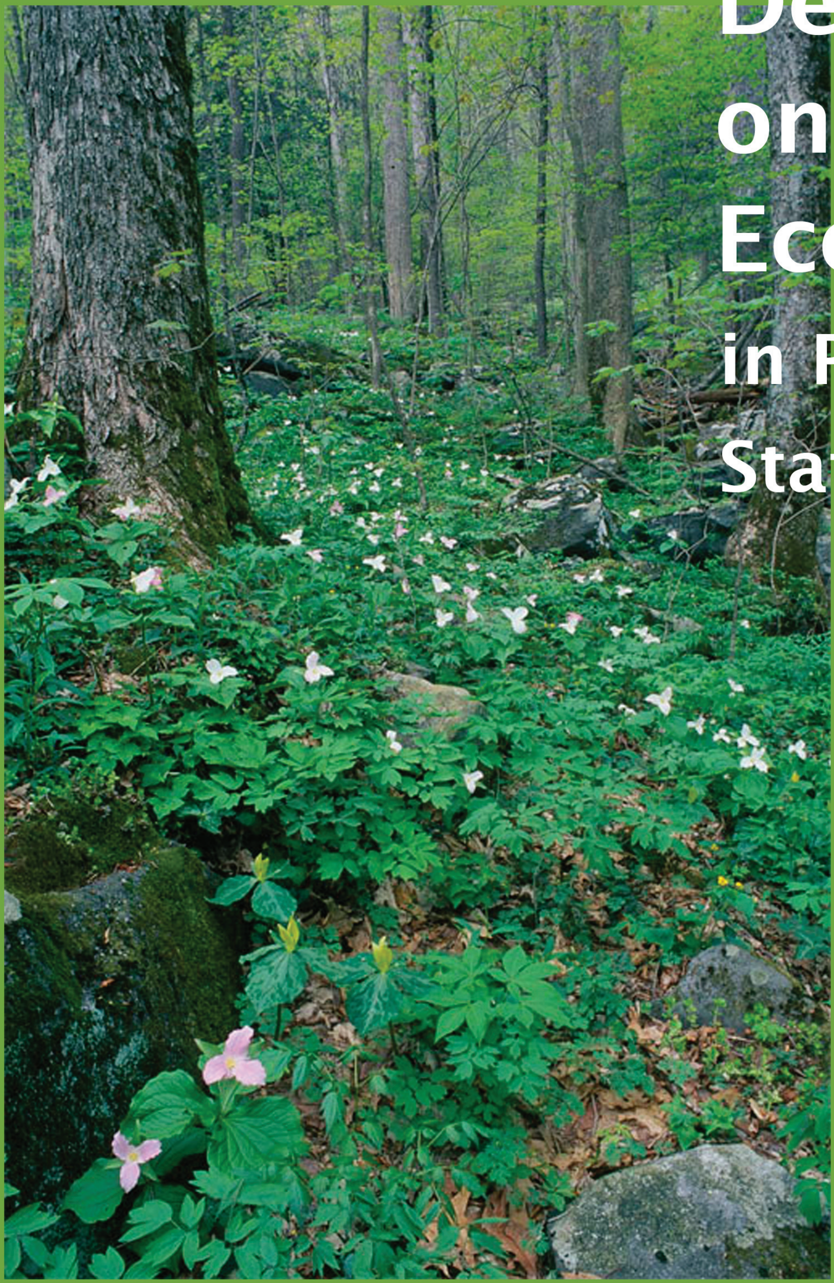
Monitoring Deer Effects on Forest Ecosystems in Pennsylvania State Forests



Roger Earl Latham (editor)
Marrett D. Grund
Stephen B. Horsley
Benjamin C. Jones
William H. McWilliams
Clayton K. Nielsen
Christopher S. Rosenberry
Robert S. Seymour
Bryon P. Shissler
Donald M. Waller

Research peer review and recommendations for
Pennsylvania Department of Conservation and Natural Resources, Bureau of Forestry

Cover photo by Bill Lea (used by permission of the U.S.D.A. Forest Service)

The photo depicts levels of native understory species diversity, structural complexity, and wildlife habitat value that have rarely been seen in Pennsylvania forests since the mid-twentieth century (this scene is in the mountains of western North Carolina). The authors of this report believe that a carefully considered, science-based, sustainable program of management can enable citizens all across the Commonwealth to enjoy the benefits of healthy forests once again.

Monitoring Deer Effects on Forest Ecosystems

in Pennsylvania State Forests

Roger Earl Latham (editor), Marrett D. Grund,
Stephen B. Horsley, Benjamin C. Jones,
William H. McWilliams, Clayton K. Nielsen,
Christopher S. Rosenberry, Robert S. Seymour,
Bryon P. Shissler and Donald M. Waller

Research peer review and recommendations
for the Pennsylvania Department of
Conservation and Natural Resources,
Bureau of Forestry

18 February 2009

Preface

This document summarizes a scientific peer review of three recently completed research studies commissioned by Dan Devlin, the State Forester of Pennsylvania, addressing issues pertinent to effects of white-tailed deer in forest ecosystems in the Commonwealth:

Benner, J. Merlin. 2007. Browsing and regeneration monitoring report for Pennsylvania's state forests, 2007. Pennsylvania Department of Conservation and Natural Resources, Bureau of Forestry. 21 pp.

deCalesta, David S. 2008. Deer density and impact [on the Kinzua Quality Deer Cooperative Area]. U.S.D.A. Forest Service, Northeastern Research Station, Forestry Sciences Laboratory. 27 pp.

Diefenbach, Duane R. and Richard S. Fritsky. 2007. Developing and testing a rapid assessment protocol for monitoring vegetation changes on state forest lands. U.S. Geological Survey, Cooperative Fish and Wildlife Research Unit, Pennsylvania State University. 47 pp.

A team of peer review organizers led by Roy Brubaker (Forest Resources Planner, D.C.N.R. Bureau of Forestry) circulated the manuscripts to ten reviewers, a mix of wildlife biologists and forest ecologists including academic researchers, government agency biologists, and consultants from Pennsylvania, Illinois, Maine, Minnesota, and Wisconsin (see list of participants, opposite).

The reviewers were asked to submit written comments consisting of:

- a critique of the methods and conclusions of each research effort, and
- recommendations on how the Bureau of Forestry should move forward, both in the short and long term, to achieve the greatest benefit in fulfilling D.C.N.R.'s commitment to ecosystem management, including:
    - suggested improvements in data collecting approaches and methods, and
    - advice on managing the state forests from an ecosystem perspective, focusing on deer effects but considering all pertinent forest ecosystem factors.

Reviewers' comments were compiled and summarized by theme, with the authorship of individual contributions kept anonymous. Summaries of critiques and recommendations were then distributed to the reviewers and researchers.

Eight of the reviewers met in conference on 6 May 2008, with Roy Brubaker acting as facilitator, assisted by Sara Nicholas (Policy Specialist, D.C.N.R. Office of Policy and Planning). The meeting began with short presentations by the three principal research investigators followed by a discussion among the reviewers, researchers, and organizers.

State Forester Devlin addressed the group, outlining the background of D.C.N.R.'s interest in effects of high deer density on the Commonwealth's forests and recapping his directive to the conferees. He requested practical recommendations on how the agency can enhance its ability to make informed management decisions, in particular, about where the Pennsylvania Game Commission's Deer Management Assistance Program (DMAP) should be targeted.

A lengthy roundtable discussion followed the researchers' departure. At the end of the group discussion, State Forester Devlin departed and the conferees split into two breakout sessions to discuss forest ecosystem and deer management issues separately. The meeting concluded with a wrap-up discussion among the entire group of the two smaller groups' conclusions.

Earlier versions of this document were circulated to all participants. Most contributed comments and suggested revisions, which are reflected in the present document.

List of Participants

Reviewers/Conferees

Roy Brubaker
(organizer, facilitator, conferee)
Pennsylvania Department of Conservation and Natural Resources
Bureau of Forestry
P.O. Box 8552
Harrisburg, PA 17105-8552
717-783-0382
robrubaker@state.pa.us

Marrett D. Grund, Ph.D.
(reviewer, conferee)
Minnesota State University
Department of Biological Sciences
308 Woodhaven Lane
Mankato, MN 56001
507-388-6163
marrettgrund@yahoo.com

Stephen B. Horsley, Ph.D.
(reviewer, conferee)
U.S. Forest Service
Forestry Sciences Laboratory
Old Route 6
Irvine, PA 16329-0267
814-563-1040
shorsley@fs.fed.us

Benjamin C. Jones, Ph.D.
(reviewer, conferee)
Pennsylvania Game Commission
Bureau of Wildlife Habitat Management
2001 Elmerton Avenue
Harrisburg, PA 17110
717-787-9613
benjjones@state.pa.us

Roger Earl Latham, Ph.D.
(co-organizer, reviewer, conferee, report editor)
Continental Conservation
P.O. Box 57
Rose Valley, PA 19086-0057
610-565-3405
rel@continentalconservation.us

William H. McWilliams, Ph.D.
(reviewer, conferee)
U.S. Forest Service
Forest Inventory and Analysis Program
11 Campus Boulevard, Suite 200
Newtown Square, PA 19073-3294
610-557-4075
wmcwilliams@fs.fed.us

Sara Nicholas
(co-organizer, conferee)
Pennsylvania Department of Conservation and Natural Resources
Office of Policy and Planning
P.O. Box 8767
Harrisburg, PA 17105-8767
717-772-9087
snicholas@state.pa.us

Clayton K. Nielsen, Ph.D.
(reviewer)
Southern Illinois University
Cooperative Wildlife Research Laboratory
124 Madison Road
Carbondale, IL 62901
618-453-6930
kezo92@siu.edu

Christopher Rosenberry, Ph.D.
(reviewer, conferee)
Pennsylvania Game Commission
Bureau of Wildlife Habitat Management
2001 Elmerton Avenue
Harrisburg, PA 17110
717-582-2988
chrosenber@state.pa.us

Robert S. Seymour, Ph.D.
(reviewer)
University of Maine
School of Forest Resources
5755 Nutting Hall
Orono, ME 04469-5755
207-581-2860
seymour@umenfa.maine.edu

Bryon Shissler
(co-organizer, reviewer, conferee)
Natural Resource Consultants, Inc.
1723 Fort Hill Road
Fort Hill, PA 15540
814-395-5335
nrc@qcol.net

Donald M. Waller, Ph.D.
(reviewer, conferee)
University of Wisconsin–Madison
Department of Botany
430 Lincoln Avenue
Madison, WI 53706
608-263-2042
dmwaller@wisc.edu

Researchers/Presenters

J. Merlin Benner
(formerly Pennsylvania Department of Conservation and Natural Resources, Bureau of Forestry)
Micro Management Resource Consultants
2780 Hillscreek Road
Wellsboro, PA 16901
570-376-2255
merlinb@ptd.net

David S. deCalesta, Ph.D.
(formerly U.S. Forest Service, Forestry Sciences Laboratory)
Wildlife Analysis PC
P.O. Box 621
Hammondsport, NY 14840
607-292-6078
daviddecalesta@yahoo.com

Duane R. Diefenbach, Ph.D.
U.S. Geological Survey
Cooperative Fish and Wildlife Research Unit
Pennsylvania State University
School of Forest Resources
404 Forest Resources Building
University Park, PA 16802
814-865-3992
ddiefenbach@psu.edu

Acknowledgments

D.C.N.R. extends special thanks to the executive officers at cooperating natural resource agencies for fostering the success of this scientific peer review by consenting to participation by distinguished scientists on their staffs—two representatives from the Pennsylvania Game Commission and two from the U.S. Forest Service.

Table of Contents

Executive Summary

In the early 2000s, the Pennsylvania Department of Conservation and Natural Resources (D.C.N.R.) began to address white-tailed deer effects on its state forest lands by intensifying recreational hunting effort in areas where impacts have been most severe. It did this by enrolling select areas of state forest in the Pennsylvania Game Commission's (P.G.C.) Deer Management Assistance Program (DMAP). Although D.C.N.R. does not set deer management regulations—the purview of the Game Commission—its mission to conserve multiple resource values on its forest lands has led D.C.N.R. into a partnership with the P.G.C. to concentrate hunters' efforts where they are most needed. The agency's goal is to sustain a healthy and functioning forest ecosystem, including the ability of forests to regenerate with the full array of native species, by balancing the deer herd with its habitat across state forest lands.

To assess the effectiveness of its targeted deer management efforts, D.C.N.R. commissioned three research studies of deer effects on the Commonwealth's forests. The aim was to help develop and test effective monitoring protocols for tracking key indicators of management success and of the need for further management action. D.C.N.R. undertook a formal evaluation of its commissioned studies, using outside experts as reviewers. By subjecting its research to the objective scrutiny of qualified biologists, including experienced forest and wildlife management professionals from outside D.C.N.R., the agency aspires to get the most out of its investment through short-term and long-term improvements to its ongoing habitat monitoring efforts.

The white-tailed deer is a keystone species, not only ecologically, but also politically and socially. The Game Commission's great success in bringing deer back from the brink of extinction in the early twentieth century had unforeseen and unintended consequences. With ample food and no effective predators other than recreational hunters constrained by seasons and bag limits, deer populations recovered quickly and then skyrocketed, resulting in severe impacts on Pennsylvania's forests. At the same time, a deer hunter culture arose with an expectation of limitless hunting opportunities, and enough political clout to override concerns about forest health. The Game Commissions's adoption of DMAP reflects a growing recognition that some responsibility for deer management lies with individual landowners, including D.C.N.R. The three research projects that are the subject of this peer review germinated out of D.C.N.R.'s acknowledgment of its share of deer management responsibility. The primary goal in supporting research of this kind is to enhance the agency's ability to make informed management decisions, in particular, about where DMAP should be targeted in the state forest system.

D.C.N.R.'s commitment to ecosystem management will entail restoring and sustaining, in addition to timber regeneration:

- native species diversity (including noncommercial and nongame species),
- habitat quality for the full range of locally indigenous wildlife species,
- vertical structure (robust ground, shrub, and subcanopy layers),
- patch or stand type diversity,
- fine-scale structural and age class heterogeneity, and
- other essential forest ecosystem qualities and processes.

Forest inventory and monitoring is a huge investment of time and effort. Recognition of this fact was the impetus for D.C.N.R. to contract with the Pennsylvania Cooperative Fish and Wildlife Research Unit to devise a more efficient sampling scheme—a rapid habitat assessment method using the indicator approach (the subject of one of the three reviewed studies). The agency is operating under numerous constraints in meeting the need for ecosystem monitoring in its transition from silvicultural management to forest ecosystem management. The staff consists mainly of foresters, with few wildlife biologists or plant ecologists. The operational emphasis is still on timber management as a first step toward achieving ecosystem management goals while the policymakers grapple with how to address ecosystem-level problems that cannot be solved by harvest scheduling and other timber management strategies alone. Staff time and funds are increasingly limited, in part because timber markets have declined and much of D.C.N.R.'s budget is tied to timber revenues. Steps taken toward managing the state forests from an ecosystem perspective must be practical (financially and otherwise) and efficient. The success of the forest monitoring program will depend in large measure on the degree to which district staff find the data useful in carrying out or improving the rest of their work.

The three research efforts that are the basis of this review are landmark studies, the first to address the connection between deer population reduction and forest ecosystem recovery since the Pennsylvania Game Commission began offering landowners additional tools, including DMAP, to reduce deer numbers on their properties if they desire to do so. The objective of "Browsing and regeneration monitoring report for Pennsylvania's state forests" (Benner 2007), conducted internally by D.C.N.R. staff, was to assess the current levels of deer browsing and tree regeneration across the entire 2.1-million-acre state forest system in order to allocate and prioritize the use of DMAP and other deer management options. "Deer density and impact on the Kinzua Quality Deer Cooperative Area" (deCalesta 2008) compared yearly deer population estimates and surveys of browsing intensity indicators on a 115-square-mile forested area enrolled in DMAP, seeking trends and correlations that might link changes in density with changes in impact. In "Developing and testing a rapid assessment protocol for monitoring vegetation changes on state forest lands" (Diefenbach and Fritsky 2007), the objective was a forest vegetation survey protocol that could be completed quickly across large areas. The protocol was designed to measure vegetation characteristics likely to respond to changes in deer browsing intensity whose measurement is cost-effective; it was tested in DMAP areas in eleven forest districts, totaling 311 square miles.

