An Optimization Approach to Selecting Research Natural Areas in National Forests

Stephanie A. Snyder, Lucy E. Tyrrell, and Robert G. Haight

ABSTRACT. The USDA Forest Service has a long-established program to identify areas in national forests for designation as protected Research Natural Areas (RNAs). One of the goals is to protect high quality examples of regional ecosystems for the purposes of maintaining biological diversity, conducting nonmanipulative research and monitoring, and fostering education. When RNA designation conflicts with other land uses, difficult choices must be made about the best number and location of sites. We addressed this problem by adapting a classic optimization formulation from the location science literature. The formulation was an integer optimization model for selecting the set of RNAs that maximized the number of regional ecosystems and natural communities represented subject to an upper bound on the total area covered by the sites in the selected set. We applied the formulation using 33 potential RNAs in the Superior National Forest in northeastern Minnesota. The 33 potential RNAs were chosen for our case study because they had been mapped and field-surveyed for the presence of natural communities. The use of those sites does not imply that other areas in the Superior National Forest do not merit further study as RNA candidates. The model quickly generated information about the trade-offs between different protection goals. We found multiple sets of potential RNAs, ranging from all 33 sites to a much smaller set of 21 sites, that attained the specified goals for natural community representation. Thus, the decision-maker can choose among sets of sites with a wide range of total areas without compromising the representation goal. We also found that requirements to choose a set of sites that represents a range of locally defined ecosystems or priority natural communities can limit the total number of natural communities that can be represented within a set of sites of a given area. Average solution times for different problems were less than 5 seconds on a personal computer, suggesting that integer optimization can readily facilitate investigation of the impacts of RNA selection goals. For. Sci. 45(3):458–469.

Additional Key Words: Forest planning, integer optimization, land-use planning, optimization modeling, reserve selection, Research Natural Areas, Superior National Forest.

The USDA Forest Service has a long-established program to identify areas in national forests for designation as protected Research Natural Areas (RNAs) (USDA 1992). Between 1927 and 1998, the Forest Service established 427 RNAs covering over 208,960 ha of land. Ranging in size from less than 10 ha to more than 4,000 ha, RNAs are lands that are permanently protected and maintained in their natural condition for the purposes of maintaining biological diversity, conducting nonmanipulative research and monitoring, and fostering education. RNAs are designed to include assemblages of species and ecosystems of scientific interest, as well as high-quality examples of

Stephanie A. Snyder (correspondence author) is a Research and Systems Analyst, Minnesota Department of Transportation, 1110 Centre Pointe Curve, Mendota Heights, MN 55120—E-mail: snyder01@yahoo.com. Lucy E. Tyrrell is Regional Research Natural Area Coordinator, USDA Forest Service, North Central Research Station, 5985 Highway K, Rhinelander, WI 54501; and Robert G. Haight is Research Forester, USDA Forest Service, North Central Research Station, 1992 Folwell Ave., St. Paul, MN 55108.

Acknowledgments: This research was supported by the North Central Research Station and the Forest Service Research Natural Area Program. The authors thank Chel Anderson and Carmen Converse of the Minnesota Department of Natural Resources for providing field survey information for potential Candidate Research Natural Areas in the Superior National Forest, Mark Nelson of the North Central Research Station for assisting with site mapping, and Kristin Snow of The Nature Conservancy for providing crosswalks between natural community classifications.

Manuscript received July 16, 1998. Accepted February 4, 1999.

Copyright © 1999 by the Society of American Foresters

Reprinted from *Forest Science*, Vol. 45, No. 3, August 1999. Not for further reproduction.

widespread species and ecosystems that are not included in other protected areas. By protecting ecosystems, RNAs complement endangered species programs that focus on protecting single species and ecosystem management programs that focus on biodiversity protection on actively managed lands.

One of the goals of the RNA Program is to continue to identify and designate RNAs that include aquatic and terrestrial ecosystems not adequately represented by the current set of RNAs (USDA Forest Service 1992). For example, in national forests in the Eastern Region of the United States, ecosystems targeted for inclusion in RNAs can be identified using a proposed framework (see Appendix of Faber-Langendoen et al. 1999a) that incorporates the Forest Service National Hierarchy of Ecological Units (McNab and Avers 1994, Keys et al. 1995) and The Nature Conservancy's National Vegetation Classification (Grossman et al. 1994). Another goal for RNA selection is to increase the size of protected areas so that they include, to the extent possible, several ecosystems and landscape-level processes and mitigate the adverse effects of management activities outside their boundaries.

Once a set of potential RNAs that includes the targeted ecosystems has been identified, selecting areas for RNA designation is not an easy or straightforward task. A haphazard selection of RNAs may not contribute to biodiversity protection or monitoring goals if key land types or natural communities are not protected. Furthermore, because RNA designation competes with alternative land uses, some of which may produce marketable outputs, an ad hoc selection of sites may incur greater economic losses than actually necessary. Consequently, the number, size, and location of RNAs should be determined with care.

The purpose of this article is to demonstrate how optimization models can be used to systematically evaluate alternative sets of potential RNAs and associated trade-offs. Although there is a large literature on the application of optimization models to reserve site selection, only recently have these approaches been applied to the problem of selecting RNA sites (Stoms et al. 1998). Our study focuses on the Superior National Forest in northeastern Minnesota. The Superior is a good national forest to analyze for RNA selection because over 90 relatively undisturbed areas have been identified as a broad pool of candidates for RNA or other special designation (Vora 1997), the natural communities in over 30 of these sites are known (Anderson 1997), and the goals for ecosystem representation are well defined (Faber-Langendoen et al. 1999b). The modeling approach described here for the Superior National Forest can be used to guide RNA establishment in other national forests.

Literature Review

To aid in the selection, design, and evaluation of protected areas, the reserve selection problem has been examined by researchers in a number of disciplines. Although a variety of ecological or biological protection goals exist (Pressey et al. 1993), two reserve selection problems are commonly addressed in the conservation biology literature: (1) maximize the number of species or ecosystems that can be represented within a specified number of sites, or (2) identify the smallest number of sites sufficient to represent all of the species or ecosystems of concern.

The first quantitative methods developed to solve these problems were scoring and ranking procedures based on metrics such as reserve size or species richness (e.g., Kirkpatrick 1983, Margules et al. 1988, Cocks and Baird 1989). Sites are selected for protection in a sequential fashion based on rank until a resource constraint, such as cost or total area protected, is reached (e.g., Margules and Usher 1981, Terborgh and Winter 1983, Pressey and Nicholls 1989). A significant drawback of this approach is that sites are scored and selected independently of the composition of previously ranked sites. As a result, strictly choosing the highest ranked sites may lead to solutions that are ineffectual or inefficient; for example, when some species or ecosystems are not represented at all or when other species are represented more than once.

A second approach to solving reserve selection problems involves greedy heuristics (Margules et al. 1988, Rebelo and Siegfried 1990, Vane-Wright et al. 1991, Bedward et al. 1992, Nicholls and Margules 1993, Pressey et al. 1993). Like the scoring and ranking methods, greedy heuristics identify a prioritized sequence of sites for reserve status. The first step is to select the best site in terms of the principal selection criterion (e.g., species richness). Next, the value of each remaining site is calculated, accounting for the species already represented. The site that best supplements the species represented in previously chosen sites is selected. This recalculation and selection continues until an appropriate resource constraint (e.g., total cost) or a stopping rule (e.g., all species of concern are represented) is met. In contrast to scoring and ranking methods, greedy heuristics avoid redundancy or omissions of representation by accounting for species represented in previously selected sites and the species still in need of representation. The principal drawback of greedy heuristics is that they do not guarantee optimal solutions (e.g., finding the maximum number of species that can be represented by a specified number of sites or finding the smallest number of sites sufficient to represent all of the species or ecosystems of concern), and there is no way to determine the degree of suboptimality (Underhill 1994). Work is continuing in the development of more sophisticated heuristics, including simulated annealing and genetic algorithms (Pressey et al. 1996), which may provide better approximations of optimal solutions.

