A mixed-modes approach for estimating hiking on trails through diverse forest landscapes: the case of the Appalachian Trail

Stanley J. Zarnoch, J.M. Bowker, and H. Ken Cordell

Abstract: Many hiking trails traverse the forests and public lands across North America. It has therefore become important for federal management to gain an understanding of total use on these trails. However, there has never been a formal attempt to estimate hiking on these long, backcountry trails. This paper presents an approach that utilizes two survey instruments (exit-site tallies and a trail-user questionnaire) and two primary estimation components (standard and augmented sites) to estimate hikers over a spatial and temporal span. For illustrative purposes, the methodology is applied to a 175 km segment of the Appalachian Trail from 1 June through 14 August 2007. Two alternative estimation methodologies are presented and compared. The model-based approach may be preferred to the design-based approach when sample size is small because it smoothes erratic strata estimates and yields smaller standard errors. However, the design-based approach relaxes an assumption and is more appropriate as sample size increases. In our survey of the Appalachian Trail, there was a 5.6% difference between the visitation estimates based on these two approaches, and such stability reinforces confidence in the methodology.

Résumé : Plusieurs sentiers de randonnée pédestre traversent les forêts et les terres publiques partout en Amérique du Nord. Il est par conséquent devenu important pour la gestion fédérale d'acquérir une compréhension de l'utilisation totale de ces sentiers. Cependant, il n'y a jamais eu de tentative formelle pour estimer la randonnée pédestre dans ces longs sentiers d'arrière-pays. Cet article présente une approche qui utilise deux instruments de sondage (des relevés de sortie du site et un questionnaire destiné aux utilisateurs des sentiers) pour estimer les randonneurs pédestres sur un horizon temporel et spatial. À titre d'exemple, la méthodologie a été appliquée à un segment de 175 km du sentier des Appalaches du 1 juin au 14 août 2007. Deux méthodes alternatives d'estimation sont présentées et comparées. L'approche basée sur un modèle peut être préférable à l'approche basée sur un plan lorsque la taille de l'échantillon est petite parce qu'elle adoucit les estimations des strates irrégulières et produit de plus petits écarts types. Cependant, l'approche basée sur un plan assouplit une hypothèse et est plus appropriée à mesure que la taille de l'échantillon augmente. Dans notre enquête sur le sentier des Appalaches, il y avait un écart de 5,6 % entre les estimations de fréquentation basées sur les deux approches et une telle stabilité renforce la confiance dans la méthodologie.

[Traduit par la Rédaction]

Introduction

Outdoor recreation, in addition to timber, water, wildlife, and grazing, comprise the multiple-use concept of modern forest management. Outdoor recreation is very important across the North American continent as natural resource managers strive to provide a diverse set of recreational opportunities to the public and the value of recreation to local communities becomes more apparent. Thus, accurate estimates of recreational use are required for national, regional, and forest-level decision making and planning. Specifically, they are needed to determine benefits from recreational use and its impacts on other forest resources and local economies (Frantz 2007). In addition, visitation estimates are needed to determine outdoor recreational trends and to quantify the effectiveness of federal programs. Nevertheless, Loomis and Walsh (1997, p. 28) maintain that obtaining accurate measures of visitor use continues to be a problem. Loomis (2000) notes further that government agencies that supply outdoor recreational opportunities have been slow to recognize the importance of consistently collected and defensible use data.

Outdoor recreation occurs on a multitude of trails that traverse the forests across the North American continent. In the United States (US), the best known and arguably the most popular is the Appalachian Trail (3505 km), which travels through 14 states from Maine to Georgia, traversing eight national forests. The North Country Trail (7400 km) is the longest US trail and travels from New York to North Dakota, traversing seven states, 10 national forests, and over 150 public lands. Other major trails include the Continental Divide Trail (5000 km) along the Rocky Mountains from Canada to Mexico, the Pacific Crest Trail (4260 km) from southern Cal-

Received 27 July 2011. Accepted 27 September 2011. Published at www.nrcresearchpress.com/cjfr on xx November 2011.

S.J. Zarnoch. USDA Forest Service, Asheville, NC 28804, USA.
J.M. Bowker and H.K. Cordell. USDA Forest Service, Athens, GA 30602, USA.

Corresponding author: Stanley J. Zarnoch (e-mail: szarnoch@fs.fed.us).

ifornia to northern Washington, and the Pacific Northwest Trail (1900 km) from Montana to Washington. Canada is currently developing the Trans Canada Trail, which is currently 16 500 km but is scheduled to be 22 000 km, making it the longest trail of its kind in the world. It will eventually pass through every province and territory and will connect the Atlantic, Pacific, and Arctic oceans. Other noteworthy Canadian trails include the National Trail, currently 3000 km and planned to be 10 000 km from coast to coast, the International Appalachian Trail (1045 km) from Mt. Katahdin, Maine, to the Gaspé Peninsula in Quebec, and the Bruce Trail (770 km) in Ontario.

A multitude of ecological impacts can result from hiking pressure, and improved methods of estimating visitation can help manage these impacts. Hiking, horseback riding, and ATV use on trails can cause negative impacts to ecosystems, forests, and wildlife, including trampling, soil compaction and erosion, disturbance, pollution, nutrient loading, introduction of non-native invasive plant species, habitat fragmentation, and edge effects (Jordan 2000). Trampling can cause compaction of leaf litter and soil and reduction of delicate herbaceous plants and brittle woody plants. Habitat disturbance through noise and motion can affect bird behavior and movement, resulting in nesting loss. Competition from introduced exotics may be enhanced at or along trails. It has been shown that certain plant species, some non-native and exotics, occurred only at trail edges in the Rocky Mountains (Benninger et al. 1992; Dale and Weaver 1974). In addition, trails could potentially lead to habitat fragmentation and increased edge effects that can alter the microclimate by increasing rain (due to less canopy interception), sunlight, and wind, decreasing humidity, and altering temperature (Cole 1978; Dale and Weaver 1974), resulting in changes in plant and wildlife composition (Hickman 1990; Miller et al. 1998). Moreover, trails may also impede movement and dispersal of animals across openings, especially bare soil, which is common on trails.

In addition to the ecological carrying capacity issues discussed in the previous paragraph, management also faces social carrying capacity issues such as crowding and conflict. Social carrying capacity, which is determined by visitor density in space and time, has been identified as a key management issue in both national parks and protected areas (Manning 1997, 2002) and wilderness (Freimund and Cole 2001). Accurate estimating and monitoring of use levels is fundamental to effective management of social carrying capacity (Muhar et al. 2002).

Although the level of use and, thus, the economic and ecological impact from trails can be high, no survey methodology has been developed to estimate the number of hikers that use long trails through remote forested landscapes, which is vital information to efficient management of this forest resource. The primary objective of this paper was to develop a prototype survey methodology to estimate the number of hikers over a spatial and temporal span on such trails, specifically applying it to the Appalachian Trail (AT). In addition, two alternative estimation methodologies, the design-based and model-based approaches, are presented and compared using the same survey data. A secondary objective was to expand this estimate to total annual visitation for the entire AT. Despite the high profile of the AT, only very limited information on hiker visitation is known, which includes hiker characteristics, attitudes, and preferences (Kyle et al. 2004; Manning et al. 2000).

There are several reasons why this research is of interest to land managers, survey statisticians, and the various publics who hike or are interested in natural resources – environmental issues. First, given the very high profile of national trails, it is surprising that no statistically rigorous methodology has been developed to estimate recreational use. Second, it uses two survey instruments (exit-site tallies and a trail-user questionnaire) to obtain information for the visitation estimator. Third, three distinct types of estimating components (standard sites, augmented sites, and special events) are used in the visitation estimator. Fourth, designed-based and model-based estimation approaches are used. Finally, the paper demonstrates the application of a wildlife-based capture–recapture procedure to an urbanized portion of the trail during an annual festival.

Material and methods

Overview

The primary objective of this research emphasized the methodological development of an efficient design for estimating recreational visitation on a long-corridor hiking trail with urban segments. The visitation metric is defined as the total number of visits to the AT for recreation during a specific number of days. A visit is defined as one person recreating on the AT for one or more consecutive days and nights. Thus, a person who hikes for 10 days without ever leaving the AT contributes one visit. However, if this hiker spent each of the nine nights off the AT, then it would be 10 visits. Conversely, a person who hikes for only one day but leaves the AT one or more times during the day (e.g., lunch, phone call, shopping, etc.) and returns to the AT before going home contributes one visit. Thus, a person leaving the AT for the final time in any given day terminates a visit. Such a person is hereafter referred to as a last-exiting recreationist (LER).

