# MODEL FORMULATION FOR FLEET SIZE DETERMINATION OF A UNIVERSITY MOTOR POOL

William W. Williams, The University of Tennessee
Oscar S. Fowler, The University of Tennessee

## ABSTRACT

A GPSS simulation model was developed to investigate the behavior of a university's motor pool dispatch fleet. Data were collected and frequency distributions were constructed for request rates and trip duration times. Regression analysis was performed to formulate mileage generator(s) as a function of service time. Statistical tests were conducted to assess degree of congruence between actual system behavior and theoretical expectations. Analysis of simulated response variables indicates that the model's internal configuration reflects reality to a high degree. Results of fifteen years of simulated experience seem to suggest that reductions in fleet capacity could be effected without undue impairment of service levels.

## INTRODUCTION

Many public agencies, including universities, maintain a fleet of vehicles for institutional-related travel by employees. Typically, requests are made to a dispatching office which determines availability and assigns cars to fill the requests. Decisions as to the size of the fleet have a significant impact on the resources from operating funds required to provide this service. In an era of increasingly austere budgets, the accurate determination of an optimum size for the fleet becomes even more important. Reduction in fleet size without undue impairment of service would permit funds required for replacement and maintenance to be released for alternative purposes. During fiscal 1975, the year from which the data was accumulated, The University of Tennessee-Knoxville dispatch fleet consisted of 145 vehicles with an effective capacity of approximately 135.

University policy encourages the use of these vehicles for official travel but does allow employees some discretion. Employees using personal vehicles for their own convenience are reimbursed operating costs only (i.e., the average rate departments are charged for university-owned vehicles). If the travel is in a personal vehicle for university convenience (i.e., a dispatch vehicle was not available upon request) reimbursement must be on a full cost basis (i.e., including some depreciation, etc.).

The determination of the minimum-cost fleet size can be visualized as a classical optimization problem. Increasing the size of the available fleet reduces the number of unsatisfied requests and concomitantly the amount of travel which must be reimbursed at the higher "personal inconvenience" rate. However, such increments to capacity magnify imputed costs of opportunities forgone as well as increase the fixed investment in the fleet. Conversely, decreasing the fleet has the opposite impact on costs.

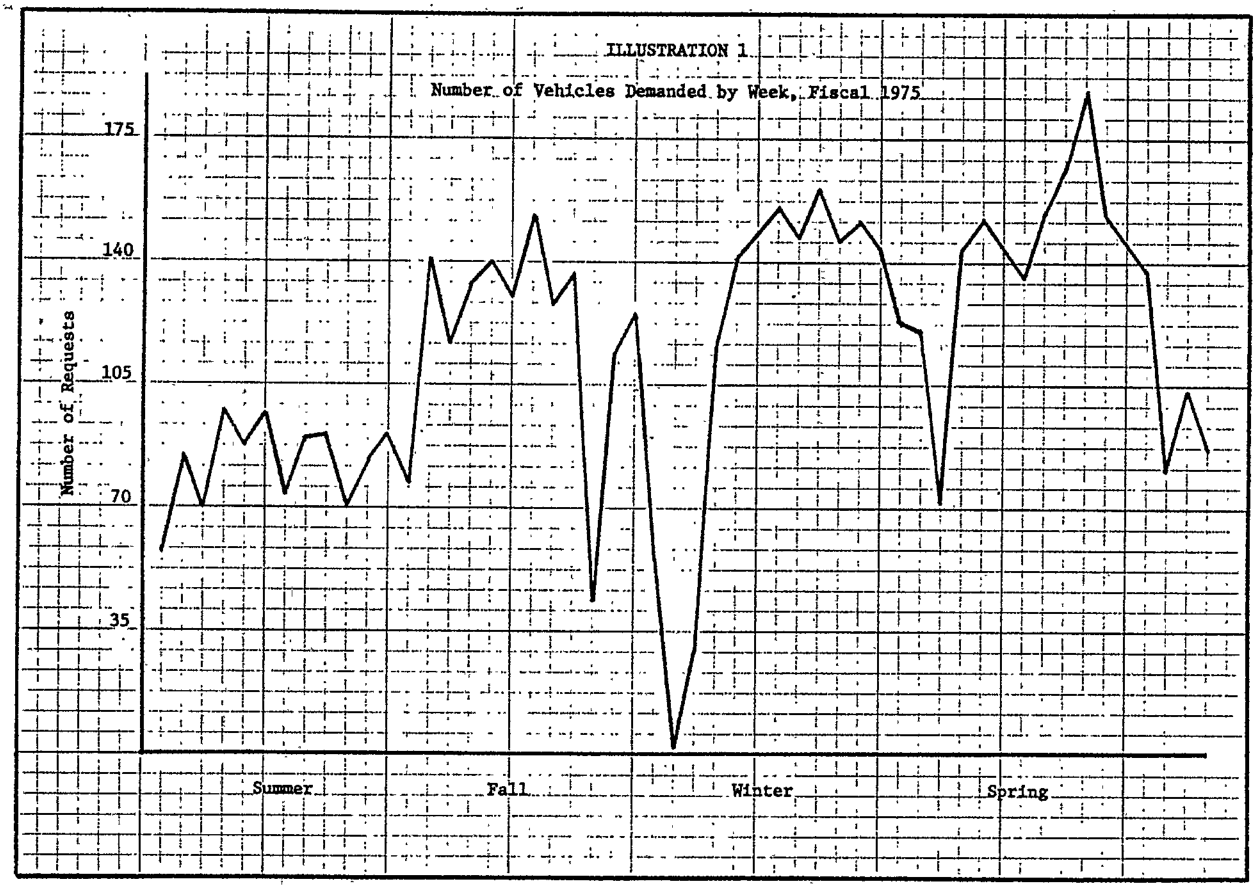
Accurate specification of the cost structure by inspection becomes exceedingly improbable, however, given the randomness inherent in the requests for vehicles and the length of trip duration. Consequently, a Monte Carlo simulation of the dispatch fleet's behavior over time seemed to offer the greatest analytical promise.

