United States
Department of
Agriculture

Forest Service


**Southern
Research Station**

General Technical
Report SRS–57

# Forest Service National Visitor Use Monitoring Process: Research Method Documentation

**Donald B.K. English, Susan M. Kocis, Stanley J. Zarnoch, and J. Ross Arnold**

## The Authors

**Donald B.K. English**, Research Social Scientist, USDA Forest Service, Southern Research Station, 320 Green Street, Athens, GA 30602; **Susan M. Kocis**, NVUM Project Coordinator, USDA Forest Service, 1645 Highway 20 East, Bend, OR 97701; **Stanley J. Zarnoch**, Mathematical Statistician, USDA Forest Service, Southern Research Station, 200 Weaver Blvd., Asheville, NC 28804; and **J. Ross Arnold**, Resource Economist, USDA Forest Service, Strategic Planning and Resource Assessment, Washington, DC 20250.

July 2002

Southern Research Station
P.O.Box 2680
Asheville, NC 28802

# Forest Service National Visitor Use Monitoring Process: Research Method Documentation

**Donald B.K. English, Susan M. Kocis, Stanley J. Zarnoch, and J. Ross Arnold**

**Abstract**

In response to the need for improved information on recreational use of National Forest System lands, the authors have developed a nationwide, systematic monitoring process. This report documents the methods they used in estimating recreational use on an annual basis. The basic unit of measure is exiting volume of visitors from a recreation site on a given day. Sites are stratified by type. Days are stratified by expected volume of exiting recreation visitors. A double sampling strategy is the primary means used to obtain measures of exiting recreation traffic. Where possible, observable counts of other measures highly correlated with visitation, such as fee envelopes, ski lift tickets, or concessionaire reports, are used to reduce variation in visitation estimates. In addition to showing how sampling units were defined, the authors also provide calculations they used in developing estimators for the mean and variance of visitation.

**Keywords:** Monitoring, National Visitor Use Monitoring, NVUM, recreation use, research methods, sampling, wilderness.

## Introduction

### Overview

Both the U.S. Congress and the General Accounting Office have questioned the credibility of recreational visitation estimates reported by the U.S. Department of Agriculture Forest Service (USDA Forest Service). Accurate information about the amount of recreation occurring on National Forest System (NFS) lands is necessary for a variety of reasons, including forest planning and budget allocation; however, methods used to estimate recreational use have been inconsistent across reporting units and often have yielded results of questionable validity. In response to the need for improved information, the USDA Forest Service has begun a new, ongoing process of measuring recreation visitation on the national forests and grasslands. The Agency will incorporate this process into its standard inventory and monitoring efforts.

As a first step, in 1998 the Agency completed a pilot study to test a statistically valid method for estimating visitor use. Building on that effort, a team of research scientists and NFS personnel developed a standard method that could be used nationwide. Implementation and refinement of the method are being accomplished through a partnership of Agency recreation staff, as well as personnel from strategic planning and resource assessment, research and development, ecosystem management, inventory andmonitoring, and the Missoula Technology and Development Center. The method is called National Visitor Use Monitoring (NVUM), and it is designed to provide statistically reliable estimates of recreation visitation on the national forests, national grasslands, and designated wilderness areas.

### What This Process Provides

The NVUM process is designed to provide an estimate of recreation visits. Moreover, it will help ensure USDA Forest Service-wide consistency in data collection and will establish a minimum standard of statistical accuracy. A recreation visit is defined as ". . . one person entering and exiting a national forest, national grassland or designated wilderness area for the purpose of recreation." In a single visit, an individual may participate in any number of recreation activities. Also, a single visit may last only 15 minutes or 15 or more days; it might be one individual visiting only one recreation site, or one individual visiting every recreation site on a given Federal ownership. Counts or estimates of recreation users exiting individual sites are referred to as recreation site visits. The NVUM is used to sample and develop estimates of recreation site visits—an intermediate step in estimating recreation visits.

Using the NVUM, managers can generate visitation estimates for individual national forests or grasslands, USDA Forest Service regions, and for the NFS as a whole. The primary reporting unit is the national forest or national grassland. Recreation visitation estimates will be developed annually for about one-fourth of the reporting units in each region; once the cycle is established, each unit will be resurveyed every 5 years. Total visitation estimates will be made by summing the estimates from each region and the Nation. In the first 3 years of implementation, regional and national estimates will be made by extrapolating data from the surveyed reporting units. In the fourth and subsequent years, regional and national recreation use will be calculated as the sum of the most recent estimate for each

reporting unit in the region. Statistically, visitation estimates for every level of reporting will be within 15 percent of actual visitation, at the 80-percent confidence level.

Collaborative efforts among administrative branches of the Agency will ensure quality control in data collection, sampling design, and statistical accuracy. In addition, the process will incorporate ongoing research results to improve the accuracy of estimates and reduce costs. Consistency among reporting units will be key—all will follow the same national protocol.

In addition to total visitation estimates, to some extent the annual reports will provide a profile of visitors. Descriptions will be averages for the sampled population or percentage distributions across several categories. For example, sample averages for length of stay, number of annual visits to the forest, and party size will be available. Percentage distributions will include proportion of visitors that engaged in different recreation activities, proportion of visitors from various distance zones, and proportion of visitors who used designated wilderness areas. Each reporting unit will obtain an estimate of the number of recreation visits to wilderness areas, the percentage that are overnight visits, and the percentage of visitors who use outfitter or guide services.

### What This Process Does Not Provide

The data collection and reporting processes will not estimate recreation visits to particular sites or ranger districts, nor will they make any description of visitors to a particular site or district. Results will describe the size and composition of the overall recreation visitor population for a national forest or grassland. Descriptive information for particular subgroups of recreation users, e.g., campers, dispersed users, local users, generally will not be available.

Within the NVUM process framework, opportunities for more detailed sampling of particular user groups, or for special data collection for a national forest or grassland, will be limited. A major goal of NVUM is to ensure the consistency of methods used to estimate recreation visits across reporting units. More intensive sampling of particular user types could compromise methodological consistency and increase the difficulty of calculating the statistical accuracy of visitation estimates.

