PROCEEDINGS —

Twenty-third Annual Meeting

Volume XXIII • Number 1 1982

TRANSPORTATION RESEARCH FORUM

PROCEEDINGS —

Twenty-third Annual Meeting

Theme:

"Developing Concinnity in Transportation"

October 28-30, 1982
Fairmont Hotel
New Orleans, LA

Volume XXIII • Number 1 1982

TRANSPORTATION RESEARCH FORUM

Evaluation Models to Assess Fleet Vehicle Utilization

*by Robert E. Stammer, Jr.**

INTRODUCTION

IN TIMES PAST, fleet management consisted mainly of traditional motor pool activities such as the acquisition and maintenance of the organization's vehicles. Changes in employee travel demands, reduced fuel supplies, spiraling operating expenses, and double digit inflation are all factors which have contributed to added responsibilities for fleet managers. The idea of merely maintaining fleet vehicles is being replaced by management's desire to improve the productivity of each vehicle. A tendency to acquire a vehicle on the basis of its ability to handle a "peak" demand is no longer acceptable. Thus, fleet managers can benefit from evaluation techniques which quantify a vehicle's utilization and identify those vehicles which are insufficiently utilized.

The Union Carbide Corporation-Nuclear Division (UCC-ND) is a large private employer in Oak Ridge, Tennessee. UCC-ND has a contract to operate and manage three separate facilities on the Department of Energy (DOE) reservation. The three plants have approximately 15,000 employees and 1,400 vehicles are used to make numerous trips within and between the three plants. While intraplant trips are frequently short, interplant trips may be as short as 3 miles or longer than 12 miles.

A recent study of travel patterns and vehicle use for these 1,400 vehicles revealed that many vehicles appeared to be insufficiently utilized. Thus, quantifying the subjective term "insufficiently utilized" and developing vehicle utilization evaluation models could, and did, lead to vehicle policy changes resulting in capital, operating, and fuel cost savings.

The initial fleet management study (1) included a variety of research activities. Specific tasks included:

1. a literature review of fleet management,
2. interviews with fleet management consultants,
3. interviews with fleet managers,
4. collection of origin-destination data for approximately 1,400 fleet vehicles for one month,
5. development of vehicle utilization evaluation models,
6. assessments of various fleet management strategies, and
7. presentation of final study recommendations.

Individuals interested in much more detailed and quantitative information should consult the original study.

Facts presented in this paper will highlight the negative and positive aspects of four vehicle utilization evaluation models. The four models use mathematical techniques such as set theory, discriminant analysis, linear regression, and logarithmic transformations. The paper concludes with a review of the benefits resulting from using the recommended analytical procedures.

THE PURPOSE OF VEHICLE UTILIZATION EVALUATION MODELS

Vehicle utilization evaluation models allow fleet managers to evaluate and more effectively use the vehicles for which they are responsible. It should be emphasized that vehicle utilization evaluation procedures are intended to be only screening tools or methods for targeting insufficiently utilized vehicles which are candidates for further management actions. The vehicle utilization evaluation models merely identify the so-called low utilization vehicles. Once insufficiently utilized vehicles are identified, fleet managers must apply other criteria to make final decisions regarding the continued role of each vehicle. Essentially, quantitative data regarding vehicle utilization can assist fleet managers in making decisions about both individual vehicles and appropriate total fleet size, and can also help justify the recommendations presented to upper level management.

DATA COLLECTION

Figure 1 shows the research instrument used to collect origin-destination data and other travel information from employees for one month. The effort resulted in over 145,000 vehicle utilization trip records. The trip information was keypunched, tested, and determined to

*Vanderbilt University, Nashville, TN

UNION CARBIDE VEHICLE USAGE LOG

License No. ______________________ Division Vehicle is Assigned To: ______________________

Please complete the following log accurately and completely to assist us with our mandatory energy conservation program. Record each leg of any trip from one location (Trip Origin) to any other location (Trip Destination) for all trips. Thank you for your assistance.