The survey was applied to a 175 km section of the AT extending from Harpers Ferry, West Virginia (WV), to Boiling Springs, Pennsylvania (PA), from 1 June through 14 August 2007. This section of the AT provided the most complete array of site types of any similar length segment along the AT, traversing multiple states and settings representing the diversity of the entire trail from remote areas to multiple-use state parks, urban towns, a national historic park, and areas that had special event attractions. No other section of the trail would have provided an opportunity to collect data on all of these components, which were required (*i*) to demonstrate the full potential and all aspects of the survey design and (*ii*) to provide the required data for a total trail expansion, which was highly desired by the participating agencies. In addition, it was convenient to offices of study cooperators and had a history of strong local hiking club affiliations helpful for recruiting volunteers for fieldwork. The time period for the survey was selected to coincide with a period of expected high AT visitation, which would potentially provide more survey data for analysis, expose any unforeseen problems that needed to be addressed in future work, and coincide best with recruiting volunteers to administer the survey.

Days were sampled at selected exit sites to obtain average daily tallies of LERs, which were then expanded to the total visitation estimate. The methodology is based on the concept that if all exit sites are identified and the number of LERs is counted for each day of the survey period, then the sum will be the total visitation for that time period (Bergstrom et al. 1996; Bowker et al. 2007; English et al. 2002; Gregoire and Buhyoff 1999; Zarnoch et al. 2002). Some trail use studies have incorporated a sampling design based on trail segments, rather than exit sites, wherein visitors were counted with either visual or electronic means as they passed (Lindsey and Lindsey 2004). Others were based on a combination of trail censuses and electronic trail counts (Jacobi 2003). However, by counting LERs, we ensure that only recreationists will be included in the visitation estimate and that they will not be "double counted" because they are not returning to the trail during the same day and, thus, will not be counted elsewhere. An alternative approach that should give similar visitation estimates is to count only first-entering recreationists. However, when additional information such as current trip attributes, opinions about the visit, satisfaction with the facilities, etc., are desired by management, an LER approach is preferred.

Sampling frame

The survey was based on a stratified random sampling design. The sampling frame consisted of all site days from Harpers Ferry, WV, to Boiling Springs, PA, during the 75-day study period. The sampling unit was the site day, defined as any day that a given site was available for exiting by LERs. The formation of the population of site days required the identification of all exit sites along the survey area where a recreationist could exit the AT. Using GIS data, trail maps, and guide books, as well as interaction with local hiking clubs, Appalachian Trail Conservancy members, and NPS managers, 120 exit sites were identified. All were open for visitation during the 75-day period, thus implying a total of 9000 site days.

Stratification using five categories of trail type and three relative levels of LER volume resulted in 15 strata. Within each stratum, a random sample of site days was selected for the field survey. Strata were formed such that all site days within a given stratum were similar with regard to LER volumes. Before sampling frame construction, a review of the sites identified along the survey area was performed by project personnel in conjunction with members of local hiking clubs knowledgeable about the relevant segments of the AT. This resulted in identification of three general site types. An exit site consisting of a trail or road intersection across the AT was considered a trail–road (TR) site type. If a defined parking lot was proximal, as happened frequently, where the AT intersected paved roads, the site was classified as a parking (P) site type. Both of these site types would presumably exhibit almost exclusive use by AT recreationists. At other places along the AT (e.g., state parks), there was a complex network of sites, some not clearly defined, with potential for considerable non-AT use. These site types were categorized as multiple-use (MU) site types.

There were also several exit sites identified that were subsumed within Harpers Ferry National Historic Park, which comprises much of downtown Harpers Ferry, WV. The AT meanders through Harpers Ferry, creating an extremely complex set of exit sites that were dissimilar to the TR, P, and MU site types. This required the creation of a Harpers Ferry (HF) site type. Here, although AT use was significant, the proportion of non-AT users was high due to tourists visiting the many other attractions in Harpers Ferry. A final site type (ATCH) defined within Harpers Ferry consisted of the Appalachian Trail Conservancy (ATC) Headquarters office. This is a very popular site for exiting hikers because the office contains information, news, and historical aspects pertinent to the AT, as well as providing comradeship for fellow hikers.

The site days in each site type were further stratified into three use levels (low, L; medium, M; high, H) depending on the anticipated LER volume on the specific day. The boundaries for the use levels did not consist of cut points because no true visitation estimate was available for the sites. Instead, the use levels were relative and ordinal in scale, where site days within a site type were categorized into L, M, or H depending on perception of last-exiting volume by the classifiers. All site days were classified into the 15 possible strata. Although all of the strata were present on the entire AT, there were no site days in the TR-H or ATCH-L strata for the 175 km survey area, thus resulting in only 13 strata for sampling. The complete sampling frame by strata is shown in Table 1.

Standard sites, augmented sites, and special event sites

The exit sites were classified as either standard sites, augmented sites, or special event (SE) sites. Standard sites were those for which there was no information about visitation available from any sources except the survey itself. Augmented sites had objective information from an auxiliary source that could be used, alone or in conjunction with survey information, to estimate recreational visitation. For example, the ATCH office maintained daily tallies of visitors throughout the year. Such information could be combined with sample estimates of the percentages of ATCH visitors who were AT LERs to yield an estimate for that site type. All TR, P, and MU site types were standard sites, and the HF and ATCH site types were augmented sites (Table 1). In general, it is possible that a given physical site may be considered a standard site for certain times during a sample period and an augmented site for other times during the sample period, depending on the availability of auxiliary information.

While performing the stratification process, it became evident that some site days may have extremely high visitation due to special events in the vicinity. For this occurrence, a special event category was created in which visitation was estimated in a different manner and added to the final estimate. The only special event for the survey was Foundry Day on 2 June 2007 at Boiling Springs, PA. On this day, five exit sites in Boiling Springs were subsumed into the special event sampling, eliminating three site days from stratum MU-L and two from MU-H.

Augmented sites can increase the efficiency of the survey by incorporating auxiliary information or variables. An auxiliary variable is one for which information is available prior to sampling (Sarndal et al. 1992, p. 219). Incorporating auxiliary information can yield variances of the estimates associated with augmented sites that are less than those of the standard sites, thus improving the overall visitation estimate. A study of US National Forest visitation found that augmented sites can lead to cost savings and variance reduction

Table 1. The total site days in each of the survey strata based on site type and use level, the original designed allocation of the sample of site days for the sampling periods from 0800 to 1400 hours (AM) and from 1400 to 2000 hours (PM) (actual achieved sample days in parentheses), and the total site days for the entire Appalachian Trail (AT) for the whole year.

		Allocated sample size (achieved)			
Strata	Survey total site days[a]	AM	PM	Total	AT total site days[a]
TR-L	3 624	4 (4)	6 (8)	10 (12)	184 988
TR-M	651	4 (3)	6 (6)	10 (9)	8 112
TR-H	0	—	—	—	9 490
P-L	1 781	4 (4)	6 (7)	10 (11)	81 429
P-M	689	5 (6)	10 (7)	15 (13)	19 140
P-H	80	9 (8)	16 (12)	25 (20)	8 625
MU-L	1 161	4 (4)	6 (3)	10 (7)	13 279
MU-M	542	5 (5)	10 (10)	15 (15)	3 028
MU-H	167	9 (7)	16 (14)	25 (21)	2 883
HF-L	156	1 (0)	3 (2)	4 (2)	791
HF-M	23	2 (2)	3 (2)	5 (4)	156
HF-H	46	4 (2)	7 (8)	11 (10)	148
ATCH-L[b]	0	—	—	—	120
ATCH-M[b]	14	—	—	2 (2)	157
ATCH-H[b]	61	—	—	4 (4)	88
Total	8 995	51 (45)	89 (79)	146 (130)	332 434

Note: Allocated sample size may be slightly more or less than the achieved sample size due to uncontrollable field circumstances. The actual achieved sample days are provided in parentheses. Standard site types: TR, trail–road; P, parking; MU, multiple use. Augmented site types: HF, Harpers Ferry; ATCH, Appalachian Trail Conservancy Headquarters office. Use levels: L, low; M, medium; H, high.

[a]Does not include Foundry Day at Boiling Springs, Pennsylvania, on 2 June 2007, which included five site days.

[b]ATCH was sampled for approximately 7–8 h each day so no AM and PM is indicated. Sample sizes are shown in the other columns.

(English et al. 2003). Despite their advantages, augmented sites are difficult to identify and may require unique methods of estimation that add complexity to the sampling methodology.