## SIMULATION DESIGN

### INPUT SPECIFICATION: VEHICLE REQUEST RATE, SERVICE TIME, AND MILEAGE DISTRIBUTIONS

Vehicle request rates were determined by examining motor pool dispatch records and identifying temporal demand loads on fleet capacity. Illustration 1 depicts the aggregate demand for vehicles by week for fiscal 1975. Inspection of this time series suggests that the pattern of vehicle demand is closely associated with the University's academic calendar. Although some variation exists, it appears that average weekly demand is lowest during the summer quarter and successively increases in aggregate terms throughout the fall, winter, and spring quarters. More extreme fluctuations coincide with University holidays and quarterly class breaks. Consequently, it was decided that a unitary vehicle request density function would be inappropriate in light of the magnitude of discrete changes in average demand.

Data were then dis-aggregated into daily requests and patterns were analyzed for each identifiable academic "season"; a daily demand time series for the fall quarter is shown in Illustration 2. Upon perusal of the daily time series it became apparent that the distribution of vehicle requests could not be considered completely random. Rather, peak weekly demand typically occurs early in the week (i.e., Monday and Tuesday), with the request rate subsequently declining later in the week (i.e., Thursday, Friday, and Saturday). A runs test was performed on the data for each period; the hypothesis posited that demand sequences would not be




random. Since demand decreases over the week, fewer runs than theoretically expected would be anticipated. For two of the four periods, the hypothesis of random demand sequences was rejected at the $\alpha$=.05 level; the remaining two periods evidenced fewer runs than the expectation.

Next, to test for the strength of autocorrelation in day-to-day demand sequences, zero-order correlations were computed for the daily request pairs $D_i$ and $D_{i+k}$[1]. High positive correlations (within the context of weekly day pairs) would suggest some type of "peaking" behavior, whereas high negative associations would be indicative of some type of "compensatory" demand pattern. Of the 60 lagged correlation coefficients computed, less than 15% proved significant at the $\alpha$=.05 level; furthermore, for the day parings that emerged with significant correlations, no consistent sequential pattern was discernible. Consequently, the requests for automobiles across time were assumed essentially independent but the "call" or mean daily demand rate was seen to be a function of the day of the week as well as the academic "season."

Finally, frequency distributions for vehicle requests were constructed for each weekday within each discrete demand period. These empirical distributions were compared to theoretical distributional expectations, estimated on the basis of the mean and variance of the observed data. The Kolmogorov goodness of fit test was employed since this test is generally considered more exact and powerful than the chi-square test for small sample sizes. For all twenty tests conducted (five weekdays for each of the four major demand periods) the hypothesis that actual daily demand was described by a Poisson process could not be rejected (See Illustration 3 for an example). Hence, vehicle request rates were assumed to exhibit a non-homogeneous Poisson configuration, i.e., requests would be randomly distributed across a given day of the week (e.g., all "Winter" Mondays), but each day would reflect a singular mean demand rate appropriate to the specific academic "season."

Investigation of the service or trip time observations proceded anolagous to that performed on the request rate distributions. Examination of motor pool records indicated the time of vehicle dispatch and return. In order to develop trip time distributions it was first necessary to ascertain the length of "turnaround" time or the amount of time required to service an incoming vehicle prior to a subsequent dispatch. Although motor pool personnel believed this time varied, primarily as a function of system load, a conservative approach was adopted for purposes of simulation design. It was decided that if a vehicle was requested for departure after 4 p.m. an incoming vehicle could be serviced and