				TRIP START			TRIP END				
DATE	DISPATCHER'S SIGNATURE (if dispatched)	ESTIMATED TIME OF RETURN	OPERATOR'S SIGNATURE	START TIME	BEGINNING ODOMETER	TRIP ORIGIN	END TIME	ENDING ODOMETER	TRIP DESTINATION	TOTAL NO. OF VEH. OCCUPANTS	TRIP PURPOSE CODE*

*1 - Personnel or Materials Pickup
2 - Personnel or Materials Delivery
3 - Work Assignment
4 - Meeting
5 - Security Patrol
6 - Return Trip
7 - Other (Please Specify)

FIGURE 1

be an excellent record of fleet vehicle utilization. The original UCC-ND study (2) should be reviewed if additional facts are desired about data collection and analysis procedures.

Three vehicle utilization measures were selected for testing with various analytical procedures. The three utilization variables used in subsequent models were:

1. vehicle trips,
2. vehicle miles, and
3. vehicle use (i.e., minutes a vehicle was in service).

Various vehicle utilization evaluation models used the number of trips, miles traveled, and minutes of use over a month's period to determine if individual vehicles were being sufficiently utilized.

SELECTION OF ANALYTICAL TECHNIQUES

Techniques such as simulation, linear programming and others were considered with the available vehicle utilization variables, and the development of useful evaluation models using these techniques appeared unlikely. Therefore, they were eliminated from further considerations.

The four evaluation techniques selected for further study were:

1. variable rankings and set theory,
2. discriminant analysis,
3. regression analysis, and
4. logarithmic transform functions.

Although these analytical techniques are certainly not new, there are several reasons why these procedures appeared appropriate for evaluating vehicle utilization. A procedure of ranking independent variables and using set theory to identify underutilized vehicles was attractive because it is straightforward, appears to be relatively simple, and possesses inherent flexibility. Discriminant analysis procedures appeared promising because they use statistical analyses of independent variables (i.e., utilization variables) to distinguish between sufficiently and insufficiently utilized vehicles. Regression analysis was selected for further study because of its prospective use with vehicle utilization indexes and proven success in other analyses.

The ease with which regression analyses can be performed, the flexibility of using various input variables, and the potential use with existing fleet data are other attractive features of regression analyses. Similarly, logarithmic transform functions were also selected because of their potential use with vehicle utilization indexes. Since logit models have been used extensively to forecast transit riderships and logarithmic transform functions can dampen the effects of widely varying utilization variables, further investigations were appropriate.

DEVELOPMENT OF VEHICLE UTILIZATION EVALUATION MODELS

The "1/n by m Ranking" Technique

This technique consists of both numerical rankings and comparison procedures to identify low utilization vehicles. First, selected vehicle utilization factors are ranked by ascending magnitude for each vehicle. Then comparisons are performed and vehicles ranking below designated percentile values for all utilization factors become candidates for removal or reassignment because of their low utilization.

The term "1/n by m Ranking" technique is a generalized expression indicating a wide range of analyses. The "1/n" or "one over n" portion signifies the cutoff point for a ranked utilization factor. Regardless of whether the lower 1/5, 1/4, 1/3, or some other portion is used, the proper cutoff value for each utilization factor should be determined by analyzing the ranked original values. This investigation is performed in an attempt to locate an appropriate cutoff point where a discrete break occurs in the ranked values for each utilization factor. Graphical techniques may be especially helpful in deciding if noticeable changes exist. The "m" in the expression indicates the number of utilization factors used in the analysis technique. The number of utilization variables used will depend on the availability and appropriateness of existing variables.

Thus, the "1/n by m Ranking" technique classifies a vehicle as being insufficiently utilized if:

$$(r_1 < R_1) \text{ and } (r_2 < R_2) \text{ and } \ldots \text{ and } (r_m < R_m)$$

where: m = number of utilization variables;
r_j = observed rank of any variable j; and
R_j = maximum allowable rank or cutoff value determined for variable j.

A "1/n by m Ranking" analysis can be performed manually for smaller fleets. However, these procedures become unwieldy as fleet sizes increase and are accomplished more quickly and precisely through the proper use of existing statistical computer programs.

The "1/n by m Ranking" technique uses set theory and can be visualized graphically with a Venn diagram. Each utilization factor can be described as a set of ranked vehicles. The intersection of the sets of vehicle rankings by utilization factor, or the darker, shaded are shown in Figure 2, identifies those vehicles with insufficient utilization according to all utilization factors. Thus, these vehicles are identified as candidates for removal or reassignment.