Sample selection

After adjusting for the Foundry Day special event, the sampling frame consisted of 8995 standard and augmented site days from which 146 sample days were selected randomly within the specified strata according to the sample allocation shown in Table 1. Neyman allocation (Cochran 1977) was used in conjunction with judgment to determine allocation of the limited labor and financial resources. The stratum sizes, which are required information for Neyman allocation (Table 1, second column), were obtained by classifying all exiting sites in the 175 km survey area into their appropriate site type and use level. No previous information was available on the strata standard deviations. Thus, relative standard deviations that increased with use level were used in the Neyman allocation process.

Due to the typical 8 h work day, the survey sampling was based on a 6 h interviewing period for each selected sample site day. This 6 h sampling period for the TR, P, MU, and HF site types was allocated randomly, with approximately one-third from 0800 hours to 1400 hours (AM) and two-thirds from 1400 hours to 2000 hours (PM). This disproportionate sampling allocation was used because exiting visitation was believed to be higher after noon. The augmented ATCH site type was open for visitation from 0900 to 1700 (1600 on weekends), where daily visitor tallies were kept by staff at the ATCH office. For these sampled site days, the interview period was extended to 8 h (7 h) to coincide with the daily visitor tally data. A calendar of sampling days (with backup days in case the scheduled ones were missed) was developed by randomly selecting the allocated size sample from all site days in each stratum. The sample sizes that were actually achieved during the survey process are shown in Table 1.

Data collection procedures

A complex stratified cluster sampling design was used in which the primary sampling unit was the site day within a stratum and the secondary sampling unit was the group interviewed on a given site day. A mixed-mode data collection design was used to efficiently obtain the required survey data (de Leeuw 2005; de Leeuw et al. 2008). One mode was a simple tally of all groups of recreationists as they exited from the survey site. The other mode was a more intensive, face-to-face interview of a random sample of exiting groups. This mixed-mode design allowed us to obtain both the exit tallies and the interview data without the need to contact all groups, which would have been not only costly, but also very cumbersome at some sites that had many recreationists.

The data collection procedure at each sampled site day consisted of a 6 h on-site tally of either all people or all groups that were exiting the site. A group is typically an identifiable collection of travelers (e.g., lone individual, family, or friends) who come to hike together, engage in similar activities, and leave together in a vehicle or by walking. The survey was based on interviewing groups, and thus group tallies were the appropriate unit because the group was the sampling unit for the interviews. At some sites, the group was easily tallied because it was contained in a vehicle. At other sites, exiting groups of people may mix together causing tally problems. Therefore, the people-tally sample days were converted to vehicle-tally sample days by dividing the people tally by the average group size. Survey interviews were conducted on a randomly selected person (most recent birthday) from a random sample of groups that were exiting. Interviewees were asked initial questions that gathered basic information needed for the estimation process. Questions included whether the respondent used the trail, was recreating, and was exiting the trail for the last time that day. This information was sufficient to determine LER status (yes, no) of the group. The rest of the questionnaire was administered only to LERs who were using the trail for recreation. These questions addressed arrival time, hiking distance, frequency of previous visits, demographics, management preferences, and other trip attributes. In addition, the number of people in each group was determined for both LER and non-LER groups. Details on the survey procedures, including the questionnaires, can be obtained from the authors.

Total trail estimate

A secondary objective was to obtain an annual visitation estimate for the entire AT, even though the survey only sampled within a restricted spatial and temporal space. It is best to sample the entire trail throughout the entire year, but this was unfeasible due to time and financial constraints. So although this total trail estimate is not based on sampling throughout the entire year, all site days on the whole trail were first identified and then classified into strata for the entire year. This approximation, while subject to criticism, is a potential method that is useful in many situations. It must be emphasized that the stratification of the entire trail for the entire year was an intensive process that was performed in conjunction with personnel associated with the Appalachian Trail who knew of the temporal and spatial patterns in visitation. Generally, sites during winter months were more typically classified in the lower use strata than during the summer. This allowed for a reasonable expansion for the total trail estimate based on the summer data, because the summer data also contained estimates for the lower use levels by stratum.

The annual sampling frame for the entire AT consisted of 953 sites yielding 332 434 site days distributed by site type and use level (Table 1). The augmented survey data for site types ATCH and HF were obtained for the entire year from the ATCH and Harpers Ferry National Historic Park monthly visitation estimates. Classification of the site days for the AT outside the 175 km survey space and time was accomplished by the project manager meeting with 33 different local area representatives throughout the entire AT who were familiar with the use-level patterns in their areas. No additional sample days were selected for actual field sampling in any of the strata. This sampling frame provided the strata weights (Table 1) that were required for the total trail visitation estimate. Unlike the survey, this population contained site days in all 15 strata, including TR-H and ATCH-L. No additional augmented site days or special events were identified in this sampling frame. By not identifying additional augmented site days (if they existed), auxiliary information potentially valuable in reducing the variance may have been lost. Moreover, no additional special events like Foundry Day were identified by either ATC or NPS staff.

The total trail estimate was based on appropriate strata weights but lacked the required spatial and temporal tally and interview data. Thus, major assumptions were imposed. One assumption was that certain parameter estimates for the population were the same as for the sample. The stratification process of classifying sites days on the total AT for the entire year into site types and use levels should help to satisfy this assumption. However, it must be emphasized that the spring, fall, and winter did not contribute any site data to the estimates, and thus, some differences could exist temporally and spatially. However, a site with a low use level in the summer should theoretically be similar to a site classified as low use level in the winter, because in the stratification process, use level was invariant to time of year, day of week, or region of trail. A second assumption is that the exit tallies are the same for the sample and the population. Although the stratification process may mitigate this problem somewhat, the survey strata means may be biased upwards because they are based only on summer data when there may be more opportunity for larger groups with children to use the AT. Alternatively, one could argue that groups may be larger during the school year as sites along the AT are often used for school outings.

Estimation methodology

Overview

Two approaches were used to obtain the weighting adjustment factors ($\overline{P}_h$, $\overline{G}_h$, and $\overline{G}_h^{\mathrm{a}}$, which are explained in the next section) for the visitation estimators in the strata. The traditional design-based approach uses estimators for the weighting adjustment factors based on sample data from each stratum according to a cluster survey design (Cochran 1977), which is the optimal method given sufficient sample data. However, when resources are limited, a model-based approach may be more appropriate. Here, a linear model was developed by using the data from all sampled strata to model the weighting adjustment factors. The fixed components of the model were site type and use level, specifically,

$$[1] \qquad y_{hijk} = \mu + s_h + u_i + \varepsilon_{hijk}$$

where y_{hijk} is the weighting adjustment factor ($\overline{P}_h$, $\overline{G}_h$, or $\overline{G}_h^{\mathrm{a}}$) for group k on site day j in site type h and use level i, μ is the overall mean, s_h is site type h (h = TR, P, MU, HF, or ATCH), u_i is use level i (i = L, M, or H), and ε_{hijk} is the normally distributed error for group k on site day j in site type h and use level i. The model-based approach smoothes out the design-based estimates, which can be erratic when strata sample sizes are small. This is accomplished by not including the typical interaction term su_{hi} in the model. The clustering of the observations within the sample days could be addressed by treating them as repeated measures and specifying a covariance structure such as compound symmetry (CS). However, the above model will yield identical estimates (with slightly different variances) to the design-based approach if the interaction term su_{hi} is included and the variance components (VC) covariance structure is used, implying a common variance and zero covariance for the observations within a cluster. Thus, the model-based approach assumed the VC covariance structure to produce the weighting adjustment factors. Alternative covariance structures are possible and may be studied in future research, but for simplicity and to be compatible with the covariance structure of the design-based estimates, the VC covariance structure was used, which is equivalent to a model with no repeated measures.

The model-based approach also allows prediction of estimates for strata unrepresented in the sample data as long as all site types and use levels are represented in the data. Thus, the modeling approach is flexible and accommodates such issues. In this situation, the design-based approach would have to use a more arbitrary method to obtain an estimate. For the 175 km segment, all of the identified 13 strata were sampled, and both approaches were used. However, the total trail estimate required estimates for two strata, TR-H and ATCH-L, which were not represented by survey data. In such situations, one alternative for the design-based approach is to use the estimates and standard errors from the nearest use level in that site type. Another approach to creating TR-H might be to add the difference (or some percentage thereof) between the TR-L and TR-M use levels to the TR-M use level. Similarly, to create ATCH-L, one could subtract the difference (or some percentage thereof) between the ATCH-H and ATCH-M use levels from ATCH-M. Nevertheless these simple alternatives are arbitrary. The model-based approach directly predicts these unrepresented strata and has the potential to yield smaller variances because the estimates are based on data from all of the sampled strata and not just one stratum, as is done in the design-based approach.