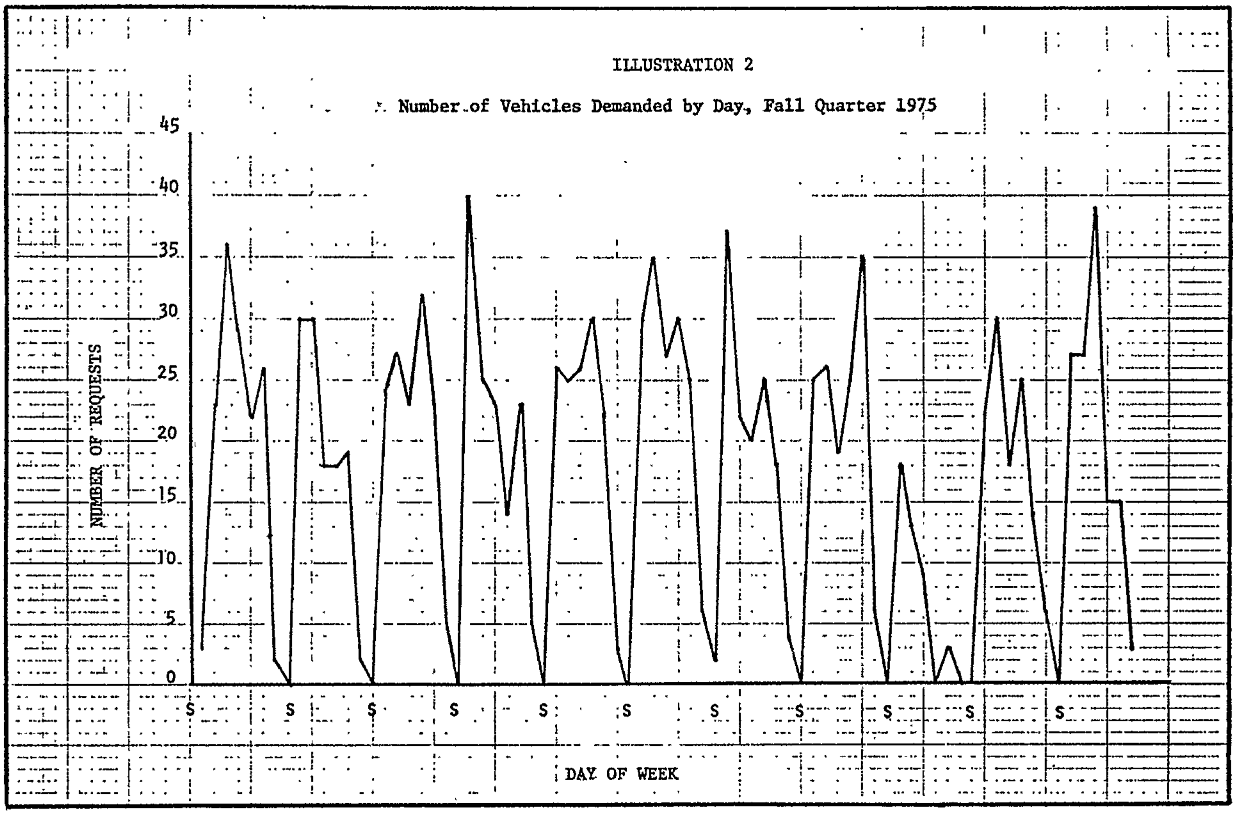



ILLUSTRATION 2

Number of Vehicles Demanded by Day, Fall Quarter 1975

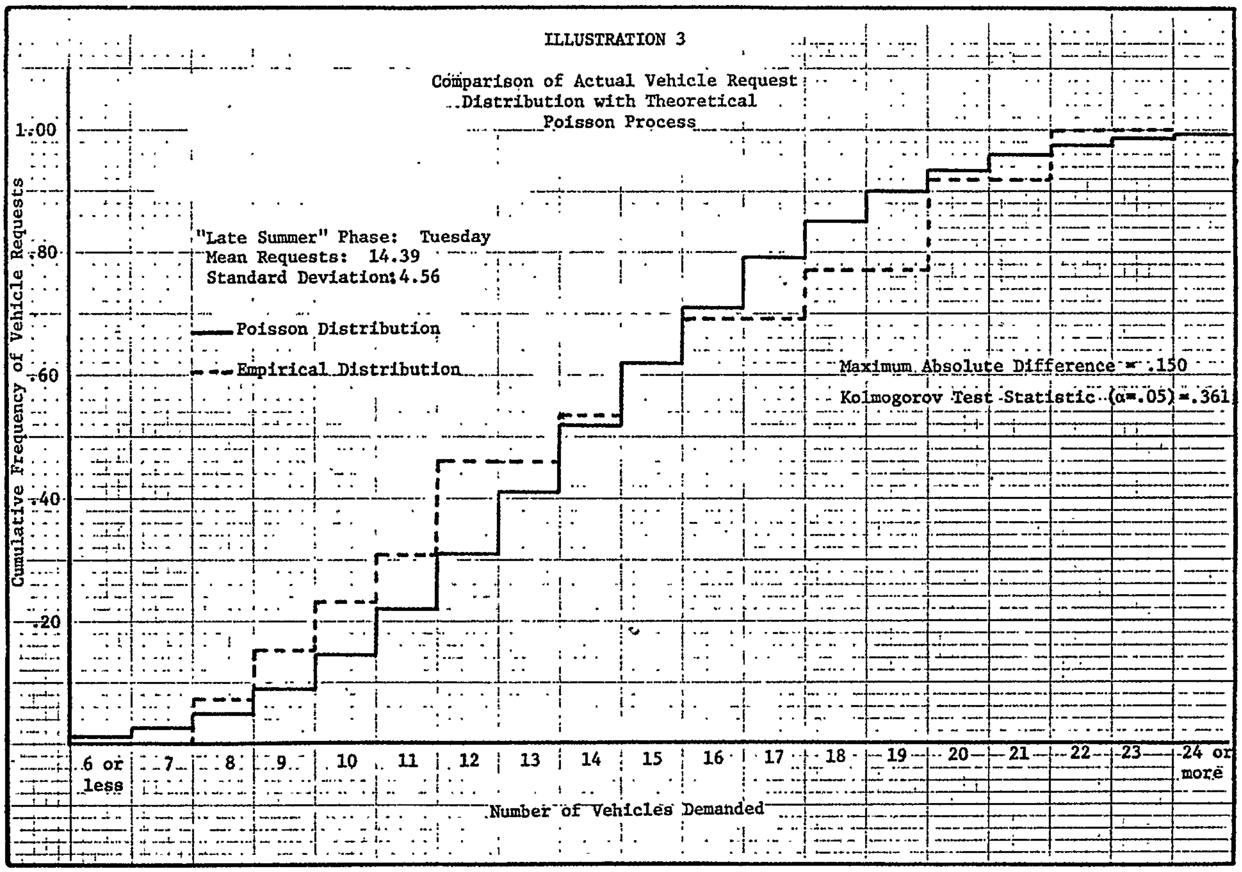



ILLUSTRATION 3

Comparison of Actual Vehicle Request Distribution with Theoretical Poisson Process "turned" to meet the dispatch.[2] A request for departure before 4 p.m. would be the equivalent of "reserving" a vehicle for an entire day. The impact of these decisions is to upwardly bias trip time durations and admit a degree of conservatism into the simulation output streams.

Tests were conducted to ascertain whether day of vehicle dispatch was associated with length (duration) of trip. Due to exigencies encountered in the data collection process, however, the expenditure of research effort into examining every dispatch day/trip time combination became prohibitive. Accordingly, for two representative months (i.e., October and March) trip time distributions were developed for each weekday. Goodness of fit tests were conducted between each weekday trip time distribution employing the Smirnov test. For October, although the hypothesis of equivalent trip time distributions was rejected for three sets of day pairings at the $\alpha=.05$ level, none could be rejected at the $\alpha=.01$ level. Likewise for the March data, no significant differences existed in trip time distributions except for Fridays.[3] Furthermore, an inspection of each daily trip time distribution with a composite/average trip time distribution revealed little conspicuous deviation.