The "1/n by m Ranking" technique provides considerable user flexibility. For example, the user has the option to use two, three, or more utilization factors and cutoffs or breakpoints in the percentile rankings are also determined by the analyst.

By studying the data before proceeding with detailed analyses, the analyst will obtain a better understanding of which combination of utilization factors and percentiles should be used in the final model. Trends and variations in data distributions may also be detected. In fact, the fleet manager can perform a sensitivity analysis of various models by varying the number of utilization

VENN DIAGRAM OF "1/n by m Ranking" TECHNIQUE

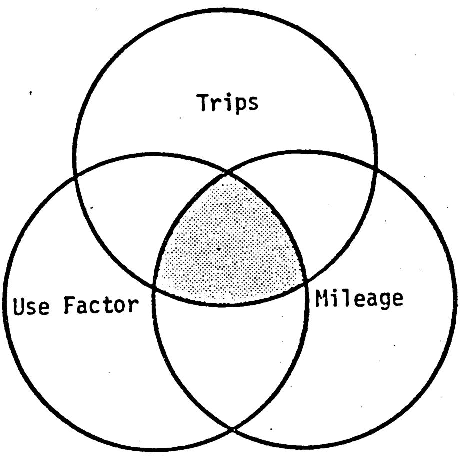

FIGURE 2

factors, weighting utilization factors differently, and changing percentile breakpoints. This provides the manager with increased knowledge about the existing relationships between a model's input variables and the identification of fleet vehicles being insufficiently utilized.

Discriminant Analysis Techniques

Discriminant analysis attempts to statistically distinguish between two or more groups of cases. Researchers want to be able to "discriminate" between the groups in the sense of being able to tell them apart and classify likely group membership of a case when the only information known is the case's values for the discriminating variables.

Discriminant analysis seeks to determine statistically how one or more independent variables can be used to discriminate among different categories of a nominal dependent variable. Linear combinations of independent variables that best distinguish between cases in the categories of the dependent variable are found. These "discriminant functions" are of the form:

$$D_i = d_{i1} Z_1 + d_{i2} Z_2 + \dots + d_{ip} Z_p$$

where D_i is the score on discriminant function i, the d's are weighting coefficients, and the Z's are the standardized values of the p discriminating variables used in the analysis.

Two discriminant analysis programs, BMDP7M from the BMD (Biomedical) Computer Programs Package and the DISCRIMINANT procedure of SPSS (Statistical Package for the Social Sciences) were used in this research. Although the discriminant analysis calculations are similar in both programs, there are slight variations in data output and graphic displays between the two programs. Therefore, in order to more completely understand discriminant analysis, the characteristic results unique to each program were used to obtain answers to individual questions. Both BMDP7M and SPSS were run by interfacing them with SAS (Statistical Analysis Systems) job control language. Programming BMDP7M and SPSS via SAS was reasonable since the initial survey data from UCC-ND was coded as a SAS dataset. Thus, three programs were needed to perform the required analyses.

To be consistent with earlier research of "1/n by m Ranking" techniques, both two and three variable classifications were performed and vehicles were classified as insufficiently utilized or low utilization vehicles (i.e., placed in Group I) if and only if each discriminating utilization variable was contained within the 33rd percentile of its respective ranking. Any vehicle having one or more utilization variables greater than the 33rd percentile was placed in Group II which contained vehicles classified as being sufficiently utilized.

Ideally, these cutoff points should result from observed distribution differences rather than from arbitrarily chosen values. The selected points were feasible, however, since clearly observable cutoffs were not present and agreement with other analysis techniques was desired.

For discriminant analysis to minimize incorrect classifications, the product of two prior probabilities (i.e., the first prior probability associated with initially selecting or identifying a case from a specific group multiplied by the second prior probability associated with assigning a particular case to a specific group of known size) must correspond to the model's data. Thus, the model's resulting tendency to misclassify a certain number of cases relates to the product of the two prior probabilities.

A series of tests on both two and three variable classifications were performed. The two most logical models were selected for further testing. Analyses of misclassified vehicles by the two and three variable discriminant analysis models suggested possible reasons for the misclassification of certain vehicles. Misclassifications appear to be much more attributable to deviant or atypical cases than to a lack of discriminatory power by the models. In fact, the two models apparently misclassify only 1.94 percent of all tested vehicles once atypical cases are removed from consideration.