Estimator accuracy is a function of its bias and variance, and thus, when evaluating various estimators, trade-offs between the bias and variance must be considered. The design-based estimators are unbiased, whereas the model-based estimators have a potentially unknown bias. In contrast, the variance for the design-based approach is considerably larger than that for the model-based approach. Thus, it is difficult to compare the accuracy of the design-based and model-based estimators. However, it is often more desirable to have a biased estimator that has a smaller variance, like the model-based estimators, than an unbiased estimator with a larger variance. Moreover, management is often more interested in change estimates for visitation between two points in time as opposed to absolute estimates. When this is the case, the potentially biased, lower variance, model-based estimators may be preferred to the unbiased, higher variance, design-based estimators. This is especially true when repeated surveys are not an option, i.e., the researcher has only one opportunity to collect data.

The visitation estimator

Total visitation for the survey from 1 June through 14 August 2007 was defined as

$$[2] \qquad \text{VISITS} = \text{SS} + \text{AS} + \text{SE}$$

where SS is the total number of standard site visits, AS is the total number of augmented site visits, and SE is the total number of visits from the special events at Boiling Springs, PA. Each of these three components required a different estimation methodology.

Standard site component

The standard site component consisted of all sites in the TR, P, and MU site types and all three use levels (L, M, and H) and is estimated as

$$[3] \qquad \widehat{\text{SS}} = \sum_{h=1}^{8} N_h \overline{P}_h \overline{C}_h \overline{G}_h$$

The correlation between the variables in eq. 3 was small, so independence was assumed, resulting in the estimated variance (Goodman 1960):

$$[4] \qquad \widehat{V}(\widehat{\text{SS}}) = \sum_{h=1}^{8} N_h^2 \left\{ \overline{P}_h^2 \overline{C}_h^2 \widehat{V}(\overline{G}_h) + \overline{P}_h^2 \overline{G}_h^2 \widehat{V}(\overline{C}_h) + \overline{C}_h^2 \overline{G}_h^2 \widehat{V}(\overline{P}_h) \right\} - \sum_{h=1}^{8} N_h^2 \left\{ \overline{P}_h^2 \widehat{V}(\overline{C}_h)\widehat{V}(\overline{G}_h) + \overline{C}_h^2 \widehat{V}(\overline{P}_h)\widehat{V}(\overline{G}_h) + \overline{G}_h^2 \widehat{V}(\overline{P}_h)\widehat{V}(\overline{C}_h) \right\} + \sum_{h=1}^{8} N_h^2 \widehat{V}(\overline{P}_h)\widehat{V}(\overline{C}_h)\widehat{V}(\overline{G}_h)$$

where N_h is the total number of site days in stratum h, weighting adjustment factor $\overline{P}_h$ is the proportion of exiting groups in stratum h that are LERs, $\overline{C}_h$ is the average daily number of exiting groups of people (LERs and non-LERs) tallied in stratum h, weighting adjustment factor $\overline{G}_h$ is the average size of the LER group in stratum h, and $\widehat{V}(\overline{P}_h)$, $\widehat{V}(\overline{C}_h)$, and $\widehat{V}(\overline{G}_h)$ are the estimated variances of $\overline{P}_h$, $\overline{C}_h$, and $\overline{G}_h$, respectively. Note that the index of summation does not include strata TR-H because it does not exist in the survey, leaving only eight sampled strata.

An estimate for eq. 3 required four components. The N_hs were the known strata sizes (Table 1), whereas the $\overline{P}_h$, $\overline{C}_h$, and $\overline{G}_h$ were estimated from the tally and interview survey data. Let n_h be the number of sample days in stratum h, m_{hi} be the number of groups interviewed on sample day i in stratum h, $p_{hij} = 1$ if group j on sample day i in stratum h was an LER, otherwise $p_{hij} = 0$. In addition, let g_{hij} be the number of people in interview group j on sample day i in stratum h for an LER group, and g_{hij}^{a} be the number of people in interview group j on sample day i in stratum h for any type of group (the superscript "a" refers to "all" groups). Let c_{hi}^{v} be the number of vehicles (or groups) tallied exiting the AT from sample day i in stratum h during the 6 h interview period.

The survey data can be used to obtain a ratio of means estimator for $\overline{P}_h$ and $\overline{G}_h$ by using p_{hij} and g_{hij}, respectively. The design-based estimators and variances are defined as

$$[5] \qquad \overline{P}_h = \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} p_{hij}}{\sum_{i=1}^{n_h} m_{hi}}$$

$$[6]\qquad \widehat{V}(\overline{P}_h) = \frac{1}{n_h(n_h-1)\left(\frac{1}{n_h}\sum_{i=1}^{n_h} m_{hi}\right)^2} \times \left\{ \sum_{i=1}^{n_h}\left(\sum_{j=1}^{m_{hi}} p_{hij}\right)^2 + \overline{P}_h^2 \sum_{i=1}^{n_h} m_{hi}^2 - 2\overline{P}_h \sum_{i=1}^{n_h}\left(m_{hi}\sum_{j=1}^{m_{hi}} p_{hij}\right)\right\}$$

$$[7]\qquad \overline{G}_h = \frac{\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} g_{hij}}{\sum_{i=1}^{n_h} m_{hi}}$$

$$[8]\qquad \widehat{V}(\overline{G}_h) = \frac{1}{n_h(n_h-1)\left(\frac{1}{n_h}\sum_{i=1}^{n_h} m_{hi}\right)^2} \times \left\{ \sum_{i=1}^{n_h}\left(\sum_{j=1}^{m_{hi}} g_{hij}\right)^2 + \overline{G}_h^2 \sum_{i=1}^{n_h} m_{hi}^2 - 2\overline{G}_h \sum_{i=1}^{n_h}\left(m_{hi}\sum_{j=1}^{m_{hi}} g_{hij}\right)\right\}$$

These estimators consider the clustering of the data on a site-day basis. The sample allocation of one-third AM and two-thirds PM was taken into account by appropriate weighting in the estimation process. Missed sampling days and sample days without any interviews (i.e., no p_{hij} or g_{hij} available) had an additional affect on the weighting. Designed-based and model-based estimates were computed using PROC SURVEYMEANS and PROC MIXED, respectively (SAS Institute Inc. 2004). The model-based estimates use a complicated iterative methodology, which will not be presented here (see SAS Institute Inc. (2004, pp. 2731 to 2739)).

To obtain the average daily tally $\overline{C}_h$, the arithmetic mean of $2c_{hi}^{\mathrm{v}}$ was used because the daily count data were not clustered. The constant "2" expands the 6 h tally to a full 12 h recreation day. To account for the unequal AM and PM sampling distribution, separate estimates for a stratum were computed for AM and PM and then combined by simple averaging, which achieved the appropriate weighting.

Augmented site component

Site type ATCH

The augmented site type ATCH data contained 75 daily visitor tallies obtained by personnel at the ATCH office in Harpers Ferry, WV, from 1 June to 14 August 2007. This was combined with the estimates $\overline{P}_h$ and $\overline{G}_h$ obtained from the six site days randomly sampled during the 75-day period. Let N_h be the number of days that the ATCH office has days in use level h (h = M, H) during the survey and $\mathrm{AS}_{hi}^{\mathrm{ATCH}}$ be the ATCH visitation tally (this counts people, not groups) on day i in use level h, then the average daily augmented site-type visitation tally $\overline{\mathrm{AS}}_h^{\mathrm{ATCH}}$ in use level h is the arithmetic mean. The estimate of the total augmented site-type $\mathrm{AS}^{\mathrm{ATCH}}$ for the ATCH for the survey is then defined as

$$[9]\qquad \widehat{\mathrm{AS}}^{\mathrm{ATCH}} = \sum_{h=M}^{H} N_h \overline{P}_h \left(\frac{\overline{\mathrm{AS}}_h^{ATCH}}{\overline{G}_h^{\mathrm{a}}}\right)\overline{G}_h$$

where weighting adjustment factor $\overline{P}_h$ is the proportion of groups exiting the ATCH that were LERs, weighting adjustment factor $\overline{G}_h^{\mathrm{a}}$ is the average group size for all groups exiting the ATCH, and weighting adjustment factor $\overline{G}_h$ is the average group size for all LER groups exiting the ATCH. The $\overline{P}_h$, $\overline{G}_h$, and $\overline{G}_h^{\mathrm{a}}$ were estimated using design-based and model-based approaches.

