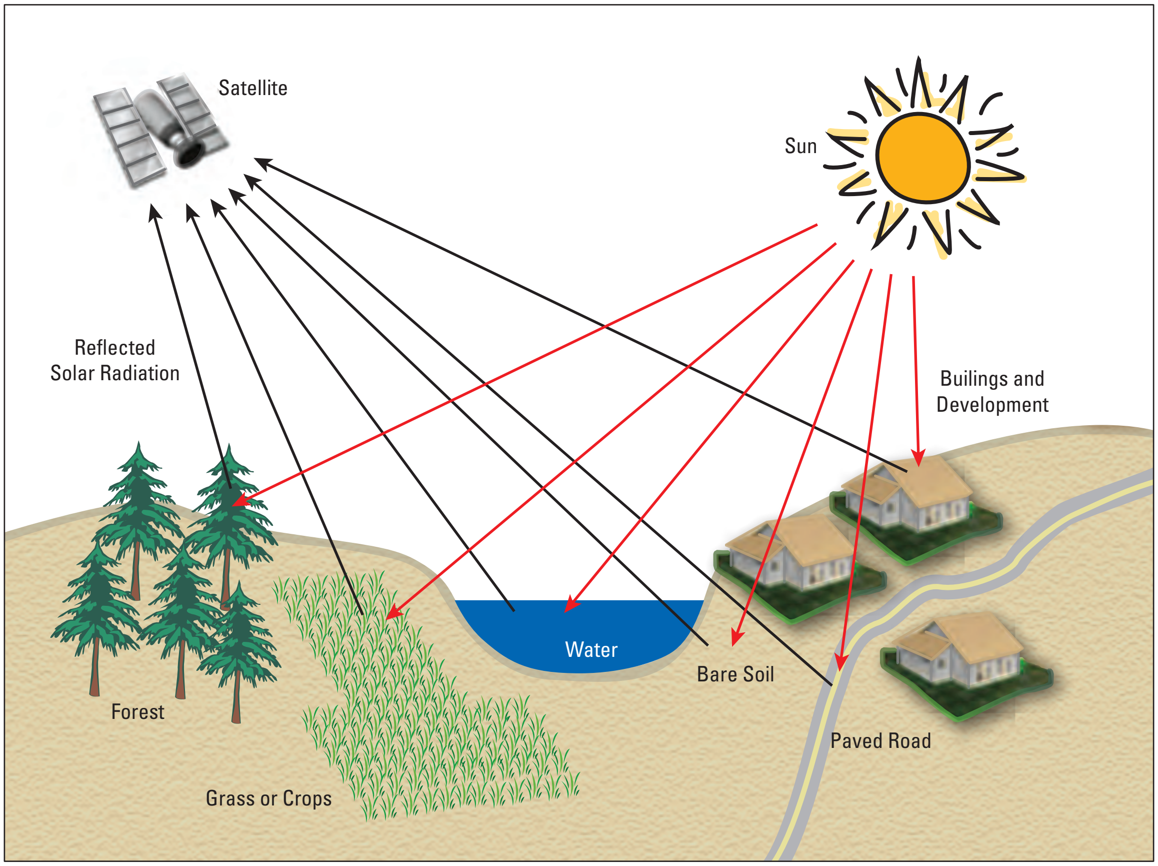

Figure 6.1. Reflected solar radiation, for example, electromagnetic energy, captured by space or aircraft-borne sensors.

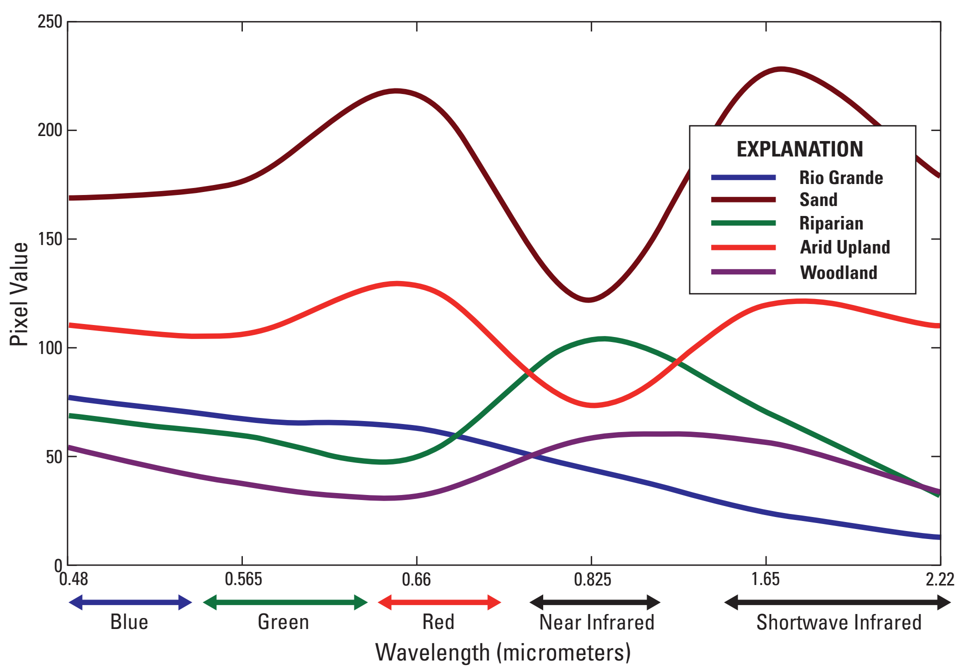

Figure 6.2. Spectral reflectance curves derived from Landsat 7 Enhanced Thematic Mapper Plus for vegetation types in Big Bend National Park, Texas.

data has increased substantially (Chudamani and others, 2004). Information can be found in a variety of journals, technical reports, symposia and workshops, and on the Internet. Journals that support remote sensing publications for invasive species include: Weed Technology, Weed Science, Photogrammetric Engineering and Remote Sensing, Remote Sensing of Environment, International Journal of Remote Sensing, Journal of Ecology, Journal of Range Management, and Journal of Arid Environments.

We provide a review of the utility of using remotely sensed data in modeling the potential distribution of invasive plant species. A complete description of all remote sensing applications for invasive species is beyond the scope of this document. Chudamani and others (2004) and Lass and others (2005) also provide recent overviews on the utility of using remotely sensed data for early detection of invasive plants. For an application of remote sensing techniques to invasive plant early detection in Big Bend National Park, see Chapter 12.

When to Use Remotely Sensed Data

Remotely sensed data may not always be the most effective approach for early detection of invasive plants. There are still limitations associated with the data resolution, processing, and costs. Nevertheless, our ability to detect, map, and monitor invasive plant populations or suitable habitats with remotely sensed data has greatly increased over the last decade. Remotely sensed data could assist in early-detection protocols if:

- The area of consideration is large (park, network, region, or area too large for effective ground surveys).
- Resource managers or contractors have access to GIS capabilities.
- One or a few invasive plants of interest exhibit unique phenological or habitat associations.
- There is a need to prioritize ground survey efforts. Remote sensing can gather information over a wide geographic area in a short amount of time to assist ground surveys.
- There is a need to estimate the likelihood of invasive plant or suitable habitat presence in areas not easily accessible by ground.
- There is a need to evaluate areas where an existing population may spread (early detection in new or adjacent areas). This may arise if an invasive plant occurs outside of a park, network, or other management unit and resource managers need to locate where the plant may start to occur inside the park.
- There is a need to understand which land parcels are most at risk to plant invasion. Risk analyses using GIS and remotely sensed data sets allow for estimates that can cross jurisdictional boundaries.
- There is a need to describe landscape trends prior to invasions or the initiation of monitoring programs. Multiple years of imagery can be analyzed to create a multitemporal data set.

Conversely, resource managers may wish to consider other modeling approaches that assist in early detection (see Chapter 7) if:

- Invasive plant populations are known to be sparse, small, or diffuse patches, which may be the case for early-detection programs. Remote sensing techniques may not be cost effective.
- Complete census for invasive plants is feasible for the area of concern.
- Degree or severity and location of the invasive population are already well known.
- Invasive plant populations are obscured by the overstory vegetation (for example, canopy trees).
- Invasive plant populations do not exhibit unique phenological differences from the surrounding landscape or unique habitat associations.
- There is no access to GIS capabilities or contractors.

NASA Office of Earth Science and the U.S. Geological Survey are developing a National Invasive Species Forecasting System. This system is for early detection and management of invasive species and includes the use of satellite data for invasive species modeling (accessed March 25, 2014, at *http://earthdata.nasa.gov/our-community/community-data-system-programs/reason-projects/invasion-species*). Initiatives such as this will help develop methodologies and models that will overcome existing challenges in using remotely sensed data for invasive plant detection and management.

Spatial, Spectral, and Temporal Scale Issues

The efficacy of remote sensing data for detecting invasive plants or associated habitat is a function of the sensors' spatial and spectral (bandwidth) resolution, and the sensors' repeat cycle. When planning remotely sensed projects, these factors need to be considered with respect to project objectives (Hobbs, 1990). It is often a challenge to balance the scale and resolution of the source data with the information need.

Spatial Considerations

Spatial resolution describes the amount of detail an image contains across a given distance, typically a cell size. The ability to "resolve" or describe small details or objects is one way of describing spatial resolution. As such, smaller objects are typically better "resolved" or detected with high-resolution images. IKONOS and QuickBird images have spatial resolutions less than 5 meters and thus are considered high-resolution images. Conversely, Advanced Very High Resolution Radiometer (AVHRR) is considered a low-resolution image with a spatial resolution of 1.1 kilometers.

Spectral Considerations

There are three generalized categories for sensors with different bandwidths. Panchromatic sensors are sensitive to radiation within a broad wavelength range. The physical quantity measured is the brightness of the object. When the wavelength ranges coincide with the visible range, the resulting image resembles a "black-and-white" photograph. In this case, "color" information is lost. IKONOS, SPOT, and Landsat 7 Enhanced Thematic Mapper Plus (ETM +) each have a panchromatic bandwidth. Multispectral sensors are sensitive to radiation within several narrow wavelength bands. These sensors register reflectance in a number of spectral bands throughout the visible, near- to far-infrared portions of the electromagnetic spectrum. The result is a multilayer image that contains both the brightness and spectral color information of the landscape. These broadband scanners have been successfully applied to discriminate between broad land-cover types. Multispectral sensors include Landsat 7 Enhanced Thematic Mapper Plus (ETM +), SPOT, IKONOS, and QuickBird, among others. Hyperspectral sensors acquire data in more (10 to several 100) but narrower (from 10 to a few nanometers wide) spectral bands than broadband sensors. This precise spectral information allows for capturing finer spectral characteristics (less variability per bandwidth) that yield better identification of objects. NASA Jet Propulsion Laboratory Airborne Visible/Infrared Imaging Spectrometer and the Probe-1 are examples of hyperspectral systems.

Sensors that yield high spatial-resolution data and have hyperspectral capabilities have the highest likelihood of detecting microhabitats or rare plants (Marcus, 2002; Marcus and others, 2003; Lass and others, 2005; Lawrence and others, 2006). However, these data also tend to be expensive, and their relatively small swath size (ground area of the image) requires extensive computer processing time and storage for analyses of large areas. As such, large parks will likely need to compromise bandwidth and spatial resolution to model invasive species occurrences across large landscapes.

Temporal Considerations

Although remotely sensed data are often used for mapping vegetation and general land-cover types, mapping individual plant species imposes many challenges. Similar spectral signatures between the target plant and the surrounding environment, changes in soil color or moisture, and low plant densities hinder discrimination efforts. However, seasonal differences in plant phenology may help in detecting invasive plants. Some invasive plants flower or green-up at a different time than the surrounding vegetation. Multiple image dates allow for detecting these phenological differences between target plants and the surrounding landscape. For example, remote sensing data that coincided specifically with flowering events aided in the detection of yellow starthistle (*Centaurea solstitalis*) (Lass and others, 1996) and leafy spurge (*Euphorbia esula*) (Anderson and others, 1996; Parker-Williams and Hunt, 2002). Likewise, Peters and others (1992) found that broom snakeweed (*Gutierrezia sarothrae*) could be differentiated from grassland species because of its distinct phenological characteristics. As such, the repeat cycle of the sensor may be a key criterion to consider when selecting imagery. Imagery dates should correspond to critical recognition phases of the target plant (Hobbs, 1990; McGowen and others, 2001; Chudamani and others, 2004).

Types of Remotely Sensed Data

There are a variety of remotely sensed data that may be used to detect invasive plants or potential invasive plant habitats (table 6.1). The data chosen for individual projects or parks will depend on several variables, including:

- fiscal considerations,
- goals and objectives,
- park size,
- distribution and patch sizes of current (known) invasive plants,
- invasive species (that is, species phenologically different than the surrounding land cover),
- availability of computing resources, and
- availability of a Remote Sensing Analyst or consultant for image manipulations.

Figure 6.3 provides a general guide to selecting a potential group of sensors to detect invasive species or model their habitats based on park size, differences in plant phenology (target plants that exhibit unique physical characteristics compared to surrounding vegetation), amount of area infested by invasive plants, and the amount of canopy cover of invasive plants. Other considerations may apply when selecting an appropriate remote sensor.

Multispectral, Low Spatial Resolution Sensors

Multispectral, low spatial resolution sensors have limited ability to detect individual invasive plants or small populations. AVHRR sensor is an example. This sensor is a broadband, 4- or 6-channel scanning radiometer, sensing in the visible, near-infrared, and thermal infrared portions of the electromagnetic spectrum. Ground resolution is 1.1 kilometers, which precludes its ability to detect small invasive plant populations.

The AVHRR sensor has useful temporal data, with fairly continuous global coverage since 1979. Two acquisitions are available daily (morning and afternoon).

Table 6.1. Spatial and spectral resolution of remote sensors that may be useful for detecting potential invasive plants or modeling their habitats. This list is not inclusive of all possible remote sensors available as of December 2006.

[µm, micrometers; km, kilometers; m, meters; <, less than]

Sensor	Spatial resolution	Bands	Wavelength (µm)	Color	Swath	Repeat path
Advanced Very High Resolution Radiometer (AVHRR)						
	1.1 km	1	0.58–0.68	Red	2,399 km	2 per day
	1.1 km	2	0.73–1.10	Near infrared	2,399 km	
	1.1 km	3a	1.58–1.64	Mid infrared	2,399 km	
	1.1 km	3b	3.55–3.93	Mid infrared	2,399 km	
	1.1 km	4	10.30–11.30	Thermal infrared	2,399 km	
	1.1 km	5	11.50–12.50	Thermal infrared	2,399 km	
Multispectral Scanner (MSS)						
	80 m	1	0.45–0.52	Blue	185 x 170 km	16 days
	80 m	2	0.52–0.60	Green	185 x 170 km	
	80 m	3	0.63–0.69	Red	185 x 170 km	
	80 m	4	0.76–0.90	Near infrared	185 x 170 km	
Landsat 7 Enhanced Thematic Mapper Plus (ETM +)						
	15 m	Panchromatic	0.52–0.90		185 km	16 days
	30 m	1	0.45–0.52	Blue	185 km	
	30 m	2	0.53–0.61	Green	185 km	
	30 m	3	0.63–0.69	Red	185 km	
	30 m	4	0.75–0.90	Near infrared	185 km	
	30 m	5	1.55–1.75	Shortwave infrared	185 km	
	60 m	6	10.40–12.50	Thermal infrared	185 km	
	30 m	7	2.09–2.35	Shortwave infrared	185 km	
SPOT (5)						
	2.5 or 5 m	Panchromatic	0.48–0.71		60–80 km	26 days
	10 m	1	0.50–0.59	Green	60–80 km	
	10 m	2	0.61–0.68	Red	60–80 km	
	10 m	3	0.78–0.89	Near infrared	60–80 km	
	20 m	4	1.58–1.75	Mid infrared	60–80 km	
IKONOS						
	1 m	Panchromatic	0.53–0.93		70.3 km	Various
	4 m	1	0.45–0.52	Blue	23.9 km	
	4 m	2	0.51–0.60	Green	23.9 km	
	4 m	3	0.63–0.70	Red	23.9 km	
	4 m	4	0.76–0.85	Visible and near infrared	23.9 km	
QuickBird						
	60 cm	Panchromatic	0.45–0.90		16.5 km	3–7 days
	2.4 m	1	0.45–0.52	Blue	16.5 km	
	2.4 m	2	0.52–0.60	Green	16.5 km	
	2.4 m	3	0.63–0.69	Red	16.5 km	
	2.4 m	4	0.76–0.90	Near infrared	16.5 km	
NASA Jet Propulsion Laboratory Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)						
	20 m	224	0.40–2.45 inch	Blue-shortwave infrared	11 km	By request
PROBE-1 Hyperspectral Instrument						
	1–10 m	128	0.40–2.45 inch	Blue-shortwave infrared	<1–6 km	By request

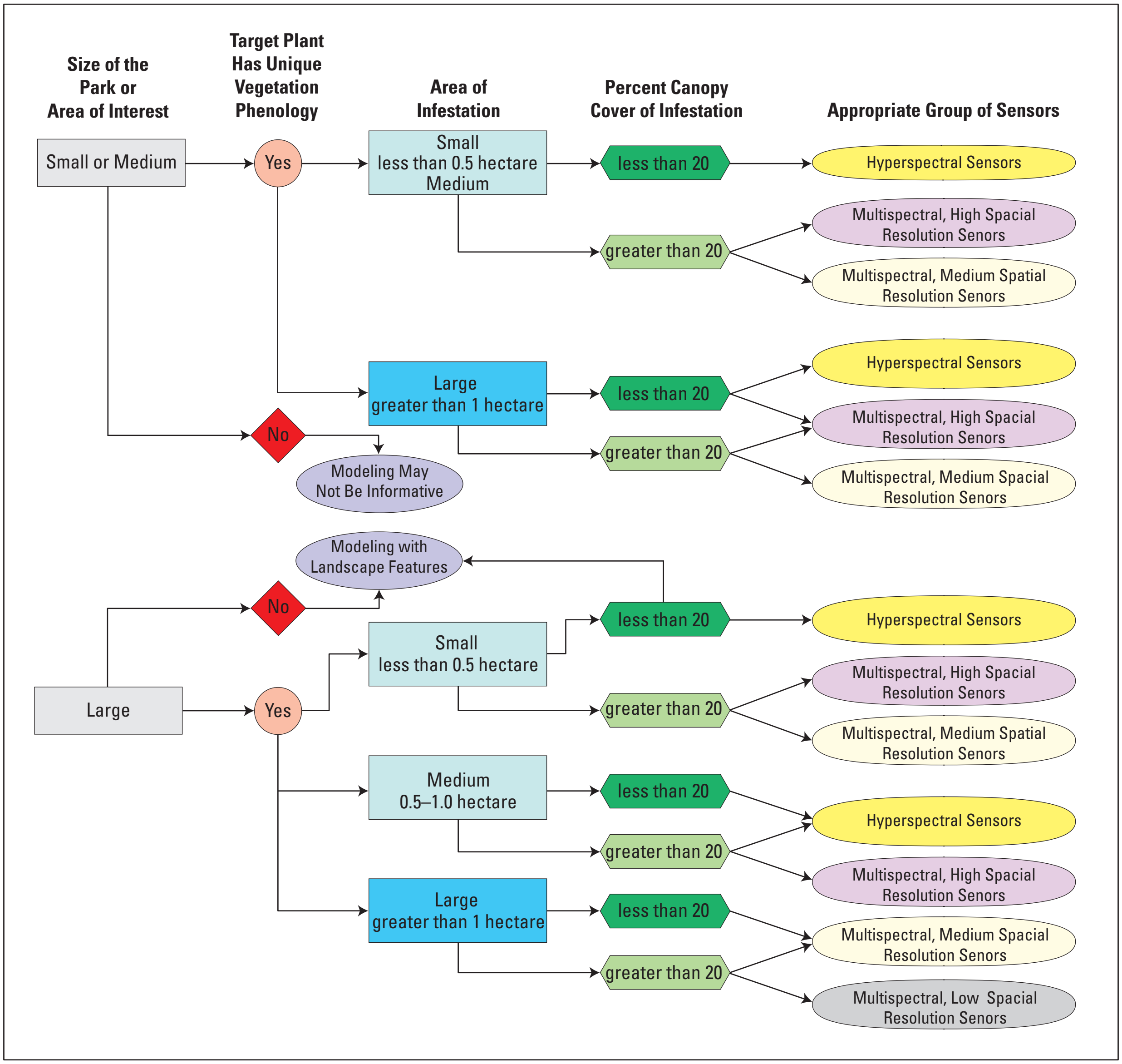

Figure 6.3. General guide to selecting a potential group of sensors to detect invasive species or model their habitats based on park size, differences in plant phenology (target plants that exhibit unique phenology compared to surrounding vegetation), amount of area infested by invasive plants, and the amount of canopy cover of invasive plants. Other considerations may apply when selecting an appropriate remote sensor.

Although AVHRR data cannot detect small invasive plant populations, these data may be useful for detecting landscape changes or evaluating degraded landscapes (Eve and others 1999). Bradley and Mustard (2005) used AVHRR data, time series analyses, to estimate cheatgrass (*Bromus tectorum*) area in the Great Basin. Further, large dense patches of broom snakeweed were detected using this sensor (Peters and others, 1992).AVHRR data are primarily used for investigation of clouds, land-water boundaries, snow and ice extent, cloud distribution, and land and sea surface temperatures. AVHRR data can also be used for studying and monitoring vegetation conditions in ecosystems, including forests, tundra, and grasslands, with applications that include land cover mapping and production of large-area image maps.

Multispectral, Medium Spatial Resolution Sensors

Multispectral, medium spatial resolution sensors such as Multispectral Scanner (MSS), Landsat TM and ETM+, and SPOT have been used extensively to model landscape vegetation types and conduct landscape change analyses. MSS is a multispectral scanning radiometer that was carried onboard Landsats 1 through 5. The instruments provided temporal coverage from July 1972 to October 1992. MSS image data consist of four spectral bands. The resolution for all bands was 79 meters, and the approximate scene size was 185 x 170 kilometers (115 x 106 miles). Due to the spatial resolution, MSS data cannot detect small populations of invasive plants. However, this dataset may be useful for detection of landscape changes and attributes that promote invasive plant establishment (Pickup and others, 1993; Chavez and MacKinnon, 1994; Lass and others, 2005).

The sensors onboard Landsat satellites have varied as technologies have improved and certain types of data proved more useful than others. The sensor on Landsat 7, the Enhanced Thematic Mapper Plus (ETM+), replicates the capabilities of the Thematic Mapper (TM) instruments on Landsats 4 and 5. Landsat 7 ETM+ has the same spectral bands as the previous Landsat sensors, but the ETM+ also features a panchromatic band with 15-meter spatial resolution and a thermal IR channel with 60-meter spatial resolution. A mechanical failure of the scan line corrector onboard Landsat 7 in May 2003 resulted in data gaps from this sensor. Landsat TM and ETM+ data are still available and reasonably priced.

SPOT satellite imaging system was launched in 1986, and currently collects data in green, red, near-infrared, and mid-infrared spectrum. By combining imagery from other SPOT satellites, data are generated at four levels of resolution (20, 10, 5, and 2.5 meters). This multiresolution approach offers users the geospatial information for regional and local analyses. The SPOT imaging system can collect stereo image pairs that contain topographic (3-D) information. SPOT satellites can also be programmed to revisit a given geographic area at any specific time. Imagery required specifically for a vegetation phenology event may be obtained from this service.

Landsat TM and ETM+ have been extensively used to model broad vegetation types. In some cases, these sensors can identify individual plant species with unique spectral or temporal characteristics (Parker-Williams and Hunt, 2002). Dewey and others (1991) compared dyers woad (*Isatis tinctoria*) locations with 60 spectral classes created from Landsat 5 TM data in northern Utah. The authors found strong associations between 10 spectral classes and dyers woad locations. They concluded that their remotely sensed predictive model provided resource managers with a powerful tool for estimating potential dyers woad distributions. Several authors have suggested that Landsat TM and ETM+ data are best used for detecting invasive plants that have patch sizes around 0.5 hectare (1 acre) or larger (Anderson and others, 1993; Everitt and Deloach, 1990; Everitt and others, 1992). McGowen and others (2001) used Landsat 5 TM to map serrated tussock (*Nassella trichotoma*) and scotch thistle (*Onopordum spp*.) in Australia. Detections were limited to areas with infestations greater than 20 percent groundcover. Cheatgrass in the Great Basin was modeled using Landsat TM and ETM+ data (Bradley and Mustard, 2005).

Multispectral, High Spatial Resolution Sensors

Multispectral, high spatial resolution sensors such as IKONOS and QuickBird have less than 5-meter spatial resolution. QuickBird's panchromatic band has a spatial resolution of 60 centimeters. These high spatial resolution sensors show promise for detecting individual species and capturing plant phenological state (Asner and Warner, 2003; Turner and others, 2003; Wang and others, 2004) and mapping shallow aquatic habitats (Mumby and Edwards, 2002). Tsai and others (2005) used QuickBird imagery to accurately map the spatial extent of the invasive horse tamarind (*Leucaena leucocephala*) in southern Taiwan.

Data from these sensors are more expensive than medium spatial resolution sensors. IKONOS and QuickBird also have relatively small swath sizes (around 20 kilometers). Thus, analyses of large areas require extensive computer processing time and storage.

Aerial Photography

Perhaps the oldest remote sensing method is aerial photography (Sabins, 1987; Lillesand and Kiefer, 1994). Historically, the use of aerial photographs was limited to small areas because of the high cost of data acquisition (Lass and others, 2005). Advances in digital aerial photography have improved both the spectral and spatial resolution. Digital cameras can be attached to a variety of aircraft, providing greater flexibility with resolution and timing. However, image preprocessing of raw digital photography presents many challenges (Lass and others, 2005).

There are wide choices of photography with varying degrees of spectral sensitivity (visible and infrared part of the spectrum). Color infrared photography is often called "false-color" photography. Surface objects that are normally red appear green; green objects (except vegetation) appear blue; and "infrared" objects, which typically are not seen with the human eye, appear red. A major use of color infrared photography is for vegetation studies. Green vegetation with active photosynthesis is a strong reflector of infrared radiation and appears bright red on color infrared photographs.

Digital Orthophoto Quadrangles (DOQs) are aerial images produced by the U.S. Geological Survey (USGS). These computer-generated images have been corrected for image displacement caused by terrain relief and camera tilt. DOQs are either grey-scale, natural color, or color-infrared images with 1-meter ground resolution. They cover an area approximately 8 kilometers on each side and have between 50- and 300-meter overlap with adjacent images. This overlap facilitates tonal matching and mosaicking of adjacent images. DOQs have been used for georegistering other imagery or GIS data, visual image interpretation, and on-screen digitizing of landscape features (Coulter and others, 2000; Lawrence and others, 2006). The National Agriculture Imagery Program (NAIP) acquires imagery during the agricultural growing seasons, which enables DOQ acquisition within the same year. NAIP imagery has a

1-meter spatial resolution and is available at *http://www.apfo.usda.gov*. There are also a variety of commercial developers that specialize in obtaining aerial photographs and videography (Lass and others, 2005). The cost of aerial photographs and videography varies depending on the type and resolution of the image and the amount of area surveyed.

Aerial photographs or videography have been found to be useful in detecting saltcedar (*Tamarix ramosissima*) (Everitt and others, 1992, 1996) and leafy spurge (Everitt and others, 1995; Anderson and others, 1996). Further, aerial photographs were used as a tool to estimate the rate of spread for giant hogweed (*Heracleum mantegazzianum*) in the Czech Republic (Mullerova and others, 2005) and for mimosa (Mimosa pigra) in northern Australia (Lonsdale, 1993). Color digital aerial photographs flown in a helicopter 100 meters above ground level were used to census Haleakala silversword (*Argyroxiphium sandwicense*) within a remote volcano crater in Hawaii (Landenberger and others, 2003). The authors found higher errors of commission and omission with aerial photograph census compared to ground census. However, aerial photograph analyses were favorable because they allowed for a larger area to be surveyed with less time than ground-based surveys (Landenberger and others, 2003).

Hyperspectral Sensors

Hyperspectral sensors are perhaps the most helpful group of remote sensors for detecting small populations of invasive plants. These sensors sample the electromagnetic spectrum in narrow, continuous increments, which allows for improved identification of species. There are many hyperspectral sensors available from both governmental and commercial use. Sensors are airborne and may be attached to a variety of aircraft. NASA Jet Propulsion Laboratory Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) (*http://aviris.jpl.nasa.gov/*) has 224 spectral bands that measure light from the blue through the shortwave infrared (0.4 to 2.5 µm with 0.01-µm increments), with a 20-meter resolution. As such, each 20-meter pixel in the image has 224 spectral attributes of data. Likewise, the Hyperion sensor has 220 spectral bands (0.4 to 2.5 µm), with a 30-meter resolution, and the Probe-1 sensor (accessed November 26, 2010 at *http://www.earthsearch.com/index.php*) has 128 spectral bands ranging from 0.4 to 2.5 µm. Spatial resolution varies between 1.0 and 10 kilometers depending on altitude of the aircraft.

Several invasive plants have been detected with the hyperspectral sensors. DiPietro (2002) used AVIRIS data to discriminate riparian vegetation from giant reed (*Arundo donax*) in California. Likewise, AVIRIS data were used to estimate leafy spurge canopy cover and distribution in Wyoming (Parker-Williams and Hunt, 2002, 2004). Lawrence and others (2006) experienced high accuracy using the Probe-1 sensor to detect spotted knapweed (*Centaurea maculosa*) and leafy spurge. Other hyperspectral sensors have been used to detect spotted knapweed (Lass and others, 2002, 2005), yellow starthistle (Miao and others, 2006), and babysbreath (*Gypsophila paniculata*) (Lass and others, 2005). Small infestations of leafy spurge were identified in southeastern Idaho using HyMap (accessed March 22, 2012 at *http://www.hyvista.com/?page_id=440*) hyperspectral data (Glen and others, 2005).

Fiscal and Technical Considerations

Our ability to detect invasive plants over large landscapes is greatly improved by the use of remotely sensed data. The optimal remote sensing data, or combination of data, would have characteristics of hyperspectral sensors and high spatial resolution sensors. Although hyperspectral data facilitates detection of individual plants, hyperspectral data has approximately 75 times greater data volume than an equivalent area using Landsat ETM+ (Thenkabail and others, 2004). Likewise, multispectral, high spatial resolution sensors (for example, IKONOS or QuickBird) also show promise in detecting invasive plants with spatial resolutions less than 5 meters. These sensors are also encumbered by large data volumes over large areas. The new challenge is to develop methods that integrate the required spectral resolution with the ideal spatial resolution and are efficient with the high-dimensional data sets for large area analyses. Remote sensing data sets also come with fiscal and technical expertise considerations. Higher spectral and spatial resolution data are substantially more expensive than multispectral, medium spatial resolution sensors and require greater technical expertise for image processing (fig. 6.4).

Data Processing Considerations

Preprocessing Considerations

Box 6.2. Computational power needed for data processing.

- Fast processor (minimally a 2.0 GHz processor)
- At least 1.0 GB of RAM
- Multiple hard drives (2 or 3 hard drives optimize efficiency)
- At least 5 times the amount of hard drive storage space needed to store 1 copy of the imagery
- Virtual memory directory should have 5 times the amount of space as the RAM.

One of the first considerations to evaluate when beginning a remote sensing project is the availability of ample computation power and storage capacity. Image processing is computationally intensive, in terms of storage (hard drive space and RAM), and CPU(central processing unit) usage. There are no specific rules for the correct processor speed or amount of RAM (random access memory) needed to process and analyze

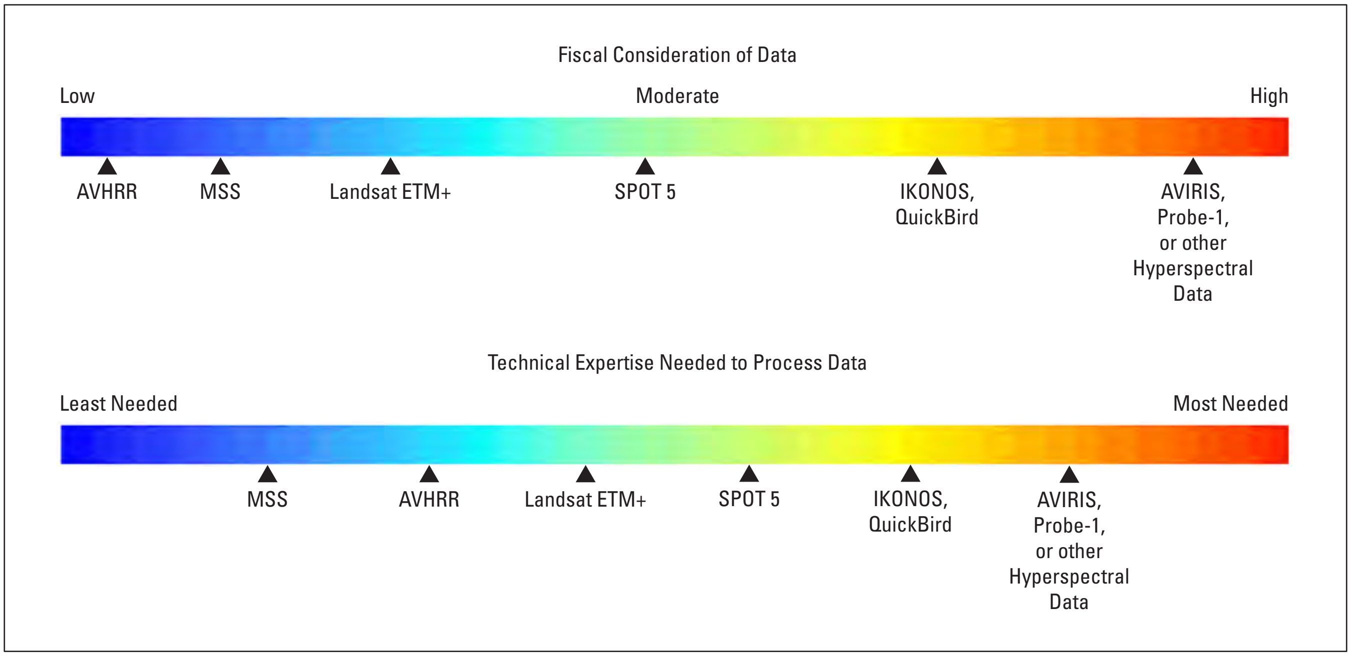

Figure 6.4. Fiscal and technical consideration associated with various types of remotely sensed data (accurate in 2006) (AVHRR, Advanced Very High Resolution Radiometer; MSS, Multispectral Scanner; ETM+ Landsat 7 Enhanced Thematic Mapper Plus; SPOT, Satellite Pour l'Observation de la Terre; AVIRIS, Visible/Infrared Imaging Spectrometer).

remotely sensed data. Generally, the computer should have at least a 2.0 GHz (gigahertz) processor (higher capacity CPUs are better) and at least 1.0 GB (gigabyte) of RAM. To allow for sufficient processing space, the computer used for image analyses should have at least five times the amount of hard drive storage space needed to store one copy of the imagery needed for the area of interest. Computers that contain more than one hard drive will optimize efficiency. One hard drive should be used as the "read from" drive and another as the "write to" drive. If possible, the operating system and virtual memory location should be on a separate hard drive from the processing drives. This would require a third hard drive but would increase processing efficiency. Further, the amount of space on the virtual memory directory should be increased to at least five times the amount of space as the physical memory (RAM).

Box 6.3. Preprocessing steps.

- Extract imagery from storage format
- Consolidate files
- Georeference imagery
- Standardize images, for example, atmospheric standardization
- Mosaic or reduce imagery to area of interest

No matter what remote sensing platform/imagery is chosen to meet research needs, several processing steps may be needed to facilitate the use of the imagery for invasive species analyses. These steps, commonly referred to as "preprocessing" techniques, place imagery in a format that facilitates analyses and reduces errors. Procedures used will vary depending on imagery selected and software used for analyses and processing. Most software packages have designed specific tools to facilitate preprocessing satellite images. Therefore, we discuss general preprocessing steps instead of specific procedures for any one software package. More information regarding the specific preprocessing procedure can be found in the software manuals. Typically, imagery "metadata" will describe processes that have been completed for each dataset and, therefore, indicate steps that remain to be accomplished. Explanations of preprocessing procedures can be found in Lillesand and Kiefer (1994) and Jensen (2005). Aspinall and others (2002) provide an overview to pre-processing and processing considerations with hyperspectral data.

Acquire and Extract Imagery

Box 6.4. Common image software packages.

- ArcView and ArcGIS - *http://www.esri.com/*
- Erdas Imagine - *http://www.hexagongeospatial.com/products/ERDAS-IMAGINE/details.aspx*
- ENVI - *http://www.exelisvis.com/ProductsServices/ENVIProducts.aspx*
- PCI - *http://www.pcigeomatics.com/*
- Idrisi - *http://www.clarklabs.org/*

Satellite imagery can be acquired from numerous sources in various file formats. Common methods of distribution include CD–ROM, DVD, and FTP or HTTP download. Imagery for a particular satellite sensor can be available in several formats that are directly compatible with various imagery manipulation software packages. Flat binary multiband formats are the most common form of storing remotely sensed data. This "generic binary" format is generally accepted by all software packages. The correct choice of format is dictated by the requirements and capabilities of the software used for analyses. Multiple steps commonly are required to place imagery in a format that is suitable for analyses. Typically, the first step is to "import" or extract imagery from a compressed transportation format into a format that the software can understand and display. Some software packages may require an intermediate format, requiring a second conversion from the storage format into a readable format, and then into an analysis format. For example, imagery may arrive in a .bil format (binary, band interleaved by line). This format can be imported and saved (first conversion) in an .img format (Erdas Imagine format). Depending on the software capabilities, the .img format may need a second conversion to a grid (ArcGIS format) for analyses. Although these steps are often trivialized, a considerable amount of time can be expended to extract imagery from its storage media or transportation format into a usable form.

Consolidate and Georeference Imagery

Satellite images may be received in several discrete files, one file for each band of data represented in a scene. For example, one file may represent the red spectrum of an image, another file representing the blue spectrum, and another file representing the green spectrum. Many software packages allow users to combine or "layer stack" these discrete bands of data into one file that contains all layers of information simultaneously. For example, Landsat ETM+ data often come with each band in a separate file. Once these files are combined, the resulting image contains attributes for each band. This consolidation of several data bands into one usable file allows users to efficiently manipulate multispectral data in subsequent processing steps.

Georeferencing is a generic term for the process of defining a spatial coordinate system to each pixel of an image so that it is precisely represented in the location where it is intended to be. The process of defining coordinates to pixels will vary in complexity depending on the precision required for analyses and the type and format of imagery used. In some cases adequate precision can be achieved throughout the image by defining coordinates to each of the corners of the image, such as in Landsat ETM+ imagery. In other cases, such as in the use of some aerial photography, several "ground control points" will need to be acquired to achieve precision within the entire image. Often DOQs are used to geo-reference other remotely sensed data. Defining a coordinate system accurately is essential if several image dates or other GIS data sets are to be analyzed simultaneously.

Atmospheric Standardizations

Procedures to compensate for variations in atmospheric conditions across multiple images and dates are referred to as atmospheric standardization. There are many types of atmospheric standardization routines available to users that, like georeferencing, vary in complexity depending on the precision desired and the type and format of imagery used. Most of the effective procedures attempt to compensate for differences in sun illumination intensity, atmospheric effects (for example, absorptions and refraction), and instrument calibration. These procedures typically normalize pixel values among adjacent image scenes or across multiple dates. Atmospheric standardization procedures can often improve the ability to conjoin adjacent imagery as well as enhance comparisons of multiple images in the same location (time series).

Mosaic or Reduce Imagery

Often, several satellite images will have to be joined because they were acquired in spatial sizes that did not completely represent the desired area for analyses. Several small adjacent scenes can be combined into one large, seamless mosaic image for analyses. The specific routines used for this procedure vary by software package; but in general, this task is accomplished by comparing overlapping pixels in adjacent images and adjusting the values so that each will closely coincide when combined. For this procedure to be effective, adjacent imagery should be acquired within the same day or within the shortest possible time period. It is also advisable to perform some form of image standardization prior to this procedure for the best results.

When one satellite image represents more land area than is required for analyses, it is often desirable to reduce the spatial extent of the image through a "clipping" process. Smaller images will expedite processing and analyses and reduce storage requirements. Clipping can be accomplished with GIS vector data or other raster data. Other related procedures, referred to as "masking," will also reduce required file sizes and subsequent processing time for analysis. Masking is the elimination of unwanted or unneeded pixels within an image. For example, areas that are contaminated by excessive cloud cover or cloud-shadowed areas, or areas that may corrupt analysis, such as snow cover or sun glint on water bodies, can be removed through a masking process. Areas that are removed to enhance image processing time and analyses are dependent on specific goals and objectives of each project.

Box 6.5. File formats.

- Flat binary multiband formats (common form of storing raster data)
  - .BIL (band interleaved by line)
  - .BIP (band interleaved by pixel)
  - .BSQ (band sequential)
- HDF-EOS format (standard data format for all NASA Earth Observing System)
- GeoTIFF format (commonly used generic image format that has spatial information) has spatial.
- ERDAS Imagine (one of the standard remote sensing formats)
  - .LAN format (old image format)
  - .IMG format
- Grid (ArcGIS, Arc/Info)
  - ArcGrid (integer data)
  - GridFloat (Arc floating point data)
- CEOS format (standard data format for radar data)
- MrSID format (Multi-resolution Seamless Image Database)
- Fast L7 (used for Landsat data where each band is self-contained in its own file)
- CAP or DIMAP format (used for SPOT data; may require software package with SPOT data interface)
- SDTS format (Spatial Data Transfer Standard)
- ESRI shape files (standard for vector data)

Attributes for Predictive Models

Box 6.6. Potential attributes for predictive models.

- Spectral reflectance
- Vegetation indices
- Elevation
- Slope
- Aspect
- Habitat associations
- Soil information
- Climatic conditions
- Landscape heterogeneity or degradation
- Anthropogenic locations (roads, fences, building)
- Species traits (flowers, early or late green-up, or senescence)

Many of the examples previously discussed used spectral characteristics as a means to identify invasive plants. Some authors used straightforward visual photograph interpretation methods (Mullerova and others, 2005; Anderson and others, 1996), while others used computer technology and various classification algorithms to classify remotely sensed images. These algorithms are based on spectral reflectance values. Because the spectral reflectance of a given pixel is influenced by the mixture of ground components in that pixel, remote sensing scientists have designed a wide variety of analyses to discriminate spectrally distinct vegetation types. Parker-Williams and Hunt (2002, 2004) applied the Mixture Tuned Matched Filtering (MTMF) classification algorithm to AVIRIS hyperspectral imagery to discriminate leafy spurge with as little as 10 percent canopy cover. Likewise, Glen and others

(2005) used the MTMF algorithm using HyMap hyperspectral imagery. Aspinall and others (2002) provide a summary of hyperspectral imagery classification algorithms used for discriminating invasive plants.

Vegetation indices are classification algorithms commonly used to discriminate plant populations. Most indices assess landscapes in terms of vegetation greenness, soil reflectance, and soil moisture. Examples of vegetation indices that may be useful in discriminating plant populations or attributes that favor invasive plant establishment include:

1. the Normalized Difference Vegetation Index (NDVI) (Lauver and Whistler, 1993),
2. Soil Adjusted Vegetation Index (SAVI) (Huete, 1988),
3. Modified Soil Adjusted Vegetation Index (MSAVI) (Qi and others 1994),
4. tasseled cap transformation,
5. simple band ratios, and
6. the wide dynamic range index (Andries and others, 1994; Tanser and Palmer, 1999; Schmidt and Karnieli, 2000; Gitelson, 2004).

Unfortunately, there is not one index that works best for all vegetation communities. Parks will need to evaluate multiple indices to determine which is most useful for their area and objectives. Vegetation indices can be derived from a variety of sensors, including AVHRR (Bradley and Mustard, 2005), Landsat TM (McGowen and others, 2001; Goward and others, 2003), IKONOS and QuickBird (Asner and Warner, 2003; Goward and others, 2003; Tsai and others, 2005), and hyperspectral data (Thenkabail and others, 2004).

Recent technologies in discriminating landscape features include texture analysis and object-oriented analysis. Image texture analysis investigates the structural and statistical properties of spatial patterns on images (Tsai and others, 2005). Object-oriented analyses classify an image based on attributes of an image object rather than attributes of individual pixels (Benz and others, 2004). Homogeneous pixels are aggregated into image objects using their spatial (size, shape, location) or spectral characteristics (Laliberte and others, 2004). These relatively new approaches show great promise for detecting invasive plants, especially when combined with spectral data.

While remotely sensed data can detect plant species and populations, some researchers have combined remotely sensed data with other spatial data sets to enhance predictive models. Anderson and others (1996) used aerial photograph interpretation, roads, trails, hydrographic data, slope, and aspect to model leafy spurge in Theodore Roosevelt National Park. The authors noted that using GIS and remote sensing data together proved to be a powerful combination of tools. Yellow starthistle was predicted using landscape variables such as elevation, slope, aspect, and Landsat land-use classifications (Shafii and others, 2003).

Remotely sensed data can also be used to assess attributes associated with invasive plant presence. For example, many invasive plants are associated with fragmented or degraded landscapes (Sakai and others, 2001; With, 2004). Tanser and Palmer (1999) used a measurement of landscape heterogeneity to assess degradation. A moving standard deviation filter was passed over Landsat TM imagery creating a moving standard deviation index (MSDI). Degraded landscapes exhibited higher MSDI values than undisturbed landscapes. Other environmental variables that are associated with landscape heterogeneity and appear to have high potential for remote sensing include total vegetation cover, relative proportion of grass and shrub cover, and organic soil cover (Warren and Hutchinson, 1984; Schmidt and Karnieli, 2000).

Species characteristics have been used to explain invasion patterns (Rejmanek and Richardson, 1996; Goodwin and others, 1999). For example, vegetative propagation, leaf size, flowering period, and wind dispersal were associated with invasive plant abundances on five Mediterranean islands (Lloret and others, 2005). However, inclusion of species characteristics into spatially explicit predictive models is hindered by the lack of strong associations between species characteristics and spatial data (see Chapter 7). Some species traits may be detected by remotely sensed data. For example, invasive species that have flowers and (or) bracts, green-up or senesce at a different time than plants in the surrounding environment, have a unique canopy architecture or growth form, or have a unique coloration are good candidates for using species-related traits in predictive modeling.

Attributes chosen to predict invasive species distributions may be direct, indirect, or "models of models" (see Chapter 7). In terms of landscape studies, ecological parameters are generally sampled from GIS or remotely sensed data. Close scrutiny of coarse resolution variables may be warranted, as these variables may introduce spatial uncertainties from interpolation errors, lack of sufficient ground data, and poor associations with causal factors (Guisan and Zimmermann, 2000). Variables that have little to no direct physiological relevance for a species' performance (slope, aspect, elevation, or topographic position) are easily measured from field or spatial data sets and are often used because of their good correlation to observed species patterns. Models constructed with these resource variables are typically general but are applicable over larger areas. Given the complexity of natural landscapes, spatially explicit predictive habitat models are generally a compromise between precision and generality (Guisan and Zimmermann, 2000). Chapter 7 provides greater discussion on attributes for species distribution models and their interpretation limitations.

Building Predictive Models

A variety of analytical methods could be used to construct predictive models for invasive species, ranging from simple overlays to more statistically driven models. Simple models can be developed directly within a GIS by using overlays of environmental variables. Boolean approaches are modeling methods that use overlay rules (where individual layers are added, subtracted, or multiplied together). These approaches have been widely used in wildlife habitat relation (WHR) models (for example, GAP models). WHR models describe resources and conditions present in areas where a species persists and reproduces or otherwise occurs. These modeled relations predict, and spatially depict, areas of potentially suitable habitat. While these types of models are informative, they may lack the ability to make statistical inferences.

Statistical approaches to constructing species distribution models (SDMs) will be discussed in Chapter 7. These approaches also apply to remote sensing and GIS data and can be more desirable than simple overlay models. Statistical approaches to model building allow enhanced accuracy and predictive power. Unfortunately, GIS software packages still lack important statistical procedures for predictive purposes. Some software packages have modules for classification and regression trees (CART), clustering analyses, logistic regression, and various other supervised classification procedures. However, statistical analyses in the GIS environment may be limited by the current lack of model selection procedures available in the software (for example, stepwise selection procedure for logistic regression) (Guisan and Zimmermann, 2000). Many statistical software analyses can be easily implemented into a GIS. For example, Generalized Linear Models (GLMs) and logistic regression analyses can be placed in a GIS by multiplying each regression coefficient with its related predictor variable layer. Most GIS software packages allow users to write algorithms for image manipulations. Alternatively, spatial data sets can be exported as ASCII files and analyzed in a variety of spatial programs. Guisan and Zimmermann (2000) and Scott and others (2002) provide additional insights into predictive model approaches. Unfortunately, there is no one analytical method that works for all scenarios. The appropriate analytical method for constructing predictive species models will be a function of park goals and objectives, the type and structure of the data, software availability, and expertise in statistical and GIS modeling.

The analytical method used to construct predictive models will dictate the type of model produced. Predictive models may display:

- probabilities of occurrence (derived from logistic GLM analyses),
- the most probable abundance (derived from ordinal GLM analyses),
- predicted occurrence (based on nonprobabilistic metrics), or
- the most probable entity (from hierarchical analyses) (Guisan and Zimmermann, 2000).

Regardless of the type of model, assessing model performance is essential in preparing adequate models. There are three forms of assessment for spatial models:

- draft models that are verified,
- committed (final) models that are assessed for accuracy, and
- final models subjected to user validation.

Draft models are produced during an evolutionary and refinement process that involves iterative collection and testing with verification data. **Model verification** does not examine the accuracy of the model or the usefulness of the model. Model verification examines only the model's internal consistency (Conroy and Moore, 2002). This is analogous to "measures of model fit" discussed in Chapter 9. Subsequent to this iterative process, models are committed to a final form, which is the version that is subjected to accuracy assessment. No further alteration of a committed model is permitted after accuracy assessment; if further alterations are performed, then the model is a new version that requires additional assessment to provide an applicable accuracy statement. Final models and associated accuracy statements are published for use by others. Model validation is performed by, and arises from, judgments of intended users. Model validation depends primarily on the goals of the users rather than on statistics alone (Guisan and Zimmermann, 2000). See Chapter 9 for a greater discussion on model validation.

Model performance can be evaluated by:

- using two independent sets of data for building and evaluation (often called "training" and "evaluation" data),
- cross-validation procedures where the dataset is separated into two sets, a training set, and evaluation set,
- jack-knife procedures that resample the dataset based on deleting a portion of the original observations or input model variables in subsequent samples,
- bootstrap techniques that perform repeated random sampling with replacement from an original sample,
- randomization procedures where random samples are obtained by sampling without replacement, and
- resubstitution procedures where the same dataset is used for training and testing, with no partitioning of data.

Detailed explanations on these procedures, and others, are provided by Efron and Tibshirani (1993), Fieldings and Bell (1997), and Guisan and Zimmermann (2000). See also Chapter 9 for more information on model evaluation and assessment.

