U. S. FOREST SERVICE
RESEARCH NOTE PSW- 42

PACIFIC SOUTHWEST
FOREST & RANGE EXPERIMENT STATION
Berkeley, California 1964

A Method for Estimating Current Attendance on Sets of Campgrounds . . . a pilot study

RICHARD L. BURY AND RUTH MARGOLIES

ABSTRACT: Statistical models were devised for estimating both daily and seasonal attendance (and corresponding precision of estimates) through correlation-regression and ratio analyses. Total daily attendance for a test set of 23 campgrounds could be estimated from attendance measured in only one of them The chances were that estimates would be within 10 percent of true attendance, 2 out of 3 times. Specific procedures for field use will rest on each user's requirements for precision and cost.

Public agencies often need estimates of campground attendance for entire administrative units rather than for individual campgrounds. As part of an effort to improve the efficiency and precision of present methods, we wondered whether reliable estimates of total campground attendance for an administrative unit could be obtained directly without estimating attendance at each campground.

Obviously, this could be done if attendance at one or several campgrounds were reliably related to total campground attendance within the unit. Changes in daily total attendance owing to weather, holidays, and other general influences would be reflected by daily counts of attendance at the "bellwether" or indicator campgrounds. Such a method would be much simpler and cheaper than estimating attendance at each campground and then totaling for the administrative unit.

THE STUDY

To check this idea, we ran a pilot study in the Summit Ranger District of the Stanislaus National Forest.[1] The test included all 23 campgrounds in the District, an area of heavy

[1]A detailed discussion of this study, its results, and the statistical models devised is given in: Bury, Richard L., Margolies, Ruth and Navon, Daniel I. Estimating past and current attendance on a set of campgrounds: details of 1961 test on Summit Ranger District, Stanislaus National Forest. 1963. (Unpublished report on file at Pacific Southwest Forest and Range Expt. Sta., U.S. Forest Serv., Berkeley, Calif.)

recreation use located north of Yosemite National Park on the west side of the central Sierra Nevada in California. Most of the campgrounds in this unit lie beside a stream or creek; the two largest are within one-half mile of a 300-acre lake.

Basic data for the test were already available from another study in which we had counted attendance each night in each campground during 48 consecutive days in the summer of 1961. The attendance unit was the visitor group, that is, the group of persons who arrived at the campground in a single vehicle. As a follow-up, we counted attendance similarly for 15 days during 1962.

We used both correlation and ratio analyses to test for a reliable relationship between the daily number of visitor groups at each campground and the total daily attendance for the set of campgrounds. From these analyses, we could describe the strength of relationships between attendance at each campground and total attendance, and could rate each campground as an indicator of total campground attendance.

Statistical models for estimating seasonal attendance and the precision of such estimates were also derived. We did not, however, use the field data to calculate a numerical example.

THE METHOD

This method consists of three major steps: grouping campgrounds into sets, calibration, and estimation of attendance. Campgrounds may be grouped on any reasonable natural or administrative basis. Calibration consists of counting attendance in each campground in the group, testing for relationships by correlation or ratio analysis, selecting indicator campgrounds, and computing equations for estimating attendance. The last step is our objective--estimation of attendance during periods after calibration.

Total daily attendance can be estimated simply by inserting the observed attendance in the indicator campground(s) in the proper equation and solving. If ratio analysis is used, the estimating equation is:

$$\text{estimated daily attendance} = \hat{y} = \hat{R}x$$

in which x is known daily attendance on the indicator campground(s) and $\hat{R}$ is the ratio between total attendance and attendance in the indicator campground(s).

The estimating equation with correlation-regression analysis is:

$$\text{estimated daily attendance} = \hat{y} = a+bX$$

in which X is known daily attendance in the indicator campground, and a and b show the relationship between total attendance and attendance in the indicator campground. Additional indicator campgrounds are included in this estimating equation by adding more (bX) terms. Values for $\hat{R}$, a, and b will be different for each set of campgrounds, and are determined through calibration. For example, the regression equation for estimating total daily attendance from one indicator campground in our test set was:

$$\text{number of groups} = \hat{y} = 142 + 9.04X_{41}$$

in which X_{41} was the observed number of groups staying overnight at Sand Flat Campground.

Total seasonal attendance can be estimated by adding estimates of total daily attendance, or can be calculated directly from the total seasonal attendance at the indicator campgrounds.

RESULTS AND EVALUATION

Results indicate great promise for this method. On the test unit, we could estimate total attendance on an average day within 10 percent of true attendance, from counts of daily attendance on only one of the campgrounds. This level of precision could be expected in 2 out of 3 estimates. By adding a second and then a third indicator campground to the estimating equations, we improved the precision of estimated daily attendance to within 8 percent and 5 percent of true attendance, respectively--again at two-thirds probability. The manager can continue adding indicator campgrounds until he gets the desired precision.

Strong relationships of attendance within the District supported this good performance. For example, the correlation coefficient between total attendance and attendance on the best indicator campground yielded an r^2 of .85; correlation increased to an R^2 of .98 when we tested estimating equations that included three indicator campgrounds. The larger campgrounds were usually better as indicators than the small ones.

We tested the between-year stability of attendance patterns by running a correlation-regression analysis on attendance data from the same set of campgrounds during the summer of 1962. Two of the top three indicators were the same for 1961 and 1962. From this evidence and percentage analysis of each year's data, we inferred that patterns of attendance were quite stable for the two years. Relationships were stronger within the 1962 sample than those already reported for 1961.

