A Cost Effective and Efficient Way to Assess Trail Conditions: A New Sampling Approach

Rachel A. Knapp, Graduate Assistant, University of New Hampshire Department of Natural Resources and the Environment, Nesmith Hall, 131 Main Street, Durham, NH 03824; ral39@unh.edu

Mark J. Ducey, Professor of Forest Biometrics and Management, Department of Natural Resources and the Environment, Nesmith Hall, 131 Main Street, Durham, NH 03824; mjducey@cisunix.unh.edu

Introduction

Trails are important recreational facilities, and subsequently trail maintenance is a high priority for many park managers. Trails in poor repair, or trails that experience excessive use, can lead to poor visitor experiences and affect park resources through erosion, widening, and vegetation changes. Monitoring trail conditions is an important step for ensuring trail quality while controlling costs for expensive repairs and relocation (Cole 1983). However, budgets are often tight, so trail assessment and monitoring techniques must be time efficient and cost effective. Many trail networks are too long to evaluate in their entirety; here we explore a method for assessing trail conditions without examining the whole network.

Cole (1983) classifies trail assessment techniques into three major categories: replicable measurements (quantitative measures, usually implemented at a small number of points on a trail), rapid surveys (usually conducted at a larger number of points), and complete censuses of trail problems or conditions. With both replicable measurements and rapid surveys, sample points are usually distributed systematically. Similarly, Marion, Leung, and Nepal (2006) divide assessments into "sampling-based" techniques (encompassing Cole's [1983] replicable measurements and rapid surveys) and "census-based" techniques.

Currently, most trail assessment literature continues to include three particular trail impact indicators: changes in trail width (either the tread or the entire disturbed zone); changes in trail depth or incision; and general trail conditions which are often referred to as condition classes (Cole 1983; Marion, Leung, and Nepal 2006; Nepal 2003). Marion, Leung, and Nepal (2006) note that since the entire trail network must be walked to implement "sampling-based" techniques, combining those with "census-based" techniques requires little additional effort; thus, quantitative measurements, like trail width and depth, may be combined with qualitative measurements, such as descriptive classes, in order to get a more complete picture of the overall trail condition. The common assumption is that the entire trail network must be assessed and characterized to estimate overall trail conditions.

When trail networks are extensive, or time or financial resources are limiting, it may not be practical to measure the entire network. The challenge then becomes one of sampling the network. As Cole (1983) points out, subjective location of measurement points or segments makes it impossible to scale up estimates to the entire trail system, while ordinary techniques force crews to take a large number of measurements on trail segments that are in relatively good condition, and convey little information of interest to managers. Trail networks could be sampled by dividing the network into segments, and then selecting those segments at random. However, this technique may compound the problem by requiring excessive non-measurement travel time, often through problem or other highly informative areas that are not to be measured, simply to reach short segments that are designated for measurement.

In this paper, we briefly introduce randomized graph sampling (Ducey 2009), or RGS, a new but simple technique that allows efficient trail assessment. The technique takes advantage of the connectedness of trail networks, allows the use of auxiliary information to focus effort where problems are more likely, and also permits unbiased estimates of trail characteristics across entire networks.

Randomized graph sampling

RGS is a variable-probability sampling technique (Ducey 2009). The mathematical definition of a graph is a set of junctions connected by lines or pathways and as such a trail network matches the mathematician's definition quite well. For our purposes, we consider any trail network as being defined by the unique pathways that connect pairs of trail intersections, and we call any single pathway from one intersection to the next intersection a trail segment. Rather than sampling segments individually, we connect segments into reasonable sampling walks, and then sample from a list of potential walks.

RGS can be described as a six-step process:

1. Identify trail segments. The first step is to identify the sample population—essentially, whatever network or sub-network of trails is of concern—and to develop a list of the trail segments that comprise that network. An ordinary map or a GIS can be used to develop the list. Trail segments will not typically have equal length. In general, it is advantageous to include the segment lengths when developing the list.