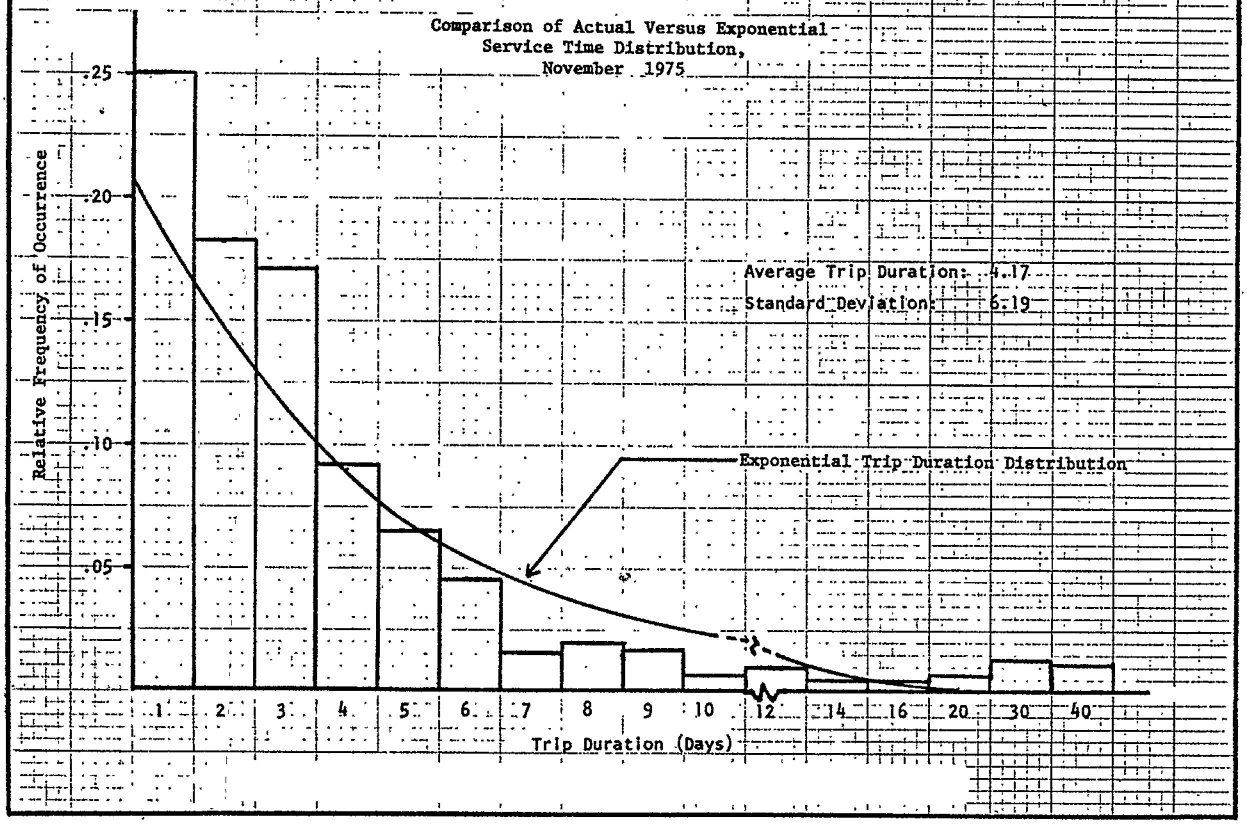
Hence, assuming, in the main, a null relationship between day of vehicle dispatch and service time, a composite trip time distribution was formulated for each month (See Illustration 4 for example). Two-sample comparisons of the composite monthly service time distributions via the Smirnov test resulted in significant deviations, however, and a higher level, synthetic annual trip time distribution was considered unacceptable. Furthermore, Kolmogorov and chi-square tests between individual monthly distributions and theoretically expected exponential and gamma distributions also failed to support congruence. Since it has been shown elsewhere that incorrect distributional specification may have a salient impact on service level requirements as a function of system intensity [3], it was decided that rather than risk the errors attendant to misspecification, the empirical trip time distributions would be retained in the simulation analysis.

The final phase of the simulation design process was concerned with the development of a trip mileage generator. Since a primary component of the equations describing total fleet costs were based on accumulated vehicle mileage, it was imperative that a reasonably precise trip distance function could be formulated.

Data were gathered for total trip mileage by length of trip (days). Scattergrams were first constructed to obtain a preliminary visual guide to function fitting; inspection of the plots suggested a generally linear relationship between total trip mileage and trip duration. A simple linear regression of total

ILLUSTRATION 4

Comparison of Actual Versus Exponential Service Time Distribution, November 1975



miles on trip days resulted in the following equation: $\hat{Y} = 54.9 + 106.3X$.[5] Although the scatter-diagrams indicated some tendency for the slope to decline with increasing trip time, numerous regression runs incorporating transformations of one or both variables as well as quadratic and cubic functions did not result in a better fit to the data.

Finally, it was decided to attempt to secure the best fit for the greatest proportion of trips. On an annual basis, over 85% of all dispatches resulted in trip lengths of five days or less. When trip mileage was regressed on this subset of trip data, the proportion of variance explained increased markedly: $\hat{Y} = 29.7 + 105.1X$.[6] Similarly, four other piecewise regressions were fit to discrete intervals of the remaining trip duration observations. These regressions generally resulted in poorer fits, primarily because of the large variance inherent in trips of long duration. It was felt, however, that since trips of short duration comprised the overwhelming majority of fleet dispatches, the least squares fit obtained for the "trip $\leq$ 5 days" subset was more than adequate for simulation purposes.[7]

## MODEL LOGIC AND VALIDATION

Once the various input distributions had been specified, a GPSS/360 (General Purpose Simulation System) model was developed to simulate the behavior of the dispatch fleet over time. GPSS was chosen primarily because of programming ease considerations. A brief description of the modeling logic follows.