Site type HF

The augmented site type HF consisted of three sites in Harpers Ferry, WV. Here, the official monthly NPS recreational visitation for Harpers Ferry National Historic Park was available as augmented site data (National Park Service 2008). These visitation estimates based on 75 days were considered superior to estimates derived from only the 20 site days in the survey. However, this augmented site data (*i*) were not stratified by use level and (*ii*) were only available on a monthly basis. To resolve these problems, the weighting adjustment factors $\overline{P}_h$, $\overline{G}_h$, and $\overline{G}_h^{\mathrm{a}}$ were converted to a weighted monthly average. The three weights for each of these estimators were the number of site days in use level L, M, and H. The weighted monthly average was then obtained by weighting the individual three use-level estimates obtained from the survey for site type HF by these strata weights. Because the NPS tallied people, units were converted to groups. Thus, the augmented site-type visitation estimate for site type HF is

$$[10]\qquad \widehat{\mathrm{AS}}^{\mathrm{HF}} = \sum_{i=6}^{8} k\overline{P}_i \frac{\mathrm{AS}_i^{\mathrm{HF}}}{\overline{G}_i^{\mathrm{a}}}\overline{G}_i$$

where k = 1 if i = 6 or 7 (June or July) and k = 14/31 if i = 8 (August), and $\mathrm{AS}_i^{\mathrm{HF}}$ is the official NPS visitation at Harpers Ferry, WV, for month i, i = 6, 7, or 8. Month 8 $\mathrm{AS}_i^{\mathrm{HF}}$ is multiplied by (14/31) to reflect that the survey terminated on 14 August and only that proportion of the monthly August augmented site-type data should be included.

Special event component

Foundry Day at Boiling Springs, PA, on 2 June 2007 was deemed a special event because that annual event draws thousands of visitors, thus increasing AT visitation on that day to a level greatly exceeding the defined strata for standard site days. The special event estimator $\widehat{\mathrm{SE}}$ is the total number of AT recreational visitors in Boiling Springs, PA, on 2 June 2007 and is

$$[11]\qquad \widehat{\mathrm{SE}} = \overline{P}^{\mathrm{SE}}\widehat{\mathrm{NG}}^{\mathrm{SE}}\overline{G}^{\mathrm{SE}}$$

where $\overline{P}^{\mathrm{SE}}$ is the proportion of all groups interviewed that were AT recreating groups, $\widehat{\mathrm{NG}}^{\mathrm{SE}}$ is the number of groups of visitors (AT and non-AT) in Boiling Springs, PA, and $\overline{G}^{\mathrm{SE}}$ is the average group size for AT recreating groups of visitors.

To obtain an estimate for $\widehat{\mathrm{NG}}^{\mathrm{SE}}$, an initial investigation indicated that visitors primarily came to Boiling Springs via shuttle buses that operated out of a school parking lot in Boiling Springs, whereas others came in private vehicles or walked to town. A mark–recapture method used for estimating animal abundance (Seber 1982) was modified to estimate $\widehat{\mathrm{NG}}^{\mathrm{SE}}$ by simply adding an interview question asking if the respondent used the shuttle buses. Based on their response, the visitors were indirectly "marked." The Lincoln–Petersen estimator (Seber 1982), a simple mark–recapture estimator, defined as

$$[12] \qquad \widehat{\mathrm{NG}}^{\mathrm{SE}} = \frac{(B_1 + 1)(I_2 + 1)}{(M_2 + 1)} - 1$$

was applied, where B_1 is the number of groups that took the shuttle bus, I_2 is the number of groups interviewed in Boiling Springs, PA, and M_2 is the number of groups interviewed that took the shuttle bus ("marked" groups).

Results

Weighting adjustment factors

A comparison of the weighting adjustment factors $\overline{P}_h$, $\overline{G}_h$, and $\overline{G}_h^{\mathrm{a}}$ revealed that the model-based approach produced more stable and plausible estimates than the design-based approach. For example, the design-based approach yielded $\overline{P}_h = 0.500$ for strata TR-L and $\overline{P}_h = 1.000$ for strata TR-M (Fig. 1). It is unlikely that the underlying true values diverged to this degree and the difference is probably an artifact of the extremely low sample size of only eight and two interviewed groups within each stratum, respectively. Thus, even though 12 and nine site days were sampled in these two strata, very few visitors were observed. The problem recurs for stratum HF-L when compared with strata HF-M and HF-H. Alternately, the model-based approach produces more plausible estimates in these situations by incorporating the data from all site types and use levels (Fig. 1). For all site types, use level L is substantially less than M and H, which are practically identical. In addition, the highest $\overline{P}_h$ is for site type P followed by TR, which is expected because these site types are used predominantly by AT hikers. The ATCH and MU site types have lower $\overline{P}_h$ because of the high proportion of non-AT visitors at these sites, i.e., less than half are knowingly using the AT. Similarly, the lowest $\overline{P}_h$ is for site type HF, with an even higher proportion of visitors that are recreating at Harpers Ferry National Historic Park, but not using the AT. The standard errors (not shown) were also more plausible for the model-based approach, typically being less than half those from the design-based approach. The estimated $\overline{G}_h$ and $\overline{G}_h^{\mathrm{a}}$ revealed similar patterns (Figs. 2 and 3); although their standard errors were usually lower for the model-based approach, the differences were less than for $\overline{P}_h$.

The average strata tallies per sample day, computed for the standard site types TR, P, and MU, represent the average number of all exiting groups (LERs or not) based on a 12 h recreation day. The results follow expectations provided by local trail experts (Table 2). The use-level relationship was also logical, with L being less than M, which was less than H. This provides further evidence that the stratification process was appropriate.

Visitation estimates

The visitation estimates for each stratum used the weighting adjustment factors obtained by the design-based and model-based approaches and the estimators previously defined. The special event estimate for Foundry Day in Boiling Springs was $\widehat{\mathrm{SE}} = 3032$, based on the mark–recapture estimate of $\widehat{\mathrm{NG}}^{\mathrm{SE}} = 1798$ and the other survey components $\overline{P}^{\mathrm{SE}} = 0.5455$ and $\overline{G}^{\mathrm{SE}} = 3.0909$. The strata level and total visitation estimates are shown in Table 3. Converting total visits in each stratum to visits per day reveals trends for the use levels, with L, M, and H reflecting monotonically increasing daily visitation for all site types except TR. These strata estimates indicate that the bulk of the visitation was from the P and M site types. Although there were many TR site days, the average visits per day were so low that their totals were about half or less than those for MU or P site types. The augmented site types, ATCH and HF, had relatively large average daily visitation, but few site days in the survey, resulting in a low share of total visitation.

Comparison of the total visit estimates by strata for the designed-based and model-based approaches revealed they

Fig. 1. The relationship of estimated $\overline{P}_h$ and use level for the five site types using the design-based approach (top graph) and the model-based approach (bottom graph). The circles represent the standard site-type strata, where TR is represented by the continuous line, P, the long-dash line, and MU, the short-dash line. The stars represent the augmented site-type strata, where HF is represented by the continuous line, and ATCH, the long-dash line.

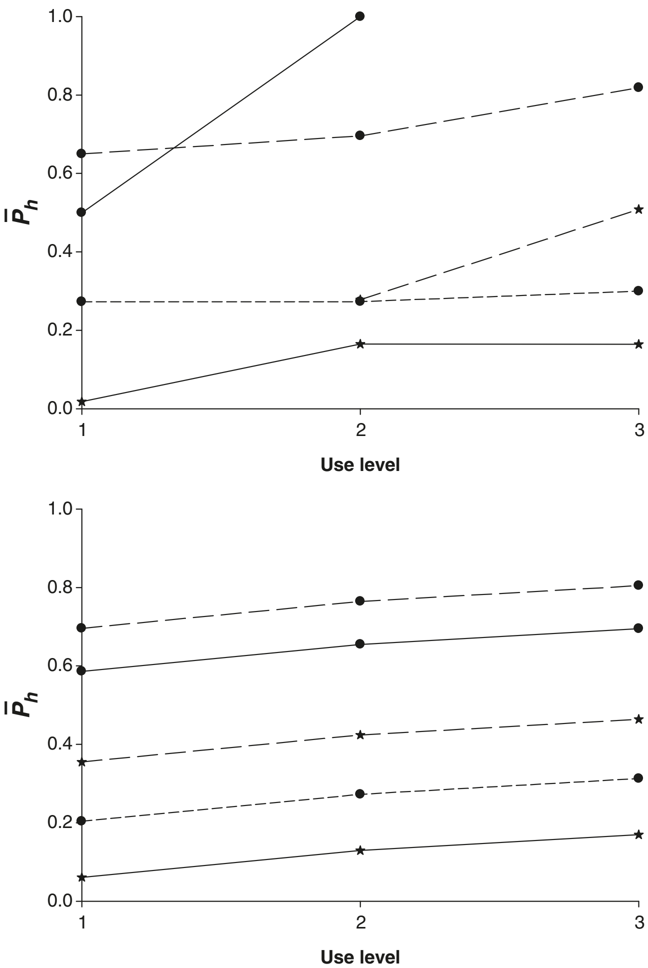


Fig. 2. The relationship of estimated $\overline{G}_h$ and use level for the five site types using the design-based approach (top graph) and the model-based approach (bottom graph). The circles represent the standard site-type strata, where TR is represented by the continuous line, P, the long-dash line, and MU, the short-dash line. The stars represent the augmented site-type strata, where HF is represented by the continuous line, and ATCH, the long-dash line.