This report summarizes the three studies, the reviewers' written comments, the discussions at the review meeting among reviewers, researchers, and D.C.N.R. staff, and the review team's recommendations for near-term and long-term enhancement of state forest monitoring from an ecosystem perspective. It does not dwell on the reviewers' critiques of the researchers' methods or conclusions but mentions them in cases where a specific criticism may help to clarify a larger issue. The peer reviewers' comments and recommendations reflect a small number of recurring themes, each of which is the subject of a section of the report.

Agree on a vision for the future with measurable objectives.

Recommendation: Establish science-based objectives quantifying the desired future condition of the state forests. D.C.N.R.'s overall vision and goals for the state forest system and how they translate into quantitative management objectives need to be clarified first, because they are the foundation for everything else. Currently, D.C.N.R.'s ecosystem management (including deer management) goals are general, qualitative, and subjective and vary to some extent from document to document. If the intent is to sustainably manage state forests from an ecosystem perspective, these goals need to be made consistent and translated into objectives that are specific, quantitative, and can be applied uniformly. The objectives must be firmly rooted in a detailed, quantitatively focused consensus on the desired future condition of the 2.1 million acres of state forest land. There was wide agreement among reviewers that day-to-day and year-to-year actions to restore and sustain healthy forests should be based on a clear, scientifically sound vision of the desired outcome. The first goal in the current state forest deer management plan is

> Sustain a healthy and functioning forest ecosystem, including the ability of forests to regenerate with desirable species, by balancing the deer herd with its habitat across state forest lands.

However, there is currently no consensus on what "a healthy and functioning forest ecosystem" means quantitatively. In order for coordinated, forward progress to be made in ecosystem restoration and management, D.C.N.R. management staff—in collaboration with scientists and other stakeholders—need to define what the target ecosystems should look like and how they should function, and then link those qualitative goals with quantitative objectives.

Desired future condition (D.F.C.) analyses are part of an emerging science-based approach to ecosystem management by the U.S. Forest Service, National Park Service, and other large-scale land management agencies. The process ties together adaptive resource management, ecological restoration, integrated planning, ecosystem monitoring, and condition reporting. D.F.C. analysis may be defined as the qualitative and quantitative description of *ecosystem attributes* that are expected to be present at some point as an outcome of deliberate management policies, strategies, and practices. Ecosystem attributes include individual resources, communities, ecosystems, and the natural processes that sustain them. Formulating a D.F.C. analysis helps resource managers to be proactive rather than reactive. The process spurs them to identify trade-offs between competing resources or goals. D.F.C analysis is not an attempt to return to the past. It takes into account both what is known about the pre-degradation condition and important influences that have been added or taken away since then and are beyond managers' control. Striving toward ultimate goals and the quantitative objectives related to those goals calls for many steps—small, medium, and large—before those goals are met. Setting those goals and objectives provides the yardstick to measure the value of steady progress. Success comes not only when a quantitative objective is fully achieved, but also in reaching various mileposts along the way. For instance, when areas no longer need deer-deterrent fences to foster regeneration, D.C.N.R. and its partners can celebrate a major success. Reaching objectives related to certain other goals, such as the return of understory species diversity and habitat quality, will take longer.

Take stock before going on.

Recommendation: Evaluate and reanalyze existing monitoring data before pursuing the next round of data collection in the field. Some members of the review team suggested that a "starter" set of indicators might be extracted from the mountain of data that has already been gathered in Pennsylvania's forests and similar forests in nearby states, using analysis methods not previously applied to this problem. They argued that a comprehensive analysis of existing data from multiple studies is the most credible and realistic way to determine whether existing data offer useful insights and, if so, which indicators are effective and which should be dropped. Several members of the panel expressed serious reservations about spending further time and money on indicator data already collected in the state forests (summarized in Benner 2007), maintaining that the effort would not translate into any meaningful management actions. Diefenbach and Fritsky (2007) made a useful start at an appropriate choice of indicators by weighing the costs and benefits of using certain herbaceous understory species. Other studies in the state forests such as the Bureau of Forestry's continuous forest inventory (C.F.I.) program were not designed to focus on indicators of deer effects on forest ecosystems but may nonetheless have produced pertinent data. Studies conducted elsewhere may be more relevant, and should be revisited in any case in the course of designing a new monitoring protocol. If reanalysis of existing data yields useful results, it could provide a basis for DMAP prioritization on state forests and lay the groundwork for standardizing, in the near future, which indicators and threshold values should be used to determine the need for DMAP enrollment. If the data fail to provide useful insights, then assessing the effectiveness of DMAP in meeting D.C.N.R.'s forest management goals will be delayed by the need for additional effort in designing, implementing, and testing a new, effective monitoring protocol.

Upgrade to "Forest Monitoring 2.0."

Recommendation: Design the next monitoring protocol, incorporating additional indicators of deer effects and forest recovery. A well-designed monitoring protocol is one of the keys to a successful ecosystem management program. A crucial part of an effective, highly sensitive monitoring protocol is the choice of indicators. The next iteration should cast a wider net beyond the indicators used in prior studies. Additional indicators from the scientific literature and expert opinion will strengthen the effectiveness of monitoring. The choice of indicators should be widened to include measurements that reflect vertical structure, patch diversity, noncommercial species regeneration, and other essential forest ecosystem qualities and processes in addition to timber regeneration and native species diversity. Certain non-deer influences that may modify how deer affect forest ecosystems also should be included: overstory conditions, recent timber management, deer fencing history, soil buffering capacity, and cover of competing vegetation that is disproportionately abundant as a legacy of prolonged high deer density, such as hay-scented fern. A mixture of rapid-response and slow-response indicators is vital. Monitored sites should include a network of existing and new deer exclosures to provide benchmark data. Given a scarcity of funds and time, the Bureau of Forestry may need to prioritize relatively few sites in the short term for intensive deer management—where early success is judged most likely—and add more sites later. Likewise, a triage approach may be appropriate in prioritizing where and when monitoring is conducted in the short term. Districts that appear to have the biggest problems might be dealt with first and the lessons learned there applied in other areas later.

Discriminate deer from non-deer effects.

Recommendation: Design monitoring methods to distinguish effects of high deer density from other major influences on forest recovery. Some individuals use the issue of non-deer influences on forest regeneration to question the credibility of forest monitoring and management in Pennsylvania. The reviewers agreed that a vital first step in improving "social credibility" is to acknowledge potentially important non-deer factors that affect regeneration and incorporate them into monitoring procedures. Potentially important factors include shade at the forest floor, forest type, soil buffering capacity, depauperate seed bank, scarcity of live seed sources, disproportionate understory abundance of unpalatable plants, and disproportionate abundance of nonnative, invasive species. These factors can be taken into account most efficiently via a mixed approach of manipulation and well-chosen locations for monitoring. Considering influences on forest recovery unrelated to effects of current deer density can improve operational efficiency in some cases; for instance, it may be possible to reduce the data collection effort by about half by omitting plots without sufficient light for regeneration to occur.

Measure deer effects, not deer numbers.

Recommendation: Focus on indicators of deer effects on forest ecosystems and avoid estimating deer density, but consider developing site-specific relative indices of annual deer density *change*. The importance of estimating deer numbers is regularly debated. It is the subject that generated the most disagreement among participants in this review. All agreed that indicators of deer effects are essential. However, wildlife biologists on the panel recommended that D.C.N.R. discontinue efforts to estimate deer density. They maintained that current methods have low scientific credibility and pose a high risk of diverting scarce time and resources from deer impact monitoring. The plant biologists and forest ecologists would like D.C.N.R. to collect data for a site-specific *relative index of deer density change* over time—percent change from a baseline survey of pellet-groups or aerial thermal image “hotspots,” or reported annual harvest rates by hunters. They believe indexing annual percent change on a site-by-site basis will be a key part of determining whether a non-response by indicators is due to non-deer effects, or legacy effects of prolonged high deer numbers, or inadequate deer herd reduction (although exclosures and herbicide-treated plots, where practical, can also be used to distinguish among these alternatives). Reviewers agreed that focusing on deer density is unsound using the current methods, which are based on flawed assumptions and, when tested, have been shown to have low accuracy. More fundamentally, according to several reviewers it is counterproductive to focus on deer density in any case because it diverts attention away from the core issue—deer-ecosystem interaction—and fosters the false impression that there is a broadly applicable relationship between specific ranges of deer density and forest health.

Bring politics and science into closer alignment.

Recommendation: Frame and test key management questions in an adaptive resource management (A.R.M.) context. The review team strongly supports shifting to A.R.M. in state forest management in general and deer management in particular. Resolving policy controversy and management impasse is a focus of A.R.M. It provides a framework for making science-based decisions in the face of critical uncertainties, and a formal process for reducing uncertainties so management performance can improve over time. Effective monitoring is a critically important part of A.R.M. Furthermore, A.R.M. requires setting target ranges of indicator values that differentiate “acceptable” from “unacceptable,” recognizing every indicator has a natural range of variation. Uncertainty can then be expressed as a set of testable models describing how the deer-forest system might respond under different management scenarios. Monitoring at multiple scales provides the data to test these models’ predictions. Implementing A.R.M. requires a strong commitment by executive leadership to a transparent decision-making process, built on measurable objectives and feedback from a credible and sustainable monitoring system. A.R.M. also requires engagement by all major stakeholders, including the staunchest opponents. Conferees noted that D.C.N.R.’s Ecosystem Management Advisory Committee (EMAC) already includes representatives of many major stakeholder groups.

Get the most out of D.C.N.R.’s monitoring investment.

Recommendation: Foster confidence in the results of monitoring and adaptive resource management by achieving the highest practical standards of scientific rigor. The key operational differences between A.R.M. and scientific research involve standards of evidence, degree of experimental control, and thoroughness of replication; however, no clear line demarcates the two approaches. There is a gradation between controlled, fully replicated experiments, which tend to be more expensive, and trials within A.R.M., which tend to be less expensive because they are incidental to funds and time expended on management anyway. Several reviewers expressed the opinion that both A.R.M. and rigorous experiments are needed and should work together to advance knowledge of the best practices for managing forests from an ecosystem perspective. The more closely the design and monitoring of management trials approach the rigor of a scientific experiment, the higher the quality of information. There are some tradeoffs between short-term cost and long-term effectiveness; however, it is not necessarily a linear relationship. Instituting small tweaks can sometimes result in large gains. Several reviewers maintained that major improvements could be made at little added cost by configuring monitoring plots and management trials across the landscape in ways that better meet experimental design standards.

Put D.C.N.R.'s use of DMAP to the test.

Recommendation: Test the assumption that D.C.N.R.-administered DMAP is an effective tool for sustainable forest management to the degree required to meet ecosystem management goals. D.C.N.R.'s goal for state forest lands is to "sustain a healthy and functioning forest ecosystem," therefore they may require a lower threshold of deer density compared to some other public and private landholders. The capability of existing deer management tools, DMAP in particular, to meet these stringent needs has not yet been demonstrated. Because the state forest management program is predicated on having effective tools, it is critical to test and, if necessary, improve the efficacy of DMAP and how it is applied in the state forests. D.C.N.R. needs to determine (1) how effective DMAP is at increasing antlerless deer harvests in the state forests, (2) how its effectiveness can be increased, (3) whether, when applied effectively, it reduces deer impacts to the degree necessary to meet D.C.N.R.'s goals, and (4) if not, then what its specific limitations are and how they can be remedied. One reason to question DMAP's effectiveness in meeting D.C.N.R.'s forest management goals concerns the issue of hunter access. Hunting effort typically varies from place to place within a DMAP area, as anywhere else, depending on ease of access. Reviewers recommend that D.C.N.R. explore ways of enhancing the spatial distribution of hunting opportunity in order to improve the hunter access situation in DMAP areas.

Most reviewers support the recommendation that D.C.N.R. promote a forest restoration study at the scale of a fully replicated and controlled experiment (in parallel with implementation of A.R.M. on all state forest lands not designated as part of the study), arguing that the stakes are too high and the need too urgent to wait for the typically lengthier, somewhat less powerful and riskier process of A.R.M. to answer the most basic question—Can current tools meet D.C.N.R.'s ecosystem management objectives? The forest restoration study could be essentially similar to Diefenbach and Fritsky's study but on a larger scale. It would include replicated areas (with exclosures) in which deer density is reduced using DMAP and an equal number of control areas (with exclosures) where ordinary hunting regulations apply, distributed widely across the state forest system. In order to weigh the relative magnitudes of deer and non-deer effects and address interactions among significant effects, both types of areas would need to be stratified by additional management treatments besides deer management (e.g., augmenting the seed supply of native understory species, herbiciding dense rhizomatous fern cover, burning to foster oak regeneration, or liming to restore depleted soil calcium). The forest restoration study will be a significant expense but its advocates, the majority of the panel, believe it is key to D.C.N.R.'s timely success in meeting its ecosystem management goals and thus will save resources in the long term. The knowledge gained will have significance beyond management of the Pennsylvania state forest system and so funding from sources in addition to state appropriations may be accessible through partnerships.

Next steps

State Forester Devlin asked the review team for a set of short-term recommendations, emphasizing the need to recognize funding limitations and the importance of cost efficiency. Since then, the economic downturn has made funding constraints even more severe. However, making significant progress toward D.C.N.R.'s ecosystem management goals in the state forests depends on procuring resources and help from partners to carry out key tasks and implementing certain changes to present methods. Some of the tasks will be somewhat costly in the short term but are considered likely to increase cost efficiency appreciably in the long term. Some of the procedural changes represent both immediate and long-term cost savings.

- Elevate the process of conducting a desired future condition analysis for Pennsylvania's state forests to top priority. Begin by establishing a task force to deliberate on potential sources of funds, personnel needs, and a timetable.

- Undertake a desired future condition analysis consisting of:
  - forest-type-specific descriptions of the conditions of resources, communities, ecosystems, and the natural processes sustaining them that are the target outcomes of management; and
  - a comprehensive tally of quantitative, measurable thresholds of indicator responses that will signal success in achieving target outcomes.
- Commission qualified statisticians to evaluate existing data on potential indicators from all available sources, including one or more of the studies that are the subject of this report, the U.S. Forest Service's Pennsylvania Regeneration Study (P.R.S.), D.C.N.R.'s continuous forest inventory (C.F.I.) program, other pertinent studies including long-term monitoring of deer exclosures, and any relevant data on the costs associated with the various monitoring tasks performed in all of the studies. If the data meet requisite standards, conduct analyses to weigh each indicator's (or group of indicators') predictive power and measurement cost using methods such as path analysis and ordination.
- Use the results of existing data reanalysis to choose the most promising indicators; incorporate additional candidate indicators suggested by the scientific literature and expert opinion, including a mixture of rapid-response and slow-response indicators; and design a monitoring protocol to measure them as efficiently as possible.
- Change the monitoring period to summer, including browse impact monitoring.
- Reduce the amount and types of monitoring data collected in areas where the forest canopy or fern cover is too dense to allow significant indicator responses.
- Inventory, establish, maintain, and monitor permanent deer exclosures to distinguish between deer and non-deer effects and to provide benchmark data needed to refine protocols and set quantitative goals.
- Stratify plots in the next cycle (and later monitoring cycles) to enable comparisons of indicator response inside and outside of exclosures and inside and outside of areas where deer densities are reduced using DMAP or other population-control measures.
- Suspend efforts to estimate deer density until such time as scientifically defensible methods have been developed that are cost-effective.
- Explore ways of improving hunter return rate of DMAP harvest reports to a consistently high percentage of the total harvest across all units in the state forest system. Use the data to develop a scientifically credible site-specific relative index of annual deer density change.
- Facilitate the learning process by configuring monitoring plots and management trials as near to rigorous experimental design standards as can be achieved with available resources.
- Commit to shifting to an adaptive resource management approach in the state forests by establishing a team of D.C.N.R. administrators, managers, scientists, and experienced A.R.M. practitioners to plan, coordinate, and oversee the transition.
- Establish a deer and state forest A.R.M. stakeholder group to help assess challenges, design management activities to address them, and implement and monitor those activities, and to participate in evaluation of results.
- Foster the political will and access to additional funds that will be necessary to carry out a forest restoration study as a rigorous scientific experiment by persuading policymakers of the urgency of resolving fundamental questions about relative magnitudes of deer and non-deer effects in forest degradation, effectiveness of DMAP at increasing antlerless deer harvests in the state forests, and practical means for D.C.N.R. to meet its goal to "sustain a healthy and functioning forest ecosystem" across the state forest system.
- Launch a research partnership among scientists and managers at D.C.N.R., P.G.C., U.S. Forest Service, and Pennsylvania Cooperative Fish and Wildlife Research Unit to collaborate on a forest restoration

study, to include examination of DMAP's effectiveness at reducing deer impacts in state forests and ways of enhancing the spatial distribution of hunting opportunity in order to improve the hunter access situation in DMAP areas.