From an applied perspective, there are two ways a habitat model can be inaccurate:

- the model can overpredict, rating locations suitable although the location is unsuitable, or the species has not been detected in the predicted location (type I error),

- or the model can fail to predict the presence of a species where it is indeed present (type II error) (Shrader-Frechette and McCoy, 1993).

The first represents an error of commission. The second case represents an error of omission. Both types of errors undermine the defensibility of the model. However, errors of omission are less acceptable than errors of commission if we wish to implement a precautionary principle in biological conservation (Shrader-Frechette and McCoy, 1993). For example, if the predictive model and field assessment conclude an area is suitable, and the species was not present, then the area would be a good candidate to monitor for potential expansion of the species. Errors of commission and omission are often summarized in confusion or error matrices that cross-tabulate the observed and predicted presence/absence patterns (see Chapter 9). These matrices can be summarized into percent agreement and disagreement or further analyzed to evaluate the association between observed and predicted. For example, the correlation coefficient often is used to measure the strength of a relation between two variables that are normally distributed. The kappa statistic (Cohen, 1960) provides a measurement of the proportion of chance or expected agreement. The gamma coefficient, Spearman *R*, or Kendall *tau* are nonparametric equivalents to the standard correlation (Goodman and Kruskal, 1979; Zar, 1984). Other model performance assessments are discussed in Chapter 9, including threshold-dependent and threshold-independent measures.

Nature is heterogeneous and often difficult to predict accurately from a single, although complex, model. As such, model assessments over a wider range of situations (in space and time) will provide a better definition of the range of applications for which the model predictions are suitable (Guisan and Zimmermann, 2000).

Landscape Risk Assessments to Prioritize Areas for Conservation Efforts

Understanding where to concentrate survey efforts to find new species or expanding populations of existing species is paramount to early-detection protocols. Most natural resource managers prioritize survey efforts based on management considerations and documented predictors of invasiveness. However, given the variability associated with species traits, climatic events, and landscape characteristics, no set of conditions or traits exists that can be universally applied to accurately characterize all successful invasions (Alpert and others, 2000). Risk-assessment procedures can assist natural resource managers in prioritizing areas for conservation efforts. Risk assessment for invasive species is the process of obtaining quantitative or qualitative measures of risk levels by incorporating a broad array of information describing factors that may influence the distribution of invasive species (Allen and others, 2006). There are several approaches to modeling risk of invasions. For example, neutral landscape models, which evaluate flows through spatially heterogeneous landscapes, were used to assess the risk of invasions in fragmented landscapes (With, 2004). Landis (2004) describes a relative risk model that incorporates a system of numerical ranks and weighting of factors that may influence the distribution of invasive species. The Landis (2004) risk model takes into account the spatial relations of the locations of species introductions, migration paths, and the habitat structure or suitability. As such, modeling potentially suitable habitat and migration paths (potential vectors and pathways) for the introduction or spread of invasive species is an important component in conducting risk assessments in parks and other natural areas. Vectors refer to the mechanism of plant introduction, while pathways refer to the route taken. Examples of vectors include wind, water, and animals (Sakai and others, 2001). Examples of pathways include roads, trails, and waterways. Discussions on invasive plant vectors and pathways are presented in Chapter 2. Remote sensing and GIS data sets can model potential habitats, vectors, and pathways to allow for landscape-scale risk assessments.

Few communities are impenetrable to invasion by non-native species, but communities differ in their susceptibility to invasion (Sakai and others, 2001). Although it is difficult to generalize about invasive species dispersal across landscapes (Tackenberg, 2003), repeated introductions increase the chances of establishment (Sakai and others, 2001; Perrings and others, 2002). In many parks and other natural areas, roads and waterways are perhaps the pathways of most concern. These pathways enhance species invasions by acting as dispersal corridors, providing suitable habitat, and containing reservoirs of propagules (Parendes and Jones 2000). Disturbance along roads by vehicle traffic and maintenance activity (for example, road grading, ditch clearing, and trimming of overhanging vegetation) is often the source of repeated introductions. Waterway disturbances occur from floods and associated transport of sediment.

GIS data sets for invasive species pathways are readily accessible. Many parks and natural areas already have road, trail, and waterway GIS layers. Road and hydrologic layers are available from a variety of GIS data clearinghouses, Federal agencies, and private companies. Digital line graphs (DLGs) (digital vector data derived from USGS maps), Digital Raster Graphics (DRGs) (scanned digital images of USGS topographic quadrangles), and National Hydrography Data sets (NHD) can be downloaded from the USGS Website *http://eros.usgs.gov/*.

Spatial data sets on potential invasive species vectors are not as readily accessible and would likely have to be created specifically for the target species and the area of interest. Many authors have analytically modeled potential vectors, especially seed dispersal by wind. Schurr and others (2005) created a mechanistic model for secondary seed dispersal by wind (the wind-driven movement of seeds along the ground surface). The authors found a relation between seed dispersal and seed size but noted that the model tended to underestimate dispersal rates. Tackenberg (2003) also modeled seed dispersal by wind and found that long-distance dispersal was primarily influenced by weather conditions that yielded thermal turbulence and convective updrafts. Tackenberg (2003) noted that the inclusion of topography in estimating dispersal rates is important, even in landscapes that exhibit only small differences in elevation and slight slopes. In addition, Campbell and others (2002) simulated landscape-scale invasions of plants that use rivers to transport propagules.

Box 6.7. Benefits of remote sensing.

- Remote-sensing technologies look beyond the human view in the electromagnetic spectrum, which allows better detection of vegetation.
- Remote sensors allow for regional analyses.
- Cost savings for large parks.
- Provides a means for prioritizing ground surveys.
- Facilitates repeat analyses and change detections.

Few studies have created spatial models of vectors or pathways. Favorable predictive spatial models incorporate:

- invasive species distribution data,
- population rates,
- factors influencing the number of propagules,
- dispersal modes,
- landscape structure,
- ecological processes, and
- statistically explained patterns (Moody and Mack, 1998; Higgins and others, 1996; Higgins and Richardson, 1999; Wadsworth and others, 2000; Sakai and others, 2001).

Summary

Remotely sensed images have a number of features that make them ideal for predicting invasive species in parks and other natural areas. Remote-sensing technologies look beyond the human view, in the electromagnetic spectrum, which allows better detection of vegetation. Remote sensors allow for regional analyses that would be cost-prohibitive using ground-based visits. Regional analyses allow for prioritizing ground-reconnaissance visits to survey, control, or eradicate potential invasive plant populations. Further, the ease of securing temporal data allows for repeat analyses and change detections.

Our ability to detect invasive plants using remotely sensed data has increased with improved sensors, computer technology, and classification techniques (Lass and others, 2005). Although integrating remotely sensed data with other spatial data sets enhances our abilities to model invasive plants, detecting small or sparse plant populations is still hampered by spatial and spectral resolution and by our limited ability to analyze large data sets. Data sets that have the highest likelihood of detecting invasive plants come with high fiscal and technical considerations. Overall, the use of remotely sensed data will be most appropriate for species whose biological characteristics, habitat composition and structure, and landscape context combine to offer a data quality and logistical advantage to ground-based methods (Landenberger and others, 2003).

Parks and other natural areas need predictive models that can help in setting priorities for control of invasive species and predicting the potential of future invasions. Remotely sensed data can aid in the development of spatially explicit predictive habitat models and estimates of distributional vectors and pathways (see Chapter 12 for an example). This information will provide land managers with early-detection tools, a means to evaluate current and future control needs, and a means to prioritize conservation efforts. Early-detection methods increase our ability to eradicate invasive plants and reduce costs of control (Rejmanek and Pitcairn, 2002).

Recommended Reading

- Introductory digital image processing—A remote sensing perspective (Jensen, 2005).
- Introduction to Remote Sensing (Campbell, 2002).
- Remote sensing and image interpretation (Lillesand and Kiefer, 1994).

References Cited

Allen, C.R., Johnson, A.R., and Parris, L., 2006, A framework for spatial risk assessments—Potential impacts of nonindigenous invasive species on native species: Ecology and Society, v. 11, p. 39.

Alpert, P., Bone, E., and Holzapfel, C., 2000, Invasiveness, invasibility and the roll of environmental stress in the spread of non-native plants: Perspectives in Plant Ecology, Evolution and Systematics, v. 3, p. 52–66.

Anderson, G.L., Hanson, J.D., and Hart, G.F., 1993, Using satellite data to map false broomweed (*Ercameria austrotexana*) infestations on south Texas rangelands: Weed Technology 7, p. 865–871.

Anderson, G.L., Everitt, J.H., Escobar, D.E., Spencer, N.R., Andrascik, R.J., 1996, Mapping leafy spurge (*Euphorbia esula*) infestations using aerial photography and geographic information systems: Geocarto International, v. 11, p. 81–89.

Andries, A.M., Gulinck, H., and Herremans, M., 1994, Spatial modeling of the barn owl (*Tyto alba*) habitat using landscape characteristics derived from SPOT data: Ecography, v. 17, p. 278–287.

Asner, G.P., and Warner, A.S., 2003, Canopy shadow in IKONOS satellite observations of tropical forest and savannas: Remote Sensing of Environment, v. 87, p. 521–533.

Aspinall, R.J., Marcus, W.A., and Boardman, J.W., 2002, Considerations in collecting, processing and analysis of high spatial resolution hyperspectral data for environmental investigations: Journal of Geographical Systems, v. 4, p. 15–29.

Atkinson, P.M., and Tate, N.J., 1999, Techniques for the analysis of spatial data, *in* Atkinson, P.M., and Tate, N.J., eds., Advances in remote sensing and GIS analysis: New York, John Wiley & Sons, Inc., p. 1–5.

Benz, U., Hofmann, C., Willhauck, G., Lingenfelder, I., and Heynen, M., 2004, Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information: Photogrammetry and Remote Sensing, v. 58, p. 239–258.

Bradley, B.A., and Mustard, J.F., 2005, Identifying land cover variability distinct from land cover change—Cheatgrass in the Great Basin: Remote Sensing of Environment, v. 94, p. 204–213.

Campbell, J.B., 2002, Introduction to remote sensing: New York, The Guilford Press, 620 p.

Campell, G.S., Blackwell, P.G., and Woodward, F.I., 2002, Can landscape-scale characteristics be used to predict plant invasions along rivers?: Journal of Biogeography, v. 29, p. 535–543.

Chavez, P.S., and MacKinnon, D.J., 1994, Automatic detection of vegetation changes in the southwestern United States using remotely sensed images: Photogrammetric Engineering and Remote Sensing, v. 60, p. 571–582.

Chudamani, J., de Leeuw, J., and van Duren, I.C., 2004, Remote sensing and GIS applications for mapping and spatial modeling of invasive species *in* XXth ISPRS Congress, July 12–23, 2004, Commission VII: Istanbul, Turkey, Geo-Imagery Bridging, p. 6679–667.

Cohen, J., 1960, A coefficient of agreement for nominal scales: Educational and Psychological Measurement, v. 20, p. 37–46.

Conroy, M.J., and Moore, C.T., 2002, Wildlife habitat modeling in an adaptive framework—The role of alternative models, *in* Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., and Samson, F.B., Predicting species occurrences—Issues of accuracy and scale: Washington, D.C., Island Press, p. 205–218.

Coulter. L., Stow, D., Hope, A., O'Leary, J., Turner, D., Longmire, P., Peterson, S., and Kaiser, J., 2000, Comparison of high spatial resolution imagery for efficient generation of GIS vegetation layers: Photogrammetric Engineering and Remote Sensing, v. 66, p. 1329–1336.

Dewey, S.A., Price, K.P., and Ramsey, D., 1991, Satellite remote sensing to predict potential distribution of dyers woad (*Isatis tinctoria*): Weed Technology, v. 5, p. 479–484.

DiPietro, D.Y., 2002, Mapping the invasive plant *Arundo donax* and associated riparian vegetation using hyperspectral remote sensing: Master's thesis, Davis, University of California, 45 p.

Efron, B., and Tibshirani, R., 1993, An introduction to the bootstrap: New York, Chapman and Hall, 456 p.

Eve, M.D., Whitford, W.G., and Havstad, K.M., 1999, Applying satellite imagery to triage assessment of ecosystem health: Environmental Monitoring and Assessment, v. 54, p. 205–227.

Everitt, J.H. and Deloach, C.J., 1990, Remote sensing of Chinese Tamarisk (*Tamarix chinensis*) and associated vegetation: Weed Science, v. 38, p. 273–278.

Everitt, J.H., Anderson, G.L., Escobar, D.E., Davis, M.R., Spencer, N.R., and Andrascik, R.J., 1995, Use of remote sensing for detecting and mapping leafy spurge (*Euphorbia esula*): Weed Technology, v. 9, p. 599–609.

Everitt, J.H., Alaniz, M.A., and Escobar, D.E., 1992, Using remote sensing to distinguish common goldenweed (*Isocoma coronopifolia*) and Drummon goldenweed (*Isocoma drummondii*): Weed Science, v. 40, p. 621–628.

Everitt, J.H., Escobar, D.E., Alaniz, M.A., Davis, M.R., and Richerson, J.V., 1996, Using spatial information technologies to map Chinese tamarisk (*Tamarix chinensis*) infestations: Weed Science, v. 44, p. 194–201.

Fieldings, A.H., and Bell, J.F., 1997, A review of methods for the assessment of prediction errors in conservation presence absence models: Environmental Conservation, v. 24, p. 38–49.

Gitelson, A.A., 2004, Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation: Journal of Plant Physiology, v. 161, p. 165–173.

Glen, N.F., Mundt, J.T., Weber, K.T., Prather, T.S., Lass, L.W., and Pettingill, J., 2005, Hyperspectral data processing for repeat detection of small infestations of leafy spurge: Remote Sensing of Environment, v. 95, p. 399–412.

Goodman, L.A., and Kruskal, W., 1979, Measures of association for cross classifications: New York, Springer-Verlag, 146 p.

Goodwin, B.J., McAllister, A.J., and Fahrig, L., 1999, Predicting invasiveness of plant species based on biological information: Conservation Biology, v. 13, p. 422–426.

Goslee, S.C., Beck, K.F., and Peters, D.P.C., 2003, Distribution of Russian knapweed in Colorado—Climate and environmental factors: Journal of Range Management, v. 56, p. 206–212.

Goward, S.N., Davis, P.E., Fleming, D., Miller, L., and Townshend, J.R., 2003, Empirical comparison of Landsat 7 and IKONOS multispectral measurements for selected Earth Observation System (EOS) validation sites: Remote Sensing of Environment, v. 88, p. 80–99.

Guisan, A., and Zimmermann, N.E., 2000, Predictive habitat distribution models in ecology: Ecological Modelling, v. 135, p. 147–186.

Higgins, S.I., and Richardson, D.M., 1999, Predicting plant migration rates in a changing world—The role of long-distance dispersal: The American Naturalist, v. 153, p. 464–475.

Higgins, S.I., Richardson, D.M., and Cowling, R.M., 1996, Modeling invasive plant spread—The role of plant-environment interactions and model structure: Ecology, v. 77, p. 2043–2054.

Hinton, J.C., 1999, Image classification and analysis using integrated GIS, *in* Atkinson, P.M., and Tate, N.J., eds., Advances in remote sensing and GIS analysis: New York, John Wiley & Sons, Inc., p. 207–218.

Hobbs, R.J., 1990, Remote sensing of spatial and temporal dynamics of vegetation, *in* Hobbs, R.J., and Mooney, H.A., eds., Remote sensing of biosphere functioning: New York, Springer-Verlag, p. 203–219.

Huete, A.R., 1988, A soil-adjusted vegetation index (SAVI): Remote Sensing of Environment, v. 25, p. 53–70.

Jensen, J.R., 2005, Introductory digital image processing: A remote sensing perspective, (3d ed.): Upper Saddle River, New Jersey, Prentice Hall, 526 p.

Laliberte, A.S., Rango, A., Havstad, K.M., Paris, J.F., Beck, R.F., McNeely, R., and Gonzalez, A.L., 2004, Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico: Remote Sensing of Environment, v. 93, p. 198–210.

Landenberger, R.E., McGraw, J.B., Warner, T.A., and Brandtberg T., 2003, Potential of digital color imagery for censusing Haleakala silverswords in Hawaii: Photogrammetric Engineering and Remote Sensing, v. 69, p. 915–923.

Landis, W.G., 2004, Ecological risk assessment conceptual model formulation for nonindigenous species: Risk Analysis, v. 24, p. 847–858.

Lass, L.W., Carson, H.W., and Callihan, R.H., 1996, Detection of yellow starthistle (*Centaurea solstitialis*) and common St. Johnswort (*Hypericum perforatum*) with multispectral digital imagery: Weed Technology, v. 10, p. 466–474.

Lass, L.W., Prather, T.S., Glenn, N.F., Weber, K.T., Mundt, J.T., and Pettingill, J., 2005, A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (*Centaurea maculosa*) and babybreath (*Gypsophila paniculata*) with a hyperspectral sensor: Weed Science, v. 53, p. 242–251.

Lass, L.W., Thill, D.C., Shafii, B., and Prather, T.S., 2002, Detecting spotted knapweed (*Centaurea maculosa*) with hyperspectral remote sensing technology: Weed Technology, v. 16, p. 426–432.

Lauver, C.L., and Whistler, J.L., 1993, A hierarchical classification of Landsat TM imagery to identify natural grassland areas and rare species habitat: Photogrammetric Engineering and Remote Sensing, v. 59, p. 627–634.

Lawrence, R.L., Wood, S.D., and Sheley, R.L., 2006, Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest): Remote Sensing of Environment, v. 100, p. 356–362.

Lillesand, T.M., and Kiefer, R.W., 1994, Remote sensing and image interpretation: New York, John Wiley & Sons, Inc., 750 p.

Lloret, F., Medail, F., Brundu, G., Camarda, I., Moragues, E., Rita, J., and Hulme, P., 2005, Species attributes and invasion success by alien plants on Mediterranean islands: Journal of Ecology, v. 93, p. 512–520.

Lonsdale, W.M., 1993, Rates of spread of an invading species *Mimosa pigra* in northern Australia: Journal of Ecology, v. 81, p. 513–521.

Miao, X., Gong, P., Swope, S., Pu, R., Carruthers, R., Anderson, G.L., Heaton, J.S., and Tracy, C.R., 2006, Remote Sensing of Environment, v. 101, p. 329–341.

Marcus, W.A., 2002, Mapping of stream microhabitats with high spatial resolution hyperspectral imagery: Journal of Geographical Systems, v. 4, p. 113–126.

Marcus, W.A., Lefleiter, C.J., Aspinall, R.J., Boardman, J.W., and Crabtree, R.L., 2003, High spatial resolution hyperspectral mapping of in-stream habitats, depths, and woody debris in mountain streams: Geomorphology, v. 55, p. 363–380.

McGowen, I.J., Frazier, P., and Orchard, P., 2001, Mapping serrated tussock and Scotch thistle in temperate pastures using Landsat TM imagery, *in* Institute of Electrical and Electronics Engineers 2001 International Geoscience and Remote Sensing Symposium—Scanning the Present and Resolving the Future: July 9–13, 2001, Sydney, Australia, University of New South Wales, p. 1267–1279.

Moody, M.E., and Mack, R.N., 1998, Controlling the spread of plant invasions—The importance of nascent foci: Journal of Applied Ecology, v. 25, p. 1009–1021.

Mullerova, J., Pysek, P., Jarosik, V., and Pergl, J., 2005, Aerial photographs as a tool for assessing the regional dynamics of the invasive plant species *Heracleum mantegazzianum*: Journal of Applied Ecology, v. 42, p. 1042–1053.

Mumby, P.J., and Edwards, A.J., 2002, Mapping marine environments with IKONOS imagery—Enhanced spatial resolution can deliver greater thematic accuracy: Remote Sensing of Environment, v. 82, p. 248–257.

Osborn, S., Wright, V., Walker, B., Cilimburg, A., and Perkins, A., 2002, Linking wilderness research and management—Volume 4. Understanding and managing invasive plants in wilderness and other natural areas—An annotated reading list: General Technical Report RMRS–GTR–79–Vol 4: Fort Collins, Colo., U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 65 p.

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

Parker-Williams, A.E., and Hunt, E.R., 2002, Estimation of leafy spurge cover from hyperspectral imagery using mixture tuned matched filtering: Remote Sensing of Environment, v. 82, p. 446–456.

Parker-Williams, A.E., and Hunt, E.R., 2004, Accuracy assessment for detection of leafy spurge with hyperspectral imagery: Journal of Range Management, v. 57, p. 106–112.

Perrings, C., Williamson, M., Barbier, E.B., Delfino, D., Dalmazzone, S., Shogren, J., Simmons, P., and Watkinson, A., 2002, Biological invasion risks and the public good—An economic perspective: Conservation Ecology, v. 6, 7 p., accessed June 30, 2006, at *http://www.consecol.org/vol6/iss1/art1*.

Peters, A.J., Reed, B.C., and Eve, M.D., 1992, Remote sensing of broom snakeweed (*Gutierrezia sarothrae*) with NOAA-10 spectral image processing: Weed Technology, v. 6, p. 1015–1020.

Pickup, G., Chewings, V.H., and Nelson, D.J., eds., 1993, Estimating changes in vegetation cover over time in arid rangelands using Landsat MSS data: Remote Sensing of Environment, v. 43, p. 243–263.

Qi, Y., Chenbouni, A., Huete, A.R., Kerr, Y.H., and Sorooshian, S., 1994, A modified soil adjusted vegetation index: Remote Sensing of Environment, v. 48, p. 119–126.

Rejmanek, M., and Pitcairn, M.J., 2002, When is eradication of exotic pest plants a realistic goal?, *in* Veitch, C.R., and Clout, M.N, eds., Turning the tide—The eradication of invasive species: Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, Species Survival Commission, Invasive Species Specialist Group, International Union for Conservation of Nature, Gland, p. 249–253.

Rejmanek, M., and Richardson, D.M., 1996, What attributes make some plant species more invasive? : Ecology, v. 77, p. 1655–1660.

Sabins, F.F., 1987, Remote sensing, principles and interpretation: New York, Freeman and Co., 449 p.

Sakai, A.K., Allendorf, F.W., Holt, J.S., Lodge, D.M., Molofsky, J., With, Baughman, S., Cabin, R.J., Cohen, J.W., Ellstrand, N.C., McCauley, D.E., O'Neil, P., Parker, I.M., Thompson, J.N., and Weller, S.G., 2001, The population biology of invasive species: Annual Review of Ecology and Systematics, v. 32, p. 305–332.

Schmidt, H., and Karnieli, A., 2000, Remote sensing of the seasonal variability of vegetation in a semi-arid environment: Journal of Arid Environments, v. 45, p. 43–59.

Schurr, F.M., Bond, W.J., Midgley, G.F., and Higgins, S.I., 2005, A mechanistic model for secondary seed dispersal by wind and its experimental validation: Journal of Ecology, v. 93, p. 1017–1028.

Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., and Samson, F.B., 2002, Predicting species occurrences—Issues of accuracy and scale: Washington, D.C., Island Press, 840 p.

Shafii, B., Price, W.J., Prather, T.S., Lass, L.W., and Thill, D.C., 2003, Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics: Weed Science, v. 51, p. 748–751.

Shrader-Frechette, K.S., and McCoy, E.D., 1993, Method in ecology—Strategies for conservation: London, Cambridge University Press, 344 p.

Tackenberg, O., 2003, Modeling long-distance dispersal of plant diaspores by wind: Ecological Monographs, v. 73, p. 173–189.

Tanser, F.C., and Palmer, A.R., 1999, The application of a remotely-sensed diversity index to monitor degradation patterns in a semi-arid heterogeneous South African landscape: Journal of Arid Environments, v. 43, p. 477–484.

Thenkabail, P.S., Enclona, E.A., Ashton, M.S., Legg, C., and De Dieu, M.J., 2004: Remote Sensing of Environment, v. 90, p. 23–43.

Tsai, F., Chou, M.J., and Wang, H.H., 2005, Texture analysis of high resolution satellite imagery for mapping invasive plants: Geoscience and Remote Sensing Symposium 4, p. 3024–3027.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., and Steininger, M., 2003, Remote sensing for biodiversity science and conservation: Trends in Ecology and Evolution, v. 18, p. 306–314.

Wadsworth, R.A., Collingham, Y.C., Willis, S.G., Huntley, B., and Hulme, P.E., 2000, Simulating the spread and management of alien riparian weeds—Are they out of control?: Journal of Applied Ecology, v. 37, p. 28–38.

Wang, L., Sousa, W.P., Gong, P., and Biging, G.S., 2004, Comparison of IKONOS and QuickBird images for mapping mangrove species on the Caribbean coast of Panama: Remote Sensing of Environment, v. 91, p. 432–440.

Warren, P.L., and Hutchinson, C.F., 1984, Indicators of rangeland change and their potential for remote sensing: Journal of Arid Environments, v. 7, p. 107–126.

With, K.A., 2004, Assessing the risk of invasive spread in fragmented landscapes: Risk Analysis, v. 24, p. 803–815.

Zar, J.H., 1984, Biostatistical analysis: Englewood Cliffs, New Jersey, Prentice-Hall, 718 p.

Predicting Risk of Invasive Species Occurrence—Plot-Based Approaches

By Thomas C. Edwards, Jr., Richard Cutler, and Karen Beard

Chapter 7 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figures

Insets

Chapter 7. Predicting Risk of Invasive Species Occurrence—Plot-Based Approaches

By Thomas C. Edwards, Jr.,[1] Richard Cutler,[2] and Karen Beard[3]

Overview

This chapter discusses methods by which a class of statistical tools called predictive models can be developed using inventory and survey data to estimate existing and potential locations of invasive plant species as part of a long-term early-detection program. Such models would be spatial in nature and be directly linked to a park's GIS database. The models can have several outcomes, depending on management and research objectives.

These same predictive models commonly are used to examine relations between ecological parameters related to the presence (or absence) of an invasive plant species and whether such relations exist (that is the strength of the relations). These ecological models may or may not be spatially explicit, depending on the stated objectives. Both approaches are statistical in nature. Creation and use of these types of statistical models is best accomplished with input from ecologists with existing biological knowledge of the invasive plant species.

Undertaking the effort to build and evaluate predictive models can require significant investment of resources in the beginning, including time and personnel for field and corresponding office work, such as data entry, plant identification, and analysis. Depending on the nature of the model, outside consultation with competent statistical authorities may also be required.

This initial investment, if conducted properly, pays substantial, long-term dividends, especially for large parks and (or) areas that may experience significant numbers of introductions. Resources can actually be saved by focusing efforts on high-risk sites identified during the modeling process, particularly protected, valued, and (or) vulnerable sites where invasive species may inflict the most damage if they go unchecked.

All invasive-plant management efforts start with a model or series of models representing the invading species, the target resources and environs, and the effects associated with invasion. In most cases, these models remain in the minds of those with local knowledge of the resources and who have the responsibility to manage them. Sustainable, long-term management of the resources, however, requires managers to capture these models and articulate them clearly so that future management decisions are well informed. Whether management decisions are predicated upon mental models or more complex predictive models depends on the scope and intent formulated in management objectives (see Chapter 4).

There is no easy way to articulate all possible paths and variables by which the decision to proceed with a predictive modeling approach to early detection occurs. However, consideration of the questions below may help determine if the effort necessary to build a predictive model is a worthwhile step toward combating the invasive species problems in a park or other conservation area.

- Is the land area that you manage large and difficult to access?
- Are nonnative invasive plants a major stressor of the park ecosystem?
- Does your park have an existing GIS?
- Do you have inventory or survey data for invasive plants that include spatial data (for example, explicit point locations, or polygons) where invasive plants have been documented?
- Do these data contain both presence (that is where the species was found) and absence (that is where the species was not found) observations?
- Were these data collected in a probabilistic or non-probabilistic framework?
- Do you also have data on a variety of potential predictors for invasive plants, and are these data already organized in a GIS?
- Do you have sufficient resources to allocate field work for up to two field seasons, with the first devoted to collection of data for model building and the second to model evaluation?

[1] U.S. Geological Survey, Utah Cooperative Fish and Wildlife Research Unit, Utah State University, Logan, Utah.

[2] Mathematics and Statistics Department, Utah State University, Logan, Utah.

[3] Department of Wildland Resources, Utah State University, Logan, Utah.

- Do you have access to statistical consulting, if needed?
- Do you face time constraints for management decisions that preclude the more methodical, time-intensive, and structured predictive modeling approach?

To answer these questions, the reader may wish to delve deeper into this chapter and other chapters in this document. Figure 7.1 and the document's Decision Tree (see the front matter of this document) offer initial pathways through which the modeling decisions may be focused; however, the eventual decision to incorporate predictive modeling into a park's invasive plant early-detection program will be based on park needs and available resources.

What You Will Learn

This chapter will provide you with an overview of:

- A decision tree representing alternate analytical pathways to model building;
- How predictive models are used in species distribution modeling, including discussion of:
    - The ecological basis for such models and
    - The statistical basis of these models;
- The basic data structure required to build a predictive model, including distinctions between presence-only data and presence/absence data;
- An understanding of the data required to build a model and how these data can be collected;
- A description of a generalized analytical pathway for model building; and
- What is meant by a classification model and descriptions of major types of classification models, including:
    - Linear discriminant and quadratic discriminant analysis,
    - Logistic regression,
    - Additive logistic regression,
    - Classification trees, and
    - Random Forests.

Overview of Species-Based Predictive Modeling

Box 7.1. Species distribution models (SDMs) are statistical representations relating the likelihood of occurrence of a species to a set of predictor variables. They can, but do not have to be, spatial in design or application.

Central to attempts to generate predictive models of invasive plants is the concept of species distribution modeling (SDM). In general, SDMs are statistical representations relating the likelihood of occurrence of a species to a set of predictor variables. Although the statistical relations underlying SDMs are increasingly translated into spatial representations, SDMs do not necessarily have to be spatial in design or application. Being statistical representations, SDMs rely heavily on the collection of site-specific data from the spatial and environmental ranges of a target species (see Box 7.2). Once the data are collected, many different types of analytical methods (for example, generalized linear models (GLM), McCullagh and Nelder,1989; generalized additive models (GAM), Hastie and Tibshirani, 1990; and Yee and Mitchell, 1991; classification trees, Breiman and others, 1984; and De'ath and Fabricius, 2000) can be applied to the data to determine if statistical relations exist between the species and the various data sets collected. Data that have a statistical relation with the species may be used as predictors in the modeling process, indicating the possible occurrence of the target species within the environmental range examined. These models are then frequently linked to an existing GIS database and spatially explicit distributions (maps) generated for use in management.

A common form of analysis applied to data of this type is discriminant (hereinafter, classification) modeling, which produces models that are used to distinguish between nominal response values (for example, species presence or absence) using a set of environmental predictors. Applications of classification models to species distribution modeling abound in the published literature and range from predicting the distribution or characteristics of plant species (Austin and others, 1983, 1990; Frescino and other, 2001; Zimmermann and Kienast, 1999) to habitat relations of terrestrial animal species (McNoleg, 1996; Jaberg and Guisan, 2001; Lawler and Edwards, 2002; Welch and MacMahon, 2005). Excellent overviews of the state of SDMs in ecology and conservation can be found in Scott and others (2002) and Guisan and Thuiller (2005).

Although the emphasis of this document is on invasive species, it is important to understand that the "invasive" label attached to the modeled species is largely irrelevant to the SDM processes this chapter describes. The analytical processes described here all have general relevance to a broad set of ecological questions, which can range from spatially explicit predictive SDMs for conservation purposes, to tests of ecological theory, or to attempts to understand species-ecological relations. Many of these issues are also relevant to the management of both invasive and rare (threatened and endangered) species. The SDM process is merely a commonly used approach for addressing these types of management and conservation issues, although it is not the only process for addressing them.

Box 7.2. Clear articulation of the question being posed is the first, and most important, step of any invasive plant modeling effort. Some other important considerations include:

- Understanding the ecological, conservation, and management, context behind a proposed invasive model,
- Knowing if the intended use is for ecological explanation, prediction, or some combination of both,
- Determining the characteristics of the data structure, including how the data are to be collected, and their properties,
- Identifying potential analytical pathways, such as likely statistical tools for building the model and whether or not validation of the model is necessary,
- Knowing what the intended use of the final model is to be—local, regional, other, for example a model could be constructed to target a management area of a park, the entire park, an NPS network (if ecologically reasonable), a watershed (with neighbors), or an entire region. It is critical to understand the extent of desired inference.

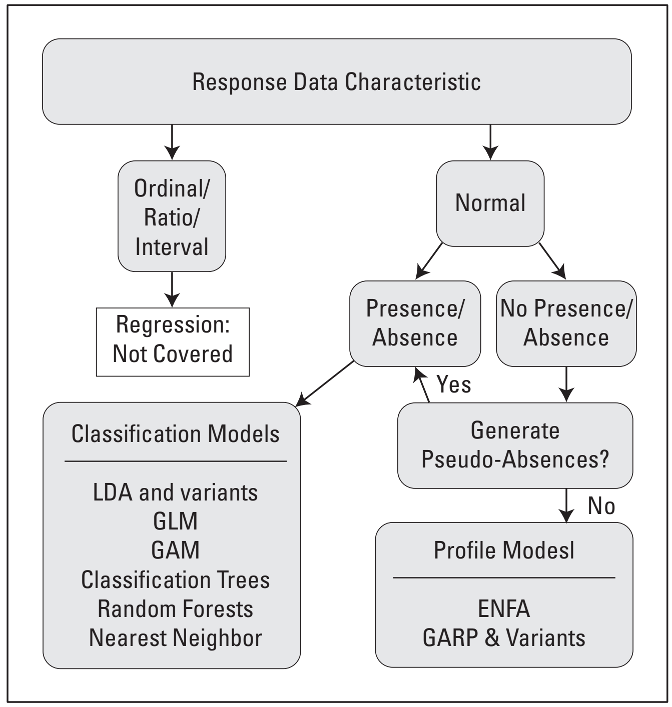

Figure 7.1. Various analytical pathways for species distribution models and associated decision points. Options are initially filtered by data type (response data characteristics). Data types and model types are discussed in this chapter (LDA, linear discriminant analysis; GLM; Generalized Linear Model; GAM Generalized Additive Model; ENFA, Ecological Niche Factor Analysis; GARP, Genetic Algorithm for Rule-set Production).

There exists a diverse array of analytical pathways for SDMs (fig.7.1). How these paths, and the statistical tools each contains, are organized for presentation in this chapter is, in part, a function of the intended audience. We have opted for an organizational approach based on our perceptions of managers' and conservationists' needs within—and outside of—the NPS, as well as our understanding of the NPS data organization. This presentation structure is designed to help formulate the selection of an appropriate analytical pathway given a park's objectives (see Chapter 4), existing information needs, and a prioritization scheme (Chapter 5). This chapter presents a set of stand-alone descriptions organized around some simple decision rules that logically lead to commonly used statistical tools in SDMs.

This chapter cannot be considered a complete discourse of each approach to the SDM process; many of these approaches, and the tools they employ, are the subject of exhaustive textbooks and rich publication histories. Thus, complete descriptions leading to fuller understanding and implementation of the method may require additional reading outside of this chapter (see Recommended Reading at the end of this chapter and cited literature). It is our expectation that once an approach is identified, more in-depth review of an approach by NPS biologists will be necessary. Some of the analytical tools described here may also require consultation with competent statistical authorities.

Ecological Setting for SDMs

For the purpose of discussion, SDMs can be organized into two basic classes. The first are those that seek explanation, typically through associative analyses, as to the ecological factors that explain why species are located within subsets of environmental gradients. Historically much of this work has been couched in terms of niche theory, with distinction made between the fundamental and realized niches (for example, Austin and others, 1990). With the advent of increased computing facilities available to ecologists, and the rise of GIS, a second class of SDMs that seek to predict where species are located has risen to the forefront (for example, Guisan and Zimmermann, 2000). This approach transforms the statistical relation underlying an SDM into a spatially explicit representation within a GIS, using predictors that must be spatially explicit in origin as well (for example, topographic variables). There are tradeoffs between the two approaches to SDMs—nonspatial explanation or spatially explicit predictions—that must be considered (see section “Characteristics of SDM Training Data”).

The typical process for creating a spatially explicit predictive model is to:

- Collect training data on a response of interest,
- Relate each response datum to digital predictor variables within a GIS,
- Build a statistical relation with these predictor variables, and
- Use the resultant model for extrapolation across the extent of the digital data (for example Lawler and Edwards, 2002; Beard and others, 1999).

When linked with statistical models of an invasive plant’s presence, for example, these digital coverages allow for extrapolation to unsampled space. Such coarse-resolution, spatially explicit models are now being used to predict plant invasions (Shafii and others, 2003; Gillham and others, 2004). However, it must be noted that any spatially explicit model of an invasive species does not necessarily represent the end-point of the analytical process, nor is it intended to be used in isolation from other biological information. Instead, such models could serve as the starting point for more in-depth approaches toward invasive species management.

For example, a first-pass spatially explicit model could be used to develop a more intensive, probabilistic field-sampling effort to locate invasive species. In this sense, the SDM first serves to determine what relations may exist between an invasive species and a set of spatially explicit predictors. These predictors are then viewed as strata in a sampling framework and are used to increase detection likelihoods and precision of estimates, such as abundances or frequencies on the landscape (see Edwards and others, 2004, 2005).

Another approach might be to view the spatial models as “coarse filters” (such as in Noss, 1987) that identify general locations where invasive species might gain a foothold. Once identified, a second, nonspatial field-sampling effort could be applied to determine if the microsite characteristics conducive to establishment and persistence occur in the area. Here, both spatial and nonspatial models are linked in a two-stage modeling process, with the first stage being spatial while the second is not.

The nature of most digital, environmental coverages precludes tight coupling of those layers with the underlying ecological reasons or processes determining presence of a species. Furthermore, existing digital coverages can often be of resolutions too coarse grained to specify precise locations (see Chapters 6 and 12 for more information on digital imagery applications). Consequently these first-stage models may identify coarse-grained spatial locations having high likelihood of invasive plant occurrences, but they will not be able to identify specific locations of microsites suitable for an invasive.

For example, few avian ecologists would deny the importance of fungal conks on aspen (*Populus tremuloides*) as an indicator of likely nesting trees for cavity-nesting species (see Daily, 1993). Presence of a fungal conk is indicative of heart rot, making the wood softer and more likely to be excavated by a primary cavity-nesting species. However, our ability to generate spatially explicit maps of fungal conks on individual trees is unlikely now or in the near future. Similar arguments can be made for some of the microsite characteristics that favor establishment by an invasive species. For example, yellow starthistle (*Centaurea solstitialis*), a recognized invasive threat in northern California, prefers silty loam and loamy soils at least 1.5 m deep and having few coarse fragments (Sheley and Petroff, 1999). However, parks are not likely to have digital coverages containing such detail on soils, effectively precluding any spatial extrapolation using these ecological predictors for yellow starthistle.

For these reasons, it may be beneficial to augment spatially explicit SDMs with nonspatial models that further refine the attributes leading to likely establishment by an invasive. These models will be more tightly linked to underlying ecological factors indicative of invasive establishment. The ecological literature is replete with examples of such models, especially for plant species (Austin and others, 1990; Brown, 1994; Leathwick, 1995; Austin and Meyers, 1996), and the analytical processes used in those studies can be easily adapted for invasive plants.

Assuming the decision has been made to engage in a predictive modeling exercise, the extent to which models for any single invasive will emphasize that spatial or nonspatial characteristics will be determined largely by the nature of the proposed predictor variables. Those having spatial characteristics and that are available in digital formats for use in a GIS can be used for the construction of spatial models. Other variables not available in spatially explicit representations, whether derived from literature or extant or primary data, can be used for the second-stage, nonspatial models.

Statistical Setting for SDMs

Once the decision to develop an SDM has been made, it is necessary to consider several statistical attributes before beginning any modeling exercise. These include understanding:

- The basic data structure necessary for SDMs,
- The characteristics of the data used to construct the SDM, often referred to as “training” data,
- How those data were or are to be collected, and
- What the proposed analytical pathway might be, given the data, and how the data were collected.

Basic SDM Data Structure

Virtually all SDMs involve a measured response (for example, species presence) that is related to a set of predictor variables, usually environmental in nature, although increased attention is being paid to sociological variables as well. The response can have several forms. One might be simple “presence.” If distance to road is assumed an important predictor of presence, that variable is measured in concert with the response. Thus, a simple data structure might consist of known present locations of the targeted species, plus a distance to road measurement for each observed presence. Note here that only “presence” has been determined, and the associated distances to road have been measured. No “absences” have been recorded. This distinction–do you have both presence and absence, or just presence only?—is an important first step in determining the eventual analytical pathway for an SDM. The alternative data structure would also contain locations where the species is not found (absences) and the distance to road for each location.

In situations where only species presence data are available, several authors have suggested generating so-called pseudoabsences to augment observed presences as an alternative to profile-type models (Ferrier and Watson, 1997; Stockwell and Peters, 1999; Zaniewski and others, 2002). The most common technique for generating pseudoabsences involves randomly selecting absence points from the entire study region, excluding where the species is found (Parra-Olea and others, 2005; Stockman and others, 2006). These techniques do not typically incorporate ecological knowledge of the species-habitat relation; for the most part, pseudoabsences are selected from broadly defined areas, such as the study region or simple range maps.

It is also important to know the underlying nature of both the response and predictor variables. The most commonly used description of data characteristics consists of four basic types:

- nominal
- ordinal
- ratio
- interval

The distinctions among these four data types become important farther down the analytical pathway, particularly in the selection of the appropriate statistical tool for modeling. Nominal data are those which are classifications or labels only. There are no numerical comparisons among different categories. Examples might include presence (yes, no) or habitat type (for example, forest, grassland, wetland) data. Analytically, we may wish to compare the numbers of observations (tallies) in each category, but the categories themselves are merely descriptive in nature.

Box 7.3. Understanding the basic **characteristics of your data structure** is a vital first step to consider before initiating an invasive species modeling effort. It is important to know:

- If sample locations where a species is absent will be available as well as locations where it is present. Much better models can be built if both presence and absence data are available.
- Whether your response and predictor data are nominal, ordinal, or interval/ratio data types.
- If you are collecting a response on more than one species at any sample plot.

A second type of data is ordinal. Ordinal scales represent classification or categorizations that can be compared by ordering. Examples might include habitat quality (excellent, good, poor) or invasion likelihoods (high, medium, low). Like nominal data, differences between the ordered categories cannot be assumed numerically consistent. Consequently both nominal and ordinal data are often referred to as nonmetric or categorical data types.

The last two categories—interval and ratio—differ only in that interval data have no true zero. For both of these data types, differences between two measurements have meaning, which is not true for nominal and ordinal data types. Interval and ratio data are sometimes called metric or continuous data. Examples of interval and ratio data include temperature (degrees Celsius) and distance (for example, meters) or area (for example, square kilometers), respectively. The two data types can be considered the same for the vast majority of modeling exercises. One major distinction is with percentage and frequency data, which, because they have true zeros, are considered ratio data. Both of these data types are frequently used as predictors in SDMs and can, depending on the statistical model selected, require special handling during analysis.

Another characteristic important in most SDMs is the nature of the observation. Typically, only a single response measure (for example, presence/absence, forest/nonforest) is made at any sample plot. This differs considerably from circumstances where multiple response measurements or observations are made at the same sample plot. While observations of multiple species at the same sample plot can be used in SDMs, they do not lend themselves well to most of the statistical tools discussed in this chapter, especially if the goal is to translate any resultant statistical model into a spatial depiction. Multiple responses are best analyzed by a class of multivariate techniques generally referred to as ordination techniques, and readers are referred to texts by Pielou (1984), Kent and Coker (1992), McGarigal and others (2000), and Manly (2004) as starting points.

Box 7.4. "Training data" is the name typically applied to the data from which a species distribution model (SDM) is constructed. These data consist of a measured response–say the presence or absence of the invasive plant species–and a set of predictor variables associated with response. The nature of these predictor data, referred to as direct, indirect, or "model" data, can affect the conservation and management utility of the predictive model.

- **Direct predictors** are variables having a known or hypothesized relation with the presence of a species. Examples for plants might include soil moisture, temperature, and solar radiation.
- **Indirect predictors** are surrogates that are clearly related to direct predictors, but for which an explicit relation is not known. Thus, aspect in a mountainous region may act as a surrogate for soil moisture, temperature, and solar radiation. Note, however, that all three direct predictors are collapsed into a single surrogate, aspect, potentially confounding model interpretation.
- **"Model" data** typically combine one or more direct and indirect variables with a mathematical relation. Thus, creation of the predictor variable "potential evapotransipration" might require solar radiation as modeled by aspect, temperature, and a recognized mathematical relation such as that by Jensen-Haise (1963) or Turc (1963).

To the extent possible, SDMs should rely on direct followed by indirect predictors, with "model" data as the least desirable.

Characteristics of SDM Training Data

Belovksy and others (2004) recently criticized the ecological modeling community for using predictor surrogates too far removed from the underlying ecology of the organism(s) being modeled. One obvious example is the overreliance on topographical variables for spatially explicit, predictive models, even though many species, including some invasives (Shafii and others 2003), are clearly associated with topography. While the criticism of Belovsky and others, (2004) is valid at simple face value, it does not negate the use of SDMs as conservation or management tools. It should, however, foster a more methodical, comprehensive approach toward any invasive SDM, one that better links the selected predictors to the underlying species ecology.

To better link selected predictors to species ecology, it is convenient to think of predictors as being one of three possible types (Box 7.4): direct, indirect, or "models of models."

Direct predictors (as in Austin, 1980) are those with known or hypothesized relations with the presence of a species. Soil moisture and temperature are directly linked to the presence of most plant species, for example. While it is easy a priori to develop reasons for the relation between species distributions and direct predictors, the relation is less certain when dealing with indirect predictors (Austin and Smith, 1989). For example, it is logical to argue that aspect in a mountainous region is related to plant species presence, but the linkage is indirect in that aspect most likely serves as a surrogate for soil moisture, with southerly facing aspects being drier than those facing north. Thus, a predictor like aspect is considered indirect, for it is removed from direct measurement of the actual ecological attribute (that is soil moisture) affecting presence, but is clearly related.

Further removed from the ecological underpinnings of why a species is (or is not) present are "models of models," a catchall phrase used to describe all predictors not labeled direct or indirect. Consider potential evapotranspiration, an important ecological variable associated with the presence or absence of plant species. If the resources exist to record and use evapotranspiration measurements in an SDM, then it could be argued the SDM would have better ecological underpinnings. However, the extrapolation of evapotranspiration as a predictor for spatially explicit SDMs relies almost exclusively on models involving complex mathematical relations among many different direct and indirect variables. In addition, there often exist different formulations for actually estimating the same variable (for example, evapotranspiration, as estimated by Jensen and Haise, 1963 compared to Turc (1963). Thus, any resultant predictor is a "model of a model" and may be far removed from the true ecological factor determining species presence.

The use of different predictor types (that is direct, indirect, or "model of a model") in the modeling process may result in distinctly different SDMs. The nature of these distinctions is largely unknown, however. Both early studies (Swan, 1970; Noy-Meir and Austin, 1970; and Austin and Noy-Meir, 1971) and more recent work (Bio, 2000; Moisen and Frescino, 2002) using artificial data for simulating plant SDMs used mixtures of the three types of predictors. Most of these studies were focused on the use of artificial data to evaluate the performance of different statistical tools and were not concerned with the effect of the predictor characteristics on ecological explanation or prediction. Nevertheless, reliance on indirect or "model of a model" predictors could lead to erroneous ecological conclusions if the goal of the invasive model is ecological explanation or understanding of the factor(s) associated with presence or absence. However, use of surrogates is of less concern if the analytical goal is a spatially explicit predictive model, as long as it is understood that ecological interpretation may be limited, and that statistical relations are largely correlative.

Collection of Training Data

Training data are data collected and used to build (that is "train") the predictive model. Training data for SDMs are typically obtained by survey sampling. Numerous texts exist on the types of sampling designs that could be applied to the collection of training data for an SDM (see Cochran, 1977, and Schreuder and others, 1993), and a complete review is beyond the scope of this chapter. However, an important distinction to understand is the difference between probabilistic and nonprobabilistic forms of sampling (see also Chapter 8).

Ideally, survey sampling involves the random selection and measurement of samples from a defined target population referred to as the sampling frame (Chapter 8). Sampling frames can have many different characteristics. Often sampling frames are a defined spatial extent considered a finite population, such as a national park, that is divided into N smaller units of some set size (for example 90-meter grid cells, watershed hydrologic units), from which a subset n_i is randomly selected and surveyed for the species of interest (see Edwards and others, 2004). These area samples (as in Nusser and others, 1998) may or may not correspond with the actual sample unit, which may be something as simple as the presence or absence of a specific species within the area being sampled. Estimates derived from these types of sampling designs are considered probabilistic and carry with them the power of inferential statistics (that is, inferences can be made to the other unsampled units within the sample frame; see Chapter 8).

Many ecological studies, however, deal with attempts to survey and build SDMs for ecological events best described as rare or uncommon on the landscape (Engler and others, 2004). Rare ecological events may not be truly amenable to randomization procedures as such, especially when the goal is to generate sufficient observations for an SDM (see Edwards and others, 2005). For example, the random selection of small areas within some defined spatial extent like a national park, which are then surveyed to locate bird nests, might not be a fruitful exercise if too few bird nests are found. While such a design would allow for inferential statistics to be determined for, say percentage of sites occupied (see Edwards and others, 2004), it may not provide sufficient tallies of nest presence, thereby precluding attempts to build an SDM. This type of circumstance typically requires the collection of training data involving the active searching for the "targeted" species of interest.

In the case of "targeted" species, ecologists actively search for the event of interest by using nonprobability sampling procedures (see Cochran, 1977). One of the more common of these nonprobability sampling efforts, termed purposive sampling, occurs when ecologists actively seek the event of interest, such as an active breeding nest of a bird or a specific plant species. Even when the species to be modeled is not considered "rare," such as is the case for the vast majority of invasive plant species, it nonetheless represents a "targeted" species for which observations are collected in a nonrandom fashion.

Training data used to build SDMs are rarely obtained from probability-based sample designs, but are more frequently obtained through nonprobabilistic (for example purposive) sampling. Reliance on nonprobabilistic training data clearly injects biases into any modeling effort, and these biases will consequently cascade through the entire analytical process and affect, for example, any derived spatially explicit map products or attempts to understand ecological relations (see Schreuder and others, 2001; Edwards and others, 2006). In some instances, probability-based samples will be available (through Fire Effects monitoring plots, USDA Forest Service vegetation monitoring plots, and so forth) for some or all of the vegetation types of interest in a park, providing existing training data which can be used to build SDMs.

Realistically this issue is difficult to overcome, although biases associated with the training data can be easily assessed through a variety of validation procedures (Chapter 9). Ideally, any models and their resultant spatial extrapolations should be validated (Chapter 9). Unfortunately, few classification models in ecology–be they spatial or not–are validated, with approximately 83 percent (n=334) of published papers not even undergoing any form of internal validation (T.C Edwards, unpub. data, 2008). Only about 5 percent are validated with independent field data.