Although discriminant analysis models appear to be reliable when classifying a vehicle into utilization groups, there are several drawbacks to using such models. For example, a fleet manager must have the following conditions before performing a discriminant analysis:

1. computer capabilities;
2. statistical knowledge or available consulting expertise; and
3. low intercorrelations between each pair of utilization variables.

Regardless of these drawbacks, discriminant analysis techniques may be used with increased confidence that distinguishable differences in vehicle utilization, in fact, do exist. The following section describes how two utilization indexes were developed for use with regression and logarithmic transform models to identify a vehicle's degree of utilization.

Indexing Techniques

The "1/n by m Ranking" and discriminant analysis models use individual criterion for each utilization variable. Two other vehicle utilization evaluation models using linear regression and logarithmic transformations were also developed. The latter two models classify the utilization of a vehicle by determining a combined index score using all utilization variables and then comparing the combined index score against a single, previously determined cutoff value to classify vehicles as being sufficiently or insufficiently utilized.

Two index scores, INDEX2 and INDEX3, were developed from vehicle utilization data using two and three independent variables, respectively. By using integer values representing the ranking of three utilization variables instead of the actual number of trips, miles traveled or minutes of vehicle use, the variability between differing values is decreased to a linear relationship. Thus, INDEX2 and INDEX3 represent monotonically increasing variables which are formed from each of the original independent variables. The variability is dampened further by dividing by the sum of all maximum variable rankings so that values lie between 0 and 1. Analysis results illustrated that vehicles classified as low utilization vehicles by a "1/n by m Ranking" technique will also be identified as low utilization vehicles by comparable indexing procedures. The converse, however, is not always true. The phenomenon whereby a vehicle with a low utilization index score was not classified as a low utilization vehicle by the "1/n by m Ranking" or discriminant analysis models has some logical justification. Intuitively, a technique which requires each variable to satisfy an individual criterion will be more limiting than an indexing technique which only requires a combined index value for all variables to satisfy a single, comparable criterion.

Even though the indexing techniques appear to be less stringent, INDEX3 and INDEX2 illustrate several inherent characteristics which make the two indexes useful as measures of vehicle utilization. In addition to index scores being between 0 and 1, the index scores closely parallel percentile values. This close association with definable values provides researchers with an immediate feel for the model's sensitivity to variations in cutoff values. Information about the distribution and normality of each index was also studied. Therefore, the two indexes were determined to be useful indicators of vehicle utilization and they were used in the development of multiple linear regression and logarithmic transform models presented in the next two sections.

Regression Analysis Models

Regression analysis models seek to describe the extent, direction, and strength of relationships between one or more independent variables and a single, continuous dependent variable. The continuous dependent variable represents a numerical expression of events or conditions which researchers desire to explain through existing knowledge of an independent variable or variables.

Regression analysis is a general term which includes many types of regression equations. For example, a regression analysis model may be linear or nonlinear, depending on whether original first power or transformed independent variables are used in the regression equation.

Each of the previously defined vehicle utilization indexes, INDEX3 and INDEX2, was shown to be a continuous variable which could serve as an indicator of an individual's utilization. Thus if values of independent variables from the collection of sample data could be used to replicate a vehicle's utilization index by a regression equation, fleet managers and other decision makers would have a useful technique for assessing vehicle utilization. A highly predictive regression equation could be used to continually monitor fleet vehicles. Once a regression examination based on existing vehicle usage is established, it may be possible to substitute projected values of the independent values into the equation to determine the anticipated level of utilization of an additional vehicle. Such forecasts, however, are contingent upon both the continued existence of the same fleet characteristics present during model calibration and the accuracy of current values for the independent variables.

Simultaneous investigations of both INDEX3 and INDEX2 were performed. Thirteen independent variables were tested initially. Approximately fifty regression runs with various variable combinations were performed using the General Linear Model (GLM) procedure from the SAS battery of programs. Residual plots were studied to evaluate nonrandom trends and some models were rerun in an attempt to improve them. Eight regression equations were produced with squared multiple correlation coefficients (R^2) equal to or greater than .600.