Besides the between-year stability of indicator campgrounds, we found that most campgrounds maintained a consistent percentage of total daily attendance throughout the 1961 sample. This indicated a highly stable pattern of attendance among campgrounds on Summit Ranger District. In such cases, reliable estimates of attendance on individual non-indicator campgrounds can be obtained from estimating equations keyed to the indicator campgrounds.

We expect that total attendance can be estimated with less error for an entire season than for single days if the error is expressed as a percent of the actual attendance. When expressed in visitor-days or other such attendance units, however, the error actually increases as each day is added to the season. The percentage error decreases because the error of estimate increases proportionately less than actual attendance.

DISCUSSION

The chief expense of this method would normally occur during calibration, because attendance at each campground would usually be counted daily by direct observation. In our test set, this job took about 4 hours and 90 miles of driving to cover the 562 campsites in the 23 campgrounds. In units where campgrounds are widely scattered, the time might be reduced by using several decentralized observers. Cost might also be lowered if attendance in individual campgrounds during calibration could be reliably estimated by some other method. However, such estimation would decrease the precision of final estimates.

Statistical processing of calibration data will require several man-weeks unless electronic computers are used. Those in charge of the work must have enough statistical knowledge to design and supervise the sample, to perform the statistical tests, and to interpret the results. If several units are to be calibrated by the same agency, costs might be lowered by training a special calibration team.

Each unit or set of campgrounds must be calibrated separately. It should also be recalibrated periodically--say, at 5-year intervals--and after every change in recreation facilities which might markedly affect the relationship of attendance among campgrounds.

Attendance relationships within the set may be different on weekdays than on weekend days. Relationships during the summer will probably be quite different than those of other seasons. Separate estimating equations could be derived and used if the difference seemed large enough. In such cases, however, the statistical model must be reworked and precision of estimates will be harder to calculate.

Attendance during calibration and estimation must be counted at a regularly scheduled time. Counts should not be made late in the morning or early in the evening because of bias to underreporting campers and the possibility of confusing day-visitors with campers.

Other researchers have reported equally favorable results from tests of similar methods. Using a heavy sample with subsets of campgrounds as indicators, Wagar could estimate seasonal attendance on a set of 8 campgrounds within 2.4 percent of an unbiased estimate of seasonal attendance, with two-thirds probability.[2] Similarly, Wenger found a high correlation ($r^2 = .85$) between daily attendance in one campground and total attendance in a set of 26 campgrounds on a National Forest in the Pacific Northwest.[3]

If direct observation is used for both calibration and indicator counts, bias in the data can be controlled by proper calibration design and specifications for counting visitors. However, the possibility of data bias should be carefully considered if attendance on individual campgrounds is estimated rather than observed directly.

This method is compatible with all common measures of attendance. We used group-days as the attendance unit in our test. But attendance can be estimated in other measurement units either directly, or indirectly through conversion factors derived from sampling the character of use on campgrounds in the set. By finding the average length of stay, for example, we could convert from group-days to group-visits. In turn, group-visits could be converted to visitors through expansion factors tied to the average visitors per group and the frequency of repeat groups.

CONCLUSION

We expect this method to be cheaper than those based on estimating attendance at individual campgrounds. After regression or ratio equations are derived, one solution will estimate total campground attendance--and possibly attendance on individual campgrounds--from attendance observed on only one or a few campgrounds. Furthermore, one solution will provide estimates for any desired length of time, and equations can be solved in only a few minutes.

The main drawback of this method is that attendance must be counted each day on indicator campgrounds. Actually, this job may be easy if the better indicators include a campground for which attendance is already being counted daily--for instance, a campground run on a fee basis. The difficulty could also be lessened by two sampling methods--first, by counting only on sample days during the season; and second, by measuring attendance in the indicator campgrounds through some means other than direct observation. In both cases, additional error could occur and the statistical model would need adjustment accordingly.

[2]Wagar, J. Alan. Sampling overnight use on unsupervised campgrounds in the White Mountain National Forest. 1963. (Unpublished report on file, Northeastern Forest Expt. Sta., U.S. Forest Serv., Upper Darby, Pa.)

[3]Personal correspondence with W. D. Wenger, Pacific Northwest Forest and Range Expt. Sta., U.S. Forest Serv., Portland, Oregon, May 13, 1963.

We have not yet devised standard procedures for applying this method. Specific procedures must rest on examination (and possible modification) of statistical models for conformance with field conditions and the planned use of attendance estimates. The statistical models now require calibration samples of at least 20 random days throughout the season for correlation-regression and 30 random days throughout the season for ratio analysis. Also, the statistical models underlying estimates of seasonal use should be refined; this step may permit a smaller sample and therefore lower cost. Refinement of statistical models will require information on the geographic and administrative areas to be covered, the levels of precision acceptable to potential users, the period for which estimates are desired, and how much can be spent on the job.

RICHARD L. BURY has been conducting forest recreation studies, with headquarters in Berkeley, since 1960. A graduate of Purdue University, he holds a master's degree from Yale, and a doctorate from the University of Connecticut. RUTH MARGOLIES was formerly a social science analyst at this Station.

U. S. DEPARTMENT OF AGRICULTURE
FOREST SERVICE
PACIFIC SOUTHWEST FOREST AND RANGE EXPERIMENT STATION
POST OFFICE BOX 245
BERKELEY, CALIFORNIA 94701

OFFICIAL BUSINESS

POSTAGE AND FEES PAID
U. S. DEPARTMENT OF AGRICULTURE

ROCKY MT FOR EXP STA 01 5
ARIZONA STATE COLL CAMPUS
FLAGSTAFF ARIZONA