# NAVAL RESEARCH LOGISTICS QUARTERLY



MARCH 1980
VOL. 27, NO. 1


OFFICE OF NAVAL RESEARCH

NAVSO P-1278

# NAVAL RESEARCH LOGISTICS QUARTERLY



The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics will publish research and expository papers, including those in certain areas of mathematics, statistics, and econom relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.


The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, Ju September, and December and can be purchased from the Superintendent of Documents, U.S. Government Prin Office, Washington, D.C. 20402. Subscription Price: $11.15 a year in the U.S. and Canada, $13.95 elsewhere. Cos individual issues may be obtained from the Superintendent of Documents.





Issuance of this periodical approved in accordance with Department of the Navy Publications and Printing Regulati P-35 (Revised 1-74).

# ANALYZING AVAILABILITY USING TRANSFER FUNCTION MODELS AND CROSS SPECTRAL ANALYSIS*

Nozer D. Singpurwalla

*The George Washington University*
*School of Engineering and Applied Science*
*Institute for Management Science and Engineering*
*Washington, D.C.*

**ABSTRACT**

If we look at the literature of reliability and life testing we do not see much on the use of the powerful methods of time series analysis. In this paper we show how the methods of multivariate time series analysis can be used in a novel way to investigate the interrelationships between a series of operating (running) times and a series of maintenance (down) times of a complex system. Specifically, we apply the techniques of cross spectral analysis to help us obtain a Box-Jenkins type transfer function model for the running times and the down times of a nuclear reactor. A knowledge of the interrelationships between the running times and the down times is useful for an evaluation of maintenance policies, for replacement policy decisions, and for evaluating the availability and the readiness of complex systems.

## 1. INTRODUCTION AND SUMMARY

The investigation reported here was undertaken to determine if a stochastic interrelationship exists between the runnng times and the down times of the Robinson Nuclear Power Plant. The data was provided to us by the Probabilistic Analysis Staff of the Nuclear Regulatory Commission (NRC). Our goal was to understand what the data were telling us about the relationship between the series of running times and the series of down times. One way of achieving this goal is to obtain a Box-Jenkins [1] type of "transfer function model" between the running times and the down times. The interpretation and uses of a transfer function model for the situation considered here are discussed in Section 1.1.

A first step in the analysis of the data was its careful screening. This was done in order to eliminate those observations that were judged to be questionable or that had arisen under unusual circumstances. Such observations introduce spurious autocorrelations and cross correlations, and thus tend to obscure the identification of a simple relationship that may exist between the running times and the down times.

We would like to emphasize that, for an analysis of data of the type discussed here (often referred to as "messy data"), an examination and screening of the data prior to model building are important preliminary operations. If one neglects to perform these operations, one may face the frustrating task of attempting to fit several transfer function models, none of which may be satisfactory.

*Jointly Sponsored by Contract AT(49-24)-0211 Nuclear Regulatory Commission and Program in Logistics Contract N00014-75-C-0729 Project NR 347 020 Office of Naval Research.

In Figures 1.1 and 1.2, we display a time sequence plot of the screened down times $X_t$ and the corresponding screened running times $Y_t$, $t = 1, 2, \ldots, 28$. Note that $X_1$ represents the first down time, $X_2$ the second down time, and so on, whereas $Y_1$ denotes the first running time, $Y_2$ the second running time, and so on. Note also that the two plots are not drawn to the same scale. In Figure 1.3 we indicate the relative positions of the $X_t$'s and the $Y_t$'s, $t = 1, 2, \ldots$.