Introduction

Background and charge to reviewers

In the early 2000s, the Pennsylvania Department of Conservation and Natural Resources (D.C.N.R.) began to address effects of high white-tailed deer density on its state forests by intensifying recreational hunting in areas where impacts were most severe. It did this by enrolling select areas of state forest in the Pennsylvania Game Commission's (P.G.C.) Deer Management Assistance Program, commonly known as DMAP. Although D.C.N.R. does not set deer management regulations—the purview of the Game Commission—its mission to conserve multiple resource values on its forest lands has led D.C.N.R. into a partnership with the P.G.C. to concentrate hunters' efforts where they are most needed. The agency's goal is to

> Sustain a healthy and functioning forest ecosystem, including the ability of forests to regenerate with desirable species, by balancing the deer herd with its habitat across state forest lands. [Pennsylvania D.C.N.R. Bureau of Forestry 2006, page 2]

Seeking to improve the effectiveness of monitoring and targeted deer management efforts, D.C.N.R. commissioned three research studies of deer effects on the Commonwealth's forests, an internal study using D.C.N.R. field staff and two by outside consultants. The aim was to help develop and test effective monitoring protocols for tracking key indicators of management success and of the need for additional management action. D.C.N.R. undertook a formal evaluation of its commissioned studies in April 2008, using outside experts as reviewers. By subjecting its research to the objective scrutiny of qualified biologists, including experienced forest and wildlife management professionals from outside D.C.N.R., the agency aspires to get the most out of its investment through short-term and long-term improvements to its ongoing habitat monitoring efforts.

The white-tailed deer is a keystone species, not only ecologically, but also politically and socially. The P.G.C.'s great success in bringing deer back from the brink of extinction in the early twentieth century had unforeseen and unintended consequences. With ample food and no effective predators other than recreational hunters constrained by seasons and bag limits, deer populations recovered quickly and then skyrocketed, resulting in severe impacts on Pennsylvania's forests. At the same time, a deer hunter culture arose with an expectation of limitless hunting opportunities, and enough political clout to override concerns about forest health. In the late 1990s and early 2000s, P.G.C. administrators made more progress on introducing forest health into the deer management equation in one decade than in many decades before then.

The P.G.C.'s adoption of DMAP reflects a growing recognition that some responsibility for deer management lies with private landowners, including D.C.N.R. With this responsibility come expectations, namely, that D.C.N.R. should have the knowledge, skills, and ability to make credible decisions. Dan Devlin, the State Forester of Pennsylvania, acknowledges that this is an uncomfortable position; it is easy to be a critic but much harder to make progress in solving the problems. The three research projects that are the subject of this peer review germinated out of D.C.N.R.'s acknowledgment of its share of deer management responsibility. D.C.N.R.'s goal in supporting research of this kind is to enhance its ability to make informed management decisions, in particular, about where DMAP should be targeted in the state forest system.

Currently, deer management in the state forests is geared almost entirely toward allowing forest trees to regenerate. In addition to its participation in DMAP, D.C.N.R. now spends approximately $3 million each year on temporary deer fencing to allow tree regeneration on state forest lands following timber harvest. D.C.N.R.'s commitment to ecosystem management will also entail restoring and sustaining:

- native species diversity (including noncommercial and nongame species),
- habitat quality for the full range of locally indigenous wildlife species,
- vertical structure (robust ground, shrub, and subcanopy layers),
- patch or stand type diversity,
- fine-scale structural and age class heterogeneity, and
- other essential forest ecosystem qualities and processes.

Forest inventory and monitoring is a huge investment of time and effort. Recognition of this fact was the impetus for D.C.N.R. to contract with the Pennsylvania Cooperative Fish and Wildlife Research Unit to devise a more efficient sampling scheme—a rapid habitat assessment method using the indicator approach (the subject of one of the three reviewed studies). Given the agency's budgetary constraints, in developing its forest ecosystem monitoring protocol D.C.N.R. is, in State Forester Devlin's words, "not looking for a Cadillac; we're looking for a Chevy." The agency needs information to make decisions now. Among the challenges in translating monitoring data into information that is useful on the ground is the vast area and large amount of variability in the 2.1-million-acre state forest system. Statewide information is useful for some purposes but it is too broad-scale to help managers in the field. By some measures Pennsylvania's state forests may be in fairly good shape from a statewide perspective, but in several regions the picture is not so bright.

D.C.N.R. is operating under numerous constraints in meeting the need for ecosystem monitoring in its transition from silvicultural management to forest ecosystem management:

- The staff consists mainly of foresters, with few wildlife biologists or plant ecologists. Data collection and management are likely to be conducted mainly by foresters for the foreseeable future. The operational emphasis is still on timber management as a first step toward achieving ecosystem management goals while the policymakers grapple with how to address ecosystem-level problems that cannot be solved by harvest scheduling and other timber management strategies alone. Meeting timber regeneration goals has long been the top priority for policymakers and field staff alike. It will be a significant challenge to refocus institutional policy and resources toward addressing the complex problems affecting forest ecosystems, including those that Wild Areas, Natural Areas, and Wild Plant Sanctuaries were set up to protect.
- Staff time is limited. D.C.N.R.'s Bureau of Forestry has been mobilizing its entire forestry and technician staff for three full weeks per year to carry out the monitoring surveys.
- Funds are increasingly limited. Timber markets have declined and much of D.C.N.R.'s budget is tied to timber revenues. A surplus built up in earlier, more prosperous times is gone. Recently, outlays have exceeded income for first time in many years. The agency is expected to run a deficit or be forced to downscale in 2009.

Steps taken toward managing the state forests from an ecosystem perspective must be practical (financially and otherwise) and efficient. "Buy-in" by district staff is of key importance to the success of the forest monitoring program. Staff members' enthusiasm for the extra labor depends in large measure on the degree to which they find the data to be useful in carrying out or improving the rest of their work.

The three studies

The three research efforts that are the basis of this review are landmark studies, the first to address the connection between deer population reduction and forest ecosystem recovery since the early 2000s, when the Pennsylvania Game Commission began offering landowners additional tools, including DMAP, to reduce deer numbers on their properties if they desire to do so. The manuscripts (Benner 2007; deCalesta 2008; Diefenbach and Fritsky 2007) and the researchers' presentations highlighted some valuable lessons learned from their work, which were acknowledged by members of the peer review team.

Browsing and regeneration monitoring report for Pennsylvania's state forests (Benner 2007). The objective of Benner's study, conducted internally by D.C.N.R. staff, was to assess the current levels of deer browsing and tree regeneration across the entire state forest system in order to allocate and prioritize the use of DMAP and other deer management options. Surveys were conducted in late winter 2006 and 2007 in nearly 75,000 vegetation sampling plots (113 square feet) at 200-foot intervals along transects spaced two miles apart, and at deer pellet-group survey plots (113 square feet) at 100-foot intervals along the same transects. Vegetation data consisted of the woody species present, browsing intensity category, and a subjective classification of the seedling density of "desirable" tree species as *adequate* or *inadequate* for regeneration.

About one-quarter of plots system-wide were classified as having adequate regeneration of desirable tree species. On the 74% of plots that had woody species present in the understory, an average of 85% of woody stems were either not browsed or classed as "lightly" browsed, although browsing intensity varied considerably among and within forest districts. Comparisons of 2007 data with 2006 data to detect one-year trends were problematic because of changes in data collection methods between the two years and the fact that different (but overlapping) sets of plots were surveyed in each year, which confounds any year-to-year change with variation due to spatial heterogeneity. Deer pellet-count data were not analyzed.

Benner and the many Bureau of Forestry staff members who cooperated in carrying out his study demonstrated that large-scale monitoring of deer effects and ecosystem recovery indicators is feasible over a large area, specifically the entire 2.1-million-acre state forest system. Benner underscored that achieving sustainability and measuring success of forest ecosystem recovery depend on monitoring the habitat itself and it is a lower priority to spend time monitoring related factors that are one or more steps removed from forest health, such as deer density. Regarding methodology, he concluded that there is a need to consolidate transects to minimize travel time so that more plots can be monitored for a given amount of time and effort. He identified weaknesses in the study's implementation, including inconsistencies among the approximately 150 data collectors, different intensities of data collection among districts, and some shifting of monitored areas from year to year by local staff. Benner identified the goal to develop a model that each year's indicator measures can be plugged into, annually updating the basis for decision making. He envisioned the model also helping to pinpoint what unknowns still need to be resolved and what additional data need to be collected in order for the Bureau of Forestry to be effective and efficient in carrying out its mission of ecosystem management in the state forests.

Deer density and impact on the Kinzua Quality Deer Cooperative Area (deCalesta 2008). This study's objective was to compare yearly deer population estimates and surveys of browsing intensity indicators on a large forested area enrolled in DMAP, seeking trends and correlations that might link changes in density with changes in impact. The study area was the 115-square-mile Kinzua Quality Deer Cooperative Area (K.Q.D.C.) in western McKean County. Results of six annual surveys, 2002–2007, were compared. Deer density data consisted of early spring counts of fecal pellet groups deposited over winter. Counts were conducted on plots (50 square feet) at 100-foot intervals along five 1-mile-long transects in 26 randomly located 1-square-mile survey blocks spread across the K.Q.D.C. Deer impact data were collected at the same time on one-half of the pellet-group survey plots (at 200-foot intervals along the transects). Six kinds of tree seedlings—American beech, striped maple, red maple, black cherry, yellow and black birches combined, and eastern hemlock—were categorized by five levels of deer browsing intensity. Estimates of deer density were calculated from the pellet-group data using an assumed deposition rate of 25 pellet groups per animal per day, the time elapsed since a day in the previous fall designated as the average date of leaf-off, and the ground surface area surveyed.

Estimated deer density was highly variable among survey blocks but annual means across the entire K.Q.D.C. showed a decreasing trend from 21–32 deer per square mile in the three years before the area was enrolled in DMAP in 2004 to 10–16 deer per square mile in the first three years of the program (ranges are aggregate 95% confidence intervals over three years). Trends in measured indicators of deer effects were less pronounced, ranging from 52%–67% of plots with no tree regeneration before DMAP to 47%–57% during the program's first three years. Plots designated as "no impact"—that is, lacking signs of recent deer browsing—showed an increasing trend from 6%–20% of total plots before DMAP to 33%–39% in the three years after DMAP implementation (it is not known to what degree these data were influenced by the scarcity or absence of woody plants on which to assess browsing intensity on some proportion of the plots, a legacy of over 50 years of high deer densities).

DeCalesta's K.Q.D.C. study was the first attempt at a multi-year program to monitor deer population changes in response to DMAP over a large landscape and to relate the population changes to changes in a browsing indicator. This and a concurrent study by the same author (deCalesta 2007) highlighted the complexity and enumerated the many points of vulnerability of the pellet-group counting method for estimating deer population density.

Developing and testing a rapid assessment protocol for monitoring vegetation changes on state forest lands (Diefenbach and Fritsky 2007). This study's objectives were to develop a forest vegetation survey protocol that could be completed quickly across large areas and to test it in state forest areas that are enrolled in DMAP. The protocol was designed to measure vegetation characteristics likely to respond to changes in deer browsing intensity whose measurement is cost-effective. The field tests assessed whether using the protocol would yield data with enough precision to detect reasonable levels of change over time. They were carried out in DMAP areas in eleven forest districts, totaling 311 square miles. Data included tree basal area and diameter at breast height by species; stem densities of shrubs, tree seedlings, and saplings by species; presence or absence of browsing on each tree seedling; counts and maximum heights of four herbaceous plant taxa (Indian cucumber-root, trilliums, Canada mayflower, and jack-in-the-pulpit); and percent cover of each of four categories of plants (*Rubus*, grasses and sedges, ferns, and forbs). Two-person teams spent the summer of 2006 surveying over 2,000 sampling points (each comprising one plot of 646 square feet and two of 67 square feet) in 234 1-square-mile blocks, intentionally oversampling to enable evaluation of statistical precision for each category of data as a basis for improving protocol efficiency.

The researchers recommended several cost-saving measures and protocol enhancements, including reducing the number of blocks surveyed while doubling the number of sampling points within each block and dropping a few of the measured indicators (however, one reviewer pointed out that the authors' use of precision analysis and comparison of within-block with between-block variation as the basis for reducing the number of blocks represents faulty logic). The estimated cost per two-person team per summer to survey 50–60 one-square-mile blocks using the revised protocol was $15,000–$20,000, plus additional costs for database management and data analysis.

Diefenbach and Fritsky's field and statistical methods provide a model for later researchers to use in testing the utility of other potential indicators of forest ecosystem recovery. They showed by example how to assess a candidate indicator's potential to provide a reliable, clear signal of deer effects on forest ecosystems while keeping required funds, time, and effort within reasonable limits. They pointed out that the relevant independent (causal) variable in studying deer effects on forest ecosystems is intensity of deer management—an applied treatment or set of treatments—and not deer density. They concluded that the key challenges in integrating forest and deer management (multiple temporal and spatial scales, high uncertainty, reluctance to take on the expense of full-scale controlled experiments, and the customary pre-eminence of politics in driving decision making) can be addressed most effectively by using the adaptive resource management (A.R.M.) approach.

Outcome of the review process

The three studies and the reviewers' written comments served as a springboard for the discussions at the review meeting. This report summarizes those comments and discussions and the review team's recommendations for near-term and long-term enhancement of state forest monitoring from an ecosystem perspective. The report does not dwell on the reviewers' critiques of the researchers' methods or conclusions but mentions them in cases where a specific criticism may help to clarify a larger issue.

Most of the ten peer reviewers' comments in written reviews and during discussion reflect a small number of recurring themes, each of which is the subject of a section of this report. Here the main themes are worded in the form of recommendations to D.C.N.R.:

1. Establish a science-based set of objectives quantifying the desired future condition of the state forests.
2. Evaluate and reanalyze existing monitoring data before pursuing the next round of data collection in the field.
3. Design the next monitoring protocol, incorporating additional indicators of deer effects and forest recovery.
4. Design monitoring methods to distinguish effects of high deer density from other major influences on forest recovery.
5. Focus on indicators of deer effects on forest ecosystems and avoid estimating deer density, but consider developing site-specific relative indices of annual deer density *change*.
6. Frame and test key management questions in an adaptive resource management context.
7. Test the assumption that D.C.N.R.-administered DMAP is an effective tool for sustainable forest management to the degree required to meet ecosystem management goals.

1 AGREE ON A VISION FOR THE FUTURE WITH MEASURABLE OBJECTIVES

Establish a science-based set of objectives quantifying the desired future condition of the state forests.