A third approach involves integer optimization models (Cocks and Baird 1989, Saetersdal et al. 1993, Camm et al. 1996, Church et al. 1996, Davis and Stoms 1996, Willis et al. 1996, Williams and ReVelle 1996, 1997, 1998). Church et al. (1996) pointed out that the two reserve selection problems commonly addressed in the conservation biology literature are applications or modifications of two classic formulations from the location science literature: the maximal covering location problem (Church and ReVelle 1974), which maximizes the number of entities or amount of demand which could be covered or represented by a specified number of

facilities, and the location set covering problem (Toregas and ReVelle 1973), which minimizes the number of facilities necessary to cover or represent all demand nodes. Both of these formulations are amenable to integer optimization, which guarantees optimal mathematical solutions. Optimization differs from scoring and heuristic approaches by identifying and evaluating entire sets of sites according to the selection criteria, rather than sequentially selecting sites based upon the characteristics of the sites chosen in previous iterations. Furthermore, in contrast to scoring and heuristic approaches, the solutions derived from optimization models in no way depend on nor are they sensitive to starting conditions or order of site selection.

The drawback of integer optimization models is that reserve selection problems can be difficult to solve. For example, applications of the reserve selection problem with large numbers of potential sites or complex representation or coverage requirements sometimes require excessive solution times (e.g., Pressey et al. 1996, 1997). Church et al. (1996) suggested, however, that most specifications of reserve selection problems should be small enough to readily solve to optimality in a reasonable amount of time. Performance comparisons of heuristic and integer optimization models of the reserve selection problem indicate that the ultimate choice of solution procedure should depend upon the problem size and complexity and the importance of achieving the true, optimal solution (Csuti et al. 1997, Pressey et al. 1997). If problem size prohibits the use of integer optimization, then smaller problems can be solved using both approaches and compared, allowing assessment of the performance, or distance from optimality, of the heuristic. We used integer optimization to solve our RNA selection problem because we wanted to generate optimal solutions and use those solutions to estimate the trade-offs between different management goals. Optimization readily allows determination of the changes in optimal sets of reserves and ecosystem representation as a function of changes in resource constraints. The ability to quickly and accurately produce optimal sets of reserves under alternative resource constraints is of great use to decision makers who require explicit, well-documented justification for land-use decisions. The computational limitations of integer optimization did not apply in our case because the problems were relatively small.

Methods

Study Area

The Superior National Forest is the largest national forest in the eastern United States, covering over 850,500 ha in northeastern Minnesota (Figure 1) (USDA Forest Service 1986).

The major landforms include glaciated plains with low mountains, 180-670 m in elevation. Dominant vegetation is transitional between boreal forests to the north and the largely deciduous forests to the south. Major forest types include jack pine (*Pinus banksiana*), white pine–red pine (*Pinus strobus–Pinus resinosa*), white spruce–balsam fir (*Picea glauca–Abies balsamea*), aspen–paper birch (*Populus tremuloides–Betula papyrifera*), tamarack–black spruce, (*Larix laricina–Picea mariana*), and sugar maple–paper birch (*Acer saccharum–Betula papyrifera*).

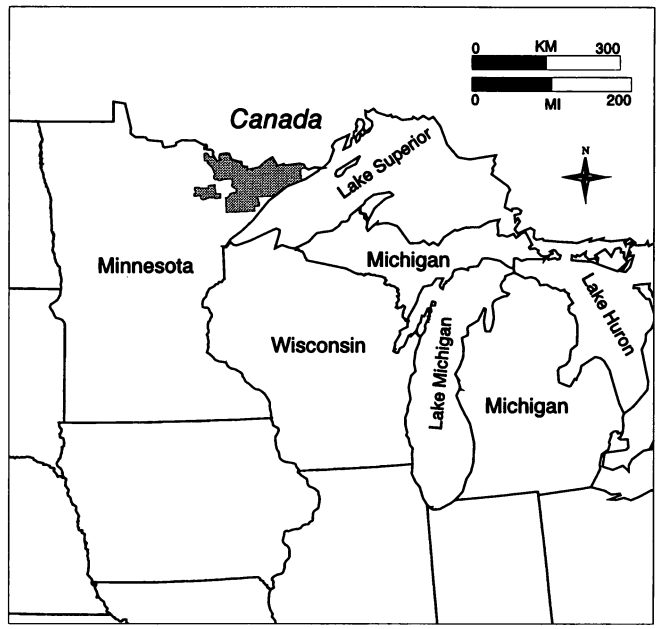

Figure 1. Location of the Superior National Forest (shaded area).

The Boundary Water Canoe Area Wilderness covers the northern third of the forest. Outside of the wilderness area, the primary uses of the forest are commercial timber production, semiprimitive and developed recreation, and wildlife habitat (USDA 1986).

Research Natural Area Selection Goals

The proposed draft framework for representing in RNAs the range of ecosystems found in national forests in the Eastern Region of the United States contains two kinds of representation goals (see Appendix of Faber-Langendoen et al. 1999a). The framework includes a goal for ecological landtype representation based on the National Hierarchy of Ecological Units, a system of mapped land units used in national forest planning at multiple spatial scales (McNab and Avers 1994, Keys et al. 1995). The National Hierarchy is incorporated into the framework so that the system of RNAs provides baseline, high quality, relatively intact ecosystems based on the same ecological units that are used in many aspects of national forest planning. In addition, the framework includes a goal for natural community representation to assure representation of the full range of vegetation types that occur within each ecological land unit.

We addressed both of those goals in our basic model formulation. First, we wanted the model to select a set of high quality sites to represent each regional landtype. Regional land types were defined using the Subsection level of the National Hierarchy of Ecological Units. A Subsection is a unit of land distinguished by climate and geologic structure covering thousands of square kilometers. Five Subsections are present in the Superior National Forest, shown by the shaded areas with alphanumeric labels in Figure 2. Brief descriptions of the Subsections can be found in Keys et al. (1995). Subsection boundaries were modified to be consis-

