SIMULATION AND OUTDOOR RECREATION

by

Frank J. Cesario
Social and Systems Science Research Section
Battelle Memorial Institute
Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201

Abstract

Outdoor recreation systems are comprised of three basic components: (1) a set of N population centers, (2) a set of M recreation sites, and (3) a set of at least NM connecting links. An important problem facing planners in the allocation of resources is how to estimate the magnitude of travel per unit time flowing in systems of varying characteristics. Simulation is viewed to be one possibility. A gravity-type model written in GPSS III was developed to simulate travel in a system comprising most of the Commonwealth of Pennsylvania in order to explore the possibility of examining the effects of:

(a) adding new parks
(b) constructing new highways
(c) a rise in the level of participation.

Data collected in Pennsylvania during the summer of 1968 were used as applicable.

INTRODUCTION

The factors which shape the level of attendance at outdoor recreation sites are not well understood. As a result, park planners and administrators are forced to make important planning and policy decisions on fragmentary information or personal whim. Considering the millions of dollars spent annually on recreational development by all levels of government – money which has alternative uses – this situation is very undesirable from the standpoint of economic efficiency. If recreation planning and policy decisions are to be made rationally, models of recreation systems must be developed which make explicit the relevant variables and interactions, and their effect on park use with respect to alternative policies. Such tools are not currently part of the recreation planners' stock in trade. In this paper, one possible approach to model building, simulation by GPSS III, is studied.

It is appropriate at this time to issue a disclaimer of sorts. The effort described in this paper is considered by the author to be a "demonstration" project. The purposes of conducting this study were to (1) illustrate what kinds of problems might be investigated by a simulation approach, (2) illustrate some typical policy experiments, and, (3) make a preliminary judgment on the potential usefulness of this method of analysis in analyzing outdoor recreation policy problems. The model is very basic and lacks technical sophistication. It is hoped, of course, that this paper will stimulate further efforts to construct more realistic models in the future.

DESCRIPTION OF THE RECREATION SYSTEM AND RELEVANT FACTORS

A recreation system is composed of:

(1) A set of "origins", population centers $C_i; i = 1, N$

(2) A set of "destinations", recreation sites $P_j; j = 1, M$

(3) A set of "travel links", highways $R_{ij}; i=1, N; j = 1, M$.

The population unit selected in this study is a county. Participation in outdoor recreation can be expected to vary from county to county. The major factors which determine participation are (1) population magnitude, (2) socioeconomic characteristics such as income distribution and age structure, and (3) available recreation opportunities.

The recreation sites considered in this study are of the state park or intermediate variety, although it is not necessary to make such a restriction. State parks are usually located within an hour or two driving time from the cities, are large well-defined areas offering the full gamut of recreational activities, and are homogeneous with respect to external factors such as advertising and pricing. However, parks differ with respect to their "attractiveness" to the general public. People generally prefer sandy beaches to rocky ones, for instance. These differences, since they affect user decisions, have to be taken into account.

In reality, there are many possible routes over which recreationists can travel from their homes to the recreation sites. In this study, the route requiring the minimum outlay of time was selected as the county-recreation site link.

The problem that develops when a county is used as the population unit is where to take the origination point of the highways that supposedly connect the county with each park. The largest city in each county was used as the origination point because (1) generally speaking, each county has only one main city or urban area toward which most county economic activity is focused,aand (2) a large majority of visitors from the county will emanate from its major urban area. In other words, the major urban area is taken to be the weighted mean population location.