D.C.N.R.'s overall vision and goals for the state forest system and how they translate into quantitative management objectives need to be clarified first, because they are the foundation for everything else. Currently, D.C.N.R.'s ecosystem management (including deer management) goals are general, qualitative, and subjective and vary to some extent from document to document. If the intent is to sustainably manage state forests from an ecosystem perspective, these goals need to be translated into objectives that are specific and quantitative and can be applied uniformly. The objectives must be firmly rooted in a detailed, quantitatively focused consensus on the desired future condition of the 2.1 million acres of state forest land. There was wide agreement among reviewers that day-to-day and year-to-year actions to restore and sustain healthy forests should be based on a clear, scientifically sound vision of the desired outcome.

The first goal in the state forest deer management plan (Pennsylvania D.C.N.R. Bureau of Forestry 2006) is

> Sustain a healthy and functioning forest ecosystem, including the ability of forests to regenerate with desirable species, by balancing the deer herd with its habitat across state forest lands.

However, there is currently no consensus on what "a healthy and functioning forest ecosystem" means quantitatively. Furthermore, "desirable" species is ambiguous; perhaps what is meant is *native* species appropriate to specific site conditions. In order for coordinated, forward progress to be made in ecosystem restoration and management, D.C.N.R. management staff—in collaboration with scientists and other stakeholders—need to define what the target ecosystems should look like and how they should function, and then link those qualitative goals with quantitative objectives.

Desired future condition (D.F.C.) analyses are part of an emerging science-based approach to ecosystem management by the U.S. Forest Service, National Park Service, and other large-scale land management agencies. The process ties together adaptive resource management, ecological restoration, integrated planning, ecosystem monitoring, and condition reporting. A desired future condition analysis may be defined as a qualitative and quantitative description of *ecosystem attributes* that are expected to be present at some point as an outcome of deliberate management policies, strategies, and practices. Ecosystem attributes include individual resources, communities, ecosystems, and the natural processes that sustain them.

Formulating a D.F.C. analysis helps resource managers to be proactive rather than reactive. The process spurs them to identify trade-offs between competing resources or goals. In addition, a D.F.C. analysis:

- identifies expected outcomes that are derived from goals (but does not prescribe or compel specific management actions or projects);
- implies a dynamic range of conditions and processes, not a static ecosystem;

- treats ecosystems at multiple scales (site, landscape, region);
- establishes a framework and purpose for management actions and programs (but does not focus only on removal or mitigation of ecosystem stressors);
- provides a documented analytical framework that examines assumptions;
- includes measurable mileposts for management objectives (i.e., thresholds of monitored indicator responses) that signal goal achievement;
- makes use of existing condition assessments (but does not merely document or predict trends based on current conditions and passive management);
- takes account of irreversible ecosystem changes and limitations imposed by ownership or other variables; and
- aims to be both realistic and sustainable, given the capacities of the target ecosystems and the socio-political system.

A desired future condition analysis is not an attempt to return to the past. It takes into account both what is known about the pre-degradation condition and important influences that are beyond managers' control, for instance, introduced diseases and pests that are now endemic, extinct animals and plants or those that have been are extirpated but are impractical to reintroduce, and climate change. Predictive models, parameterized using data collected in the course of adaptive resource management (see section 6, p. 37) or in a forest restoration study (see section 7, p. 41), will help in judging what is possible to achieve.

Reference sites (intact natural areas) and historical records are consulted as part of the D.F.C. analysis process. A member of the review team from out of state asked what the benchmark areas are in Pennsylvania similar to the Menominee Indian Reservation in Wisconsin and the Joyce Kilmer Memorial Forest in western North Carolina (see photo on this report's cover). In fact, most of the few small old-growth remnants in Pennsylvania appear to be substantially more degraded from their pre-European-settlement condition than those examples—in the understory by prolonged deer overbrowsing and in the overstory by tree-killing diseases and pests introduced from Eurasia (e.g., Hearts Content in Warren County), and in some cases by soil acidification or fire exclusion. Nonnative, invasive plants are increasingly usurping space in the understory. The distribution of Pennsylvania old-growth remnants is skewed toward minor forest types (e.g., hemlock – white pine) and rocky, steep, logging-unfriendly terrain, further limiting their usefulness as a basis for generalization. Botanical records from the nineteenth and early twentieth centuries give many clues about understory plant species composition, but by the time records began to be kept, the Commonwealth's forests had already endured a century or more of unsustainable timbering practices, severe post-logging slash fires, and unrestrained exploitation of other forest resources such as tannin, charcoal, distillation products (alcohol, turpentine), pharmaceut-icals (e.g., ginseng), and flavorings (e.g., wintergreen). Nonetheless, the methods of historical ecology have advanced considerably in recent years and will have much to contribute to a D.F.C. analysis of Pennsylvania's many forest types. Reference sites in Pennsylvania and in the surrounding region where similar forest types occur, even though degraded to various degrees, also will play a vital role.

The quantitative objectives will need to be forest type-specific and will also vary by region within a forest type (e.g., northern hardwoods in northwestern Pennsylvania versus northeastern Pennsylvania). The D.C.N.R. publication *Terrestrial and Palustrine Plant Communities of Pennsylvania* (Fike 1999) describes 54 distinct forest and woodland community types.

Some of the general categories of ecosystem attributes that should be included in a D.F.C. analysis were spelled out in the corrective action request made by Scientific Certification Systems in its audit of the Bureau of Forestry's management practices, namely, to

> continue taking strategic and political actions until the deer herd is reduced to levels that will, over time, result in recovery of understory vertical structure, species

composition, and abundance of vegetation across all relevant state forests.

In the review discussion, State Forester Devlin also listed a few general D.F.C. components, including a diverse mix of native species appropriate to each ecosystem type, well-developed understory and mid-canopy layers, and no more need for deer fencing.

Characterizing a desired future condition (including specifics on the meaning of "healthy and functioning forest ecosystem") is a first step toward knowing whether forward progress is being made. The second step is measuring progress toward that D.F.C., quantifying what might be called a positive vector of change. Perhaps more important than reaching a definite end point is heading in the right direction (i.e., achieving improvements in indicators). Striving toward ultimate goals and the quantitative objectives related to those goals calls for many steps—small, medium, and large—before those goals are met. Setting those goals and objectives provides the yardstick to measure the value of steady progress. Success comes not only when a quantitative objective is fully achieved, but also in reaching various mileposts along the way. For instance, when areas no longer need deer-deterrent fences to foster regeneration, D.C.N.R. and its partners can celebrate a major success. Reaching objectives related to certain other goals, such as the return of understory species diversity and habitat quality, will take longer.

Summary of short-term recommendations—measurable objectives for ecosystem management

- Elevate the process of conducting a desired future condition analysis for Pennsylvania's state forests to top priority. Begin by establishing a task force to deliberate on potential sources of funds, personnel needs, and a timetable.
- Undertake a desired future condition analysis consisting of:
    - forest-type-specific descriptions of the conditions of resources, communities, ecosystems, and the natural processes sustaining them that are the target outcomes of management; and
    - a comprehensive tally of quantitative, measurable thresholds of indicator responses that will signal success in achieving target outcomes.

Summary of long-term recommendations—measurable objectives for ecosystem management

- Set deadlines for reaching quantitative objectives, including interim milestones of partial achievement.

2 TAKE STOCK BEFORE GOING ON

Evaluate and reanalyze existing monitoring data before pursuing the next round of data collection in the field.

Despite the research reviewed in this report and other published scientific studies on indicators of forest health and recovery (reviewed in Latham *et al.* 2005, pages 135-144), there is still no definitive answer to the question—What should we be monitoring to measure success? It is clear that there is no single indicator that can act as a surrogate for the multitude of factors implicated in forest health. Furthermore, indicators cannot be exactly the same in different regions or different forest types. A published review of the research literature often cited for relative ranking of browse species by palatability nevertheless cautions that

> preferred species frequently differ between regions in the same forest type, within regions over long periods of time, at different times during a growing season, and at different deer densities in the same forest type. [Latham *et al.* 2005, page 71]

Benner (2007) confirmed regional differences in preference ranking among browse species.

Some members of the review team suggested that a "starter" set of indicators might be extracted from the mountain of data that has already been gathered in Pennsylvania's forests and in similar forests in nearby states, using analysis methods not previously applied to this problem. They advocated a full inventory of pertinent existing data to be conducted as soon as possible, with qualified statisticians assigned the task of evaluating its quality. If the data meet requisite standards, these experts would then conduct analyses to weigh each indicator's (or group of indicators') predictive power and measurement cost. Several members of the panel expressed serious reservations about spending further time and money on indicator data collected in the past in the state forests (summarized by Benner 2007), maintaining that the effort would not translate into any meaningful management actions. However, other potentially relevant data exist. If the decision is made to evaluate and reanalyze existing data, possible sources include:

- one or more of the studies that are the subject of this report (Benner 2007; Diefenbach and Fritsky 2007; deCalesta 2008);
- the Pennsylvania Regeneration Study (P.R.S.), part of the U.S. Forest Service's Forest Inventory and Analysis Program (F.I.A.);
- the continuous forest inventory (C.F.I.) program of the D.C.N.R. Bureau of Forestry, Resource Inventory and Analysis section;
- other pertinent studies, especially those that include long-term monitoring of deer exclosures; and
- data on the costs associated with the various monitoring tasks performed in all of these studies.

Diefenbach and Fritsky's (2007) counsel,

> To further refine a vegetation monitoring program based on the recommendations presented in this report, changes in deer density are required during which repeated vegetation measurements are collected

struck many of the reviewers as being right on target. Until the behavior of candidate indicators is compared among many sites, including some in which deer browsing pressure is dramatically reduced, little progress can be made in selecting effective indicators.

Of the three reviewed studies that deal with deer impact indicators, Diefenbach and Fritsky's (2007) made a useful start at an appropriate choice of indicators by weighing the costs and benefits of using certain herbaceous understory species, recognizing that forest ecosystem recovery goes beyond timber regeneration. The indicators used by the other researchers are almost exclusively concerned with tree regeneration and in some cases are highly problematic. For instance, Benner's (2007) "adequately stocked" versus "not adequately stocked" is not a quantitatively measured indicator but is determined solely by foresters' judgments in the field, and deCalesta (2008) equated "deer impact" with a subjective classification of visible signs of browsing on understory shrubs and tree seedlings, which may be absent or nearly so on sites with a history of prolonged high deer density.

Other studies in the state forests such as the Bureau of Forestry's continuous forest inventory (C.F.I.) program were not designed to focus on indicators of deer effects on forest ecosystems but may nonetheless have produced pertinent data. Studies conducted elsewhere may be more relevant, and should be revisited in any case, in the course of designing a new monitoring protocol (see section 3, page 23).

An excerpt from *Managing White-tailed Deer in Forest Habitat from an Ecosystem Perspective: Pennsylvania Case Study* (Latham *et al.* 2005) sums up some of the challenges and suggests an approach to selecting rapid-response indicators:

> Direct sampling of the most vulnerable components—shrubs and understory plants—is problematic in the short term, because recovery in forests that have been severely overbrowsed will likely take many years. Given this problem, [the authors] have looked at supplementing direct measures of herbaceous and understory plants that recover rapidly, such as *Rubus* [blackberry, raspberry, dewberry] species, with a more rapidly responding surrogate for herbaceous vegetation. The surrogate is a subset of the tree species, namely, those that can regenerate successfully only if browsing pressure is low enough also to permit recovery of shrub and herbaceous plant diversity. The assumption is that seed sources remaining in the canopy are available to initiate recovery of this component of the woody flora quickly, even where the reappearance of most shrubs and herbaceous species will take longer because of the deer-induced decline of local seed sources. Whether deer management policies that enhance the regeneration of the suite of indicator trees will actually enhance the regeneration of understory plants will need to be tested in the years ahead.

Note that the use of tree seedling indicators is suggested as a complement to, not a replacement for, monitoring herbaceous and shrub species.

A comprehensive analysis of existing data compiled from multiple studies is the most credible and realistic way to determine whether existing data offer useful insights and, if so, which indicators are effective and which may be dropped. Suggested analytical approaches for ranking the predictive power of candidate indicators or groups of indicators include path analysis and ordination methods such as detrended correspondence analysis and non-metric multidimensional scaling. If existing data yield useful results, they could provide a basis for DMAP prioritization on state forests and lay the groundwork for standardizing, in the near future, which indicators and threshold values should be used to determine the need for DMAP enrollment. If the data fail to provide useful insights, then assessing the effectiveness of DMAP in meeting D.C.N.R.'s forest management goals will be delayed by the need for additional effort in designing and implementing a new, effective monitoring protocol.

Summary of recommendations—evaluation and reanalysis of data on hand

- Commission qualified statisticians to evaluate existing data on potential indicators from all available sources, including one or more of the studies that are the subject of this report, the U.S. Forest Service's Pennsylvania Regeneration Study (P.R.S.), D.C.N.R.'s continuous forest inventory (C.F.I.) program, other pertinent studies including long-term monitoring of deer exclosures, and any relevant data on the costs associated with the various monitoring tasks performed in all of the studies. If the data meet requisite standards, conduct analyses to weigh each indicator's (or group of indicators') predictive power and measurement cost using methods such as path analysis and ordination.

3 UPGRADE TO "FOREST MONITORING 2.0"

Design the next monitoring protocol, incorporating additional indicators of deer effects and forest recovery.

A well-designed monitoring protocol is one of the keys to a successful ecosystem management program. An inadequate protocol can only result in wasted time, funds, and effort and missed opportunities for management to adapt, self-correct and improve over time. A crucial part of an effective, highly sensitive monitoring protocol is the choice of indicators. There was consensus among reviewers that the next iteration should cast a wider net beyond the indicators used in prior studies (see section 2, page 19). Incorporating additional indicators suggested by the scientific literature and expert opinion will strengthen the effectiveness of monitoring.

D.C.N.R. has committed to responding to the corrective action request regarding deer management in the state forests made by Scientific Certification Systems in its audit of the agency's management practices (part of the Green Certification process), which calls for the Bureau of Forestry to

> continue taking strategic and political actions until the deer herd is reduced to levels that will, over time, result in recovery of understory vertical structure, species composition, and abundance of vegetation across all relevant state forests. [Pennsylvania D.C.N.R. Bureau of Forestry 2006]

The Bureau of Forestry uses a variety of information to shape DMAP decisions, such as previous years' DMAP harvest rates, the numbers and success of antlerless deer licenses in the surrounding wildlife management unit (W.M.U.), the level of deer browsing on shrubs and tree seedlings, saplings, and root suckers, and the tree species regenerating in the target area. Members of the review team pointed out, however, that the previous years' harvest rates in the target area and in the surrounding W.M.U. are not pertinent to addressing the corrective action request. They argued that D.C.N.R. may be monitoring parameters that are irrelevant to the goal, and incorrect management decisions may result because the quantitative thresholds of indicator measures that trigger management actions are not well defined. One of the most important goals of refining the monitoring protocol is to standardize indicator threshold-value triggers of DMAP enrollment within a few years. To achieve this goal, the Bureau of Forestry needs to:

- establish specific, quantifiable values for indicators as objectives, with deadlines for reaching them in specific areas;
- collect appropriate data to regularly assess existing conditions relative to the quantified objectives; and
- formulate an appropriate decision-making model that will achieve each quantified objective in the allotted time.

Diefenbach and Fritsky's rapid assessment protocol is a rigorous framework for assessing forest vegetation change. However, as the authors themselves recommended as a conclusion of that study, further development and refinement of the monitoring protocol is needed and measurement of additional environmental variables—factors other than deer browsing itself that are likely to influence vegetation responses to changes in levels of herbivory by deer—should be added. They suggested, for example, that data also be collected on soil conditions, seed banks, and silvicultural treatments at the same survey plots where vegetation is monitored.

Reviewers further recommended using the monitoring results to parameterize complex models (e.g., LANDIS-II, described in Scheller *et al.* 2007), with deer density and herbivory effects as model subunits, and applying the models to predict forest change. Deer and forest management objectives could then be developed based on predicted impacts of deer on forest vegetation. A 3–4-year time frame should be adequate to produce the monitoring data needed to parameterize the models.