General Analytical Pathway

If the objective has been clearly articulated (see Box 7.2 and Chapter 4) and the data structure has been determined, the selection of specific statistical tools for modeling is largely determined. Two of the most commonly applied types of statistical analysis in ecology and related areas are regression and discriminant analysis, or classification. In regression analysis, one or more quantitative response variables are explained or predicted as a function of one or more quantitative ecological explanatory variables. For example, one might model regional plant species richness as a function of soil type, soil moisture, measures of temperature, and topographic variables. Yet another application is explaining and predicting abundances of animal and bird species by using quantitative information on available habitat and other ecological variables.

Classification involves identifying "rules" or models for classifying observations into two or more known classes using quantitative variables. One of the simplest ecological applications of discriminant analysis is in the analysis of presence/absence data for plants or animals. Here, a set of predictors is hypothesized to have a relation between the likely presence (or absence) of the targeted species (Franklin, 2002). A more complex example might involve classifying landscapes into vegetation classes on the basis of reflectance values from satellite imagery and geographical variables such as elevation, aspect, and slope (Franklin, 1995; Moisen and others 1996; see also Chapter 6). Again, existing data sets may prove invaluable for these purposes, saving time and money.

At first glance, regression and classification seem to be quite different conceptually. However, one way to view classification is that it is like regression—predicting or explaining a response variable using one or more explanatory variables—but that the response variable(s) is (are) categorical in nature. A consequence of the regression view of classification is that many of the statistical models used for regression may be used, or modified for use, in classification.

What follows is a review of statistical model families for classification, including discussion on how the model families relate to one another. Included are several new and flexible tools that have the potential to change how ecological data are commonly analyzed and interpreted. Most of the techniques discussed have been developed in the last 20 years and are at the cutting edge of statistical research and applications.

Classification Methods for Presence-Absence Data

Here we present a number of statistical classifiers that may be used with ecological data. Some of these classifiers are currently in wide use in ecology, while others have been used in both ecology and remote sensing. One of the classifiers is relatively new (Random Forests) but has already been shown to be superior to other methods for some applications.

As discussed previously, a classifier is a rule for assigning observations to one of two or more known classes on the basis of measured numerical and categorical variables. Usually the choice of a classifier is driven by the desire for classification accuracy. Observations that are classified into the correct class are said to be correctly classified, while observations that are classified into the wrong class are said to be misclassified. Obviously, we would like to choose a classifier that has a high correct classification rate or, equivalently, a low misclassification rate. In generally, more complex classifiers have higher classification accuracies, but their results can be harder to interpret. Consequently, the selection of a particular classifier is based on finding balance between the simplest (that is easiest to interpret) classifier that has high classification accuracy. The particular choice of a classifier will depend on objectives, the nature of the data collected or in hand, and a basic understanding of the underlying statistics of each.

Linear Discriminant Analysis and Quadratic Discriminant Analysis

Fisher (1936, 1938) first proposed the use of linear combinations of the measure variables as classification rules. He also coined the phrases linear discriminant analysis (LDA) and linear discriminant function to describe his methodology. The basic idea of the method is illustrated in figure 7.2 for a simple case of two classes (represented by the ellipses full of dots) and two variables, x and y. The vertical line at the x-value "c1"separates the two groups: observations with x-value less than c1 are classified as being in the first group and observations with x-values greater than c1 are classified as being in the second group. Some of the observations in the intersection of the two ellipses will be misclassified by this simple rule.

For higher dimensional data (that is more complex data sets), the classifiers are more complicated linear combinations of the measured variables. With three or more classes, LDA involves multiple intersecting linear combinations of the measured variables. Prior information about the probability of membership of observations in the different classes (for example, prior probability) may also be used in the selection of the linear discriminant functions. Also, different numbers of observations in different classes may also be suitably handled using prior probabilities proportional to the numbers of observations in the different classes.

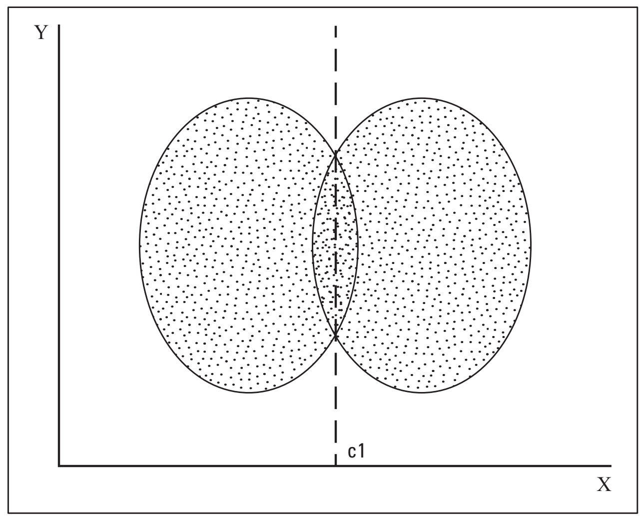

Figure 7.2. Linear discriminant function.

If the measured variables are approximately multivariate normal in distribution, and the clouds of data values for the different classes have roughly the same size and shape, then LDA is a good classifier. Remotely sensed spectral data for vegetation are an example of data that fit this description (see Cutler and others, 2003).

In practice, the rather important assumption of equality of size and shape of the data clouds for observations in different classes is rarely met for most ecological data. Quadratic discriminant analysis (QDA) is a generalization of LDA in which the assumption of equality of shape and size of the data clouds for the different groups is relaxed. In figure 7.3, the data clouds for the two classes are approximately equal in size but have different orientations, and hence "shapes." In this case, if the measured values are multivariate normal in distribution, then the classifier that minimizes the expected misclassification rate is a quadratic curve, represented as the dashed curve. With three or more classes the QDA classifier involves several intersecting quadratic functions.

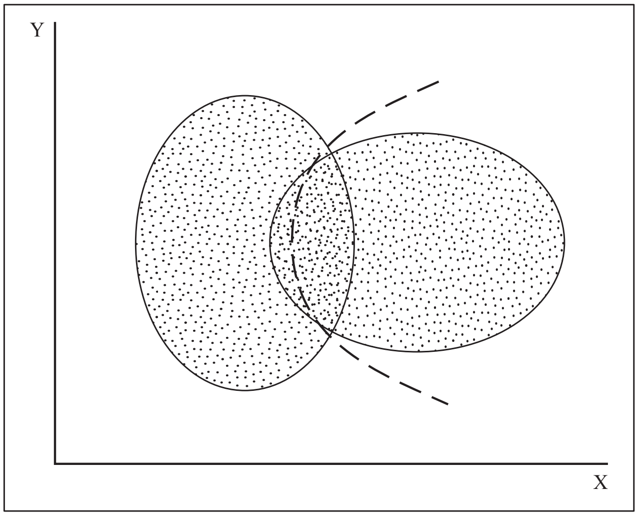

Figure 7.3. Quadratic discriminant function.

For measured variables that are multivariate normal or approximately multivariate normal in distribution, LDA and QDA are accurate classifiers that are simple to interpret. LDA and QDA are both robust to modest departures from multi-variate normality, such as one sees in remotely sensed spectral data, but not to extreme departures. In particular, both LDA and QDA are known to perform poorly if categorical variables are used.

Logistic Regression

Logistic regression (Hosmer and Lemeshow, 2000) is a regression like procedure for relating a binary response (again, presence or absence) to a linear combination of numerical and categorical variables. Specifically, let p_i denote the probability of presence of the species of interest at site i. The logistic regression model is:

$$p_i = \exp\{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}\} / (1 + \exp\{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}\})$$

where the x_{ij}'s are the values of the different measured values (indexed by j) at the different sites (indexed by i), and the β_j's are regression coefficients that are estimated using the data. Observations at different sites are assumed to be statistically independent with binomial distributions with success probabilities μ_i. The function $f(u) = \exp\{u\}/(1 + \exp\{u\})$ is called the logistic function, and hence the name of the procedure. The logistic function is needed here to ensure that the estimated (or predicted) probabilities of presence lie between 0 and 1.

Logistic regression has been widely used in ecological modeling, in part because the regression structure of the model lends itself to easy interpretation in most applications. The set of β_j's, for example, are interpreted in the same fashion as in linear regression, with positive and negative signs indicating the direction of the relation. Those β_j's that differ from 0 can be assumed statistically related to the response, depending on a selected level of α. The logistic regression model also makes absolutely no assumptions about the distribution of the measured variables. Consequently it is robust to departures from the assumption of multivariate normality, and categorical variables may easily be used in the model by coding up a number of zero-one variables, one for each category of each categorical variable. If the measured variables are approximately multivariate normal in distribution, LDA and QDA will outperform logistic regression; in all other cases, logistic regression is superior.

The logistic regression model is a special case of a class of regression like models known as Generalized Linear Models (GLMs), which also include ordinary linear regression and Poisson regression. Consequently, many of the best references for logistic regression are books on GLMs (for example, McCullagh and Nelder, 1989). Logistic regression models are subject to the same kinds of problems that bedevil ordinary linear regression, including influential points and multicollinearity of the measured variables. A complete logistic regression analysis entails addressing these issues. Another shortcoming of logistic regression is the restriction to linear functions of the measured variables. In ecological and other applications of classification methods, the assumption of linearity may be highly suspect.

Additive Logistic Regression

Additive Logistic Regression (Austin and others, 1994) is an extension of logistic regression that includes additive and possibly nonlinear combinations of numerical measured variables. Additive logistic regression is one of a class of models called Generalized Additive Models (GAMs), developed by Hastie and Tibshirani (1986). The formulation of additive logistic regression is similar to logistic regression. Letting p_i denote the probability of presence at a particular site, i, the additive logistic regression model is:

$$p_i = \exp\{s_0 + s_1(x_{i1}) + s_2(x_{i2}) + \ldots + s_k(x_{ik})\} / (1 + \exp\{s_0 + s_1(x_{i1}) + s_2(x_{i2}) + \ldots + s_k(x_{ik})\})$$

where, as with logistic regression, the x_{ij}'s are the values of the different measured values at the different sites. s_0 is a constant estimated from the data, and the other s_j's are smoothing functions that are also estimated nonparametrically from the data by using scatter plot smoothers. That is, the s_j's are not specified in advance of the analysis. So, for example, it may be that the probability of presence of a particular species initially increases (smoothly) with elevation, but then decreases abruptly. This kind of relation would result in an estimated s_j for elevation that might look like an inverted parabola (fig. 7.4).

The principal advantage of additive logistic regression is its ability to deal with highly nonlinear, and even nonmonotonic, relations between the response and the predictor variables. Additive logistic regression and other GAMs are sometimes described as data driven rather than model driven.

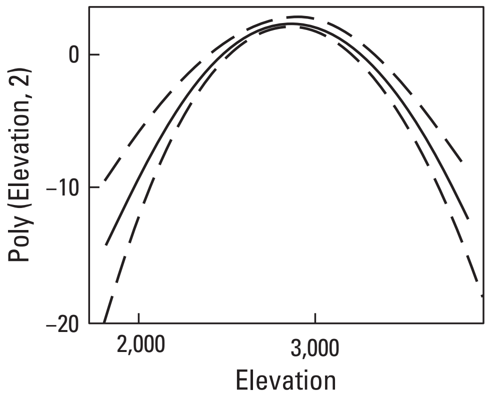

Figure 7.4. Estimated s_j for a species response to elevation.

This is because the data determine the nature of the dependency of the response variable on each of the explanatory variables, as opposed to assuming a parametric form for the relations (See and Mitchell 1991).

Some ecological applications of additive logistic regression include Norton and Mitchell (1993), who found additive logistic regression to be useful for modeling the occurrence of the Greater Glider (*Petauroides volans*) in southeast Australia using climatic variables, and Moisen and others (1996), who used additive logistic regression to model the distribution of forest vegetation based on forest inventory data and explanatory variables of elevation, slope, aspect, and satellite data incorporated in the GAP Analysis vegetation cover map. Leathwick (1995) used additive logistic regression to determine the relation of climatic variables, such as mean annual temperature and mean annual solar radiation, to the distribution of forest tree species in New Zealand, and concluded that additive logistic regression was more appropriate for fitting monotonic and plateautype distributions than other models.

One shortcoming of additive logistic regression, shared by logistic regression, LDA, and QDA, is the inability to model complex interactions among the measured variables. Additive logistic regression is also a "data hungry" procedure in the sense that it requires much more data than the simpler logistic regression model. In applications in which only small amounts of data are available, additive logistic regression may not be a viable option. With logistic regression it is possible to write down a model for the predicted or estimated probability of occurrence of a species at a site using the estimated regression coefficients, β_j. It is not possible to do the same with additive logistic regression because the s_j's are nonparametrically estimated and therefore do not have a (parametric) functional form.

In response to this concern, Hastie and Tibshirani (1986) suggest that one use of additive logistic regression (and GAMs in general) is as an exploratory tool for discovering parametric models for data relations. Several studies have followed this suggestion (see Yee and Mitchell, 1991; Brown,1994). In a comparison study of additive and ordinary logistic regression for predicting the distribution of *Eucalyptus cypellocarpa*, additive logistic regression overestimated the probability of occurrence beyond the range of observations but was, in general, found to be advantageous for nonlinear data structures due to the flexibility of the additive structure (Austin and Meyers, 1996).

Classification Trees

Classification trees were developed by Breiman and others (1984). They work by recursively partitioning the data into groups that are as homogeneous as possible with respect to the response variable, using the measured variables. For example, in an analysis to model presences of the lichen *Lobaria oregana* as a function of geographic and ecological variables such as stand age, precipitation, and elevation, classification trees first selected an elevation of 4,280 feet as a split point, with the species more likely to be found in stands at lower elevations (fig. 7.5). Then, among the stands at lower elevations, classification trees chose the variable stand age and a split point of 79 years. At lower elevations, the species was more likely to be found in the older stands than younger.

For the higher elevation stands, the classification tree also chose stand age as the next variable to split on, but with a split point of 120 years. The full tree, with every data point perfectly classified, overfits the data in the sense that many of the splits at lower levels of the tree are modeling noise, so methods for "pruning" the tree or determining the appropriate number of splits are typically applied, resulting in trees with between 6 and 12 total splits.

Classification trees are already widely used in ecology and are gaining popularity (see De'ath and Fabricius, 2000, for an overview of ecological applications). In many applications classification trees have superior classification accuracies to the other methods discussed so far. The ability of classification trees to characterize complex interactions among measured variables may be one reason for this popularity. Classification trees can accommodate both numerical and categorical measured variables and make no assumptions about the distributions of the measured variables. They are also easy to interpret, with binary decisions at each split, and hence are intuitively appealing for understanding structure in ecological situations. Classification trees may also be used with response data having more than two classes. For example, Cutler and others (2003) used classification trees to classify satellite spectral data into 11 vegetation classes.

Classification trees have two shortcomings. First, they require more data than simpler procedures, such as LDA, QDA, and logistic regression. Second, classification trees are quite unstable in the sense that small perturbations of the data may result in substantially different fitted classification trees (Edwards and others, 2006). This phenomenon is closely related to the problem of multicollinearity in regression like models, and it raises questions about the validity of interpreting the splits in a classification tree.

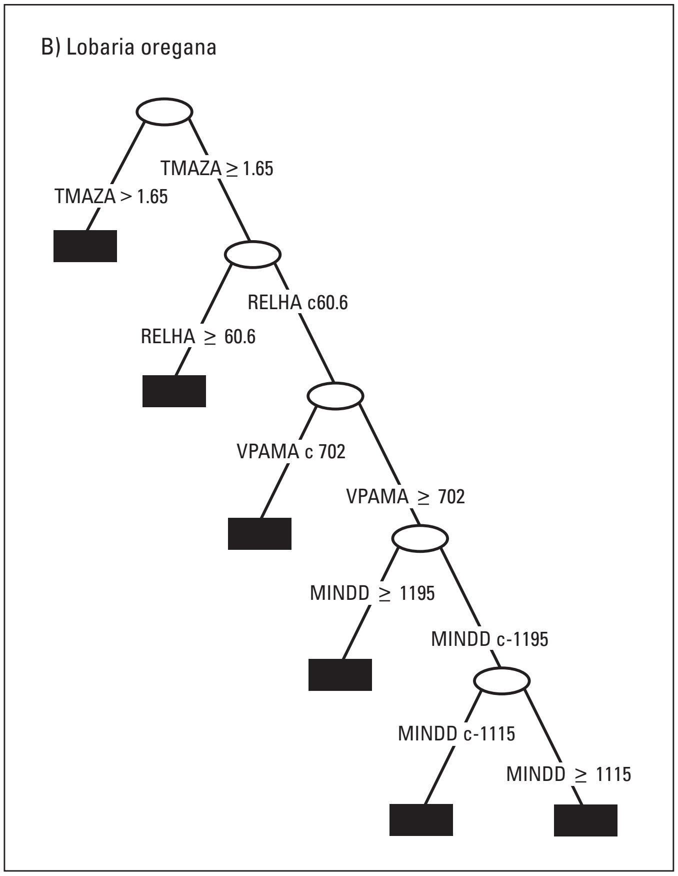

Figure 7.5. Example of a classification tree for predicting presence of *Lobaria oregana*.

Random Forests

As the name suggests, random forests (Breiman, 2001) are collections of classification trees. Breiman's original motivation for developing Random Forests was to address the issue of instability of classification trees, as noted previously. The idea is to randomly select subsets of the original data by sampling with replacement, fitting classification trees to each sampled data set, and then combining the results by averaging or "voting." The basic algorithm for Random Forests is:

- Many bootstrap samples (that is randomization procedure in which the dataset is resampled with replacement) of the original data are drawn.
- On each bootstrap sample, a classification tree is fit.
- At each node, in each classification tree, a randomly selected subset of the variables is made available for splitting.
- The trees are fully grown and no pruning takes place.
- For each tree, fitted on a single bootstrap sample, predictions are generated for all data values that were in the original data set, but which were not in the bootstrap sample (that is "out-of-bag" data values; Breiman, 1996).
- For a given data value, the predicted class is the class with the highest count among out-of-bag predictions for that data point. Ties are split randomly.

The main drawback with Random Forests is in interpretation. No simple formulation, such as for logistic regression or classification trees, is available. Measures of variable importance in the classification are available and are much better than commonly used regression like variable selection methods, and graphical representations are currently under development. Graphical representations of the marginal or partial relation between each individual predictor and the predicted probabilities can be constructed, but these kinds of plots necessarily cannot display complex interactions and multivariate structure that Random Forests is particularly well suited for exploiting.

Additionally, Random Forests is not a tool for traditional statistical inference. It is not suitable for ANOVA or hypothesis testing. It does not compute P-values, or regression coefficients, or confidence intervals. The variable importance measure in Random Forests may be used to subjectively identify ecologically important variables for interpretation, but it does not automatically choose subsets of variables in the way that variable subset selection methods do. Rather, Random Forests characterizes and exploits structure in high dimensional data for the purposes of classification and prediction.

Quantities produced by Random Forests may be used as inputs into traditional multivariate statistical methods, such as cluster analysis and multidimensional scaling. Unlike many traditional statistical analysis methods, Random Forests makes no distributional assumptions about the predictor or response variables, and can handle situations in which the number of predictor variables greatly exceeds the number of observations. With this range of capabilities, Random Forests offers some interesting and powerful alternatives to traditional parametric and semiparametric statistical methods for the analysis of ecological data. In terms of classification accuracy Random Forests is probably the best all-purpose classifier available at this time.

Software Availability

Linear and quadratic discriminant analysis, logistic and additive logistic regression, k-nearest neighbor classifiers and classification trees are widely available in major statistical analysis packages including SAS, SPSS, S-PLUS (Venables and Ripley, 1997), and R. Random Forests is available as a function in R and through FORTRAN code at the Random Forests Website (accessed November 28, 2010 at *http://www.stat.berkeley.edu/~breiman/RandomForests/*). Both Random Forests and Classification Trees are available as proprietary software from Salford Systems. There are differences in the implementations of classification trees and Random Forests in the different sources. The Salford Systems software is the most user-friendly with a GUI interface. The R package is freeware (accessed November 28, 2010 at *http://www.r-project.org*) and is most likely to have new statistical procedures as they are developed. SAS remains the most widely used major statistical software system in the United States and worldwide.

The list of classifiers presented here is not a comprehensive list of classifiers. In particular, we have not discussed support vector machines and artificial neural networks, which are rather complex methodologies. Random Forests is the best known method in a class of classifiers known as ensemble classifiers, which involve fitting many classifiers to the same data and combining the results. Steele (2000) and Steele and Patterson (2002) have used ensembles of k-nearest neighbor classifiers for classifying remotely sensed vegetation data.

Summary

Despite the importance and ubiquity of the invasive species problem in our national parks and natural areas, guidance has been lacking on how to conduct invasive species early-detection monitoring. Knowing where to find the proverbial needle in the haystack is critical to effective management of nonnative plant species early in the invasion process, requiring predictive tools—simple or complex—to assist in this process. In fact, the recent ESA report on policy and management of biological invasions (Lodge and others, 2006) recommended the use of modern statistical classification procedures for the risk assessment of potential habitat for invasive species (Recommendation 2 and discussion) and also suggests the use of environmental and meteorological data in such models.

The purpose of building these predictive models is to assist park personnel in prioritizing sites to search for invasive species. More specifically, managers can develop models for extrapolating to unsampled regions the likely locations of invasive plant species and assess whether these locations are suitable for the establishment of invasive plants. The former modeling effort relies on statistical models linked to a GIS, thereby providing spatially explicit depictions (strata) of where invasive plant species are likely to occur. The latter, unlike the former objective which is centered on developing spatially explicit strata of invasive likelihoods, is of a nonspatial nature and relies more on identifying the microsite characteristics necessary for the establishment of an invasive species. This chapter addresses these objectives by providing an overview of species distribution models, a general approach to discern which modeling procedures are best suited to one's data structure (fig. 7.1), and a number of statistical classification methods for carrying out modeling procedures.

We acknowledge that the different characteristics of ecosystems and different species in different parks require flexibility in designing appropriate models for stated management objectives. We also note that there are elements of commonality among most parks and conservation areas. Many invasive species use roads and trails as their main vectors of access to parks, and disturbed sites are most susceptible to invasion by nonnative species. Consequently, we have presented an overview of the methods to invasive species distribution modeling rather than being prescriptive.

Once models have been built, they also need to be validated (Chapter 9) to examine the adequacy of the models for predicting the likelihood of occurrence of invasive plants and to accurately predict the locations of invasive plants in selected sites. Sampling strategies (Chapter 8) also need to be formulated to assist in the validation procedure and for early-detection monitoring. Subsequently, models will need to be updated as new data become available and (or) priorities (that is, management objectives) change. The following chapters address these issues as separate topics (see Chapter 10 for a synthesis and discussion of implementation issues).

Recommended Reading

An excellent overview of the state of species distribution models in ecology and conservation can be found in Scott, J.M., Heglund, P.J., Samson, F., Haufler, J., Morrison, M., Raphael, M., and Wall, B. (2002).

For a more recent synopsis of the state of species distribution models in ecology and conservation, see also Guisan, A., and Thuiller, W. (2005).

For an example of nonspatial models linked to underlying ecological factors in an applied invasive plant management setting, see Austin, M.P., and Meyers, J.A. (1996). The application of pseudoabsences in species distribution models can be found in Zaniewski, A.E., Lehmann, A., and Overton, J. McC. (2002). Manly (2004) presents an excellent introduction and overview to multivariate statistical methods not covered in this chapter.

References Cited

Austin, M.P., 1980, Searching for a model for use in vegetation analysis: Vegetatio, v. 42, p. 11–21.

Austin, M.P., Cunningham, R.B., and Good, R.B., 1983, Attitudinal distribution in relation to other environmental factors of several Eucalypt species in southern New South Wales: Australian Journal of Ecology, v. 8, p. 169–180.

Austin, M.P., and Meyers, J.A., 1996, Current approaches to modeling the environmental niche of eucalypts—Implications for management of forest biodiversity: Forest Ecology and Management, v. 85, p. 95–106.

Austin, M.P., Nicholls, A O., Doherty, M.D., and Meyers, J.A., 1994, Determining species response functions to an environmental gradient by means of a beta function: Journal of Vegetation Science, v. 5, p. 215–228.

Austin, M.P., Nicholls, A.O., and Margules, C.R., 1990, Measurement of the realized qualitative niche—Environmental niche of five Eucalyptus species: Ecological Monographs, v. 60, p. 161–177.

Austin, M.P., and Noy-Meir, I., 1971, The problem of nonlinearity in ordination—Experiments with two-gradient models: Journal of Ecology, v. 59, p. 763–773.

Austin, M.P., and Smith, T.M., 1989, A new model for the continuum concept: Vegetatio, v. 83, p. 35–47.

Beard, K.H., Hengartner, N., and Skelly, D.K., 1999, Effectiveness of predicting breeding bird distributions using probabilistic models: Conservation Biology, v. 13, p. 1108–1116.

Belovksy, G.E., Botkin, D.B., Crowl, T.A., Cummins, K.W., Franklin, J.F., Hunter, M.L. Jr., Joern, A., Lindenmayer, D.B., MacMahon, J.A., Margules, C.R., and Scott, J.M., 2004, Ten suggestions to strengthen the science of ecology: BioScience, v. 54, p. 345–351.

Bio, A.M.F., 2000, Does vegetation suit our models? Data and model assumptions and the assessment of species distribution in space: Utrecht, The Netherlands, Knag/Netherlands Geographical Studies, 195 p.

Breiman, L., 1996, Bagging predictors: Machine Learning, v. 24, p. 123–140.

Breiman, L., 2001, Random forests: Machine Learning, v. 45, p. 5–32.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J., 1984, Classification and regression trees: Monterey, California, Wadsworth and Brooks/Cole, 368 p.

Brown, D.G., 1994, Predicting vegetation types at treeline using topography and biophysical disturbance variables: Journal of Vegetation Science, v. 5, p. 641–656.

Cochran, W.G., 1977, Sampling techniques: New York, John Wiley & Sons, Inc., 428 p.

Cutler, D.R., Brown, L., Powell, J., Bentz, B., and Cutler, A., 2003, Identifying ‘Redtops’—Classification of satellite imagery for tracking mountain pine beetle progression through a pine forest: Salt Lake City, Utah, Proceedings of the 35th Symposium of the Interface of Computing Science and Statistics, March, 2003.

Daily, G.C., 1993, Heartwood decay and the vertical distribution of red-naped sapsucker nest cavities: Wilson Bulletin, v. 88, p. 575–581.

De’ath, G., and Fabricius, K.E., 2000, Classification and regression trees—A powerful yet simple technique for ecological data analysis: Ecology, v. 81, p. 3178–3192.

Edwards, Jr., T.C., Cutler, D.R., Geiser, L., Alegria, J., and McKenzie, D., 2004, Assessing rarity and seral stage association of species with low detectability—Lichens in western Oregon and Washington forests: Ecological Applications, v. 14, p. 414–424.

Edwards, Jr., T.C., Cutler, D.R., Zimmerman, N.E., Geiser, L., and Moisen, G.G., 2006, Effects of sample survey design on the accuracy of classification tree models in species distribution models: Ecological Modelling, v. 199, p. 132–141.

Edwards, Jr., T.C., Cutler, D.R., Zimmermann, N.E., Geiser, L., and Alegria, J., 2005, Use of model-assisted designs for sampling rare ecological events: Ecology, v. 86, p. 1081–1090.

Engler, R., Guisan, A., and Rechsteiner, L., 2004, An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data: Journal of Applied Ecology, v. 41, p. 263–274.

Ferrier, S., and Watson, G., 1997, An evaluation of the effectiveness of environmental surrogates and modeling techniques in predicting the distribution of biological diversity: Armidale, Australia, Environment Australia, New South Wales National Parks and Wildlife Service, 184 p.

Fisher, R.A., 1936, The use of multiple measurements in taxonomic problems: Annals of Eugenics 7, p. 179–188.

Fisher, R.A., 1938, The statistical utilization of multiple measurements: Annals of Eugenics, v. 8, p. 376–386.

Franklin, J., 1995, Predictive vegetation mapping: geographic modeling of biospatial patterns in relation to environmental gradients: Progress in Physical Geography, v. 19, p. 494–519.

Franklin, J., 2002, Enhancing a regional vegetation map with predictive models of dominant plant species in chaparral: Applied Vegetation Science, v. 5, p. 135–146.

Frescino, T.S., Edwards, T.C., Jr., and Moisen, G.G., 2001, Modeling spatially explicit forest structural attributes using generalized additive models: Journal of Vegetation Science, v. 12, p. 15–26.

Gillham, J.H., Hild, A.L., Johnson, J.H., Hunt, E.R., and Whitson, T.D., 2004, Weed invasion susceptibility prediction (WISP) model for use with geographic information systems: Arid Land Research and Management, v. 18, p. 1–12.

Guisan, A., and Thuiller, W., 2005, Predicting species distribution—Offering more than simple habitat models: Ecology Letters, v. 8, p. 993–1009.

Guisan, A., and Zimmermann, N.E., 2000, Predictive habitat distribution models in ecology: Ecological Modelling, v. 135, p. 147–186.

Hastie, T.J., and Tibshirani, R.J., 1986, Generalized additive models: Statistical Science, v. 1, p. 297–310.

Hastie, T.J., and Tibshirani, R.J., 1990, Generalized additive models: New York, Chapman & Hall, 352 p.

Hosmer, D.W., Lemeshow, S., 2000, Applied logistic regression (2d ed.): New York, John Wiley & Sons, Inc., 392 p.

Jaberg, C., and Guisan, A., 2001, Modeling the influence of landscape structure on bat species distribution and community composition in the Swiss Jura Mountains: Journal Applied of Ecology, v. 38, p. 1169–1181.

Jensen, M.E., and Haise, H.R., 1963, Estimating evapotranspiration from solar radiation: Journal of Irrigation and Drainage Engineering—American Society of Civil Engineers, v. 89, p. 15–41.

Kent, M., and Coker, P., 1992, Vegetation description and analysis—A practical approach: Boca Raton, Florida, CRC Press, 348 p.

Lawler, J.J., and Edwards, T.C., Jr., 2002, Landscape patterns as predictors of nesting habitat—A test using four species of cavity-nesting birds: Landscape Ecology, v. 17, p. 233–245.

Leathwick, J.R., 1995, Climatic relationships of some New Zealand forest tree species: Journal of Vegetation Science, v. 6, p. 237–248.

Lodge, D.M., Williams, S., MacIsaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, P.B., Smith, M., Andow, D.A., Carlton, J.T., and McMichael, A., 2006, Biological invasions—Recommendations for U.S. policy and management, Ecological Applications, v. 16, p. 2035–2054.

Manly, B.F.J., 2004, Multivariate statistical methods—A primer (3d ed.): New York, USA, CRC/Chapman and Hall, 208 p.

McCullagh, P., and Nelder, J.A., 1989, Generalized linear models (1st ed.): London, Chapman and Hall, 208 p.

McGarigal, K., Cushman, S., and Stafford, S., 2000, Multivariate statistics for wildlife and ecology research: New York, Springer, 283 p.

McNoleg, O., 1996, The integration of GIS, remote sensing, expert systems and adaptive co-kriging for environmental habitat modeling of the Highland Haggis using object-oriented, fuzzy-logic and neural network techniques: Computers and Geosciences, v. 22, p. 585–588.

Moisen, G.G., Cutler, D.R., and Edwards, T.C. Jr., 1996, Generalized linear mixed models for analyzing error in a satellite-based vegetation map of Utah, *in* Mowrer, H.T., Czaplewski, R.L., and Hamre, R.H., eds., Spatial accuracy assessment in natural resources and environmental sciences: U.S. Department of Agriculture, Forest Service General Technical Report RM—GTR–277, p. 459–466.

Moisen, G.G., and Frescino, T.S., 2002, Comparing five modelling techniques for predicting forest characteristics: Ecological Modelling, v. 157, p. 209–225.

Norton, T.W., and Mitchell, N.D., 1993, Application of generalized additive models to wildlife modeling and population viability analysis, *in* Proceedings of the 7th Annual Symposium on Geographic Information Systems in Forestry, Environments and Natural Resources, Vancouver, Canada: Vancouver, Canada, Ministry of Supply and Services, 507–516 p.

Noss, R.F., 1987, From plant communities to landscapes in conservation inventories—A look at the Nature Conservancy (USA): Biological Conservation, v. 41, p. 11–37.

Noy-Meir, I., and Austin., M.P., 1970, Principal components ordination and simulated vegetational data: Ecology, v. 51, p. 551–552.

Nusser, S.M., Breidt, F.J., and Fuller, W.A., 1998, Design and estimation for investigating the dynamics of natural resources: Ecological Applications, v. 8, p. 234–245.

Parra-Olea, G., Martinez-Mayer, E., and Perez-Ponce de Leon, G., 2005, Forecasting climate change effects on salamander distribution in the highlands of central Mexico: Biotropica, v. 37, p. 202–208.

Pielou, E.C., 1984, The interpretation of ecological data—A primer on classification and ordination: New York, John Wiley & Sons, Inc., 288 p.

Schreuder, H.T., Gregoire, T.G., and Weyer, J.P., 2001, For what applications can probability and nonprobability sampling be used?: Environmental Monitoring and Assessment, v. 66, p. 281–291.

Schreuder, H.T., Gregoire, T.G., and Wood, G.B., 1993, Sampling methods for multiresource forest inventory: New York, John Wiley & Sons, Inc., 464 p.

Scott, J.M., Heglund, P.J., Samson, F., Haufler, J., Morrison, M., Raphael, M., and Wall, B., eds, 2002, Predicting species occurrences—Issues of accuracy and scale: Covello, Calif., Island Press, 840 p.

Shafii, B., Price, W.J., Prather, T.S., Lass, L.W., and Thill, D.C., 2003, Predicting the likelihood of yellow starthistle (*Centaurea solstitialis*) occurrence using landscape characteristics: Weed Science, v. 51, p. 748–751.

Sheley, R.L., and Petroff, J.K., 1999, Biology and management of noxious rangeland weeds: Corvallis, Oregon State University Press, 438 p.

Steele, B.M., 2000, Combining multiple classifiers—An application using spatial and remotely sensed information for land cover type mapping: Remote Sensing and the Environment, v. 74, p. 545–556.

Steele, B.M., and Patterson, D.A., 2002, Land cover mapping using combination and ensemble classifiers, *in* Proceedings of the 33rd Symposium on the Interface, June 13–18, 2001 Costa Mesa, Calf.: Interface Foundation of North America, Fairfax Station, Va., accessed December 16, 2010, at *http://www.interfacesymposia.org/I01/I2001HTML/BSteele.htm.*

Stockman, A.K., Beamer, D.A., and Bond, J.E., 2006, An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species: Diversity and Distributions, v. 12, p. 81–89.

Stockwell, D.R.B., and Peters, D., 1999, The GARP modeling system—Problems and solutions to automated spatial prediction: International Journal of Geographic Information Science, v. 2, p. 143–158.

Swan, J.M.A., 1970, An examination of some ordination problems by use of simulated vegetation data: Ecology, v. 51, p. 89–102.

Turc, L., 1963, Evaluation des besoins en eau d'irrigation, évapotranspiration potentielle, formulation simplifié et mise ä jour: Annales Agronomíques, v. 12, p. 13–49.

Venables, W.N., and Ripley, B.D., 1997, Modern applied statistics with S-PLUS (2d ed.): New York, Springer, 405 p.

Welch, N.E., and MacMahon, J.A., 2005, Identifying habitat variables important to the rare Columbia spotted frog in Utah (USA)—in an information-theoretic approach: Conservation Biology, v. 19, p. 473–81.

Yee, T.W., and Mitchell, N.D., 1991, Generalized additive models in plant ecology: Journal of Vegetation Science, v. 2, p. 587–602.

Zaniewski, A.E., Lehmann, A., and Overton, J. McC., 2002, Predicting species spatial distributions using presence-only data—A case study of native New Zealand ferns: Ecological Modelling, v. 157, p. 261–280.

Zimmermann, N.E., and Kienast, F., 1999, Predictive mapping of alpine grasslands in Switzerland—Species versus community approach: Journal of Vegetation Science, v. 10, p. 469–482.

Sampling and Survey Design

By Paul H. Geissler

Chapter 8 of
Early Detection of Invasive Plants—Principles and Practices
Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figure

Tables

Insets

Chapter 8.

Sampling and Survey Design

By Paul H. Geissler[1]

Overview

This chapter starts with a nontechnical discussion of various approaches to searching for invasive species, the objectives of early detection, and the importance of probability sampling. Small parks or other natural areas may not have the resources for extensive early-detection efforts, and some less extensive approaches are suggested for those. Whereas smaller parks are concerned with simplicity, larger parks are concerned with efficiency because they have the resources for the additional planning required to increase efficiency and a large enough effort to make that investment in planning worthwhile. After the general discussion, I will provide an introduction to some of the statistical survey design methods with examples. The reader may wish to read the introductory material to select an appropriate technique or techniques, and then either proceed to the technical description or seek statistical advice. A comprehensive discussion of all situations is beyond the scope of this chapter, but it is hoped that it will provide a useful introduction.

Search Strategies

Incidental observations by park staff and visitors as they pursue their normal activities around the park or other natural area can yield useful information for the early detection of invasive species. Often they are the first to notice a new invasive species, and a procedure is needed for collecting and verifying these incidental observations. Educational displays and brochures would improve this information by informing people how to identify the most likely invasive species. Although little effort is required to collect incidental observations, their usefulness is limited. The coverage is uncertain, and it is not known which areas have been searched and which areas have not. The ability of visitors and some staff to identify some invasive species is uncertain.

While recording and verifying incidental observations is an important first step, early detection should not stop there. With small parks and other natural areas, it may be feasible to completely search the entire park annually with trained observers (see Chapter 13 for an application). If the whole park cannot be searched annually, it may be feasible to establish a random rotation so the whole park is searched every so many years. Using a random rotation assures that a representative sample is available each year, allowing valid estimates of severity and trends to be extended to the whole park.

If it is infeasible to search the entire park, trails and roads should be searched first because invasive species commonly are first introduced to these areas by visitors and park operations. It is important to have trained observers search designated areas so that the area searched is known and there is reasonable confidence that the targeted invasive species can be found. The area searched should be recorded to avoid duplication and to allow extrapolation to larger areas. Ideally, all trails and roads should be searched annually at times that correspond with the displays of recognizable phenological traits by the target species. If this is infeasible, a random rotation should be established such that all trails and roads are searched every few years. If some trails or roads are more likely than others to have invasive species, they could be searched annually, while less likely trails and roads could be searched every few years in a random rotation. Again, the introduction of randomness to the survey design allows managers to extrapolate trends to all of the roads and trails in the park, including those that have not been sampled.

Search strategies for invasive species.

1. Collect and verify incidental observations
    a. Prepare educational materials describing potential invasive species and distinguishing them from similar, “look-alike” native species.
    b. **Advantages**: an important first step, low cost, many people looking.
    c. **Disadvantages**: uncertain coverage and identification, cannot extrapolate.
2. If practical, completely search the entire park or other natural area with trained observers.

[1]U.S. Geological Survey, Ecosystems, Status and Trends Program, Fort Collins, Colorado, Emeritus Retired.

a. If annual searches are impractical, set up a random rotation so the whole park is searched every few years.

b. **Advantage**: complete coverage.

c. **Disadvantage**: only practical for small parks because it could be expensive.

3. If resources are available, search trails and roads (areas most likely to harbor invasive species) with trained observers in addition to #1.

   a. If all trails and roads cannot be searched annually, establish a random rotation.

   b. Trails and roads most likely to be invaded could be searched annually, with others searched in a random rotation.

   c. **Advantages**: Area searched is known, reasonably confident that targeted species will be found if present, the areas most likely to be invaded areas are searched, can extrapolate to trail and road system and estimate severity there.

   d. **Disadvantages**: No information on invasive species away from trails and roads, ineffective for species dispersed by birds, wind, or water.

4. If some target species are dispersed by birds, wind, or water and if resources are available, search areas away from trails and roads with trained observers.

   a. If it is infeasible to access back-country areas because of costs, difficulty, or danger, search random transects perpendicular to trails and roads with trained observers, in addition to #1 and #3.

      i. **Advantage**: Some areas away from trails and roads are searched and can extrapolate to trail and road corridors.

      ii. **Disadvantages**: Parkwide estimates are not available because areas outside of trail and road corridors have no opportunity to be observed, and areas inside and outside of trail and road corridors are likely to differ because trails and roads are purposefully located relative to topographic features to facilitate ease of construction.

   b. If it is feasible to access back-country areas, search plots or transects randomly located in suitable habita throughout the park, in addition to #1 and #3.

      i. **Advantage**: unbiased parkwide estimates are available.

      ii. **Disadvantage**: Some areas of the park may be difficult or expensive to access.

Searching trails and roads covers some of the most likely areas to be invaded. However, some invasive species are spread by birds, wind, water, and other methods that do not depend on trails or roads. One approach to finding these invasive species is to search transects at right angles (perpendicular) to trails and roads. This will in effect widen the trail and road corridors by twice the length of the transects, covering the adjacent area. While this approach will search some areas away from trails and roads, it is likely that large areas of the park or other natural area will not have any chance of being searched. Therefore, the sample would not be representative of the park because trails and roads are purposefully located relative to topographic features to facilitate construction and to ease hiking. Consequently, the habitat away from the trail and road corridors is likely to differ from that within the corridors and require additional search efforts. On the other hand, access to the backcountry away from trails and roads may be difficult and expensive, and transects perpendicular to roads and trails may be the only feasible method of searching away from roads and trails.

If it is feasible to search the backcountry, search plots or transects randomly could be located in suitable habitat throughout the park or other natural area. Random sampling does not imply simple random sampling, and it is often advantageous to select relatively more samples in more likely and more accessible areas (see Box 8.2 for definitions). This approach will provide unbiased estimates of severity and trends for the entire park and will provide the best protection against infestations. Accessing the backcountry areas may be difficult and expensive, but it may be necessary to find invasive species dispersed by birds, wind, or water.

Selecting Search Strategies to Meet Objectives

There are usually four objectives for surveying invasive species:

1. To find and eradicate as many clusters of these species as possible in as short a time as possible,

2. To assess the severity of the problem in the park so that one can judge how much effort to put into eradication efforts,

3. To learn to be more effective in finding invasive species by improving predictive models used to develop search strategies, and

4. To evaluate the effectiveness of control strategies.

Regardless of the objectives, it is in the manager's best interest to find invasive species as quickly, efficiently, and effectively as possible. To do so, a model of the plant's distribution is used to guide the search. The model may be a simple mental model of the plant's ecology that includes searching all the roads and trails first because of potential seed dispersal along these corridors. It may also be a sophisticated mathematical species distribution model (see Chapters 6 and 7). The amount of time spent on developing a model depends on how much time will be allocated to search efforts. If only a day or two are available for searching, targeting trails and roads without developing a formal model may be the best available option. However, if searches will be more extensive or continuing, it makes sense to take the extra time to develop a better model to guide the search. In general, the more time spent searching, the more time one should spend developing a model to improve search efficiency. If the objective is solely to find and eradicate invasive species (objective 1), a probability (random) sample is not needed because nothing is being estimated. However, a probability sample is required for the other objectives. Frequently, one will need a probability sample to assess the severity of the infestation or improve the search efficiency.

For the second objective, where assessing the severity of the problem in the park is desired, an estimate of the distribution and abundance of invasive species in the park provides some information on how many infestations of invasive species have not been found. If searches were initiated in the most heavily infested part of the park, the estimate of invasive species prevalence cannot be directly extrapolated to the rest of the park. To assess conditions across the park, one needs to know the relation between the sample and the population for which estimates are required (the park or other natural area). This assessment requires a random (probability) sample.

For the third objective, information on the presence and absence of invasive species is used to improve predictive models so they become more efficient in finding invasive species. This case also benefits from a random (probability) sample because the model should reflect the natural processes occurring in the population. If the sample does not adequately represent the population, the resulting model will be flawed.

For the fourth objective, evaluating the effectiveness of control techniques does not require searches because the locations of the control areas are known. However, it is included here because it is a valuable followup activity. The objective shifts from finding invasive species and estimating severity to estimating the proportion of sites that have remained free of invasive species or to estimating prevalence of invasive species pre- and post-treatment. Often one will want to compare control methods. Statistically, this is an experimental design situation, not a survey design question.

Invasive Species Detection Surveys Compared to Monitoring Surveys

Searching for invasive species requires a different survey design from monitoring surveys because the objectives are different. Monitoring surveys primarily seek to detect trends–changes over time–whereas invasive-species surveys primarily seek to find invasive species and estimate their abundance or status. To detect changes over time, it is important to go back to the same sites and to estimate the change on those sites, removing the site-to-site variation from the error variance. Monitoring strata should never be changed because the strata specify the selection probabilities for the sites, which are remeasured. Invasive-species surveys focus on estimating current status rather than on estimating trends (changes over time). Surveys require a new, independent sample for each estimation period (for example year). It is advantageous always to go to new sites in order to cover as many new areas as possible and find as many infestations as possible rather than remeasuring old sites. The exception is for areas with a high likelihood of infestation such as trails, where all units may be included in the sample each estimation period or where the independent sample selects the same unit again. The strata should be updated for each estimation period (for example year) if there is new information available that would make sampling more efficient. However, strata must remain fixed within each estimation period.

Probabilities of selection for unequal probability sampling should be updated whenever units are selected, if there is new information. Here the focus is on optimizing the estimation of the current status, although this results in less precise estimates of trends. Independent surveys should be conducted for each estimation period (for example year) and comparisons made among the independent estimates. This is analogous to using an independent t-test instead of a paired t-test over time.

Comparison Between Surveys for Detecting Invasive Species and Monitoring Surveys

Invasive Species Surveys	Monitoring Surveys
Emphasizes status: severity, distribution	Emphasizes trend: change over time
New sites are selected each year to increase spatial resolution and estimate distribution	Sites are remeasured to increase temporal resolution and estimate change
Change over time estimated from independent surveys	Change over time estimated from repeat measurement of same sites
Restratify each year if there is updated information	Never change strata

Small Parks and Natural Areas

The procedures outlined in the rest of this chapter are relatively complex and attempt to answer the question of what can be done with the limited resources and time available to small parks. Invasive species commonly are first located by park staff during their normal activities or by park visitors. Recording these incidental observations provides an important source of information for all parks. It is well worth following up on these reports.

One cannot, however, rely on incidental observations alone because coverage is likely to be incomplete both because some likely areas may not be searched and because invasive species may not be recognized. If resources are limited, the visitor center and campgrounds would be good locations to start searching because many invasive species commonly are spread by human activity. If more resources are available, search along trails and roads, which also are likely invasion pathways. If all trails and roads cannot be searched each year, set up a rotation to cover them over several years. Ideally, though, park resources should allow for surveys to cover the broad range of habitats and conditions existing in the parks.

In general, if there is money and (or) the desire to initiate a comprehensive early-detection program, parks will want to prioritize search strategies in the following order (most desirable first) in accordance with the availability of resources:

- Small parks
    - Complete surveys of roads, trails, and other high-priority sites
    - Rely on incidental reports
- Medium-sized parks
    - Complete surveys of roads, trails, and other park areas, using rotating panel designs as required
    - Rely on incidental reports
- Large parks—More complex designs are covered in the remainder of this chapter, including the use of predictive models (Chapters 6 and 7).

Probability Samples

A random (probability) sample is the only way of assuring a representative sample. A probability sample is not necessary, however, if one only wants to find and eradicate invasive species without either assessing the severity of the problem in the natural area or improving the ability to find them. In this case, one is not using statistics to make inferences about the park or to improve search efficiency. Again, this approach may be sufficient if time is limited. For example, one may just search around the visitor center and campgrounds if only a few hours are available.

Box 8.1. Random sampling is the only way one can be assured of a representative sample and valid estimates.

With a probability sample, each unit in the population has a known probability of selection, and random selection is used to select the specific units to be included in the sample (see Box 8.2; Lohr, 1999; Scheaffer and others, 1990). Many people incorrectly equate random sampling with simple random sampling, where every unit of the population has the same probability of selection. With probability sampling, units that are more likely to have invasive species can and should be selected with greater probability to focus the search on areas that are most likely to have invasive species. Species distribution models (SDMs) and other modeling techniques can assist managers in defining focus areas that are more likely to be invaded (see Chapters 6 and 7).

Some have tried alternatives to probability samples, often with erroneous results. Failures are most obvious with surveys that predict the winner of an election because the true answer is known as soon as the election is held. With natural resource surveys, one seldom knows the right answer, so one is often happy with the wrong answer. Good examples of problems from natural resource surveys do not exist because unambiguous assessments of the true situation are lacking. Because of the difficulty of selecting a probability sample, some have used a convenience sample of those units which were easiest to measure. Two classic examples are the *Literary Digest* election poll of 1936 (Lohr, 1999) and Shere Hite's book *Women in Love: a Cultural Revolution in Progress* (Lohr, 1999). The *Literary Digest* mailed questionnaires to 10 million voters and 2.3 million were returned. The results of the poll were: Landon 55 percent, Roosevelt 41 percent, whereas the election results were: Landon 37 percent, Roosevelt 61 percent. Hite made a number of unbelievable claims including that 70 percent of all women married 5 or more years are having sex outside of their marriages. She mailed 100,000 questionnaires, but less than 5 percent were returned. Both examples had huge sample sizes, but used convenience samples and had high nonresponse rates, resulting in incorrect results.

Box 8.2. Random (probability) **sampling**, where each unit in the population has a known probability of selection and random selection is used to select the specific units to be included in the sample, is often confused with **simple random sampling**, where every unit of the population has the same probability of selection. With probability sampling, units that are more likely to have invasive species can and should be selected with greater probability to focus the search on areas that are most likely to have **invasive species**.

In the early years of the last century, statisticians debated the relative advantages of random sampling and purposive sampling, also known as quota sampling. Kruskal and Mosteller (1980) and Bellhouse (1988) provide interesting histories of survey sampling. Purposive and quota sampling select a sample that will match key characteristics of the population. An extensive test of purposive sampling of the records of the 1921 Italian census found the results to be unacceptable. After Neyman (1934) demonstrated the superiority of random samples in all but a few restrictive situations that are unlikely to occur in practice, statisticians have favored random samples over purposive ones. However, quota sampling persisted for a time in election polls (Scheaffer and others, 1990). There is a wonderful picture of a beaming President Truman holding up a newspaper with the headline "DEWEY WINS." Despite all the demonstrated problems with purposive and quota sampling, these methods are still used by some investigators. One can only hope that the theory of spontaneous generation is not still being used to explain the spread of invasive species.

Selecting Random Search Units

Divide the roads and trails into segments that can be conveniently searched. Enter these segments in a spread sheet and assign a random number to each segment by entering "=rand()" into the cell to the right of the segment identification. Then sort the segments in order according to the random numbers by highlighting the list and then selecting "Data" from the menu and then "Sort" from the menus. Note that the random numbers will change whenever a change is made to the spreadsheet, so the sorted list may not appear to be in order of the random numbers because they will have changed. Search the segments in this order, each year continuing with the next segment after the last one searched the previous year. Any sequential group of segments on this list is a random sample from the park roads and trails. The same procedure can be used with off-trail areas using convenient-sized search units that can be identified in the field instead of segments.