### Opportunities for Special Studies

Special studies that gather in-depth information from or about selected subgroups of recreation visitors (including those engaged in particular activities or using individual recreation sites) will be pursued through other survey efforts, not as a part of the NVUM process. The effects on field personnel workload will be reduced by the narrow focus of sampling required for NVUM. Federal law requires that such surveys be administered within the guidelines of information collections approved by the Office of Management and Budget (OMB). Special surveys must be fully funded by the requesting unit. Technical assistance in developing survey instruments and sampling plans that conform to OMB guidelines is available from the NVUM research team.

## Research Design

The research design for NVUM uses the double sampling technique developed over 35 years ago to measure recreation use on national forests (James 1967, James and Henley 1968, James and Ripley 1963). The first stage of the process involves selecting a stratified random sample of times and locations where recreational visitors can be counted. For each time and location, survey personnel compile traffic counts for a 24-hour period. Concurrently, they conduct interviews of a random sample of visitors to calibrate traffic counts to the number of unique visits. Each reporting unit will have done some presampling to identify its population of recreation sites and the days each is open for public use. Appendix A provides a flow chart that summarizes the prework steps and data collection processes.

### Defining Recreation Sites

It is helpful to categorize a recreation visit with reference both to where and when it occurs. Recreation sites and the days they are open form the population of sampling units. Five strata are used to categorize location types:

1. Day-use developed sites (DUDS) include sites with facilities that meet the INFRA[1] definition development scale for moderate, heavy, or high degrees of modification. Generally, such facilities provide for visitor comfort, convenience, and education opportunities. Within the NVUM framework, sites with facilities that provide only for the safety and health of

[1] INFRA (Infrastructure Application) is the USDA Forest Service's corporate database that tracks the resources and infrastructure on Agency lands as well as the outputs that come from them. Additional information about INFRA data and definitions is available on the web at http://www.fs.fed.us/eng/infomgmt/infra.htm.

visitors are not sufficiently developed to be included in this category. The DUDS include picnic areas, fish viewing sites, fishing sites, interpretive sites, observation sites, playground-park sport sites, ski areas (alpine and Nordic), some wildlife viewing sites, caves, visitor centers, museums, and swimming areas. The DUDS do not include overnight sites. Boat launches, trailheads, and ranger stations that provide only minimal information services are also not included.

2. Overnight-use developed sites (OUDS) meet the INFRA definition for development scales of moderate, heavy, or high degrees of modification. They include campgrounds (family and group), fire lookouts and cabins, hotels, lodges and resorts (both publicly and privately owned), horse camps, organization sites (both publicly and privately owned), and any other overnight-developed sites within USDA Forest Service jurisdiction, whether managed by the Agency or by a concessionaire. Organizational camps are not included in this category, nor are recreational residences; they typically are sampled as part of the general forest area.

3. Wilderness (WILD) includes lands and waters that are part of the National Wilderness Preservation System. Wilderness study areas, research natural areas, or other roadless areas are not included in this category. Interviews with WILD visitors may be conducted at trailheads and other access points.

4. General forest area (GFA) includes all national forest not included in DUDS, OUDS, or WILD categories. Generally, sample points will be at trailheads or on NFS roads where users exit the national forest. These are the portals though which visitors engaging in dispersed activities such as hiking, hunting, and dispersed camping can access undeveloped areas. In some cases, a GFA entry point will be a river, lake, boat harbor, or airport. Sample points may be on both high-speed and low-speed roads, which are managed by the USDA Forest Service or another jurisdiction.

5. On-forest viewing corridors: viewing scenery on USDA Forest Service lands from public roads, ferries, scenic trains, cruise ships, airplanes, trams, or other travel corridors is a popular recreation activity. In many rural communities, money spent by sightseeing tourists is a vital economic force. Managing the viewsheds of travel corridors is an ongoing element of Agency stewardship. In measuring use levels, the USDA Forest Service long has included sightseers, who constitute a significant proportion of recreation visitation. Nonetheless, little of this activity fits the strictest definition of a national forest recreation visit. For example, people traveling on roads not under USDA Forest Service jurisdiction may not actually enter a national forest. Still, recognizing the significance of such use, we have incorporated into the NVUM process a sampling of this stratum to roughly estimate its magnitude. Estimates of the number of people who travel these corridors to view the forest are not included as national forest recreation visitation, but are reported separately. Nor have we made an effort to establish a preassigned level of statistical accuracy for such counts.

This last stratum allows reporting units to identify important travel corridors—beyond those listed in GFA—that pass through or close to NFS lands, but where the speed, layout, jurisdiction, and/or location of the corridor do not provide physical access to the forest. Individuals who stay within these travel corridors are not included in sampling at sites of the four other types. The stratum includes places from which visitors may view scenery, wildlife, and other natural objects while traveling through or near NFS lands. Most sites of this type are travel corridors (usually a road) owned or maintained by some other public agency.

We identified two classes of travel corridor that provide opportunities for viewing forest scenery. For each, we developed criteria to define characteristics of the travel route and of the people who might use them to view the forest landscape. We selected the criteria to ensure consistency across reporting units in categorizing travel ways and travelers.

**Class 1—**Highways, roads, or other public conveyances designed for viewing scenery that are within NFS ownership but not under NFS jurisdiction. These include interstate, State, or local highways, as well as rivers or other waterways. All air travel routes are excluded. Corridors must be: (a) displayed on secondary, series-base maps (forest visitor maps) as directly traversing continuous USDA Forest Service ownership; and (b) of sufficient length to require at least 15 continuous minutes of travel across USDA Forest Service ownership at normal speed.

Visitors must be: (a) on a recreational trip, or (b) traveling for another purpose, but specifically state that viewing forest scenery was the main reason for choosing that particular travel route.

It is problematic to get an unbiased estimate of the proportion of people using a scenic corridor for the purpose of viewing forest landscapes. Rest areas and scenic view

sites are suitable and safe locations to conduct interviews, although such locations invite a biased sample because the interviewer must select vehicles and individuals to be interviewed. Nonetheless, because stopping people at random on high-speed or busy roads is unsafe for interviewer and traveler alike, we chose interview locations that were safe and would introduce a minimum amount of bias.