THE MODEL

Description of Study Area. The system chosen for study is in Northeastern Pennsylvania, in the vicinity of the Delaware Water Gap. The study area is shown in Figure 1. The dynamics of day use (i.e., trips taken from home lasting one day or less) in the 5 Pennsylvania state parks listed in this Figure is to be studied. Some relevant descriptive data on each park is given in Table 1. Table 1 also gives the attractiveness "indexes" of these parks. These indicate relative attractiveness only, e.g., Ricketts Glen is 10 times as attractive as Big Pocono, and have no physical meaning. The method by which these indexes are derived is somewhat lengthy and is not given here. *

In simulating visitor flows to these parks, it is necessary to consider all the enclosed counties in Figure 1. Visitors from these counties will make up the bulk of the total visitors since the rate of decay of visitation with respect to travel time is very rapid for these types of parks. ** Population data for the study area counties is given in Table 2.

The period simulated is one day (8 a.m. to 8 p.m.). One clock unit equals one minute. Field data were used in the model where possible, but many subjective factors had to be included since the data were not collected originally for this purpose.

TABLE 1. PERTINENT DATA ON PENNSYLVANIA STATE PARKS IN SAMPLE

Number	Park	County	Capacity[a]	Acreage	1968[b] Attendance	Attractiveness Index
1	Big Pocono	Monroe	230	1,306	204,643	1
2	Gouldsboro	Monroe-Lackawanna	800	2,800	183,729	8
3	Hickory Run	Carbon	850	15,500	561,812	6
4	Ricketts Glen	Luzerne	1,000	13,134	465,170	10
5	Tobyhanna	Monroe	600	4,188	303,014	4

[a] In terms of number of parking spaces.
[b] Includes both day use and overnighters.

*See Cesario, F., Goldstone, S., and Knetsch, J., Outdoor Recreation Demands and Values, a report to The Middle Atlantic Utility Group, Battelle Memorial Institute, Columbus Laboratories, December 30, 1969 for a discussion of this issue.