Reviews of independent variable intercorrelations, the logicality of mathematical signs, the size of intercept values, F-tests for each equation, partial F-tests relating to the contribution of each additional independent variable added to a model, and residual plots resulted in two final equations being selected. These two equations are:

$$\text{INDEX3} = .001 \text{ TRIPS} + 1.70 \text{ MILES/10K} + 4.04 \text{ MINS/100K}$$

$$\text{INDEX2} = .001 \text{ TRIPS} + 3.13 \text{ MILES/10K}$$

where the units of two independent utilization variables, mileage and time usage, were modified slightly. There are noticeable similarities between these two final equations.

Results obtained from these two regression equations, as well as results from logarithmic transform equations presented in the next section, are compared with earlier model results in a later section of this paper. Before these model comparisons are presented, however, the next section presents information about the development of logarithmic transform models.

Logarithmic Transform Models

Logit models or multiple logistic functions are used to obtain a probability between 0 and 1 which relates to the likelihood of an event occurring. Logit models can be expressed by the following general equation:

$$P(Y) = \frac{e^{U(x)}}{1 + e^{U(x)}}$$

where:

$P(Y)$ = probability associated with the occurrence of a particular event,

$U(x)$ = a function of an independent variable or variables which is used to forecast the probability of occurrence of the dependent event, Y.

In the field of transportation, logit models have been used specifically to forecast transit ridership. Studies have investigated the probability that trip makers will choose transit or automobile travel, depending on their knowledge of costs, levels of service, and other value tradeoffs. According to the values trip makers associate with each independent variable in the function U(x), it will vary from $-\infty$ to $+\infty$ while the value of P(Y) increases monotonically from 0 to 1.

Since INDEX3 and INDEX2 are continuous variables whose values also vary between 0 and 1 depending on vehicle utilization, several logarithmic transform models were developed and tested to determine their appropriateness for predicting vehicle utilization.

Normal probability plots of INDEX3 and INDEX2 values produced by SAS computer runs were compared with values resulting from approximately 15 logarithmic transform functions. From these comparisons, a single function appeared to be the most promising. The function selected for further testing was:

$$\frac{1 - e^{-U(x)}}{1 + e^{-U(x)}}$$

U(x) was again defined to be a function of an independent variable or variables.

After performing a linear transformation on the above equation to expedite analyses and reduce computer expenses, six logarithmic transform models from an initial set of ten were selected for further testing.

ASSESSMENT OF FOUR VEHICLE UTILIZATION EVALUATION MODELS

To briefly review, the four techniques selected for model development were "1/n by m Ranking" procedures, discriminant analyses, linear regression equations, and logarithmic transform equations. Earlier sections have presented an overview of the procedures used in developing each type of model. This section evaluates the appropriateness of model results and compares the various models. It is difficult if not impossible, to compare all four models simultaneously due to the differing assumptions and criteria associated with each. Therefore, model results are presented using paired comparisons of individual models or types of models.

There was little analytical evidence to distinguish between linear regression and logarithmic transform models. It was noted, however, that all models relying on combined indexes appear to be less stringent in terms of identifying low utilization vehicles than models requiring vehicles to satisfy individual criterior for each utilization variable.

This assessment is based on tests of

the four types of models using UCC-ND vehicle data for one large industrial plant, K-25. Linear regression and logarithmic transform models classified 208 and 90 vehicles, respectively, as being low utilization vehicles while "1/n by m Ranking" and a discriminant analysis model identified only 59 and 43 low utilization vehicles, respectively. Thus, linear regression and logarithmic transform models were inexacting and considered inappropriate for the most accurate evaluations.

No independent "correct" groupings of sufficiently and insufficiently utilized vehicles exist that can be used as standards for comparing results from the two remaining techniques, discriminant analysis and the "1/n by m Ranking" technique. Intuitively, statistical analyses using discriminant functions should provide what might be considered the "best" vehicle utilization classifications. Since standard classification groupings do not exist, however, insufficiently utilized vehicles were defined a priori as those found in the 33rd percentile ranking for each utilization criterion. All other vehicles were classified as being sufficiently utilized. The results of these comparable analyses were informative.