2. Identify possible walks. A walk is defined as a reasonable and feasible sampling path connecting one or more segments. What is "reasonable" or "feasible" is defined operationally. For example, walks do not have to start and end in the same place, but that is often simplest for the data collector; and walks will typically begin and end at trailheads or other reasonable parking locations. The walk length should be determined by the time available for the inquiry. For example, if you have two people and four days to cover 16 walks then you would want the walks to be 4 to 5 hours each. This would allow each person to cover 2 walks per day, or 8 walks each, over the 4 days, for the total necessary samples. It is not necessary to develop a list of all possible walks. In general, walks will overlap and some segments will appear in multiple walks. The only requirement is that each segment in the network appears in *at least* one walk.

3. Determine sampling probabilities. Determining sampling probabilities involves two substeps: first, determining the selection probabilities for each walk, and second, calculating the inclusion probabilities for each segment. The selection probabilities for walks are used operationally for choosing which walks to sample; the inclusion probabilities for segments are needed for calculating the final estimates after measurements have been taken.

The simplest method of selecting walks would be to assign each walk equal probability. For example, if there are 4 possible walks, each walk would have a 1 in 4, or 25%, chance of being selected (Table 1). We call this approach the "equal probability design."

Walk (Selection Probability, q_j)	A (.25)	B (.25)	C (.25)	D (.25)	
Segment					Totals (p_i)
1	.25		.25		.50
2	.25	.25			.50
3	.25	.25		.25	.75
4			.25	.25	.50
5				.25	.25

Table 1. Equal probability design.

Walk (Selection Probability, q_j)	A (.18)	B (.23)	C (.27)	D (.25)	
Segment					Totals (p_i)
1	.18		.27		.45
2	.18	.23			.41
3	.18	.23		.25	.80
4			.27	.25	.59
5				.25	.32

Table 2. Probability-proportional-to-length design.

Walk (Selection Probability, q_j)	A (4/16)	B (2/16)	C (5/16)	D (5/16)	
Segment					Totals (p_i)
1	.25		.31		.56
2	.25	.12			.37
3	.25	.12		.31	.68
4			.31	.31	.62
5				.31	.31

Table 3. Assigned probability design.

However, equal probabilities are neither necessary nor (usually) efficient. A simple alternative is to assign probabilities proportional to walk length. With the probability-proportional-to-length design, each walk's probability of selection is based on its length, as a fraction of the total lengths of all walks in the list. Suppose our 4 walks have lengths of 18 km, 23 km, 27 km, and 32 km, for a total of 100 km. Then the respective walk selection probabilities would be .18, .23, .27, and .32 (Table 2).

A third alternative is to incorporate auxiliary information and assign the probabilities directly. Auxiliary information could come from a sophisticated model or, more likely, involve a simple scoring system based on local expertise. The goal of assigning probabilities is to focus sampling effort where problems are likely to occur, and as a result to decrease the final sample variance. The probability of selecting a particular walk becomes that walk's score, divided by the sum of the scores for the entire list of walks (Table 3). Each segment must have a probability greater than zero and positive. Rather, no segment or walk can have a zero probability of being sampled and all probabilities must be positive in order to calculate selection probabilities as described above.

No matter how the individual walk selection probabilities are assigned, calculation of the inclusion probabilities for segments is identical and straightforward. Let q_j be the selection probability of the j^{th} walk. Let d_{ij} be an "indicator variable" that tells whether the i^{th} segment occurs in the j^{th} walk; $d_{ij} = 1$ if segment i is in walk j, and $d_{ij} = 0$ if it is not. Then the inclusion probability p_i for the i^{th} segment is

$$p_i = \sum_j d_{ij} q_j$$

Examples of p_i for the equal probability, probability-proportional-to-length, and assigned probability designs are shown in the totals columns of Tables 1–3.

4. Select walks randomly. Regardless of which design is used, one or more walks must be selected randomly, either with or without equal probability. The number of walks to select can be based on a sample size equation (cf. Thompson 2002, 53–56; description of sample size equations is beyond the scope of this paper), but it may have to be based on the available time or financial budget for trail assessment. Typically, the more walks sampled, the lower the standard error of the estimate(s) of the feature(s) of interest.