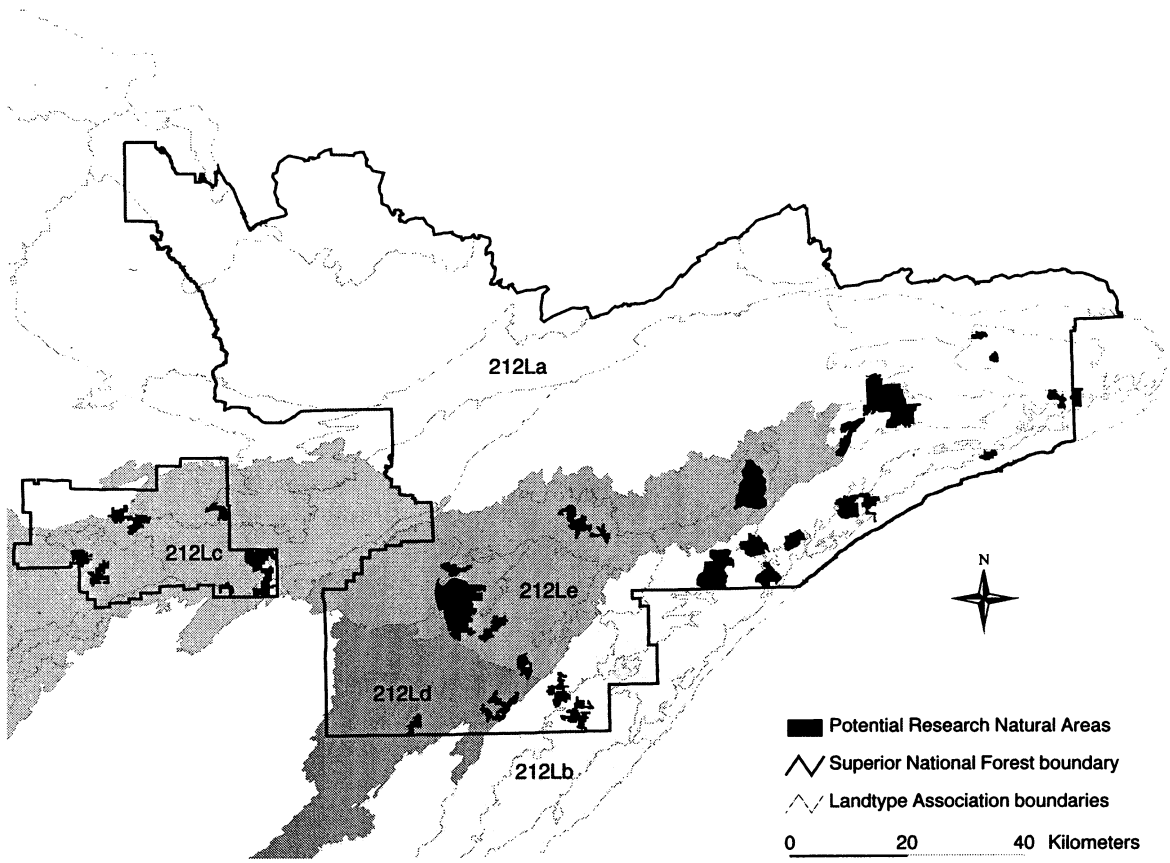

Figure 2. Ecological units in the Superior National Forest and potential Research Natural Areas used in the optimization models. Lightly shaded areas with alphanumeric labels represent Subsections.

tent with those used in national forest planning based on a 1997 agreement between the Minnesota Department of Natural Resources and the Superior National Forest.

Second, we wanted the model to select sites that included the range of vegetation communities within each Subsection. Communities were defined using The Nature Conservancy's (TNC) National Terrestrial Vegetation Classification, which classifies communities hierarchically based upon both forested and nonforested vegetation (Grossman et al. 1994). The community level of the hierarchy that was used for RNA selection was the Alliance. Alliances were defined by the dominant species of the uppermost level or canopy of vegetation. For forested types, Alliances were analogous to Society of American Foresters' cover types. Thirty-seven Alliances were known to be present in one or more Subsections in the Superior National Forest (Table 1) (Faber-Langendoen et al. 1999b). For descriptive purposes, the alliances were grouped into three categories based on their current or past landcover in an ecoregion. Matrix alliances cover (or historically covered) extensive, unfragmented parts of the landscape. Large and small patch alliances cover less area and occur in discontinuous patches.

Combining the goals for landtype and plant community representation, we cross-tabulated those Alliances present in the Superior National Forest with Subsections to give 132 Subsection-Alliance pairs represented by cells with black squares in Table 1. Each Subsection-Alliance pair was an ecological feature in need of representation in our basic model.

In addition to the goals for Subsection and Alliance representation, the proposed framework considers representation of local land types. Therefore, we modified the basic model to include a third goal of representing the range of Landtype Associations (LTAs) within each Subsection. LTAs are divisions of Subsections based on phases of soil subgroups, local climate, and plant associations. Each Subsection contained a number of LTAs whose boundaries are indicated by the narrow lines in Figure 2.

A fourth goal of RNA selection in our model was to represent priority natural communities identified by a regional RNA assessment (Faber-Langendoen et al. 1999a,b). When resource limitations prevent the immediate protection of examples of all plant communities and land types, a long-range approach can be adopted in which natural communities are prioritized and assigned time frames for protection. We identified a set of high-priority Subsection-Alliance pairs for representation (cells with checkmarks in Table 1) based on recommendations for the Superior National Forest (Faber-Langendoen et al. 1999b), and we modified our basic model to require representation of these priority communities. Priority was given to all of the matrix Alliances in Subsections 212Lc and 212Le, and the White Pine–(Red Pine)–Quaking Aspen Forest Alliance in Subsection 212Lb. These matrix Alliances were dominant landscape features in the past and are not currently represented in protected areas in these Subsections, which occupy a large portion of the Superior National Forest.

Table 1. Alliances and Subsections in the Superior National Forest.

	Subsections				
	212La	212Lb	212Lc	212Ld	212Le
Matrix Alliances					
Jack Pine Forest Alliance					☆ ✔
Red Pine Forest Alliance	☆ ■	■	☆ ■✔	■	✔
White Pine-(Red Pine)-Quaking Aspen Forest Alliance	■	✔	☆ ■✔	■	■
White Pine Forest Alliance	☆ ■	■		■	☆ ■✔
White Spruce-Balsam Fir-Quaking Aspen Forest Alliance	☆ ■	☆ ■	☆ ■✔	☆ ■	☆ ■✔
White Spruce - Balsam Fir Forest Alliance	■	■		■	
Sugar Maple-Yellow Birch-(American Beech)Forest Alliance	☆ ■	☆ ■	☆ ■✔	☆ ■	☆ ■✔
Quaking Aspen-Paper Birch Forest Alliance	☆ ■	■	☆ ■✔	☆ ■	☆ ■✔
Large Patch Alliances					
Black Spruce Forest Alliance	☆ ■		■	■	■
Eastern White Cedar-Yellow Birch Forest Alliance	☆ ■		■		■
Eastern White Cedar Forest Alliance	☆ ■	☆ ■	☆ ■	☆ ■	☆ ■
Red Oak-Sugar Maple-(White Oak) Forest Alliance			☆ ■		■
Paper Birch Forest Alliance	☆ ■	☆ ■	☆ ■	■	☆ ■
Silver Maple Temporarily Flooded Forest Alliance			■	■	■
Black Spruce Saturated Forest Alliance	☆ ■	☆ ■	☆ ■	☆ ■	☆ ■
Eastern White Cedar Saturated Forest Alliance	☆ ■	☆ ■	☆ ■		☆ ■
Black Ash-Red Maple Saturated Forest Alliance	■	☆ ■	☆ ■		
Tamarack Saturated Forest Alliance	■				
Black Spruce Saturated Woodland Alliance	■		☆ ■		☆ ■
Speckled Alder Seasonally Flooded Shrubland Alliance	■	☆ ■	☆ ■	☆ ■	☆ ■
Leatherleaf Saturated Dwarf-Shrubland Alliance	☆ ■	☆ ■	☆ ■	■	☆ ■
Cattail-(Bulrush) Semipermanently Flooded Herbaceous	■	■	■	■	☆ ■
Few-Seeded, Wiregrass Sedge Saturated Herbaceous	☆ ■	☆ ■		■	
Jack Pine / Lichen Nonvascular Alliance	■				
Rock Outcrop/Butte Sparse Vegetation		☆ ■			
Small Patch Alliances					
(Jack Pine, Red Pine) Woodland Alliance	■				
Red Osier Dogwood-Willow Seasonally Flooded Shrubland	■				
Bog Birch-(Willow) Saturated Shrubland Alliance	☆ ■		■	■	■
Sedge (*C. rostrata, uticulata*) Seasonally Flooded Herbaceous	■				☆ ■
Wild Rice Semipermanently Flooded Herbaceous	■		■	■	■
Wiregrass Sedge Saturated Herbaceous	■	■	■	■	■
Yellow-Whte Water Lily Permanently Flooded Herbaceous	■	■	■	■	■
Open Bluff/Cliff Sparse Vegetation		☆ ■			
Open Pavement Sparse Vegetation	■	■	■		■
Cobble/Gravel Shore Sparse Vegetation	■		■		☆ ■
Inland Strand Beach Sparse Vegetation	■	■	■	■	■
Non-Tidal Mud Flat Seasonally/Temporarily Flooded Sparse	■	■	■	■	■