**Class 2—**Routes over which commercial trips or tours are taken to view national forest scenery, but where individuals do not enter onto routes defined as class 1. This class provides the visitor who has made a conscious decision to pay for commercial services an opportunity to view scenery on USDA Forest Service lands. Reporting units document the number of passengers on such trips or tours, but do not conduct visitor surveys. Trips or tours include scenic boat rides on waters adjacent to NFS ownership, scenic overflights, and scenic bus, train, or tram travel.

**Site Days**

Some developed sites, trailheads, designated wilderness entry points, forest roads, or other access are open and available for public recreation use year round; others are open only seasonally. Recreation use can be measured any day a site is open. A site day describes the spatial-temporal combination of one location open for 1 calendar day.

Any recreation visit to any site necessarily involves the individual entering the site, engaging in one or more recreation activities, and exiting the site. It is important to count each visitor only once per site visit. The risk of double counting is reduced by making counts either when visitors first enter the site, or at their final exit. If the goal is simply a visitor count, either will suffice. However, James (1967) showed that better information may be collected by surveying visitors as they leave. Exiting visitors are able to provide information about length of stay, facilities used, and activities; those just beginning their visit may be less willing to be interviewed. For these reasons, we count and interview visitors as they exit recreation sites and areas.

**Basic Data Sampling Process**

On each randomly selected sample day, the reporting unit collects two types of data. Personnel first take a 24-hour count of exiting visitor traffic, usually using a mechanical counter; however, where available they may use other routinely collected data, such as ski lift tickets, mandatory permits, or fee envelopes. During the same 24-hour sample period, they conduct 6 hours of on-site interviews. On-site interviews occur during one of two randomly selected 6-hour interview periods during the sample day.

Interviews are done with a sample of exiting visitors, but full interviews are administered only to visitors who recreated at the site and are leaving for the last time that day. By identifying the proportion of last-exiting recreation visitors to total exit traffic, the reporting unit can calibrate the 24-hour count to estimate the number of unique recreation visits that occurred.

Interview protocol for on-forest viewing corridors is slightly different. Installing 24-hour traffic counters on some travel routes, especially major Federal and State highways, may be unsafe. It may be most practical to make manual traffic counts using a hand-tally recorder. Interview locations are moved to an adjacent rest area or similar off-road site. Due to the difficulty of simultaneously interviewing and conducting a hand tally, personnel are encouraged to alternate duties according to a specified schedule. The person conducting the hand tally counts all traffic on the side of the roadway with the rest area and records commercial and private vehicles to obtain a proportional sample.

**Proxy (External) Information**

Information collected while operating some recreation sites is closely related to the amount of recreation visitation. Although such information is often collected for an entire use season or year, it is possible to have it for only one or a few days. If the information meets certain criteria, it can serve as a proxy for the amount of site visitation (Yuan and others 1995). Site days for which such information exists are called proxy site days.

Incorporating proxy information should improve the accuracy of site-visitation estimates and reduce the error of total visitation estimates at the national forest level. Ultimately, there is a reduction in the amount of sampling necessary to reach targeted accuracy levels. Some visitor sampling is needed to estimate conversion rates from the observed proxy measure to the desired visitation units. However, proxy site days can be sampled at a lower rate, because conversion coefficients are more easily estimated.

Several criteria are used to determine the feasibility of visitation proxy information. First, the information must represent all users of the site. Proxy data that pertain only to a particular segment of users, e.g., number of visitors using outfitters, but not individuals who do not use outfitters, is not acceptable. Second, the proxy count must be an exact tally; it cannot be an estimate. Third, only a few types of proxy information are acceptable: fee receipts, fee envelopes, mandatory permits, permanent traffic counters, and ticket sales. Fee sites or sites where a

fee-envelope system is used primarily include campgrounds, ski areas, fee demo sites, and some other day-use recreation sites. Some designated wilderness, backcountry areas, and rivers require permits for all users. Such areas are acceptable for gathering proxy information. Voluntary permit and trail register systems are not. Using these criteria, a reporting unit can determine which proxy information can be used.

**Use-Level Strata**

For either proxy or standard sites, it is possible to enumerate all site days in which recreation could occur. Days on which a site is closed or recreation use is expected to be zero are classified as closed/zero days. Such days are not sampled. The remaining site days are stratified to reduce the variance of estimated annual visitation. Previous USDA Forest Service research stratified sites by expected annual use and by weekday versus weekend or holiday day types (Gregoire and Buyhoff 1999, James and Henley 1968, James and Rich 1966, Lucas and others 1971, Yuan and others 1995).

However, analysis of a pilot study indicated that exit volume for many trailheads and overnight sites was near zero on some Saturdays, but quite high on Sundays and Monday holidays. This resulted in high variance in exiting recreation volume for weekend strata. We felt that more homogeneous strata could be developed by asking reporting units to stratify site days by the expected level of exiting visitor traffic, relative to all site days in that site type. Stratification of day-use sites results from identifying site days that have the highest and lowest level of last-exiting recreation traffic. We stratify site days in each site-type stratum into four classes—high, medium, low, and zero/closed exit volume—in order to most efficiently use the limited number of available sampling days. This divides site days into classes that minimize exit-volume differences within a class, and maximize differences across classes.

**Sample Allocation**

Units involved in the first year of sampling had an average of 64,000 site days. Based on results from the 1998 pilot study, an average of about 200 sampling days per reporting unit (more on those with larger populations of site days, less on those with smaller populations) are needed to obtain the target level visitation-estimate accuracy. Each USDA Forest Service region is allotted 200 sampling days per surveyed reporting unit. Within regions, sampling days are allocated to reporting units in a series of stages. The allocation of days to a reporting unit is a three-step process:

Step 1. Assign to each reporting unit 8 sample days[2] in on-forest viewing corridors, 3 each for high- and medium-use days, and 2 for low-use days.