5. Conduct measurements. As discussed above, there are several common trail impact indicators including trail width, incision, and qualitatively-defined trail condition classes (Cole 1983; Marion, Leung, and Nepal 2006). Whatever indicators are appropriate to local conditions and stated study objectives can be measured on all segments in the selected walks. The appropriate indicators should be decided in advance, as they may affect sampling time per km of trail, and thus impact both what constitutes a "feasible" walk and also the number of walks that can be sampled within the project budget.

6. Calculate estimates. Using the data from the sampled walks, it is straightforward to compute unbiased estimates of network totals (or means) of a variety of attributes for each sample walk. If Y is the network total of some attribute (such as the total number of meters in a particular condition class), then the estimate of Y calculated from the j_{th} walk is

$$\hat{Y}_j = \sum_{i \in j} \frac{y_i}{p_i}$$

where y_i is the amount of the attribute in the i^{th} segment; only those segments appearing in the j^{th} walk are used to form the estimates. If multiple walks have been selected, then the mean of the $\hat{Y}_j$ provides the best estimate of the total, and the standard error and confidence limits can be calculated from the $\hat{Y}_j$ using the usual formulae (Ducey 2009).

For networks of moderate size and complexity, steps 1–6 can all be completed using a simple paper map and ordinary spreadsheet software.

Kingman Farm sample study

To demonstrate RGS, and to compare sampling designs, we conducted a complete survey of nearly 5 km of forest trails at Kingman Farm in Durham, N.H. (Figure 1). We did not include trails on town or private land, trails in fields, or farm roads.

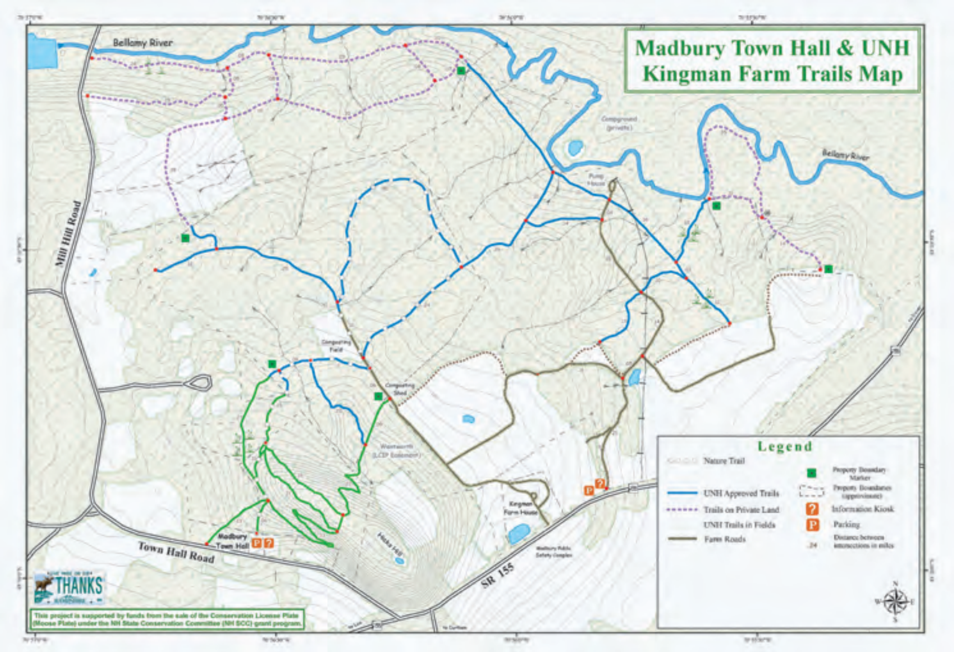

Figure 1. Kingman Farm trails.