LEGEND

■	Present in the Superior National Forest
☆ ■	Present in the Superior National Forest and represented by the set of potential RNAs
☆ ■✔	High-priority Subsection-Alliance pair

Potential Research Natural Areas

Our analysis was conducted using 33 potential RNAs in the Superior National Forest. The sites were part of a larger set of 93 sites that were identified in 1996 using maps and high-altitude aerial photography as potential representatives of the highest quality remaining examples of characteristic ecosystems in each LTA (Vora 1997). A rapid assessment using aerial and ground surveys was conducted in 1997 to map boundaries and inventory natural communities in at least two sites in each LTA (Anderson 1997). We selected 33 sites for our analysis because mapping and field surveying had been completed at the time of our study. The use of these 33 potential RNAs does not imply that the other sites do not merit further study as RNA candidates or that the 33 sites provide examples of all the natural communities in need of representation.

Locations and natural communities of the 33 potential RNAs were described by Anderson (1997). The sites range in size from 269–7,524 ha (Table 2) and include one or more high quality, relatively undisturbed examples of plant com-

Table 2. Potential Research Natural Areas in the Superior National Forest.

Potential RNA	Subsection	LTA(s)	Area (ha)
Cascade Lake	212La	13,36,37	6,867
Locket Lake		35	320
Lullaby Creek		35	269
Rice Chain		36	1,574
Trout Lake		37	671
Barker Lake	212Lb	8,10	1,746
Beaver River		10	510
Cabin Creek		5,11	3,936
Fall River Fen		2,10	421
Heartbreak Creek		8,11	1,031
Lillian Creek South		3	1,391
Lutsen Addition		2,8,10	835
Pearl Lake		5,8,11	1,636
Schroeder Addition		8,11	1,415
South Brule River		10	637
Watertank Lake		10	1,097
Candle Lake	212Lc	5	758
Deepwater Lake		20	895
Loka Lake		5	1,652
Pike Mountain		5,10	363
Rice Lakes		6,7	723
Slow Creek		20	1,136
Watercress Lake		5,7	1,051
Wynne Creek		5,10	1,294
Sullivan Creek	212Ld	1	1,633
Wolf Lake		1	595
Big Lake - 7 Beavers	212Le	8,10,11	7,524
Dragon Lake		1	1,530
Dunka		1,8,11	789
Little Isabella River		1	445
South Greenwood Creek		10	1,051
Timber-Frear		2,3,4	4,408
White Pine Picnic		10	860

munities. Collectively, the 33 sites cover a total of 51,030 ha. Each Subsection contains 2–11 potential RNAs (Figure 2). The five sites in Subsection 212La are clustered in the northeast corner of the Forest outside the already protected, extensive Boundary Waters Canoe Area.

The natural communities within each site were identified (Anderson 1997) according to Minnesota natural community types as described by the Minnesota Natural Heritage Program (1993). These types were based on vegetation, hydrology, landform, soil, and natural disturbance regime. To be consistent with The Nature Conservancy's National Vegetation Classification, we translated the Minnesota natural community types into appropriate, corresponding TNC Alliances using relational tables (Kristin Snow, The Nature Conservancy, Midwest Regional Office, Minneapolis, MN, personal communication). Each site contained 1–7 Alliances. In total, the 33 potential RNAs included 63 of the 132 unique Subsection-Alliance pairs known to exist in the Superior National Forest (indicated with stars in Table 1).

At the time of our study, four relatively small RNAs (847 ha total) had previously been established in the Superior National Forest (Faber-Langendoen et al. 1999b). Our analysis did not take into account the Subsection-Alliance pairs that occurred in those sites or any other RNA-equivalent sites within the Subsections but not necessarily on Forest Service land (e.g., scientific and natural areas established and administered by the Minnesota Department of Natural Resources).

The Basic Model

To address the first two goals of the proposed framework for selecting RNAs, we formulated a discrete, 0-1 integer optimization model to select the set of reserves that maximized the total number of Subsection-Alliance pairs protected given an upper bound on the total area covered by the selected set. The model was based on the maximal covering location problem (Church and Revelle 1974, Church et al. 1996) with the following notation:

j, J = index and set of Subsections,

k, K = index and set of Alliances,

m, M = index and set of potential RNAs,

T = upper bound on total area covered by the selected set of RNAs,

A_m = area of RNA m,

N_{jk} = set of RNAs that contain the Subsection-Alliance pair jk,

$$X_m = \begin{cases} 1 & \text{if RNA } m \text{ is selected for protection} \\ 0 & \text{if RNA } m \text{ is not selected,} \end{cases}$$

$$Y_{jk} = \begin{cases} 1 & \text{if Subsection - Alliance pair } jk \text{ is represented by the selected set of RNAs} \\ 0 & \text{if Subsection - Alliance pair } jk \text{ is not represented.} \end{cases}$$

The model was formulated as follows:

Maximize

$$\sum_{j \in J} \sum_{k \in K} Y_{jk} \tag{1}$$

Subject to

$$\sum_{m \in M} A_m X_m \leq T \tag{2}$$

$$Y_{jk} \leq \sum_{m \in N_{jk}} X_m \qquad \forall j \in J, \forall k \in K \tag{3}$$

$$X_m, Y_{jk} \in \{0,1\}. \tag{4}$$

The objective (1) maximized the number of unique Subsection-Alliance pairs represented or covered by the set of selected RNA sites. The first constraint (2) ensured that the total area of the selected set of RNAs did not exceed T, the upper bound on total area of sites selected for protection. The upper bound on total area was viewed as a proxy for opportunity cost because alternative land uses such as logging or developed-site recreation were assumed to be lost when sites were designated for RNA status. We assumed that these costs

were proportional to RNA area and not differentiated by the quality of the site for other land uses.

The second set of constraints (3) defined the conditions under which a Subsection-Alliance pair was considered represented or covered. This constraint stipulated that for any Subsection-Alliance pair to be represented, at least one of the eligible RNAs that contained that pair must be selected for protection. This constraint also ensured proper accounting of representation of the pairs. If

$$\sum_{m \in N_{jk}} X_m = 0,$$

then the corresponding Y_{jk} variable must also equal zero, indicating that none of the selected sites included the Subsection-Alliance pair jk. If

$$\sum_{m \in N_{jk}} X_m \geq 1,$$

then $Y_{jk} = 1$, indicating that the Subsection-Alliance pair jk was represented. The last set of constraints (4) defined the integer restrictions for the variables.

The basic optimization model found the set of reserves that protected the greatest number of Subsection-Alliance pairs for a given upper bound on total area protected. We estimated the trade-offs between the number of Subsection-Alliance pairs protected and total area of the selected set of sites by decreasing the upper bound on total area in increments of 405 ha (1,000 ac) from the maximum of 51,030 ha and re-solving the optimization model.