**Cesario, F, et al., op. cit. ibid.

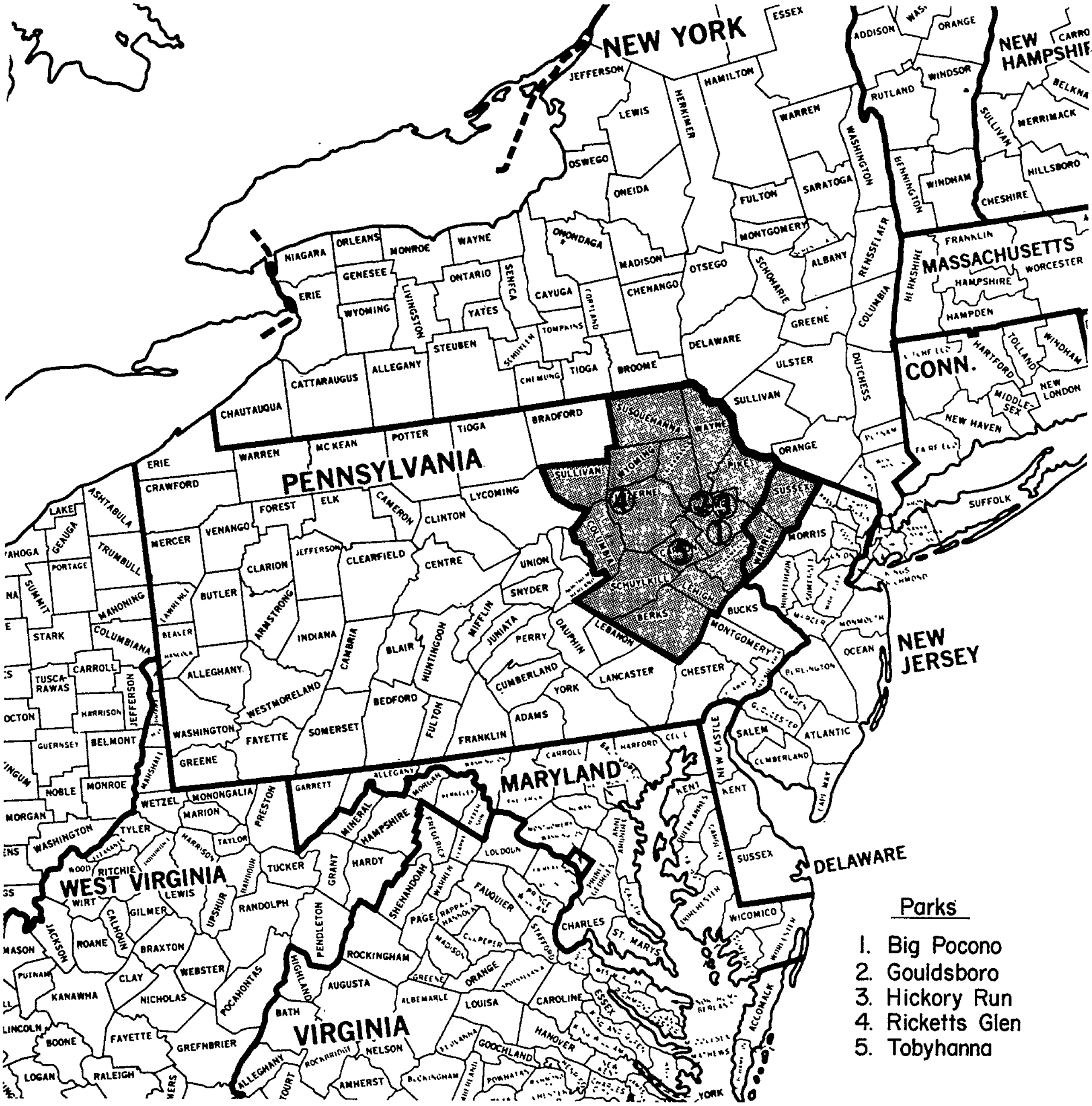


FIGURE 1. STUDY AREA

TABLE 2. POPULATION DATA FOR COUNTIES IN RECREATION SYSTEM,[a, b] IN THOUSANDS

Number	County	1968 Population, persons	1968 Population, households
1	Berks	288.1	91.7
2	Carbon	51.3	16.1
3	Columbia	53.6	16.5
4	Lackawanna	222.4	67.4
5	LeHigh	241.4	74.4
6	Luzerne	332.2	102.1
7	Monroe	43.2	13.4
8	Northampton	213.2	64.8
9	Pike	10.0	3.3
10	Schuylkill	159.1	50.3
11	Sullivan	6.3	1.8
12	Sussex (NJ)	61.8	18.3
13	Susquehanna	34.2	9.8
14	Warren (NJ)	69.3	21.3
15	Wayne	27.1	7.8
16	Wyoming	18.8	5.5

[a] These figures are estimates obtained from Sales Management-Survey of Buying Power, E. Stroudsberg, Pa., June 10, 1969.
[b] Counties in Pennsylvania, unless otherwise noted.

Model Description. Household trips are generated from each of the 16 counties to the 5 parks in question. The scaled number of trips is an estimate of the actual number taken on a typical summer Sunday.* The trips are distributed to each of 5 parks based on the formula:

$$T_{ij} = N_i \left[\frac{\dfrac{A_j}{D_{ij}^{\alpha}}}{\sum_{\ell=1}^{5} \dfrac{A_\ell}{D_{i\ell}^{\alpha}}} \right] \qquad \text{(Eq. 1)}$$

where,

T_{ij} = number of trips from county i to park j

N_i = total number of trips taken to parks in system from county i

A_j = "attractiveness" of park j

D_{ij} = expected travel time from main city in county i to entrance to park j

α = parameter expressing elasticity of visitation with respect to travel time.

*Since the number of transactions in process at any one time would exceed 750 the system had to be scaled down approximately in a 1:5 ratio. This means that every transaction generated in the model represents 5 households in the real world.

