NPS Inventory and Monitoring Program
Development of Protocols and Sampling Strategies for Invasive Species

Title: GIS-Based Adaptive Sampling for Early-Detection Monitoring of Invasive Species in National Parks: Incorporating Habitat Models and Incidental Reports into Sample Selection

USGS principal contacts: David Smith, USGS – Leetown Science Center, 304-724-4467, david_r_smith@usgs.gov; John Young, Leetown Science Center, Aquatic Ecology Lab, 304-724-4469, john_a_young@usgs.gov; Frank van Manen, Leetown Science Center, Southern Appalachian Field Lab, 865-974-0200, vanmanen@utk.edu

NPS principal contact: Wendy Cass, Botanist, Shenandoah National Park, Luray, Virginia

Overview: We propose to combine adaptive sampling with auxiliary information from habitat models and incidental reports for effective early-detection monitoring of invasive species in National Parks. Adaptive sampling is a probability-based sampling design that allows sampling effort to increase in response to observing rare and clustered organisms. The design was first proposed by Thompson (1990) and has been applied to a wide variety of biological populations since then, including rare plants and trees (Brown 1994, Acharya et al. 2000, Smith et al. 2004). Benefits of adaptive sampling include increased counts of rare individuals and detections of rare species (Smith et al. 2003) and increased precision of abundance estimates (Brown 1994, Acharaya et al. 2000). Adaptive sampling is a promising design for early-detection monitoring of invasive species because the objectives are to detect presence and quantify abundance while the invasive species is rare and before it becomes widespread.

Information from habitat model predictions and incidental reports can be useful for sample selection by allocating sampling effort to areas where the invasive species are likely to first appear and spread. For example, Young et al. (2003) developed habitat models for plant species that are illegally harvested in eastern national parks for the medicinal plant trade. Predictive GIS models of species occurrence were developed based on known plant locations. Field searches in Shenandoah National Park based on the habitat models were 3.3 to 11.2 times more likely to result in finding the species of interest than were random searches. Although the approach substantially improved the probability of finding the species in question over random searches, sampling procedures could be substantially improved. Moreover, new methods are needed to incorporate incidental reports of species presence in areas where data gaps limit model building. Assessments of invasive species occurrence and spread can be approached using similar methods making development of new sampling approaches especially timely and relevant.

Auxiliary information can be incorporated into an adaptive sampling design either in the selection of the initial sample or in setting the condition to adapt. A major challenge in application of adaptive sampling is presented by the need to sample what are called 'edge units', which help define the edges of ecological clusters, but do not contribute to abundance estimation. One strategy to reduce the cost of sampling edge units is to base

the condition to adapt on predictions from habitat models. All of the sampling units (initial and adaptive) can be selected within a GIS, which would save time and money because edge units would not have to be visited in the field. For surveys that are repeated, as in monitoring programs, habitat models can be updated and improved using data collected in past surveys and recent incidental reports.

We propose to develop and evaluate a GIS-based adaptive sampling design for early-detection of invasive species in National Parks by using rare plant species as a surrogate. Rare plants are an appropriate surrogate to address early detection of invasive species because in both instances the species are relatively difficult to find and population abundance is difficult to estimate precisely. Data on plant distributions in Shenandoah National Park, along with fully developed habitat models and GIS coverages, will provide a reliable basis for the simulation study (Young et al. 2003). Various adaptive and conventional sampling designs will be constructed that make use of predictions from habitat models and incidental reports. Performance of the designs in terms of ability to identify hot spots, estimator precision, and yield of rare individuals and species will be compared to a baseline design (e.g., conventional simple random sampling) to identify designs with the greatest efficiency and yield. After establishing through the simulation study that a GIS-based adaptive sampling design is effective for early-detection monitoring, we propose additional 'out-year' phases of work including field trials focusing on particular invasive species and implementation of a monitoring program (see the "Out-year Plans" section of this proposal).

Objectives:

1) Develop a GIS-based adaptive sampling design that incorporates information on habitat value and incidental reports for early-detection monitoring of invasive species.
2) Evaluate the performance of GIS-based adaptive sampling designs relative to other adaptive and conventional sampling designs by using rare plant species as a surrogate.

Need:

The Union of Concerned Scientists, on behalf of the National Environmental Coalition on Invasive Species (NECIS), recently announced the release of the Call To Action on Invasive Species, which identifies the need for early detection/rapid response programs.

The NPS Inventory and Monitoring Program has identified the development of protocols and sampling strategies for invasive species as a major research need. By providing an effective sampling design for detection of rare organisms and species, the proposed research addresses the needs for early-detection monitoring. Because the sampling design will result in estimates of abundance, the proposed research addresses the needs for monitoring the effectiveness of management actions to control or eradicate invasive species.

Identifying effective and efficient sampling designs is important not only to the study of invasive species, but also to the study of rare, at-risk organisms in general. Products of

this research will be broadly useful to those involved in conservation of rare species and will be of interest to biometricians working on efficient sampling designs for rare and clustered organisms.

Procedures/methods:
Adaptive sampling begins by selecting an initial sample according to a probability-based design (e.g., simple random sample [SRS], systematic, stratified random, probability proportional to size [PPS]). Using the initial sample as a starting point, sampling units are selected adaptively if a predefined condition is met. The condition is very flexible and could be based on auxiliary variables, such as habitat model predictions, as we propose here. Adaptive sampling continues until the condition is no longer met. Thompson (1990) modified the Hansen-Hurwitz and Horvitz-Thompson estimators for unbiased estimation of abundance from adaptively-collected data.