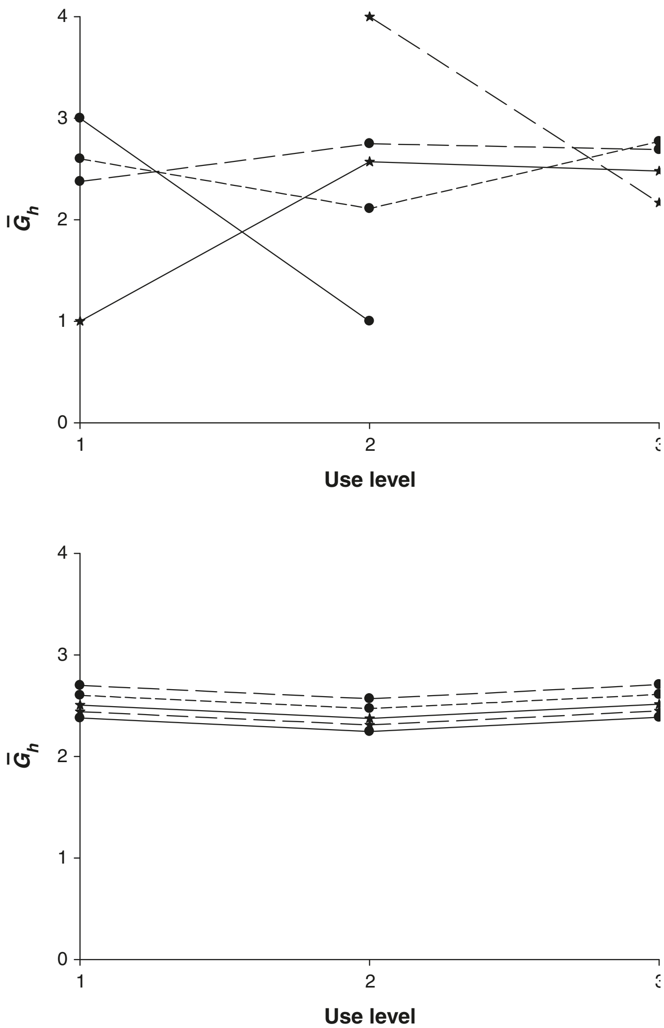
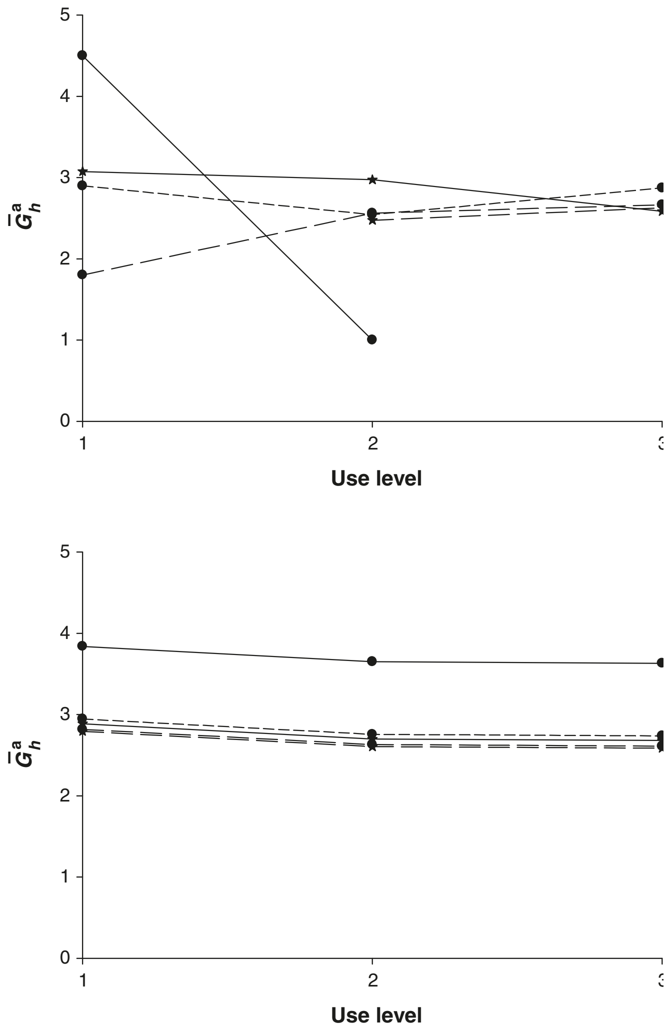


Fig. 3. The relationship of estimated $\overline{G}_h^{\mathrm{a}}$ and use level for the five site types using the design-based approach (top graph) and the model-based approach (bottom graph). The circles represent the standard site-type strata, where TR is represented by the continuous line, P, the long-dash line, and MU, the short-dash line. The stars represent the augmented site-type strata, where HF is represented by the continuous line, and ATCH, the long-dash line.



were similar for site types P, MU, and ATCH for use level H. For use levels L and M, they were substantially different with no discernable pattern. The ATCH site type produced similar estimates for both approaches, but the HF site type with the model-based approach yielded almost three times as much visitation as the design-based approach. In nearly all strata, the coefficient of variation for the model-based approach was less than that for the design-based approach. However, both have many confidence intervals that are relatively wide (normality was assumed for all 95% confidence intervals). Thus, strata-level estimates of visitation should be viewed with caution because of variability due to small sample size.

Using the model-based approach, the total visitation for the survey from 1 June through 14 August 2007 on the AT from Harpers Ferry, WV, to Boiling Springs, PA, was 70 912, with 95% confidence intervals of 48 678 to 93 146. The coefficient of variation was 16%. Eighty-three percent of the visitation was from standard sites. The three HF sites at Harpers

Table 2. The average number of all exiting groups (LERs or not) per sample day for each stratum adjusted to a 12 h recreation day.

Site type	Use level	Sample days	Average	SE
TR	L	12	1.88	0.76
TR	M	9	0.51	0.38
TR	H	0	0.51[a]	0.38[a]
P	L	11	2.52	1.54
P	M	13	6.75	2.14
P	H	20	26.70	4.62
MU	L	7	13.69	9.60
MU	M	15	31.20	6.97
MU	H	21	51.94	17.44

Note: See Table 1 or text for explanation of site-type and use-level codes.

[a]The TR-H stratum had no sample days so the average and standard error (SE) are based on the TR-M strata.

Table 3. The design-based and model-based visitation estimates, standard errors (SE), coefficients of variation (CV), and 95% confidence intervals (site type HF represents the summation over all use levels and all three sites because monthly augmented site-type data were used).