Step 2. Each reporting unit is allotted no more than 50 sample days for proxy site days.

Allocations within a reporting unit are made as follows:

a. Initial allocation of 4 sample days for each proxy sample cell (defined by a combination of site type and proxy type).

b. If any of the 50 days remain, allocate a fifth sample day to each sample cell with more than three different sites.

c. If any of the 50 days still remain, assign a sixth sample day to cells with more than five different sites.

d. If any of the 50 days still remain, assign a seventh sample day to cells with eight or more different sites.

e. If any of the 50 days still remain, assign an eighth sample day to cells with ten or more different sites.

f. Return any unallocated proxy sample days to the unobligated regional total.

Step 3. Each reporting unit will have up to 12 sample cells (combinations of site type and use level) for site days without proxy information. Sample days are allocated among cells within a region according to the following rules:

a. Initial minimum allocation of 8 sample days per nonproxy sample cell for each reporting unit, to ensure that visitation variance for each cell is computed on at least eight observations (unless the cell has fewer than 8 site days).

b. Any sample days still available for the region are allocated across all nonproxy cells in all survey units. Allocation proportions are determined by the number of site days in each cell (provided in appendix A); weighted by the product of: (1) the cell's estimated standard error from the previous year's survey, and (2) a factor that reflects the relative importance of each use-level stratum in visitation estimates. The

[2] In NVUM's first year, each reporting unit was allocated 5 sample days for this site.

importance weight factors are: high-20, medium-10, and low-1.

That is, high-use site days (regardless of site type) have a weight 20 times that of low-use site days, and twice that of medium-use site days. (Note: In NVUM's first 2 years, no standard errors estimates will be available, so weights equal the importance factor alone.)

**Selection of Interviewing Days and Times**

The set of days for sampling visitation is drawn at random for each stratum, but with a small adjustment. We encountered a logistical problem when the same calendar day was selected for more sites than survey crews could cover. To avoid that problem, we take the following steps in developing a sampling calendar:

1. Group by calendar day the site days selected for sampling.

2. Identify calendar days that have more than 3 site days scheduled; select 2 at random and retain them in the sample.

3. Determine the number of site days in each proxy and nonproxy stratum that need to be replaced.

4. Draw replacements at random from the set of unused site days on calendar days that have fewer than 2 site days already selected for sampling.

Previous research (Yuan and others 1995) indicates that the percent of visitors exiting a site for the last time varies by time of day. To ensure unbiased estimates of the volume of last-exiting recreation traffic, we sample over as much of the day as is practical. For nonproxy sample days, one of two interview periods is selected at random with equal probability. The a.m. survey period begins at 08:00 and concludes at 14:00; the p.m. period runs from 14:00 to 20:00. From late fall to early spring, schedules are adjusted to ensure that interviews are completed before dark. For example, if the sun sets between 17:00 and 17:30, an appropriate afternoon interview period may be 11:00 to 17:00.

For proxy site days, a particular time period is not identified. The purpose of surveying is to obtain information to convert proxy counts to site-visit estimates. Interviews are conducted during the 6-hour daylight period with the greatest level of exiting recreation traffic.

**Closed Sites**

Unpredictable weather, precipitation patterns, fires, or other natural phenomena may change the dates a site is open or closed, and some changes may continue indefinitely. Unfortunately, no ex-post adjustment for such changes really is possible. Reporting units must try to determine a priori a site's most likely opening and closing dates. The set of open site days is defined before the survey year begins, and it is to that set that the visitation estimate applies. For all days that a site will be closed due to reconstruction, restoration, or any other reason, that site should be listed as closed.

Sites may be closed administratively, in whole or part, for unforeseen reasons such as fire, flood, heavy snowfall; construction (including unanticipated repairs); or resource protection (too little snow, wildlife protection). In all such cases, the site will be shown as open but with below normal visitation. If a site is partially closed, interviewing should continue as scheduled but below normal visitation should be noted in the daily summary form. If a site has been closed completely and unexpectedly, the interviewer should record zero traffic counts on the daily summary form. Conversely, sites that are open when they were expected to be closed are treated as closed.

**Selecting Interviewees**

Interviewers will conduct as many surveys as possible, but normally fewer than 60 interviews per day. It is important, nonetheless, to spread out interviews over an entire sampling day. For example, at a busy developed site where many interviews could be conducted quickly, interviewers should time them to cover the whole sampling period. To determine the individual in a vehicle or group to survey, use a random selection process, e.g., the person over age 15 with the most recent birthday.

**Survey Forms**

Each interviewee is asked basic questions from a survey form. One or two additional questions are asked at proxy sites to convert proxy counts to site-visit estimates. One-fourth of the sample is asked a set of questions about economic benefits and trip-related spending. Another one-fourth is asked questions about satisfaction with recreation services and facilities.

## Data Analysis

### Part I—Daily Site Visits at Nonproxy Site Days

Although traffic counters tally either exiting vehicles or exiting persons, the method for estimating daily site visits is essentially the same for both. Most site-visit estimators follow standard formulae for stratified random samples (Cochran 1977). For a given stratum $h$ ($h$ = 1, 2, 3, . . . , H) and sampled site day $i$ ( $i$ =1, 2, 3, . . . , $n_h$ ), within stratum $h$ let:

$C_{hi}$ = total car (or person) count (obtained from traffic counter) during the 24-hour sampling period for day $i$ in stratum $h$,

$V_{hij}$ = number of persons in the $j^{th}$ ( $j$ = 1, 2, 3, . . . $J$) sampled vehicle on site-day $i$ [obtained from the onsite questionnaire (note that when traffic counts are of exiting visitors, $V_{hij}$ = 1 is a constant)],

$LR_{hij}$ = indicator variable in the onsite questionnaire, coded as:

= 1 if the $j^{th}$ vehicle sampled on site-day $i$ is a last-exiting recreation vehicle,

= 0 otherwise.