Representing Landtype Associations and Priority Alliances

Additional constraints were added to the basic model to address the second two goals of the proposed framework for RNA selection. One of those goals (goal 3) was to pick at least one site within each Landtype Association (LTA). Using site location information in Table 2, we defined the set of RNAs within each LTA and represented the set by the notation L_i. Using this notation, we defined a constraint that required selection of at least one RNA within each LTA:

$$\sum_{m \in L_i} X_m \geq 1 \qquad \forall i \in I \tag{5}$$

where i and I were the index and set of LTAs, respectively. This set of constraints was added to the basic model, which was solved using the full range of upper bounds on total area of the selected set of sites. The results were compared to those from the basic model to estimate the impacts of requiring complete LTA representation.

The last goal for RNA selection (goal 4) was representation of priority Subsection-Alliance pairs (cells with check marks in Table 1). To ensure representation of priority Subsection-Alliance pairs, we added a constraint to the basic formulation for each priority pair jk :

$$Y_{jk} = 1$$

The model was solved using the full range of upper bounds on total area of the selected set of sites, and the results were compared to those obtained from the basic model to estimate the impacts of requiring representation of priority Subsection-Alliance pairs.

Software

All of the models were solved on an IBM300PL personal computer using the integrated solution package GAMS/OSL 2.25 (GAMS Development Corporation 1990), which was designed for large and complex linear and mixed integer programming problems. Input files were created using GAMS (General Algebraic Modeling System), a program designed to generate data files in a standard format that optimization programs can read and process. The models were solved using the primal simplex algorithm in conjunction with the branch and bound algorithm for integer-variable problems. Both these algorithms were obtained from IBM's OSL (Optimization Subroutine Library), a Fortran-based subroutine library designed to solve optimization problems.

Results

Trade-Offs Between Ecosystem Representation and Protected Area

We used the basic model to produce a trade-off curve showing maximum numbers of Subsection-Alliance pairs represented for decreasing upper bounds on the total area of the selected set of sites (Figure 3). Each point on the trade-off curve represents a unique set of selected sites. Although we generated solutions using upper bounds on total RNA area in increments of 405 ha (1,000 ac), we showed only those points in which there was a change in the set of selected sites. A different solution was not generated each time the upper bound was decreased. The flat portion of the trade-off curve between points A and B shows that all 63 Subsection-Alliance pairs can be represented by sets of sites covering a wide range of areas (51,030–35,235 ha). If all of the 33 potential RNAs were selected, the protected set of sites would cover 51,030 ha (point A, Figure 3). Because many Subsection-Alliance pairs were found in more than one potential RNA, the model was able to select a smaller set of sites without reducing representation. For example, the entire set of 63 Subsection-Alliance pairs was represented with 21 RNAs covering 35,235 ha (point B, Figure 3), a 31% reduction in total area.

When the upper bound on total area was less than 35,235 ha, not all of the Subsection-Alliance pairs could be represented. Nevertheless, the upper bound on total area could be reduced by a substantial amount without great reduction in Subsection-Alliance representation. For example, the model showed that 57 Subsection-Alliance pairs, more than 90% of the total of 63 pairs, could be represented within a set of 18 RNAs covering 16,605 ha (point C, Figure 3), a 67% reduction in total area from the maximum of 51,030 ha.

Lowering the upper bound on total area below 16,605 ha accelerated the reduction in numbers of Subsection-Alliance pairs represented. For example, with an upper bound of 8,100 ha, 45 pairs could be represented, a 28% reduction from the

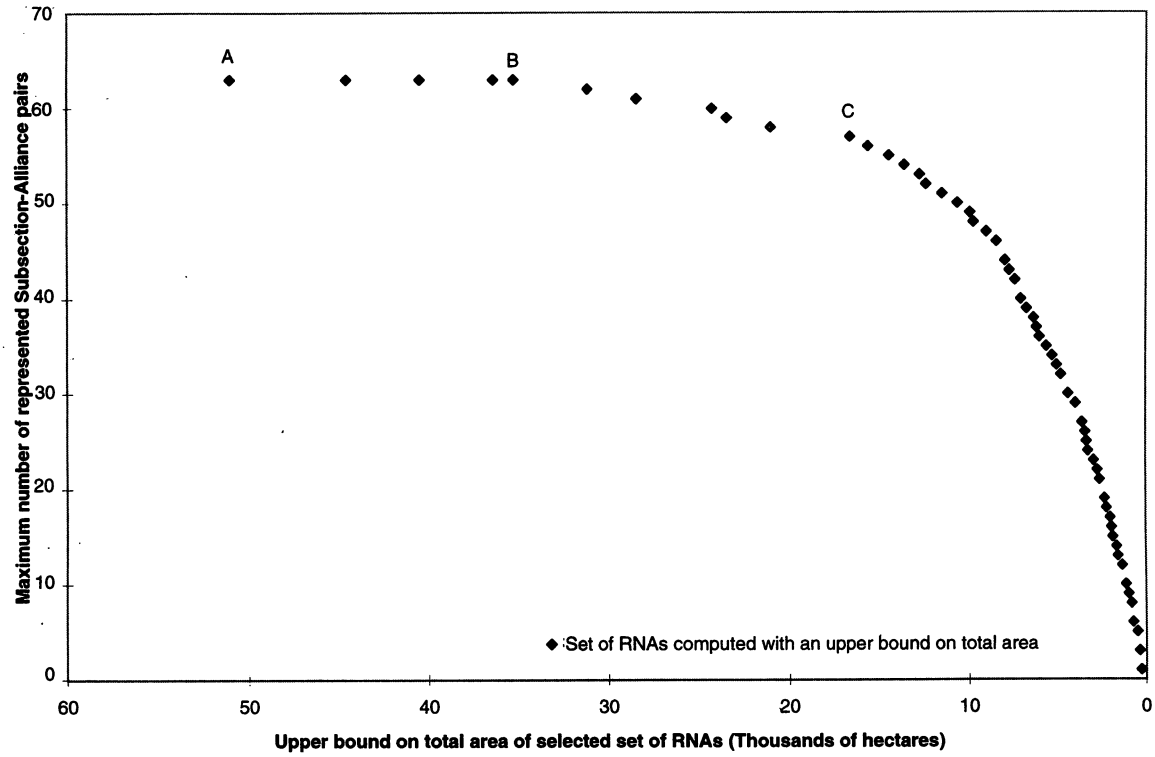

Figure 3. Trade-off curve showing maximum number of Subsection-Alliance pairs represented as a function of a decreasing upper bound on the total area of the set of sites selected for protection. Points A, B, and C are described in the text.

maximum of 63 pairs. With an upper bound of 4,050 ha, 25 pairs could be presented, a 60% reduction in representation.