Estimator	Strata	Total site days	Visits/day	Total visits	SE	CV	Lower 95%	Upper 95%
Design based	TR-L	3 624	2.8	10 217	8 505	83	–6 453	26 887
Model based	TR-L	3 624	2.6	9 493	5 318	56	–931	19 918
Design based	TR-M	651	0.5	329	244	74	–150	808
Model based	TR-M	651	0.7	484	390	80	–279	1 248
Design based	TR-H	0	—	—	—	—	—	—
Model based	TR-H	0	—	—	—	—	—	—
Design based	P-L	1 781	3.9	6 923	4 792	69	–2 469	16 315
Model based	P-L	1 781	4.7	8 422	5 355	64	–2 073	18 917
Design based	P-M	689	12.9	8 893	4 710	53	–339	18 125
Model based	P-M	689	13.2	9 127	3 094	34	3 063	15 191
Design based	P-H	80	58.8	4 704	1 072	23	2 603	6 806
Model based	P-H	80	58.2	4 653	861	19	2 965	6 341
Design based	MU-L	1 161	9.7	11 269	8 778	78	–5 936	28 475
Model based	MU-L	1 161	7.3	8 430	6 217	74	–3 756	20 616
Design based	MU-M	542	18.0	9 732	4 331	45	1 243	18 221
Model based	MU-M	542	21.0	11 382	3 148	28	5 213	17 551
Design based	MU-H	167	43.1	7 205	3 241	45	852	13 557
Model based	MU-H	167	42.4	7 078	2518	36	2 144	12 012
Design based	ATCH-L	0	—	—	—	—	—	—
Model based	ATCH-L	0	—	—	—	—	—	—
Design based	ATCH-M	14	17.3	242	121	50	5	479
Model based	ATCH-M	14	14.5	203	59	29	87	318
Design based	ATCH-H	61	29.6	1 806	521	29	785	2 827
Model based	ATCH-H	61	31.0	1 892	469	25	973	2 811
Design based	HF	75	34.8	2 614	499	19	1 636	3 591
Model based	HF	75	89.5	6 716	1 740	26	3 306	10 125
—	Special	—	—	3 032	1 765	58	–427	6 491
Design based	Total			66 967	15 122	23	37 328	96 605
Model based	Total			70 912	11 344	16	48 678	93 146

Note: See Table 1 or text for explanation of strata (site type and use level) codes.

Ferry contributed 9%, whereas the ATCH accounted for 3%. Foundry Day accounted for 4% of the visitation. For the survey, the design-based approach was consistent with the model-based approach, yielding total visitation of 66 967, but its confidence intervals and coefficient of variation were substantially larger (Table 3).

Discussion

Survey design

The major objective of this research was to develop a prototype survey design that could be used for estimating visitation on long hiking trails. The prototype was designed for and applied to a 175 km section of the Appalachian Trail. The survey design produced a sampling frame based on exit sites stratified by site type and use level. This structure could easily be applied to other trails with only minor modifications such as additional site types and (or) use levels.

The distinction between standard and augmented sites proved useful and resulted in more efficient estimators. Augmented sites are appealing because they contain auxiliary information that can be exploited to produce a less variable visitation estimate based on data collected outside the survey. In addition, this can lead to more efficient use of limited resources. In this study, there were only two augmented site types, HF and ATCH, but on other trail surveys, there may be numerous augmented site types depending on the level of auxiliary information that is known about the exit sites. The HF augmented site type illustrated use of auxiliary data consisting of monthly visitation records, whereas the ATCH augmented site type was based on daily visitation tallies. Each required a different method to convert the respective auxiliary data to visitation estimates. Undoubtedly, other trails could have fee tickets, parking lot traffic counters, or mandatory registration that would require further adaptation of these methods for conversion.

One atypical "special event" was identified in the survey. A wildlife-based mark–recapture method was adapted to estimate visitation on Foundry Day, an annual festival in Boiling Springs, PA, the center of which was traversed by the AT. Isolating a special event instead of simply including the five site days in their appropriate strata (three in MU-L and two in MU-H) resulted in 3032 visits instead of 3(7.3) + 2(42.4) = 106.7 visits. This large difference emphasizes the importance of identifying special events. Special events require knowledge of the site(s) and may require innovative techniques besides mark–recapture methods. In some cases, they may provide complete censuses of LERs based on the administrative and (or) coordination activities associated with the particular events.

Low sampling intensity

A major problem encountered with backcountry trail visitation estimation is that on-site sampling invariably produces sample days with low or no visitation, thus yielding limited data for estimation of $\overline{P}_h$, $\overline{G}_h$, and $\overline{G}_h^{\mathrm{a}}$. Low sampling intensity not only results in erratic and variable weighting adjustment factor estimates, as shown in this paper, but also is the reason why the visitation estimators are based on the product of their means instead of the mean of their individual daily product.

The model-based approach attempts to mitigate the effect of small sample sizes on the design-based estimator. Results for the weighting adjustment factor estimates reveal several justifications for preferring the model-based approach over the design-based approach. First, limited resources can result in low sampling intensity, which increases the risk of erratic estimates for the design-based approach. This is alleviated somewhat with the model-based approach because the data are combined, the relationship between site type and use level is modeled, and then the individual strata estimates are obtained from the model. Second, the model-based approach smoothes the weighting adjustment factors so that inconsistencies are eliminated or mitigated. For instance, if a parameter increases with increasing use level for a given site type, it probably exhibits this pattern for the other site types. The design-based approach does not have this property because the individual strata estimates are not linked via a common model. Third, the standard errors of the weighting adjustment factor estimates are substantially smaller with the model-based approach because the data are combined and used jointly in the estimation process. Alternatively, the design-based approach estimates a stratum's weighting adjustment factors based only on the data observed in that stratum, resulting in a smaller sample size and, consequently, a larger standard error. Although the overall visitation estimate for both approaches are similar, if individual strata estimates are desired, the model-based approach appears more appropriate because the weighting adjustment factors do not fluctuate as widely as they do for the design-based approach.

Despite the advantages of the model-based approach when the sampling intensity is low, the design-based approach is preferred when adequate sampling is achievable. In recreational use studies, particularly with backcountry trails, this is rarely the case. With the design-based approach, the estimate for each stratum is always independent of the other strata and reflects its individual characteristics and properties. The model-based approach relies on a model and assumes no interaction between site type and use level, which can lead to biased estimates. Nevertheless, accepting some degree of bias may be a small trade-off for decreased variability, particularly if individual strata estimates are desired.

Toward a total trail estimate

In the absence of time and resources necessary to adequately sample the entire AT, study weighting adjustment factor estimates, along with complete classification of all exit sites throughout the year, can be used to produce an annual visitation estimate for any segment of the trail for any time span. An interesting example of this is the application of the model-based estimates to the entire Appalachian Trail, with strata appropriately classified for the year, which yields an annual visitation estimate of 1 948 701, with 95% confidence intervals of 1 172 146 to 2 725 256 and a coefficient of variation of 20%. Most of the visitation (99%) could be attributed to the standard sites because no special events or augmented site types such as HF and ATCH were identified by ATC or NPS trail experts outside of the survey area. Prior to this project, the only available annual visitation estimate was reported to be 3 to 4 million (Appalachian Trail Conservancy 2009; National Park Service 2009). However, there is no documentation for this estimate, and neither NPS staff nor Appalachian Trail Conservancy personnel knew how, or by whom, the estimate was generated. Thus, the reliability of this previous estimate cannot be established.

Although our total trail estimate is subject to criticism, it is one way in a resource constrained environment to arrive at such an estimate without sampling over a wide spatial and temporal range, provided that careful stratification of all sites is performed. Clearly, the annual total trail estimate produced by our methodology is dependent on a number of testable assumptions; however, it is transparent and more scientifically defensible than previous efforts. Hopefully, funding in the future will allow additional sampling across the entire spatial and temporal range of the trail so that this estimate can be revisited and perhaps revised.

Future research

Using the Appalachian Trail as an example, this research has provided and demonstrated a general survey framework for backcountry trails that traverse long distances through diverse forest landscapes that include urban area. The extension of the visitation estimate for the 175 km segment of the trail to the total trail could be improved by conducting a future survey collecting data throughout the total spatial and temporal dimensions of the trail for the entire year. It must be emphasized that expansion of visitor characteristics (gender, race, activity, etc.) from a trail segment to the entire trail as was done for visitation should not be performed because stratification based on site type and use level is likely not appropriate for visitor characteristics. In addition, there are undoubtedly other augmented site types and numerous special events that were not considered that should be identified and utilized to improve the total trail visitation estimate. It is hoped that this survey design will be used on other trails by other researchers who will help to improve and extend the methodology that has been presented here.

Future research should also focus on methods to reduce and better estimate the variances of the visitation estimate and the weighting adjustment factors. The variance of the visitation estimate is a complex product of ratio of means and arithmetic estimators in which independence is assumed. One of the resampling variance estimators such as the bootstrap (Lehtonen and Pahkinen 2004) may be more appropriate under these conditions. The problem of small sample sizes and their effect on the estimates and variances may be alleviated by smoothing alternatives such as generalized variance functions used in small area estimation (Valliant 1987; Wolter 1985). This may be particularly valuable if strata-level estimates are desired. A model-assisted approach (Sarndal et al. 1992) for the weighting adjustment factors may also hold promise if meaningful predictor variables related to weather and local recreational characteristics could be isolated. However, their effects may be accounted for by the site types and use levels that already have been used in the model. In addition, to incorporate other site variables requires additional information on these variables, which may not be available for remote trails that traverse through backcountry forests. Although our approach in this research was only to model the weighting adjustment factors, a Poisson or negative binomial model for the actually daily tally counts may also prove useful (McCullagh and Nelder 1989).