Data collection

We took a series of qualitative and quantitative measurements at systematically-located points every 20 m along the entire trail network; this interval is considered adequate for this length of trail network (Leung and Marion 1999; Hawes et al. 2006). Distances were determined using a measuring wheel. At each sample point we measured trail width and trail incision (depth) as well as assigned an appropriate condition class rating. Trail width was defined as width of visibly trampled ground, and was measured using a tape to within 10 cm increments. Trail depth was defined as greatest depth of visible depression between noticeable banks on both sides of the trail. This was measured to the nearest centimeter. Condition classes were adapted from Nepal (2003) and Marion, Leung, and Nepal (2006) to local conditions: class 0, no apparent damage; class 1, lightly damaged; class 2, moderately damaged; class 3, highly damaged; class 4, severely damaged, or "hotspot."

Results

In total, we took measurements at 229 sites along 19 different trail segments. The total network consisted of 4,817 m of trails. For the purpose of this example, we focused on trail incision and condition class as primary impact indicators. According to the 229 sample points, approximately 840 m, or 17%, of the trail had incision deeper than 4 cm. Also, approximately 220 m, or 5%, of the trail was assigned a condition class greater than 1. Overall the Kingman Farm trails were in good condition with 3 of the 19 segments showing most of the damage.

We developed a list of 12 possible "walks" that included a combination of trail segments that would take approximately 3 to 5 hours to sample. Each walk was assigned a selection probability based on the design method. For example, for the equal probability method, each walk's selection probability was 1/12. For the assigned probability design, we adopted a very simple scoring procedure. Each segment was assigned one point as a base value, an additional one point if it crossed a mapped stream, and an additional two points if it crossed mapped wetlands. For example, if a single loop passed through two wetlands and had three river crossings, its total point assignment would be eight. In the assigned probability design, each walk's probability was its point score (obtained by adding the points on each of its segments) divided by the sum of the point scores of all walks.

Since, RGS is an unbiased method, the average estimate for each measured attribute is identical (and equals that which would be obtained by exhaustive measurement of the network) no matter which design is employed. However, the design does affect the variability of estimates obtained from single walks. Even the relatively crude assigned probability method employed in this study substantially reduced the sampling variance. The standard error of the average incision estimate of a single walk was 642 for probability proportional to length, 608 for the equal probability design, and a low 482 for the assigned probabilities design. This design also provided a lower standard error of the condition class risk estimate where probability proportional to length was highest at 284, equal probability at 271, and assigned probability at 243.

Conclusions

Trail assessment can be very time-consuming, and for most organizations, time is money. Having a design that reduces the total length of trails to sample in complex networks could improve the effectiveness and cost-efficiency of trail assessment and monitoring. Although RGS is a new technique, this preliminary study suggests that it may have several applications in the area of recreational ecology. RGS is an unbiased sampling method, but it takes advantage of the connectedness of trail networks while respecting operational constraints on sampling routes. The ability to reduce standard errors by using auxiliary information—even crude qualitative scoring techniques—should further improve the efficiency of RGS.

References

Cole, D.N. 1983. *Assessing and Monitoring Backcountry Trail Conditions.* Research Paper INT-303. Intermountain Forest and Range Experiment Station. Ogden, Utah: USFS.

Ducey, M.J. 2009. Randomized graph sampling. Working Paper 2009-01. Forest Biometrics Lab, University of New Hampshire.

Hawes, M., S. Candy, and G. Dixon. 2006. A method for surveying the condition of extensive walking track systems. *Landscape and Urban Planning* 78, 275–287.

Leung, Y.-F, and J.L. Marion. 1999. The influence of sampling interval on the accuracy of trail impact assessment. *Landscape and Urban Planning* 43, 167–179.

Marion, J.L., Y.-F. Leung, and S.K. Nepal. 2006. Monitoring trail conditions: New methodological considerations. *The George Wright Forum* 23:2, 36–49.

Nepal, S.K. 2003. Trail impacts in Sagarmatha (Mt. Everest) National Park, Nepal: A logistic regression analysis. *Environmental Management* 32, 312–321.

Thompson, S.K. 2002. *Sampling*. 2nd ed. New York: Wiley-Interscience.