The proportion of vehicles on site day $i$ that were last exiting $(P_{hi})$ is:

$$P_{hi} = \sum_{j=1}^{J} \frac{LR_{hij}}{J},$$

The mean persons per recreation vehicle for last-exiting recreation vehicles $(V_{hi})$ is:

$$V_{hi} = \frac{\sum_{j=1}^{J} LR_{hij} V_{hij}}{\sum_{j=1}^{J} LR_{hij}}.$$

Therefore, the estimate for total exiting site visits $(SV_{hi})$ on site day $i$ is:

$$SV_{hi} = C_{hi}\ P_{hi}\ V_{hi}\ .$$

An estimate of mean daily site visits for stratum $h$ is simply:

$$\overline{SV}_h = \sum_{i=1}^{n_h} \frac{SV_{hi}}{n_h},$$

with estimated variance

$$V(\overline{SV}_h) = \sum_{i=1}^{n_h} \frac{(SV_{hi} - \overline{SV}_h)^2}{n_h(n_h - 1)}.$$

This formula implies a simplifying assumption. The sampling frame for nonproxy days is actually a two-stage design. The first stage is a random sample of site days, and the second is a random sample of visitors within each first-stage sampling unit. However, the first-stage sampling rates are relatively small. Preliminary information from the presampling work for reporting units surveyed in the first year indicates that across all sampled forests, the average first-stage sampling rate was about 0.25 percent. Hence, the second term in the sample variance equation for a two-stage sample (Cochran 1977, p. 278) will be negligible and could be eliminated, yielding the above formula.

Expansion to the stratum total $(SV_h)$ is:

$$SV_h = N_h \overline{SV}_h,$$

with estimated variance

$$V(SV_h) = N_h^2 V(\overline{SV}_h),$$

where

$N_h$ = total number of site days in stratum $h$.

### Part II—Proxy Site-Visit Estimates

Estimation of recreation site visits for the proxy component differs from that of the nonproxy in several ways. First, strata are defined differently. For a given site type, there may be several different types of proxy information. Each unique combination of site type and proxy type constitutes a stratum. Second, a component of the proxy site-visit estimation equation is not based on a sampling survey but is obtained by direct census (a count known without error) of the proxy count, which yields a reduction in estimator variance. Here, the purpose of sampling is to obtain the information needed to convert proxy counts to site visits.

Let $P_{hk}$ be the annual total proxy count for site $k$ in stratum $h$, and let $CR_{hk}$ be the known compliance rate of visitors with respect to the proxy count at that site. For example, not all campground users may pay the required fee, or not all designated wilderness users may obtain mandatory permits. Then, the compliance adjusted proxy count for site $k$ $(PC_{hk})$ is the proxy count that would have been observed with 100 percent compliance, and

$$PC_{hk} = \frac{P_{hk}}{CR_{hk}} .$$

For a given stratum *h* the mean daily proxy count is

$$\overline{PC}_h = \frac{\sum_{k=1}^{K} PC_{hk}}{\sum_{k=1}^{K} N_{hk}},$$

where

$N_{hk}$ = number of site days whose use is represented by proxy count for site *k* in stratum *h*.

Conversion coefficients are needed to obtain site-visit estimates from proxy counts. Coefficients for each site type-proxy type stratum are obtained by sampling individuals on randomly selected site days within the stratum. Two different variables are needed to make the conversion. Consider a campground that collects fee envelopes from campers. Payment is required for each night a campsite is occupied. However, a person, group, or family camping for a week can pay in one envelope (one proxy) or as many as one envelope per day (seven proxies). The first conversion variable (*R*) measures the proxy count per group recreation visit. The second conversion variable (*G*) measures the number of people per group recreation visit. For example, up to five people are allowed to use the same campsite. Let

$R_{hij}$ = number of proxies per group recreation site visit for the $j^{th}$ group surveyed in sample day *i* in stratum *h*, $\left(R_{hij} \geq 1.0\right)$,

$G_{hij}$ = group size for the $j^{th}$ surveyed group on the $i^{th}$ sample day in stratum *h*.

Because individual surveys are clustered within sample days, the conversion coefficient for each stratum is calculated using a ratio of means approach. Let the sum of $R_{hij}$ and $G_{hij}$ over all surveys on the $i^{th}$ day be $SR_{hi}$ and $SG_{hi}$, respectively. Then the estimate of the proxy conversion coefficient $\left(A_h\right)$ for stratum *h* is

$$\overline{A}_h = \frac{\sum_{i=1}^{n_h} SG_{hi}}{\sum_{i=1}^{n_h} SR_{hi}},$$

where

$n_h$ = number of days sampled in stratum *h*, and which has a variance of

$$V(\overline{A}_h) = \frac{1}{(n_h)(n_h - 1)(\sum_i SR_{hi}/n_h)^2}\left[\sum_i SG_{hi}^2 + \overline{A}_h^2 \sum_i SR_{hi}^2 - 2\overline{A}_h \sum_i (SG_{hi} SR_{hi})\right],$$

Then the estimator for mean daily site visits in stratum *h* is

$$\overline{SV_h} = \overline{A_h}\ \overline{PC_h} .$$

with a variance of

$$V(\overline{SV}_h) = \overline{PC}_h^2\ V(\overline{A}_h) .$$

An exception is if the proxy count is a permanent traffic counter that may count nonrecreation use as well as recreation users, such as along a scenic roadway or at a day-use site where people may enter just to use the bathroom. In these cases, the mean daily proxy count is adjusted by the proportion of all of the individuals surveyed on site days in that stratum who were recreating at the site $\left(\overline{RP}_h\right)$. Here, the mean daily site-visit estimate is given by

$$\overline{SV}_h = \overline{A}_h \overline{PC}_h \overline{RP}_h$$

and its variance by

$$V(\overline{SV}_h) = \overline{RP}_h^2\, \overline{PC}_h^2 V(\overline{A}_h) + \overline{A}_h^2\, \overline{PC}_h^2 V(\overline{RP}_h) - \overline{PC}_h^2 V(\overline{A}_h) V(\overline{RP}_h) .$$