Searching randomly selected segments or search areas allows one to obtain an unbiased estimate of the number of segments or search areas with invasive species within the portion of the park that was subject to search, including those that remain undiscovered. This estimate provides an indicator of the severity of the problem. Random sampling is the only way to be assured of a representative sample and valid estimates. If searching is limited to roads and trails, that estimate only applies to the roads and trails in the park. Valid inferences cannot be made to other areas. If off-trail searches are limited to those areas frequented by visitors, inferences cannot be made concerning other areas of the park.

Estimating Trends for All Search Units

The number of segments or search areas with invasive species in the portion of the park subject to search is estimated from

$$\hat{y} = \hat{p}N \quad \hat{p} = (x/n) \quad v(\hat{y}) = N^2\left(\frac{N-n}{N}\right)\left(\frac{\hat{p}(1-\hat{p})}{n-1}\right)$$

Where n segments were searched out of a total of N segments and x segments were found to have invasive species. If there are $N = 50$ segments, $n = 20$ were searched and $x = 2$ had invasive species, then the proportion of segments with invasive species would be $\hat{p} = 2/20 = 0.1 = 10$ percent and the estimated number of segments with invasive species is $\hat{y} = (0.1)50 = 5$. The variance of this proportion is

$$v(\hat{y}) = (50)^2\left(\frac{50-20}{50}\right)\left(\frac{0.1(1-0.1)}{20-1}\right) = 7.1.$$

The finite population correction, or FPC = (N–n)/N reduces the variance so that it is zero when all segments have been searched ($N = n$), because there is no longer any sampling variation. A confidence interval can be estimated as

$$\hat{y} \pm t\sqrt{v(\hat{y})} = 5 \pm 2.1\sqrt{7.1} = 5 \pm 5.6 = (0,10.6),$$

where t is the t value for n–1 degrees of freedom and a 5 percent error rate (95 percent confidence). Note that zero is used as the lower confidence limit because the number of segments cannot be negative. To project the number of sample segments that are needed for the required confidence interval, note that the confidence interval half width, ignoring the FPC, is

$$w = t\sqrt{N^2\,\frac{\hat{p}(1-\hat{p})}{n-1}}$$. Rearranging $$n = \left(\frac{tN}{w}\right)^2 p(1-p)+1$$

Timed Searches

Timed searches are sometimes used to estimate trends, assuming that search efficiency of the observer and that the detectability of the plants have not changed over time. However, with invasive species, assessing the severity of the infestation is usually more important than monitoring its changes over time. Timed searches are not useful with invasive species because the relation between the sample and the population is unknown, and consequently the severity of the infestation cannot be assessed.

Survey Design

The first step in planning a search for invasive species is to develop a **sampling frame**, showing all the areas that may be sampled. This may be a map of the natural area, showing the boundaries. Managers may wish to include some areas adjacent to it. It may be necessary to exclude some areas of the park that are too dangerous to sample (for example where the slopes are too steep). Some areas may be excluded because they are inaccessible, and inferences cannot be made to any areas excluded from sampling. Often a difficult decision will have to be made, balancing the wish to protect the entire park from invasive species with the feasibility of sampling inaccessible areas.

Steps in Designing a Survey

1. Develop a sampling frame, showing all the areas that may be sampled.
2. Decide on appropriate sampling units.
3. Develop a conceptual or predictive model to identify invasion pathways, areas where invasion is likely, and sensitive or consequential areas.

Stratified Sampling

4. Define strata without skips or gaps with similar probabilities of having invasive species and (or) with similar sensitivity or consequence.

5. Set stratum sample sizes that will provide adequate precision within the available budget.

6. Select sampling units to be observed using GRTS (Generalized Random-Tessellation Stratified, recommended) or systematic sampling.

Unequal Probability Sampling

7. Estimate the probability that each sampling unit contains invasive species.

8. Increase that probability for sensitive or consequential areas to increase the probability of selecting those sampling units.

9. Select sampling units to be observed using GRTS (recommended) or systematic sampling with probability proportional to the modified probability that they contain invasive species.

The second step is to decide on an appropriate **sampling unit**. These are the independently selected units that will be searched. A sampling unit may be a stretch of trail, an off-trail transect, an area to be searched, or other convenient unit. To reduce travel costs, a cluster of several subunits may be selected within each primary sampling unit. Note that the subunits are not independent and that the sample size is the number of primary units, but sampling subunits will be beneficial in reducing the variance of the primary units. In a stratified survey (see below), different sampling units may be used in different strata. For example, a linear sampling unit may be appropriate along trails, and an area sampling unit may serve better off trails.

Focused Searches

> **Box 8.3. Stratified** and **unequal probability sampling** can be used to increase the probability of selecting units in areas where the invasive species is likely to occur and in especially sensitive or consequential areas.

At a minimum, searches should be focused on invasion pathways, routes along which invasive species are likely to move and to occur, and on especially sensitive areas with important native plants. A model, either conceptual or predictive (see Chapters 6 and 7), is needed to identify these areas.

Focusing searches would exclude simple random sampling, equally spaced grid sampling, and other approaches where all sampling units have the same probability of selection. Stratified and unequal probability sampling can be used to increase the probability of selecting units in areas where the invasive species is likely to occur and in especially sensitive or consequential areas. With stratified sampling, the sampling frame is divided into strata, without any skips or gaps. Then different sampling rates can be specified for each stratum. For example, all the trails may be searched, one sampling unit may be selected for every 5 hectares where the species is predicted to occur, and one sampling unit may be selected for every square mile in other areas. Unequal probability sampling can be viewed as a generalization of stratified sampling, where the probability of selection varies continuously across the park instead of remaining constant within strata and only changing among strata. Unequal probability selection is more flexible but is somewhat more complex and less familiar than stratified sampling. Stratification does not reduce the variance substantially unless there are large differences among homogeneous strata, so stratification is best viewed as a technique for focusing the sampling on critical areas rather than as a variance reduction technique.

Stratified Sampling

To select a stratified systematic sample, the model is used to predict the probability that invasive species will occur in different areas of the park. The model may be a predictive mathematical model (see Chapters 6 and 7) or a conceptual mental model, depending on the time and resources available. Clearly, better models will yield better results. Divide the park into strata with similar predicted probabilities of occurrence and with boundaries that are recognizable in the field. To estimate the number of samples necessary in each stratum, an estimate of the mean and variance of the response variable is needed, often from a pilot or similar study. The estimate and variance of the number of sampling units with an invasive species in stratum, h, including those that have not been found, from the binomial distribution are

$$\hat{y}_h = N_h \hat{p}_h \quad \hat{p}_h = x_h / n_h$$

$$\hat{v}(\hat{y}_h) = N_h^2 \left(\frac{N_h - n_h}{N_h} \right) \left(\frac{\hat{p}_h (1 - \hat{p}_h)}{n_h - 1} \right),$$

where p_h is the proportion of sampling units in which the invasive occurs, and where n_h out of N_h sampling units are observed in stratum h and of those, the invasive species was found in x_h units (Scheaffer and others, 1990; Thompson, 2002). In many situations, the finite population correction (FPC= $\left(\frac{N_h - n_h}{N_h} \right)$) will be close to 1 and can be ignored.

The stratified estimate and variance for the number of sampling units with the invasive species in the park are

$$\hat{y} = \sum_{h=1}^{L} \hat{y}_h \quad \hat{v}(\hat{y}) = \sum_{h=1}^{L} \hat{v}(\hat{y}_h)$$

The optimal number of units to select in each stratum to minimize the park variance is

$$n_h = round \left[c \left(N_h \sqrt{p_h (1 - p_h) / c_h} \Big/ \sum_i^H N_i \sqrt{p_i (1 - p_i)\, c_i} \right) \right]$$

(Levy and Lemeshow, 1991) where c is the total parkwide budget excluding fixed costs, c_h is the per unit cost in stratum h ($c = \sum n_h c_h$). Note that one divides the expression under the radical by cost in the numerator and multiplies by it in the denominator. The probability of finding at least one invasive in stratum h is $w_h = 1 - (1 - p_h)^{n_h}$. Note that the probability of not finding an invasive in one sampling unit (SU) is (1–ph) and the probability of not finding it in nh SUs is $(1 - p_h)^{n_h}$. Similarly, the probability of finding at least one invasive in the park is $1 - \prod_{h=1}^{H}(1 - w_h)$. One can set up a spreadsheet using these equations for a park, using an estimate of ph from the predictive models, Nh from the sampling frame, and nh from optimal allocation. This approach will allow one to try and evaluate alternate stratifications.

For example, say there are three strata as defined in table 8.1. Stratum A is a small area near campgrounds, roads, and trails where 10 percent of the sampling units (SU) are predicted to have the invasive species. Stratum C is a large area where only 0.1 percent of the sampling units are predicted to have the invasive, and stratum B is an intermediate area with 1 percent predicted to have the invasive. Because the invasive is expected to occur close to human activity, the cost to observe an SU in stratum A is less than in other strata. The optimal stratum sample sizes to minimize the variance of the park estimate of the number of SUs with invasive species are given in line 8 in table 8.1, using the formulas previously provided. Relatively similar sample sizes are projected for all strata in spite of the large differences in the projected occurrence of the invasive species because of the large differences in their areas. About 3 percent of the stratum A SUs will be observed, while fewer than 1 percent will be observed in the other strata. The prediction is that two SUs with invasive species will be found in stratum A and none in the other strata on average. The probability of finding at least one invasive is 83 percent in stratum A but only 15 percent and 1 percent in strata B and C, respectively, illustrating the difficulty of finding invasive species when they are rare.

Another strategy (line 15 in table 8.1) is to put almost all of the effort into stratum A, where the invasive is most likely to occur and to sample only two SUs in strata B and C. At least two SUs are needed to estimate the stratum variance. On average, one would find nine SUs with invasive species compared with two in the first allocation, but the park variance increases by an order of magnitude. This example illustrates the compromises between the objectives of:

1. finding and eliminating as many invasive infestations as possible and
2. estimating the number of infestations in the park and gathering information on a range of habitat conditions for use in improving the predictive model.

A small sample in a large stratum with low density has little chance of finding an invasive species, but if one is found, that stratum will have a large estimate and variance. For this reason, selecting only two SUs in a large stratum is not a good idea, and it may be better to exclude that stratum. If, for example, two units are selected in stratum C and one is found to have an invasive, the estimated number of SUs with invasive species will be (50 percent of SUs with invasive species)(7,500 SUs)=3,750 SUs with invasive species, when only eight SUs actually have them. The variance is $(7{,}500)^2(0.5)(1–0.5)/1 = 14{,}062{,}500$, without the FPC.

A third strategy (line 22 in table 8.1) is to exclude stratum C and allocate samples optimally to strata A and B. Although the estimate is biased low because the invasive species in stratum C are excluded, the variance is considerably less than the variance in the other two allocations. On average, four infestations will be found compared with two with the first allocation.

Box 8.4. It is useful to **explore sampling alternatives by using a spreadsheet**, examining the projected standard errors, number of infestations found, and the probability of detecting an invasive species, given that it is present.

Table 8.2 gives the results of simulating the sample allocations in table 8.1 with 5,000 replications. The first (optimal) allocation with 17, 16, and 13 SUs observed in strata A, B, and C, respectively, is reasonably well behaved. The distribution has a long right tail, as shown by the mean greater than the median and the presence of a large maximum, because the number of SUs with invasive species follows a binomial not a normal distribution. However, only a median of two infestations of invasive species (SUs with invasive species) is found. To find more invasive species, one is tempted to put almost all the observations in stratum A where the invasive density is highest, and only observe two SUs in strata B and C. At least two observations are needed to estimate a variance. The second allocation shows that this strategy occasionally results in extremely large estimates because a small sample is observed in a large stratum, as predicted in table 8.1. The maximum estimate is 3,801 SUs with invasive species, when only 78 SUs actually have them. However, this approach finds a median of nine infestations, compared to two with the first allocation. The third allocation does not observe any SUs in the large stratum C where the invasive density is the lowest and uses an optimal allocation for strata A and B.

The estimates are biased low because the invasive species in stratum C are excluded. However, the standard errors are smaller than those with the other allocations, and a median of four infestations was found. This example illustrates some of the compromises that are required in allocating stratum sample sizes. It is useful to explore alternatives by using a spreadsheet, examining the projected standard errors, the number of infestations found, and the probability of detecting an invasive species, given that it is present.

Table 8.1. Example of allocation of samples to strata.

Line	Quantity	Stratum A	Stratum B	Stratum C	Park
1	N population size	500	2,000	7,500	10,000
2	p proportion of sampling units with invasive species, in percent	10.0	1.0	0.1	
3	y number sampling units with invasive species	50	20	8	78
4	c_h cost per sampling units	1	2	4	
5	c total budget				100
6	$N\sqrt{p(1-p)/c}$	150	141	119	409
7	$N\sqrt{p(1-p)c}$	150	281	474	906
	First allocation				
8	n optimal sample size	17	16	13	46
9	Percentage of sampling units observed	3.40	0.80	0.17	0.46
10	Projected cost	17	32	52	101
11	$v(y)$ projected variance of y	1,358	2,619	4,675	8,652
12	$se(y)$ projected standard error	37	51	68	93
13	Projected number of invasive species found	1.70	0.16	0.01	1.87
14	Probability of finding at least one, in percent	83.3	14.9	1.3	86.0
	Second allocation				
15	n sample size	88	2	2	92
16	Percentage of sampling units observed	17.60	0.10	0.03	0.92
17	Projected cost	88	4	8	100
18	$v(y)$ projected variance of y	213	39,560	56,179	95,952
19	$se(y)$ projected standard error	15	199	237	310
20	Projected number of invasive species found	8.80	0.02	0.00	8.82
21	Probability of finding at least one in percent	100.0	2.0	0.2	100.0
	Third allocation				
22	n sample size	35	33	0	68
23	Percentage of sampling units observed	7.00	1.65	0.00	0.68
24	Projected cost	35	66	0	101
25	$v(y)$ projected variance of y	615	1,217	0	1,833
26	$se(y)$ projected standard error	25	35	0	43
27	Projected number of invasive species found	3.50	0.33	0.00	3.83
28	Probability of finding at least one, in percent	97.5	28.2	0.0	98.2

Table 8.2. Simulation of allocations in table 8.1, with 5,000 replications. Fifty percent of the values are between the 25th and 75th percentiles, and 90 percent are between the fifth and 95th percentiles.

	Minimum	5th percentile	25th percentile	Median	Mean	75th percentile	95th percentile	Maximum	True
First allocation (table 8.1 lines 8–13, with stratum sample sizes 17, 16, 13									
Estimated number of sampling units with invasive species	0	0	29	59	78	88	213	856	78
Standard error	0	0	29	40	54	47	133	602	
Number of sampling units found with invasive species	0	0	1	2	2	3	4	8	
Second allocation (table 8.1 lines 14–19), with stratum sample sizes 88, 2, 2									
Estimated number of sampling units with invasive species	0	23	40	51	79	63	80	3,801	78
Standard error	0	10	13	15	43	16	18	3,750	
Number of sampling units found with invasive species	0	4	7	9	9	11	14	19	
Third allocation (table 8.1 lines 20–25), with stratum sample sizes 35, 33, 0									
Estimated number of sampling units with invasive species	0	14	43	57	69	89	150	317	78
Standard error	0	14	23	26	36	62	69	127	
Number of sampling units found with invasive species	0	1	3	4	4	5	7	13	

Cluster Sampling

If a sampling unit (SU) is relatively large, one may want to record the number of subunits with invasive species instead of just noting whether or not the whole SU has invasive species to obtain a finer resolution measurement of the prevalence of invasive species. For example, if the SU is a 1-km segment of trail, the presence or absence of invasive species in each 100-meter subsegment could be observed and the number of subsegments with invasive species recorded. This approach changes the response variable from the number of SUs with invasive species to the number of subunits with them.

$$\bar{y} = \frac{1}{n}\sum_{i}^{n} y_i \quad \hat{y} = N\,\bar{y}$$

$$v(\hat{y}) = N^2\left(\frac{N-n}{N}\right)\sum_{i=1}^{n}(y_i - \bar{y})^2 \Big/ n(n-1)$$

where y_i is the number of subunits with invasive species in sampling unit *i*. For stratified sampling, these quantities would apply to each stratum.

The second use of cluster sampling is to increase efficiency when travel time is an issue. Travel time to get to an SU is often a large part of the cost of a survey. If it takes a long time to get to an SU, it makes sense to make a number of observations during each visit. An approach to reducing travel costs is to measure a cluster of subplots once one has arrived at an SU. This is another example of cluster sampling, which can be used with any sampling design. Cluster samples include subplots located within a larger plot and points along a transect. In cluster sampling, the mean of subplots is the "observation" for analyses and not responses from the individual subplots because the subplots are not independently selected and have artificially reduced variability compared to the whole population. However, the variance of the cluster "observation" is reduced because the variance is based on multiple responses from the subplots in the cluster. One should spread out the subplots as much as practical to maximize the variance within an SU. Because the total variance is constant, maximizing the variance within an SU minimizes the variance among SUs, reducing the variance of the estimate.

When using cluster sampling to reduce travel time, the optimal number of subplots depends on the relative similarity of the subplots within the cluster as measured by the Pearson correlation coefficient among the subplots within a cluster (intraclass correlation). Correlation can be estimated from an analysis of variance of the clusters using pilot data. For illustration, suppose that the pilot data measured two clusters, each with three subplots. The observations for the first cluster were 17, 15, and 13 and for the other cluster were 21, 18, and 24. Their analysis of variance is given in table 8.3, where n = 2 is the number of randomly or systematically selected points and m = 3 is the number of subplots at each point. Clusters are significantly different because the "between" cluster variation is larger than the "within" cluster variation by more than would be expected by chance alone. This implies that the clusters are more homogeneous than the park as a whole. It is expected because points that are close

Table 8.3. Analysis of variance of clusters.

Source	Sum of squares	Degrees of freedom	Mean square	Sum of squares	Degrees of freedom	Mean square	Probability
Between clusters	SSB	n-1	MSB	54	1	54	0.045
Within clusters	SSW	n(m-1)	MSW	26	4	6.5	
Total	SST	nm-1	MST	80	5	16	

together tend to be more similar than points that are widely separated because of environmental differences.

The intraclass correlation coefficient (ignoring the FPC) is estimated (Lohr, 1999) as

r = 1 – [m/(m–1)] (SSW/SST) = 1 – (3/2)(26/80) = 0.51.

The adjusted R^2 is $R_a^2 = 1 - MSW/MST = 1 - 6.5/16. = 0.594$
The optimal number of subplots to use in each cluster (Lohr, 1999) is

$$m = \sqrt{\frac{c_1}{c_2}\left(\frac{1}{R_a^2} - 1\right)} = \sqrt{\frac{4}{1}\left(\frac{1}{0.594} - 1\right)} = 1.7 \approx 2$$

which would round to two subplots, where c_1=4 is the cost or time required to travel to a point and c_2=1 is the cost or time to measure a subplot. The total cost excluding fixed costs is $c = c_1 n + c_2 n m$, where n is the number of plots (independently selected points) and m is the number of subplots in each plot. Note that more subplots are used when the travel cost or time to get to a point is larger and when there is a low correlation among the subplots.

Adaptive Cluster Sampling

Adaptive cluster sampling will increase the precision of estimates of severity of infestations when invasive species are clustered and the clusters are rare. This situation is likely to occur during the early stages of invasion. When a rare invasive species is found, it is prudent to search the surrounding area to see if there are other plants, as the original population could have dispersed propagules. Adaptive cluster sampling takes advantage of this additional information to improve the estimate of severity. It should be noted that Adaptive Cluster Sampling does not improve the ability to detect the total number of invasive populations in the area of interest, only the size/density of the one(s) that was detected by the original sample.

For example, consider the cluster occurrences of invasive species shown in figure 8.1. A systematic grid sample of four sampling units (SUs) finds a member of two of the clusters. The SUs adjacent in the cardinal directions to the unit with invasive species are observed. The process is repeated until all the SUs adjacent to all the SUs with invasive species have been observed. Clearly, one cannot directly extrapolate the proportion of units with invasive species to the entire park because one has artificially increased the proportion of units observed with invasive species by searching adjacent units.

The estimation must account for the adaptive cluster sampling. Each group of adjacent SUs with invasive species is called a network. Observe the mean number of subunits with invasive species in each network $w_i = y_i / m_i$, where y_i is the number of invasive species found in network i (asterisks in fig. 8.1) and m_i is the number of subunits with invasive species (grid cells in fig. 8.1). For the example, there are two networks with invasive species and two without them (w_1=3/2, w_2=0/1, w_3=0/1, w_4=3/4).

$$\bar{y} = \sum_{i=1}^{n} w_i \Big/ n \quad \hat{y} = N\bar{y} = N\sum_{i=1}^{n} w_i \Big/ n \text{ and}$$

$$v(\hat{y}) = N^2\left(\frac{N-n}{N}\right)\frac{\sum_{i=1}^{n}(w_i - \bar{y})^2}{n(n-1)},$$

where $\hat{y}$ is the estimated total number of subunits with invasive species in the park, $\bar{y}$ is the mean, n is the number of SUs in the initial sample, and N is the total number of SUs. For the example,

ȳ = ˆ(3/2 + 0/1 + 0/1 + 4/3)/4 = 0.708   ŷ = 100(0.708) = 70.8

$$v(\hat{y}) = (100)^2\left(\frac{100-4}{100}\right)$$

$$\frac{(3/2 - 0.708)^2 + (0/1 - 0.708)^2 + (0/1 - 0.708)^2 + (4/3 - 0.708)^2}{4(4-1)}$$

$$= 1{,}617$$

$se(\hat{y})$ = 40.2. The 95 percent confidence interval for the number of subunits in the park with invasive species is $\hat{y} \pm t_{\alpha,df}\, se(\hat{y})$ = 70.8 ± 2.78(40.2) = [0,182.6], where $t_{\alpha,df}$ is the t value for the 5 percent significance level and 3 degrees of freedom. Note that the confidence limit is not negative because the number of subunits cannot be negative. For more information see Thompson (2002) and Smith and others (2003 and 2004).

Sample Selection

With stratified sampling, a sample is selected independently in each stratum. **Random**, **systematic**, or **GRTS** sampling may be used to select the sample. **Systematic samples** are the most precise in the presence of environmental gradients and have the advantage of being the easiest to implement

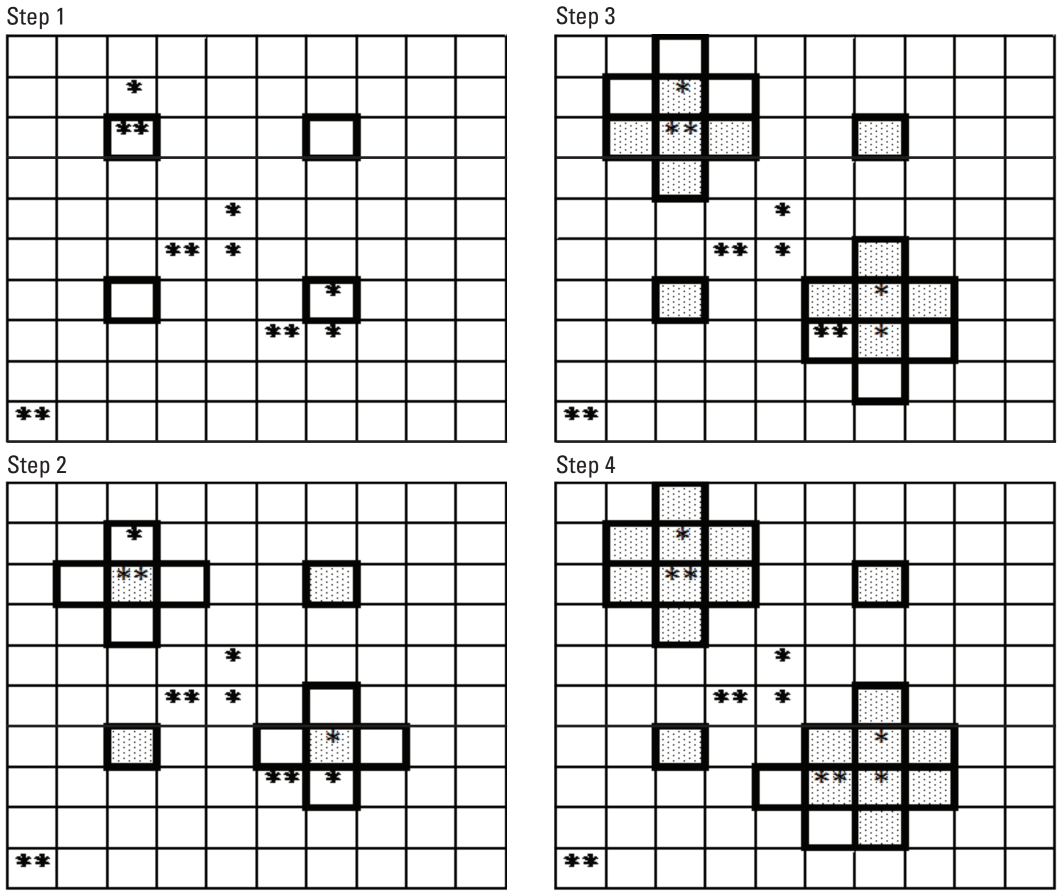

Figure 8.1. Example of adaptive cluster sampling, starting a with systematic sample of four sampling units (SU) and selecting the SUs with invasives north, south, east and west of the SU with invasives. SUs selected in previous steps are shaded. Asterisks indicate subunits with invasives.

in the field. **Random samples** tend to clump and leave gaps. Because of the spatial autocorrelation, widely spaced systematic samples may be negatively correlated, leading to a reduction in the variance compared to a random sample of the same size (Lohr, 1999; Thompson, 2002). This contrasts with compact cluster samples, where the correlation is positive, leading to an increase in the variance. However, the simple random-sampling variance estimator overestimates the variance for systematic samples in the presence of environmental gradients. If more than one (for example m=3) random start is used, an unbiased estimate of the variance can be calculated among the independent samples defined by the random starts (Strayer and Smith, 2003;Lohr, 1999, Scheaffer and others, 1990). This variance estimate does not appear to be as stable for small *m*. **GRTS sampling** (Stevens and Olsen, 2004) can also be used to select samples with stratified sampling.

Box 8.5. Generalized Random-Tessellation Stratified (GRTS) sampling has the advantages of allowing one to easily adjust the sample size if plans change or if "spare" sample units (SUs) are needed in case some SUs cannot be observed.

Although this method is more complex, it has the advantages of allowing one to easily adjust the sample size if plans change or if "spare" SUs are needed in case some SUs cannot be observed. GRTS maintains much of the spatial balance advantages of a systematic sample while the sample size changes. One cannot change the sample size with an equally spaced systematic sample without disturbing the spatial balance. GRTS also has a good associated variance estimator (Stevens and Olsen, 2003), which is better than the replicated survey variance estimator for systematic samples. GRTS sampling is recommended because the sample size can easily be changed while maintaining the spatial balance needed for precise estimates and because it has an excellent local variance estimator.

To illustrate the selection of a systematic sample, consider a population consisting of 25 sampling units, numbered 1, 2, …, 25. A sample of six units is required, with m = 3 random starts, each selecting n" = 2 units. The sample interval k = int(N/n′) = int(25/2) = 12. Pick m = 3 random numbers (0<r<1), say 0.42, 0.02 and 0.88. The starting SUs are int(r*k)+1 = 6, 1 and 11. If 0<r<(1/k), the first sampling unit is selected; if (1/k)<r<(2/k), the second sampling unit is selected and so forth. Add k = 12 successively to each starting number. Stop when the number is greater than N = 25. The systematic samples are 6, 18; 1, 13, 25; and 11, 23. Note that n′ is either 2 or 3, depending on the starting unit. Here n = ∑n′ = 7, not 6 as planned. To select a grid sample, select a systematic sample of the rows and a separate systematic sample of the columns to give a two-dimensional systematic sample.

A program is needed to select a GRTS sample. Download and use one of the following programs:

- S-DRAW (*http://www.west-inc.com/programs/S-Draw1c.zip*, For instructions, *http://www.west-inc.com/reports/grts.pdf)*
- R-GRTS (*http://www.epa.gov/nheerl/arm*)
- RRQRR (*http://www.nrel.colostate.edu/projects/starmap/rrqrr_index.htm*).

S-DRAW by Trent McDonald, WEST Inc., has been used here to illustrate the selection of a GRTS sample because it may be the simplest to use and does not require other software. Run S_DRAW, selecting two-dimensional, coordinates, randomize, and output sample in random order and provide the following values: sample size = 12, population size = 25, pixel size = 1, random seed =–1 and input frame = *c:\temp\sample.* txt. Defining the population as a 5 by 5 grid, the input frame in the text file was:

1 1
1 2
. . .
5 5

To allow “spare” SUs in case some selected SUs cannot be observed, select a sample size of 12, although only 9 are needed. The first nine units listed is the GRTS sample and the rest are spares. One can use as many SUs as needed, taking them sequentially from the start of the list. The largest pixel size that will allow each selected SU to have its own pixel is recommended. Selecting too large or too small a pixel size will yield samples without the desired spatial balance. Random seed =–1 indicates that the computer should generate its own seed. The GRTS sample on the grid follows with the order of the points indicated as 1, 2, …, c. Here a, b, and c are the spare points. If a sample size of five is desired, use the first five points (labeled 1 through 5).

. 2 3 . .
. . c . .
a 9 1 7 5
b . 8 . 6
. . . 4 .

Unequal Probability Sampling

Unequal probability sampling can be thought of as a generalization of stratified sampling that allows the probability of selection to vary continuously across the park, without sharp jumps at strata boundaries. For example, one may want to sample with probability proportional to the predictive model estimates of the probability that invasive species occur in an SU (see Chapters 6 and 7). The additional flexibility comes at the cost of some additional complexity, and the sample allocation formulas for stratified sampling are not available.

Table 8.4. Unequal probability selection of SUs. See Unequal Probability Sampling section for definitions.

Sampling unit	q	p	p′	Selected SUs
1	0.0090	0.0928	0.0928	X
2	0.0140	0.1443	0.2371	
3	0.0200	0.2062	0.4433	X
4	0.0180	0.1856	0.6289	
5	0.0100	0.1031	0.7320	
6	0.0080	0.0825	0.8144	X
7	0.0050	0.0515	0.8660	
8	0.0060	0.0619	0.9278	
9	0.0070	0.0722	1.0000	
Total	0.0970	1.0000		

As an example, consider a population of nine units (table 8.4). Calculate the probability of selection (q_i) proportional to the model estimates of the probability that the SU has invasive species (q_i) by dividing that probability by their sum

$$(p_i = q_i \big/ \sum q_j)$$

and then calculate the cumulative probabilities

$$p_i^{'} = \sum_{j=1}^{i} p_j \ .$$

To select a systematic unequal probability sample, select a number *r* from a random number table, where 0<r<1, and divide it by the sample size n = 3, say r/n = 0.236/3 = 0.079. This is the first point. Then add 1/n = 1 / 3 successively to select the other points (0.412, 0.745). Select the first SU with a cumulative probability larger than these numbers. For example, cumulative probabilities 0.000 to 0.093 are associated with the first SU. It is selected because 0.079 falls in this range. Cumulative probabilities 0.094 to 0.237 are associated with the second SU. It is not selected because no sample numbers fall in this range. In table 8.5, y_i is the observed value (1 if invasive species were found, 0 if none were found), and p_i is the selection probability from table 8.4. The inclusion probabilities, π_i, that point *i* is in the sample and π_{ij}, that both points *i* and *j* are in the sample are needed. Sampling with replacement,

$$\pi_i = 1-(1-p_i)^n \quad \pi_{ij} = \pi_i + \pi_j - [1-(1-p_i-p_j)^n] \, .$$

Inclusion probabilities with random sampling can be extremely difficult or impractical to calculate unless selection is made with replacement, but it is not a problem with systematic selection. Note that the probability of not selecting point *i* once is (1—p_i) and of not selecting it *n* times is $(1–p_i)^n$. Then the probability that it is in the sample is $1–(1–p_i)^n$. The probability of including both points *i* and *j* is the probability of including *i* plus the probability of including *j* minus the probability of including either *i* or *j*.

No single estimator for unequal probability sampling is uniformly better than the others (Thompson, 2002). The Horvitz-Thompson Estimator seems to work well where there is a proportional relation between y_i and π_i. Although it is unbiased, it can have a large variance if π_i and y_i are not well related.

$$y = \sum_{i=1}^{v} \frac{y_i}{\pi_i}$$

$$v(y) = \sum_{i=1}^{n} \sum_{j<i} \left(\frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}} \right) \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2$$

For the example in table 8.5,

$$y = \frac{0}{0.2533} + \frac{1}{0.4998} + \frac{0}{0.2276} = 2.0$$

$$v(y) = \left(\frac{(0.2533)(0.4998) - 0.0976}{0.0976} \right) \left(\frac{0}{0.2533} - \frac{1}{0.4998} \right)^2$$

$$+ \left(\frac{(0.2533)(0.227+) - 0.0419}{0.0419} \right) \left(\frac{0}{0.2533} - \frac{0}{0.2276} \right)^2$$

$$+ \left(\frac{(0.4998)(0.2276) - 0.0873}{0.0873} \right) \left(\frac{1}{0.4998} - \frac{0}{0.2276} \right)^2$$

$$= 1.188 + 0 + 1.212 = 2.400$$

If there is not a proportional relation between *yi* and *πi*, the generalized Horvitz-Thompson Estimator (aka Horvitz-Thompson Ratio Estimator) is recommended

$$\bar{y} = \sum_{i=1}^{n} \frac{y_i}{\pi_i} \Big/ \sum_{i=1}^{n} \frac{1}{\pi_i}$$

$$y = N\bar{y}$$

$$v(y) = \sum_{i=1}^{n} \left(\frac{1-\pi_i}{\pi_i^2} \right) (y_i - \bar{y})^2 + 2 \sum_{i=1}^{n} \sum_{j>i}$$

$$\left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right) \left(\frac{(y_i - \bar{y})(y_j - \bar{y})}{\pi_{ij}} \right) = S_1 + 2S_2$$

For the example in table 8.5,

$$\bar{y} = \left(\frac{0}{0.2533} + \frac{1}{0.5000} + \frac{0}{0.2276} \right) \Big/$$

$$\left(\frac{1}{0.2533} + \frac{1}{0.5000} + \frac{1}{0.2276} \right) = (2/10.342) = 0.1935$$

$$y = 9(0.1934) = 1.7405$$

$$S_1 = \left(\frac{1-0.2533}{(0.2533)^2} \right) (0-0.1934)^2 + \left(\frac{1-0.4998}{(0.4998)^2} \right)$$

$$(1-0.1934)^2 + \left(\frac{1-0.2276}{(0.2276)^2} \right) (0-0.1934)^2$$

$$= 0.4354 + 1.3024 + 0.5579 = 2.2958$$

$$S_2 = \left(\frac{0.0976 - (0.2533)(0.4998)}{(0.2533)(0.4998)} \right) \left(\frac{(0-0.1934)(1-0.1934)}{0.0976} \right)$$

$$+ \left(\frac{0.0419 - (0.2533)(0.2276)}{(0.2533)(0.2276)} \right) \left(\frac{(0-0.1934)(0-0.1934)}{0.0419} \right)$$

$$+ \left(\frac{0.0873 - (0.4998)(0.2276)}{(0.4998)(0.2276)} \right) \left(\frac{(1-0.1934)(0-0.1934)}{0.0873} \right)$$

$$= 0.3658 - 0.2442 + 0.4153 = 0.5369$$

$$V = 2.2958 + 2(0.5369) = 3.3696$$

Table 8.5. Observations with selection and inclusion probabilities. See Unequal Probability Sampling section for definitions.

Sampling unit	y_i	p_i	π_i	π_{i1}	π_{i2}	π_{i3}
1	0	0.0928	0.2533		0.0976	0.0419
4	1	0.2062	0.4998	0.0976		0.0873
5	0	0.0825	0.2276	0.0419	0.0873	

A computer package to calculate Horvitz-Thompson estimates and variances is available for the open source (free) R software environment for statistical computing and graphics (*http://www.r-project.org/*). The *psurvey.analysis* package for R by Tony Olsen and Tom Kincaid, USEPA, is available at *http://www.epa.gov/nheerl/arm/documents/psurvey.analysis_2.9.zip* (overview at *http://www.epa.gov/nheerl/arm/documents/intro.pdf*). After installing R, from the console menu select "Packages," then select "Install package(s) from local zip file," then browse for and select the zip file that was downloaded. Enter "help(*category.est*)" for help. To analyze the sample in table 8.5, enter library(*psurvey.analysis*); wgt= 1/c(0.0928, 0.2062, 0.0825); catvar=c(0,1,0); est=*category.est*(catvar, wgt, vartype="SRS", popsize=9).

The weights (wgt) are the reciprocal of the inclusion probability, and the response (catvar) is either a 1 indicating that an invasive was found or a 0 indicating that it was not. It estimates 1.57 units with invasive species, using a slightly different estimator, with a standard error of 1.95 (variance 3.80).

Again, S_DRAW is used to illustrate the selection of an unequal probability GRTS sample with the one-dimensional population in table 8.4. Check one-dimensional, coordinates, randomize, and output sample in random order and provide the following values: sample size = 3, population size = 9, pixel size = 1, random seed = –1 and input frame = *c:\temp\sample1*.txt. The input text file included the numeric values in the first two columns of table 8.4. The selected units (inclusion probabilities) are 4 (0.557), 9 (0.216), and 3 (0.619). A local variance estimator for GRTS samples is available in the *psurvey.analysis* package.

Detectability

During the early stages of invasion when invasive species are rare, it may not be appropriate to assume that invasive species are always detected when they are present in a sampling unit (SU). If they are not always detected, estimates of severity are biased low, and the bias will not be reflected in the confidence intervals. If estimating severity in this situation is an important priority, the bias can be removed by making repeated observations on the SUs, by using double observers, or by observing the perpendicular distance from an invasive to the observer's path. Accounting for detectability will take more time, but these methods are necessary if it is important to document the severity of an invasive that cannot reliably be detected.

Observing Perpendicular Distances

Box 8.6. It may not be appropriate to assume that invasive species are always detected when they are present in a sampling unit (SU), especially during the early stages of invasion when invasive species are rare. Accounting for detectability is necessary if it is important to document the severity of an infestation of an invasive that cannot reliably be detected.

One method of estimating and adjusting for the imperfect detectability of invasive species is to observe the perpendicular distance from an invasive species to the observer's path while recording the path length. This information is used to estimate a curve showing the decrease in detectability with distance. This approach estimates the number of individual plants or clumps of plants in the park, and it is necessary to have an operational definition of an individual plant or clumps of plants to clearly define the unit being observed. This method assumes that all plants exactly on the observer's path are observed, that the plants do not move in response to the observer, and that the perpendicular distance is measured accurately. There are several methods of estimating an "effective transect width" so the number of individuals or clumps can be estimated as (area of park) $n/(2wL)$ where n is the number observed, w is the effective width, and L is the length of the observer's path. For more information see Williams and others (2002) or Buckland and others (2001). A computer program (DISTANCE) is available at *http://www.mbr-pwrc.usgs.gov/software.html*.

Two Observers

Another method of estimating detectability is to use two observers. Information on plants found by one observer and missed by the other can be used to estimate the number missed by both observers. This method can be used with either independent or dependent observers. With dependent observers, if the primary observer finds an invasive species in a subunit, he or she tells the secondary observer. The secondary observer records whether the primary observer found an invasive in the subunit and whether the secondary observer found an invasive missed by the primary observer in the subunit. The secondary observer should not point out missed invasive species to the primary observer until the primary observer has passed the plant and it is clear that the primary observer has not detected it. The data for each sampling unit includes the number of subunits searched, the number of subunits found with invasive species by the primary observer, and the number of subunits where the secondary observer found an invasive when the primary observer did not find any. Observers should alternate in the role of primary and secondary observer. See Nichols and others, (2000) for more information.

With independent observers, each observer records the plants they detect without telling or indicating to the other observer. After they have finished, the observers compare notes and count the number of plants detected by both observers, the number detected by the first observer and not the second, and the number detected by the second observer and not the first. A computer program (DOBSER) is available at *http://www.mbr-pwrc.usgs.gov/software.html* to analyze data from dependent and independent observers.

Multiple Observations

Another method of estimating and adjusting for the imperfect detectability of invasive species is to use multiple observations of the same sampling units, either by the same observer at different times or by several observers independently at the same time. The presence or absence of a species is recorded, avoiding the problem of identifying individuals. Instead of estimating density, the proportion of area (or plots) occupied by the species is estimated. Observations should be over a suitably short time so it is reasonable to assume that there has not been any change in occupancy. Record the detection history of each sampling unit or subunit as a series of 0 s and 1s, where 0 indicates no detections and 1 indicates a detection. More information is available in MacKenzie and others (2006) and MacKenzie and others (2002). A computer program is available at *http://www.mbr-pwrc.usgs.gov/software/presence.html*.

Summary of Sampling and Survey Design

Survey design is similar to architecture, where one selects various components and designs a structure that meets the design objectives. A wide variety of components are available, and more than one way can be used to accomplish an objective. The nature of the design depends on the objectives, and different types of surveys are designed for different purposes. The design of an invasive-species survey differs from that of a monitoring survey, just as the design of a residence differs from that of an office building. The criterion for selecting one design over another is which one provides the most information and accomplishes the objective with minimal cost.

All information on invasive species is important, and invasive species commonly are first detected through incidental observations. However, planning can substantially increase the efficiency and effectiveness of invasive-species surveys, as it can with many endeavors. The amount of planning effort depends on the resources available. It is wasteful either to spend all the time planning or to start a vast unplanned effort. Simplified procedures for small parks and other areas with few resources and limited expertise have been suggested. If resources are available for a substantial effort, it is worthwhile to develop a probability risk assessment model (see Chapters 6 and 7) to predict where invasive species are likely to occur and to use that model to focus the searches on those areas. If the objective is only to find and eradicate invasive species, a probability sample is not needed because statistics are not used to make any estimates or inferences. However, if the objectives include an assessment of the severity of the infestation or a desire to improve the ability to predict where invasive species are likely to occur, a probability (random) sample is essential because only a probability sample can ensure that the sample is representative of the park.

To improve the efficiency of finding invasive species, searches should be focused on areas that are most likely to have invasive species, as predicted by the model. Unbiased estimates of the severity of the infestations are also needed as well as information to improve predictive models to their occurrence. Stratified and unequal probability sampling can concentrate sampling on the most likely area, while still providing probability samples that will allow unbiased estimates of severity. Stratified sampling is simpler and more familiar, whereas unequal probability sampling provides more flexibility by allowing the probability of selection to vary continuously across the park. Cluster sampling can be used in two ways. First, sampling units may be divided into subunits and the presence or absence of invasive species observed on the subunits to provide a finer resolution than would be available by just making the observation on the sampling unit. Second, cluster sampling can reduce traveltime by observing a cluster of subunits when one reaches the sampling unit. In this case, the subunits should be spread out as much as practical to minimize the variance among sampling units. During the early stages of invasion, invasive species may be clumped and rare. In this situation, the precision of severity of infestation estimates can be improved by using adaptive cluster sampling to search for invasive species in sampling units adjacent to those found to have invasive species. GRTS is recommended to randomly select the sampling units to be searched because it increases precision in the presence of environmental gradients by providing a spatially balanced sample, allowing the sample size to be easily adjusted, and having an excellent variance estimator; also, computer programs are readily available. In the early stages of invasion, searches may not always detect invasive species that are present in a sampling unit. If it is important to estimate the severity of the infestation, one needs to estimate the detectability and use it to adjust the severity estimate to account for the invasive species that have been missed during the searches.

Recommended Reading

- For an easy introduction to statistical surveys, consult Scheaffer and others, 1990, Elementary survey sampling: Duxbury Press, Pacific Grove, California, USA.
- Intermediate statistical surveys can be found in Lohr, (1999).
- A more advanced coverage of statistical surveys is provided by Thompson, (2002).

References Cited

Bellhouse, D.R., 1988, A brief history of random sampling methods, *in* Krishnaiah, P.R., and Rao, C.R., eds., Handbook of statistics: Amsterdam, The Netherlands, Elsevier, v. 6, p. 1–14.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., and Thomas, L., 2001, Introduction to distance sampling: Oxford, United Kingdom, Oxford University Press, 448 p.

Kruskal W., and Mosteller, F., 1980, Representative sampling, IV—The history of the concept in statistics, 1895–1939: International statistical review, v. 48, p. 169–195.

Levy, P.S. and Lemeshow, S., 1991, Sampling of populations—Methods and applications: New York, John Wiley & Sons, Inc., 768 p.

Lohr, S.L., 1999, Sampling: design and analysis: Pacific Grove, Calif., Duxbury Press, 450 p.

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A., and Langtimm, C.A., 2002, Estimating site occupancy rates when detection probabilities are less than one: Ecology, v. 83, p. 2248–2255.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.P., Bailey, L.L., and Hines, J.E., 2006, Occupancy estimation and modeling—Inferring patterns and dynamics of species occurrence: San Diego, Calif., Academic Press, 344 p.

Neyman, J., 1934, On the two different aspects of the representative method: Journal of the Royal Statistical Society, v. 97, p. 558–625.

Nichols, J.D., Hines, J.E., Sauer, J.R., Fallon, F.W., Fallon, J.E., and Geglund, P.J., 2000, A double-observer approach for estimating detection probability and abundance from point counts: The Auk, v. 117, p. 393–408.

Scheaffer, R.L., Mendenhall, W., and Ott, L., 1990, Elementary survey sampling: Pacific Grove, Calif., Duxbury Press, 486 p.

Smith, D.R., Brown, J.A., and Lo, N.C.H., 2004, Application of adaptive cluster sampling to biological populations, *in* Thompson, W.L., ed., Concepts, designs, and techniques for estimating population parameters: Covelo, Calif., Island Press, p. 77–122.

Smith, D.R., Villella, R.F., and Lemarie´, D.P., 2003, Application of adaptive cluster sampling to low-density populations of freshwater mussels: Environmental and Ecological Statistics, v. 10, p. 7–15.

Stevens, D.L., and Olsen, A.R., 2003, Variance estimation for spatially balanced samples of environmental resources: EnvironMetrics, v. 14, p. 593–610.

Stevens, D.L., and Olsen, A.R., 2004, Spatially balanced sampling of natural resources: Journal of the American Statistical Association, v. 99, p. 262–278.

Strayer, D.L., and Smith, D.R., 2003, A guide to sampling freshwater mussel populations: Bethesda, Md., American Fisheries Society Monograph, v. 8, 103 p.

Thompson, S.K., 2002, Sampling: New York, John Wiley & Sons, Inc., 400 p.

Williams, B.K., Nichols, J.D., and Conroy, M.J., 2002, Analysis and management of animal populations: San Diego, Calif., Academic Press, 817 p.

Process of Model Assessment and Evaluation

By Thomas C. Edwards, Jr., Richard Cutler, and Karen Beard

Chapter 9 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figures

Insets

Chapter 9.
Process of Model Assessment and Evaluation

By Thomas C. Edwards, Jr.,[1] Richard Cutler,[2] and Karen Beard[3]

Purpose

The investment in building a predictive model to facilitate the early detection of invasive plants can be significant (Chapter 7). However, before any such model can be applied to a management or conservation issue, it must undergo some form of assessment and evaluation. The purpose of this chapter is to present a methodology by which predictive models can be evaluated. It describes three distinct assessment and evaluation steps that should be undertaken in concert with model building, each of which provides different measures of model accuracies. Each of these steps should be considered integral to the model building processes described in Chapters 6 and 7, and no predictive model should be implemented without this assessment and evaluation. In addition, a set of specific metrics that are commonly used in model assessment and evaluation is described. These metrics represent a minimum set that should be applied to any invasive plant predictive model.

Overview

This chapter will provide you with an overview of:

- A three-step process of model assessment and evaluation (hereinafter validation), including:
    - Step 1: Measures of model fit
    - Step 2: Internal validation
    - Step 3: External validation
- The metrics of validation and
- What is meant by threshold-dependent versus threshold-independent metrics.

Overview of Predictive Model Assessment and Evaluation

Assessment and evaluation are crucial elements of any analytical pathway leading to the development of a predictive model of an invasive plant species. Unfortunately, different people equate model assessment with different processes, and no single authoritative source exists to combine all approaches or to suggest a generalized process to be followed. Rather, most effort has been expended on developing specific assessment metrics instead of formulating a systematic process for assessment and evaluation, with several edited texts (for example, Lowell and Jaton, 1999; Mowrer and Congalton, 2000; Hunsaker and others, 2001) and numerous papers (for example, Verbyla and Litaitis, 1989; Congalton 1991; Fielding and Bell, 1997) serving as general guidance.

Many of the commonly used approaches and associated metrics were developed as use of GISs exploded in the early 1990s, and there was a concomitant recognition that techniques for estimating the amount of error found in GISs was necessary (see Goodchild and Gopal, 1989), especially for remote-sensing applications (see Chapter 6). Given the rapid assimilation of GIS technologies into predictive species distribution models (SDM), the transfer of techniques and metrics developed for remote-sensing applications (see Congalton and Green, 1999) to SDMs was a logical next step, especially as managers and conservationists realized SDMs could easily be portrayed as spatially explicit representations (see Chapter 7).

The dynamic nature of plants and animals, however, has led to modifications of many of the metrics, including the need to account for prevalence (Manel and others, 2001), species rarity (Engler and others, 2004), and for many animals, detectability or occupancy (MacKenzie and Royle, 2005). Although many of the concerns noted for plants and animals are important aspects of model assessment and evaluation, most of the assessment metrics still rely on the same basic information.

[1]U.S. Geological Survey, Utah Cooperative Fish and Wildlife Research Unit, Utah State University, Logan, Utah.

[2]Mathematics and Statistics Department, Utah State University, Logan, Utah.

[3]Department of Wildland Resources, Utah State University, Logan, Utah.

Prevalence, rarity, and detectability (see also Chapter 8) are simply considered scalars that adjust the assessment metrics.

For this document, we suggest that validation of predictive models be centered on three basic steps and their associated metrics. These can be labeled:

- model fit
- internal validation
- external validation.

The first, **measures of model fit**, consists of those metrics used to determine how well the training data (see Chapter 7 for definition) fit the selected statistical model. Such measures are also often referred to as resubstitution metrics. **Internal validation** consists of those techniques that still rely on the training data, but through various randomization approaches split the data into small subsets for model building and subsequent evaluation. Its principal outcome is a determination of the amount of bias found in the training data, and how these biases may affect model performance. Last is **external validation**, which refers to an independent assessment of the model, typically accomplished by the collection of new data to which the model is applied and then evaluated.

This chapter begins with a description of the metrics of validation, followed by discussion of each of these three components: model fit and internal and external validation. Last, we propose a process that we feel should be applied to all invasive plant models. Note that it is not necessary to perform all the suggested steps for model validation. For example, the cost of an external validation effort can easily exceed the cost of creating the initial model. In contrast, internal validation only requires additional computing time, and costs relative to the collection of a new, external dataset are miniscule. Thus, the scale and extent of the invasive model to be validated could determine whether an internal validation effort only is warranted. In cases where invasive plant inventories have been planned but not yet conducted, it is possible to include field validation as part of a well-planned inventory effort.