Representing Landtype Associations

The trade-offs between Subsection-Alliance representation and total area of the selected set of sites were computed with an additional requirement that each Landtype Association (LTA) be represented within at least one RNA (Figure 4). When the upper bound on total area was relatively small and restrictive (< 24,300 ha), requiring LTA representation significantly reduced the number of Subsection-Alliance pairs

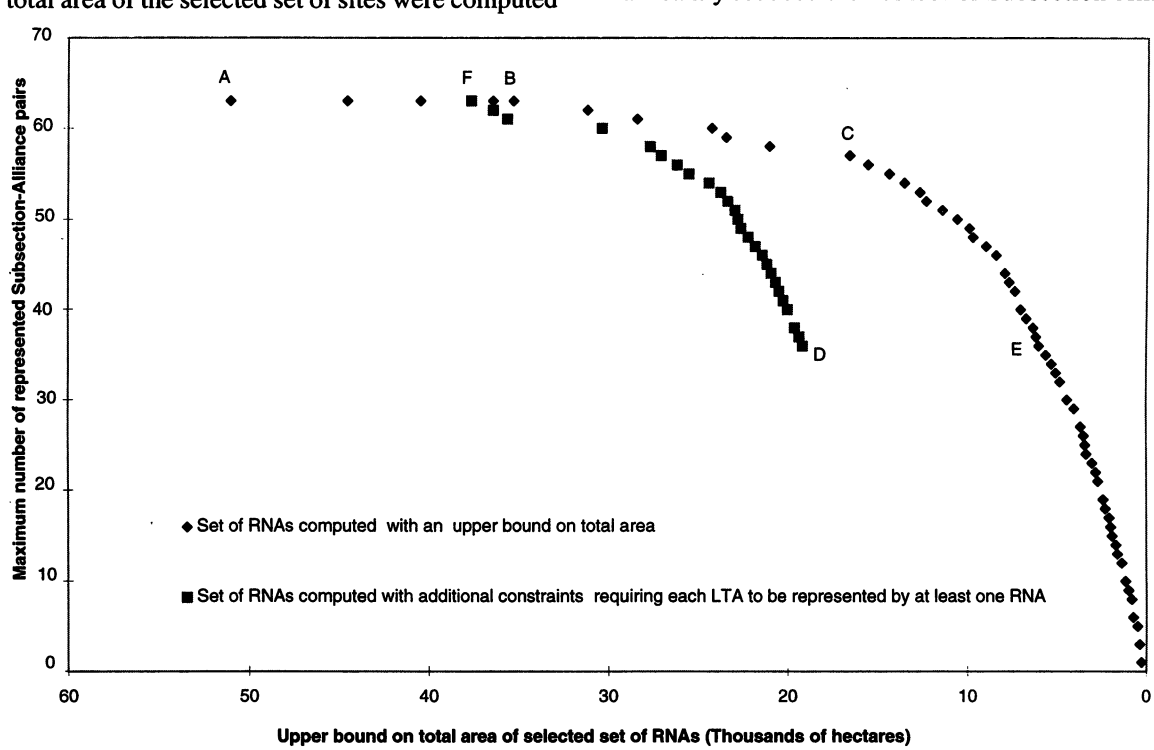

Figure 4. Trade-off curves showing maximum number of Subsection-Alliance pairs represented as a function of a decreasing upper bound on the total area of the set of sites selected for protection. The curves were computed with and without requirements for Landtype Association (LTA) representation. Points A, B, C, D, E, and F are described in the text.

that could be represented in a set of sites of a given area. For example, the minimum area required to represent each LTA with at least one RNA was 19,035 ha (point D, Figure 4). This set of sites represented 36 Subsection-Alliance pairs, which was a 37% reduction from the 57 Subsection-Alliance pairs that were included in the set of sites that had roughly the same area but was not required to include all the LTAs (point C, Figure 4). Looking at the effects another way, requiring LTA representation significantly increased the area required to represent a given number of Subsection-Alliance pairs. For example, in comparison to point D, a different set of 36 Subsection-Alliance pairs could be represented with 6,075 ha without full LTA representation (point E, Figure 4). This is a 68% reduction in total area compared to the minimum area (19,035 ha) required to represent all of the LTAs.

Adding the LTA constraint had less impact on the trade-offs between Subsection-Alliance representation and total area when the upper bound on total area was relatively large and less restrictive (i.e., > 24,300 ha). For example, the minimum area required to represent all 63 Subsection-Alliance pairs without full LTA representation was 35,235 ha (point B, Figure 4). This set of sites represented all but 4 LTAs. When the LTA restriction was added, the minimum area required to represent all of the LTAs and all 63 Subsection-Alliance pairs was 37,665 ha (point F, Figure 4), an increase of 2,430 ha or roughly 7% from point B. It is useful to note that this solution, represented by point F, still had 26% less area than the set composed of all of the potential RNAs (point A, Figure 4).

Representing Priority Alliances

Requiring representation of priority matrix Alliances significantly affected the trade-offs between Subsection-Alliance representation and total area of the selected set of sites (Figure 5). When the upper bound on total area was relatively small and restrictive (< 24,300 ha), requiring representation of priority matrix alliances reduced the total number of Subsection-Alliance pairs that could be represented in a set of sites of a given area. For example, with an upper bound of 17,010 ha, the set of sites that maximized representation without the requirement to include priority matrix Alliances represented a total of 57 Subsection-Alliance pairs (point C, Figure 5). The set of sites that maximized representation with the requirement to include priority matrix Alliances represented a total of 36 Subsection-Alliance pairs (point G, Figure 5), a 37% reduction in representation.

Comparing the matrix Alliances represented by the different sets of sites associated with points C and G further illustrates the effects of the constraints (Table 3). Although point C included sites that represented 19 matrix Subsection-Alliance pairs (cells with circles in upper portion of Table 3), this set of sites did not represent two priority Subsection-Alliance pairs (cells with check marks and no circles in the upper portion of Table 3). Point G included sites that covered slightly more area than the set of sites associated with point C (17,010 ha compared to 16,605 ha) but included different sites. Although those sites represented all 12 priority Subsection-Alliance pairs (cells with checkmarks and triangles in the lower portion of Table 3), they represented no other matrix Subsection-Alliance pairs. In fact, point G included no sites in Subsections 212La or 212Ld. These results demonstrate how a requirement to represent priority matrix Alliances may reduce overall Subsection-Alliance representation when the upper bound on total area is limiting.

Requiring representation of priority matrix Alliances had less impact on the trade-offs between Subsection-Alliance representation and total area when the upper bound on total

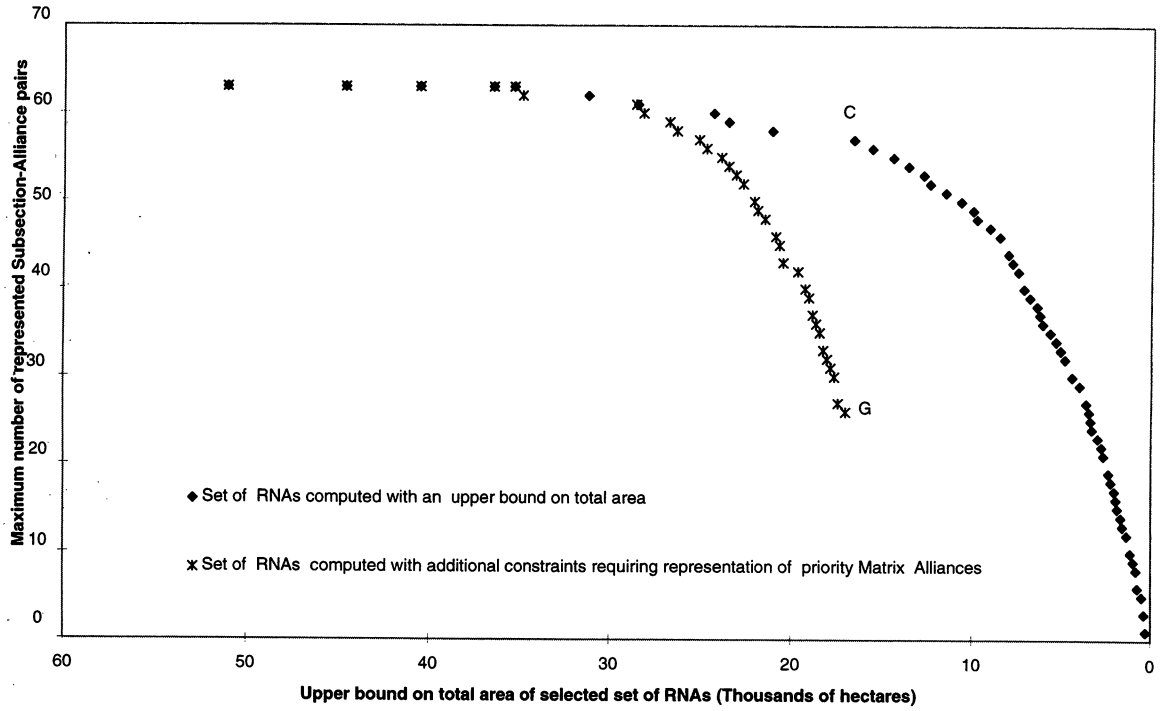

Figure 5. Trade-off curves showing maximum number of Subsection-Alliance pairs represented as a function of a decreasing upper bound on the total area of the set of sites selected for protection. The curves were computed with and without requirements for representing priority matrix Alliances. Points C and G are described in the text.