The model was constructed around an implicit time interval of one day, with each run comprising one year of simulated time. Basically, as the simulation proceeds dynamically, vehicle requests are generated from distributions specific to the major demand periods as well as day of week. GPSS FUNCTION statements are employed to access the appropriate overarching demand which subsequently "calls" the day of week distribution via another FUNCTION statement. Where necessary, special holiday and quarterly class break functions are incorporated. Similarly, service times are sampled from appropriate monthly trip duration functions for each vehicle dispatched. Concurrently, operating mileage is calculated and accumulated for dispatch vehicles. Likewise, reimbursement mileage data are computed and maintained for demand requests that cannot be satisfied by the available fleet. Basic statistical data for average and maximum fleet utilization, simulated request rate and service time means and variances, etc., are dynamically collected, stored, and updated by the GPSS processor and MSAVEVALUE Blocks integral to the program.

To preclude potential bias in the response variables due to the "empty" conditions inherent at the start of the simulation, a number of pilot runs were conducted to determine the length of the transient state. It was found that approximately two months of simulated time was necessary before the system converged to the steady state; at this point in simulated time the output variables of interest generally tended to be symmetrically distributed about the mean response rate [2].

ILLUSTRATION 5

Actual Versus Simulated Number of Vehicles Requested by Month, Fiscal 1975

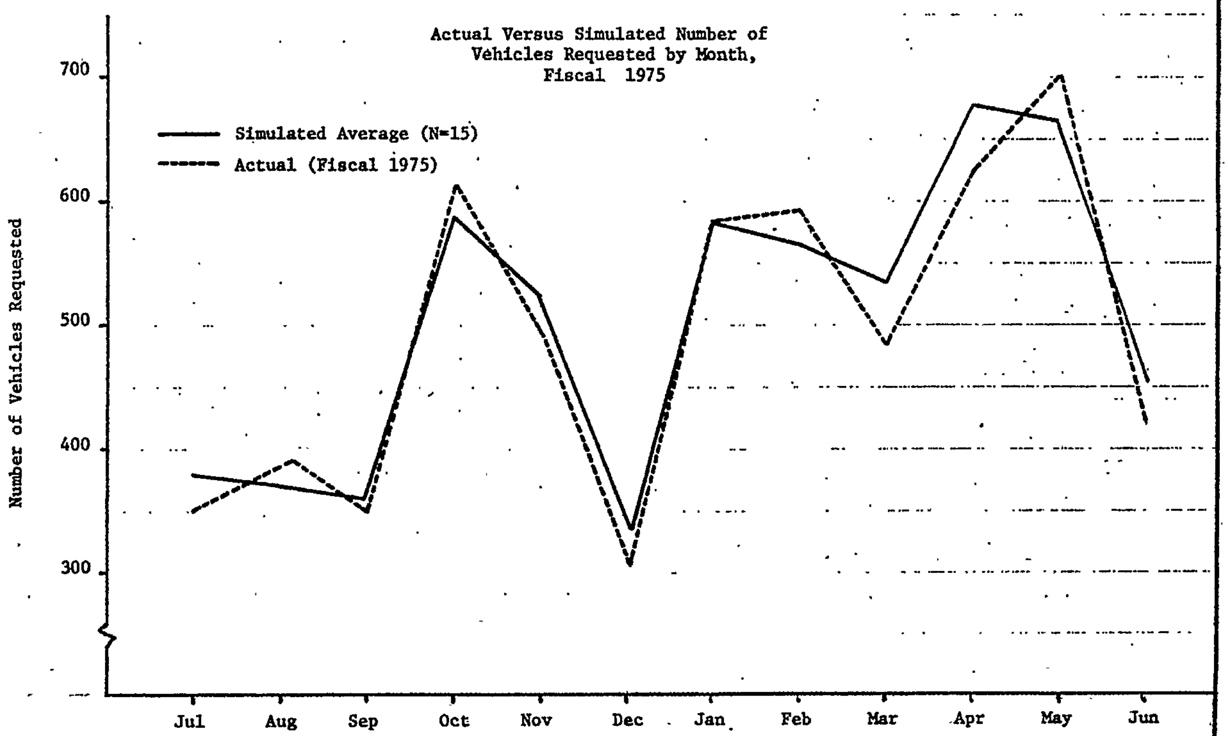
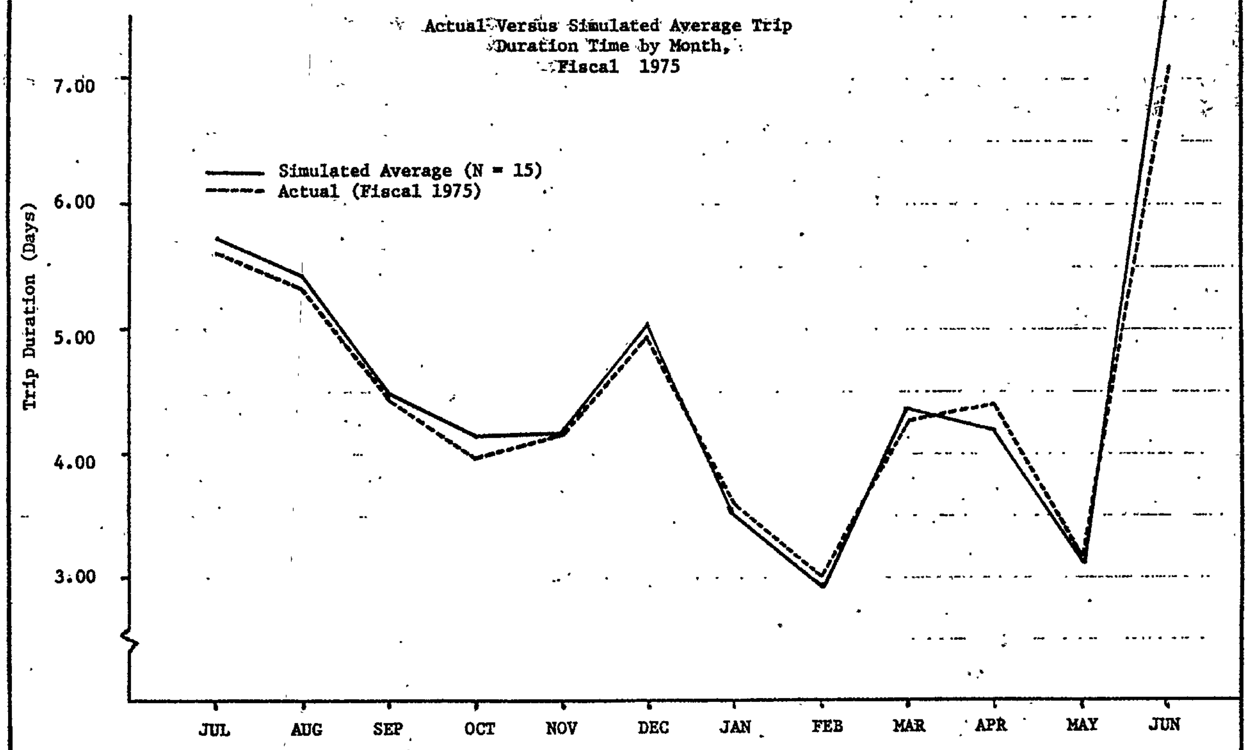