Metrics of Validation

The basic data structure for all three validation steps (that is, model fit, internal validation, and external validation), and associated validation metrics, is termed the confusion matrix. For a 2-case nominal response, this is typically portrayed as a 2x2 table (fig. 9.1), of which the columns and rows refer to the observed and predicted values, respectively. The matrix is populated by tallying the number of model observations that fall into each of the four possible categories. Thus, a sample predicted as "1" (for example, presence) and observed as a "1" would be tallied as a [1,1], or true positive, assuming "presence" equates with positive. A true negative would similarly be tallied as a [0,0].

Note that the selection of "1" as present and "0" as absent is purely arbitrary; any coding scheme, such as "+" and "–", could be used as well. In terms of error, it is common to speak of false positives and false negatives, where [1,0] and [0,1] represent the appropriate coding schemes, respectively. False positives and false negatives are sometimes referred to as omission and commission errors, or Type I and II errors, respectively.

A variety of assessment metrics can be derived from the basic confusion matrix (Fielding and Bell, 1997, their table 2). Before selecting which metrics to calculate and report, however, it is important to understand how model-based "predictions," which are frequently continuous, are translated into the nominal codes of "0" or "1." Virtually all of the statistical tools described in Chapter 7 determine some sort of mathematical or "rule-based" algorithm based on the training data. This relation is then used to estimate a likelihood – or predictive value – that an observation noted as a "1" is actually a "1." For example, prediction likelihoods derived from a logistic model range from 0 to 1.0. Classification trees, which are still bounded by 0 and 1.0, result in discrete prediction likelihoods at the terminus of each tree node.

Each tool, however, still requires the translation of a continuous value ranging from 0 to 1.0 into a discrete classification of either 0 or 1. Thus, various thresholds can be established that convert a continuous range into a discrete value. How these thresholds are set is an important determinant of the eventual accuracies obtained from any predictive model. For example, a commonly used threshold is 0.5, with values greater than 0.5 assigned the discrete value of "1," while those

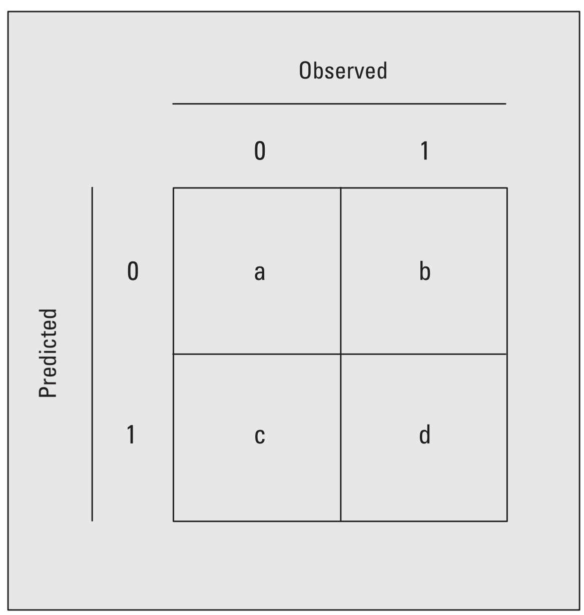

Figure 9.1. Basic confusion matrix, where "1" is assumed to be the event of interest, such as the presence of an invasive plant species, and the letters "a" through "d" are the observed frequencies of observations in each of the possible cells.

less than or equal to 0.5 are assigned a 0. Predicted values greater than 0.5 whose observed nominal value is "1" are considered correctly classified; those less than 0.5 but observed as "1" would be incorrect. Obviously, as you scale the threshold higher, say up to 0.7, some number of estimated values between 0.5and 0.7 would now be incorrect, or misclassified, as well. Accuracy measures whose values are determined by a threshold are considered threshold-dependent metrics. Those that do not rely on thresholds are considered threshold-independent metrics. Fielding and Bell (1997) provide an excellent overview of the basic metrics that can be used to assess SDMs, and readers are directed to that publication for definitions of each and their specific formulations.

Threshold-Dependent Metrics

Assuming "0" represents absence and "1" presence, values closer to "1" represent presence while those closer to "0" represent absence. The cutoff between presence and absence is, in the majority of cases, arbitrary. The selection of this cutoff threshold has implications for how the confusion matrix is populated, and consequently for all metrics subsequently derived from the confusion matrix. Because metrics derived from the confusion matrix are all based on a cutoff value that assigns the observation into one of the (usually two) possible nominal classes, they are sometimes referred to as threshold-dependent metrics (after Fielding and Bell, 1997).

Effects of the cutoff threshold are manifest in several areas. The first of these is prevalence, or knowledge of the true presence/absence ratio of the species being modeled. A threshold of 0.5 implicitly assumes that presence/absence is roughly equivalent. In some applications of classification models, such as those used to classify an individual as male or female based on morphometric characteristics (see Edwards and Kochert, 1986), a cutoff of 0.5 represents a sex ratio of 1:1, a valid assumption for many species. In other cases, where the model is attempting to identify locations of rare species on landscapes (for example, Engler and others, 2004; Edwards and others, 2005), an expected ratio of 1:1 (threshold = 0.5) may not be valid. This need to account for prevalence is well documented, as are its effects on several of the metrics derived from the confusion matrix (see Manel and others, 2001).

Threshold-Independent Metrics

Box 9.1. What values for assessment metrics indicate "good" models? No definitive guidance exists, although models where all or some subset of the assessment metrics exceeds 0.7 are common in the published literature. From a practical perspective, the **best measure is whether the model helps resolve the management or conservation issue,** even though this is a clearly subjective valuation.

In addition to threshold-dependent metrics, several different approaches for assessing model fit exist that are not dependent on an arbitrary threshold. Principal among these is the receiver operator characteristic (ROC) (Hanley and McNeil, 1982). Others include graphical methods, such as plots of the cumulative observed proportion for occurrences against the likelihood of occurrence (fig. 9.2, after Ferrier and Watson, 1997). These threshold-independent metrics do not rely on a user-specified threshold, but provide a model assessment based on different threshold levels.

One of the difficulties when evaluating assessment metrics is determining what values indicate "good" models. No definitive guidance exists, although models where all or some subset of the assessment metrics exceeds 0.7 are common in the published literature. From a practical perspective, the best measure is whether the model helps resolve the management or conservation issue, even though this is a clearly subjective valuation. For example, a management issue may require greater emphasis be made on the ability to predict the event of interest (that is, sensitivity; see Core Threshold-Dependent Metrics) rather than the nonevent (that is, specificity; see Core Threshold-Dependent Metrics), with management goals better realized by accepting greater error in the nonevent than the event. Thus, an accuracy value of 0.8 or greater for predicting the event and an accuracy of only about 0.5 for predicting the nonevent may be an acceptable management scenario, but not necessarily the best from a pure modeling perspective given the model's weak ability to correctly predict the nonevent. Ideally, such decision thresholds should be determined prior to engaging in any modeling effort.

Assessing Model Fit

Measures of model fit can be broken into two elements:

- The measures used to assess statistical fit, and
- Those metrics that provide an assessment of how well the predictions derived from the statistical model predict on the training data (often referred to as resubstitution accuracy or error, depending on which is reported).

In general, measures of fit describe the statistical variance or deviance reduction when the observed responses are fit to the selected predictors. In GLMs, such as logistic regression (see Chapter 7), a measure of fit, deviance (D^2), is obtained from the relation D^2=(null deviance – residual deviance)/null deviance, where maximum likelihood estimation is used to calculate deviance. In a perfectly fitting model, there is no residual deviance and D^2=1.0. McCullagh and Nelder (1983) recommend that D^2 be adjusted to account for the number of observations and predictors. As a measure, D^2_{adjusted} better represents the explanatory power of a fitted model than does simple D^2. It is also better suited to compare among different models having different observation sizes and predictor variables.

Box 9.2. Akaike information criterion (AIC) is best used for comparing among competing models developed from the same training data and a fixed set of predictors.

In practice, however, a model that has a low D^2, indicative of a poor fit, is likely to also have poor classification ability. This is no different than a simple linear regression model with a low R^2 having poor predictive capability. Consequently, in models for which the objective is purely prediction (Chapter 7) the reporting of D^2 has little utility. However, if a desire for explanation exists as well, the D^2 serves as an analog to the least-squares R^2 estimate and provides an overall measure of how good the predictors are at reducing model variance. From a pragmatic standpoint, D^2 may be useful to managers interested in understanding if they are targeting (or correctly predicting) areas most susceptible to invasion and (or) the key vectors and pathways leading to invasion. For nonparametric methods, such as classification trees, the Gini index also provides a measure of information content (Breiman and others, 1984). This metric is rarely reported.

Use of the Akaike information criterion (AIC) as a measure of information content (Burnham and Anderson, 2002) has increased in the last several years, allowing one to rank competing models. However, AIC as a measure is not standardized for comparison among models that use different predictor variables. It is best used for comparing among competing models developed from the same training data, and for which model difference is based on different subsets of a fixed set of predictors. Ideally, these competing models should represent a series of hypotheses that are being tested to determine which contains the best information for explaining or predicting the response. Several environmental variables have been linked to the presence of yellow starthistle, for example. AIC can assist managers with reducing the suite of possible predictor variables to a manageable subset upon which search strategies can be based.

Core Threshold-Dependent Metrics

Box 9.3. Core threshold-dependent metrics include **percent correct classification** (PCC) (a measure of the overall capability of the model to predict both "0" and "1"), **prevalence** (a measure of the frequency of the response of interest), **sensitivity** (a measure of the ability to predict the event of interest), **specificity** (a measure of the ability to predict the nonevent), and the **kappa statistic** (a measure of the agreement for the main matrix diagonals after the probability of chance has been removed; see fig. 9.1).

We concentrate here on a small subset of threshold-dependent metrics we feel are core to assessing an SDM (Box 9.3). These include percent correct classification (PCC), prevalence, sensitivity, specificity, and the kappa statistic (Cohen, 1960, 1968). PCC is a measure of the overall capability of the model to predict both "0" and "1." Prevalence is a measure of the frequency of the response of interest (for example, presence or "1") relative to the total N. Sensitivity and specificity are the ability to correctly model "1" and "0," respectively, assuming "1" is the response of interest, such as the presence of the invasive species of concern.

It should be obvious that as a measure of model performance, PCC can be greatly influenced by the number of 0's relative to 1's, or circumstances where prevalence of the targeted species (for example, presence of the modeled invasive) is markedly less than absence.

Assume, for example, a sample of 100 training points, where 90 of 100 observations are absence and only 10 are presence. It is possible to obtain a PCC = 90 percent by predicting all absences correctly and presences incorrectly. Such a model may have little to no utility as a management tool, although reporting PCC may lead one to believe the model is highly accurate. Nonetheless, PCC is a useful metric as long as prevalence is calculated and reported, too.

Given the example above, a sensitivity of 0 percent and specificity of 100 percent provide information beyond simple PCC = 90 percent and would allow for interpretation indicating that the model is quite good at predicting where the invasive plant is not found, but of no use for predicting where the invasive occurs.

A final threshold-dependent metric to consider is kappa. It is a measure of the agreement for the main diagonals (a and d in fig. 9.1) after the probability of chance has been removed (Cohen, 1960). Kappa ranges from $-\infty$ to 1.0, with values closer to1.0 indicating a better model.

Threshold-Independent Metrics

As the name implies, threshold-independent metrics are not based on the selection of a specified threshold for classifying the predicted observation as either of the two possible binomial outcomes (for example, [0,1], [present, absent]). Instead, model performance is evaluated across the continuum of thresholds from 0 to 1.0.

The ROC area under curve (AUC) is a measure of model performance obtained by plotting 1-specificity (false positive) on the x-axis versus sensitivity (true positive) on the y-axis for varying thresholds (fig. 9.2) (Hanley and McNeil, 1982). Chance is represented by the 45-degree diagonal. A good model will achieve a high true positive rate while the false positive rate is still relatively small; thus the ROC plot will rise steeply at the origin, then level off at a value near the maximum of 1 (fig. 9.2: PICO). The ROC plot for a poor model (whose predictive ability is the equivalent of random assignment) will lie near the diagonal, where the true positive rate equals the false positive rate for all thresholds (fig. 9.2: JUSC2). Visually, ROC plots indicating good model discrimination tend to "push" the plot into the upper left corner, away from chance.

The AUC, calculated from the ROC plot, is generally considered a better overall assessment metric than most threshold-dependent metrics given that it is not dependent

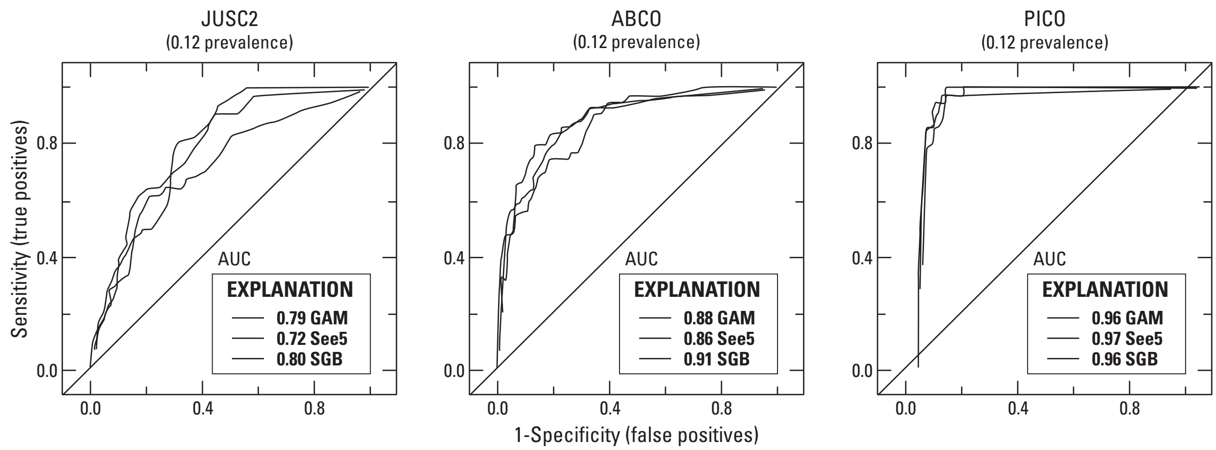

Figure 9.2. Example ROC plots. Note increase in AUC scores from JUSC2 to PICO as the ROC plot moves farther from the 45-degree line (from Moisen and others, 2006). See Threshold-Independent Metrics for definitions.

on a specified threshold, but is instead based on all possible thresholds. AUC ranges from 0.5 to 1.0, with values closer to 1.0 indicating a model with good classification power not explained by chance alone. Sometimes it might be desirable to compare AUCs from different models or to perform a formal test of whether the estimated AUC differs from chance. Procedures and formulae for estimating standard error (SE) about the AUC for use in such comparisons can be found in Hanley and McNeil (1982) or DeLong and others (1988).

Less quantitative, more descriptive methods also exist for assessing model predictive capabilities independent of a specified threshold. One approach is to simply calculate one or more threshold metrics, such as PCC or sensitivity, at different thresholds and then plot the assessment metric as a function of the threshold. Flatter curves would indicate a model less sensitive to threshold selection, although interpretation is clearly more subjective than that provided by the AUC. Another approach is to plot the cumulative observed proportion of occurrences against the likelihood of occurrence. One particular benefit of this graphic is that it allows for easy interpretation of whether model likelihoods of the event (for example, presence) are accurate.

Summary

Whether using complex computational models or less complex mental models to describe expected patterns of invasion, predictive models need to be validated to ensure that resources are not wasted on early-detection strategies built upon erroneous assumptions. A formal validation procedure can assist this process. A comprehensive validation process includes (1) measures of model fit, (2) an internal validation step (using training data), and (3) an external validation step (incorporating new, independent data). In each instance, sets of specific metrics of model assessment can be applied that generally fall into two categories: threshold-dependent metrics and threshold-independent metrics. While it is desirable to approach model validation using all three of these steps, it is also recognized that constraints on fiscal and staffing resources may prevent adoption of a comprehensive validation procedure. In most instances, internal validation will be the preferred option because collection of additional field data for external validation may prove to be cost-prohibitive.

Recommended Reading

For information covering metrics of model validation specific to remotely sensed data applications, see Congalton and Green, (1999).

For similar coverage of validation metrics for conservation models, see Fielding and Bell, (1997).

For a good discussion of issues of prevalence for ecological models, read Manel and others (2001).

References Cited

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone C.J., 1984, Classification and regression trees: Monterey, Calif., Wadsworth and Brooks/Cole, 368 p.

Burnham, K.P., and Anderson, D.R., 2002, Model selection and multimodel inference (2d ed.): New York, Springer, 488 p.

Cohen, J., 1960, A coefficient of agreement for nominal scales: Educational and Psychological Measurement, v. 41, p. 687–699.

Cohen, J., 1968, Weighted kappa—Nominal scale agreement with provision for scaled disagreement or partial credit: Psychological Bulletin, v. 70, p. 213–220.

Congalton, R.G., 1991, A review of assessing the accuracy of classifications of remotely sensed data: Remote Sensing of Environment, v. 37, p. 35–46.

Congalton, R.G., and Green, K., 1999, Assessing the accuracy of remotely sensed data—Principles and practices: Boca Raton, Fla., Lewis Publications, 183 p.

DeLong, E.R., DeLong, D.M., and Clarke-Pearson, D.L., 1988, Comparing the areas under two or more correlated receiver operating characteristic curves—A nonparametric approach: Biometrics, v. 44, p. 837–845.

Edwards, T.C., Jr., and Kochert, M.N., 1986, Use of body weight and length of footpad as predictors of sex in golden eagles: Journal of Field Ornithology, v. 57, p. 317–31.

Edwards, T.C., Jr., Cutler, D.R., Zimmermann, N.E., Geiser, L., and Alegria, J., 2005, Use of model-assisted designs for sampling rare ecological events: Ecology, v. 86, p. 1081–1090.

Engler, R., Guisan, A., and Rechsteiner, L., 2004, An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data: Journal of Applied Ecology, v. 41, p. 263–274.

Ferrier, S., and Watson, G., 1997, An evaluation of the effectiveness of environmental surrogates and modeling techniques in predicting the distribution of biological diversity: Armidale, Australia, Environment Australia, New South Wales National Parks and Wildlife Service, 184 p.

Fielding, A.H., and Bell, J.F., 1997, A review of methods for the assessment of prediction errors in conservation presence/absence models: Environmental Conservation, v. 24, p. 38–49.

Goodchild, M., and Gopal, S., 1989, Accuracy of spatial databases: Philadelphia, Pa., Taylor and Francis, 308 p.

Hanley, J.A., and McNeil, B.J., 1982, The meaning and use of the area under the Receiver Operating Characteristic (ROC) curve: Radiology, v. 143, p. 29–36.

Hunsaker, C.T., Goodchild, M.F., Friedl, M.A., and Case, T.J., eds., 2001, Spatial uncertainty in ecology—Implications for remote sensing and GIS applications: New York, Springer, 402 p.

Lowell, K., and Jaton, A., eds., 1999, Spatial accuracy assessment—Land information uncertainty in natural resources: Chelsea, Mich., Ann Arbor Press, 450 p.

MacKenzie, D.I., and Royle, A., 2005, Designing occupancy studies—General advice and allocating sample effort: Journal of Applied Ecology, v. 42, p. 1105–1114.

Manel, S., Williams, H.C., and Ormerod, S.J., 2001, Evaluating presence-absence models in ecology—The need to account for prevalence: Journal of Applied Ecology, v. 38, p. 921–931.

McCullagh, P., and Nelder, J.A.,1983, Generalized linear models (1st ed.): London, Chapman and Hall, 532 p.

Moisen G.G., Freeman, E.A., Blackard, J.A., Frescino, T.S., Zimmermann, N.E., and Edwards, T.C., Jr., 2006, Predicting tree species presence and basal area in Utah—A comparison of stochastic gradient boosting, generalized additive models, and tree-based methods: Ecological Modelling, v. 199, p. 176–187.

Mowrer, H.T., and Congalton, R.G., eds., 2000, Quantifying spatial uncertainty in natural resources—Theory and applications for GIS and remote sensing: Chelsea, Mich., Ann Arbor Press, 350 p.

Verbyla, D.L., and Litaitis, J.A., 1989, Resampling methods for evaluating classification accuracy of wildlife habitat models: Environmental Management, v. 13, p. 783–787.

Data Management and Management Response Strategies

By Penelope Latham, Sean Mohren, Susan O'Neil, and Bradley A. Welch

Chapter 10 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figure

Tables

Insets

Chapter 10.
Data Management and Management Response Strategies

By Penelope Latham,[1] Sean Mohren,[2] Susan O'Neil,[3] and Bradley A. Welch[4]

Overview

This chapter reviews the data management and management response strategies needed to implement an integrated invasive species early-detection program for a park or network of parks. The process of implementing an invasive species early-detection program is multifaceted and requires a great deal of coordination, both internally and externally. Key steps in this implementation process are outlined here.

1. A critical first step requires initial meetings with participating parks to determine their information needs, management objectives, and project scope (see Chapters 4 and 11).

2. Communication with adjacent landowners, especially potential Federal and State partners, follows logically, given that a primary goal of early detection is to locate new invasive species populations before they become entrenched and (or) serve as source populations for subsequent invasions (see Chapters 4 and 11).

3. After an initial assessment, the development of a system to serve stakeholders' needs and objectives will involve a series of operations that include:

    - collecting information about individual species and prioritizing species and (or) sites,
    - canvassing relevant databases in the parks and beyond to determine sources and quality of existing data,
    - developing strategies and systems to manage data strategically and incidentally collected,
    - field sampling and then analyzing, synthesizing, and disseminating key observations and findings to a rapid response network, and
    - finally, activating management responses.

In this chapter, we focus on the third step, or series of steps, in this process. The first two steps are covered in principle as part of Chapter 4 and in practice as the focus of the Klamath Network case study (Chapter 11).

Although the general sequence of steps runs from systems planning to data acquisition and analysis to dissemination and response (see Chapter 3, fig. 3.1), these steps are meant to be integrated. Effective integration of operations depends upon consistent and frequent communication between managers and the technical professionals developing the field sampling, data management, and response systems. Effective integration and communication help maintain clearly defined roles and responsibilities throughout the early-detection and rapid-response processes.

Data Management

Box 10.1. Data Management and Reporting. Well-articulated **data management and reporting schemes** not only improve internal management efficiency within a park but also foster communication with external stakeholders and partners. Many resource-management agencies and organizations are **standardizing invasive-plant databases** to facilitate data sharing and data comparability and to improve report integration at varying scales.

Before implementing a well-designed early-detection program, it is important to have data management and reporting procedures in place as well as having identified potential management options in the event that a new invader is detected (see figs. 3.1 and 3.2). All too frequently search strategies are implemented without careful consideration of what data need to be collected, how the information will be stored and reported, and how the current search event may affect future early-detection efforts. Well-articulated data-management and reporting schemes not only improve internal management efficiency within a park but also foster communication with external stakeholders and partners. If designed properly, data-management and reporting schemes also serve as the conduit for reporting at the regional and national scales, including, at the Federal level, the Government Performance and Results Act (GPRA) Goal 1AI for invasive-plant management. Many

[1] Pacific West Region, Inventory and Monitoring Program, National Park Service.

[2] Klamath Network, National Park Service.

[3] Puget Sound Partnership, Tacoma, Washington (United States).

[4] Natural Resources Program Center, National Park Service (United States), now Brecon Beacons National Park Authority (United Kingdom).

resource-management agencies and organizations, including the National Park Service, are beginning to standardize invasive-plant databases to facilitate data sharing and data comparability and to improve report integration at varying scales. This section highlights the key points to consider when developing a data-management and reporting plan.

A well-designed early-detection monitoring plan will generate a wealth of data and information that will prove useful to current and future resource managers. Compilation and management of presence and absence data alone will require metadata, consideration of taxonomic nomenclature, spatial context, fields related to sampling strata, intensity of effort, and ecological data, not to mention what may be required to track and evaluate the rapid-assessment and rapid-response efforts outlined in figure 3.1. Data management should be viewed as a process that begins with the conception and implementation of a project, continues through data collection and analysis, and culminates with data storage, archiving, and distribution (Michener and Brunt, 2000). A solid and comprehensive data-management system that takes into account the complexities of early detection of invasive plants is essential to accession, storage, and dissemination of quality information to support the management of park ecosystems. Effective data management for long-term monitoring will anticipate and accommodate changing technology, developing field methodologies, and most importantly, turnover in personnel.

Data Documentation (Metadata)

The standard for data management was set for the National Park Service in Director's Order (DO) #19. The DO states:

> "The National Park Service also has a strong business need for excellent records management, since the mission of the NPS is to care for natural and cultural resources so that they are 'unimpaired' for future generations. This requirement for managing resources in perpetuity sets a high standard for record keeping, as no resources can be managed well into the future without complete records of how they were managed in the past."

"Data" is broadly used to refer to all types of records and information collected or developed to meet project goals. Quality data can be rendered nearly useless for long-term needs if not documented in a manner that lets future users understand its content, purpose, and limitations. Consequently, metadata (information about the data) are essential for future users and interpreters of the data (Mohren, 2007). This is particularly true for those elements describing data quality and use, which form the basis of making informed decisions regarding the fitness of a particular data source (Chrisman, 1994). The metadata, at the least, should reference locations of key information about a project, usually found in project tracking databases, protocols, reports, and field notes.

Ideally, metadata documentation should encompass all data-related products such as photographs, spreadsheets, GPS data, products created during data analysis and reports, not just the raw data stored in a database or a GIS layer. The degree of documentation will vary by product type, but it should be done in a manner that provides current and future data users with the "who, what, when, where, why, and how" of the data or information. Providing sufficient documentation will ensure the data and information are available for future projects or analyses. To ease the task of metadata development for tabular and geospatial data, a number of software packages are currently available to NPS employees, including "NPS Database Metadata Extractor" (see *http://science.nature.nps.gov/im/gis/metadata.cfm*), and Environmental Systems Research Institute, Inc. (ESRI) ArcCatalog.

Data Quality

One of the most important aspects of data management is ensuring that data (and metadata) are of known quality. Quality is an encompassing term that includes objectivity, utility, and integrity. Where objectivity refers to the accuracy and reliability of the information, utility is the usefulness of the information to the target audience, and integrity refers to the security of the product. In order to detect a change in natural-resource trends or patterns over time, the acquired data need to be of high quality with minimal amounts of error or bias. Data of inadequate quality can lead to loss of sensitivity, which may result in misinterpretation of the information.

The level of quality desired will vary and depend on purpose, budget, available equipment, and personnel. Understanding the level of accuracy of a given dataset will allow the user certain levels of confidence when applying the data to management purposes. While it is always the goal to obtain 100 percent accurate data, errors inevitably occur from a range of sources and should be anticipated, minimized, and corrected where possible. Well-conceived and proven quality assurance and control (QA/QC) methods used from the planning phase through the archiving phase ensure that data are held to the highest possible standards for accuracy and precision.

Quality assurance (QA) is the planned and systematic pattern of actions needed to provide adequate confidence that the project fulfills expectations (that it is problem-free and able to perform the task for which it was designed). **Quality control** (QC) should be independent of the collection procedures and is the process of examining the data after they have been produced to make sure they are compliant with programmatic data quality standards.

Some standard (QA/QC) practices include:

- creating metadata files during project planning that are updated through every stage of the project,
- incorporating verification and validation methodologies to protect information being collected, recorded, and processed,

- incorporating domain values, pick lists, and various other quality-control methods into the databases prior to data entry,
- checking multiyear data sets for consistency to ensure that they can be integrated, and
- verifying data at multiple levels of collection and processing to correct errors and identify missing values.

Box 10.2. Methods of Data Verification 1. Visual review of data at time of data entry. 2. Visual review of data after data entry. 3. Duplicate data entry for a subset of data and compare data sets. 4. Review of randomly selected data subsets.

Data verification is the process of ensuring the data entered into a database correspond with the data recorded on the hardcopy field forms and data loggers. In general, there are four recommended methods for conducting proper data verification (Box 10.2). Visual review, at the time of data entry, is completed after each record is entered into the database by field staff. Visual review after the data have been entered is accomplished by comparing a printout of the data entered into the database to the hardcopy data sheets. Duplicate data entry is accomplished by entering all the records to the database and then entering a predetermined random number of records into a blank database. A query is then used to compare the records and report any mismatches. This method is more time consuming, but it provides a measure of the accuracy of data entry. Finally, a designated staff member who did not enter the data can review a predetermined subset of records and compare them to the original hardcopy forms. A timeline should be developed during the planning phase to outline the methods used to verify data, the number of records that will be checked, and a timeframe for verification.

While data verification can be completed by someone with little knowledge of the data, data validation requires a reviewer to have extensive knowledge about what the data mean and how they were collected. **Data validation** is the process of reviewing the finalized data to make sure the information presented is logical and accurate. The accuracy of the validation process can vary greatly and is dependent on the reviewer's knowledge, time, and attention to detail. Data-validation procedures include data-entry application programming, outlier detection, and general review. When possible, it is advisable to build filters for data that exceed logical values into the application used to store the data. For example, logic filters can prevent entry of 60 m for a tree diameter instead of the correct value of 0.6 m. Not all fields in a database will have appropriate domains; in these cases, data verification may suffice.

An **outlier** is an unusually extreme value for a variable, given the statistical model being used to analyze the data. While some outliers are a result of data contamination, they may also be indicators of important thresholds or extremes in variation of the parameter of interest. Statistical tests such as Grubbs' test and regression mapping can be used to examine the data for outliers (Michener and Brunt, 2000). It is generally advisable to flag and retain nonerror-associated outliers, allowing those conducting data analysis to make determinations regarding subsequent inclusion or rejection of specific data points. Lastly, individual review of data ranges and relations through tabular or graphical displays by someone intimately familiar with the types of data being collected is useful.

Depending on the type of data collected for early detection, some or all of these recommendations may be suitable. Presence/absence data are somewhat simpler to manage, verify, and validate than quantitative data. Even the simplest data sets require an organized data-management system and attention to detail.

Existing Data

Despite a compelling need for more scientific information about invasive plants in nearly all units of the National Park Service, any new monitoring efforts would be ill-served without first evaluating the historical information collected and archived in the parks and in other data repositories. In particular, early-detection efforts require species-specific life-history information that has, in many instances, been collected and summarized in data sets outside the NPS domain (see Chapter 5 for resources); the nationwide need to address invasive species has created a wealth of information and data already available that could support the objectives of many early-detection programs. Note however, that historical data may require additional processing and may be poorly documented, stored in a software format that is no longer supported, or only available in hardcopy format.

Efforts to determine potential high-risk species or areas where invasive species may enter the park will benefit greatly by acquiring available data from areas that surround the parks. The nature of the data, whether qualitative or quantitative, will restrict the range of subsequent options for data analysis, modeling, and hypothesis testing and may affect management options. It may become clear that the data available are not sufficient to meet the stated objectives. At this point, a decision must be made to acquire the needed data or to modify the objectives. For example, if the priority assessments identified require quantitative data and little or none are available, prioritization methods may need modification. If computer-generated predictive modeling is identified as desirable for hard to access, susceptible, or high-value sites, plans must be made to acquire appropriate data. Depending upon their resolution, remote-sensing products acquired for land-use change monitoring may also be useful.

Data Collection and Storage

Fundamental issues to consider when developing a data-collection and storage strategy are encompassed by the following four basic questions:

- How quickly will the data need to be made available, that is, certified, summarized and analyzed?

- What methods are available for data storage and who is the primary client (data user)?
- What information needs to be extracted from the raw data? What types of questions will routinely be asked?
- What is the skill level of personnel needing access to the data and how user-friendly is the tool (for example, the database) for delivering the data?

Climate and funding typically limit access to data in the field. If field access is of primary importance, a strategy that maximizes field time may be the best option. Under this scenario, field forms are used to collect the data with digital data entry occurring after the fieldwork is complete. When using this method, several QA and QC issues need to be addressed, such as identification of transcription errors, missing data or datasheets, and illegible datasheets. However, if the data need to be analyzed more quickly, electronic data collection might be a better method. Electronic data collection avoids some of the QA/QC issues associated with using paper datasheets and provides the opportunity to conduct analysis as data are collected or shortly thereafter. Additional funding may be needed to purchase the hardware and software and additional training is needed to train field staff to use electronic equipment. While other problems can be associated with electronic data collection such as equipment failure with potential data loss and difficulty in reading small screens, **with care and training, most problems can be avoided.** In addition to processing data collected during scheduled surveys, a method for periodically incorporating and evaluating incidental reports of invasive species is essential.

National Park Service Invasive Plant Database Standards

The National Park Service recognizes the need to develop nationally consistent tools for managing invasive species information that will meet park's objectives relative to resource management activities and scientific questions regarding plant invasions. In particular, early-detection and rapid-response efforts rely on predictable and transparent communication tools to engage an appropriate management response. At this time, the NPS Natural Resource Stewardship and Science Directorate (NRSS) is transitioning NPS data systems to a Service Oriented Architecture and XML (Web-based) services approach to data management and delivery. The project, called Integrated Resource Management Applications (IRMA), initially integrates three former NPS data systems—the NR Bibliography Inventory System (NatureBib), the NPS Data Store, and the NR Biological Inventory System (NPSpecies)—in a common Web portal. The new NPS Data Store includes data originally housed in the former Data Store plus data formerly housed in NatureBib (aka. References Application). Eventually, integration of these data systems with other NPS applications is planned. The NPS database that currently houses data about nonnative plants (APCAM) is not currently accessible through IRMA. NRRS is in the process of reviewing future needs for other NPS data systems through user need surveys, user boards, and systems analysis to understand how future development can be leveraged to better take advantage of existing systems. For more information about IRMA see *http://irma.nps.gov/App/*, accessed August 4, 2014.

Future revisions will, at a minimum, incorporate North American Invasive Species Management Association (NAISMA) standards (tables 10.1 and 10.2; for more information, see *http://www.naisma.org/*, accessed March 24, 2014). However, these fields are considered the minimum fields needed to help integrate data from multiple projects by various groups. Stohlgren and others, (2005), in their paper entitled "Beyond North American Invasive Species Management Association Standards," outline steps that include some simple study design suggestions and field methods that will make the NAISMA standards more statistically sound and increase the power of the data collected by allowing for greater inference across unsampled areas (see Chapter 8 for other options).

Compatibility with Other Database Systems

Ready transfer of information among land-management agencies and organizations is essential for effective rapid response to invasive species early-detection information. The new NPS data system based on service-oriented architecture (SOA) will allow data exchange and integration among different data systems through a system of agreed-upon data standards. Although currently only available to NPS users and for a limited number of applications, the NPS data portal will eventually be open to partners and the general public, allowing access to non-sensitive publications, data and information.

Taxonomic Standards

For NPS users, taxonomic nomenclature should be compatible with NPS taxonomic standards. After confirming and assigning management response priorities to the species identified during early detection, a process should be developed to transfer or make available certified species data from the invasive-species database to NPSpecies which is now retired as a stand-alone data system and integrated into IRMA. NPSpecies is the Park Service's master database for documenting the occurrence of species in more than 270 national park units that contain significant natural resources. While the initial focus of this database was on vertebrates and vascular plants, it is now designed to include all taxa and all parks. NPS adopted the Integrated Taxonomic Information System (ITIS) as the interagency taxonomic standard, but allows parks to use other taxonomic authorities to generate species lists with preferred scientific and common names to meet local needs. Any scientific name, including outdated names or names from various taxonomic authorities that might be listed on a voucher or in a report, can be entered into the database, but ideally the record has a valid ITIS taxonomic serial number, or TSN.

Table 10.1. Standard fields for invasive species as outlined by the North American Weed Management Association (NAWMA). For more information, see *http://www.naisma.org/* accessed March 24, 2014.

Field	Description
Collection date	The date the weed infestation was observed in the field. It does not refer to the date information was entered into the computer.
Examiner	The individual who collected the information in the field, at the site of the infestation.
Genus, species, intraspecific (optional), authority	These fields will contain the scientific or species name of the weed. The scientific name consists of the genus name followed by the species name, in Latin. Some plants are further classified into subspecies or variety. Lastly, the individual who first classified the plant and assigned the scientific name is called the authority.
Common name	The English or Spanish name for the plant.
Plant code	3 – 10 digit codes for scientific names.
Infested area, unit of measure	Area of land containing one weed species. An infested area of land is defined by drawing a line around the actual perimeter of the infestation as defined by the canopy cover of the plants, excluding areas not infested. Areas containing only occasional weed plants per acre do not equal one acre infested. Generally, the smallest area of infestation mapped will be 1/10th (0.10) of an acre or 0.04 hectare.
Gross area, unit of measure	This field is intended to show general location and population information. Like "*Infested area,*" it is the area of land occupied by a weed species. Unlike "*Infested area,*" the area is defined by drawing a line around the general perimeter of the infestation not the canopy cover of the plants. The gross area may contain significant parcels of land that are not occupied by weeds.
Canopy cover	Canopy cover will be estimated as a percentage of the ground, covered by foliage of a particular weed species. Cover will be recorded as a numeric value. If inventory procedures include the use of cover classes such as the Greater Yellowstone Area 10-point codes, or Daubenmire codes, the midpoint of the cover class will be entered as the cover value.
National ownership	The ownership of the land where the infestation is located. Ownership will consist of two, tiered groups. The first tier, "National ownership," will identify broad categories of land ownership, such as Federal, Provincial, State, county, city, and private lands. Codes are available for the various Federal agencies and should be entered here. Individual private landowners will not be identified. Individual State and Provincial land-management agencies will not be coded in this field. The second ownership field, "Local ownership" is reserved for these codes and is described in the following section.
Local ownership	This second ownership field is reserved for State and local users. There is no consistency in the naming of State and Provincial agencies, nor is there consistency in which branch of government manages these lands. It would therefore be difficult to create useful coding conventions for these entities at this time. This field will be available to regional or local entities to define and establish useful codes.
Source of the data	This field refers to the owner or manager of the data. This may be a different person or entity from the landowner or the person who collected the data. It may be an office manager or a database specialist. This entity that will be responsible for answering questions about the data or be responsible for data requests.
Country	The nation or country in which the infestation is located. Separate records or mapping polygons will be created for infestations that cross international boundaries.
State_Province	The State or Province where the infestation is located.
County_Municipality	The county (United States, Mexico and Canada) or municipality (Canada) where the infestation is located.
HUC_ Number	The Hydrological Unit Code, or HUC number, is a unique number assigned to the 2,000 major watersheds in the United States and Puerto Rico. The U.S. Geological Survey (USGS) has divided all the water systems in the United States into watersheds using the following system.

Table 10.1. Standard fields for invasive species as outlined by the North American Weed Management Association (NAWMA). For more information, see *http://www.naisma.org/* accessed March 24, 2014.—Continued

Field	Description
Legal, Latitude and Longitude (Lat-Longs), Universal Transverse Mercators (UTMs)	The location of an infestation will refer to the center of the infestation or the center of the polygon, which defines it. Today, location can be described using a variety of tools; any of the following methods may be used: legal, metes and bounds, Lat-Longs, and UTMs.
Quad number	This is the identification number, which appears on the corner of the quadrangle (quad) map. In the United States this refers to maps published by the U.S. Geological Survey (USGS). In Canada these maps are part of the National Topographic System maintained by the Geological Survey of Canada.
Quad name	This is the name that appears on the quadrangle map. It often refers to a prominent geographic feature, town or identifiable point in the area.

Table 10.2. Standard survey fields for invasive species as outlined by the North American Weed Management Association (NAWMA). For more information, see *http://www.naisma.org/*.

Field	Description
Area surveyed	The field refers to the entire land area that was surveyed for weeds, whether weeds were found or not. Information will be recorded in two data fields, the area surveyed and a unit of measure.
Type of survey	This field refers to the survey method. At this time only two survey methods are recognized: observed and remote. Observation refers to surveys that were conducted by direct observation or visiting the site of infestation. The observations can be made in many ways: helicopters, vehicles, horseback, or on foot. The second option is remote sensing. This refers to any survey that was conducted by using aerial photography, satellite imagery, or any method where the infestation was not directly observed.
Survey begin date	This field refers to the date the survey was started. It does not refer to the date that information was entered into the database.
Survey completion date	This field refers to the date the survey was completed. It does not refer to the date that information was entered into the database.
Legal, Latitude and Longitude (Lat-Longs), Universal Transverse Mercators (UTMs)	The location of a survey will refer to the center of the polygon that defines it. Today location can be described using a variety of tools; any of the following methods may be used: legal, metes and bounds, Lat-Longs, and UTMs.
Quad number	This is the identification number, which appears on the corner of the quadrangle (quad) map. In the United States this refers to maps published by the U.S. Geological Survey (USGS). In Canada these maps are part of the National Topographic System maintained by the Geological Survey of Canada.
Quad name	This is the name that appears on the quadrangle map. It often refers to a prominent geographic feature, town or identifiable point in the area.

Database Functions

Consideration of the questions the database will be used to answer is a critical preliminary step in developing the invasive-species database. Invasive-species data collected during early-detection monitoring will likely be used in a variety of different ways. Some possibilities include providing species data to help support early-detection prioritization (Chapter 5), locating and prioritizing sites for treatment, contributing to and improving predictive modeling efforts whether conceptual or computer generated (Chapters 6 and 7), tracking treatments, and monitoring spread of populations identified through early-detection surveys. Establishment of a user board to evaluate and prioritize these needs is essential. To ensure that the database meets user needs, a good practice is to develop diagrams of table and field relations, data entry screens, and associated field forms, and to have a fairly clear idea of what the standardized reports will look like.

To successfully implement network-scale early-detection invasive-species monitoring, the needs of the parks and network will need to be clearly understood as well as the benefits of standardizing data-collection methods.

Care should be taken when prioritizing the tools and reporting elements of a database. Having requirements, beyond what is absolutely necessary, can make a database or data collection process more complex which may require a database developer, extended training time for project leads and crew leaders, and more advanced hardware and software.

Data Management Roles and Responsibilities

Developing a data-management program where every person involved in the project understands the roles and responsibilities they have toward data management throughout all phases of a project is essential to obtaining quality information. As long-term monitoring projects can outlive the staff that are currently dedicated to those projects, documenting responsibilities is crucial. In the case of invasive species, it is essential to report and distribute the early-detection information acquired through a rapid-response network in a timely manner that will meet management priorities. Prior to implementing monitoring, it is recommended that staff involved in the project document what the final products for the year will be, where they will be stored, how and where they will be distributed, and who is the primary person responsible for each task and product. It is useful to make a list and sequence of all the tasks (with time frames) that will be conducted by the field crew, crew leaders, project leaders, data managers, rapid-response coordinators, taxonomic experts, statisticians, and any other staff directly involved in the project.

Reporting

Reporting involves both intraorganizational recording of activity and conveyance of information to external audiences, such as participating weed management partners, other parks, and other agencies. Reporting assists managers with:

- justifying management actions,
- documenting the state of natural resources,
- improving predictive models and search strategies,
- identifying emerging issues,
- quantifying successes,
- isolating problems,
- communicating with and alerting stakeholders, and
- directing rapid-response efforts (National Invasive Species Council, 2003).

Most agencies and organizations require internal tracking and evaluation of invasive-species programs. Many organizations, especially Federal agencies, require data and reporting to be rolled up to regional and national levels for cost and performance evaluation across the organization (for example, Government Performance Reporting Act). Reporting can be incorporated into the data-management and reporting plan such that standardized forms are produced periodically, evaluated, and disseminated to target audiences. This plan should also identify who is responsible for creating the reports, how the reports will be distributed, and where they will be stored. In the case of early-detection programs, it will be necessary to review previous reports to determine the status of invasive species over time. In the case of rapid-response efforts and identifying emerging issues, a process should be in place to create reports that can be delivered to the manager in a short period of time. Consequently, having a designated location to store the reports will improve efficiency and reduce frustration when an individual is looking for information about the project. The USGS nonindigenous aquatic species database (*http://nas.er.usgs.gov/AlertSystem/,* accessed March 24, 2014) provides a good example of an early-alert initiative.

Management Response Strategies

Once early-detection reports are generated, an appropriate management response strategy must be determined. Questions arise regarding how rapid the response should be, how imminent the threat is, and how severe the potential ecological damage will be. Consideration of these questions is necessary because budgets and staff time can be just as restrictive for response measures as they are for search strategies. Consequently, priorities must be set for response strategies just as they are for the search strategies used (Chapter 5), giving due consideration to the resources at hand (see Chapter 11) and the defined management objectives (see Chapter 4). Options available to resource managers will vary with the stage of invasion exhibited by the target species (Chapter 2, fig. 2.1) and with site conditions (Chapter 5). As invasive-plant populations become more entrenched, management strategies become progressively more limited and substantially more expensive to implement (fig. 10.1). Concurrently, the ecological impacts realized become more severe. If for no other reasons than these, it is advisable to respond quickly and comprehensively to invasive plant populations that are detected early.

The following information provides one approach of how early-detection and rapid-response strategies can be integrated. This example is based on a process and document produced by Spencer (2005, written comm.) to set invasive-plant management priorities and determine appropriate actions for Dinosaur National Monument in Colorado and Utah. We include this example because it integrates the early-detection principles detailed in this document and applies them in a management response framework that is broadly applicable.

Setting Management Response Priorities

Because it is infeasible to control every invasive plant that occurs in a natural area, it makes sense to focus management efforts on those species that have or could have the greatest impact to natural resources and to the highest value, at-risk habitats (see Chapter 5).

Ideally, managers prioritize invasive species and affected habitats on the basis of several aspects of the species' relative invasiveness, relative importance (the sum or average of measures of relative abundance such as relative cover, frequency, or density), or quality of affected habitat. Information to be assessed will vary somewhat depending on the stage of invasion being addressed. Chapter 5 covers prioritization in more detail, but examples of pertinent information include:

1. ecological impact (risk to regional biodiversity, adverse impacts to soil resources, capacity to alter forage availability, and so forth),
2. current distribution and abundance,
3. trend in distribution and abundance,
4. control feasibility/management difficulty, and
5. value of ecological areas or management units.

Persons conducting the assessment should understand resource management priorities and possess knowledge of the resource and (or) sites being addressed. Good taxonomic skills are important, too. In addition, knowledge of the availability of data that will be needed to address some of the issues, such as distribution, trends, or even species locations will be essential.

Initiating on-the-ground management action will then be determined by evaluating inventory data in combination with local priorities that can be site (location) and (or) species driven. If the site and (or) species of focus are identified as priorities, management action is deemed necessary. The decision process that follows will consider the potential actions to be taken to address a particular species on a particular site for a particular time period. The proposed project and site will be reviewed by the National Environmental Policy Act (NEPA) and Environmental Impact interdisciplinary team annually to determine (1) if the project falls under the parameters of relevant integrated pest-management plans and environmental assessments, and (2) if sensitive natural or cultural resources or the human environment could be adversely affected as a result of management actions (or continuing management).

Determining Management Response Options

Once a particular species and (or) site is chosen and management action is deemed necessary, a desired outcome, or management response, must be established. The management response for a particular species on a particular site will be determined by circumstances and practical realities (fig. 10.1). Alternatives include:

1. **Eradication**: reducing the reproductive success of an invasive species or specified populations in largely uninfested regions to zero and eliminating the species or population within a specified period of time. Once all specified populations are eliminated or prevented from reproducing, intensive efforts continue until the existing seed bank is exhausted; may be legally mandated or desirable for a new invader or new site.
2. **Control:** reducing the vigor of invasive plant populations within an infested region, decreasing the propensity of invasive species to spread to surrounding lands, and mitigating the negative effects of populations on infested lands. This strategy inflicts some damage on the target species with the goal of lessening the rate of spread, with the intent to reduce the current infestation at least marginally.
3. **Containment:** maintaining an intensively managed buffer zone that separates infested regions, where containment activities prevail, from largely uninfested regions where eradication activities prevail. Management actions do not usually reduce the current infestation. As better techniques are made available or environmental circumstances render a species more susceptible to control or eradication strategies, areas identified for containment may be upgraded to control or eradication status.

In order to select the most effective management response for a specified invasive-species problem, managers should always start at the beginning of the decisionmaker rubric (fig. 10.1) with the assumption that eradication is a possibility. Whether or not the decisionmaker(s) revert(s) to a control or containment response depends on local information known about the species itself and about the site it occupies. For example, one may automatically assume that eradication or control options for a widespread species are not feasible and that the appropriate response is containment. However, if the specific site in the natural area is a high-value habitat and the target species is present only in small and isolated infestations, then a more appropriate goal may be control (for example, suppression) or even eradication at the particular site, depending on other site considerations.

On-the-ground management requires a review of site considerations and available techniques. Selection of tools and treatment methods will depend on many different biotic and abiotic variables that, if not considered properly, are likely to adversely affect the success of the specific treatment and restoration strategy selected. Examples of factors to consider include distance from roads and trails, proximity to water bodies, weather conditions, soil types, prevailing winds, plant vigor and developmental stage. Full consideration of all the possible variables and scenarios is beyond the scope of this chapter. Rather, it is our goal in this chapter to establish linkages between the early-detection process and subsequent rapid response options within a contiguous framework (see fig. 10.1). Several handbooks, documents, and Web sites offer selective and thorough coverage of appropriate site

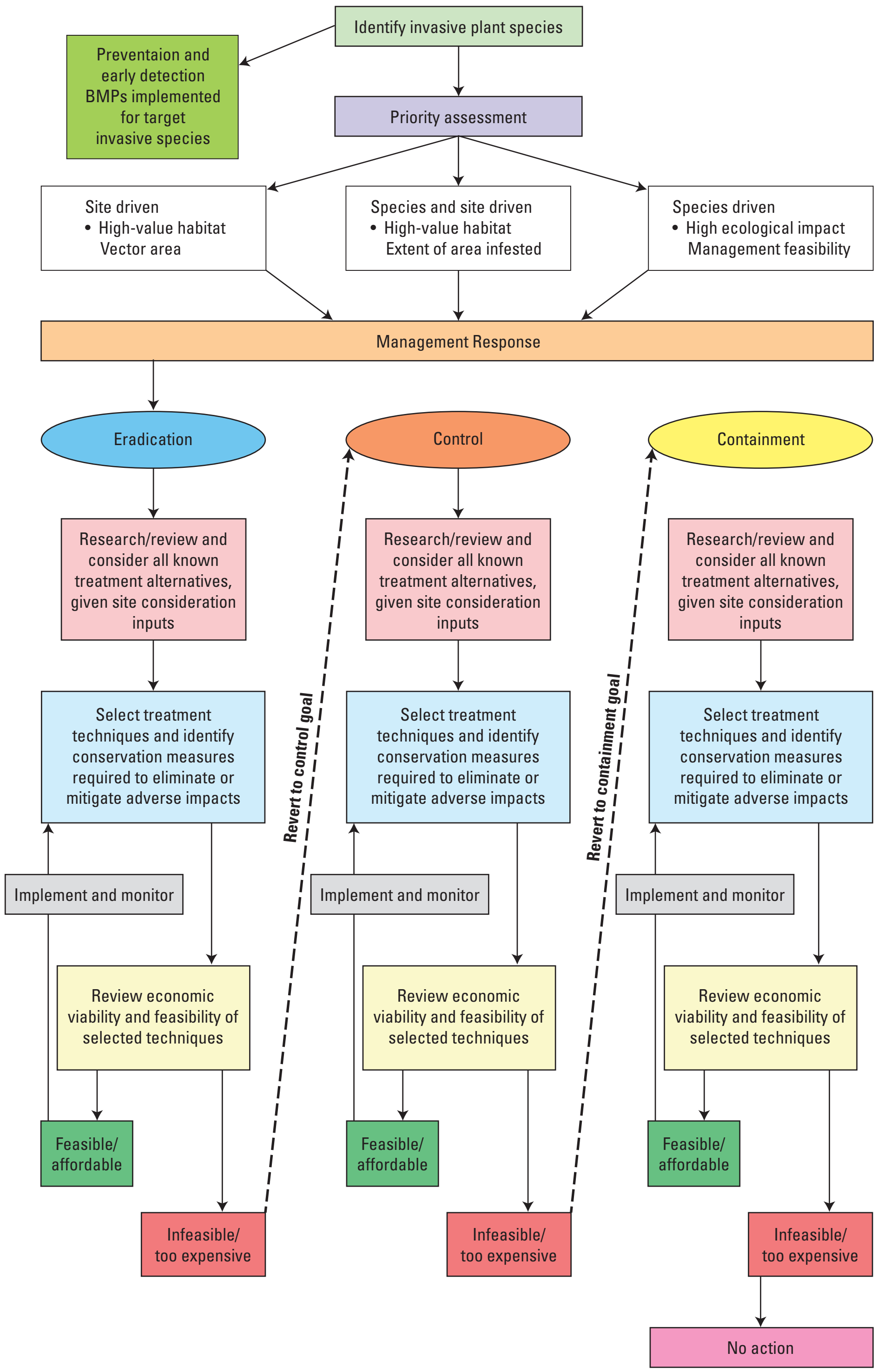

Figure 10.1. Ideal flow of management response strategies for invasive plant species (after Spencer 2005).

considerations and treatment options for general approaches as well as species-specific strategies. Sources include:

- The Nature Conservancy's "Weed Control Methods Handbook" (accessed July 18, 2011, at *http://www.csu.edu/cerc/documents/WeedControlMethodsHandbook.pdf*),
- National Invasive Species Information Center, "Manager's Tool Kit—Control" (accessed July 18, 2011, at *http://www.invasivespeciesinfo.gov/toolkit/control.shtml*),
- "Creating an Integrated Weed Management Plan: A Handbook for Owners and Managers of Lands with Natural Values," Volume IV. (accessed August 4, 2014, at *http://cpw.state.co.us/Documents/CNAP/IntegratedWeedManagement/IWMApp5.pdf#search=Creating%20an%Integrated%20Weed%20Management%20Plan%3A*), and
- Miller, J. H. Exotic Pest Plants and Their Control. Revised February 25, 2002. USDA Forest Service (accessed July 18, 2011, at *http://www.bugwood.org/weeds/forestexotics.html*).