Several reviewers pointed out the importance of widening the field of indicators to include measurements that reflect vertical structure, patch diversity, noncommercial species regeneration, and other essential forest ecosystem qualities and processes in addition to timber regeneration and native species diversity. For most indicators, target ranges of measured values will eventually need to be set that differentiate "acceptable" from "unacceptable," recognizing that every indicator has a natural range of variation (see *Establish a science-based set of objectives quantifying the desired future condition of the state forests*, page 15).

Direct-response indicators of the effects of high deer density on forest health (dependent variables in the data analyses) are those that a deer population affects directly. These are the indicators that respond to deer population density, feeding preferences, and seasonal movement. Categories of direct-response indicators include:

- tree seedling/sapling counts by size class and species;
- native shrub density/cover by species;
- nonnative, invasive shrub species cover;
- native herbaceous species composition and cover;
- nonnative, invasive herbaceous species cover; and
- herbaceous indicator plant height and fruit production relative to the density of stems per unit of area (Augustine *et al.* 1998).

Response-modifying indicators (independent—or causal—variables in the data analyses) are the major non-deer influences (see section 4, page 29) and other factors that are likely to modify how deer affect the indicators above. Categories of response-modifying indicators include:

- overstory conditions (forest type, tree size class distribution, species composition, percent canopy closure);
- recent timber management history, including harvest levels and time since harvest;
- deer fencing history;
- oak mast production;
- forest size (contiguous area) in which the monitoring site is embedded;
- proximity of edge and type of surrounding landscape (e.g., agricultural lands);
- soil characteristics, especially moisture regime, calcium availability, and buffering capacity;
- soil seed bank and seed rain from bird, wind, and gravity dispersal;
- indices of deer harvest rate or deer population density; and
- cover of competing vegetation that has become disproportionately abundant as a legacy of prolonged high deer density, such as rhizomatous ferns (hay-scented fern, New York fern, bracken), certain native woody plants (e.g., striped maple, American beech), and nonnative, invasive shrubs or herbaceous plants.

Browse impact data, such as were collected for Benner's (2007) and deCalesta's (2008) studies, are also direct-response indicators, but some experts are skeptical of their value. Because browse impact assessments appear to be a major thrust of past and ongoing monitoring efforts, resolving this issue is of high importance. Skeptics argued that browse impact data are inherently flawed because the ground- and shrub-layer woody vegetation that such data are collected from is often scarce or absent in forests that have a long history of high deer abundance and severe overbrowsing. Furthermore, browse impact data are usually gathered in winter when twigs are most visible but the more reliable metrics—the abundance of healthy advance regeneration and herbaceous vegetation—are collected in late spring and summer; if browse impact monitoring is continued, an additional visit could be avoided by doing it in summer. In any case, it is clear that browse impact data are valid only relative to total twig abundance within deer's reach, which must be measured or estimated at the same time.

Surveying benchmark forest stands is crucial to the success of monitoring. Aside from the historical record, which is extremely sketchy, there is no other way to know for sure what constitutes recovery or even what a healthy forest ecosystem looks like. During the peer-review discussion, the question arose whether any state forest Wild Areas might serve as benchmarks. The answer was no; even the Hammersley Wild Area, 51 square miles of roadless forest in the Susquehannock State Forest that is in many ways D.C.N.R.'s premier Wild Area, has understory and regeneration conditions as degraded as in silvicultural areas. Until the deer herd is reduced and sustained at low levels in a significant fraction of each major forest type long enough to allow recovery, which might be decades in many areas, exclosures have the best potential to provide the critical benchmark data needed to refine monitoring protocols and set quantitative goals (quantitative goals are discussed in section 1, page 15).

However, one factor that can complicate exclosure data is variation in legacy effects of prolonged high deer numbers on seed availability. Forest understory species have notoriously short-lived seeds and many are ant- or gravity-dispersed. Where deer numbers have been high for more than a few years, some understory species may have been extirpated, even from the seed bank. Plants with bird and wind-dispersed seeds will recolonize, although slowly. That is the rationale for suggesting seed bank and seed rain assays as potential indicators.

How quickly direct-response indicators change as deer density changes varies greatly from one indicator to another. Indicators fall along a response-time spectrum. For instance, ginseng, hobblebush, American yew, and native honeysuckles are typically near the slow-response extreme, in part because the number of unbrowsed plants that go to seed each year in the entire region has fallen to record low levels. Seedling densities of common, wind-dispersed tree species such as red maple are near the rapid-response extreme. It is vital to use a suite of indicators along this entire spectrum to reflect whole-ecosystem response in both the short and long term.

Rapid-response indicators can serve as early warning signs to assess growing impacts and impending regeneration problems before local deer populations exceed thresholds where impacts become severe and long lasting. Conversely, as impacts decrease in response to deer management efforts, rapid-response indicators reflect initial recovery trends. Slow-response indicators are the key to monitoring ecosystems fully and over long time periods to ensure the recovery and sustainability of slower-growing, more sensitive species that take time to recolonize or rebuild populations.

One reviewer pointed out in written comments that slow-response indicators are also crucial to understanding time-lag effects and threshold effects, which are common in the dynamics of both degradation and recovery. Time-lag and threshold effects involve nonlinear responses of ecosystem attributes to a change in environmental conditions. One way in which such effects affect management is that they often cause the trajectory of recovery to be different from that of degradation. In other words, where time lags or thresholds are significant in ecosystems, recovery is not simply degradation in reverse; the recovery threshold may be very different from the degradation threshold. In such a situation, the desired ecosystem state may have lower resilience than its degraded counterpart, and recovery will depend on reducing the major cause of degradation to well below the level that originally triggered the shift to a degraded state.

In such cases recovery will require a long time or additional management action besides deer reduction. Restoring low deer density does not bring about understory recovery and tree regeneration where deer have been abundant for so long that they have all but eradicated the more-palatable competitors of the least-preferred species, such as rhizomatous ferns and nonnative, invasive shrubs, whose cover often approaches 100%

(Augustine *et al.* 1998; Latham *et al.* 2005). In such situations there is little space left for diverse native understory shrubs and herbaceous plants to reestablish, and dormant seeds and sources of fresh seeds within dispersal distance are likely to have died out long ago. Likewise, understory recovery will be delayed where the tree canopy is closed and shade is dense (these issues are discussed further in section 3, p. 23).

Challenges that must be faced in retooling the monitoring protocol include the scarcity of time, funds and staff with specialized training. Quantitative monitoring of numerous plant species requires time and a particular skill set, both of which are limited in availability. For Benner's study, about 150 Bureau of Forestry staff members monitored nearly 75,000 sampling plots over two years but did not engage in the more time-intensive and specialized collection of data on herbaceous plants. In addition, monitoring has tended to be focused primarily on areas under silvicultural management. Wild Areas, Natural Areas, and Wild Plant Sanctuaries are just as important to manage from an ecosystem perspective, and it is just as important to monitor indicators of forest recovery and health in them.

There was wide agreement among reviewers that monitoring needs to be conducted in June–August because of the importance of herbaceous plant indicators. Earlier in the spring, a great many later-flowering species are unidentifiable or even invisible. By June several spring ephemerals are no longer visible, but they are vastly outnumbered by species that are difficult or nearly impossible to identify until mid-summer. After August, many more spring-flowering species have declined to the point where they would be severely undercounted.

One member of the group pointed out that, in some cases, similarly responding species may be lumped into groups for more efficient cover estimation instead of surveying individual species and summing their cover by groups later. Besides increasing efficiency, using functional groups is desirable because any one individual indicator species is likely to be present at only a fraction of monitoring locations. Bracketing species into functional groups greatly increases the likelihood of a sampling plot having a non-zero value for an indicator, which increases the indicator's sensitivity and discrimination power relative to indicators that are absent from a high proportion of samples.

Another member cautioned that the spatial scale at which deer effects are measured and analyzed to guide the deer management effort should be as fine or coarse as the scale of significant spatial variation in the data. If monitoring data were averaged over an entire district, which in most cases would be far too coarse a scale, then management intensity in some parts would be too much, in other areas too little, and only in a fraction of the district just right. The data collected to date for Benner's (2007) and Diefenbach and Fritsky's (2007) studies could be analyzed to provide a preliminary estimate of the scale of heterogeneity, for at least a few indicators in some forest districts. The appropriate scale of measurement may differ from one district to another, depending on the heterogeneity of the state forest landscape. The aim should be to stratify monitoring plots within districts so that there are several plots within each major forest type or combination of environmental characteristics, which may correspond to obvious features such as ridgetops and valleys, northern hardwoods and oak-mixed hardwood types, or deep woods and forest fragments with cropland edges.

The temporal scale of monitoring should be data-driven eventually, but the review team consensus was to start out by rotating surveys among monitoring plots within a district or group of districts so that one-fifth, one-fourth, or one-third of plots are monitored each year (i.e., each monitoring plot's survey would occur at 3, 4 or 5-year intervals). Annual surveys at every monitoring plot are not necessary because even rapid-response indicators tend to change over longer periods. For example, in a 10-year deer enclosure study in northwestern Pennsylvania of deer density effects on vegetation, indicators of deer density were tracked annually but several years elapsed between vegetation surveys (Horsley *et al.* 2003). Timing of indicator surveys in particular areas could be contingent on relevant events; for instance, surveys might be scheduled for set intervals (e.g., on the third and seventh years) after a timber cut or an oak mast year. Not every indicator needs to be monitored in every survey period; rapid-response indicators might be monitored every 2–4 years and slow-response indicators every 6–8 years. Eventually, using several years' data to parameterize predictive models could lead to refinement of monitoring frequency by running the models to see how rapidly they predict changes for specific sets of indicators. During the interval between deer impact surveys in an area, deer management should be held constant. Otherwise, it will be difficult to interpret indicator responses.

Given a scarcity of funds and time, the Bureau of Forestry may need to prioritize relatively few sites in the short term for intensive deer management—where early success is judged most likely—and add more sites later. Likewise, a triage approach may be appropriate in prioritizing where and when monitoring is conducted in the short term. Districts that appear to have the biggest problems might be dealt with first and the lessons learned there applied in other areas later.

Ideally D.C.N.R. should work toward developing a monitoring protocol that can be used on other public lands and on private lands across the state.

Summary of short-term recommendations—monitoring-protocol upgrade

- Use the results of existing data reanalysis (described in section 2, page 19) to choose the most promising indicators; incorporate additional candidate indicators suggested by the scientific literature and expert opinion, including a mixture of rapid-response and slow-response indicators; and design a monitoring protocol to measure them as efficiently as possible.
- Change the monitoring period to summer, including browse impact monitoring.
- Inventory, establish, maintain, and monitor a network of permanent deer exclosures to provide benchmark data needed to refine protocols and set quantitative goals.

Summary of long-term recommendations—monitoring-protocol upgrade

- Weigh indicators' predictive power and measurement cost after several years of data collection.
- Use several years' monitoring data to parameterize complex models to predict forest change.
- Use the predictive models to refine the return interval between indicator surveys on each monitoring plot by seeing how rapidly the models predict changes for various sets of indicators.
- Analyze the data after every monitoring cycle and use the predictive models to further refine indicators and revise monitoring protocols to maximize efficiency.
- Standardize indicator thresholds that trigger DMAP enrollment based on quantitative objectives (see section 1, page 15) and predictive models. Use indicator thresholds to formulate a decision-making process that will achieve each quantitative objective in the allotted time.

4

DISCRIMINATE DEER FROM NON-DEER EFFECTS

Design monitoring and research methods to discriminate the effects of reduced deer browsing from other major influences on forest recovery.

Improving credibility in any natural resource management endeavor includes both scientific approaches (appropriate and rigorous study designs and data analyses) and social approaches (clear and accurate communication with the public and policymakers). Aside from limited resources that constrain all management programs, improving credibility with the public and policymakers is often the greatest challenge. Some use the issue of non-deer influences on forest regeneration to question the credibility of forest monitoring and management in Pennsylvania. Many hunters, and consequently some elected officials, are concerned that state agencies overemphasize the negative effects of deer. They often promote non-deer factors as more important contributors to observed declines in forest health. The three studies in this review vary in their approach to this problem, from little or no consideration of other factors to incorporation of procedures to limit their influence (e.g., monitoring regeneration plots only in areas with sufficient light). The reviewers agreed that a vital first step in improving "social credibility" is to acknowledge potentially important non-deer factors that affect regeneration and incorporate them into monitoring procedures.

Potentially important factors other than current deer abundance that vary at landscape and regional scales and are likely to influence indicator performance include:

- **shade at the forest floor**,
- **forest type** (e.g., northern/Allegheny hardwoods, oak-mixed hardwoods);
- **soil factors** (e.g., soil acidification by atmospheric deposition, degree of buffering by soil calcium, moisture regime);

They also involve legacy effects of prolonged high deer density, which may include, but are not limited to:

- **depauperate seed bank;**
- **scarcity of live seed sources within dispersal distance**;
- **disproportionate understory abundance of unpalatable plants that have proliferated because their competitors have been eradicated** (dense rhizomatous fern cover strongly inhibits native tree, shrub and herbaceous seedling establishment and certain other native species in some circumstances may do the same, e.g., striped maple and American beech); and
- **disproportionate abundance of nonnative, invasive species** (which are generally unpalatable and strongly inhibit native tree, shrub and herbaceous seedling establishment through shade, allelopathy, and effects on soils and soil microbial communities; examples include garlic mustard, Japanese stiltgrass, Japanese/giant knotweeds, multiflora rose, Tartarian/Morrow's/Amur/Japanese honeysuckles, ailanthus, Norway maple, and many others).

Various reviewers commented on appropriate ways of dealing with these sources of variation in adaptive resource management (see section 6, page 37) and in research (see section 7, page 41, and Appendix, page 47). Approaches likely to produce scientifically plausible evidence of indicator responses to changing deer densities are summarized in the following bullet points. In every case, to distinguish the effects of non-deer influences from deer effects it is essential to compare indicator responses in areas where deer populations are reduced (or within fenced deer exclosures) to areas where deer populations have not been reduced.

- Monitor factors that may be confounded with deer effects and include them in data analyses as unmanipulated variables ("natural experiments"). This can be expensive because a great many monitoring plots are required to make sure the full range of variation in each factor is covered. Measuring the extent and diversity of the seed bank and seed rain add to the demands on funds, time, and labor. In some cases, there may be little land available in an important category, for instance, stands where the seed bank or seed rain has not been depleted and approximates that of a healthy, diverse forest.
- Manipulate potentially confounded factors in management trials or experimental treatments. Examples include herbiciding hay-scented fern; thinning the canopy to increase sunlight; augmenting seed banks with native forest understory species (of locally indigenous genotypes); spreading lime on soils to increase buffering capacity; and using prescribed burning to promote regeneration.
- Control for potentially confounded factors by selecting areas to be monitored based on predetermined criteria. Sites are rejected that do not fit into relatively narrow ranges of each variable. For instance, managers could decide *a priori* that data will be collected where forest-floor light levels are 50–75% of full sun, rhizomatous fern cover is less than 20%, plentiful seed sources of native shrub and herbaceous species exist within dispersal distance, etc. The main disadvantage is that, across large areas of some state forests, land that meets some of these criteria is scarce.
- Control for potentially confounded factors using a mixed approach of manipulation and well-chosen locations for monitoring. Monitoring locations are limited to specific ranges of factors that are least practical to manipulate at the district or state scale (e.g., forest-floor light availability, soil moisture regime) and the other variables are manipulated as part of management (e.g., herbicide treatment, augmenting seed banks). Two or more levels of some of the confounded variables may also be included, without manipulation, by stratified sampling (e.g., place half of the monitoring plots in areas with high-calcium soils).