Analyses of both techniques were performed by using two and three utilization variables, respectively. The results from comparable analyses show strong similarities between vehicle classifications. Two and three variable discriminant analyses showed that 77 percent and 73 percent of the same vehicles were classified as being insufficiently utilized. The similarities between the results of the two techniques were encouraging. Secondly, the two and three variables discriminant analysis results, respectively, differed from comparable "1/n by m Ranking" results by misclassifying only 5 percent and 4 percent of all vehicles as being sufficiently utilized. These small differences in classification results between the two techniques reinforce the fact that the results from the two analysis techniques are very similar.

The comparisons between the two remaining techniques resulted in the conclusion that the "1/n by m Ranking" technique is the most universally applicable. Although discriminant analysis procedures were shown to be significantly better than mere chance in correctly classifying a vehicle according to its utilization, there are several reasons why this technique was not selected as the most applicable. Detailed analyses requiring statistical and computer knowledge, possible intercorrelations between input variables and similar results from a simpler procedure were identified drawbacks to the use of a discriminant analysis technique.

The positive aspects of the "1/n by m Ranking" technique exceed those of the other techniques. The technique provides valuable information without requiring excessive amounts of time and effort and the results are very similar to those from more complex statistical analyses. The analyst has considerable flexibility in selecting utilization variables, choosing cutoff points, and determining the number of insufficiently utilized vehicles. The underlying process is easy to understand and may be more meaningful to fleet managers than some of the more mathematical and perhaps abstract models. Due to the nature of these models, intercorrelations between utilization factors are less likely to affect the model results. Although the "1/n by m Ranking" technique does not provide a finite numerical or quantitative indicator regarding the degree of utilization as do several other models, the analyst can readily determine the impact that various changes in utilization factors will have on a vehicle's classification status. The "1/n by m Ranking" technique also shows wide versatility in the types of input data which it can use.

EFFECTIVE USE OF "1/N BY M RANKING" MODELS

General Guidelines

Regardless of the types of data used, the "1/n x m Ranking" technique is easy to adapt to any fleet operation. The number of vehicles within the fleet and utilization factors in the model will determine the appropriateness of manual or computer data management. As a guideline for the practitioner, computerized data management techniques are strongly recommended if the product of the number of vehicles times the number of utilization variables exceeds 100 or the number of utilization variables alone exceeds 3.

If computer techniques are used, there are many statistical utility programs such as SAS and others which will readily rank, assign ascending rank order values, and then compare individual vehicle data. Vehicle utilization assignments are made on the basis of these analyses. A similar sequence of events is followed for manual analyses.

An Actual Application

Since the "1/n by m Ranking" technique was found to be the most promising of the four techniques investigated, this modeling technique was used to evaluate the 1,400 vehicles of the large 3-plant industrial complex. The ease with which analyses may be performed allowed researchers to study vehicle utilization by performing a series of model runs. Changing utilization variable indexes and varying cutoff percentiles were used in various combinations to identify underutilized fleet vehicles. A total of 133 vehicles or approximately 9.5 percent of all fleet vehicles were identified as candidates for reassignment or removal.

After studying the research findings, these 133 vehicles were removed from the fleet by management. A follow-up investigation showed that removal of these vehicles:

1. had no significant effect on productivity,
2. would save approximately 6,000 gallons of fuel per year, and
3. reduce capital expenditures for replacement vehicles by 900,000 dollars over the next five years.

SUMMARY

The versatility, ease of application, and demonstrated success of "1/n by m Ranking" models are encouraging. This straightforward technique of evaluating vehicle utilization can be very beneficial to both fleet managers and other fiscal administrators.

Two other areas of potential research were also identified. First, more research is needed to study how fleet managers can use this newly developed model with existing fleet data. This would eliminate, or at least greatly diminish, the requirements of extensive data collection. Secondly, the "1/n by m Ranking" technique may have broad potential applications in answering both transportation and non-transportation related questions.

REFERENCES

1. Stammer, R. E., *Effective Use of Fleet Vehicles Through an Analysis of Vehicle Utilization Evaluation Models and Trip Characteristics*, Transportation Center, University of Tennessee, April 1982, 250 pp.
2. Stammer, R. E., and Mundy, R. A., *Improving Transportation / Energy Efficiency at Union Carbide Corporation-Nuclear Division, Oak Ridge, Tennessee*, Transportation Center, University of Tennessee, February 1981, 199 pp.