Expansion to total site visits for stratum *h* is

$$SV_h = N_h\, \overline{SV}_h ,$$

where

$N_h$ = total number of site days in stratum *h*. The estimated variance for total site visits is

$$V(SV_h) = N_h^2 V(\overline{SV}_h) ,$$

**Part III—Expansion to Forest Population Mean and Total**

To estimate mean daily site visits for the entire population of site days over all strata, combining both proxy and nonproxy, let strata weights be defined as

$$W_h = \frac{N_h}{\sum_{h=1}^{H} N_h},$$

then the mean daily site-visit estimator is

$$\overline{SV} = \sum_{h=1}^{H} W_h \, \overline{SV}_h \, ,$$

with estimated variance

$$V(\overline{SV}) = \sum_{h=1}^{H} W_h^2 \, V\left(\overline{SV}_h\right) .$$

Site-visit estimate for the total population is

$$SV = N \, \overline{SV} \, ,$$

where

N = total number of site days in the reporting unit's population, and the estimated variance is

$$V(SV) = N^2 V(\overline{SV}) .$$

**Part IV—Estimating National Forest Visits**

We based the first NVUM survey for a national forest on a stratified random sampling design of site days, with strata defined by site type and daily exit volume, i.e., DUDS low. Our objective was to estimate mean daily site visits and total annual site visits from a sample of site days randomly selected in each stratum. Site-visit estimates were obtained for each sample day, averaged by strata, and then expanded according to classical stratified random sampling methodology.

However, a primary reporting goal of NVUM is number of national forest visits (NFV). Because any single national forest visitation may include a variety of site visits, we recognized that recreation visitors might move from one to another sampling unit when visiting multiple sites. A person who goes to two sites in a single NFV might be contacted and correctly included in the last-exiting sample for either or both, even if the sites are of different site types and/or exit-volume levels. Conversely, a person visiting just one site might be interviewed only when last exiting that site. Thus, those who visit more sites per national forest will be overrepresented in the sample, leading to an upward bias in the estimator for mean number of site visits per NFV, and a downward bias in number of NFVs. This type of problem is uncommon in most classical sampling situations, but a related issue has been identified in some onsite visitor samples (Shaw 1988).

We used the following approach to reduce problems associated with estimating NFVs. Consider that each national forest visitor has a coupon that was distributed proportionately to each site visited on that trip. For instance, if three sites are visited on an NFV trip, then each site gets one-third of the coupon. The true NFV number is the total number of coupons on all site days in the forest. For a given site day, the estimator of NFV is defined mathematically as

$$\hat{N}FV_i = P_i * CARS_i * CBAR_i \, ,$$

where

$P_i$ = proportion of vehicles that were last-exiting recreationists on site day *i*,

$CARS_i$ = number of vehicles obtained from the car counter adjusted for axles and one- or two-way traffic on site day *i*, and

$CBAR_i$ = average number of coupons per last-exiting recreating vehicle on site day *i* (remembering to use proportions of a coupon if multiple sites were visited).

To further clarify the meaning of *CBAR*, let

*n* = number of last-exiting recreating vehicles interviewed,

$PEOPLE_j$ = number of people in last-exiting recreating vehicle *j*, and

$SVPNFV_j$ = number of sites visited on this NFV for the people in vehicle *j*

then *CBAR* is defined as

$$CBAR = \frac{1}{n} \sum_{j=1}^{n} \frac{PEOPLE_j}{SVPNFV_j} .$$

Note that the *NFV* estimator is identical to the *SV* estimator, except that *CBAR* is used instead of average people per vehicle. Justification for this approach for estimating *NFV* is verified by a couple of examples. If all visitors only go to one site on their NFV, then *SVPNFV* = 1 and *CBAR* reduces to the average people per car and *NFV* = *SV* as is expected. Alternatively, if all people go to two sites on their NFV, then *SVPNFV* = 2 and *CBAR* equals one-half the average people per car and *NFV* = 0.5 *SV* as expected. Logical extensions to other scenarios where the number of sites visited per visitor varies are more complicated but should be obvious. Because the NFV estimator is identical to the *SV* estimator except for *CBAR*, all the statistical methodology previously explained for the *SV* estimator is appropriate for the *NFV* estimator. In particular, the $NFV_i$'s are used to calculate strata means and variances that are then expanded to forest level estimates.

This approach can be extended to the proxy portion of the *NFV* estimator by a simple modification of the proxy conversion coefficient $A_h$ for stratum *h* defined previously as

$$\overline{A}_h = \frac{\sum_{i=1}^{n_h} SG_{hi}}{\sum_{i=1}^{n_h} SR_{hi}}.$$

For the *NFV* estimator, $SG_{hi}$ must be replaced with $SC_{hi}$ which is defined as

$$SC_{hi} = \sum_{j=1}^{m} \left( \frac{GRPSIZE_{hij}}{SVPNFV_{hij}} \right).$$

All other statistical methodologies follow as outlined previously for the proxy situation.

**Part V—Expansion to Regional and National Estimates**

All estimates at the reporting unit level, whether totals or means, have been based on a stratified random sampling design. To calculate regional-level estimates, these estimates are folded into a two-stage sampling design from which the appropriate estimators and their variances are obtained. National totals and their variances are simply the sum of all of the regional totals and their variances.

An unbiased estimator for regional totals of a variable of interest, such as total number of site days, is based on Cochran (1977, equation 11.21) and is defined as

$$\hat{Y}_u = \frac{N}{n} \sum_{i=1}^{n} M_i \overline{y}_i,$$

where

*N* = total number of reporting units in the region,

*n* = number of reporting units sampled in the region,

$M_i$ = total number of site days in reporting unit *i*, and

$\overline{y}_i$ = mean estimate per site day for reporting unit *i* based on stratified random sampling.

Due to stratified random sampling, an estimate of the variance is obtained as a slight modification of Cochran (1977, equation 11.24), defined as

$$V(\hat{Y}_u) = \frac{N^2(1-f)}{n} \frac{\sum_{i=1}^{n} (\hat{Y}_i - \overline{Y}_u)^2}{n-1} + \frac{N}{n} \sum_{i=1}^{n} M_i^2 s_{2i}^2,$$

where

$\hat{Y}_i$ = estimate of the total for reporting unit *i*,

$\overline{Y}_u$ = average total estimate for a reporting unit,

$f = (1 - n/N)$ = finite population correction for reporting units, and

$s_{2i}^2$ = variance of the mean estimate based on stratified random sampling.