Acknowledgments

This work was supported in part by an Interagency Acquisition Agreement (No. R2490050022) between the National Park Service, the Appalachian National Scenic Trail, and the USDA Forest Service. The authors thank Matt Owens, Allison Ginn, Gary Green, Lincoln Larson, and Donald B.K. English for assistance with various phases of the study.

References

Appalachian Trail Conservancy. 2009. Through hiking facts and statistics [online]. Available from www.appalachiantrail.org/about-the-trail [accessed 8 November 2011].

Benninger-Truax, M., Vankat, J.L., and Schaefer, R.L. 1992. Trail corridors as habitat and conduits for movement of plant species in Rocky Mountain National Park, Colorado. Landsc. Ecol. **6**(4): 269–278. doi:10.1007/BF00129705.

Bergstrom, J.C., Teasley, R.J., Cordell, H.K., Souter, R., and English, D.B.K. 1996. Effects of reservoir aquatic plant management on recreational expenditures and regional economic activity. J. Agric. Appl. Econ. **28**: 409–422.

Bowker, J.M., Bergstrom, J.C., and Gill, J. 2007. Estimating the economic value and impacts of recreation trails: a case study of the Virginia Creeper Rail Trail. Tourism Econ. **13**(2): 241–260. doi:10.5367/000000007780823203.

Cochran, W.G. 1977. Sampling techniques. 3rd ed. John Wiley, New York.

Cole, D.N. 1978. Estimating the susceptibility of wildland vegetation to trailside alteration. J. Appl. Ecol. **15**(1): 281–286. doi:10.2307/2402937.

Dale, D., and Weaver, T. 1974. Trampling effects on vegetation of the trail corridors of north Rocky Mountain forests. J. Appl. Ecol. **11**(2): 767–772. doi:10.2307/2402226.

de Leeuw, E.D. 2005. To mix or not to mix data collection modes in surveys. J. Off. Stat. **21**(2): 233–255.

de Leeuw, E.D., Hox, J.J., and Dillman, D.A. 2008. International handbook of survey methodology. The Psychology Press, London, England.

English, D.B.K., Kocis, S.M., Zarnoch, S.J., and Arnold, J.R. 2002. Forest Service national visitor use monitoring process: research method documentation. USDA Forest Service Gen. Tech. Rep. SRS-57.

English, D.B.K., Kocis, S.M., Arnold, J.R., Zarnoch, S.J., and Warren, L. 2003. The effectiveness of visitation proxy variables in improving recreation use estimates for the USDA Forest Service. J. Nat. Conserv. **11**(4): 332–338. doi:10.1078/1617-1381-00065.

Frantz, J. 2007. The value of hiking: economic benefits of hiking and non-motorized outdoor recreation. Available from the Washington Trails Association and University of Washington Political Science Department at www.wta.org/trail-news/publications/Value-of-Hiking.pdf [accessed 1 March 2011].

Freimund, W.A., and Cole, D.N. 2001. Use density, visitor experience, and limiting recreational use in wilderness: progress to date and research needs. *In* Proceedings "Visitor Use Density and Wilderness Experience", 1–3 June 2000, Missoula, Montana. *Compiled by* W.A. Freimund and D.N. Cole. USDA Forest Service, Rocky Mountain Research Station, Ogden, Utah, RMRS-P-20. pp. 3–8.

Goodman, L.A. 1960. On the exact variance of products. J. Am. Stat. Assoc. **55**(292): 708–713. doi:10.2307/2281592.

Gregoire, T.G., and Buhyoff, G.J. 1999. Sampling and estimating recreational use. USDA Forest Service Gen. Tech. Rep. PNW-456.

Hickman, S. 1990. Evidence of edge species attraction to nature trails within deciduous forest. Nat. Areas J. **10**: 3–5.

Jacobi, C. 2003. Developing a regression estimator for Acadia National Park carriage road use using censuses and electronic trail counters 2001–2002. Acadia National Park Natural Resources Report 2003-1.

Jordan, M. 2000. Ecological impacts of recreational use of trails: a literature review. Available from www.ohv.parks.ca.gov/pages/1324/files/ecological%20impacts%20of%20recreational%20users.pdf [accessed 8 November 2011]. The Nature Conservancy, Cold Spring Harbor, New York.

Kyle, G., Graefe, A., Manning, R., and Bacon, J. 2004. Predictors of behavioral loyalty among hikers along the Appalachian Trail. Leis. Sci. **26**(1): 99–118. doi:10.1080/01490400490272675.

Lehtonen, R., and Pahkinen, E. 2004. Practical methods for design and analysis of surveys. 2nd ed. John Wiley & Sons, London, England.

Lindsey, P., and Lindsey, G. 2004. Using pedestrian count models to estimate urban trail traffic. J. Regional Analysis and Policy, **34**: 51–68.

Loomis, J.B. 2000. Counting on recreation use data: a call for long-term monitoring. J. Leis. Res. **32**: 93–96.

Loomis, J.B., and Walsh, R.G. 1997. Recreation economic decisions: comparing benefits and costs. 2nd ed. Venture Publishing, Inc., State College, Pennsylvania.

Manning, R.E. 1997. Social carrying capacity of parks and outdoor recreation areas. Parks Recreat. **32**(10): 32–38.

Manning, R.E. 2002. How much is too much? Carrying capacity of national parks and protected areas. *In* Monitoring and Management of Visitor Flows in Recreational and Protected Areas. Proceedings of the Conference held at Bodenkultur University, Vienna, Austria, 30 January – 2 February 2002. *Edited by* A. Arnberger, C. Brandenburg, and A. Muhar. Institute for Landscape Architecture and Landscape Management, Bodenkultur University, Vienna, Austria. pp. 306–313.

Manning, R.E., Valliere, W., Bacon, J.J., Graefe, A., Kyle, G., and Hennessy, R. 2000. Use and users of the Appalachian Trail: a source book. Available from www.nps.gov/appa/parkmgmt/upload/Main_Report-2.pdf [accessed 1 March 2011].

McCullagh, P., and Nelder, J. 1989. Generalized linear models. 2nd ed. Chapman and Hall, London, UK.

Miller, S.G., Knight, R.L., and Miller, C.K. 1998. Influence of recreational trails on breeding bird communities. Ecol. Appl. **8**(1): 162–169. doi:10.1890/1051-0761(1998)008[0162:IORTOB]2.0.CO;2.

Muhar, A., Arnberger, A., and Brandenburg, C. 2002. Methods for visitor monitoring in recreational and protected areas: an overview. *In* Monitoring and Management of Visitor Flows in Recreational and Protected Areas. Proceedings of the Conference held at Bodenkultur University, Vienna, Austria, 30 January – 2 February 2002. *Edited by* A. Arnberger, C. Brandenburg, and A. Muhar.

Institute for Landscape Architecture and Landscape Management, Bodenkultur University, Vienna, Austria. pp. 1–6.
National Park Service. 2008. Harpers Ferry NHP reports: visitation by month/year [online]. Available from www.nature.nps.gov/stats/park.cfm [accessed 1 March 2011].
National Park Service. 2009. Appalachian National Scenic Trail: News [online]. Available from www.nps.gov/appa/parknews/index.htm [accessed 1 March 2011].
Sarndal, C.E., Swensson, B., and Wretman, J. 1992. Model assisted survey sampling. Springer, New York.
SAS Institute Inc. 2004. SAS/STAT® 9.1 user's guide. SAS Institute Inc., Cary, North Carolina.
Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. 2nd ed. The Blackburn Press, Caldwell, New Jersey.
Valliant, R. 1987. Generalized variance functions in stratified two-stage sampling. J. Am. Stat. Assoc. **82**(398): 499–508. doi:10.2307/2289453.
Wolter, K. 1985. Introduction to variance estimation. Springer-Verlag, New York.
Zarnoch, S.J., Kocis, S.M., Cordell, H.K., and English, D.B.K. 2002. A pilot sampling design for estimating outdoor recreation site visits on the national forests. USDA Forest Service Res. Pap. SRS-29.