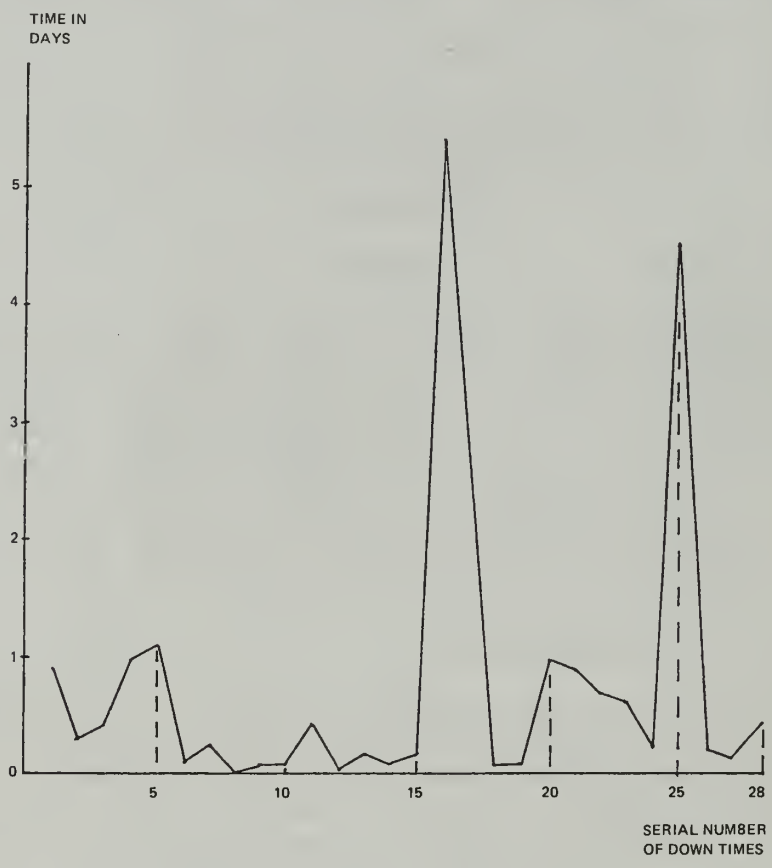


FIGURE 1.1. Time series plot of screened down times

In transfer function model building observations must be considered in pairs. In our case, the paired observations will be

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_{28}, Y_{28}).$$

In considering the above pairs, we will have to bear in mind that $X_t$ precedes $Y_t$ in chronological time.

In Table 1.1 we present the actual values of the screened down times $X_t$ and the corresponding screened running times $Y_t$, $t = 1, 2, \ldots, 28$.

If changes in a series of observations $Y_t$, $t = 1, 2, \ldots$, tend to be anticipated by changes in another series of observations, say $X_t$, $t = 1, 2, \ldots$, then $X_t$ is said to be a *leading indicator* of $Y_t$. In our case it is reasonable to assume that the down times $X_t$ are a leading indicator of the running times $Y_t$. Based upon this consideration, and together with an analysis of the available data, albeit *an insufficient amount*, the best transfer function model we have identified and fitted is given as