This is a variation of the popular "gravity" model, discussions of which can be found in many references.*

As

$$\left[\frac{\dfrac{A_j}{D_{ij}}}{\sum_{\ell=1}^{5} \dfrac{A_\ell}{D_{i\ell}}} \right] \geq 0 \text{ for all i and j}$$

and

$$\sum_{j=1}^{M} \left[\frac{\dfrac{A_j}{D_{ij}}}{\sum_{\ell=1}^{5} \dfrac{A_\ell}{D_{i\ell}}} \right] = 1.0$$

each term can represent the probability that a household from city i chooses park j. (For convenience, α was chosen to be 1.0).

The number of trips generated per unit time at each source i is assumed to be a function of population, Ti, and availability of recreation opportunities. The particular formulation is

$$Ni = k\, P_i \left(\sum_{\ell=1}^{5} A_\ell D_{i\ell}^{-1} \right) \qquad \text{(Eq. 2)}$$

The mean time between originations is allowed to change over the course of the day. It is well known that recreation travel is moderate from 8:00 a.m. to 11:00 a.m., heavy from 11:00 a.m. to 2:00 p.m., and lightest after 2:00 p.m. The mean generation times used in the basic system (without modification for experimentation) are given in Table 3. The integeneration times were assumed to be exponentially distributed.

Only trips made to the 5 study parks are considered. The "leakage" to parks outside the system is of no interest in this study.

Travelers are ADVANCEd over the highways to the parks, where they must queue to pay the entrance fee. Expected county-park travel times were estimated from road maps. An arbitrary allowance of ±10 percent is made in this expected travel time.

*For example, Oi, W. Y., and Shuldiner, P. W., An Analysis of Urban Travel Demands, Northwestern University Press, 1962, pp. 58-59; and Voorhees, A. M., "A General Theory of Traffic Movement", Proceedings, Institute of Traffic Engineers, 1956, pp. 46-56.

TABLE 3. MEAN GENERATION TIMES OF RECREATION TRIPS, BY TIME PERIOD, IN MINUTES

Number	County	Time Period of Departure, in hours from 8:00 a.m. 0-3	3-6	6-9	9-12
1	Berks	34	17	68	136
2	Carbon	22	11	44	88
3	Columbia	11	5	24	48
4	Lackawanna	3	2	5	10
5	LeHigh	13	7	26	52
6	Luzerne	2	1	3	6
7	Monroe	22	11	44	88
8	Northampton	5	11	22	44
9	Pike	0	180	0	0
10	Schuylkill	5	11	22	44
11	Sullivan	68	34	126	252
12	Sussex (NJ)	0	180	0	0
13	Susquehanna	0	180	0	0
14	Warren (NJ)	0	60	0	0
15	Wayne	0	90	0	0
16	Wyoming	68	34	126	252

Since park capacity utilization is of major importance, the contribution to attendance by county residents outside the study area cannot be overlooked. The proportion of total attendance at each park accounted for by out-of-study-area counties, as estimated from field data, is given in Table 4.

TABLE 4. PROPORTION OF ATTENDANCE ACCOUNTED FOR BY OUT-OF-STUDY-AREA COUNTIES

Number	Park	Percent of Attendance, Included Counties	Percent of Attendance, Excluded Counties
1	Big Pocono	64	36
2	Gouldsboro	92	8
3	Hickory Run	78	22
4	Ricketts Glen	71	29
5	Tobyhanna	90	10

A GENERATE block was included which represented visitors from all excluded counties. To account for travel, these transactions were not ADVANCEd over highways as were the others, but were assigned directly to the parks.* The number assigned to each park is based on a study of data on origins of visitors to each of the parks. The probability that a trip from outside the study area visits park j, $P(O)_j$; is given by

$$P\,(O)_j = \frac{V_{oj}}{\sum_{m=1}^{5} V_{om}} \qquad \text{(Eq. 3)}$$

where,

V_{oj} = visitors from outside study area to park j on days when data examined.

To account in part for travel time, the generations of visitors from this "external" GENERATE block is offset 60 minutes.