Evaluating Potential Treatment Consequences

Once appropriate treatment techniques and tools are identified, impacts resulting from their use also need to be identified. All tools and techniques will have some type of consequence, whether intentional or unintended, benefical or adverse, direct or indirect. At this point in the decisionmaking process, steps need to be identified to reduce or eliminate any potential adverse (nontarget) effects. These steps can be conservation measures that are practices incorporated into the planning phase of the treatment to prevent potential adverse effects (for example, for threatened species, control treatments should occur pre-emergence or after seed set) or they can be mitigation measures that fix or correct an adverse effect after action has occurred (for example, native trees can be planted after nonnative species are removed in riparian areas).

If the selected treatment techniques and conservation/mitigation measures are affordable, effective, and practical, then the treatment plan is approved for implementation. At a minimum, implementation of any treatment plan will include informal documentation (monitoring) of its effectiveness. More formal monitoring will occur in cases where specific biological or ecological thresholds are identified prior to treatment implementation.

Review Economic Viability and Feasibility of Selected Techniques

If the treatment or conservation/mitigation measures selected are not affordable, effective, and practical, then the treatment plan cannot be approved as it stands and the decisionmaker(s) needs to reconsider other response options (for example, control or containment if eradication was the preferred option), as indicated in figure 10.1.

No Action

There may be cases when all known treatments and conservation/mitigation practices are still not affordable, effective, or practical and a determination of "No Action" must be made. This is not necessarily a decision not to address the problem (a "live with it" decision). Rather, it is an acknowledgement that the problem may need to be monitored further and reevaluated at a later date. More data or new control technologies/strategies may become available or changes in environmental circumstances may improve the feasibility of using available techniques and strategies.

Summary

Before implementing a well-designed early-detection program, data management and reporting procedures should be in place to assist with organizing, analyzing, reporting, and storing the data and information that will be generated by the program. Likewise, due consideration should be given to potential management options that may be used as part of the rapid response process in the event that a new invader is detected. All too frequently search strategies are implemented without careful consideration of what data may need to be collected, how the information will be stored and reported, and how the current search event may affect future early-detection efforts and subsequent management actions. Well-articulated data-management and reporting schemes not only improve internal management efficiency within a park but also foster communication with external stakeholders and partners. If designed properly, data-management and reporting schemes also serve as the conduit for reporting at the regional and national scales. In fact, many resource-management agencies and organizations—including the National Park Service and the U.S. Geological Survey—are beginning to standardize invasive-plant databases to facilitate data sharing and data comparability and to improve report integration at varying scales. Similarly, use of a consistent conceptual approach to management-response options, such as the example presented in fig. 10.1, ensures that current partners and successive managers will understand the basis for decisionmaking now and in the future.

Ultimately, the key to success in combating nonnative species invasions is the development of a long-term strategy that will address immediate needs, protect important park resources, incorporate information needs (autoecological research or predictive modeling efforts), and allow parks to assess whether the strategy used is effective in accomplishing the early-detection objectives identified. In the case of invasive plants, the last criterion is challenging both functionally and financially. The outcomes of plant interactions following management actions often are uncertain, and combined with a changing climate, the outcomes will most certainly become less predictable in the future (Sutherst, 2000). Nevertheless, understanding the outcomes of current management approaches is a critical factor in the adaptive management process. A well-formulated data-management scheme and a transparent framework for consideration of management response strategies are vital steps toward developing that understanding.

Recommended Reading

For thorough coverage of topics related to on-the-ground control and management of invasive plants, including strategies and species-specific techniques, see The National Invasive Species Council's Manager's Tool Kit Web site (accessed July 18, 2011, at *http://www.invasivespeciesinfo.gov/toolkit/main.shtml*).

References Cited

Chrisman, N.R., 1994, Metadata required to determine the fitness of spatial data for use in environmental analysis, *in* Michener, W., Brunt, J., and Stafford, S., eds., Environmental information management and analysis—Ecosystems to global scale: Philadelphia, Taylor & Francis, p. 177–190.

Michener, W.K., and Brunt, J.W., 2000, Ecological data—Design, management and processing: New York, Wiley-Blackwell, 192 p.

Mohren, S.R., 2007, Data management plan, Klamath inventory and monitoring network: Fort Collins, Colo., National Park Service, Natural Resource Report NPS/KLMN/NRR-2007/012.

National Invasive Species Council, 2003, General guidelines for the establishment and evaluation of invasive species early detection and rapid response systems, version 1, accessed December 17, 2010, at *http://www.invasivespeciesinfo.gov/docs/council/GuidelineCommunication.doc.*

Stohlgren T.J., Barnett, D.T., and Simonson, S.E. 2005, Beyond NAWMA—The North American Weed Management Association mapping standards: accessed March 24, 2014, at *http://www.naisma.org/images/BeyondStandards.pdf*.

Sutherst, R.W., 2000, Climate change and invasive species: a conceptual framework *in* Mooney, H.A., and Hobbs, R.J., eds., Invasive species in a changing world: Washington, D.C., Island Press, p. 211–240.

Invasive-Plant Early-Detection Protocol Development in the Klamath Network-National Park Service

By Dennis Odion[1] and Daniel Sarr[2]

Chapter 11 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figures

Chapter 11.
Invasive-Plant Early-Detection Protocol Development in the Klamath Network-National Park Service

By Dennis Odion[1] and Daniel Sarr[2]

Overview

This chapter discusses our experience with the key issues and decisions that land managers may face in developing and implementing invasive species early detection monitoring programs for large and diverse natural areas. Our experience has been from developing such a program for the Klamath Network Inventory and Monitoring (I&M) Program of the National Park Service (*http://science.nature.nps.gov/im/units/klmn/*). The Klamath Network was chosen as a case study because it encompasses large natural areas with a diversity of habitats and invasive species concerns. The Klamath Network comprises six units managed by the National Park Service in northern California and southern Oregon: Crater Lake National Park, Lassen Volcanic National Park, Lava Beds National Monument, Oregon Caves National Monument, Redwood National and State Parks, and Whiskeytown National Recreation Area (fig. 11.1). Parks within the Klamath Network are scattered across a complex mountainous region with corresponding variation in climate, soils, vegetation, land use, and disturbance regimes. Collectively, the parks cover nearly 200,000 hectares and range in size from 196 to 73,775 hectares.

The Inventory and Monitoring Program of the Klamath Network was established in 2000. Since its inception, the Network has placed a high priority on invasive species. Inventories of invasive species were among the first network activities. Nonnative invasive species (hereafter invasive species) ranked as the top vital sign in the network's Vital Signs Monitoring Program (Sarr and others, 2007). In addition, a concern about invasive species was part of the reason that vegetation, land cover, whitebark pine, cave entrance, aquatic, and intertidal communities were selected as vital signs. We recognized that the Network would be monitoring invasives, but would not manage invasive species. In addition, we recognized that the resources devoted to managing invasives would not be under the control of the Inventory and Monitoring Program.

[1]Southern Oregon University, Ashland, Oregon.

[2]National Park Service, Klamath Inventory and Monitoring Program, National Park Service, Ashland, Oregon.

Development of the Invasive Species Protocol

The Klamath Network Inventory and Monitoring budget had to balance the needs of invasive species monitoring with the desire to monitor nine other vital signs. It is beyond the scope of this chapter to discuss how we allocated resources to different vital signs, but this topic is relevant to many land managers who must balance threats from invasives with those from other sources. Our approach to balancing multiple threats is reviewed in Sarr and others (2007). Our allocations yielded a budget of $50,000-70,000 per year for implementation of an invasive species early detection protocol.

The network followed the steps recommended herein for developing an invasive species monitoring protocol outlined in Chapter 3 and figure 3.2. We assessed the scope of our program and park management needs (**Step 1**, see also, Chapter 4) and existing data and expert opinion on invasives to pursue **Steps 2-4.** These are the steps where relevant information is compiled about species to target for monitoring. We used the information to refine the monitoring objectives (**Step 5**). The network has developed and implemented the protocol based on these objectives (Odion and others, 2010). The protocol is also consistent with Steps 6-11 in terms of design, testing, data management and documentation. The complete, approved protocol can be accessed at: *http://irmafiles.nps.gov/reference/holding/376490,* accessed March 24, 2014.

The biggest challenge we faced in developing the protocol was balancing top down (I&M program-led) and bottom-up (park-led) needs. For the I&M program-led efforts, the overarching goal is to monitor ecological integrity. Accordingly, a major priority is to model the vulnerability of different park environments to invasive species threats. In contrast, the parks have more immediate, practical needs to try to eradicate incipient biological invasions. Our approach to reconcile top-down and bottom up needs was a compromise. The protocol has been deemed a management protocol. Yet it qualifies as an I&M monitoring protocol because it contains long-term monitoring and landscape susceptibility modeling elements. Only time will tell whether the Klamath Network has struck the right balance between top down and bottom up concerns.

Figure 11.1. National Park units of the Klamath Network (KLMN) of southern Oregon and northern California.

Protocol Objectives and Rationale

We present herein each of the main objectives articulated in our protocol (Odion and others, 2010) followed by the rationale and information upon which we relied to develop the objective. We emphasize that our budget limited the scope of our objectives.

Objective 1. Detect populations of selected invasive plants throughout the network by sampling along roads, trails, and powerline corridors, and in campgrounds, where introduction is most likely.

Rationale for Selecting Invasive Plants

Managers of any large natural areas may be faced with a plethora of invasive species problems representing multiple taxa groups. This is certainly the case in the Klamath Network. The most significant invasive species problem may be nonnative pathogens like white pine blister rust (*Cronartium ribicola*), Port-Orford cedar root rot (*Phytophthora lateralis*), and Sudden Oak Death (*Phytophthora ramorum*). We were able to justify not including these in an early detection protocol because they are the focus of other ongoing monitoring and research efforts. The chief management priority for parks has been invasive plants.

Rationale for Early Detection Emphasis

Managers will face competing demands on their invasive species management resources. To address this in the Klamath Network, we worked with park resource staff intimately involved in managing invasive species to weigh early detection along with other monitoring objectives. Several of the parks desired monitoring of established invasives along with management effectiveness monitoring, and there were educational and other roles considered for the network's monitoring resources. As many managers can attest, when resources become limited, monitoring will often not be seen as the top priority. Under the I&M Program, we needed to ensure that monitoring took precedence over other services that we were capable of providing.

The Network chose early detection monitoring as a cost-effective means to identify populations for removal, control, or eradication before they become widely established within a park (OTA, 1993, Myers and others, 2000, Harris and others, 2001, Rejmanek and Pitcairn, 2002, Timmins and Braithwaite, 2002). We considered early detection and rapid response to be a second line of defense against invasives that would complement efforts to prevent the transport and spread of invasives to new areas (Westbrooks, 2004). In addition to saving money, early detection and rapid response efforts will minimize ecological damage caused by control efforts, which may become futile or counterproductive if not done early in the invasion process (Rejmanek and Pitcairn, 2002, Odion and others, 2004). Given the limited resources of the Network, early detection was considered an especially pragmatic approach and highly effective when combined with rapid response (Westbrooks, 2004). A more complete discussion of the merits of early detection is presented in Chapter 3.

Rationale for Network Wide Monitoring

For efficiency sake, land managers may wish to target particular ecosystem types that are known to be more susceptible to invasion for invasive species monitoring. For example, across the network, the most striking pattern observed in the invasive plant inventory we conducted was that invasive species declined sharply from low elevations of Whiskeytown to the higher elevations at Lassen Volcanic (fig. 11.2). This pattern has been well-established in other studies in California (Mooney and others, 1986, Rejmanek and Randall, 1994, Schwartz and others, 1996, Keeley and others, 2011).

To help predict potential vulnerability of high elevation and other ecosystems to invasion, we developed a semi-quantitative conceptual model to link species' physiological tolerances with potentially invasible park habitats (Odion and others, 2010). Each invasive species in the network was given a value for habitat preference and elevational limits. A number of sources were used to determine values for habitat

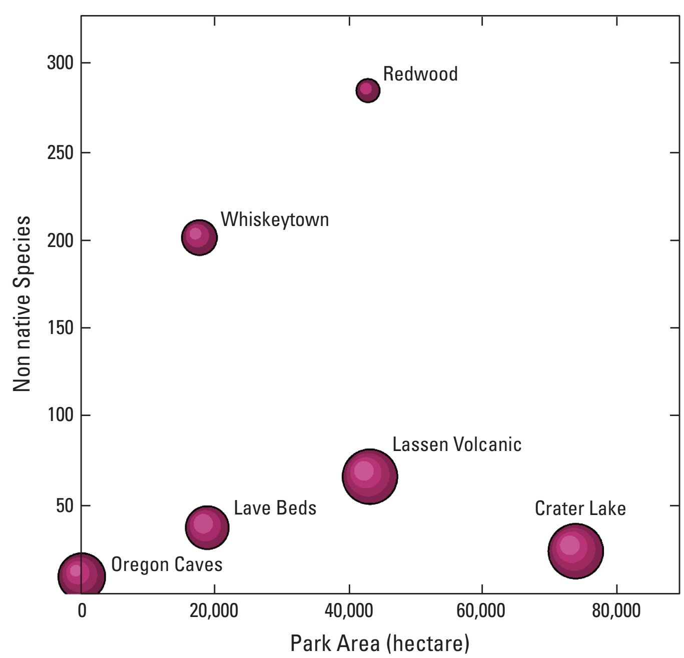

Figure 11.2. Nonnative-species richness as a function of park area and elevation in National Park Service units in the Klamath Network. The lower elevation parks have more nonnative species than expected for their size, whereas higher elevation parks have fewer recorded species. (Accessed from the National Park Service NPSpecies database, February 2007).

preferences and elevation limits, including the Jepson Manual (Hickman, 1993, D'Antonio and others, 2004), online sources such as the California Invasive Plant Council's web page (*http://www.cal-ipc.org/*) and expert opinion of park resource managers. Species were then ordinated by their tolerance values using Principal Components Analysis. The result was a diagram visually illustrating patterns of species' tolerances (fig. 11.3). Species with similar attributes clustered together in ordination space. Tolerance variables were overlaid as vectors.

There were three general groups of species with distinctive combinations of drought, nutrient, shade, and cold tolerance. Given that high elevation environments generally have been the least invaded in the Network, it was surprising that the group characterized by moderate cold tolerance contained the most species (69) and the greatest number of total occurrences of species prioritized for monitoring. These species are not only relatively cold tolerant, but they are also able to tolerate moderate drought and shade. This group of species, which includes Canada thistle (*Cirsium canadensis*), yellow starthistle (*Centaurea solstitialis*), knapweed (*Centaurea maculata*), and dalmation toadflax (*Linaria genistifolia*), was recognized as having particularly broad ecological amplitudes. Mid and upper elevations that show lower levels at this time, may have maintained their ecological integrity due to lack of

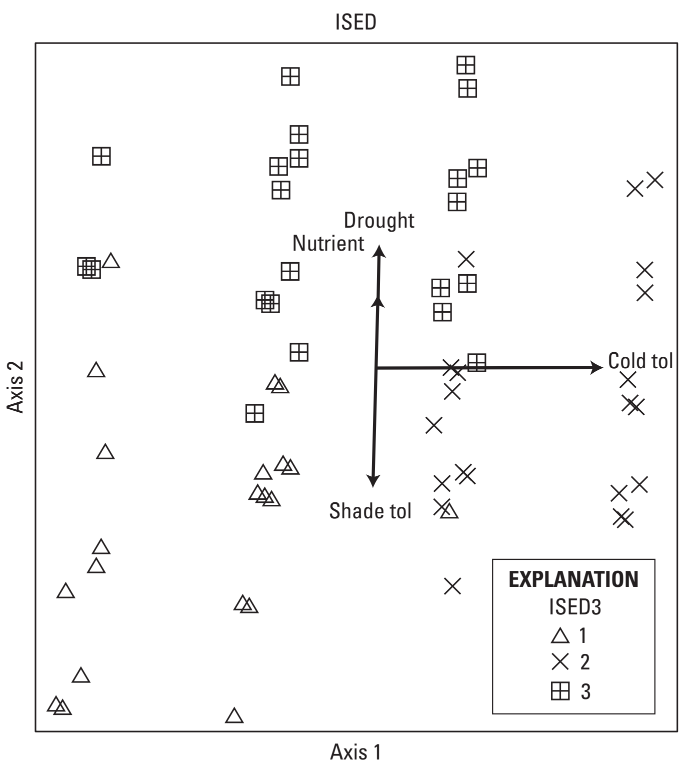

Figure 11.3. Ordination biplot of invasive species in the Klamath Network based on physiological tolerances. The first axis explains 43 percent of the variation in tolerance values among 166 species, while the second axis explains 31 percent. Some labels represent the locations of more than one species in the ordination space. Group 1 are squares, Group 2 are triangles, and group 3, potential mid- to upper elevation species, are Xs.

disturbance or other factors, not lack of potential invaders. We also recognized that climate change and increasing globalization will likely increase future susceptibility of higher elevations to plant invasions (Pauchard and others, 2009).

There were relatively few shade tolerant invasives, which was consistent with general observations that open environments in the parks have been most invaded. Thus, the analysis suggests that the network could have focused monitoring in high light ecosystems. This would have reduced the potential sampling frame considerably, as most of the network is vegetated by closed forests. The limited network monitoring budget also supported targeted monitoring to maximize early detections.

However, we also needed to consider the overriding goal of the network to monitor ecological integrity. Ecological integrity is based on the level to which an ecosystem has been degraded as a result of human activities. We considered the ecological integrity of the habitat invaded to be proportional to the risk of degradation due to an invasive species. In addition, the network had already developed an overriding philosophy for its program: to consider all broad ecosystem types equally valuable to monitor (Sarr and others, 2007). Therefore, we were reluctant to limit monitoring to high light and (or) low elevation environments, even though this would mean more intensively sampling the environments most vulnerable to plant invasions. Preventing invasions at higher elevations or shady old-growth forests could have the effect of keeping these valuable scenic areas almost completely free of invasives, maintaining their currently high ecological integrity. In an analysis of the efficiency of investments in surveillance and targeted monitoring, Wintle and others (2010) found that surveillance monitoring was justified if the expected benefits are substantially higher than those arising from a well-planned targeted design. We believe this to be true because of the fundamental NPS goal of maintaining ecological integrity.

Ultimately, we chose to monitor all parks and widespread habitats in the network. However, we included more species in the monitoring in wilderness and other remote areas with higher ecological integrity (see species prioritization objective below). Efficiency was maximized by targeting roads and trails for monitoring.

Rationale for Targeted Monitoring Along Roads and Trails

Our method employs elements of surveillance and targeted monitoring to strike a balance between broad coverage of park ecosystems and a focus on the most likely spots of invasion. Surveillance monitoring has been criticized because its benefits are harder to identify *a priori* (Wintle and others, 2010). In contrast, targeted monitoring is hypothesis-driven, with concrete relevance to managers. Moreover, for efficiency sake, land managers may wish to target particular disturbances or heavily used areas for invasive species monitoring of environments that are known to be more vulnerable to invasion.

For example, burns are well-known hot-spots for invasion (Brooks and others, 2004, Klinger and others, 2006, Mack and D'Antonio, 1998). Thus, we carefully considered monitoring burn areas. We did include burned areas in our sampling frame, but they are only included when traversed by roads and trails. The NPS has a large fire program that includes its own monitoring component, so we felt that fire effects on invasives would be covered. Monitoring burns also creates a practical difficulty because large areas would need to be sampled after episodic burn events. Our advice to managers is to try and monitor burns by obtaining outside funding when large fires occur. There is often funding for post-fire management of natural areas.

There are numerous studies of vegetation that have documented a strong local association between roads and trails and invasive species (Trombulak and Frissell, 2000, Douglas and Matlack, 2006). The occurrence of invasive plants has been found to predictably decline with distance from roads and trails (Reed and others, 1996, Greenberg and others, 1997, Parendes and Jones 2000, Silveri and others, 2001, Watkins and others, 2003). The increased abundance of invasives alongside roads has been related to road surface materials (Greenberg and others, 1997, Silveri and others, 2001), light (Parendes and Jones, 2000), and higher frequency of disturbance (for example, Parendes and Jones 2000). Gelbard and Belknap (2003) found much greater numbers of invasives along paved road verges than on 4-wheel drive tracks. Other utility corridors along with fuel breaks have many of the same features as roads (disturbance and propagule pressure), and they share similar invasion susceptibility (Merriam and others, 2006).

The ordination of invasive species habitat preferences (fig. 11.3) illustrated the preponderance of possible invaders that are intolerant of shade, and road and trail environments get more sunlight. They may be the only high light environments in many forested areas. The Network's invasive species inventory also supported a road and trail monitoring focus. At low elevations, invasives declined dramatically with distance from the road or trail. Lastly, the network has a vegetation protocol (Odion and others, 2011) that will provide a modest level of broader surveillance monitoring. Therefore, we focused our monitoring along roads and trails.

Objective 2: Provide the information to park management on a timely basis to allow effective management responses.

Rationale for Rapid Response Approach

The goal of invasive species monitoring for most managers is to provide information for invasive control efforts. A key element of early detection monitoring is to support rapid management response (Chapter 3). We explicitly recognized the need to link monitoring and rapid response (fig. 11.4), and designed a reporting scheme using briefings to quickly communicate the most urgent findings to park managers. In

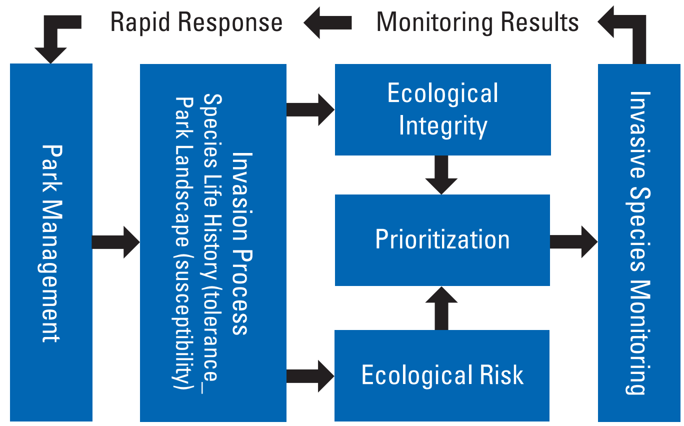

Figure 11.4. Conceptual model of the Klamath Network's invasive species early detection monitoring protocol. Park management and invasive control efforts affect the invasion process. This process places differential ecological risks across the park landscape, affecting ecological integrity. These effects determine the prioritization of species and locations to sample in the invasive species monitoring protocol. The results of this monitoring feed directly into rapid response, a key component of park management of invasives.

addition to formal reports, field crews will meet with park resource staff upon completing their seasonal field work. The purpose of these meetings will be to convey the most urgent findings verbally so that park managers can schedule more immediate treatments, if appropriate and feasible. GIS layers showing data collected during each year will also be provided to the parks no later than December 1 in the year of sampling.

Objective 3: Develop and maintain a list of priority invasive plant species with greatest potential for spread and impact to park resources for monitoring in each park (Step 6).

Rationale for Species Prioritization

How does a land manager faced with so many invasives and a limited monitoring budget determine which species to monitor? Resources are likely to be too limited to monitor and control all species. Moreover, most invasive species do not become major ecological problems and trying to control many invasives that do cause ecological problems causes more damage than it reduces (Davis and others, 2011). With these concerns in mind, we aimed to focus monitoring efforts on species that had the potential to transform entire ecosystems. In addition, we sought to emphasize potential ecological transformers that were not so well-established that they are beyond control (Rejmanek and Pitcairn, 2002). We further sought an informed and systematic prioritization process to identify ecosystem transformers that can still be controlled. This species prioritization is described in Chapter 5. Robert Klinger and Matt Brooks (USGS) undertook the prioritization process at the park level. For each

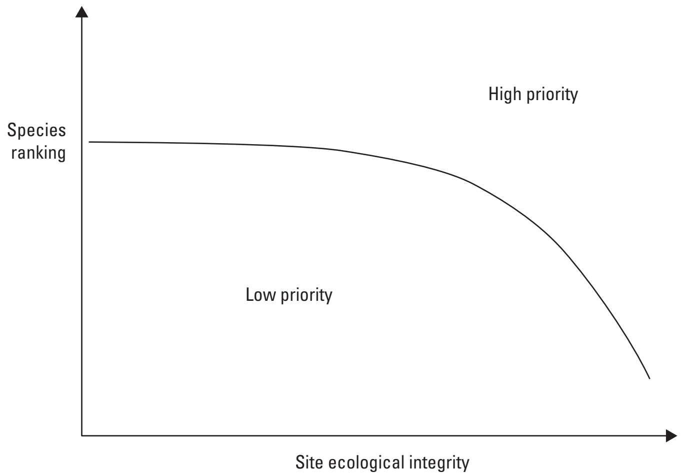

Figure 11.5. Conceptual model of how invasive species priorities may change as a function of the ecological integrity of sites where they are found.

park, a list of invasive species present, and species that could invade, which we determined based on the literature, was first developed. Then these species were classified into the stage of invasion they had reached in each park: colonization, spread, or equilibrium, or they were considered to not be an invasive that transforms ecosystems. This classification was based on all available published information on a species' ecology, expert opinion of park staff, and invasiveness rankings by California Invasive Plant Council. Only species that were unanimously considered non-threats by all managers and scientists were excluded from the ranking process. If there was any question, the species was included. Depending on the park and species' ranking score, most or all of the colonization and spread species (that is, ecosystem transformers that were not infeasible to control) were selected as the priority species to monitor throughout a park.

However, we recognized that it is difficult to apply a uniform species prioritization across heterogeneous landscapes. We considered the general relation between ecological integrity and the ranking of species via prioritization conceptually as shown in figure 11.5. We incorporated the concepts shown in this figure by adding equilibrium species to the monitoring in areas of high ecological integrity where these species are currently not found. Areas of high ecological integrity included remote and wilderness areas, as determined by park resource specialists.

Objective 4: Use monitoring data collected from this protocol and the vegetation protocol to estimate possible trends and develop and refine models of invasive species habitat requirements and of the most susceptible habitats (both along roads and trails and not).

Objective 5: Sample plots in infested and uninfested areas in an unbiased manner to provide data for species habitat modeling.

Rationale for Collecting Data for Landscape Susceptibility Modeling

Understanding the vulnerability of different environments to invasive species can be particularly valuable in developing management plans for controlling invasives. As monitoring data are collected, the conceptual understanding of the relations between invasive species and the environments in which they are found can be elucidated using spatial modeling (Chapters 6 and 7 discuss the purposes and value of such models).

We recognized the value of such modeling and built in appropriate data collection to support it. Figure 11.6 shows one type of output from predictive susceptibility modeling from data collected at Redwood using our protocol. Models such as this provide a conceptual basis for predicting beyond the current known range of invasives to new areas in the parks. Our modeling in the future will also use data from the vegetation monitoring protocol (Odion and others, 2011), so that predictive modeling is not as strictly limited to road and trailside environments.

Objective 7: Sample road and trail segments (generally 3 km) in each park, as many as possible, using a probabilistic sampling design to maximize detection of priority species.

Rationale for Probabilistically Sampling Along 3 Kilometer Road and Trail Segments

When subsampling a population, managers need to use probabilistic sampling designs to maximize detection and meet assumptions of inferential statistics. Thus, since we were interested in making park-wide inference, we randomly selected a subset from all road and segments in each park. In subsequent monitoring years, we could revisit these segments, visit only new segments, or select segments to monitor at random each monitoring season. We chose the latter option as most statistically defensible. However, this means the revisit frequency for any one segment will vary.

With sampling frames developed for each park, we allocated sampling effort to allow a statistical sample of road, trail, or powerline segments in all the parks to meet our objectives. At Oregon Caves, a complete survey was feasible, but the larger parks in the Network have between 257 and 1,448 km (160 and 900 miles) of roads, trails, or powerline segments. We set a target sample size of 25 segments in each park, except Redwood, which has an extensive road and trail network, received 35, and Oregon Caves, which received 10 (because this comprised the whole road and trail network). This yielded a seasonal sampling target of 145 segments. Assuming a 5.5 month field season, the crew had to complete just over 2 segments per day through the field season to meet the target sample across all six parks. We found through pilot studies that 3 km segments could be monitored in half of a 10 hour day, on average, by a two person crew. Figure 11.7 shows an example of the sampling frame from Lava Beds.

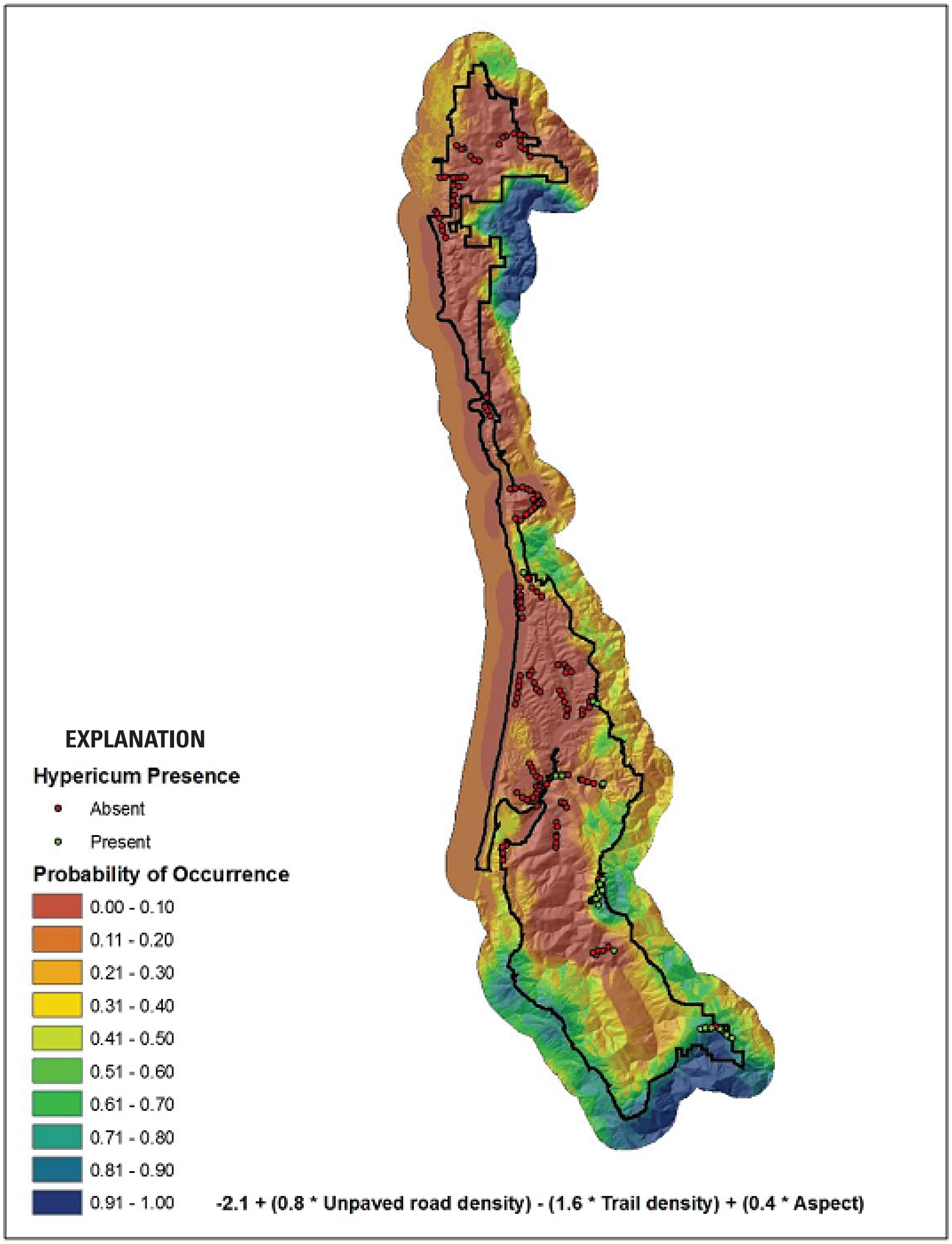

Figure 11.6. Interpolated surface showing the probability of occurrence of Klamath weed (*Hypericum perforatum*) at Redwood National Park. Data used in the modeling were collected during a pilot study at Redwood [(see Appendix A, Odion and others (2010)].

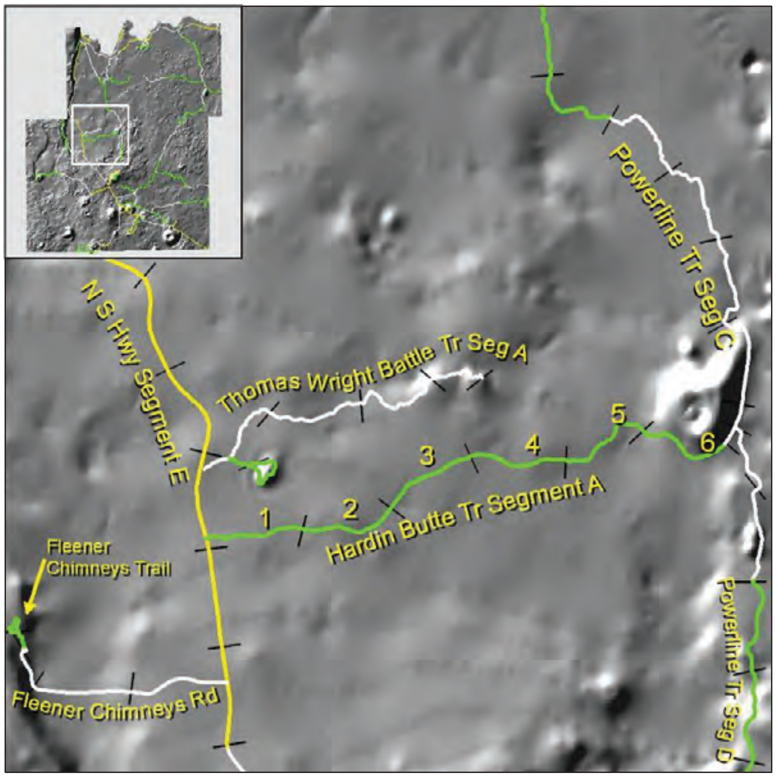

Figure 11.7. Close-up showing sampling frame at Lava Beds National Monument, 3 kilometer segments and 500 meter subsegments (black lines). Hardin Butte Trail's 500 meter subsegments at Lava Beds are labeled 1–6.

To make the sampling fit into half a day and ensure adequate sample sizes, the field work had to be efficient. We found that field crews could traverse the selected segment and record, using a TrimbleGeoXM handheld GPS, a geographic coordinate of all prioritized species visible from the road or trail, and the size of each infestation (<1 m^2, 1–25 m^2, or >25 m^2) (fig. 11.8). We chose a maximum number of four individual infestations to be recorded per segment. The logic was that more than 4 infestations would require that managers carry out substantial control efforts on the segment, and identifying more infestations would not be productive. For populations that would not be discreet enough to map as separate patches, we developed an option for recording continuous infestations along a segment. We also chose to estimate each infestation size rather than map its perimeter in the field. This was done to prevent extensive, time-consuming GPS mapping of infestations.

Upon reaching the end of the segment, the crew will return along the same route, recording environmental data at each infestation, and at 6 random points off the road or trail. This approach was designed to provide quantitative environmental information at places where infestations are both present and absent for landscape vulnerability modeling.

Lessons Learned

In 2009 and 2011, the Klamath Network hired two person crews to visit each of the six parks and implement the invasive species early detection protocol (Odion and others, 2010). In both years, cool, wet weather delayed the onset for sampling from late April into May. Nonetheless, the protocol proved efficient and informative, and the crews were able to complete 170 segments in 2009 and 146 segments in 2011. However, the roving crew format creates a heavy workload for the I&M Program, with frequent travel and camping logistics for the crew, and a need to hit the parks at precisely the right times.

Well-trained park-based staff could implement the protocol instead of or in combination with a roving network crew. The availability of park-based staff however, varies widely by park, making a single approach difficult. Nonetheless, we are exploring the possibility of supporting park staff-based efforts at implementing the protocol with the broader I&M Program chiefly responsible for training data management, analysis and reporting across parks. Having the monitoring conducted by park staff may also improve rapid response to monitoring.

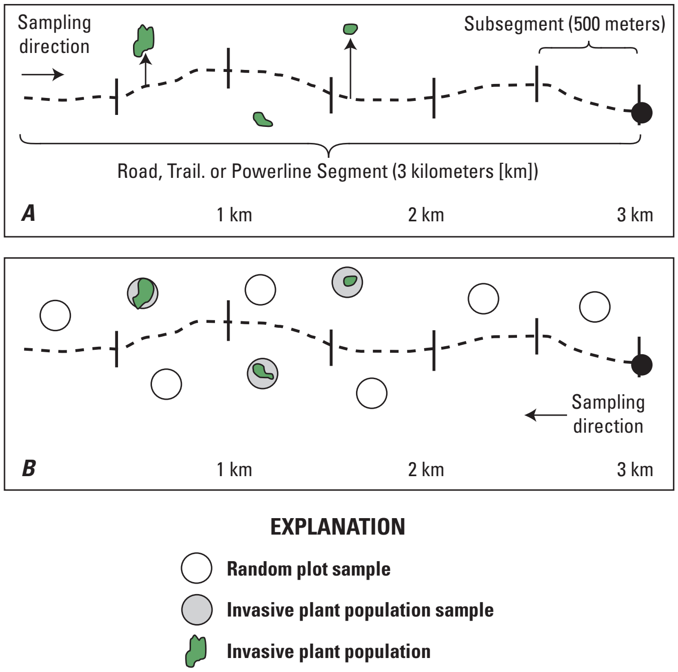

Figure 11.8. Invasive species early detection response design to be completed at each randomly selected road, trail, or powerline segment in a park: *A,* location mapping and sampling of invasive plant populations; and *B,* plot sampling of random locations and the invasive plant populations located.

The link between monitoring and rapid response needs to be strengthened. Currently, managers in Klamath Network parks attempt to respond to invasive species threats to the extent possible given existing resources. Three of the California parks, Lassen Volcanic, Redwood, and Whiskeytown, have historically had Exotic Plant Management Team (EPMT) support from the NPS, which can help with rapid response needs. However, the EPMT crews in California are stretched thin trying to provide for numerous parks, so their ability to provide rapid response is limited. Ultimately, an early detection monitoring protocol, however well designed, will fail to support park management goals unless specific arrangements are made to fully integrate the scientific findings with management actions on the ground. A complete vision of early detection and rapid response will require that additional fiscal and staffing resources are made available to support rapid response. Fortunately, there is regional support for more network level collaboration on exotic plant control within the parks.

An additional challenge that we faced during initial implementation was that park resource specialists wanted to increase the species we monitor. On one hand, we could agree to accommodate all such requests. On the other hand, this creates significant logistical problems and does not respect the rigor that went into the invasive species prioritization process. To date, we have chosen to accommodate requests to monitor more species. If this bogs us down in the future, we will have to revisit this decision and come up with additional standards to apply for adding species. Increasing the involvement of park staff in the monitoring protocol could help reduce this problem in the future.

A final challenge that arose during initial implementation is that park management efforts and monitoring were not coordinated. The result was that some segments were monitored right after invasive species control efforts had occurred. This meant that there were potentially sites that were infested that were considered uninfested, which will affect the accuracy of models of landscape vulnerability to invasion. The difficulty is that both network and park staff must change schedules with short notice due to unforeseen circumstances. This further highlights the potential value of having the monitoring be conducted by park staff.

In conclusion, there are a number of challenges in developing an invasive species monitoring program, especially one that satisfies both long-term scientific needs and more immediate park needs. Only time will tell whether the Klamath Network has struck the right balance. Having capacity to adapt through learning will be key to ensuring that we do. Thus, one additional objective that we identified was: to adapt spatial sampling as knowledge improves through monitoring.

References Cited

Brooks, M.L., D’Antonio, C.M., Richardson, D.M., Grace, J.B., Keeley, J.E., DiTomaso, J.M., Hobbs, R.J., Pellant, M., and Pyke, D., 2004, Effects of invasive alien plants on fire regimes: BioScience, v. 54, p. 667–688.

D’Antonio, C.M., Berlow, E., and Haubensak, K.A., 2004, Invasive exotic plant species in Sierra Nevada ecosystems,0 *in* Stine, P., Graber, D., and Murphy, D., eds., Proceedings from the second Sierra Nevada Science Symposium: U.S. Department of Agriculture, Forest Service, Gen. Tech. Rep. PSW–GTR, v. 193, p. 175–184.

Davis, M.A., and others, 2011, Don’t judge species on their origins: Nature, v. 474, p. 153–154.

Douglas, C., and Matlack, G., 2006, The role of roadsides in plant invasions—A demographic approach: Conservation Biology, v. 20, p. 385–391.

Gelbard, J.L., and Belknap, J., 2003, Roads as conduits for exotic plant invasions in a semiarid landscape: Conservation Biology, v. 17, p. 420–432.

Greenberg C.H., Crownover, S.H., and Gordon, D.R., 1997, Roadside soil—A corridor for invasion in xeric scrub by nonindigenous plants: Natural Areas Journal, v. 17, p. 99–109.

Harris, S., Brown, J., and Timmins, S., 2001, Weed surveillance—How often to search: Wellington, New Zealand Department of Conservation, Science for Conservation, p. 175.

Hickman, J. eds., 1993, The Jepson manual: Berkeley, University of California Press, 1,400 p.

Keeley, J.E., Franklin, J., and D'Antonio, C.M., 2011, Fire and invasive plants on California landscapes *in* McKenzie, D., Miller, C., and Falk, D.A., eds., The Landscape Ecology of Fire: Ecological Studies, v. 213, p. 193.

Klinger, R.C., Brooks, M.L., and Randall, J.M., 2006, Fire and invasive plant species, *in* Sugihara, N.G., van Wagtendonk, J.W., Shaffer, K.E., Fites-Kaufman, J., and Thode, A.E., ed., Fire in California's ecosystems: Berkeley, University of California Press, p. 499–519.

Mack, M.C., and D'Antonio, C.M., 1998, Impacts of biological invasions on disturbance regimes: Trends in Ecology and Evolution, v. 13, p. 195–198.

Merriam, K.E., Keeley, J.E., and Beyers, J.L., 2006, Fuel breaks affect nonnative species abundance in California plant communities: Ecological Applications, v. 16, p. 515–527.

Mooney, H.A., Hamburg, S.P., and Drake, J.A., 1986. The invasions of plants and animals into California, *in* Mooney, H.A., and Drake, J.A., eds., Ecology of biological invasions of North America and Hawaii: New York: Springer, p. 250–327.

Myers, J.H., Simberloff, D., Kuris, A.M., and Carey, J.R., 2000, Eradication revisited—Dealing with exotic species: Trends in Ecology and Evolution, v. 15, p. 316–320.

Odion, D.C., Alexander, J., and Swezy, M., 2004, Use of short rotation burning to combat non-natives and their seed banks in California north coastal prairie, *in* Sugihara, N.G, Morales, M.E., and Morales, T.J., eds., Proceedings of the Symposium—Emerging Policies and New Paradigms—Fire Management: Berkeley, Calif., Association for Fire Ecology, Miscellaneous Publication No. 2, p. 46–57.

Odion, D.C., Sarr, D.A., Mohren, S.R., and Klinger, R.C., 2010, Invasive species early detection monitoring protocol for Klamath Network parks: Fort Collins, Colo., National Park Service, Natural Resource Report NPS/KLMN/NRR–2010/227.

Odion, D.C., Sarr, D.A., Mohren, S.R., and Smith, S.B., 2011, Monitoring vegetation composition, structure and function in the parks of the Klamath Network Parks: Fort Collins, Colo., National Park Service, Natural Resource Report NPS/KLMN/NRR–2011/401, 183 p.

Office of Technology Assessment, U.S. Congress (OTA), 1993, Harmful non-indigenous species in the United States: Washington, D.C., Government Printing Office, OTA–F–565, 391 p.

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

Pauchard, A., and others, 2009, Ain't no mountain high enough—Plant invasions reaching new elevations: Frontiers in Ecology and Environment, v. 7, p. 479–486.

Reed, R.A., Johnson-Barnard, J., and Baker, W.L., 1996, Contribution of roads to forest fragmentation in the Rocky Mountains: Conservation Biology, v. 10, p. 1098–1106.

Rejmanek, M., and Randall, J.M., 1994. Invasive alien plants in California—1993 summary and comparison with other areas in North America: Madroño, v. 41, p. 161–177.

Rejmanek, M., and Pitcairn, M.J., 2002, When is eradication of exotic pest plants a realistic goal?, *in* Veitch, C.R., and Clout, M.N., eds., Turning the tide—The eradication of invasive species: Gland, Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, Species Survival Commission, Invasive Species Specialist Group, International Union for Conservation of Nature, p. 249–253.

Sarr, D.A., Odion, D.C., Mohren, S.R., Perry, E.E., Hoffman, R.L., Bridy, L.K., and Merton, A.A., 2007, Vital signs monitoring plan for the Klamath Network—Phase III report: Fort Collins, Colo., National Park Service, Natural Resource Technical Report NPS/KLMN/NRR–2007/016, accessed August 8, 2014, at *http://irmafiles.nps.gov/reference/holding/155272*.

Schwartz, M.W., Porter, D.J., Randall, J.M., and Lyons, K.E., 1996. Impact of nonindigenous plants, *in* University of California; SNEP Science Team and Special Consultants, Sierra Nevada Ecosystem Project—Final report to Congress, Volume II, assessments and scientific basis for management options: Davis, University of Calif., Centers for Water and Wildland Resources, p. 1203–1218.

Silveri, A., Dunwiddie, P.W., and Michaels, H.J., 2001, Logging and edaphic factors in the invasion of an Asian woody vine in a mesic North American forest: Biological Invasions, v. 3, p. 379–389.

Timmins, S.M., and Braithwaite, H., 2002, Early detection of invasive weeds on islands, *in* Veitch, C.R., and Clout, M.N., eds., Turning the tide—the eradication of invasive weeds: Gland, Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, Species Survival Commission, Invasive Species Specialist Group, International Union for Conservation of Nature, p. 311–318.

Trombulak, S.C., and Frissell, C.A., 2000, Review of ecological effects of roads on terrestrial and aquatic communities: Conservation Biology, v. 14, p.18–30.

U.S. Geological Survey, 2014, Klamath network: U.S. Geological Survey digital data, The National Map, available at: *http://viewer.nationalmap.gov/viewer/*.

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

Westbrooks, R.G., 2004, New approaches for early detection and rapid response to invasive plants in the United States: Weed Technology, v. 18, p. 1468–1471.

Wintle, B.A., Runge, M.C., and Bekessy, S.A., 2010, Allocating monitoring effort in the face on unknown unknowns: Ecology Letters, v. 13, p. 1325–1331.

Appendix A. Klamath Network Invasive Species Early Detection Briefing for Lassen National Park

Invasive Species Early Detection Monitoring: End of Season Bulletin—Fiscal Year 2009

Prepared by Sean B. Smith,[1] Daniel A. Sarr,[1] and Dennis C. Odion[2]

Lassen Volcanic National Park

Fiscal Year 2009 Accomplishments

The Klamath I&M program implemented the first season of its Invasive Species Early Detection Protocol from April to September 2009. During the season, a two person crew led by Sean Smith visited all six parks in the Klamath Network, beginning the season in Whiskeytown NRA and concluding in Redwood NSP. The crew visited 170 road and trail segments for a total of 395 km surveyed. The sample effort matched or exceeded expectations, which was particularly heartening for the first full season of implementation. This represents the first quantitative sample of all six parks in the Network using a repeatable, peer-reviewed methodology at comparable intensities. Data are under analysis and a full report, including shapefiles of invasive species locations for all six parks, is in preparation.

Park-Specific Findings

Lassen was sampled between July 10 and July16, and then again between July 24 and July 28, the height of the flowering season for most invasive species. Thirty three segments or 82.3 road and trail kilometers (fig. 11.9) were sampled. The Lassen effort recorded none of the 37 prioritized early detection invasive species across the park.

[1]National Park Service, Klamath Inventory and Monitoring Program, National Park Service, Ashland, Oregon..

[2]Southern Oregon University, Ashland, Oregon.