Shade on the forest floor is widely recognized as an important non-deer factor affecting forest recovery, therefore it is possible to combine improvements in social credibility and operational efficiency at the same time by monitoring most indicators only where there is enough light to support understory growth. One reviewer calculated, based on the figures in Diefenbach and Fritsky's study (2007; their Table 7, page 16), that it may be possible to reduce the data collection effort by about half—or redirect some of that effort to expand coverage—by omitting or reducing monitoring effort on plots with insufficient light for regeneration. Since tree regeneration and understory recovery are the ultimate measures of sustainable forest management and light reaching the forest floor is an important non-deer influence, simply recording these plots as "not enough light" concentrates the bulk of data collection efforts onto plots where regeneration and understory recovery should occur. Results from existing research, including the three studies reviewed in this report, show clearly that it is inefficient to collect data on the full range of indicators on plots with insufficient light for regeneration. Minimal tracking of such plots makes it possible to estimate how much land in a stand, tract, or district has enough light for regeneration and what proportion of plots switch between "enough light" and "not enough light" from canopy disturbance, succession, and ecosystem recovery. Monitoring indicators only on sites where they have enough light to respond will greatly boost efficiency.

Light availability at the forest floor depends in large part on forest management history (harvest intensity and time since harvest). Land zoned for timbering is about two-thirds of the state forest total and less than 1% of that area is cut every year. The rotation scheme is now driven by stand age-class considerations and not by a statewide annual percentage goal; implementing an ecosystem management approach will probably lead to the adoption of a longer rotation, further decreasing the area available to monitor deer impact and forest recovery indicators. The current percentage of state forest lands with enough understory light to allow tree regeneration is unknown, but of the 2,269 plots surveyed in Diefenbach and Fritsky's study (2007; their Table 7, page 16), 1,120 (49%) had less than 75% overstory stocking, a rule-of-thumb indicator of suitable understory light conditions for establishment and maintenance of tree seedlings (advance regeneration) and native understory shrubs and herbaceous plants. The subset of that category that has a dense cover of unpalatable interfering vegetation at ground level (hay-scented fern, American beech root sprouts, striped maple, nonnative invasive plant species)—a common legacy of prolonged high deer density—is unknown.

The most reliable way of distinguishing deer from non-deer effects is to compare indicator responses inside and outside of fenced deer exclosures and inside and outside of areas where deer densities are reduced using DMAP or other population-control measures. Such comparisons take the guesswork out of determining whether indicator responses are due to excessive deer browsing. Several reviewers noted that D.C.N.R. could get the greatest benefit from its monitoring investment and foster stronger confidence in the results by routinely collecting such comparative data, called experimental *control* data (see Appendix, page 47), in the forest monitoring program. Several reviewers stressed the importance of monitoring both unmanipulated areas—with ambient deer density where ordinary hunting regulations apply—and deer exclosures, in addition to DMAP areas. They argued that all three are needed for differences in intensity of deer management to be plausibly linked to changes in deer impact indicators. Neither deCalesta's (2008) nor Benner's (2007) studies included either type of experimental control, which makes the results less scientifically defensible and "socially credible" in terms of whether DMAP has an effect on indicator performance.

Summary of recommendations—deer versus non-deer effects on forests

- Reduce the amount and types of monitoring data collected in areas where the forest canopy or fern cover is too dense to allow significant indicator responses.
- Inventory, establish, maintain, and monitor permanent deer exclosures to distinguish between deer and non-deer effects.
- Stratify plots in the next cycle (and later monitoring cycles) to enable comparisons of indicator response inside and outside of exclosures and inside and outside of areas where deer densities are reduced using DMAP or other population-control measures.

5 MEASURE DEER EFFECTS, NOT DEER NUMBERS

Focus on indicators of deer effects on forest ecosystems and avoid estimating deer density, but consider developing site-specific relative indices of annual deer density change.

The importance of estimating deer density is regularly debated. It is the subject that generated the most disagreement among participants in this review. Perhaps surprisingly, wildlife biologists on the panel generally recommended that D.C.N.R. discontinue density estimation efforts. They maintained that current methods have low scientific credibility and pose a risk of diverting scarce time and resources from measuring indicators of deer effects on ecosystems. The plant biologists and forest ecologists suggested that D.C.N.R. continue to collect data for a site-specific *relative index* of deer density change over time (e.g., percent change from a baseline survey of pellet groups or aerial thermal image "hotspots," or reported annual harvest rates by hunters). They believe indexing annual percent change on a site-by-site basis will a key part of determining whether a non-response by indicators is due to non-deer effects, or legacy effects of prolonged high deer numbers, or inadequate deer herd reduction (although exclosures and herbicide-treated plots, where practical, can also be used to distinguish among these alternatives).

All agreed that indicators of deer effects are essential targets of monitoring. The introduction to Benner's manuscript (2007) stated that D.C.N.R. has decided to monitor habitat impacts rather than deer density. This position is consistent with the P.G.C.'s deer management philosophy. The consensus of the group was that upper management and field staff in both agencies could benefit from a consistent message from staff biologists and forest management planners that impacts, not deer densities, are the most important measures of a deer management program.

The primary independent variable in analyses of monitoring data is intensity or type of deer management, not deer density per se. One reviewer pointed out that the deer population response to different levels of management is actually one of the dependent, or response variables, in the same category as the data on impact indicators. Data on deer population response are not needed to answer the key question—Is deer management succeeding in reducing effects of high deer density to levels at which ecosystem values can be recovered and sustained? However, such data may be beneficial for testing the underlying assumption that management tools actually reduce and maintain deer populations at sustainable levels compatible with forest management from an ecosystem perspective.

Deer biologists' view: estimates of deer density such as those based on pellet-group counts should be discontinued

The wildlife biologists' consensus was that pellet-group counts are not a scientifically defensible method of estimating deer density or even relative deer density change. They maintained that there are no methods that are reasonably accurate, with the exception of the mark-recapture method in the specific case of small, isolated populations in which a large proportion is fitted with radio collars; however, this is a situation that occurs rarely, if at all, on state forest lands, and the costs are astronomical.

Deer density estimates from pellet-group counts depend critically on pellet deposition and decay rates, but the wildlife biologists pointed out that these are highly variable, both spatially and temporally. Current models are based on average defecation rates measured on captive deer in another region, where they were feeding on vegetation different from any that occurs in Pennsylvania. The models do not take into account variable influences on decay rate (e.g., faster decomposition on south slopes because of earlier snowmelt). Several members felt it would be prohibitively expensive and time-consuming to conduct experiments to parameterize a model that takes into account the variation in pellet deposition rates of deer feeding on the various arrays of food available in Pennsylvania forests in the appropriate season and the variation in pellet decay rates under a range of topographic and weather conditions.

They pointed out, furthermore, that pellet-group counts often differ between experienced and neophyte observers and from a single observer when fresh versus when fatigued. Counts also can change substantially with relatively minor tweaks in methods. For instance, according to Benner (2007), detection rates increased "astronomically" when he and his research team changed their pellet-group count design from transects to circular plots; deCalesta (2008) started his study using belt transects but later determined that data collectors were missing nearly half of the pellet groups relative to their performance when they adopted circular plots.

Deer density estimates from pellet-group counts typically have high variability from year to year at each survey point, whether or not there is reason to think that the local deer population has changed. This is attributed to high temporal and spatial variability in the way a deer herd uses various parts of the landscape. Deer congregate in different places at different times in different years, and survey plots or transects by necessity are localized. According to deCalesta, variability in movements from year to year results in wildly varying data where transects are as little as 1,000 feet apart. Precision is low because of the scale of sampling compared with the scale of spatial and temporal variability in deer utilization of the habitat. Variation in oak masting in different years also can contribute to this pattern.

Sources of error in estimates based on aerial thermal imagery include line-of-sight obstruction and counting two or more deer lying or standing together as one. Even though thermal imagery is acquired on cold winter nights, evergreen trees and shrubs, the trunks and branches of deciduous trees and shrubs, topographic features, and boulders can hide enough of a deer's profile at the angle of sight from the camera to screen out a clear signature. A recent review evaluating the application of thermal imagery technology in a variety of deciduous forest environments reported inconsistent results, with 11–69% of the deer missed in the audited surveys and an average detection rate of 56% of the total deer present in the study areas (Haroldson *et al.* 2003).

Several reviewers made the point that regardless of method, the hunting public has demonstrated a propensity to obsess over published numbers, which actually have little relevance to deer effects and cannot be interpreted in a valid manner (see *Deer counts and politics, public relations, and hunter retention*, p. 35).

Plant biologists' view: pellet-group counts or other measures should be used to develop site-specific *relative indices of annual deer density change*

Many panel members agreed that some index of the relative magnitude of deer density *change* from year to year in the area around each monitoring location would be beneficial. The plant biologists maintained that some of the data traditionally used to estimate deer density may yield useful site-specific information on relative density change over time even if they are unreliable as a basis for estimating absolute numbers. The chief benefit would come in cases where deer impact indicators are found to recover slowly or not at all for some years during which DMAP or other population-reducing tools are being applied. In such cases, the plant biologists agreed it is critical to know whether the lack of indicator response is in spite of deer reduction—due to non-deer effects or legacy effects—or because the use of DMAP itself was ineffective in population management.

Whether a site-specific index of relative change is useful depends on the validity of several underlying assumptions. One is that the unreliability of deer density estimates is due mostly to factors varying among sites and much less to factors varying within sites among years. This assumption may hold up better with respect to aerial thermal imagery than to pellet-group counts. An index based on pellet-group counts rests on assumptions that pellet deposition rates depend on winter diet and that diet is unlikely to vary from year to year in one spot (it may change over longer time periods, especially if forest restoration is successful). Another assumption, which may be less defensible, is that pellet decay rates depend on topography and other factors that do not change but not appreciably on annual variation in weather. Possibly the most serious potential pitfall is the mismatch between the scale of sampling and the scale of variability in a deer herd's movement across the landscape.

If pellet-group counts are used to generate a site-specific relative index of annual deer density change, the group agreed that researchers need to look critically at the assumptions, think about their plausibility across a variety of field conditions, and change the protocol where appropriate. To be credible as the basis for a relative index, methods must be tailored to conditions specific to various regions of the state and meet scientific standards that often are not met, as reflected in deCalesta's study (2008). Under some conditions the best method may be plot surveys; in others, distance sampling (where pellet groups are counted while walking along a line) may save time. At least one reviewer maintained that pellet groups are easy to count and efficient techniques have been developed and tested (Marques *et al.* 2001, Laing *et al.* 2003, Shi *et al.* 2006, Forsyth *et al.* 2007), which could be adapted for a site-specific relative index of annual change. One drawback is that browse impact surveys and pellet-group counts have been conducted at the same time for cost efficiency, but it is infeasible to count pellet groups, done in late winter–early spring before leaf-out when pellets are visible, at the same time as herbaceous plant indicators are monitored, which is in early June–late August when plants are identifiable and measureable.

There was some discussion of another data set that may provide a more cost-effective basis for a site-specific relative index of annual deer density change, namely, reported harvest rates by participating hunters. Drawbacks of this approach include the typically slow reporting rate and the strong dependence of analysis results on the highly suspect assumption of proportionality between reporting rates for DMAP tags and the number of deer brought to check stations. Sending hunters a survey postcard is impractical, because the state has the names and addresses only of DMAP and check-stop hunters, which make up a small percentage of the total. Several reviewers suggested comparing deer harvest reporting and hunter survey methods used elsewhere and exploring ways of improving the reporting rate in the state forests to support credible estimates of relative deer density changes from year to year.

Deer counts and politics, public relations, and hunter retention

When managers provide a deer density estimate—no matter how the numbers are couched in expressions of uncertainty such as margin of error—hunters, policymakers, and the general public tend to take that number at face value. Reviewers agreed that there are significant risks in appearing to have high certainty about deer densities when an accurate census of deer, or any free-ranging wildlife, is nearly impossible. Panel members with deer management experience pointed out that estimates with low accuracy compromise the credibility of a management program. Individuals who do not support the deer program speaking in public forums cite low-quality estimates as evidence that there is no credible science behind management actions. One reviewer said that there is widespread agreement among wildlife professionals that this has been one of the greatest mistakes made in deer management programs throughout the country. More fundamentally, according to several reviewers, it is counterproductive to focus on deer density in the first place. It diverts attention away from the core issue, which is deer influence on ecosystems, and it falsely assumes that there is a broadly applicable relationship between particular deer density numbers and forest health.

One reviewer pointed out that publicizing deer density estimates bolsters a misconception that density ranges are associated with states of ecosystem recovery. There seems to be an increasing trend of accepting 15–20 deer per square mile for tree regeneration and 5–10 deer per square mile for full forest ecosystem recovery as facts, with the aim of applying them in a wide array of situations. However, scientists believe that deer density (whether it is feasible to estimate accurately or not) interacts with a host of other factors in determining forest health, including forest-floor light level, forest type, understory species composition, landscape context, soil conditions, length of growing season, alternative food sources, patterns of seasonal movement by deer, and legacy effects of prolonged high deer numbers. These interactions limit the potential usefulness of deer density alone as a predictor of ecosystem impacts. It is likely a pointless exercise to develop a general rule for how many deer per square mile are compatible with a sustainable, healthy forest ecosystem.

Several participants maintained that, because of inherent uncertainties in estimating deer numbers, it is misleading to release any figures. Doing so risks leading to public or agency staff reactions such as "Now that we're at five deer per square mile, why aren't we seeing a vegetation response?" The appropriate management response to a lack of indicator improvement is to sustain or raise the deer harvest rate as monitoring continues. However, the hypothetical situation where targeted increases in hunter effort are accompanied by a lack of indicator response begs the question—How do we know if we've actually reduced the population if we don't see any reduction in impacts? One suggested possibility is to monitor hunter success rates area by area; a decrease in reported harvest rates might be used as a rough indicator of lowered deer density.

According to deCalesta, whereas some have argued that hunters would give up where deer density falls lower than 15 per square mile, he nonetheless asserted that enough hunters will continue hunting an area if they are told the truth about how deer densities have decreased. P.G.C. biologists in the group recommended that managers respond to lowered densities by providing services and support to hunters designed to foster higher participation and retention rates. Diefenbach cautioned that airborne thermal imaging estimates and pellet-group counts may be seen by some as politically necessary even after deer densities and impacts have been reduced. P.G.C. biologists warned that publicizing counts provokes the hunting public to obsess on the numbers, for instance, when D.C.N.R. airborne thermal image analysis in the Tioga State Forest generated an estimate of 90 deer per square mile. Benner argued that estimates should not be publicized at all, or the statistics should be expressed in a way that avoids alarmism such as not specifying the particular areas where density estimates are very high.

Summary of recommendations—estimating local deer densities

- Suspend efforts to estimate deer density until such time as scientifically defensible methods have been developed that are cost-effective.
- Explore ways of improving hunter return rate of DMAP harvest reports to a consistently high percentage of the total harvest across all units in the state forest system. Use the data to develop a scientifically credible site-specific relative index of annual deer density change.

6 BRING POLITICS AND SCIENCE INTO CLOSER ALIGNMENT

Frame and test key management questions in an adaptive resource management context.

Adaptive management of deer and forests

In presenting the results of their study, Diefenbach and Fritsky asserted that the key challenges in integrating forest and deer management could be addressed most effectively by using the adaptive resource management (A.R.M.) approach. In his presentation before the reviewers, Diefenbach further described why and how D.C.N.R. might implement such a program. Most, if not all, of the review team strongly supported shifting from the present management framework to A.R.M. in state forest management in general and deer management in particular. The only negative comments were expressions of concern about the expected operational and political challenges of such a transformation.