Note that, for simplicity, the finite population correction at the second-stage level is ignored, because the number of site days sampled within a reporting unit is negligible compared to the total number of site days in the unit.

The unbiased ratio-to-size estimator from Cochran (1977, equation 11.25) is used to estimate regional means after slight modification and is defined as

$$\overline{Y}_R = \frac{\sum_{i=1}^{n} M_i \overline{y}_i}{\sum_{i=1}^{n} M_i},$$

with estimated variance after modification from Cochran (1977, equation 11.30) as

$$V(\overline{Y}_R) = \frac{1}{M_0^2} \left( \frac{N^2(1-f)}{n} \sum_{i=1}^{n} M_i^2 \frac{(\overline{y}_i - \overline{Y}_R)^2}{n-1} + \frac{N}{n} \sum_{i=1}^{n} M_i^2 s_{2i}^2 \right),$$

where

$$M_0 = N \sum_{i=1}^{n} \frac{M_i}{n}.$$

Again, for simplicity, the finite population correction at the second-stage level is ignored.

**Part VI—Estimators for Visitor Characteristics**

The NVUM survey and reporting process also provides three types of estimates for various visitor characteristics that are important to managers. They are: (1) total estimates, such as total number of NFVs by children under 16; (2) mean estimates, such as mean trip or site-visit duration time; and (3) proportion estimates, such as proportion of visitors who camp. In addition, each of these types could be defined on one of three scales: (a) site visit, (b) NFV, or (c) annual basis. For example, an estimator may be developed for the proportion of visitors camping sometime during their NFV, or for how often the average

visitor visits the national forest per year. Thus, with three estimator types and three scale levels, there are nine different estimators for a specific visitor characteristic. The most meaningful estimator will depend on the question asked and estimator desired.

All total estimates follow the previously defined methodology for SV and NFV estimates except when making appropriate, minor changes to reflect the variable of interest. All mean and proportion estimates are developed as ratios of two total estimates. The total estimates are defined on a site-day basis as

$$\hat{T}OTAL_i = P_i * CARS_i * \overline{X}_i ,$$

where

$P_i$ = proportion of vehicles that were last exiting on site-day *i*,

$CARS_i$ = number of vehicles tallied by a counter, adjusted for axles and one- or two-way traffic on site-day *i*,

$X_{ij}$ = a variable dependent on the scale and level of the estimator for group *j* in site-day *i*

$= Z_1 / (Z_2 Z_3)$, and

$\overline{X}_i$ = average of the $X_{ij}$ over all groups.

The $X_{ij}$ is straightforwardly defined. The $Z_1$ is the variable of interest. The $Z_2$ and $Z_3$ variables are defined so as to adjust $Z_1$ to the appropriate scale. Appendix B lists some examples of common estimators that clarify how the estimators are computed for different scales and types. Given that:

*SVPNFV* = number of site visits per NFV for vehicle *j*, and

*NFVPY* = number of NFVs per year for vehicle *j*,

then, for

| | | |
|---|---|---|
| *SV* scale estimators | $Z_2 = 1$ | $Z_3 = 1$ , |
| *NFV* scale estimators | $Z_2 = SVPNFV$ | $Z_3 = 1$ , |
| annual scale estimators | $Z_2 = SVPNFV$ | $Z_3 = NFVPY$ . |

Defining a total estimator in this manner on a site-day basis allows application of the same statistical methodology as previously described for the *SV* and *NFV* estimators.

## Acknowledgments

We thank those who provided review comments on earlier versions of this report. In alphabetical order, they are: Deborah Chavez (U.S. Department of Agriculture Forest Service), Tim Green (U.S. Department of Agriculture Forest Service), Troy Hall (Virginia Polytechnic Institute and State University), Scott Jackson (U.S. Army Corps of Engineers), Dennis Propst (Michigan State University), Chris Siderelis (North Carolina State University), and Jo Tynon (Oregon State University). In addition, there were a number of U.S. Department of Agriculture Forest Service personnel who provided significant contributions to our method, including (also alphabetically) Jim Bedwell, Rich Calnan, Ken Cordell, Dave Hackett, Margaret Lincoln, Mike Noland, Stan Specht, Jerry Stokes, Greg Super, Francisco Valenzuela, and Larry Warren.

## Literature Cited

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

**Gregoire, T.G.; Buyhoff, G. J.** 1999. Sampling and estimating recreation use. GTR-PNW-456. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 39 p.

**James, G.A.** 1967. Recreation use estimation on Forest Service lands in the United States. Res. Note SE-79. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 8 p.

**James, G.A.; Henley, R.K.** 1968. Sampling procedures for estimating mass and dispersed recreation use on large areas. Res. Pap. SE-31. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 15 p.

**James, G.A.; Rich, J.L.** 1966. Estimating recreation use on a complex of developed sites. Res. Note SE-64. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 8 p.

**James, G.A.; Ripley, T.H.** 1963. Instructions for using traffic counters to estimate recreation visits and use. Res. Pap. SE-3. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 12 p.

**Lucas, R.C.; Schreuder, H.T.; James, G.A.** 1971. Wilderness use estimation: a pilot test of sampling procedures in the Mission Mountains Primitive Area. Res. Pap. 109. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 44 p.

**Mood, A.M.; Graybill, F.A.; Boes, D.C.** 1974. Introduction to the theory of statistics. 3rd ed. New York: McGraw-Hill.

**Shaw, Daigee.** 1988. On-site samples' regression: problems of non-negative integers, truncation, and endogenous stratification. Journal of Econometrics. 37: 211-223.

**Yuan, S.; Maiorano, B.; Yuan, M. [and others].** 1995. Techniques and equipment for gathering visitor use data on recreation sites. Publ. 9523-2838-MTDC. U.S. Department of Agriculture, Forest Service, Technology and Development Program. 78 p.