ILLUSTRATION 6

Actual Versus Simulated Average Trip Duration Time by Month, Fiscal 1975




A total of fifteen years of motor pool experience was generated for each of six alternative fleet size configurations. Correlated sampling procedures were employed to minimize the introduction of random variation into the simulation results and to increase the statistical confidence in alternative comparisons. Inspection of output streams suggests close congruence between actual and simulated demand and trip duration distributions (Illustrations 5 and 6). Furthermore, confidence interval construction indicates that reproduction of actual motor pool experience is accommodated within the simulation's structural capability (Table 1). Finally, although the motor pool did not normally segregate data for dispatch fleet mileage, discussions with motor pool personnel corroborated the reasonableness and general precision of the aggregate mileage generator.

## SIMULATION ANALYSIS

### OUTPUT STREAMS

Illustration 7 portrays the simulated fleet load potential based on fiscal 1975 request and trip duration data. As is apparent, the daily average number of vehicles dispatched remains well under 80 units except during April and May. Similarly, the "Monthly Maximum" plot indicates the absolute maximum number of vehicles dispatched for any day in a given month. Obviously, with a current effective fleet capacity approaching 135, the probability of exhausting the fleet is essentially nil. This is even more evident in Illustration 8 which shows the "service level" for the six fleet size configurations simulated. A reduction of effective fleet capacity by one-third still provides coverage of over 95% of all requests by the dispatch pool.

The simulated utilization for various fleet size configurations is shown in Illustration 9. As would be anticipated, the smaller the fleet, the greater the utilization. However, as the fleet is decreased the marginal increase in utilization declines. In addition, the variability in monthly utilization increases with smaller fleet sizes.

Although these output streams are useful in gaining understanding of the dispatch system's behavior, they do not, in isolation, permit identification of the minimum cost fleet size. In order to obtain an approximation of least-cost capacity, the simulation outputs must be juxtaposed with the cost structure of the motor pool's operation.

### STUDY EXTENSIONS, LIMITATIONS AND FUTURE REFINEMENTS

At the time of this writing, comprehensive analysis

TABLE 1

Confidence in Simulated Estimates of Aggregate System Parameters

**Aggregate Dispatch Vehicle Requests**

Simulated Results

Average Annual Requests = 6,029
Standard Deviation = 112
Number of Runs = 15

Confidence Interval

P (Lower Limit ≤ Parameter ≤ Upper Limit) = 1 − α
P (5964 ≤ μ ≤ 6093) = .95

Observed

Actual Annual Dispatch Requests = 5,984 (Fiscal, 1975)

**Average Fleet Service Time**

Simulated Results

Average Annual Service Time = 4.30
Standard Deviation = .13
Number of Runs = 15

Confidence Interval

P (Lower Limit ≤ Parameter ≤ Upper Limit) = 1 − α
P (4.23 ≤ μ ≤ 4.38) = .95

Observed

Actual Average Service Time = 4.30 (Fiscal, 1975)

is currently being conducted to determine the least-cost fleet size. Regressions are being fit to simulated pool mileage and reimbursement mileage as a function of fleet size. The best fit equations will then be entered into a total cost model which comprises ownership costs, operating costs, and reimbursement costs. Finally, this equation can be solved for the optimum fleet size via the calculus.

Policy decisions based on the preceeding analysis should be tempered with some caution for a number of reasons. First, since simulation inputs were based on fiscal 1975 data, generalization of output statistics to the present must remain tentative. Due to a marked decline in the rate of University budgetary growth and the invidious impact of inflation, squeezes in departmental travel budgets may have resulted in a decline in motor pool requests. Conversely, such conditions may well have stimulated a shift from air transportation to lower cost vehicular travel. In any event, only an up-to-date sampling of current experience will indicate whether the request pattern has significantly changed. Further, as noted previously, service time distributions were constructed very conservatively. The net effect of this bias is to overstate utilization statistics and understate service level. These factors should be evaluated prior to decision implementation; perhaps additional information would be warranted before significant reductions in fleet capacity are effected.