If the park of first choice is full, the party may or may not wait, depending on how many parties are in queue. If the queue size is 10 or less, the party waits. If the queue size is greater than 10, the party elects either to return home or visit another park, its decision being based primarily on the relative travel times between returning hom and visiting the next nearest park. If $D_{j,j'}$ is the expected travel time from park j to park j' and $D_{j,h}$ is the expected travel time to return home, then the probability that the party returns home from park j, $P(h)_j$, is

$$P(h)_j = \frac{D_{j,j'}}{D_{j,j'} + D_{j,h}} \qquad \text{(Eq. 4)}$$

and the probability that the party visits the next closest park, $P(p)_j$ is

$$P(p)_j = 1 - P(h)_j = \frac{D_{j,h}}{D_{j,j} + D_{j,h}} \qquad \text{(Eq. 5)}$$

Of course, if the party elects to go to another park, it is subject to being turned away at the next park as well. The same rules for selection of returning home or going to the next nearest park apply. When the selection has been made to return home, the trip is TERMINATEd.

Once inside a park, the party stays a certain length of time, the expected length of stay being longer for earlier arrivals than for late ones. As with the intercreation times, different mean lengths of stay are given for 3 different time periods. The expected length of stay varies among the parks since different activities are offered by each. The mean lengths of stay used in the model are given in Table 5. Lengths of stay are assumed to be exponentially distributed.

*This was necessary so that queuing statistics could be collected.

TABLE 5. MEAN LENGTH OF STAY FOR PARKS IN STUDY AREA IN MINUTES

Number	Park	Time Period of Arrival, in hours measured from 9 a.m.			
		0-3	3-6	6-9	9-12
1	Big Pocono	120	60	30	15
2	Gouldsboro	240	120	60	30
3	Hickory Run	300	150	75	38
4	Ricketts Glen	200	100	50	25
5	Tobyhanna	200	100	50	25

Upon completion of the visit, the trips are TERMINATEd:

Output. The output from each simulation run consists of

(1) Storage (Park) statistics at end of each hour

(2) Queue (at toll booths) statistics at end of each hour

(3) Total number of parties diverted from parks of their first choice, and number of parks to which parties were diverted

(4) Total number of parties from county i who initially set out for park j

(5) Queuing time (via QTABLE) distribution

(6) Queue length distribution

(7) Transit time (from home back to home) distribution

(8) Standard output and queuing statistics.

Other output data are occasionally of interest, but the above information is most important for the questions being investigated.

EXPERIMENTS

Three types of system alterations were performed in the investigation of likely uses of the model:

(1) A new interstate highway was "constructed" through the study area

(2) A new park was "developed" in the study area

(3) Participation was projected to increase.

The ramifications of each alteration are discussed below.

(1) Construction of a New Highway. A new highway, by decreasing travel time for visitors who are in a position to use it, serves to increase the "accessibility" of certain parks. Since participation depends in part on accessibility, we would expect an increase in participation from those counties who could make use of the new highway to visit one or more parks. The percentage by which existing participation increases is assumed, according to Equation (2), to be a linear function of accessibility, namely

$$\frac{\sum_j \left(\frac{A_j}{D_{ij}{}^{\alpha}}\right) \text{after} - \sum_j \left(\frac{A_j}{D_{ij}{}^{\alpha}}\right) \text{before}}{\sum_j \left(\frac{A_j}{D_{ij}{}^{\alpha}}\right) \text{before}} \times 100. \qquad \text{(Eq. 6)}$$

Since relative accessibility among individual parks has changed, a new visitor distribution function must be defined for each county in accordance with Equation 1. Finally, the intergeneration times for counties making use of the new highway is decreased. We want to compare the visitor distributions before and after construction of this highway.

(2) Addition of a New Park. The new park, named "Cesario's Folly" is to be located arbitrarily in Luzerne County. Addition of a new park causes two changes in visitor flows:

(a) A redistribution of visitors, and

(b) Increased participation.

The redistribution takes place according to Equation (1). Participation effects are calculated by Equation (6).