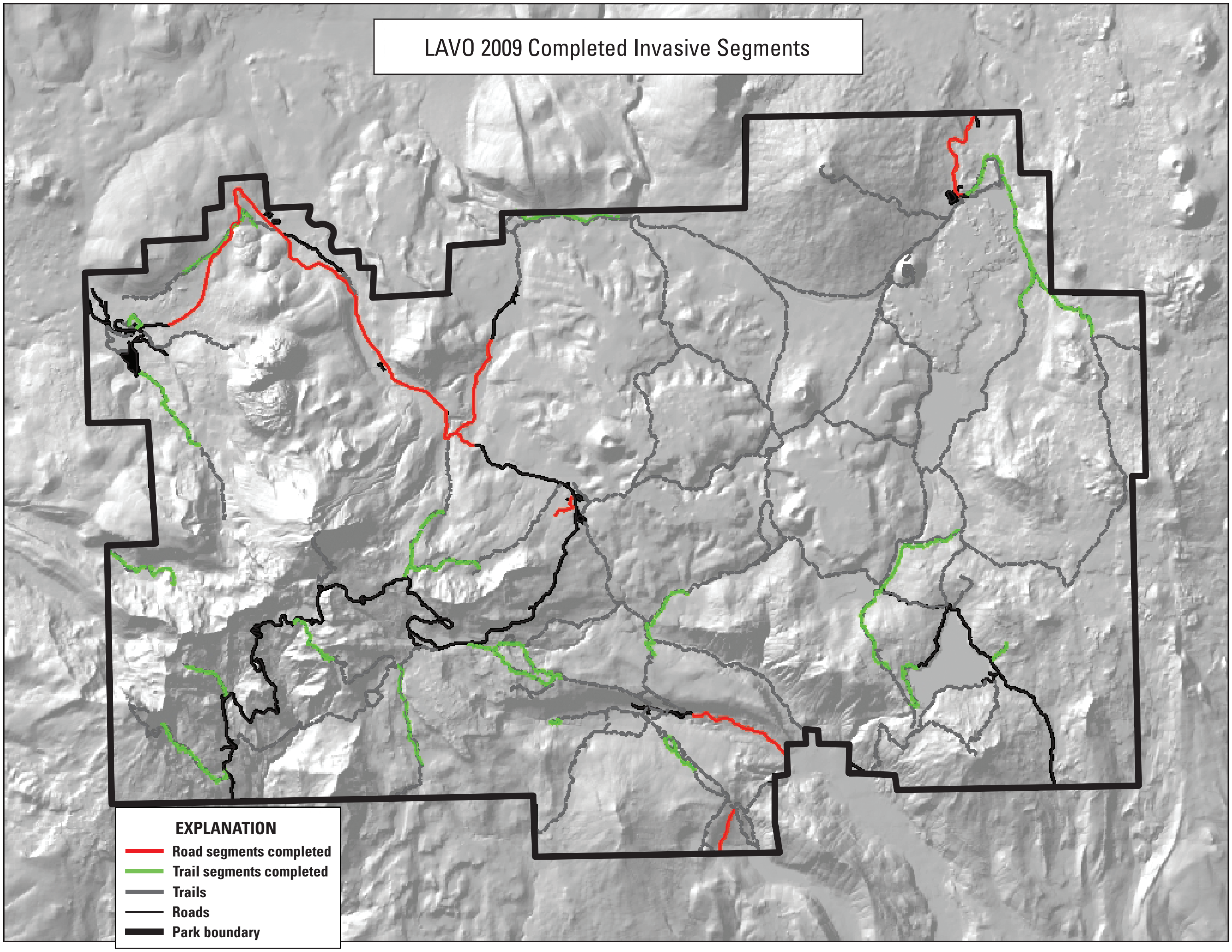

Figure 11.9. Locations of invasive plant species recorded in FY 2009 Invasive Species Early Detection monitoring. Note that not all road or trail segments are sampled each year.

Spatial Distribution and Risk Assessment of Johnsongrass (*Sorghum halepense*) in Big Bend National Park, Texas

By Kendal E. Young[1] and T. Scott Schrader[2]

Chapter 12 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figures

Chapter 12.
Spatial Distribution and Risk Assessment of Johnsongrass (*Sorghum halepense*) in Big Bend National Park, Texas

By Kendal E. Young[1] and T. Scott Schrader[2]

Overview

We present an example of the utility of using remotely sensed and GIS data to model potential habitat for johnsongrass *(Sorghum halepense*, fig. 12.1) in Big Bend National Park (BIBE).We provide a brief species profile, indicate attributes used to model potential distributions, and demonstrate how potential distribution models and vector and pathway models can provide early detection tools to prioritize conservation efforts.

Figure 12.1. Johnsongrass (*Sorghum halepense*) in Big Bend National Park, Texas.

Origins and History

Johnsongrass is native to the Mediterranean region of Europe and Africa (Holm and others, 1977; McWhorter, 1989). Johnsongrass was apparently first introduced into South Carolina around 1830 for livestock forage, but it rapidly spread across the Southern United States (Tellman, 1998; Howard, 2004). Currently, Johnsongrass is fairly widespread through the contiguous United States (Great Plains Flora Association, 1986; Wunderlin, 1998).

Description

Johnsongrass is a tall (heights may reach 3.7 meters or 12 feet) warm-season perennial (Anderson, 1961; Radford and others, 1968; Martin and Hutchins, 1980; Diggs and others, 1999). Leaves have a prominent midvein. Johnsongrass flowers from May to November in the Southwest (Martin and Hutchins, 1980; Diggs and others, 1999). Inflorescence ranges from 10 to 60 centimeters (4–24 inches) with an open panicle. Seeds are approximately 2 millimeters in length (Radford and others, 1968) and may have twisted awns that aid in seed dispersal (Diggs and others, 1999). The leaves of Johnsongrass respond distinctly to solar radiation and, therefore, can be distinguished easily from other plants by remote sensing (McWhorter, 1989).

[1] U.S. Forest Service, Hathaway Pines, California.

[2] Agricultural Research Service, Jornada Experimental Range, Las Cruces New Mexico.

Vectors and Pathways

Characteristics of Johnsongrass (fig. 12.2) that aid in its spread include:

- formation of dense rhizomes that host meristematic tissue responsible for regenerating plants (Anderson and others, 1960),
- moderate drought resistance (Anderson and others, 1960),
- salt tolerance (Yang and others, 1990),
- abundant seed production with seeds that remain viable for 2 to 5 years prior to germination (Leguizamon, 1986; Huang and Hsiao, 1987; Allen, 1990; Unger and others, 1999), and
- possible production of toxins that are allelopathic (Warwick and Black, 1983).

Figure 12.2. Large Johnsongrass patch in a riparian area in Big Bend National Park, Texas.

Wind, water, machinery, and animals disperse Johnsongrass seed (Ghersaand others, 1993; Hartzler and others, 1991). Johnsongrass seed has been carried up to 1.0 kilometer (0.62 miles) from parent plants by winds (around 31miles/hour) that occurred during thunderstorms (Ghersa and others, 1993). Seeds are dispersed along waterways by flowing water. Farming equipment also spreads seeds (Ghersa and others, 1993). Johnsongrass seed is a contaminant in hay and commercial seed (Allen, 1990).

Habitats

Johnsongrass is associated with a variety of habitats but is most common in ecosystems at elevations below 1,800 meters (6,000 feet) with moist to mesic moisture regimes, especially riparian habitats (Howard, 2004). This plant is associated with open habitats and does not persist under closed canopies. Johnsongrass can be found in irrigation canals, flood plains, springs, and stock tanks (Bendixen, 1988; Monaghan, 1979). Johnsongrass grows in various sized patches throughout BIBE in depressions, ditches, and waterways that have had historical disturbances.

Predictive Habitat Model

We used Landsat 7 ETM+ imagery to illustrate how remotely sensed data can model predicted Johnsongrass habitat. We used spectral reflectance values for three seasons of data across 5 years (fall 1999, summer and fall 2000, spring and fall 2001, spring 2002, and spring 2003) to capture Johnsongrass vegetation phenology. Johnsongrass occurrences in BIBE consisted of 147 georeferenced localities.

We used the program Maxent (*http://www.cs.princeton.edu/~schapire/maxent/*; Phillips and others, 2006) to predict Johnsongrass distributions. Maxent is a general approach for modeling species distributions using presence-only data sets (see Chapter 7 for general discussion of modeling data options). Maxent estimates a target probability distribution by finding the probability distribution of maximum entropy (the distribution that is most spread out, or closest to uniform). Maxent uses pixels with known species occurrence records to constitute the sample points. Remotely sensed and GIS data sets can provide environmental variables measured at each sample point (see Chapter 6). Analysis output includes a probabilistic interpretation, grading from least to most suitable habitat conditions.

We evaluated the Johnsongrass predicted habitat model in Maxent by withholding 10 percent of the occurrence locations for testing. Maxent evaluates model performance by testing if the model performed significantly better than random (Phillips and others, 2006).This approach, considered threshold-dependent, used a binomial test (Wilcoxon signed-rank test) based on omission and predicted area. Model performance is evaluated using extrinsic omission rate (fraction of the test localities that fall into pixels not predicted as suitable) and the proportion of all the pixels that are predicted as suitable habitat.

The second approach to evaluating model performance is considered a threshold-independent procedure and uses receiver operating characteristic (ROC) curves (Phillips and others, 2006).The advantage of ROC analysis is that area under the ROC curve (AUC) provides a single measure of model performance, independent of any particular choice of threshold. The AUC can be interpreted as the likelihood that habitat quality is correctly classified by the predictive model at randomly selected sites (Phillips and others, 2006). Chapter 9 provides additional discussion on threshold-dependent and threshold-independent metrics as well as ROC interpretations.

The predictive habitat model indicated that BIBE hosts approximately 14,137 hectares of habitat highly suitable for Johnsongrass. Large patches of potential Johnsongrass habitat exist in the northern part of the park and in major drainages throughout BIBE (fig. 12.3). Threshold-dependent evaluation using the "equal test for sensitivity and specificity" with a cumulative threshold of 23.1 indicated a 14 percent error of omission (P< 0.01) (fig. 12.4). ROC curves indicate that both the training and test data performed better than a random prediction (fig. 12.4). The AUC for test data was 0.92, standard deviation = 0.029, and 0.97 for training data, indicating that the likelihood that a random positive Johnsongrass occurrence and a random negative location were accurately predicted to 92 percent.

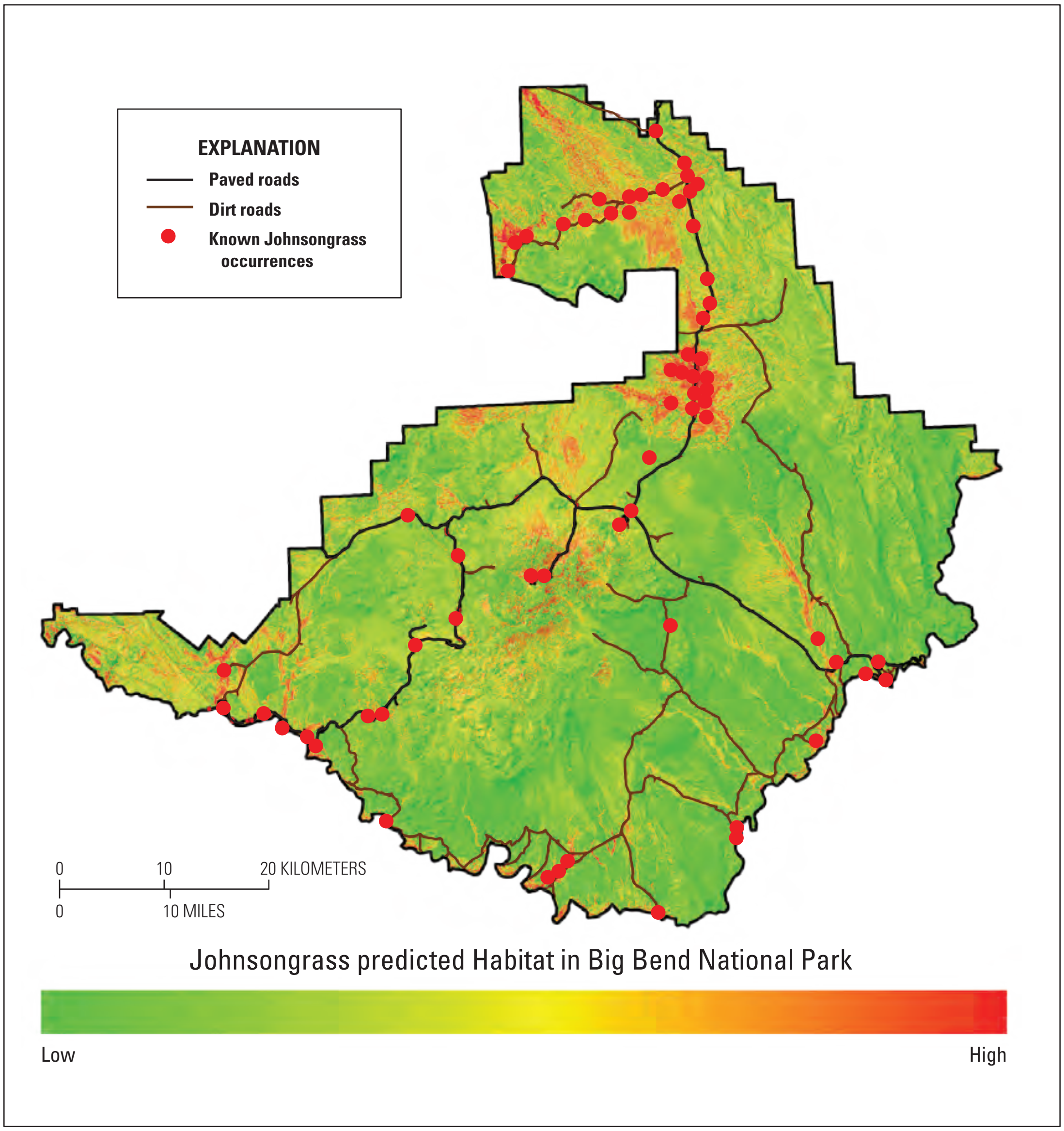

Figure 12.3. Johnsongrass predicted habitat model constructed from remotely sensed data. Predictions were generated using Maxent software (Phillips and others, 2006).

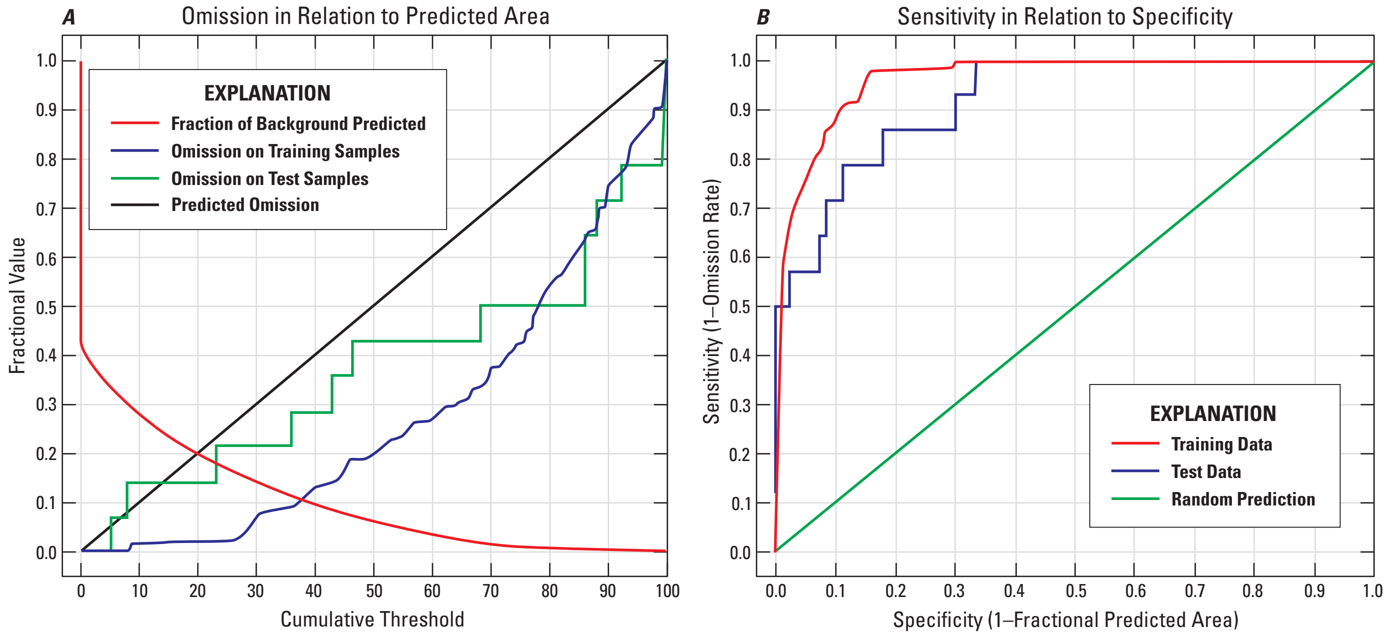

Figure 12.4. Comparison of model performance using threshold-dependent *A*, and threshold-independent; *B*, methods. Graph *A* represents omission rate and predicted area as a function of the cumulative threshold for Johnsongrass training and test data; whereas graph *B* illustrates receiver operating characteristic (ROC) curves plotted as a function of sensitivity compared to specificity. Graph *A* shows an optimal omission curve for test samples that would resemble the predicted omission curve. For graph *B*, both the training and test dataset outperformed a random prediction, indicated by the steep rise at the origin, leveling off near the sensitivity value of one (see Chapter 9).

Vector and Pathway Spatial Model

With the exception of animal movements, spatial models of vectors and pathways can be easily created. Animal movement vectors may be spatially modeled where sufficient data exist. However, no data were available for animal movement vectors in BIBE. We obtained GIS data sets from either BIBE (roads and trails) or USGS (hydrology) to represent vectors and pathways. Water flow from summer monsoons and flood events represent one of the primary vectors (mechanisms of plant introduction) for the spread of Johnsongrass seed in BIBE, with streams and rivers being the primary pathway. As such, we created a 300-meter buffer around perennial streams and springs and a 30-meter buffer around intermittent streams. A representation of a stream pathway is provided in figure 12.5. Wind dispersion of seeds is also an important vector. We created 400-meter x 250-meter ellipses around each known Johnsongrass location to model potential wind distribution. Ellipses were oriented to the direction of the prevailing winds. The top of the ellipse was positioned at the georeferenced plant location. Since roads, trails, and campgrounds are common pathways for the spread of invasive plants, we buffered paved roads and campgrounds 150 meters, dirt roads 30 meters, and trails 15 meters. Buffer distances should be adjusted for individual parks or projects.

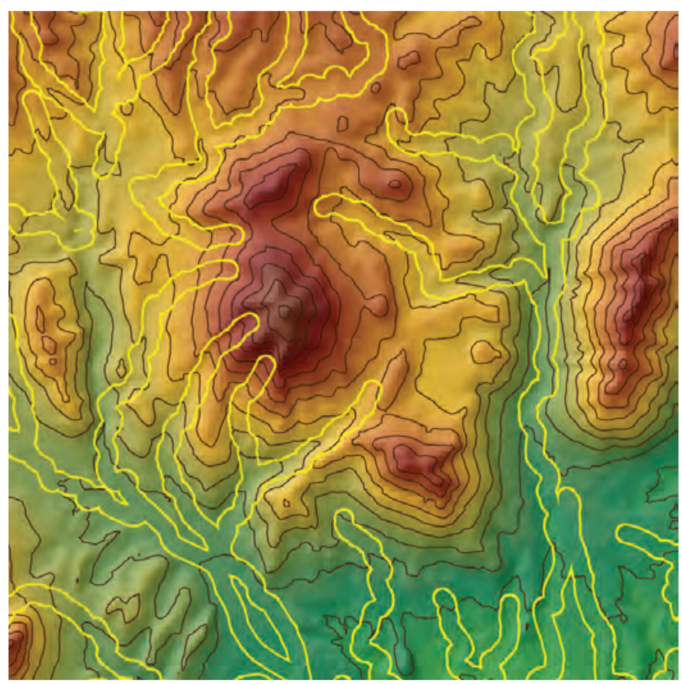

Figure 12.5. Graphical representation of a stream buffer (yellow lines) that serves as the primary pathway for the spread of Johnsongrass in Big Bend National Park, Texas.

Risk Analysis Model

Areas of potential risk for the invasion of Johnsongrass in BIBE (figure 12.6) were estimated by placing current known occurrences, potential suitable habitat, modeled vectors and pathways, areas with adequate soil moisture regimes (given desert environment), and areas with disturbances (for example, fire) into a spatial context. Areas near current populations that have pathways connecting to other potential habitat are considered at risk of invasion. These areas can be monitored to detect invasions before they become established. Further, existing plant populations can be controlled to reduce the possibility of spread to potential habitats. Barriers along pathways may help reduce the risk of spread.

Risk analyses can provide an effective approach for prioritizing areas for invasive species conservation efforts. For example, areas at risk of invasion can be placed in context to sensitive plant populations, or other management considerations, to further refine areas to be monitored. Further, risk surfaces provide a spatially explicit model that could be used to develop a ground-based sampling strategy to locate or monitor invasive plant populations (see Chapter 8).

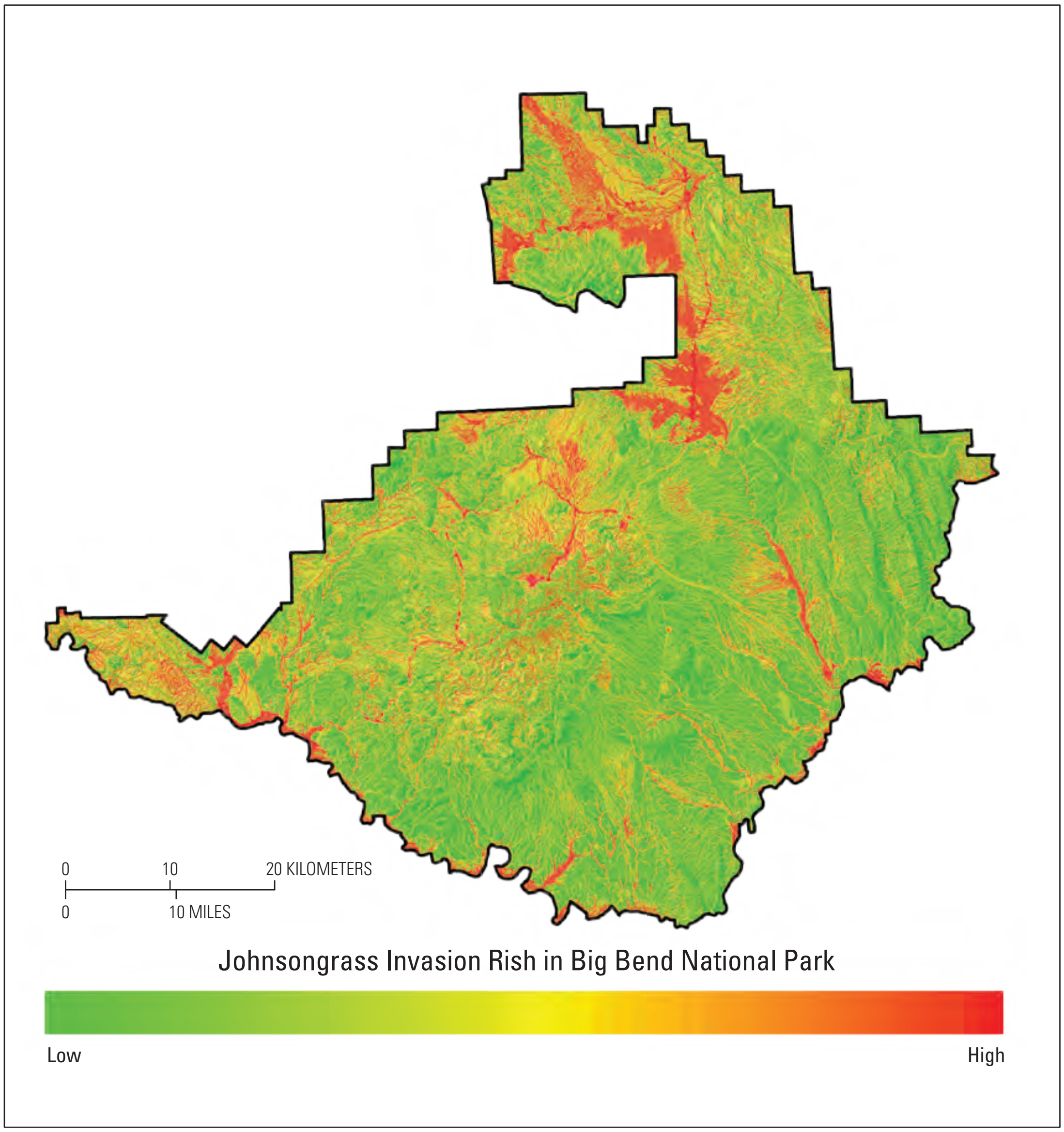

Figure 12.6. Johnsongrass risk assessment in Big Bend National Park, Texas estimated by using current known occurrences, potential suitable habitat, modeled vectors and pathways, areas with suitable soil moisture regimes, and disturbed sites.

References Cited

Allen, J.A., 1990, Establishment of bottomland oak plantations on the Yazoo National Wildlife Refuge Complex: Southern Journal of Applied Forestry, v. 14, p. 206–210.

Anderson, L.E., 1961, Johnson grass in Kansas: Kansas State University of Agriculture and Applied Science, Agricultural Experiment Station, Circular, v. 380, p. 12.

Anderson, L.E., Appleby, A.P., and Weseloh, J.W., 1960, Characteristics of Johnson grass rhizomes: Weeds, v. 8, p. 402–406.

Bendixen, L.E., 1988, Johnson grass (*Sorghum halepense*) management systems: Weed Technology, v. 2, p. 64–67.

Diggs, G.M., Lipscomb, B.L., and O'Kennon, R.J., 1999, Illustrated flora of north-central Texas: Fort Worth, Sida Botanical Miscellany no. 16, Botanical Research Institute of Texas, 1,626 p.

Ghersa, C.M., Martinez-Ghersa, M.A., Satorre, E.H., Van Esso, M.L., and Chichotky, G., 1993, Seed dispersal, distribution and recruitment of seedlings of *Sorghum halepense*: Weed Research, v. 33, p. 79–88.

Great Plains Flora Association, 1986, Flora of the Great Plains: Lawrence, University Press of Kansas, 1,392 p.

Hartzler, R.G., Gover, A., and Stellingwerf, J., 1991, Factors affecting winter survival of Johnson grass (*Sorghum halepense*) rhizomes: Weed Technology, v. 5, p. 108–110.

Holm, L.G., Plocknett, D.L., Pancho, J.V., and Herberger, J.P., 1977, The world's worst weeds—Distribution and biology: Honolulu, University Press of Hawaii, 609 p.

Howard, J.L., 2004, *Sorghum halepense*, *in* Fire effects information system: U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory, accessed December 17, 2010, at *http://www.fs.fed.us/database/feis/*.

Huang, W.Z., and Hsiao, A.I., 1987, Factors affecting seed dormancy and germination of Johnson grass *Sorghum halepense* (L.): Weed Research, v. 27, p. 1–12.

Leguizamon, E.S., 1986, Seed survival and patterns of seedling emergence in *Sorghum halepense* (L.): Weed Research, v. 26, p. 397–403.

Martin, W.C., and Hutchins, C.R., 1980, A flora of New Mexico, Volume 2: Vaduz Germany, Lubrecht and Cramer, 2,589 p.

McWhorter, C.G., 1989, History, biology, and control of Johnson grass: Review of Weed Science 4, p. 85–121.

Monaghan, N., 1979, The biology of Johnson grass (*Sorghum halepense*): Weed Research, v.19, p. 261–267.

Phillips, S.J., Anderson, R.P., and Schapire, R.E., 2006, Maximum entropy modeling of species geographic distributions: Ecological Modelling, v. 190, p. 231–259.

Radford, A.E., Ahles, H.E., and Bell, C.R., 1968, Manual of the vascular flora of the Carolinas: Chapel Hill, The University of North Carolina Press, 1183 p.

Tellman, B., 1998, Stowaways and invited guests—How some exotic plant species reached the American Southwest *in* Tellman, B., Finch, D.M., Edminster, C., and Hamre, R., eds., The future of arid grasslands—Identifying issues, seeking solutions, Proceedings, October 9–13, 1996, Tucson, Arizona: Fort Collins, Colo., Proceedings RMRS-P-3 U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, p. 144–149.

Unger, P.W., Miller, S.D., and Jones, O.R., 1999, Weed seeds in long-term dryland tillage and cropping system plots: Weed Research, v. 39, p. 213–223.

Warwick, S.I., and Black, L.D., 1983, The biology of Canadian weed 61—*Sorghum halepense*: Canadian Journal of Plant Science, v. 63, p. 997–1014.

Wunderlin, R.P., 1998, Guide to the vascular plants of Florida: Gainesville, University Press of Florida, 806 p.

Yang, Y.W., Newton, R.J., and Miller, F.R., 1990, Salinity tolerance in *Sorghum*, II—Cell culture response to sodium chloride in *S. bicolor* and *S. halepense*: Crop Science, v. 30, p. 781–785.

San Francisco Area Network Cast Study

By Andrea Williams, Susan O’Neil, Elizabeth Speith, Jane Rodgers, Maria Alvarez, and Robert Steers

Chapter 13 of

Early Detection of Invasive Plants—Principles and Practices

Edited by Bradley A. Welch, Paul H. Geissler, and Penelope Latham

Scientific Investigations Report 2012–5162

U.S. Department of the Interior
U.S. Geological Survey

Contents

Figures

Table

Insets

Chapter 13.
San Francisco Area Network Case Study

By Andrea Williams,[1] Susan O'Neil,[2] Elizabeth Speith,[3] Jane Rodgers,[4] Maria Alvarez,[5] and Robert Steers[6]

Overview

The purpose of this chapter is to provide a case study of the early detection protocol developed at the San Francisco Bay Area Network (SFAN) of National Parks, with particular emphasis on Golden Gate National Recreation Area where the protocol was initially developed and tested. The program was selected as a case study due to the unique approach of utilizing volunteers for a successful early detection program. Developing an early detection protocol that is scalable to available staff and volunteers was critical to this network of parks. Having a program that can adapt to different person-hour and skill levels allows parks to maximize volunteer and staff effectiveness. This chapter will provide background information on the park(s), details from the monitoring protocol (approved and formalized in 2009), and lessons learned since implementation. The full protocol is available at *http://science.nature.nps.gov/im/units/sfan/assets/new_docs/Protocols/Invasive_plants_monitoring_protocol.pdf*, accessed August 8, 2014.

The San Francisco Bay Area Network of National Parks—Fort Point National Historic Site, Golden Gate National Recreation Area (GOGA), John Muir National Historic Site, Muir Woods National Monument, Pinnacles National Monument, Point Reyes National Seashore (PORE), and Eugene O'Neil National Historic Site—are located on California's central coast, in one of the most significant areas for biodiversity in the nation (Stein and others, 2000) and world (Myers and others, 2000a). These parks remain significant to the conservation of endemic species and communities despite and because of their close proximity to large metropolitan centers of the San Francisco Bay Area, like San Francisco, Oakland, and San Jose, which have displaced native habitat and are projected to increase in population size from about 7 million to 8 million by 2020 (Association of Bay Area Governments, 2000). Recognizing the extraordinary significance and exposure to invasive species in the region, the United Nations Educational, Scientific, and Cultural Organization, Man in the Biosphere Program, designated the Central California International Biosphere Reserve in 1988 which includes several SFAN park units.

It is this confluence of urbanization and remarkable natural resources that enables the SFAN to utilize volunteers as citizen scientists for documenting everything from rare to invasive species. For many years Golden Gate National Recreation Area, in particular, has worked closely with the communities and landowners surrounding the park to develop park stewards. Many of these stewards receive specialized training to participate in activities while others volunteer their expertise regarding local cultural or natural resources issues. With this opportunity for additional human resources, the parks must also develop systems for training and managing volunteers and to quality check their contributions.

Parks within the SFAN are all either adjacent to or near urban settings, with private landowners along park boundaries. Many of these parks have been altered through human habitation—as home or work sites, agricultural, or working landscapes. Due to the proximity of human development, many of the invasive species in the parks are of horticultural origin that spread as an unintended consequence of agriculture, local gardening and landscaping.

Invasive plant species negatively affect natural resources in several ways, including altering landscapes and fire regimes, reducing native plant and animal habitat, and blocking views and increasing trail maintenance needs. Given the extraordinary biodiversity of the San Francisco Bay Area, and the pressure to develop private lands in the area, SFAN parks serve as crucial refugia for native species. Over 100 rare plant species can be found in SFAN parks. Invasive plants threaten many of these rare species. In Golden Gate National Recreation Area alone, 25 species of nonnative plants were noted as directly threatening rare plant populations (Golden Gate National Recreation Area, written commun., 2004).

As noted throughout this publication, the best way to prevent further spread and infestation of these species is to vigilantly monitor the wildland/urban interface. Trails, roads,

[1]Marin Municipal Water District.

[2]Puget Sound Partnership, Tacoma, Washington (United States).

[3]U.S. Geological Survey Pacific Basin Information Node and Hawaii Cooperative Studies Unit, UH-Hilo, Hawaii.

[4]Grand Canyon National Park, National Park Service.

[5]Golden Gate National Recreation Area, National Park Service.

[6]San Francisco Bay Area Network, Inventory and Monitoring Program, National Park Service.

and waterways are the main routes of infestations in most natural areas, and the SFAN is no exception. Monitoring these important corridors of spread is a key strategy utilized in the SFAN protocol for early detection, along with identifying source populations and other disturbed areas within the parks. Monitoring the likely routes of invasion and uninfested areas is the most effective way to prevent the spread of existing invasive species and the infestation of new species in SFAN parks (for example, McNeely and others, 2001).

Detecting invasive species before they become established has been a longstanding practice in agriculture, with point-of-entry and point-of-distribution inspections, insect traps, and nursery certification. In 2000, more than 33.5 million vehicles were monitored at the California border agricultural inspection stations. More than 70,000 lots of prohibited material were intercepted at the border inspection stations, including insects such as the gypsy moth (*Lymantria dispar*), imported fire ant (*Solenopsis invicta*), boll weevil (*Anthonomus grandis*), Mexican fruit fly (*Anastrepha ludens*), zebra mussel (*Dreissena polymorpha*), pecan weevil (*Curculio caryae*), Japanese beetle (*Popillia japonica*), Oriental fruit fly (*Bactrocera dorsalis*), European corn borer (*Ostrinia nubilalis*), burrowing nematode (*Radopholus similis*); and plants such as musk thistle (*Carduus nutans*), and diffuse knapweed (*Centaurea diffusa*) (California Department of Food and Agriculture, 2005). Wildland managers have been slower to implement strong prevention and early-detection programs; lack of clear regulatory oversight, low funding/staffing, and (or) unsuitable vector control, hamper such efforts.

Finding and removing invasive species before they negatively impact native species will prevent further loss of biodiversity and is crucial for successful eradication (Myers and others, 2000b; Harris and others, 2001; Timmins and Braithwaite 2002; Rejmanek and Pitcairn, 2002). At Point Reyes National Seashore (PORE), for example, removal of invasive European beachgrass (*Ammophila arenaria*) as part of the coastal dune restoration program has already resulted in reestablishment of Tidestrom's lupine (*Lupinus tidestromii*) and beach layia (*Layia carnosa*), both of which are on the Federal government list of endangered plants, and nesting of federally threatened western snowy plover (*Charadrius alexandrinus nivosus*) (Peterson and others, written commun., 2003).

Issues Related to Past Invasive Plant Work—Case Study at Golden Gate National Recreation Area

"The most critical step in addressing new invasive plant problems is to know they exist" (Federal Interagency Committee for the Management of Noxious and Exotic Weeds, 2003). Prior to the inception of the Inventory and Monitoring (I&M) Program, many SFAN parks had maps, removal programs, and even monitoring efforts in place. One challenge to invasive plant monitoring and management programs is consistency in the way data is collected and managed. Invasive plant work is often done by multiple entities, accomplished opportunistically and is dependent on funding and staffing that may not be available from year to year. The following section highlights past invasive plant work conducted at the SFAN parks, particularly highlighting Golden Gate National Recreation Area as the initial site of the early detection protocol implementation, the park unit with most active cadre of volunteers, and the largest park in the network.

Golden Gate National Recreation Area (GOGA) has several programs working exclusively on invasive species removal and restoration. The Habitat Restoration Team (HRT), funded partially by the Golden Gate Parks Conservancy, began in 1992 and has grown into a large-scale invasive-plant removal program. The team, and its early-detection/followup-focused offshoot, the Invasive Plant Patrol, have set routes and priority infestations they treat weekly in summer and monthly in fall/winter. Three to seven core volunteers are often augmented by groups of 20 or more. The Site Stewardship Program (SSP) "is a Golden Gate National Parks Conservancy volunteer program, created in 1993. SSP's mission is to bring people together to protect ecologically sensitive areas within the Golden Gate National Recreation Area" (Golden Gate National Recreation, 2004). They focus on restoration at four areas of concern for endangered species within GOGA. The Parks Conservancy runs several similarly successful volunteer groups such as Presidio Park Stewards, Trails Forever, and the Native Plant Nurseries. GOGA logs 25,000 hours of plant-related volunteer hours annually; these programs account for part of 150,000 volunteer hours worked for natural resource programs each year at the park (S. Fritzke, written commun., 2006).

Surveys of targeted invasive nonnative plants were initiated at GOGA in 1987. These surveys were conducted by qualified botanists who hand-mapped species infestations and distributions using U.S. Geological Survey quadrangle maps. In 1995 the park began collecting invasive-plant species data on the Presidio and in Gerbode Valley by using GPS equipment. Between 1996 and 1998 survey and monitoring efforts were continued, but many efforts were not well planned and were inconsistently implemented. As a result, different park programs and projects used a variety of mapping protocols based upon their variable needs and available resources, leading to a wealth of useful but disjointed information. The park developed a manual for surveying and mapping invasive species in 1999 to address this, and it was piloted as part of the Redwood Creek watershed data-collection efforts. While the manual provided consistency and a protocol for data collection that was used in a number of park watersheds, it did not take into account some of the specific GPS/GIS and data-management challenges, and few current weed workers are aware of its existence. As of 2006, the park's method of documenting weed patches and infestation areas as well as weed management was a Microsoft Access application (the Restoration Database) that did not allow the input of spatial data. As a result, spatial weed survey data were recorded in an entirely different place and manner (in scattered GIS shapefiles) than weed treatment data. Finally, only the watershed-based mapping efforts conducted in 1999–2000 recorded the actual survey areas and what was surveyed. While some historical and current locations of many invasive nonnative plants are known, there is no reason to believe that areas with no data are weed-free. Golden Gate National Recreation Area is not unique in facing these challenges with invasive species data management.

While GOGA has the most extensive and long-running invasive species management program(s), all SFAN parks have a wide variety of efforts aimed primarily at invasive species control, but information is rarely shared between parks or programs run by different organizations. Parks have varying levels of capacity and lack a reliable and comprehensive method of obtaining a bigger picture for landscape level management. The early detection protocol builds on the existing volunteer capacity of parks and focuses on helping parks target their efforts and collect and share quality information.

The goals of this monitoring strategy are to formalize and build on current knowledge (documented in GIS and databases) of species locations, to spot new infestations, and to notify park managers so they can eradicate infestations at more cost-efficient stages (that is early detection and rapid response). Given the widespread problem of invasive species in the San Francisco Bay Area and the spread of infestations across park boundaries, close coordination through local Weed Management Areas, the Bay Area Early Detection Network (BAEDN, *http://www.cal-ipc.org/WMAs/BAEDN/,* accessed August 8, 2014), and through the California Invasive Plant Council, is an essential part of this early-detection protocol.

Collaboration

As illustrated above, different programs within one park often do not collect data that can be compared or analyzed without considerable difficulty. Thus, consistent methodologies between parties working in the same area are crucial and will greatly increase efficiency of labor and funding. In California, organizations like Weed Management Areas (WMAs) and the California Invasive Plant Council (Cal-IPC) help agencies (governmental or nongovernmental), private landowners, interest organizations, and the public by serving to promote cooperation and acting as information clearinghouses. Specifically in the San Francisco Bay Area, the BAEDN serves local resource managers in this capacity by producing species lists of problematic plants, protocols and online resources for documenting and reporting, and access to knowledgeable local experts. The large SFAN parks are members in their regional WMAs: Marin-Sonoma, San Francisco, and San Benito. The SFAN large parks have worked jointly on grant-funded projects through the WMAs. SFAN I&M staff are also working collaboratively with BAEDN, the Bay Area Open Space Council, the Association of Bay Area Governments, and Cal-IPC to share protocols, methods, materials, and reporting, and to recruit and train early-detection volunteers. Lastly, the Marin-Sonoma WMA also has a nursery outreach program that will assist SFAN in identifying new and potential invasive plants and possibly prevent invasions through promoting voluntary codes of conduct for nurseries, landscapers, and landowners. This local effort also interfaces with a statewide (top-down) effort, the California Horticultural Invasives Prevention (Cal-HIP) project, which works with major growers and distributors to substitute noninvasive species for problematic nonnative species.

Monitoring Questions

Box 13.1. Monitoring questions for SFAN early detection:

1. Where are new populations of invasive-plant species becoming established in SFAN parks?
2. What are the features of road and trail corridors that make the best predictors for invasive-species establishment?
3. Are invasive species spreading from roads and trails into sensitive or critical park habitat?

There are two main components of the invasive species monitoring program for the SFAN that were adapted from the NPS Prairie cluster prototype monitoring program (Young and others, 2007). The first component, and the focus of this case study and our network's program (Williams and others, 2009), is early detection monitoring to locate new, isolated infestations before they become entrenched in the network parks. The second component, not addressed here, is monitoring existing populations of species that are known to have the ability to change the structure and function of entire ecosystems. This involves choosing priority habitats and monitoring changes in their native and exotic components over time. This type of work is integrated into the SFAN plant community change protocol. While monitoring known populations may seem less useful than removing them, understanding spread rates will help target efforts towards the fastest-spreading species, refine models and revisit intervals, and lend confidence to estimates of "acres of infestation prevented" by rapid-response programs.

Parks need to know where incipient populations of highly invasive plants are becoming established, and protect the most high priority habitats from invasion. The objectives designed to answer the monitoring questions focus on surveying road- and trail-side in priority areas. The protocol explicitly describes methods to be used by staff and trained volunteers. Field testing began in 2006 at GOGA and the protocol was reviewed and approved in 2009. Budget constraints necessitate looking in areas where it will do the most good—in high-quality and high-risk areas—along a primary vector for invasive plants and, when possible, using volunteer labor. While surveyors may spot weeds far from the trail in the open scrub and grasslands of SFAN parks, reliable presence/absence data are limited to within several meters from roads and trails.

Protocol Objectives

The approved protocol objectives are as follows:

1. Develop and revise as needed (minimally, every three years) a list of target invasive species for each park including: species that do not currently occur in the parks, that occur in localized areas of parks, or are extremely rare, but that would cause major ecological or economic problems if they were to become established in SFAN parks.

2. Prioritize SFAN subwatersheds by management importance, risk, and current infestation level. Within each park, use visual assessment and GPS technology to detect and map presence and absence of priority weed species along all roads and trails in the top ranked 25 percent of subwatersheds annually, the next 50 percent biennially, and the remaining 25 percent within 5 years (55 percent of all subwatersheds visited each year), noting presence and absence of priority weed species over the next five years.

3. Every five years, evaluate invasive-plant monitoring and mapping data collected to determine the primary pathways and predictive factors leading to new invasions along roads and trails in each park. Use these data to refine subwatershed rankings for search priority and timing. Identify possible management actions to prevent new infestations.

While only the second objective "qualifies" as a monitoring objective for I&M program purposes, all three are necessary to achieve the goals of this protocol. Overall, the current protocol can best be described as "looking for the worst plants in the best places" using volunteers to complement staff time (see Williams and others, 2009).

Using Rapid Assessment Techniques

This protocol builds on and standardizes efforts already in place in many parks including volunteer programs, active detection programs for finding invasive species, and research. We have selected early detection for its proven utility in identifying infestations while they are small and cost-effective to control. Combined with rapid-response programs, early detection helps to prevent spread to and invasion of uninfested areas. Early detection is also relatively easy to implement at several locations targeting a multitude of species with different levels of intensity. The chosen methods can be scaled according to resource availability, ranging from techniques for an opportunistic strategy with minimal staff in the field to a full volunteer/staff program with targeted and systematic efforts based on location, seasonality, ground-truthing, and removal in appropriate instances.

Qualitative techniques—such as the presence/absence data to be gathered for early-detection monitoring—are less resource-intensive, easier to analyze and explain to stakeholders, and facilitate monitoring of a larger area (Elzinga and others, 2001; Dewey and Anderson, 2004). Such low-intensity monitoring allows for a more rapid management response, as simpler data should lead to faster decisionmaking. Random plot-based sampling, even targeted to certain areas, is unlikely to capture rare occurrences. Relying on a volunteer effort for off-trail, plot-based sampling is inappropriate due to unpredictability in the number of volunteers available, their skill-levels, and safety concerns. Large-scale sweeps of road- and trail-side will help ensure broad coverage while capturing information to help inform future modeling to better target searches. Much of the park is accessible by trail. Roads and trails are major vectors for invasive plants, which tend to clump along these corridors. Searching these areas will capture the greatest amount of information for relatively less effort.

Creating Management Units

Existing nonnative species programs have often created subunits for their management areas that tend to be small (seldom exceeding a few hundred acres) and (or) have overlapping boundaries. Some are not spatially discrete or even mapped, decreasing data utility. In contrast, clearly defined management areas that correspond to ecological zones or partially closed systems have the potential to be most useful for landscape-scale management purposes. Prior to the inception of the invasive plant early detection protocol, park GIS specialists delineated all subwatersheds within each park unit by dividing CalWater watersheds into smaller, manageable areas based largely on topography. These boundaries were adopted by park resource managers for use as subunits in the early detection program (Williams and others, 2009). All subwatersheds within a park were then ranked based on the number and degree of current infestations, importance of resources present, and management priorities (fig. 13.1). Surveys are then implemented along roads and trails using an unequal sample design that is weighted by subwatershed rank.

Using Subwatersheds

Subwatershed boundaries are based on geologic features and are often more objective than other types of management units. Primary reasons for choosing subwatersheds included the presence of an existing subwatershed layer in the corporate GIS, the biological basis of subwatersheds, their tiering into watersheds which are used by other biologists when considering other management issues, and GOGA's existing practice of tracking invasive species removal by subwatershed. Using subwatersheds for nonnative-species monitoring makes biological sense. In addition to aquatic invasive plants, many terrestrial invasive species follow drainage corridors, either through lightweight seeds following up-canyon winds, vegetative spread along creeks or downstream fragment dispersal, or seed deposit by frugivorous animals in much-utilized riparian

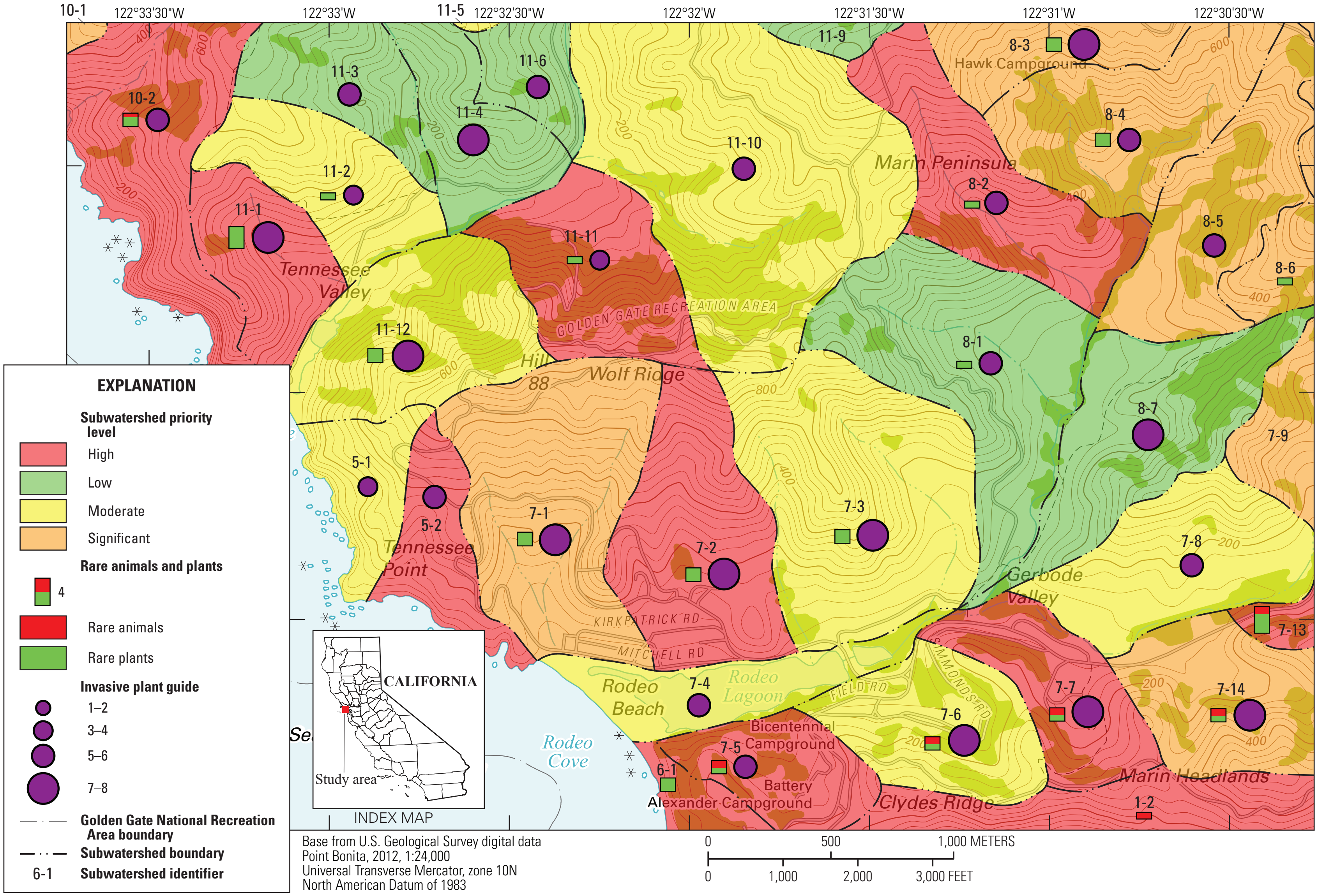

Figure 13.1. Prioritized subwatersheds in Golden Gate National Recreation Area's western Marin Headlands (Fort Cronkhite, Rodeo Beach, and Wolf Ridge). Colors show priority level, with red being the highest priority; three elements—rare plants and animals and number of work-performed guilds—are also shown.

habitats. Potential disadvantages to using subwatersheds include the fact that boundaries are not highly noticeable on the ground, transportation routes (roads, trails) generally pass through several subwatersheds, and field crews may have difficulty in using numeric subunits that lack a familiar reference name. Subwatersheds are also not of a standard size and commonly span more than one habitat type, which may pose a problem in comparisons.

Other Management Unit Types Considered but Rejected

Inventory and Monitoring staff also considered the following options in dividing parks into management units: other landscape features, habitat type, and grid. A short description of each option follows:

Park units could have been divided on the basis of landscape features. Roads, trails, waterways, rock outcrops, fences, and tree lines have been used for centuries by people to delineate areas. Most of these features are also functionally meaningful in invasion biology and serve well as survey paths. Using landscape features to form subunits is not recommended unless much of the park has already been divided that way, and documentation and GIS boundary coverages exist. Trails are often rerouted, fence lines deteriorate, and trees may fall or be cut down, limiting their utility as permanent boundary markers.