ILLUSTRATION 7

Simulated Maximum and Average Fleet Dispatches by Month, Fiscal 1975 (N = 15)

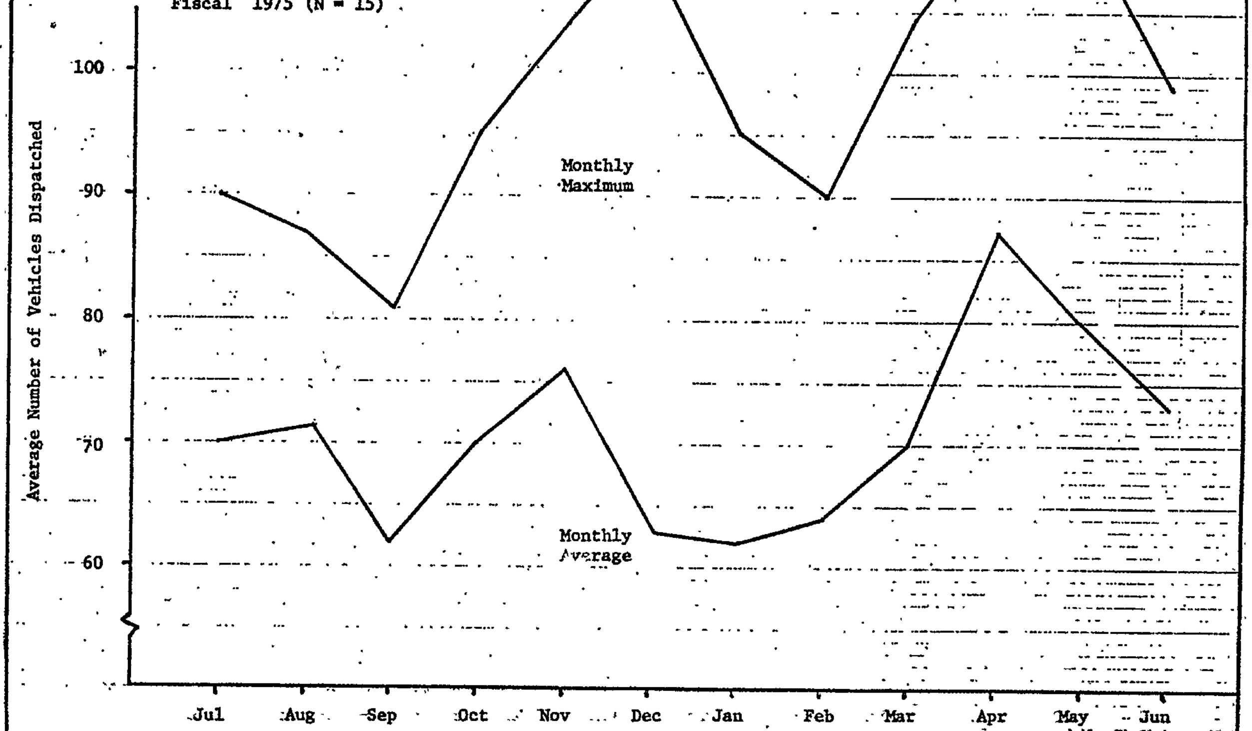
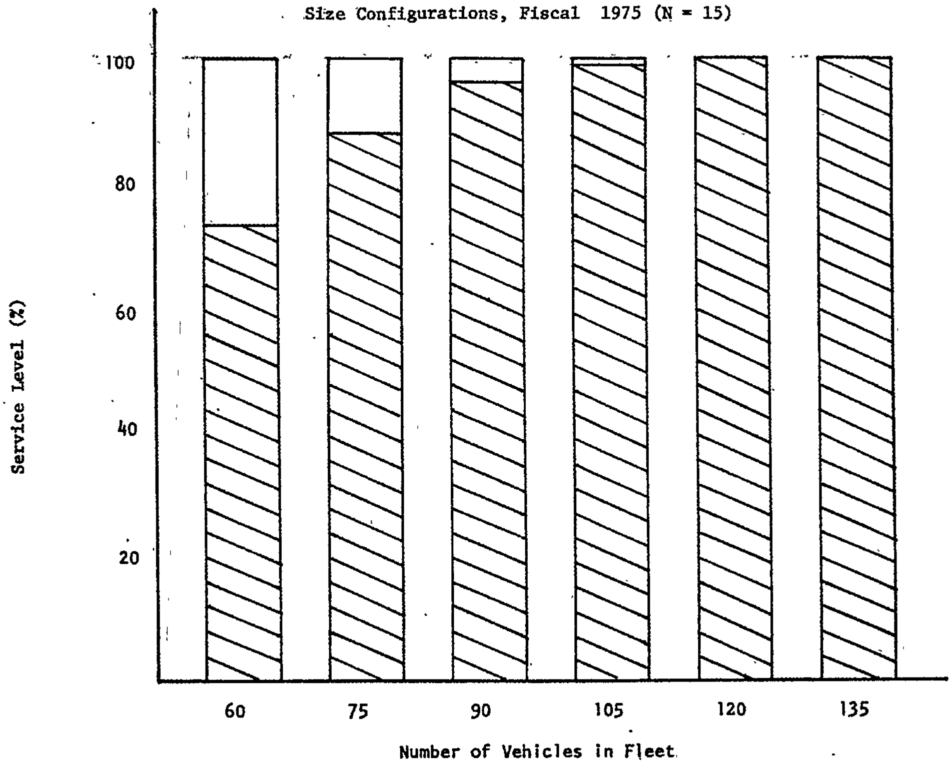


ILLUSTRATION 8

Simulated Service Levels for Alternative Fleet Size Configurations, Fiscal 1975 (N = 15)




Examination of the simulation outputs suggests greater efficiency may be realized via demand-phased acquisition and disposition of vehicles. A study of demand rate and service time patterns indicates peak fleet load occurs in the spring quarter. Procurement and disposition decisions might be timed so as to provide maximum capacity during this period. Simulation studies investigating several dynamic procurement strategies have been initiated; as yet results are inconclusive.