It is necessary to assign an attractiveness index to this new park. Based on comparisons of its characteristics with other parks, Cesario's Folly is judged to be about "average", therefore an attractiveness index of "5" was assigned. Its scaled capacity is set at 150 cars.

(3) Projection of Participation. There is presently much controversy in the United States about the future of outdoor recreation. "Experts" have projected anywhere from a two-fold to a ten-fold increase in participation by the year 2000. What effect will this increase (whether on the low or high side of the projections) have on our developed resources, the supply of which most certainly will not grow at anywhere near the same rate as participation. Where and when should capacity be added?

Analysis of this question seems like a natural use of this simulation model. A sensitivity analysis, using both high, low, and intermediate projections and combinations of these, would be appropriate.

Selected Experiments. Using the above basic set of alterations, 5 "experiments" were performed on this system. These experiments are as follows:

Experiment 1 - The existing system was run

Experiment 2 - Participation was projected to increase by about 20 percent over present levels

Experiment 3 - The new highway was added to the existing system

Experiment 4 - The new park was added to the system.

Experiment 5 - The new highway was combined with a projected increase in participation.

RESULTS

Several policy "experiments" were conducted using the model as described. These are merely illustrations of the way in which a simulation model might be used to examine likely effects of contemplated or unexpected changes in the existing system.

Each "experiment" was run only once. Obviously, if the model had to be CLEARed and STARTed several times on each run, the results would be slightly different. The emphasis was put in illustrating possibilities rather than analyzing any one possibility in depth.

Average Park Utilization. Average utilization for each run is given in Table 6. This result would support a popular contention that while parks might be operating near capacity at certain times, they are not utilized extensively, on the average, over a period of time as short as one day.

TABLE 6. AVERAGE PARK UTILIZATION

Run \ Park	1	2	3	4	5	6
1	.14	.35	.34	.42	.21	NA
2	.11	.37	.37	.50	.20	NA
3	.11	.42	.49	.50	.26	NA
4	.11	.42	.37	.39	.31	.37
5	.13	.52	.46	.64	.27	NA

Also, it can be seen that as accessibility and/or participation increases, utilization increases. This is to be expected. The new park (Number 6) seems to divert visitors only from a few parks.

Crowdedness. In only one case - Run 5 - was a park (Ricketts Glen) full to capacity at the peak time of day. No visitors were evidently diverted to other parks, so there is good possibility that the number of parties waiting to get into Ricketts Glen at the time of peak load was never greater than 10 when another party arrived. A few parties may have returned home, however, since they were not TABULATEd.

Ricketts Glen also operated at near capacity (over 90 percent) in Runs 2 and 3. In all other parks, capacity was never closely approached.

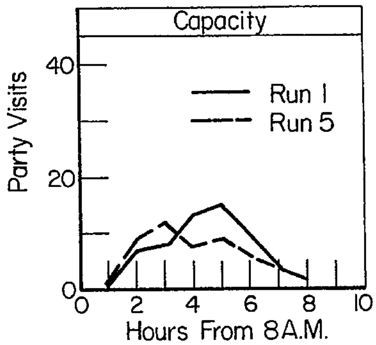
Time of Buildup to Peak Load. The time of day when peak loading occurred in the model actually agreed very closely with the time of peak loading in the actual system. Graphs of the visitor loads in each park for the two extreme runs (Runs 1 and 5) at the end of each hour are given in Figure 2.

It can be seen that the peak load is obtained in early afternoon between 1:00 p.m. and 3:00 p.m. This agrees with observations, although the load does not, in reality, drop off as quickly as it does in the model. Perhaps the length of stay should be increased after 2:00 p.m. in any future runs. The reason that peak times differ is due primarily to changes in parameter values that could not, because of technical problems, be made proportional.