Auxiliary information, habitat model predictions and incidental reports, will be incorporated at two places in the sampling design. First, the auxiliary information will be used in the selection of the initial sample through the initial sampling design by application of PPS sampling and stratification. Second, the condition to adapt will be a function of habitat model predictions and occurrence of incidental reports in the vicinity of the sampling unit. Habitat information was incorporated similarly in an evaluation of adaptive sampling of waterfowl and was shown to be efficient (Smith et al. 1995). Because this information is available prior to field sampling, a GIS can be used for sample selection and sampling effort can be directed to areas where target species are known or suspected to occur.

To evaluate the sampling designs, we will incorporate plant habitat models developed by Young et al. (2003) for Shenandoah National Park, Virginia. Locations of 3 plant species were used as the point of reference to delineate favorable habitats. In the complex topography of the Appalachian Mountains, highly variable terrain characteristics form indirect gradients that have a strong influence on the distribution and productivity of plant species. We used several topographic variables, mapped as 10- x 10-m grid cells, to represent these indirect gradients, including aspect, slope, solar isolation, terrain shape index, topographic convergence index, and topographic relative moisture index. Categorical variables included land-cover type and geology or coercion type. We used those variables in combination with a multivariate statistical technique, Mahalanobis distance, to calculate and map an index of potential habitat quality (Clark et al. 1993). We used logistic regression to test model performance with an independent dataset of plant locations from Shenandoah National Park, which indicated a significant relationship between Mahalanobis distance values and species occurrence.

We will use the plant habitat models of Young et al. (2003) to generate populations for models that are weakly to strongly predictive by manipulating error distributions in the models. The resulting populations will be the basis for simulation studies. Sampling will be simulated over a variety of conventional and adaptive techniques, focusing on designs that incorporate information on habitat and incidental reports. Incidental reports will be generated as random events. Averages over replicated simulations will be taken as

expected values for comparison of precision, bias, and cost. We envision conducting much of the simulation work using Visual Basic extensions for ArcGIS™ software.

Expected results or products:
The results of this study will be an evaluation of developed protocols for early detection of invasive species by integrating adaptive sampling designs with predictive habitat models and incidental observations. Products will include a final report with recommendations, a manuscript in a peer-reviewed journal, conference presentations, and Visual Basic extensions for ArcGIS™ software.

Technology/information transfer:
One of the products from this research would be the development of a visual basic extension for use with ArcGIS™ software. This GIS-based tool can be used by NPS Inventory and Monitoring Program personnel to develop sampling regimes based on incidental observations of newly detected invasive species.

Work Schedule:

July 04 to Sep 04	Design development, GIS data management, model building,
Oct 04 to Dec 04	Software programming, computer simulation
Jan 05 to June 05	Complete simulations, report writing, software documentation

Budget:

Activity	Cost
Personnel	
Programmer (6 pay periods.; $1,994 per pp)	$11,964
GIS technician (6 pay periods; $1,587 per pp)	$9,522
Travel	$1,500
Supplies	$1,000
Indirect costs (4%)	$960
Total	$24,950

Out-year plans: Contingent on the outcome of the evaluations, we propose to continue this work by conducting field trials of the sampling protocols based on rare or invasive species. The final phase would be to work with NPS to establish a pilot monitoring program at a chosen national park that follows the proposed protocol. Budgets and work schedules for these additional phases of work have not been developed. However, we would be interested in pursuing this work under the NPS Inventory and Monitoring Program (or similar program).

Literature cited:
Acharya, B., G. Bhattarai, A. de Gier, and A. Stein. 2000. Systematic adaptive cluster sampling for the assessment of rare tree species in Nepal. *Forest Ecology and Management* 137:65-73.

Brown, J. A. 1994. The application of adaptive cluster sampling to ecological studies. Pages 86-97 in D. J. Fletcher and B. F. J. Manly, eds., *Statistics in Ecology and Environmental Monitoring*. Dunedin: University of Otago Press.

Clark, J. D., J. E. Dunn, and K. G. Smith. 1993. A multivariate model of female black bear habitat use for a geographic information system. Journal of Wildlife Management 57:519-526.

Smith, D. R., M. J. Conroy, and D. H. Brakhage. 1995. Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. Biometrics 51:777-788.

Smith, D. R., R. F. Villella, and D. P. Lemarié. 2003. Application of adaptive cluster sampling to low-density populations of freshwater mussels. *Environmental and Ecological Statistics* 10:7-15.

Smith, D. R., J. A. Brown, and N. C. H. Lo. 2004. Application of adaptive cluster sampling to biological populations. In W. L. Thompson (ed.) Sampling rare or elusive species: Concepts, designs, and techniques for estimating population parameters. Island Press, Covelo, CA

Thompson, S. K. 1990. Adaptive cluster sampling. *Journal of the American Statistical Association* 85:1050-1059.

Young, J.A., F.T. van Manen, and C.A. Thatcher. 2003. Habitat modeling for protection of illegally harvested plants in National Parks of the Blue Ridge Mountains. Final Report to the National Park Service. U.S. Geological Survey, Aquatic Ecology Branch, Kearneysville, West Virginia and Southern Appalachian Field Branch, Knoxville, Tennessee, USA. [unpublished] 74 pp.

Paul:

Attached please find our proposal in response to your recent RFP on sampling designs for invasive species. We attempted to address the NPS needs in a cost-effective manner by proposing a simulation study on the application of adaptive sampling to early-detection monitoring. We are very excited about the prospects of working on this project and look forward to developing effective sampling strategies for invasive species.

Best regards,
Dave

Dave Smith
david_r_smith@usgs.gov