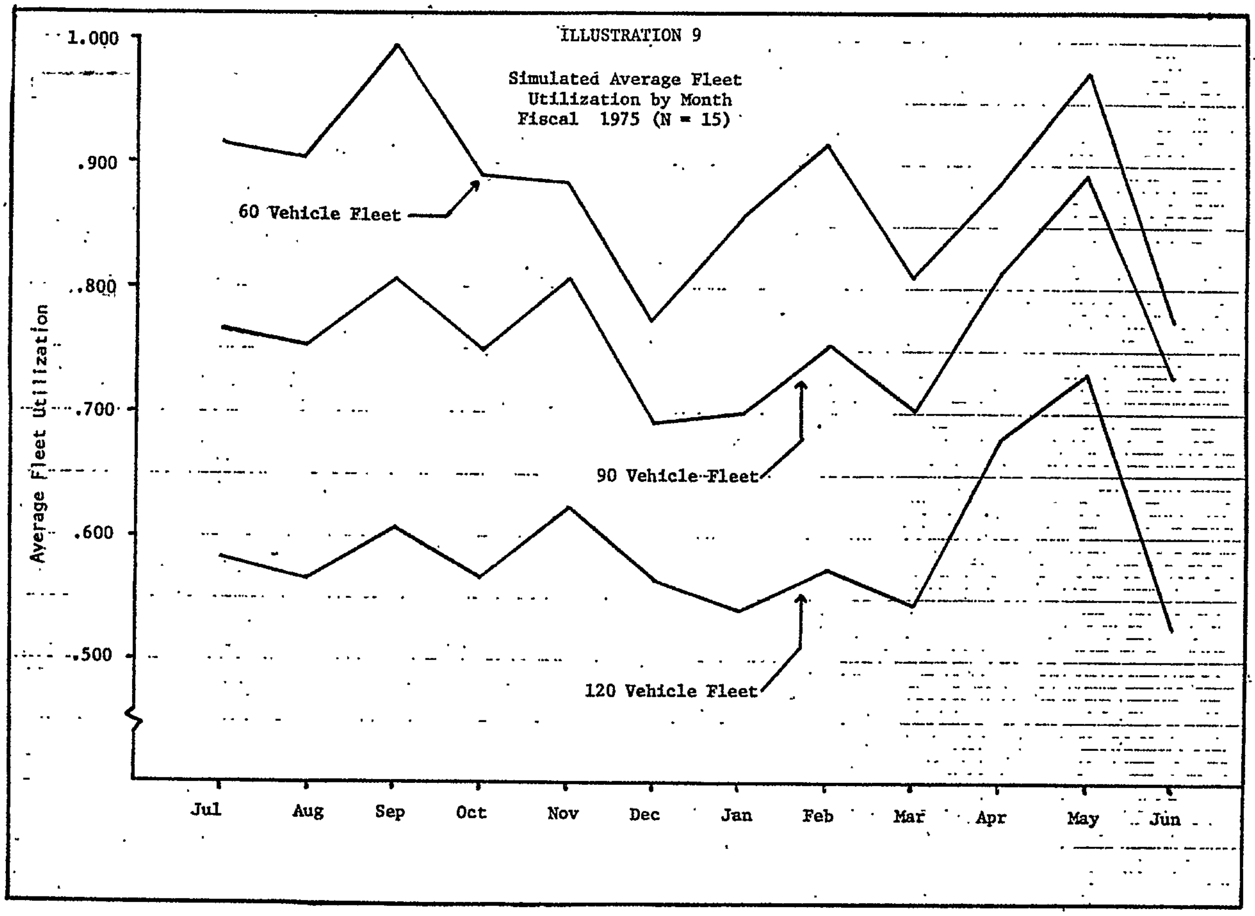
FOOTNOTES

1. Where i = 1, 2, ...5 and k = 1, 2, ...(6-i). Sundays were excluded from correlational analysis as the motor pool dispatch offices were generally closed and seldom were vehicle requests submitted.

2. An examination a sample return time data suggested that most vehicles are returned in early-to-mid-morning. Thus, at least four hours are available to service each vehicle.

3. A plausible explication for the failure of Friday's data to support the identical distribution hypothesis may be due to the high relative frequency of four-day service times. Since the vehicle dispatch office was not normally open on weekends, a trip length of four days seems intuitively more likely than service times of two or three days (i.e., returns on Saturday and Sunday). It was precisely at the four day cumulative frequency that the significant test statistic deviation occured vis-a-vis other weekdays.

4. Parameters used to generate the theoretical cumulative distribution were estimated from the observed data.




5. $R^2 = .44$; $F = 240.71$ (d.f. = 1,307); $p \leq .001$. Examination of residuals indicated that the assumption of error term homoscedasticity may have been violated. A weighted least squares, however, did not enhance the fit to the data.

6. $R^2 = .92$; $F = 681.81$ (d.f. = 1,58).

7. The four other regressions employed in the simulation:

$6 \leq \text{trip} \leq 10$: $\hat{Y} = 144.3 + 143X$; $R^2 = .12$
$11 \leq \text{trip} \leq 15$: $\hat{Y} = -4.2 + 74.5X$; $R^2 = .03$
$16 \leq \text{trip} \leq 20$: $\hat{Y} = -53.2 + 66.2X$; $R^2 = .02$
$\text{trip} \geq 21$: $\hat{Y} = -985.9 + 73.3X$; $R^2 = .46$

## REFERENCES


1. Conover, W. J. Practical Nonparametric Statistics (NY: John Wiley & Sons, 1971).

2. Emshoff, J. R., and R. L. Sisson. Design and Use of Computer Simulation Models (NY: MacMillian Co., 1970).

3. Raedels, A. R. "A Comparison of Theoretical Versus Empirical Interarrival and Service Time Distributions for Teller Staffing in Financial Institutions." Paper presented at Sixth Annual meeting, Midwest Region, American Institute of Decision Sciences.