With a good vegetation map, parks could have been divided into habitat types. Many nonnative plants invade habitats preferentially, and searches may be targeted to fewer species. For more comprehensive surveys (for example if a list of all plants seen was kept), fewer species would potentially be encountered. As with other subunits, boundaries are not highly recognizable on the ground, transportation routes (roads, trails) generally pass through several habitat types, and edges can be convoluted and change quickly with succession and disturbances (within 10–20 years), making them poor standard search areas for repeat monitoring.

Parks could have also been divided into standard search area sizes (grids). While statistically easier to deal with, grids do not function well in management, as they have no biological basis. They also are not identifiable on the ground. People may have difficulty in using numeric subunits that lack a familiar reference name, and roads and trails will pass through multiple grids. Prioritizing or stratifying grids may be more difficult because a single grid cell may include multiple habitats, aspects, or other confounding factors, depending on cell size.

Prioritizing Management Units

Matrix Methods

SFAN developed a ranking matrix containing information from three general areas: management priority, risk, and current level of infestation. Each piece of information has an associated confidence level. The ranking matrix for GOGA was run using data from GIS layers from parks and the Exotic Plant Management Team, and data from the Restoration Database and 1994 PORE-GOGA vegetation map accuracy assessment plots. Coverages containing information from three general areas were added to the project: management priority, risk, and current level of infestation. ArcView 3.3, GeoProcessing Wizard (*http://www.rockware.com/product/featuresLobby.php?id=198&category=473*), and XTools (*http://www.xtoolspro.com/tools.asp*) were used to compile spatial data for analysis. Coverages of similar type (for example, all exotics polygon files; or all roads, trails, fences, and power lines) were combined using GeoProcessing "Merge themes together" into shapefiles, and intersect files (GeoProcessing "Intersect two themes") were made for each by using subwatersheds as the overlay. XTools "Update perimeter, area, acres, and length" was run for the nonpoint intersect files to add area or length of features within each subwatershed. The resulting *.dbf files from the intersected themes were imported into an Access database and analyzed. The January 2006 version of the GOGA "Work Performed" database, which stores vegetation management information, was mined for data, including number of species and hours of work organized by subwatershed. Similar species were grouped into guilds (graminoid, herb, forb, shrub/subshrub, vine/groundcover, broom, thistle, and tree) for analysis. Results of queries from the Access database were exported to Excel for summary and presentation. Table 13.1 presents an example dataset for the ranking matrix.

Box 13.2. Using guilds for nonnative species.

The use of guilds (see text) in analysis provides several benefits, such as:

- smoothing the range of species present (more than 100 in some areas),
- avoiding double-counting from misidentifications, (for example, *Cortaderia selloana* for *C. jubata*) or generic identifications (for example, *Cotoneaster* sp.), and
- helping characterize the complexity of invasions (five species from five guilds is different from five species from two guilds).

Table 13.1. Sample of subwatersheds and matrix elements (unranked).

[Elements were standardized by acreage of subwatershed in park, ranked, and given a score (generally –1, 0, or 1). Exceptions were number of rare animal taxa and hours of work performed. Rare plant and animal taxa scores were weighted for final rankings.--, not available]

Sub-watershed	Total acres	Acres in park	Infra-structure acres	Acres of rare plants	Number of rare animal taxa	Acres of invasive alliances	Acres of at-risk alliances	Mapped invasive acres	Hours exotic work performed	Number guilds mapped	Number guilds work performed
PORE3-3	1,812.73	1,812.74	8.49	16.60	1	15.46	385.03		6	1	1
PORE3-7	2,412.33	1,760.77	10.84	0.03	1	45.97	298.00	0.08	42	3	1
PORE4-1	1,446.99	1,258.60	11.33	0.00	1	189.65	41.89	9.92	6	3	2
PORE5-10	1,447.23	1,447.28	12.60	5.58	1	293.09	1,025.00	84.59	443	4	2
PORE5-13	1,949.91	1,852.16	22.67	135.36	1	496.60	923.65	82.60	32.5	5	3
PORE5-14	1,416.23	1,333.89	20.10	81.33	1	756.87	521.71	22.16	--	2	
PORE5-2	1,121.51	1,121.51	4.14	--	--	49.62	280.99	20.62	74.2	3	1
PORE5-4	1,805.93	1,797.87	14.72	3.12	3	110.49	1,217.14	21.33	63.5	4	2
PORE5-5	2,289.17	2,283.75	17.04	84.17	2	188.16	1,411.60	19.00	267.9	4	4
PORE5-6	1,791.23	1,791.24	8.90	26.06	1	179.43	1,415.45	5.58	2406	4	3
PORE5-7	1,741.57	1,741.57	13.90	26.28	1	585.38	1,065.60	14.88	53.75	2	2
PORE5-8	1,032.66	1,032.66	9.22	11.93	1	417.94	579.63	0.01	4	2	1
PORE6-12	1,098.21	1,098.22	13.41	337.95	1	586.89	490.85	2.66	--	2	
PORE6-13	1,062.28	1,062.29	11.77	474.68	2	440.50	604.57	--	479.5	2	2
PORE6-16	2,838.30	2,838.31	43.15	132.02	2	1,713.31	941.12	3.08	9,799.3	5	4
PORE7-1	2,072.52	2,072.53	29.10	136.93	2	1,383.21	680.35	0.22	--	--	--

Matrix Elements

In-Park Only

All acreage percentages were based on the number of acres in the park (determined by subtracting “out of park” alliance acres from subwatershed acres).

Roads, Trails, Power Lines and Fence Lines

Infrastructure lines were conservatively buffered at 4 feet to convert linear features into features with a recognizable area.

Rare Plants

Rare plant coverage data were generally good and were consulted for both number of taxa and total acres. The proportion of subwatershed acres comprising rare plant acreages was used as the scoring factor.

Rare Animals

Rare animal polygons overestimated habitat due to large buffers and unioning errors (that is a single area mapped multiple times was not “collapsed” into one area but counted as multiple areas). The polygons also included historical populations. Given these factors, the number of taxa was used to represent rare animal data in the matrix instead of acres (Table 13.1).

Vegetation Map Data

The 1994 PORE–GOGA vegetation map was used to determine nonnative-dominated (for example California nonnative annual grassland or Eucalyptus) and high- or low-risk alliances (Table 13.1); accuracy assessment plots were used for scoring number of species per plot (corrected for area).

Nonnative Species Mapping

Invasive-plant data from nonnative plant mapping presented a large problem: no negative data (that is no absence data). Consequently, it is unclear whether areas without polygons are infested or not. Additionally, overlapping coverages and lack of information about what was treated and what still needs to be treated make existing acreage values unreliable. Nonetheless, the proportion of invasive plants per area mapped was used as one of the ranking criteria because data lacking information or containing bad information (not mapped or mapped but treated) still yields search time, either from a higher priority by invasives not having been mapped or from staff doing removal of invasives (treated but not mapped).

Nonnative-Species Removal Database

The number of guilds was obtained from the "Work Performed" database, as was staff work time. These data categories were also used as scoring items, though staff time was not ranked and broken up into quarters (Table 13.1).

The total score for each subwatershed was obtained by combining respective risk scores from each matrix element. Then, based on the total score, a subwatershed was placed into one of the four rankings: high, significant, moderate, and low priority (see Williams and others 2009). Subwatersheds entirely outside the managed boundaries were excluded. Confidence levels for data were relatively uniformly low and did not factor into rankings. Confidence levels will become more important with standardized data collection for nonnative species mapping, when age of data will be a driving factor in prioritizing search areas.

Confidence levels are as follows:

1. **High confidence**: Knowledge is current and well documented. Surveys are no more than two years old (for infestation level) and cover most of the management unit. Landscape features noted are from a management document no more than 10 years old.

2. **Moderate confidence**: Knowledge is slightly out of date or lacks good documentation. Surveys are two to five years old, cover less than one-half of the management unit, or are based on anecdotal information from a good source. Landscape features are from an out-of-date document or are based on anecdotal information from a good source.

3. **Low confidence**: Knowledge is out of date and (or) lacks documentation. Surveys are more than five years old, cover little of the management unit, or are based on poor sources. Landscape feature information is unsubstantiated.

Prioritizing Species

The procedure for determining which species would be targeted for early detection in each management unit was based on the state of current knowledge. The list of species would change as the program proceeds due to introduction of new species, the potential establishment and expansion of previously "new" species, better understanding of species distributions, and the methodology used in monitoring these species due to changes in staff/volunteer/funding availability.

I&M staff did the following in 2006 for GOGA and PORE:

1. Reviewed the park data sets (NPSpecies, other plant lists) and compiled a list of all nonnative species known or thought to occur in the parks (about 300 species). Removed known noninvasive species (that is not in Global Compendium of Weeds) or species locally nonnative (for example, planted coast redwood (*Sequoia sempervirens*) in the Presidio's scrub and dune habitat); 174 species remained.

2. Noted which species were listed by the California Invasive Plant Council (Cal-IPC). Noted California Department of Food and Agriculture (CDFA) ratings, and whether the Nature Conservancy (TNC) has a completed abstract for the species. Also noted if an unlisted species shared invasive characteristics with a congener.

3. Based on best available knowledge, noted if a species is an ecosystem transformer. Ecosystem transformers are those plants that effect a SYSTEM CHANGE, not just crowd out other plants; they change ecosystem processes such as nutrient cycling, hydrology, and fire regime.

4. Based on best available knowledge (internal reports, California Natural Diversity Database forms, manager knowledge), noted if a species endangers rare plants. Invasive plants that are a danger to rare plant species have been documented in SFAN parks.

5. Based on best available knowledge, noted ease of control independent of number of acres infested. **High** is easily hand-pulled (or if shrub/tree will not resprout if cut), slow spread; **moderate** is easily hand-pulled (or other non-chemical) but rapid spread, or will fragment if hand-pulled (coppice if cut), slow spread; **low** is hard to hand-pull (fragments or coppices or has deep-seated roots), spreads quickly, is similar life form as nearby plants (for example grasses).

6. Based on best available knowledge, noted feasibility of control based on number of acres, cost for removal, politics, and access. Levels used were **low**: more than 100 acres, or less but high cost (difficult access, specialized technique), politics against removal; **medium**: 25–100 acres, or less but high cost (difficult access, specialized technique), politics against removal; **high**: less than 25 acres, control straightforward, politics neutral or in favor of removal. We based "controllable" acreage on size of unit and annual area treated by invasive-species program.

7. Contacted the county Weed Management Area (WMA) to determine what species are nearby but not yet in the management unit. The WMA will most likely have countywide data, which may or may not be applicable to the management unit.

8. Contacted NPS staff in the area. People living and (or) working nearby may have witnessed new infestations.

9. Visit local nurseries and depending on the type of nursery and staff knowledge, the following data will be gathered:

- What species they are selling as ornamentals that have the potential to become invasive in the management unit.
- What species the nursery staff are familiar with in the area that may be invasive.

Box 13.3. Ease and feasibility of removal.

Species can be easy to control but still not feasible to control: Monterey pine (*Pinus radiata*) and Monterey cypress (*Cupressus macrocarpa*) are both invasive but relatively easy to control—they take several years to reach reproductive age, and will not resprout when cut. However, these trees are not feasible for eradication since they are relatively widespread. Also, some populations are considered historical and removal is arduous. Lastly, in some situations people protest their removal since they are charismatic trees.

Scores were assigned based on rankings, and then the list was sorted by feasibility of control (see Williams and others 2009). The overall list resulted in several levels of priority (fig. 13.2). Breaking the list into smaller chunks served several purposes: new surveyors were introduced to a small number of the highest priority species and could be progressively trained, while experienced observers could inventory a site for species on all lists; data collection could be restricted on the basis of lower priority level; the levels captured several types of early-detection possibilities, such as species that are rare and invasive, new populations of widespread species, and species that are present but not known to be invasive (yet). The amount of data recorded when a species was encountered also differed by what list it was placed on. High priority species had more information collected on their spatial extent and environmental setting than the lowest priority species:

- Priority 1 species are mapped with a point and polygon, with associated habitat and cover data.
- Priority 2 species receive the same level of mapping if under the threshold size of 100 square meters (denoting a satellite population); otherwise they are given a point.
- Priority 3 species are mapped with presence/absence for the search area, or a point if less than 100 square meters.
- Priority 4 species have presence-absence data collected for the search area.

In 2007, 2008 and 2009, similar reprioritization procedures were undertaken with the PORE and GOGA invasive plant early detection species lists. These changes are reflected in annual reports available on the SFAN homepage: *http://science.nature.nps.gov/im/units/sfan/index.cfm* Many high-ranked species thought to be rare by managers were actually shown to be common once surveys were performed, which is why shifts in the list should be done annually for the first few years unless an inventory was done beforehand. Downlisting more common species allows for faster searching

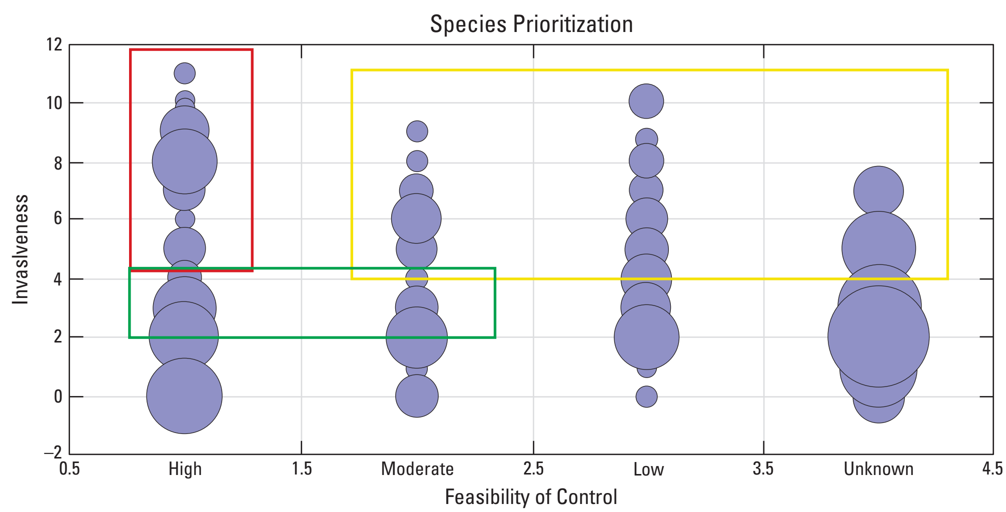

Figure 13.2. Priority 1 species (N=23; red box, upper left) scored high in invasiveness and high feasibility; Priority 2 species (N=29; yellow box, upper right) were highly invasive but lower feasibility, plus some species moderately invasive but high feasibility; Priority 3 species (N=31; green box, lower left) were moderately invasive and feasible; Priority 4 species (N=77) scored at least one point for invasiveness. Some shifts were made based on difficulty of identification (for example, grasses), and dune and aquatic species were segregated into a dedicated-search list. Size of dot represents number of species.

and ensures the appropriate amount of information is gathered. By 2009, the list included species appropriately ranked based on the methods outlined above in addition to feasibility.

Survey Methods

The number and abilities of early-detection personnel available to the SFAN parks vary widely. Having a program that can adapt to different person-hours and skill levels allows parks to maximize their effectiveness. Engaging people in detection, giving them clear direction and a point person to answer questions and receive invasive species occurrence reports, and following up with feedback on reports are essential components to a good program.

Opportunistic Sampling

Box 13.4. Observer Recruitment.

Observers can come from a variety of sources, and each park unit must take responsibility for creatively recruiting help for the program to be effective. Observers can be from any skill level, as long as they have completed minimal training in the identification of top-priority species. Observers consist of (but are not limited to) the following types of individuals:

- Interpreters leading hikes and disseminating information,
- Rangers patrolling the backcountry,
- Maintenance staff working at remote sites, road edges, and along trails,
- Resource managers, research permittees and scientists working in the backcountry,
- Park contractors,
- Park leasees,
- Special Use permittees,
- Volunteer groups (especially Native Plant Societies),
- Educational groups, and
- Park partners.

The program in SFAN parks has the flexibility to accommodate volunteers and staff members who vary in both botanical and technical skills through opportunistic sampling (OS), which has been used nationwide to increase the chances of early detection of nonnative plant species. Because staff, researchers, visitors, and volunteers travel through the SFAN parks regularly, each person is a potential set of eyes, or in other words, an "observer" who is able to make field observations with direction provided by SFAN. Opportunistic sampling is based on providing observers with the tools needed to correctly identify top-priority weed species and document new populations with the highest level of accuracy possible.

A primary difference between "passive" OS detection and "active" volunteer-based detection is the delineation of a search area. Active detection includes the search area so that an area may have negative data (absence data) on species appropriate for the search level of the observer. Passive detection likely will yield only the presence of priority species, which is still important information but not as useful as presence and absence data. In either case, having maps of different areas of the park with current infestations marked, along with "weeds to watch for" in those areas, will be useful for observers and allow one to get quality information in return. For specific details on methods and data collection, please see Williams and others (2009) for the full protocol and standard operating procedures.

Negative Data

An important component of managing invasive species is knowing where they do not occur. Surveyors use track logs to note where they searched. Species on the priority list of the observer's skill level that were not seen receive an "absent" listing in the survey area tab of the database. Advanced observers should be able to note all plant species seen within an area, at least to genus, and therefore have negative data for all other species.

Training

Informal training may be held as needed, but formal training is at the core of a volunteer-based program. Below are the different volunteer levels and their corresponding training requirements for the "Weed Watcher" program at GOGA:

- First-level observer (prerequisite participation in one guided hike and (or) the Weed ID 1 class)

    1. Train volunteers to identify Priority 1 target plants. Training takes place during a one hour orientation conducted by a SFAN I&M employee at a designated priority subwatershed. Each volunteer is exposed to search images and identifying features for each of the plants and will receive a set of "Plant-out-of-Place" ID reference cards. Identification skills are practiced during a two-hour guided hike along trails in the designated high-priority watershed.

    2. Train volunteers to use maps to record locations of target plants. Volunteers are exposed to a combination of USGS 7.5 minute quad maps, aerial photographs, and (or) GIS map layers during

a one-hour orientation at a field site. Skills are tested during a two-hour guided hike. Volunteers receive take-home paper maps for use during unsupervised "Weed Watcher" patrols.

3. Train volunteers to collect occurrence data using paper data sheets. Skills are tested during a two-hour guided hike. Volunteers are given multiple methods to report their findings via email, drop-off locations, and on-line report form (in development through Parks Conservancy or BAEDN). Volunteers receive take-home data collection sheets for use during unsupervised "Weed Watcher" patrols.

- Second-level observer (prerequisite participation in 3 guided hikes and (or) two guided hikes plus one Weed ID 1 class)

  1. Train volunteers to identify Priority 1 & 2 target plants. Training takes place during two outings with a SFAN I&M employee. Volunteers receive individual training on plant identification. Skills will be tested during a guided hike or via an on-line "Weed ID" test.

  2. Train volunteers to collect occurrence data with greater precision using paper data sheets and maps. Skills are tested during two guided hikes with a SFAN I&M employee.

  3. Train volunteers to make assessments of occurrences. Training includes determining cover class and distribution of patches.

  4. For volunteers interested in using GPS units: train volunteers to collect occurrence and assessment data using handheld GPS units programmed with the GeoWeed database interface. Training takes place during a series of guided hikes and a one-hour individual training and (or) a Biological Data Collection Using GPS class. Skills are tested during guided hikes.

- Third-level observer (prerequisite participation in a minimum of five guided hikes, one hour of GeoWeed training, and one hour of GPS training and (or) participation in a GPS biological data collection class)

  1. Train volunteers to identify the full list of high-priority target plants. Training takes place during a series of outings with a SFAN I&M employee, catered to the individual's needs. Volunteers receive a plant book for completing this requirement. Skills are tested during the guided hikes.

  2. Train volunteers to collect occurrence and assessment data using handheld GPS units programmed with the GeoWeed interface. Training takes place during a series of guided hikes and a one-hour individual training and (or) a Biological Data Collection Using GPS class. Skills are tested during guided hikes.

In general, SFAN volunteers are involved in parks and have a good base level of plant recognition. For example, many volunteers are already familiar with most of the high priority invasive species before formal training. In areas where people do not know plants as well, lists should be small—no more than 10 plants to avoid overwhelming new volunteers. Also, avoid giving volunteers species they will rarely see unless they are of high enough importance to resource management or if they are so visually distinct that identification is accurate even if rarely encountered. You may also need to train volunteers to photograph unknown species for identification.

Data Management and Reporting

To improve data quality and tracking, staff members use naming conventions and mapping standards. Search areas and weed occurrences have a similar descriptive code: name, subwatershed, and date are essential elements. The area code is SURVEYSUWAYYYYMMDDFILA, so 01/23/2006 survey in Subwatershed 7–1 by Andrea Williams would be SURVEY070120060123ANWI. The invasives-mapping naming convention substitutes a GESPXX plant code (USDA PLANTS) for SURVEY, so if Andrea found jubata grass (*Cortaderia jubata*) on her survey, she would code it COJU2X07012006012301, where the final two digits denote the number of occurrences of this species on this day (first is 01, second 02, and so on). While such coding seems cumbersome, the use of this naming convention incorporates a measure of data redundancy that can prevent user error. Mapping standards go beyond NAWMA basics to include guidance on how to draw the boundaries of plant patches based on biology and spatial distribution and defining a threshold for the patch size considered early detection.

Data acquired from surveys may be time sensitive. Acting upon new detections of highly invasive species is critical. Therefore, a feedback loop between monitoring and treatment programs must be established. On a monthly basis, new detection monitoring reports are submitted to the local park weed manager. These reports include both newly discovered species and newly discovered infestations. On an annual basis, the monitoring coordinator meets with local park weed managers to review the program, provide and receive feedback, and make program adjustments as necessary.

Next Steps and Lessons Learned

Volunteer Recruitment

Volunteer programs are most successful with a committed group, clear and consistent guidelines, and meaningful work (see *http://www.invasivespeciesinfo.gov/toolkit/arsresearch.shtml*). SFAN's most consistent volunteers come from other GOGA volunteer programs, but overall recruitment has been a slow process. The biggest increase in numbers of volunteers has occurred during endangered species 'Big Year' competitions where participants earn competition points by volunteering with our Weed Watcher program (Jen Jordan Rogers, written commun., 2011). However, once the Big Year competition is over, these volunteers drop off. Despite the high volunteerism rate in the San Francisco Bay Area, and specifically in GOGA, recruiting individuals who are interested in a long-term commitment to map invasive plants following specific guidelines is difficult. Over fifty unique volunteers have been trained since 2007, but only one has ever achieved 3rd level volunteer status and that person unfortunately moved away. All other volunteers have stayed at a 1st level status and showed no desire to advance. It is likely that there are people in the surrounding metropolis that would love to become an advanced weed mapping volunteer but are just not aware of the opportunity. Recruitment efforts have included advertising on Craigslist, community college job boards, on California Native Plant Society chapter websites, Parks Conservancy websites, and on official National Park websites. Despite these internet postings, few volunteers show up for advertised trainings or hikes.

So far the quality of data that the vast majority of our volunteers have provided is generally of poor quality compared to what is collected by park staff and interns who execute the protocol as part of their jobs. First, volunteers use a simplified species list and when they do find a priority species, they do not collect the same level of detail in regard to its size, coverage, and environmental setting. Second, confidence in species identification is low unless they provide a picture using their own camera. Third, mapping accuracy can be questionable (locations are mapped on aerial photographs) unless they provide GPS points using their own equipment. Most volunteers do not send in pictures or GPS points when they mail us their data sheets. Because of these reasons and that the amount of time spent on training and advertising has resulted in a minimal return on our investment, we have broadened our use of volunteers for the 2011 field season and beyond:

- To overcome data quality issues, in 2011 a smart phone application was designed for GOGA using the What's Invasive template (*http://whatsinvasive.com*). Without any advertising, seven people have found the GOGA phone application on their own and have contributed data, which provides GPS location and the option to upload a picture. The current strategy is to incorporate this tool into our existing training, and also to advertise its existence and invite users to try it out in the park and its surroundings. This tool will be useful for finding populations outside of and adjacent to park boundaries that we are not able to map but that could be propagule sources. BAEDN and Calflora have also developed a mobile reporting tool that feeds directly into the Calflora database; as of this writing, the tool is available in beta version for Android and iPhone. Check *http://baedn.org* for current information.

- To improve the benefit received from investing park staff time in volunteer training, the parks are shifting away from the traditional mapping courses and focusing on opportunities to pull weeds during survey hikes with program staff. This type of contribution by volunteers will greatly improve the number of weeds removed and will also not require lengthy training.

- Lastly, for those volunteers who demonstrate a strong desire and commitment to become an advanced weed mapper, park staff gladly train them to the fullest as this type of volunteerism is beneficial to the program. However, due to the time commitment this requires of park staff, it will not be done unless there is confidence in the long-term contributions of the person(s) interested.

Timing and Revisit Intervals

Matching survey time to optimal phenology is difficult with approximately 70 invasive species. More data collection is needed to determine whether multiple, seasonal visits in a single year or annual visits in different seasons over several years are more effective at detecting all of the invasive species in any one area. New Zealand researchers have modeled species behavior and detectability relative to control and budget thresholds (Harris and others, 2001), but these intervals might not fit accepted models for California. A tool that has been developed to aid our field staff in maximizing the detection of invasive species has been our "detectability index." This index tracks the phenological stage of the invasive plants on the lists and highlights what times of the year they are most visible or cryptic on the landscape, if applicable. A sudden increase in a particular species from one year to the next, for example, may have more to do with its phenological stage during each survey date and less to do with its rate of spread.

At this stage of implementation, determining revisit intervals may be difficult due to the unequal sampling among subwatersheds and the fact that a small proportion of species have been moved between different priority lists and thus, were mapped differently among years. However, most species were not moved among the different priority lists and also, high priority subwatersheds, which are mapped annually and are distributed throughout the parks, may have a high enough sample size to support such analyses at this point. Performing analyses that could aid in determining optimal revisit intervals should be performed in the next one or two years, once all subwatersheds in GOGA and PORE have finally been surveyed. Lastly, remapping priority species over the years, or revisiting previously mapped species not yet controlled, and then remapping their spatial extent and coverage, will also yield population- or patch-scale rates of change that can further aid determination of revisit intervals.

Data Difficulties

With I&M staff, park staff at several park locations, and Parks Conservancy staff all needing to contribute and access data without all having a common drive, how do you manage a database? Currently, the program uses GeoWeed, a modification of The Nature Conservancy's WIMS program (Williams and others 2009) that is stored on a network drive and is accessible to national park staff (including I&M) only. Other interested parties, like the Parks Conservancy, who perform a large amount of invasive plant work in GOGA, do not have access to the database. Also, national park staff housed in other offices have a slow connection to this database that limits its utility. Recently, advances have been made in cloud-based invasive plant database development through Calflora and BAEDN (*http://www.calflora.org/entry/wentry.html*). Yearly, all of the newly acquired geospatial data is added to this online database, but in the future, the goal is to convert GeoWeed to an online, cloud-based system so that all of the agencies working in GOGA can access it.

Rapid Response

SFAN staff will continue to work closely with park staff on program implementation, especially rapid response. Currently, qualified surveyors may remove small populations if under a threshold size (that is it would take less time to remove the plants than it would take to hike back out to remove them later), but larger populations do not have explicit rapid-response commitments from parks. For true success of the early-detection program, removal must be conducted within a certain period of time. SFAN staff, park weed managers, and the NPS California Exotic Plant Management Team are working together to find a level of feasible commitment and to look for additional funding sources for rapid-response programs. Because grantors prefer eradication programs over ongoing control, these efforts should be relatively successful.

Acknowledgments

Thanks to Sue Fritzke, Marcus Koenen, and Jeff Runde for program support. Ellen Hamingson and Brent Johnson have provided valuable assistance and oversight with supervision, implementation, and species prioritizations. Jen Jordan Rogers and Eric Wrubel have been invaluable field techs and volunteer coordinators. Last but not least, thanks to the interns and the more than 50 volunteers who have helped map invasive plant species in GOGA and PORE.

References Cited

Association of Bay Area Governments, 2000, City, county, and census tract forecasts: 1990–2020, accessed August 8, 2014, at *http://www.abag.ca.gov/abag/overview/pub/p2000/p2000ct/p2000ct-info.html.*

California Department of Food and Agriculture, 2005, Pest exclusion branch—Exterior exclusion program activities, accessed February 15, 2011, accessed September 28, 2011 at *http://www.cdfa.ca.gov/phpps/pe/InteriorExclusion/.*

Dewey, S.A., and Anderson, K.A., 2004, Distinct roles of surveys, inventories, and monitoring in adaptive weed management: Weed Technology, v. 18, p. 1449–1452.

Elzinga, C.L., Salzer, D.W., Willoughby, J.W., and Gibbs, J.P., 2001, Monitoring plant and animal populations: Malden, Mass., Blackwell Science, 368 p.

Federal Interagency Committee for the Management of Noxious and Exotic Weeds, 2003, A national early-detection and rapid-response system for invasive plants in the United States: Washington, D.C., accessed February 1, 2006, at *http://ficmnew.fws.gov/FICMNEW_EDRR_FINAL.pdf*.

Harris, S., Brown, J., and Timmins, S., 2001, Weed surveillance—How often to search?: Science for Conservation, v. 175, 27 p.

McNeely, J.A., Mooney, H.A., Neville, L.E., Schei, P.J., and Waage, J.K., eds., 2001, Global strategy on invasive alien species: International Union for Conservation of Nature, accessed February 1, 2006, at *http://www.iucn.org/*.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Gonseca, G.A.B., and Kent, J., 2000a, Biodiversity hotspots for conservation priorities: Nature, v. 403, p. 853–858.

Myers, J.H., Simberloff, D., Kuris, A.M., and Carey, J.R., 2000b, Eradication revisited: dealing with exotic species: Trends in Ecology and Evolution, v. 15, p. 316-320.

Rejmanek, M., and Pitcairn, M.J., 2002, When is eradication of exotic pest plants a realistic goal?: *in* Veitch, C.R., and Clout, M.N., eds. Turning the tide—The eradication of invasive species: Gland, Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, Species Survival Commission, Invasive Species Specialist Group, International Union for Conservation of Nature, p. 249–253.

Stein, B.A., Kutner, L.S., and Adams, J.S., eds., 2000, Precious heritage—The status of biodiversity in the United States, Oxford, United Kingdom, Oxford University Press, 399 p.

Timmins, S.M. and Braithwaite, H., 2002, Early detection of invasive weeds on islands, *in* Veitch, C.R., and Clout, M.N., eds., Turning the tide—The eradication of invasive weeds: Gland, Switzerland and Cambridge, United Kingdom, International Union for Conservation of Nature, Species Survival Commission, Invasive Species Specialist Group, International Union for Conservation of Nature, p.: 311–318.

U.S. Geological Survey, 2012, Point Bonita: U.S. Geological Survey digital data, The National Map, accessed August 18, 2014, at: *http://viewer.nationalmap.gov/viewer/*.

Williams, A., O'Neil, S., Speith, E., and Rodgers, J., 2009, Early detection of invasive plant species in the San Francisco Bay area network—A volunteer-based approach: National Park Service Natural Resource Report NPS/SFAN/NRR–2009/136, 318 p., accessed April 10, 2014, at *http://www.amazon.com/Invasive-Species-Detection-Francisco-Network/dp/1489551204*.

Young, C.C., J.L., Haack, L.W., Morrison, and DeBacker, M.D., 2007, Invasive exotic plant monitoring protocol for Heartland Network Inventory and Monitoring Program: Fort Collins, Colo., National Parks Service, Natural Resource Report NPS/HTLN/NRR-2007-018.

Glossary

Definitions of terms related to the research, management, and monitoring of invasive species vary widely in their use and are by no means consistently or accurately used within or among agencies, organizations, or academia. Indeed, there is ongoing debate over invasive species terminology in all realms. Nonetheless, there is a need to facilitate communication, cooperation, and education through the use of consistent terms. Additionally, Federal agencies are required to adhere to national standards established through legislation and policy.

The following definitions include those taken from the Invasive Species Executive Order 13112 (1999) signed under President William Clinton and instated by the National Invasive Species Council formed under said executive order. These definitions, despite continued debate, are intended for adoption by all Federal agencies including the National Park Service (NPS). Other definitions such as those from the NPS Management Policies 2001 also are listed for completeness. Given the relationship between the NPS Inventory and Monitoring Program (NPS I&M) and the construction of this handbook (see Foreword), appropriate terms used and defined by the NPS I&M Program have been included to provide context for their use throughout this document. In some cases, it is recommended that current NPS definitions be revised to meet Federal standards. Sources are listed with each definition. Some explanation is given where necessary.

A

Active detection Reliant upon planned search strategies.

Adaptive management A systematic process for continually improving management policies and practices by learning from the outcomes of operational programs. Its most effective form— "active" adaptive management—uses management programs that are designed to experimentally compare selected policies or practices by implementing management actions explicitly designed to generate information useful for evaluating alternative hypotheses about the system being managed (NPS I&M).

Alien species With respect to a particular ecosystem, any species, including its seeds, eggs, spores, or other biological material capable of propagating that species, that is not native to that ecosystem (USPEO, 1999).

Area frame A sampling frame that is designated by geographical boundaries within which the sampling units are defined as subareas (NPS I&M).

Attributes Any living or nonliving feature or process of the environment that can be measured or estimated and that provide insights into the state of the ecosystem. The term Indicator is reserved for a subset of attributes that is particularly information-rich in the sense that their values are somehow indicative of the quality, health, or integrity of the larger ecological system to which they belong (Noon 2003). See Indicator. May also be referred to in the data-management section as data associated with GIS files.

B

Biological significance An important finding from a biological point of view that may or may not pass a test of statistical significance (NPS I&M).

C

Co-location Sampling of the same physical units in multiple monitoring protocols (NPS I&M).

Conceptual models Purposeful representations of reality that provide a mental picture of how something works to communicate that explanation to others (NPS I&M).

Containment Prevention of spread (Mack and others, 2000).

Control Partial elimination of new and existing invasive plants and associated material.

D

Data validation The process of reviewing the finalized data to make sure the information presented is logical and accurate.

Data verification The process of ensuring the data entered into a database correspond with the data recorded on the hardcopy field forms and data loggers.

Diffusion coefficient, D Demographic variable used to characterize the rate of spread for a given species based on its dispersion habits (Williamson, 1996).

Driver The major external driving forces that have large-scale influences on natural systems. Drivers can be natural forces or anthropogenic (NPS I&M).

E

Early detection Involves active, planned measures and (or) passive, incidental reports used to locate newly introduced nonnative species in a given region. Populations of these species may exist as new occurrences in the region or they may be incipient populations of nonnative species that exist elsewhere in the region (NISC, 2003).

Ecological integrity A concept that expresses the degree to which the physical, chemical, and biological components (including composition, structure, and process) of an ecosystem and their relationships are present, functioning, and capable of self-renewal. Ecological integrity implies the presence of appropriate species, populations, and communities and the occurrence of ecological processes at appropriate rates and scales as well as the environmental conditions that support these taxa and processes (NPS I&M).

Economic harm Adverse impact on a person's property or business, on public property or business, or on the economy of a community, region or the State as a whole (Odell, 2004).

Ecosystem A spatially explicit unit of the Earth that includes all of the organisms, along with all components of the abiotic environment within its boundaries (Likens, 1992).

Ecosystem drivers Major external driving forces such as climate, fire cycles, biological invasions, hydrologic cycles, and natural disturbance events (for example, earthquakes, droughts, floods) that have large-scale influences on natural systems (NPS I&M).

Ecosystem management The process of land-use decisionmaking and land-management practice that takes into account the full suite of organisms and processes that characterize and comprise the ecosystem. It is based on the best understanding currently available as to how the ecosystem works. Ecosystem management includes a primary goal to sustain ecosystem structure and function, a recognition that ecosystems are spatially and temporally dynamic, and acceptance of the dictum that ecosystem function depends on ecosystem structure and diversity. The whole-system focus of ecosystem management implies coordinated land-use decisions (NPS I&M).

Environmental harm Includes loss of native biodiversity, ecological processes, natural resources or their use in the State, including, without limitation, harm to State or Federal threatened or endangered species or their habitat (Odell, 2004).

Eradicate To kill, destroy, remove and prevent the further growth or reproduction of a species or population in an ecosystem (Odell, 2004).

Exotic species "Exotic species" are those species that occupy or could occupy park lands directly or indirectly as the result of deliberate or accidental human activities. Exotic species are also commonly referred to as nonnative, alien, or invasive species. Because an exotic species did not evolve in concert with the species native to the place, the exotic species is not a natural component of the natural ecosystem at that place (NPS, 2001). [Historically, the term "exotic" has been reserved for showy and (or) tropical species and only recently associated with invasive species. Further, nonnative, exotic, and alien are not synonymous with invasive under the current national definition or most scientific definitions.]

External validation An independent assessment of the model, typically accomplished by the collection of new data to which the model is applied and then evaluated.

F

Fitness homeostasis Characteristic of a species that maintain relatively constant fitness over a range of environments (Rejmanek, 2000).

Focal resources Park ecosystems, flora, fauna, or abiotic features that, by virtue of their special protection, public appeal, or other management significance, have paramount importance for monitoring regardless of current consequences or whether they would be monitored as an indication of ecosystem integrity. Focal resources might include ecological processes such as deposition rates of nitrates and sulfates in certain parks, or they may be a species that is harvested, endemic, alien, or has protected status (NPS I&M).

G

Guilds Plants grouped into similar lifeform classifications (graminoid, herb, forb, shrub/subshrub, vine/groundcover, broom, thistle, and tree).

I

Import To bring into the State from another State or country (Odell, 2004).

Incipient Beginning to come into being and refers to newly established populations of invasive plants (as in Mack and others, 2000).

Indicators A subset of monitoring attributes that are particularly information-rich in the sense that their values are somehow indicative of the quality, health, or integrity of the larger ecological system to which they belong (Noon, 2003). Indicators are a selected subset of the physical, chemical, and biological elements and processes of natural systems that are selected to represent the overall health or condition of the system (NPS I&M).

Internal validation Techniques used to evaluate a model that use various randomization approaches to split the data into small subsets for model building and subsequent evaluation.

Intrinsic rate of natural increase, r Demographic variable used to characterize exponential population increase for a given species when the survivorship and reproductive factors have been predetermined and fixed (Williamson, 1996).

Introduction The intentional or unintentional escape, release, dissemination, or placement of a species into an ecosystem as a result of human activity (USPEO, 1999).

Invasive species An alien species whose introduction does or is likely to cause economic or environmental harm or harm to human health (USPEO, 1999). [Though there is debate over the application of the term "invasive" to native and nonnative species, the use of the term "nonnative invasive species" is redundant under this national definition.]

Inventory An extensive point-in-time survey to determine the presence/absence, location or condition of a biotic or abiotic resource (NPS I&M).

K

Kappa statistic A measure of the agreement for the main matrix diagonals after the probability of chance has been removed.

L

Lag time The time between initial establishment and initial spread of an invasive species.

M

Management objectives Realistic, clear, and measurable statements of the desired future condition of a resource or the desired outcome of a specified management action (Elzinga and others, 1998).

Measures Specific feature(s) used to quantify an indicator, as specified in a sampling protocol. For example, pH, temperature, dissolved oxygen, and specific conductivity are all measures of water chemistry (NPS I&M).

Measures of model fit Those metrics used to determine how well the a reserved subset of study data (training data) fits the selected statistical model.

Metadata Data that describe the content, quality, condition, and other characteristics of a dataset. Its purpose is to help organize and maintain an organization's internal investment in spatial data, provide information about an organization's data holdings to data catalogues, clearinghouses, and brokerages, and provide information for processing and interpreting data received through a transfer from an external source (NPS I&M).

Monitoring A collection and analysis of repeated observations or measurements to evaluate changes in condition and progress toward meeting a management objective (Elzinga and others, 1998). Detection of a change or trend may trigger a management action, or it may generate a new line of inquiry. Monitoring is often done by sampling the same sites over time, and these sites may be a subset of the sites sampled for the initial inventory (NPS I&M).

Monitoring objective As defined by the NPS I&M Program: a specific statement that provides focus about the purpose or desired outcome of a particular monitoring program. It assists in defining the scope of the observations or measurements that will be used to evaluate progress toward the management objective over time (NPS I&M).

N

Native species With respect to a particular ecosystem, a species that, other than as a result of an introduction, historically occurred or currently occurs in that ecosystem

(USPEO, 1999). For the NPS, they are defined as all species that have occurred or now occur as a result of natural processes on lands designated as units of the national park system. Native species in a place are evolving in concert with each other (NPS, 2001).

Nonnative species Species that are not naturally occurring in an ecosystem but capable of living and reproducing in the wild without continued human agency, including invasive species. Nonnative species may include genetically modified native species (Odell, 2004).

Noxious weed Any plant or plant product that can directly or indirectly injure or cause damage to crops (including nursery stock or plant products), livestock, poultry, or other interests of agriculture, irrigation, navigation, the natural resources of the United States, the public health, or the environment (USC, 2000).

P

Passive detection Reliant upon incidental reporting from staff and visitors.

Pathway The "advance or progression in a particular direction, regardless of mode (that is, conveyance) that disperses plants." The pathway has an origin, a vector, and one or more destinations (not a probability distribution of destinations) (Mack, 2003, p. 4)

Percent correct classification (PCC) A measure of the overall capability of the model to predict both "0" and "1").

Pests Living organisms that interfere with the purposes or management objectives of a specific site within a park, or that jeopardize human health or safety. Decisions concerning whether or not to manage a pest or pest population will be influenced by whether the pest is an exotic or a native species (NPS, 2001).

Prediction The potential for species to become established and spread within a biogeographic region.

Prevalence A measure of the frequency of the response of interest.

Prioritization The desire to focus management efforts, whether for control or early detection, on a reduced subset of the total species pool where these efforts will be most effective.

Propagule pressure The number of individuals of a species that is released (Lodge and others, 2006).

Protocols As defined by the NPS I&M Program: detailed study plans that explain how data are to be collected, managed, analyzed and reported and are a key component of quality assurance for natural-resource monitoring programs (Oakley and others, 2003).

Public lands Areas belonging to the Federal government or to the State or a political subdivision thereof, except for areas subject to exclusive Federal jurisdiction (Odell, 2004).

Q

Quality assurance (QA) The planned and systematic pattern of actions needed to provide adequate confidence that the project fulfills expectations.

Quality control (QC) The process of examining the data after they have been produced to make sure they are compliant with programmatic data-quality standards.

R

Random (probability) sampling Each unit in the population has a known probability of selection, and random selection is used to select the specific units to be included in the sample. With probability sampling, units that are more likely to have invasive species can and should be selected with greater probability to focus the search on areas that are most likely to have invasive plants.

Remote sensing A means to describe characteristics of an area without physically sampling the area.

Risk assessment For invasive species, the process of obtaining quantitative or qualitative measures of risk levels by incorporating a broad array of information describing factors that may influence the distribution of invasive species.

S

Sampling frame All the areas that may be sampled.

Sampling unit These are the independently selected units that will be searched. A sampling unit may be a stretch of trail, an off-trail transect, an area to be searched, or other convenient unit.

Scope The geographic and temporal frame of reference for a proposed project or activity.

Sensitivity A measure of the ability to predict the event of interest.

Simple random sampling Every unit of the population has the same probability of selection.

Spatial resolution The amount of detail an image contains across a given distance, typically a cell size.

Species A group of organisms, all of which have a high degree of physical and genetic similarity, generally interbreed only among themselves, and show persistent differences from members of allied groups of organisms (USPEO, 1999).

Species distribution models (SDMs) Statistical representations relating the likelihood of occurrence of a species to a set of predictor variables.

Specificity A measure of the ability to predict the nonevent.

Stakeholders Parties with that can influence and can be affected by management; these include, but are not limited to, State, tribal, and local government agencies, academic institutions, the scientific community, non-governmental entities including environmental, agricultural, and conservation organizations, trade groups, commercial interests, and private landowners (USPEO, 1999).

Standard Operating Procedure The components of a protocol that explain the step-by-step process to monitoring a vital sign.

Stochastic events Randomly occurring events that are not entirely predictable, such as floods, droughts, hurricanes, and so forth (Mack, 2003).

Stressors Physical, chemical, or biological perturbations to a system that are either (a) foreign to that system or (b) natural to the system but applied at an excessive [or deficient] level (Barrett and others, 1976:192). Stressors cause significant changes in the ecological components, patterns and processes in natural systems. Examples include water withdrawal, pesticide use, timber harvesting, traffic emissions, stream acidification, trampling, poaching, land-use change, and air pollution (NPS, I&M).

Susceptibility Measure of a given site's vulnerability to invasion, usually characterized by two key factors: the site's resource availability and local propagule pressure.

T

Threshold-dependent metrics Accuracy measures used in the model validation process whose values are determined by a threshold, or cutoff point that is used to distinguish between predicted data values considered correctly classified and those that have been incorrectly classified. Thresholds are often quite arbitrary in nature, so the method for choosing a threshold should be considered carefully.

Threshold-independent metrics Accuracy measures used in the model validation process whose values do not rely on a user-specified threshold but provide a model assessment based on different threshold levels.

Training data Data collected and used to build (that is, "train") the predictive model.

Trend As defined by NPS I&M: directional change measured in resources by monitoring their condition over time. Trends can be measured by examining individual change (change experienced by individual sample units) or by examining net change (change in mean response of all sample units) (NPS I&M).

U

Undesirable plants Plant species that are classified as undesirable, noxious, harmful, exotic, injurious, or poisonous, pursuant to State or Federal law. Species listed as endangered by the Endangered Species Act of 1973 (16 U.S.C. 1531 et seq.) shall not be designated as undesirable plants under this section and shall not include plants indigenous to an area where control measures are to be taken under this section (USC, 2000).

Unintentional When used with reference to the import, introduction, transport, or spread of species means the import, introduction, transport or spread of species incidental to another activity, including, without limitation, the import, introduction, transport, or spread of another species, or through pathways associated with the movement of goods, materials, or other articles in commerce and with the movement of any means of conveyance, or any other identified pathway for species invasion (Odell, 2004).

V

Valued site Any geographic location that has been designated by resource managers as having cultural and (or) natural resources worthy of protection.

Vector analysis The investigation of the supply of organisms (plants or propagules in this case) associated with particular transfer mechanisms, including variables that may influence the supply and the characterization of the organisms themselves (Ruiz and Carlton, 2003, p. 472).

Vectors The transfer mechanisms for plants or plant propagules (Ruiz and Carlton, 2003, p. 472).

Vital signs A subset of physical, chemical, and biological elements and processes of park ecosystems that are selected to represent the overall health or condition of park resources, known or hypothesized effects of stressors, or elements that have important human values. The elements and processes that are monitored are a subset of the total suite of natural resources that park managers are directed to preserve "unimpaired for future generations," including water, air, geological resources, plants and animals, and the various ecological, biological, and physical processes that act on those resources. Vital signs may occur at any level of organization including landscape, community, population, or genetic level, and may be compositional (referring to the variety of elements in the system), structural (referring to the organization or pattern of the system), or functional (referring to ecological processes) (NPS I&M).

Vulnerable site An area of interest that is likely to be invaded by nonnative species because of its disturbance history, availability of resources, proximity to propagules, location along dominant invasion pathways, and so forth. These sites may be a priority for monitoring efforts because of their elevated probability of invasion and, subsequently, their potential to serve as a source site for future invasions.

W

Weed In the broadest sense, any plant growing where it is not wanted. Weeds can be native or nonnative, invasive or noninvasive, and noxious or not noxious (Sheley and others, 1999).

References Cited

Barrett, G.W., Van Dyne, G.M., and Odum, E.P., 1976, Stress ecology: BioScience, v. 26, p. 192–194.

Elzinga, C.L., Salzer, D.W., and Willoughby J.W., 1998, Measuring and monitoring plant populations: BLM Technical Reference 1730–1, BLM/RS/ST–98/005+1730.

Likens, G., 1992, An ecosystem approach—Its use and abuse—Excellence in ecology, book 3: Oldendorf/Luhe, Germany, Ecology Institute, 147 p.

Lodge, D.M., Williams, S., MacIsaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, P.B., Smith, M., Andow, D.A., Carlton, J.T., and McMichael, A., 2006, ESA report—Biological invasions—Recommendations for U.S. policy and management: Ecological Applications, v. 16, p. 2035–2054.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M., and Bazzaz, F.A., 2000, Biotic invasions—Causes, epidemiology, global consequences and control: Ecological Applications, v. 10, p. 689–710.

Mack, R.N., 2003, Global plant dispersal, naturalization, and invasion—Pathways, modes, and circumstances, *in* Ruiz, G.M., and Carlton, J.T., eds., Invasive species—Vectors and management strategies: Washington, D.C., Island Press, p. 3–30.

National Invasive Species Council, 2003, General guidelines for the establishment and evaluation of invasive species early detection and rapid response systems, version 1, 16 p.

National Park Service, 2001, Management policies 2001: Washington, D.C., National Park Service.

Noon, B R., 2003, Conceptual issues in monitoring ecological systems, *in* Busch, D.E., and Trexler, J.C., eds., Monitoring ecosystems—Interdisciplinary approaches for evaluating ecoregional initiatives: Washington, D.C., Island Press, p. 27–71.

Oakley, K.L., Thomas, L.P., and Fancy, S.G., 2003, Guidelines for long-term monitoring protocols: Wildlife Society Bulletin, v. 31, p. 1000–1002.

Odell, T., 2004, Invasive species control—A comprehensive model state law: Environmental Law Institute, Washington, D.C.

Rejmanek, M., 2000, Invasive plants—Approaches and predictions: Australian Journal of Ecology 25, p. 497–506.

Ruiz, G.M. and Carlton, J.T., eds., 2003, Invasive species—Vectors and management strategies: Washington, D.C., Island Press, 484 p.

Sheley, R., Petroff, J., and Borman., M., 1999, Biology and management of noxious rangeland weeds: Corvallis, Oregon State University Press. 438 p.

U.S. Congress, 2000, Plant protection act: 7 U.S.C. §§ 7701–7772, June 20, 2000.

U.S. Presidential Executive Order, 1999, Executive order 13112 of February 3, 1999: Federal Register: February 8, 1999, v. 64, no. 25.

Williamson, M., 1996, Biological invasions: London, United Kingdom, Chapman & Hall, 256 p.

Publishing support provided by:
Denver Publishing Service Center

For more information concerning this publication, contact:
Program Coordinator for Status and Trends
Ecosystems Mission Area
U.S. Geological Survey
12201 Sunrise Valley Drive, MS 301
Reston, VA 20192

Or visit the Status and Trends Program Web site at:
http://www.usgs.gov/ecosystems/status_trends/

Welch and others, eds.—**Early Detection of Invesive Plants—Principles and Practices**—Scientific Investigations Report 2012–5162

ISSN 0196-1497 (print)
ISSN 2328-0328 (online)
http://dx.doi.org/10.3133/sir20125162