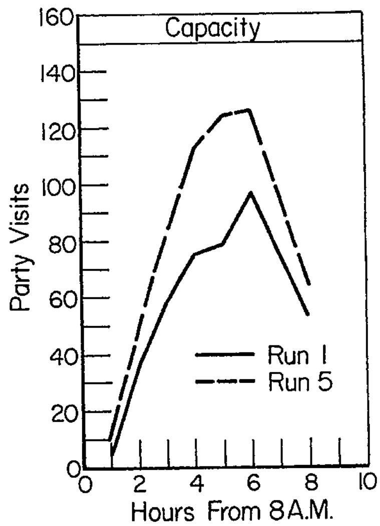
Queue Length at Park Entrance. The queue lengths attained in the simulation indicate that Ricketts Glen probably will require more than one toll booth during certain periods unless another park is added The maximum number of parties in queue for each park in each run is given in Table 7.

TABLE 7. MAXIMUM NUMBER OF PARTIES IN QUEUE

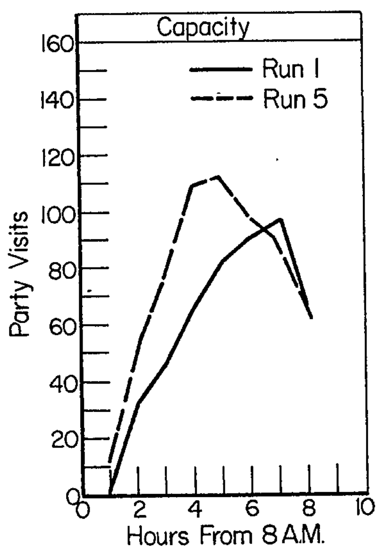
Run \ Park	1	2	3	4	5	6
1	1	14	5	57	4	NA
2	1	9	7	41	4	NA
3	1	11	12	33	5	NA
4	1	6	6	9	5	8
5	1	12	8	44	8	NA



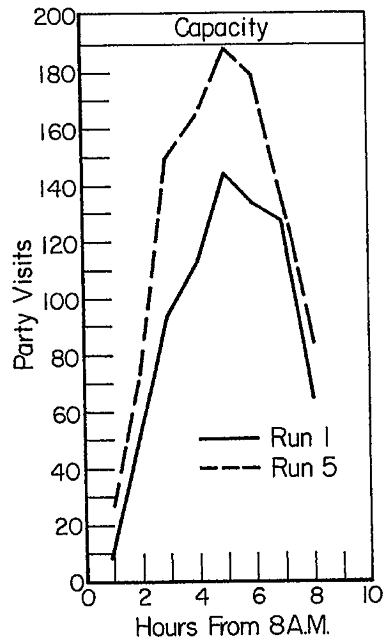
a. Big Pocono



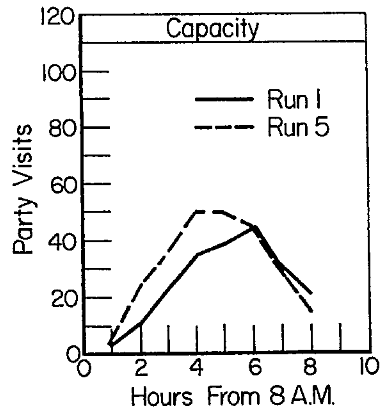
b. Gouldsboro



c. Hickory Run



d. Ricketts Glen



e. Tobyhanná

FIGURE 2. VISITOR LOADINGS ON 5 PARKS

By examining the distributions of queue lengths and time in the queue we could make a decision as to how many booths are needed at what times.

TENTATIVE CONCLUSION

It is tentatively concluded, with this limited evidence, that simulation with GPSS III could indeed be useful for analyzing the kinds of problems confronting recreation planners. The large number of assumptions made in constructing and using the model discussed in this paper and the lack of verification of the results does not allow this author to be more committal. A firm conclusion can only be reached after a more realistic model is constructed with better data and the empirical results are thoroughly tested against the real world. It is hoped that such models will follow this preliminary investigation.