Table 3. Comparison of Matrix Alliances represented by two sets of RNAs of approximately 16,600 ha. Sets C and G correspond to points C and G in Figure 5.

	Subsections				
	212La	212Lb	212Lc	212Ld	212Le
Set C (No requirement to cover priority matrix Alliances)					
Matrix Alliances					
Jack Pine Forest Alliance					●✔
Red Pine Forest Alliance	●		●✔		●✔
White Pine-(Red Pine)-Quaking Aspen Forest Alliance		●✔	●✔		
White Pine Forest Alliance	●				●✔
White Spruce-Balsam Fir-Quaking Aspen Forest Alliance		●	●✔	●	✔
White Spruce-Balsam Fir Forest Alliance					
Sugar Maple-Yellow Birch-(American Beech) Forest Alliance	●	●	●✔	●	✔
Quaking Aspen-Paper Birch Forest Alliance	●		●✔	●	●✔
Set G (Requirement to cover priority matrix Alliances)					
Matrix Alliances					
Jack Pine Forest Alliance					▲✔
Red Pine Forest Alliance			▲✔		▲✔
White Pine-(Red Pine)-Quaking Aspen Forest Alliance		▲✔	▲✔		
White Pine Forest Alliance					▲✔
White Spruce-Balsam Fir-Quaking Aspen Forest Alliance			▲✔		▲✔
White Spruce-Balsam Fir Forest Alliance					
Sugar Maple-Yellow Birch-(American Beech) Forest Alliance			▲✔		▲✔
Quaking Aspen-Paper Birch Forest Alliance			▲✔		▲✔

LEGEND

●	Subsection-Alliance pair represented by Set C
▲	Subsection-Alliance pair represented by Set G
✔	Priority Subsection-Alliance pair

area was relatively large and less restrictive (i.e., > 24,300 ha) (Figure 5). Both models yielded the same solution for the minimum amount of area (35,235 ha) needed to represent all 63 Subsection-Alliance pairs. With an upper bound of 24,300 ha, the set of sites that maximized representation with the requirement to represent priority matrix Alliances included a total of 55 Subsection-Alliance pairs, only 4 less than the set of sites computed without the priority constraints.

Discussion

We addressed the problem of selecting Research Natural Areas in a national forest to meet various goals for ecosystem and natural community representation. We used integer optimization to analyze this reserve selection problem because it provided a logical framework to incorporate site-specific information about the presence of natural communities, it gave exact solutions to optimization problems with little computational effort, and it conveyed easy-to-analyze relationships for the trade-offs between RNA-selection goals.

A strength of our model was its tractability: solution times averaged 0.8–2.3 seconds on an IBM300PL personal computer. Model tractability is important because it enables decision makers to readily determine the effects of different management goals on the optimal number and location of selected sites. Although our models solved quickly and easily, some applications might be more difficult. The classic maximal covering location problem, upon which our formulation was based, is known as *n–p* (nondeterministic polynomial) hard, which means that solution time may not be a polynomial function of the size of the problem. As a result, some relatively small problems could take an exhaustive amount of time to solve. It is difficult to specify an absolute upper bound in problem size that this formulation can be efficiently applied to. Solution time and complexity will be a function of the structure of the data as well as computing power and the choice of integer optimization algorithms. If problem complexity or computation time become restrictive, our formulation could be used to gauge the effectiveness of alternative heuristic solution algorithms in tests on smaller problems.

Another strength of our model was its ability to quickly generate trade-offs between different protection goals. These trade-offs are important to identify so that decision-makers can make informed choices. For example, using 33 potential RNAs in the Superior National Forest, we found multiple sets of sites, ranging from all 33 sites to a much smaller set of 21 sites, that attained the specified goal for Subsection and Alliance representation. Thus, the decision maker can choose from multiple sets of sites with a wide range of total areas without compromising the Subsection-Alliance protection goal. We also found that requirements to choose a set of sites that represents a range of Landtype Associations or priority matrix Alliances can significantly reduce the number of Subsection-Alliance pairs that can be represented within a set of sites of a given area. Identifying these impacts should help the decision maker evaluate these more refined protection goals.

We emphasize that the 33 potential RNAs used in our study were chosen because they had been mapped and field surveyed for the presence of natural communities. The use of those sites does not imply that other areas in the Superior National Forest do not merit further study as RNA candidates or that the 33 sites provide examples of all the ecosystems in need of representation. The difficulty here is not with the model but having sufficient information about potential Research Natural Areas.

The process of RNA selection can be influenced by a number of factors that we did not consider in our model. For example, the RNA selection process could take into account the quality and successional status of candidate natural communities, the needs of individual species, disturbances such as fires or storms that alter the structure and composition of natural communities, climatic changes that might affect the demography of protected plants and animals, and uncertainties in the data on presence and absence of natural communities. Therefore, we want to emphasize that results from models such as ours should be used in conjunction with information obtained from other sources.

There is certainly room for improving and expanding the simple model presented here. For example, we could refine the definitions of community representation by taking into account the quality of the candidate natural communities. We assumed that a Subsection-Alliance pair was represented when it was present within at least one selected RNA. Because we did not account for the areas of communities within each site, a Subsection-Alliance pair was considered represented even if the community occurred in a very small parcel within one selected site. To avoid this problem, protection goals could specify minimum areas that are required for adequate community representation. Another approach would be to grade the quality of each community parcel on the basis of features such as area and condition including species composition and landscape context (e.g., Faber-Langendoen et al. 1999b). Then, representation goals could be based on quality standards in addition to community types. Yet another option for modeling would be to reduce the area of some sites to include only the priority Subsection-Alliance pairs and not those that are represented in other sites.

Another refinement involves recognizing and incorporating uncertainty about natural community occurrence. We assumed that the presence and absence of natural communities within each site was known with certainty. However, when resources and personnel are limited, exhaustive field surveys of potential natural areas might not be possible. In this situation, the community composition of each site could be expressed using probabilities of occurrence based on expert opinion. Management goals for community representation could then be expressed in probabilistic terms (e.g., a community is represented only when the probability of its occurrence within the selected sites is greater than some risk threshold). We are developing such probabilistic formulations that can be solved using integer optimization.

Finally, it is essential to integrate RNA selection into the broader scope of national forest planning, in which simultaneous decisions must be made about a wide range of land uses and management practices. This integration requires detailed information about each potential RNA so that trade-offs between natural community representation and other resource uses could be illustrated. Further, efforts are being made to evaluate potential RNAs in light of the natural communities already represented by protected areas within and outside the National Forest System (e.g., Faber-Langendoen et al. 1999b) and at regional scales defined by ecological rather than political boundaries. The modeling framework presented here could certainly be extended to handle these situations. Obviously, RNA selection will not be done in isolation, but within the context of resource limitations and societal demands for a host of desired goods and services from national forests.