Adaptive resource management is a decision-making process that is more highly structured than traditional approaches to decision making. It emphasizes reducing uncertainty about resource responses to management actions as a means to improve management. The National Research Council defines A.R.M. as a way to promote

> flexible decision making that can be adjusted in the face of uncertainties as outcomes from management actions and other events become better understood. Careful monitoring of these outcomes both advances scientific understanding and helps adjust policies or operations as part of an iterative learning process. Adaptive management also recognizes the importance of natural variability in contributing to ecological resilience and productivity. It is not a "trial and error" process, but rather emphasizes learning while doing. Adaptive management does not represent an end in itself, but rather a means to more effective decisions and enhanced benefits. Its true measure is in how well it helps meet environmental, social, and economic goals, increases scientific knowledge, and reduces tensions among stakeholders. [Williams *et al.* 2007]

Resolving policy controversy and management impasse is a focus of A.R.M. According to *Adaptive Management: The U.S. Department of the Interior Technical Guide*, A.R.M.

> involves exploring alternative ways to meet management objectives, predicting the outcomes of alternatives based on the current state of knowledge, implementing one or more of these alternatives, monitoring to learn about the impacts of management actions, and then using the results to update knowledge and adjust management actions. Adaptive management focuses on learning and adapting, through partnerships of managers, scientists, and other stakeholders who learn together how to create and maintain sustainable resource systems. … Often the uncertainty about management impacts is expressed as disagreements among stakeholders who have differing views about the direction and magnitude of resource change in response to management. An adaptive approach explicitly articulates these viewpoints, incorporates them into the decision-making process, and uses management itself to help identify the most appropriate view about resource dynamics. … Managers have sometimes been reluctant to acknowledge uncertainty in environmental assessments and management strategies. Often there is a perception that asserting certainty as to management impacts is more convincing. … Acknowledgement of uncertain management outcomes is sometimes seen as an invitation for confrontation among different interest groups, resulting in an inability to reach

> timely agreement on a proposed action. Adaptive management forces stakeholders to confront unresolved uncertainties that can significantly influence management performance. An adaptive approach provides a framework for making science-based decisions in the face of critical uncertainties, and a formal process for reducing uncertainties so that management performance can be improved over time. [Williams *et al.* 2007]

Diefenbach pointed out in his presentation that A.R.M. integrating forest and deer management in the state forests would explicitly focus on problems long regarded as intractable:

- multiple temporal and spatial scales;
- much uncertainty;
- a tendency for scientifically rigorous experiments to be costly and time-consuming; and
- politics currently driving decision making.

Deer effects occur at multiple spatial and temporal scales. Overall forest conditions (the scale of monitoring) typically apply to land areas on the order of tens of square miles. Detectable changes may occur in a 3–5-year range but recovery of degraded areas is expected to take on the order of two decades at a minimum and, for more deer-sensitive components of forest ecosystems, 50–100 years or more. In contrast, silvicultural treatments typically apply to areas of tens or hundreds of acres; temporary fencing in lieu of herd reduction to allow timber regeneration needs to be in place for 5–10 years. All natural resource systems operate at multiple spatial and temporal scales and involve interactions among many component systems. A.R.M. protocol stipulates that alternative management actions need to be based on thorough consideration of multi-scale responses.

There are many uncertainties in the deer-forest system. Which silvicultural prescriptive tools are most effective (e.g., SILVAH versus ORSPA)? Can available deer management tools, especially DMAP, reduce deer density to the degree necessary to meet D.C.N.R.'s forest management goals beyond timber regeneration? To what degree are high deer numbers responsible for forest degradation in comparison with other factors such as acid rain or fire exclusion?

> Adaptive management openly acknowledges uncertainty about how ecological systems function and how they respond to management actions. However, adaptive management is not a random trial-and-error process. Instead, it involves formulating the resource problem, developing conceptual models based on specific assumptions about the structure and function of the resource system, and identifying actions that might be used to resolve the problem. Through the monitoring of outcomes following management interventions, adaptive management promotes improved understanding about which actions work, and why. [Williams *et al.* 2007]

Controlled experiments are the most definitive and fastest way to answer key questions, but they tend to be expensive and some management decisions need to be made sooner than the time it usually takes them to produce conclusive results. It seems paradoxical that controlled experiments are faster than A.R.M. in producing definitive results but not fast enough to inform decisions that need to be made right away. The resolution of this apparent contradiction is that less-than-definitive results will suffice in the short term in an A.R.M. framework; trends inferred from uncontrolled management trials are used to refine or modify management methods, which are then tested further. What is learned in this manner may eventually converge on the results of a full-scale scientific study, but it might take many cycles of trial, monitoring, assessment, modification, and retrial. Decisions needed this year include how many deer to harvest and where, how many acres to treat, what silvicultural prescriptions to apply, and what restoration methods to use. Because A.R.M. compares methods in the course of management that would need to be done whether in an A.R.M. context or not, it is often less expensive than embarking on a full-scale scientific study in addition to the regular management program. Scientifically rigorous research and A.R.M. typically span different parts of a spectrum—though they may overlap in the middle—of higher to lower strictness of standards of evidence, degree of experimental control, and thoroughness of replication (more on this topic in Appendix, page 48).

Political controversy often leads to management impasse. In A.R.M., proponents of competing views sit together at the table as management options are weighed, and a group consensus is sought on which specific management methods should be compared and on what specific monitoring results would support each viewpoint over the other. This process requires "buy-in" by members of groups whose opinions differ, making it more likely that the results of comparing alternative management schemes will be accepted by a wider range of stakeholders.

Effective monitoring is a critically important part of A.R.M. Furthermore, A.R.M. requires setting target ranges of indicator values that differentiate "acceptable" from "unacceptable," recognizing every indicator has a natural range of variation (the importance of quantitative goals is a subject of section 1, page 15). Uncertainty can then be expressed as a set of testable models describing how the deer-forest system might respond under different management scenarios. Predictions of a model often differ at different spatial and temporal scales. They must consider expected results in the short and long term, as well as at fine, medium, and coarse spatial scales. Monitoring at multiple scales provides the data to test these models' predictions.

Implementing A.R.M. requires a strong commitment by executive leadership to a transparent decision-making process, built on measurable objectives and feedback from a credible and sustainable monitoring system. An institutional challenge pointed out in the review discussion is the need for D.C.N.R. staff to acknowledge that effective ecosystem management goes beyond the few system attributes that the agency currently measures. Another is shifting the emphasis of information gathering from tracking activities to monitoring outcomes. Perhaps the biggest challenge is that A.R.M. requires engagement by all major stakeholders, including staunch opponents. Conferees noted that D.C.N.R.'s Ecosystem Management Advisory Committee (EMAC) already includes representatives of many major stakeholder groups.

A.R.M. is structured for decision making to become less politically contentious and more scientifically credible. Monitoring provides data on resource responses to management actions, which feed into an objective assessment process on a regular schedule. Each cycle in the feedback loop is completed when what is learned is used to improve management actions. As one reviewer pointed out, the peer review process that is the subject of this report could serve as a model for the assessment part of this loop. Assessment needs to be firmly rooted in the scientific method, centered on a rigorous, ideally annual, self-appraisal process involving qualified and largely impartial scientists and resource managers. Commitment to the scientific method does not exclude professional judgment and attention to anecdotal evidence, which often play a key role in formulating hypotheses and models to be tested in management trials. A reviewer suggested that assessment could be streamlined if periodic reviews were conducted by a committee with a low rate of turnover. With high continuity, reviewers have familiarity with ongoing research projects over time, resulting in enhanced efficiency and effectiveness.

Sustainably managing state forests from an ecosystem perspective over the long term hinges on forest monitoring results showing definitively (and continually) whether deer management tools are actually accomplishing D.C.N.R.'s ecosystem management objectives. For this reason, according to several members of the review team, the state forest monitoring effort currently lacks focus. Long-term, repeated comparison of indicator responses among areas subjected to different levels of hunting effort—with replication and controls (described in Appendix, p. 48)—would be the most effective way of gaining the necessary information.

What is the difference between A.R.M. and scientific research?

The key operational differences between A.R.M. and scientific research involve standards of evidence, degree of experimental control, and thoroughness of replication (see discussion in Appendix, page 47). However, no clear line demarcates the two approaches. There is a gradation between fully controlled, replicated experiments, which tend to be more expensive, and trials within A.R.M., which tend to be less expensive because they are part of routine management activities. Several, but not all, reviewers expressed the opinion that both A.R.M. and rigorous experiments are needed and should work together to advance knowledge of the best forest management practices from an ecosystem perspective. Scientists at D.C.N.R., the P.G.C., the Cooperative Fish and Wildlife Research Unit at Penn State, the U.S. Forest Service's Forestry Sciences Laboratory, and D.C.N.R.'s other research partners in the state could target experiments on especially difficult questions as funds are available while D.C.N.R. could make steady progress internally on answering less thorny management questions in the course of routine management, once the agency has adopted the A.R.M. approach.

Achieving appropriate standards of scientific rigor within A.R.M. will help D.C.N.R. maximize returns on its monitoring investment by fostering confidence in the results. Several reviewers maintained that major improvements could be made in this area at little added cost simply by configuring the array of monitoring plots and management trials across the landscape in ways that better meet experimental design standards. The more closely the monitoring and design of management trials approach the rigor of a scientific experiment, the faster the learning and the higher the quality of information. There are some tradeoffs between short-term cost and long-term effectiveness; however, it is not necessarily a linear relationship. Instituting small tweaks can sometimes result in large gains.

Summary of short-term recommendations—adaptive resource management

- Facilitate the learning process by configuring monitoring plots and management trials as near to rigorous experimental design standards as can be achieved with available resources.
- Commit to shifting to an adaptive resource management approach in the state forests by establishing a team of D.C.N.R. administrators, managers, scientists, and experienced A.R.M. practitioners to plan, coordinate, and oversee the transition.
- Establish a deer and state forest A.R.M. stakeholder group to help assess challenges, design management activities to address them, and implement and monitor those activities, and to participate in evaluation of results.

Summary of long-term recommendations—adaptive resource management

- Fully adopt the A.R.M. approach, in partnership with other resource management agencies, as the official framework for state forest management from an ecosystem perspective.

7 PUT D.C.N.R.'S USE OF DMAP TO THE TEST

Test the assumption that D.C.N.R.-administered DMAP is an effective tool for sustainable forest management to the degree required to meet ecosystem management goals.

The Department of Conservation and Natural Resources' goal for state forest lands is to "sustain a healthy and functioning forest ecosystem," which includes native species diversity, vertical structure, patch diversity, tree regeneration, structural and age class heterogeneity, and other essential forest ecosystem qualities and processes appropriate to each forest type. To meet its ecosystem management goals, D.C.N.R. may require a lower threshold of deer density than that required simply for tree regeneration. The effectiveness of existing deer management tools, DMAP in particular, to meet D.C.N.R.'s stringent needs has not been fully tested. Because the state forest management program is predicated on having tools that will work to meet its goals, it is critical to test and, if necessary, improve the efficacy of DMAP and how it is applied in the state forests.

No studies, including those that are the subjects of this peer review, have demonstrated an association between the level of localized deer harvesting effort (e.g., DMAP versus ordinary hunting regulations) and deer population density or deer effects on vegetation. In deCalesta's study (2008) on the Kinzua Quality Deer Cooperative Area, the estimated deer density (from pellet-group counts) across the time series from 2002-2007 appears to curve partly in tandem with the "no regeneration" trend curve and opposite from the "no impact" curve (Figure 4, page 9). However, there is a high degree of uncertainty in assuming a causal link between estimated deer numbers and tree seedling counts in this study for several reasons:

- There was no replication or control.
- Monitoring plots were assigned to the category "no impact" if they showed little or no visible signs of browsing on twigs within reach of deer, but twigs—understory shrubs and tree seedlings—are often scarce or absent on sites with a history of prolonged high deer density.
- Pellet-group counts have low scientific credibility as the basis for estimating deer density (see section 5, page 33).

It is possible that subjecting other existing data to different types of analysis than those already performed may yield evidence pertaining to this question (reanalyzing other data is the subject of section 2, page 19). The relevant analysis is to compare the responses of indicators in several DMAP areas with those in several non-DMAP areas, based on the assumption that DMAP, as it is administered in the state forests, results in higher harvest rates. No study to date has made this comparison or gathered data on key indicators including herbaceous plant species, amount of shade at the forest floor, recent timber management history, fencing history, contiguous forest area, and proximity of forest edge. According to Diefenbach, the management program at Raystown Lake most closely resembles an application of the adaptive resource management approach to deer and forests in the entire state, but even there, whether or to what degree DMAP effectively meets the U.S. Army Corps of Engineers' management objectives still has not been demonstrated.

Most, but not all, reviewers supported the recommendation that D.C.N.R. promote a forest restoration study at the scale of a fully replicated and controlled experiment (see Appendix, page 47), arguing that the stakes are too high and the need too urgent to wait for the typically lengthier, less powerful, and in some cases riskier process of A.R.M. to answer the most basic question—Can current tools meet D.C.N.R.'s ecosystem management objectives? Such a study would have several overarching goals:

- Weigh the relative magnitudes of deer and non-deer effects in forest degradation, and address interactions among significant effects.
- Determine how effective DMAP is at increasing antlerless deer harvests in the state forests; how its effectiveness can be increased; whether, when applied effectively, it reduces deer impacts to the degree necessary to meet D.C.N.R.'s goals; and if not, then what its specific limitations are and how they can be remedied.
- Provide the basis for predictive models of forest recovery, leading to standardized triggers for DMAP enrollments or other management tools.

Little research has been done to formulate working hypotheses about how Pennsylvania forests might recover from excessive herbivory. A carefully designed network of exclosures would begin to provide benchmarks and make it possible to distinguish indicator responses to current deer densities from their responses to a host of other influences (see section 4, page 29). Several reviewers expect a wealth of information from long-term monitoring of exclosures to repay in full measure the cost and effort required for their construction and maintenance. However, researchers must recognize and plan for the likelihood that recovery of many species and other ecosystem components will be slow in some areas of the state, even within exclosures.

The forest restoration study could be essentially similar to Diefenbach and Fritsky's study (2007) but on a larger scale. It would include replicated areas (with exclosures) in which deer density is reduced using DMAP and an equal number of control areas (with exclosures) where ordinary hunting regulations apply. To insure that results are broadly applicable, treatment and control areas would need to be distributed widely across the state forest system (while A.R.M. is implemented on all other state forest lands not designated for the study). In order to weigh the relative magnitudes of deer and non-deer effects and address interactions among significant effects, both types of areas would need to be stratified by additional management treatments besides deer management, for example, augmenting the seed supply of native understory species, herbiciding dense rhizomatous fern cover, burning to foster oak regeneration, or liming to restore depleted soil calcium. The point was raised in the review discussion that D.C.N.R.'s interest in DMAP originally was to address deer density "hotspots" but in order to meet its ecosystem management goals the agency will need to use it more widely. Thus, there should be no shortage of potential study areas on state forest lands.

The forest restoration study will be a significant expense but its advocates, the majority of the panel, believe it is key to D.C.N.R.'s timely success in meeting its ecosystem management goals and thus will save resources in the long term. The knowledge gained will have significance beyond management of the Pennsylvania state forest system and so funding from sources in addition to state appropriations may be accessible through partnerships. One reviewer suggested that a full-scale forest restoration study could be achieved by involving all major players (D.C.N.R., P.G.C., U.S. Forest Service, Pennsylvania Cooperative Fish and Wildlife Research Unit) in a collaborative, 5–10-year research project housed at the Pennsylvania Cooperative Fish and Wildlife Research Unit. He judged that the first 5 years' efforts would provide 3–4 years worth of data for development and refinement of the rapid assessment protocol for indicators of deer browsing effects and forest health, and perhaps for further advances on an effective method of quantifying relative deer density. The timeframe and objectives are conducive to being structured as several doctoral students' dissertation research. Participating agency biologists and technicians could provide much field support.

At least one conferee expressed doubt that DMAP could be tested in this way, saying that comparing DMAP with deer hunting under ordinary regulations is not a sufficiently controlled experiment and that the DMAP program is not designed to test the effect of hunter effort. Others countered that the Bureau of Forestry could apply DMAP in a consistent way among experimental replicate areas using additional means to facilitate hunter access and increase rates of participation.