# Appendix A—NVUM process flowchart

**Phase I—Prework**

Define forest's recreation sites
Classify into site types (DUDS, OUDS, Wilderness, GFA)
[Forest, regional coordinator][1]

↓

Identify dates each site is expected to be
Open Closed → Site visits = 0
[Forest] [1]

↓

Determine if proxy information is available for each open site day
[Forest] [1]

| ↓ | ↓ |
|---|---|
| No<br>(Nonproxy site days) | Yes<br>(Proxy site days) |
| Stratify days in each site type by expected exit volume:<br>High<br>Medium<br>Low<br>[Forest] [1] | Stratify days in each site type by expected exit volume:<br>High<br>Medium<br>Low<br>[Forest] [1] |
| Complete spreadsheet:<br>Site ID, site type, date, use class<br>[Forest] [1] | Complete spreadsheet:<br>Site ID, site type, date, use class, proxy type, proxy units<br>[Forest] [1] |
| Determine sampling schedule<br>[National team] [1] | Determine sampling schedule<br>[National team] [1] |

Train, schedule, interview staff
[Forest, regional coordinator] [1]

**Phase II—Survey Process (For Forest Staff)**

**Nonproxy site days**

One day prior to scheduled sampling, determine if the site will be open
on the sampling day or closed due to unusual circumstances, e.g. late snow

↓ Yes ↓ No → Record 0 visits and do not attempt interviews

↓

Set up traffic counter to obtain daily count

On scheduled sampling day:
- Conduct as many exit interviews as is practical for the assigned time (one interview per group)
- Retrieve traffic counter at day's end
- Record traffic count on sampling spreadsheet

One day after scheduled sampling:
- Check interview forms for completeness
- Mail completed forms to regional coordinator

**Proxy Site Days**

For each day during the interview year, record the proxy count for sites

One day prior to scheduled sampling, determine if the site will be open
on the sampling day or closed due to unusual circumstances, e.g. late snow

Yes No → Record 0 visits and do not attempt interviews

On scheduled sampling days:
- Obtain compliance check for the proxy count to determine what percent of users are measured by the proxy
- Conduct exit interviews for the assigned time frame

One Day after scheduled sampling:
- Check interview forms for completeness
- Mail completed forms to regional coordinator

[1] Brackets indicate who is primarily responsible for accomplishing the task.

**Appendix B—Example of the formulation of several estimators based on the original site visit estimator SV[1]**

| Y | Estimator description | Estimator formulation[2] | Type | Scale |
|---|---|---|---|---|
| $Y_1$ | Total number of annual site visits | X = PEOPLE | Total | SV |
| $Y_2$ | Total number of annual NFVs | X = PEOPLE/SVPNFV | Total | NFV |
| $Y_3$ | Total number of different visitors to the national forest on an annual basis | X = PEOPLE/(SVPNFV*NFVPY) | Total | Annual |
| $Y_4$ | Average number of NFVs per visitor per year | $Y_2 / Y_3$ | Average | Annual |
| $Y_5$ | Total number of annual NFVs by children under 16 | X = NKIDS/SVPNFV | Total | NFV |
| $Y_6$ | Proportion of NFVs by children under 16 | $Y_5 / Y_2$ | Proportion | NFV |
| $Y_7$ | Total number of NFVs where camping was an activity | X = (PEOPLE*CAMP)/SVPNFV | Total | NFV |
| $Y_8$ | Proportion of NFVs where camping was an activity | $Y_7 / Y_2$ | Proportion | NFV |
| $Y_9$ | Total number of group trips to the national forest on an annual basis | X = 1/SVPNFV | Total | NFV |
| $Y_{10}$ | Sum of the ages of one person per group for each group trip to the national forest | X = (1*AGE)/SVPNFV | Total | NFV |
| $Y_{11}$ | Average age of a NFV | $Y_{11} / Y_9$ | Average | NFV |

PEOPLE = the average number of people in a last-exiting recreating car; SVPNFV = number of site visits per national forest visit; NFVPY = number of national forest visits per year; NKIDS = number of people under the age of 16 years in a last-exiting recreating car; CAMP = an indicator variable that is equal to 1 if camping was an activity on the national forest trip and 0 if it was not; AGE = age of one person selected at random from a last-exiting recreating car.

[1] All estimators are based on the original site visit (SV) estimator defined as SV = P*CARS*X, where P = proportion of vehicles that were last-exiting recreating vehicles, CARS = number of cars exiting the site during the 24-hour period, and X = average number of people in a last-exiting recreating vehicle.

[2] The estimator is formulated from either the original SV estimator by substituting the specified X into the original SV estimator or as the ratio of the two previously defined estimators.

**English, Donald B.K.; Kocis, Susan M.; Zarnoch, Stanley J.; Arnold, J. Ross.** 2002. Forest Service National Visitor Use Monitoring Process: Research Method Documentation. Resch. Pap. SRS–57. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 14 p.

In response to the need for improved information on recreational use of National Forest System lands, the authors have developed a nationwide, systematic monitoring process. This report documents the methods they used in estimating recreational use on an annual basis. The basic unit of measure is exiting volume of visitors from a recreation site on a given day. Sites are stratified by type. Days are stratified by expected volume of exiting recreation visitors. A double sampling strategy is the primary means used to obtain measures of exiting recreation traffic. Where possible, observable counts of other measures highly correlated with visitation, such as fee envelopes, ski lift tickets, or concessionaire reports, are used to reduce variation in visitation estimates. In addition to showing how sampling units were defined, the authors also provide calculations they used in developing estimators for the mean and variance of visitation.

**Keywords:** Monitoring, National Visitor Use Monitoring, NVUM, recreation use, research methods, sampling, wilderness


The Forest Service, United States Department of Agriculture (USDA), is dedicated to the principle of multiple use management of the Nation's forest resources for sustained yields of wood, water, forage, wildlife, and recreation. Through forestry research, cooperation with the States and private forest owners, and management of the National Forests and National Grasslands, it strives—as directed by Congress—to provide increasingly greater service to a growing Nation.