The ability to quickly assess the effects of different resource protection goals based on the information at hand is important to decision makers who require an explicit, well-documented justification for their land-use decisions. Integer optimization models can be customized to handle many kinds of goals for Research Natural Area selection. Model results, in turn, can provide insights about the trade-offs between competing goals.

Literature Cited

ANDERSON, C.E. 1997. Evaluation of selected potential Candidate Research Natural Areas as representative of ecological Landtype Associations on the Superior National Forest. Minn. Dep. of Natur. Resour., Div. of Fish and Wildl., Bio. Rep. No. 58. 19 p. [plus maps and appendices].

BEDWARD, M., R.L. PRESSEY, AND D.A. KEITH. 1992. A new approach for selecting fully representative reserve networks: addressing efficiency, reserve design, and land suitability with an iterative analysis. Bio. Conserv. 62:115–125.

CAMM, J.D., S. POLASKY, A. SOLOW, AND B. CSUTI. 1996. A note on optimal algorithms for reserve site selection. Bio. Conserv. 78:353–355.

CHURCH, R.L., AND C. REVELLE. 1974. The maximal covering location problem. Reg. Sci. 32:101–118.

CHURCH, R.L., D.M. STOMS, AND F.W. DAVIS. 1996. Reserve selection as a maximal covering location problem. Bio. Conserv. 76:105–112.

COCKS, K.D., AND I.A. BAIRD. 1989. Using mathematical programming to address the multiple reserve selection problem: an example from the Eyre Peninsula, South Australia. Bio. Conserv. 49:113-130.

CSUTI, B., ET AL. 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. Bio. Conserv. 80:83–97.

DAVIS, F.W., AND D. STOMS. 1996. A spatial analytical hierarchy for Gap Analysis. P. 15–24 *in* Scott, J.M., et al. (eds). Gap Analysis: A Landscape Approach to Biodiversity Planning. Am. Soc. for Photogram. and Rem. Sens., Bethesda, MD.

FABER-LANGENDOEN, D., K. SNOW, AND L. TYRRELL. 1999a. Establishing a region-wide network of representative Research Natural Areas: An assessment for the U.S. Forest Service Eastern Region (Region 9). Unpubl. Rep. on file with the USDA For. Serv., North Central Res. Sta., Rhinelander, WI.

FABER-LANGENDOEN, D., K. SNOW, AND L. TYRRELL. 1999b. Research natural area assessment for the Superior National Forest. Unpubl. Rep. on file with the USDA For. Serv., North Central Res. Sta., Rhinelander, WI.

GAMS Development Corporation. 1990. General algebraic modeling system. Version 2.25.090. Washington, DC.

GROSSMAN, D.H., K.L. GOODIN, AND C.L. REUSS (EDS.). 1994. Rare plant communities of the coterminous United States: An initial survey. The Nature Conservancy, Arlington, VA. 620 p.

KEYS, J. JR., ET AL. 1995. Ecological units of the eastern United States - first approximation (map and booklet of map unit tables). Atlanta, GA. USDA For. Serv., presentation scale 1:3,500,000; colored.

KIRKPATRICK, J.B. 1983. An iterative method for establishing priorities for the selection of nature reserves: an example from Tasmania. Bio. Conserv. 25:127–134.

MARGULES, C.R., A.O. NICHOLLS, AND R.L. PRESSEY. 1988. Selecting networks of reserves to maximize biological diversity. Bio. Conserv. 43:63–76.

MARGULES, C.R., AND M.B. USHER. 1981. Criteria used in assessing wildlife conservation potential: A review. Bio. Conserv. 21:79–109.

MCNAB, W.H., AND P.E. AVERS (COMPS.). 1994. Ecological subregions of the United States: Section descriptions. Admin. Publ. WO-WSA-5. USDA For. Serv.,Washington, DC. 267 p.

MINNESOTA NATURAL HERITAGE PROGRAM. 1993. Minnesota's native vegetation: A key to natural communities, version 1.5. Minn. Dep. of Natur. Resour., Div. of Fish and Wildlife, Bio. Rep. No. 2. 111 p.

NICHOLLS, A.O., AND C.R. MARGULES. 1993. An updated reserve selection algorithm Bio. Conserv. 64:165-169.

PRESSEY, R.L., C.J. HUMPHRIES, C.R. MARGULES, R.I. VANE-WRIGHT, AND P.H. WILLIAMS. 1993. Beyond opportunism: Key principles for systematic reserve selection. Trends Ecol. Evol. 8:124–128.

PRESSEY, R.L., AND A.O. NICHOLLS. 1989. Efficiency in conservation planning—scoring versus iterative approaches. Bio. Conserv. 50:199–218.

PRESSEY, R.L., H.P. POSSINGHAM, AND J.R. DAY. 1997. Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. Bio. Conserv. 80:207–219.

PRESSEY, R.L., H.P. POSSINGHAM, AND C.R. MARGULES. 1996. Optimality in reserve selection algorithms: When does it matter and how much? Bio. Conserv. 76:259–267.

REBELO, A.G., AND W.R. SEIGFRIED. 1990. Protection of fynbos vegetation: ideal and real world options. Bio. Conserv. 54:15–31.

SAETERSDAL, M., J.M. LINE, AND H.J.B. BIRKS. 1993. How to maximize biological diversity in nature reserve selection: Vascular plants and breeding birds in deciduous woodlands, western Norway. Bio. Conserv. 66:131–138.

STOMS, D.M., M.I. BORCHERT, M.A. MORITZ, F.W. DAVIS, AND R.L. CHURCH. 1998. A systematic process for selecting representative Research Natural Areas. Natur. Areas J. 18:338–349.

TERBORGH, J., AND B. WINTER. 1983. A method for siting parks and reserves with special reference to Columbia and Ecuador. Bio. Conserv. 27:45–58.

TOREGAS, C., AND C. REVELLE. 1973. Binary logic solutions to a class of location problems. Geographical Analysis 5:145–155.

UNDERHILL, L.G. 1994. Optimal and suboptimal reserve selection algorithms. Bio. Conserv. 70:85–87.

USDA FOREST SERVICE.1986. Superior National Forest Land and Resource Management Plan. Eastern Region, USDA For. Serv.

USDA FOREST SERVICE. 1992. Preparing for the Future: Forest Service Research Natural Areas. FS-503.

VANE-WRIGHT, R.I., C.J. HUMPHRIES, AND P.H. WILLIAMS. 1991. What to protect?—Systematics and the agony of choice. Bio. Conserv. 55:235–254.

VORA, R.S. 1997. Identification of potential natural areas, including representative ecosystems, on the Superior National Forest. Unpubl. Rep. on file with the USDA Forest Service, North Central Research Station, Rhinelander, WI.

WILLIAMS, J., AND C. REVELLE. 1996. A 0-1 programming approach to delineating protected reserves. Environ. Plan. B 23:607–624.

WILLIAMS, J., AND C. REVELLE. 1997. Applying mathematical programming to reserve selection. Environ. Model. Assess. 2:167–175.

WILLIAMS, J., AND C. REVELLE. 1998. Reserve assemblage of critical areas: a zero-one programming approach. Eur. J. Op. Res. 104:497–509.

WILLIS, C., A. LOMBARD, R. COWLING, B. HEYDENRYCH, AND C. BURGERS. 1996. Reserve systems for limestone endemic flora of the Cape Lowland fynbos: Iterative versus linear programming techniques. Bio. Conserv. 77:53–62.