One reason to question DMAP's capability to meet D.C.N.R.'s forest management goals concerns the issue of hunter access. The DMAP program itself only makes extra hunting tags *available* for targeted areas. Effective distribution of tags, level of participation, and how thoroughly hunter efforts are spread throughout the DMAP area are up to the landowner and the tagholders. Hunting effort typically varies from place to place within a DMAP area, as anywhere else, depending on ease of access. Enhancing the spatial distribution of hunting opportunity could improve the hunter access situation; for instance, the Bureau of Forestry could require DMAP tags to be used in particular areas that traditionally are underhunted because of inaccessibility or other reasons.

According to Benner, overall participation by hunters in state forest DMAP areas has been around 80–85% of those issued tags. However, DMAP's effectiveness as a management tool depends on the proportion of hunters who harvest antlerless deer. A study of hunter harvest rates and attitudes (Bhandari *et al.* 2006) showed that 66% of hunters in the Sproul State Forest who responded to questionnaires ranked "to help manage deer population" as an "important" motivation for hunting (although that percentage was lower than for any other reason given). In the same survey, 42% of respondents strongly preferred to hunt bucks, 27% harvested a buck, and 34% harvested an antlerless deer. Rates vary by region and year. In the state forests, the proportion of hunters bagging an antlerless deer has varied among DMAP areas in the same year and among years in the same area, ranging from 11%–27%. It was suggested that adding time to the hunting season—a tool not made available to D.C.N.R. so far—would increase harvest rates, although several conferees were skeptical. They acknowledged that more time would be a benefit to hunters from a recreational opportunity perspective but expressed doubt that it would make any difference in terms of deer management, especially with low to moderate doe hunting rates on state forest lands. This has not been tested. It was suggested that hunters with higher skill levels, so-called "alpha hunters," would be able to visit more DMAP areas if the season were longer. All of these uncertainties could be examined as part of an evaluation of DMAP performance in the state forests.

In many areas of state forest, the key issue may not be the number of tags or length of seasons but the distribution of hunters and their kills. P.G.C. wildlife biologist Steve Liscinsky showed in the 1970s that hunter activity was concentrated around camps. An uneven distribution of hunters helps to shape an uneven distribution of deer. One conferee even suggested that deer might finally be evolving increased abilities to evade human hunters. High hunting mortality rates every year may comprise strong selection, which, if advantageous mutations were to arise, would be expected to have an effect on inherent behavior population-wide.

The Raystown Lake manager's working hypothesis in applying A.R.M. is that hunters need to bring the deer population *below* ecological carrying capacity for a while until the vegetation recovers, before allowing it to increase again to maintenance levels. This reflects what has become one of the basic tenets of restoration ecology. Experience in a broad range of ecosystems has shown repeatedly that time-lag effects and threshold effects are common in the dynamics of both recovery and degradation (see section 3, second half of page 25). Recovery rarely follows the trajectory of degradation in reverse; the recovery threshold is often very different from the degradation threshold. If the threshold deer density for recovery turns out to be lower than the optimal deer density for sustaining healthy forests over the long term, it will be even more important to look for innovative ways of enhancing DMAP's effectiveness. For instance, one way of increasing the antlerless deer harvest under DMAP might be to make it the only antlerless hunting option.

Summary of short-term recommendations—DMAP as an ecosystem management tool

- Foster the political will and access to additional funds that will be necessary to carry out a forest restoration study as a rigorous scientific experiment by persuading policymakers of the urgency of resolving fundamental questions about relative magnitudes of deer and non-deer effects in forest degradation, effectiveness of DMAP at increasing antlerless deer harvests in the state forests, and practical means for D.C.N.R. to meet its goal to "sustain a healthy and functioning forest ecosystem" across the state forest system.
- Launch a research partnership among scientists and managers at D.C.N.R., P.G.C., U.S. Forest Service, and Pennsylvania Cooperative Fish and Wildlife Research Unit to collaborate on a forest restoration study, to include examination of DMAP's effectiveness at reducing deer impacts in state forests and ways of enhancing the spatial distribution of hunting opportunity in order to improve the hunter access situation in DMAP areas.

Summary of long-term recommendations—DMAP as an ecosystem management tool

- Conduct the forest restoration study (in parallel with implementation of A.R.M.), adhering to rigorous scientific standards in experimental design and methods, to test the effectiveness of DMAP and its administration by the Bureau of Forestry in meeting D.C.N.R.'s ecosystem management objectives and to provide the basis for predictive models of forest recovery.
- Identify specific limitations of the current DMAP program based on field and human dimensions studies and implement or request changes to remedy them.

References Cited

Augustine, D. J., L. E. Frelich and P. A. Jordan. 1998. Evidence for two alternate stable states in an ungulate grazing system. *Ecological Applications* **8**: 1260-1269.

Bhandari, P. R. C. Stedman, A. E. Luloff, J. C. Finley and D. R. Diefenbach. 2006. Effort versus motivation: factors affecting antlered and antlerless deer harvest success in Pennsylvania. *Human Dimensions of Wildlife* **11**: 423-436.

Benner, J. M. 2007. Browsing and regeneration monitoring report for Pennsylvania's state forests, 2007. Unpublished manuscript. Pennsylvania Department of Conservation and Natural Resources, Bureau of Forestry. 21 pp.

deCalesta, D. S. 2007. Accuracy and precision of deer pellet group counts for estimating deer density in northwestern Pennsylvania. Unpublished manuscript. U.S.D.A. Forest Service, Northeastern Research Station, Forestry Sciences Laboratory. 15 pp.

deCalesta, D. S. 2008. Deer density and impact [on the Kinzua Quality Deer Cooperative Area]. Unpublished manuscript. U.S.D.A. Forest Service, Northeastern Research Station, Forestry Sciences Laboratory. 27 pp.

Diefenbach, D. R. and R. S. Fritsky. 2007. Developing and testing a rapid assessment protocol for monitoring vegetation changes on state forest lands. Unpublished manuscript. U.S. Geological Survey, Cooperative Fish and Wildlife Research Unit, Pennsylvania State University. 47 pp.

Fike, J. 1999. *Terrestrial and Palustrine Plant Communities of Pennsylvania*. Pennsylvania Department of Conservation and Natural Resources, Harrisburg, The Nature Conservancy, Middletown, Pennsylvania, and Western Pennsylvania Conservancy, Pittsburgh. 86 pp.

Forsyth, D. M., A. R. J. Barker, G. Morris and M. P. Scroggie. 2007. Modeling relationships between fecal pellet indices and deer density. *Journal of Wildlife Management* **71**: 964-970.

Haroldson, B. S., E. P. Wiggers, J. Beringer, L. P. Hansen and J.B. McAninch. 2003. Evaluation of aerial thermal imaging for detecting white-tailed deer in a deciduous forest environment. *Wildlife Society Bulletin* **31**: 1188-1197.

Horsley, S. B., S. L. Stout and D. S. deCalesta. 2003. White-tailed deer impact on the vegetation dynamics of a northern hardwood forest. *Ecological Applications* **13**: 98-118.

Laing, S. E., S. T. Buckland, R. W. Burn, D. Lambie and A. Amphlett. 2003. Dung and nest surveys: estimating decay rates. *Journal of Applied Ecology* **40**: 1102-1111.

Latham, R. E., J. Beyea, M. Benner, C. A. Dunn, M. A. Fajvan, R. Freed, S. B. Horsley A. F. Rhoads and B. P. Shissler. 2005. *Managing White-tailed Deer in Forest Habitat from an Ecosystem Perspective: Pennsylvania Case Study*. Audubon Pennsylvania and the Pennsylvania Habitat Alliance, Harrisburg. 340 pp. (pa.audubon.org/deer_report.html)

Marques, F. F. C., S. T. Buckland, D. Goffin, C. E. Dixon, D. L. Borchers, B. A. Mayle and A. J. Peace. 2001. Estimating deer abundance from line transect surveys of dung: Sika deer in southern Scotland. *Journal of Applied Ecology* **38**: 349-363.

Pennsylvania D.C.N.R. Bureau of Forestry. 2006. Pennsylvania State Forests—Deer Management—A Plan for the Department of Conservation and Natural Resources. Pennsylvania Department of Conservation and Natural Resources, Bureau of Forestry, Harrisburg. 27 pp.

Scheller, R. M., J. B. Domingo, B. R. Sturtevant, J. S. Williams, A. Rudy, E. J. Gustafson and D. J. Mladenoff. 2007. Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecological Modeling* **201**: 409-419.

Shi, H., E. J. Laurent, J. LeBouton, L. Racevskis, K. R. Hall, M. Donovan, R. V. Doepker, M. B. Walters, F. Lupi and J. Liu. 2006. Local spatial modeling of white-tailed deer distribution. *Ecological Modeling* **190**: 171-189.

Williams, B. K., R. C. Szaro and C. D. Shapiro. 2007. *Adaptive Management: The U.S. Department of the Interior Technical Guide*. Adaptive Management Working Group, U.S. Department of the Interior, Washington, DC. 72 pp.

Appendix: Notes on Field Methods

The key operational differences between adaptive resource management (A.R.M.) and scientifically rigorous research involve standards of evidence, degree of experimental control, and thoroughness of replication (see *Adaptive management of deer and forests*, page 37).

Standards of evidence

The standard of evidence is the acceptable level of what statisticians call the type 1 error probability. That is the likelihood of inferring that a management activity caused a measured response when it actually did not. Experimenters in ecology and wildlife biology customarily use an arbitrary cutoff of a 1 in 20 (5%) probability of such an error to decide whether the difference is significant. In other words, if the chance of inferring that a measured difference between two management methods or between management and no management is an effect of the management activity when it actually is not works out to be 5% or less, then the difference is regarded as significant (medical researchers often apply a stricter standard of 1 in 100 or 1 in 1,000). If the difference is large and highly consistent, then it is likely to meet the 5% criterion even within A.R.M. If the difference is subtle or muddled by things that vary from place to place and are beyond the experimenters' control, then experimental replication and control procedures need to be more stringent to detect it. In A.R.M. a lower standard of evidence can be accepted because A.R.M. is inherently a long-term process—management is open-ended, unlike most scientific research projects—and information gained in this way is tested further in years-long cycles of assessing and modifying management methods and continuing to collect and analyze comparative monitoring data.

In A.R.M., monitoring data may be analyzed statistically if they meet the assumptions of statistical methods or, more often, they are examined qualitatively. In either case, results of each monitoring cycle are used as the basis for weighting competing models to explain the effects of management on resources. *Model* in this context is a predictive forest model (a type of predictive ecosystem model), which is a set of rules or equations that allow quantified predictions of how specific indicators of ecosystem conditions will change over time based on a change in a variable of management interest (e.g., antlerless deer harvest effort). Such models are based on a conceptual framework of how ecosystem components interact. The input to a predictive forest model is the starting condition (e.g., deer density, native understory diversity and density, soil buffering capacity, hay-scented fern cover, nonnative plant cover) and relevant variables (e.g., antlerless deer harvest rate, herbicide treatment, soil liming). The output from the model is a quantitative prediction of how indicators will change and over what time span.

A.R.M. decision makers give an initial weighting to competing models as the basis for quantitative decisions (e.g., how many antlerless deer harvest permits will be requested in a given year). For instance, if state forest managers assigned a 90% weight to the view that high deer densities are responsible for losses of the shrub layer, herbaceous diversity and tree regeneration, and a 10% weight to theories attributing the losses to non-deer effects, then the initial number of permits under A.R.M. would end up close to the recommendation that would be made under the deer density model alone. The assigned model weights are then adjusted in future years based on how well each model did in predicting the outcome of the management action. The adjustments can be made using professional judgment, probability theory, or some combination. The exact values of the initial weights are not crucial because the system is self-correcting. The models themselves are subject to updating based on the results of monitoring and new models might be proposed for consideration.

Standards of evidence are an important consideration in how well monitoring data perform as a basis for updating model weights. The higher the standard, the more confident managers can be in assigning weights that will lead to wise management decisions.

Experimental control

Experimental control is achieved in three basic ways; often a mixture of two or all three is used in the same trial or experiment:

- Physically regulate or geographically separate factors likely to affect indicator response that vary spatially or over time and are not part of the hypothesis being tested (for examples, see section 4, page 29). Accomplish this by manipulating the extraneous variables directly to make them as consistent as possible among all trial areas or by choosing trial areas that are as similar to each other as possible, or both.
- Compare indicator responses in management trial areas with those in unmanaged areas or among areas subjected to different management treatments. For example, run management trials and monitor indicators in areas that are similar to each other, subjecting half of them to the management method under trial and leaving half of them alone (the latter are often called control plots or simply controls). Alternatively, subject equal numbers of trial and monitoring sites to each of two or more management methods or combinations of methods.
- Employ a method sometimes referred to as using "positive" and "negative" controls. In examining deer effects on forest ecosystems, this would consist of comparing indicator responses inside and outside of fenced deer exclosures, in both management trial areas and unmanaged control areas. For certain research questions, this two-tiered control approach is among the most powerful methods of producing useful information in the shortest practical amount of time.

Physically controlling variables that are not part of the hypothesis being tested is often costly in funds, time, and labor, and choosing sites for management trials that are nearly alike in those extraneous variables is seldom a luxury that managers have available. A.R.M. does not make such stringent demands. However, even in A.R.M. it is still highly desirable to compare indicator responses in areas that are managed with those in roughly similar areas that are not managed, or to compare among similar areas that are managed in different ways. Control greatly increases the chances of separating indicators' responses to the management activity from their responses to myriad factors that are beyond the managers' control. Some A.R.M. projects have been carried out without any experimental control. Although highly suspect—there is a much greater risk of misinterpreting the results—trends inferred from uncontrolled trials may be used to refine or modify management methods, which can then be tested further.

Replication

Replication is vital to interpreting the results of both controlled experiments and A.R.M., but A.R.M. typically involves fewer replicates. In the context of testing management methods to reduce deer impacts on forest ecosystems, replicates are multiple geographically defined areas in which a particular kind of management activity is applied (or areas left unmanaged to serve as controls). The average size of the Bureau of Forestry's DMAP properties is on the order of 6,000 acres; each such area could be a replicate management trial or control area.

It is worth going into detail on two points to clear up common misconceptions about replication:

- Monitoring plots are not replicates, unless there is only one monitoring plot per management trial area (which would be inadequate to reflect the responses of indicators across an entire trial area). The data from all of the monitoring plots within a replicate management trial area are averaged together and the replicate-wide averages are the only data used in analyses. Treating the data from individual monitoring plots as replicates instead of as subsamples, a common mistake, violates fundamental assumptions underlying the logic of inductive reasoning and leads to false or misleading inferences.
- True replication entails spatial interspersion of treatments, that is, replicates cannot be clumped geographically by type. Where areas with the same management activity or combination of management activities (or control areas) are clumped together on the landscape, then the areas within a clump would have to be considered as subsamples, and the entire clump of areas—not the individual areas within it—as a single replicate. Without interspersion of treatments the trivial effect of spatial autocorrelation (the tendency of nearby samples to be more similar than more distant samples) is confounded with, and cannot reliably be separated statistically from, potentially interesting effects such as those resulting from management trials. The solution is to make sure that trial areas treated with different management methods, including no management (control areas), are interspersed or alternate with each other across the entire set of areas included in the study.

Replication is mandatory for controlled experiments and at least minimal replication is highly desirable for A.R.M. Replication with interspersion (and control) is the quickest and most effective way to separate the effects of a management activity from local differences in factors that are beyond the managers' control or beyond the power of the monitoring protocol to detect. Without at least minimal replication, the risk is high that some local peculiarity of a trial area will result in an effect that will be misinterpreted as having been caused by the management activity, or in the case of a control area, by its absence. In A.R.M. a level of replication as low as $N = 3$ or even $N = 2$ (N is the smallest number of areas in *each category* of management trials, e.g., management type A, management type B, and control) may show clear enough trends for inferences to be drawn, which can be used to refine or modify management methods for further testing.