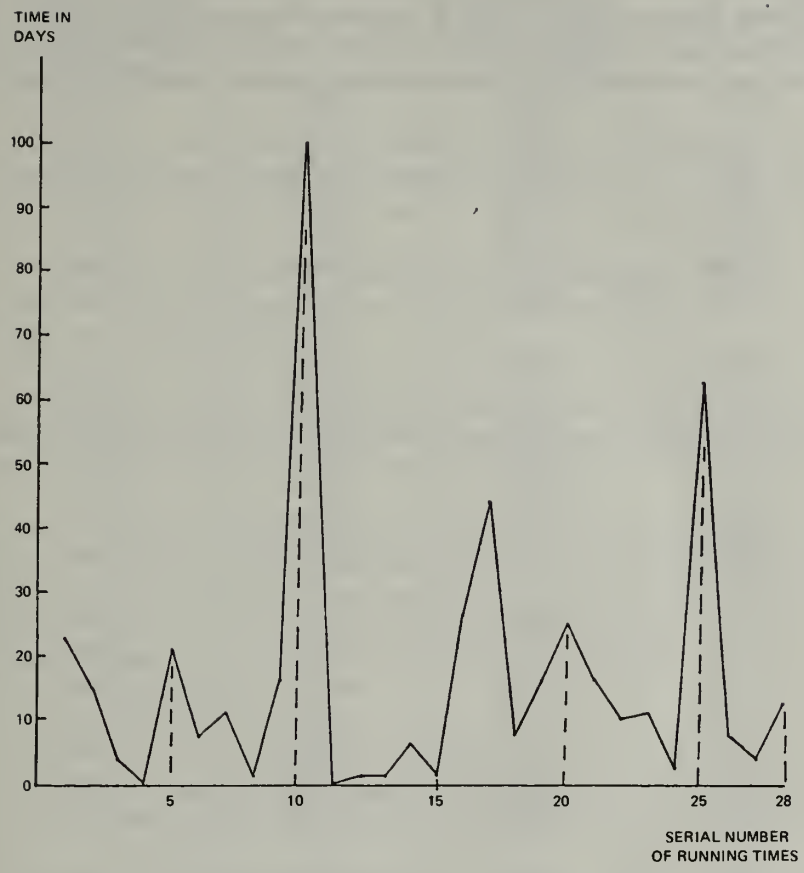


FIGURE 1.2. Time series plot of screened running times

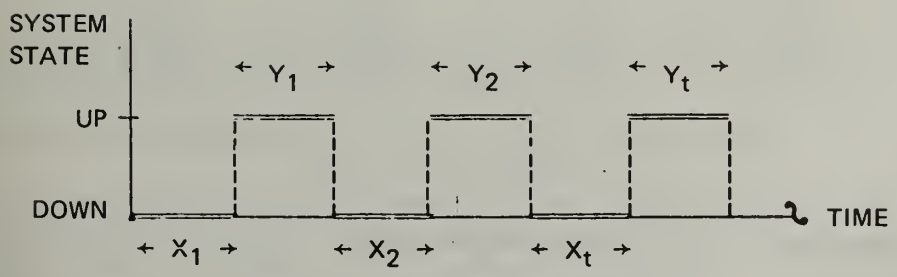


FIGURE 1.3. State of the system versus time

TABLE 1.1 — *Screened Down Times and Running Times of the Robinson Nuclear Power Plant*

| Down Times (days): $X_t$ | Prewhitened Down Times (days): $\alpha_t$ | Runnng Times (days): $Y_t$ | Prewhitened Running Times (days): $\beta_t$ |
|---|---|---|---|
| 0.87 | 0.07 | 22.83 | 6.40 |
| 0.26 | −0.54 | 15.75 | −1.43 |
| 0.42 | −0.30 | 3.50 | −11.55 |
| 0.95 | −0.09 | 0.00 | −14.50 |
| 1.13 | 0.19 | 21.00 | 3.88 |
| 0.06 | −0.78 | 8.00 | −13.17 |
| 0.19 | −0.47 | 11.00 | −5.09 |
| 0.02 | −0.84 | 1.00 | −16.82 |
| 0.08 | −0.81 | 16.00 | −1.18 |
| 0.06 | −0.86 | 101.00 | 80.84 |
| 0.40 | −0.55 | 0.58 | 27.76 |
| 0.01 | −0.99 | 1.00 | 2.97 |
| 0.15 | −0.75 | 0.83 | −10.46 |
| 0.08 | −0.89 | 6.90 | −11.44 |
| 0.17 | −0.80 | 0.83 | 15.85 |
| 5.48 | 4.51 | 27.00 | 8.86 |
| 2.81 | 1.21 | 44.00 | 22.01 |
| 0.06 | −0.13 | 8.00 | 11.43 |
| 0.08 | 0.13 | 17.00 | 7.25 |
| 0.98 | 0.30 | 25.00 | 8.48 |
| 0.87 | 0.04 | 17.00 | −0.57 |
| 0.70 | −0.03 | 10.30 | −3.64 |
| 0.63 | −0.11 | 11.36 | −3.64 |
| 0.19 | −0.59 | 2.81 | −13.93 |
| 4.60 | 3.85 | 63.00 | 47.04 |
| 0.19 | −1.19 | 8.00 | −17.05 |
| 0.14 | 0.13 | 3.90 | −3.48 |
| 0.41 | −0.12 | 12.50 | −0.50 |

(1.1) $$Y_t - 10.630 = 6.6X_t - 0.55X_{t-1}.$$

**Interpretation and uses of the Transfer Function Model**

Transfer function models are generally used to forecast the future values of a time series $Y_t$ (in our case the running times) given the previous values of the leading indicator series $X_t$ (the down times in our case). However, forecasts of the running times based upon the previous and present values of the down times, via Equation (1.1), must be used with great caution for the following two reasons. First, the occurrence of unforeseen but rare circumstances may cause the future running times to be abnormally small (or even large). An example of this is a reactor shutdown due to an unforeseen operator error. Second, since Equation (1.1) is based on only 28 observations, it may not be too reliable as a model for forecasting. We can, however, make several observations of practical interest based upon Equation (1.1).

We first note that previous values of $Y_t$, such as $Y_{t-1}$, $Y_{t-2}$, ..., etc., do not appear in Equation (1.1). This implies that the running time history gives us little information about the *individual* future running times; that is, the next running time may be unpredictable from a knowledge of the previous running times. However, future running times *on the average* may be estimated from prevous running times.

An important consequence of Equation (1.1) is that the running times $Y_t$ appear to be strongly influenced by the immediately preceding down times $X_t$. Since the down times generally correspond to maintenance actions, we can make the following conjecture:

*Conjecture*: Barring unforeseeable circumstances, and confining ourselves to the limits of the observed data, the operating times are, on the average, increased by a factor of about six per unit increase in the maintenance times.

An explanation to support the above conjecture is that the more thorough a job of repair that is performed, the longer the next running time becomes. This is perhaps one of the most important conclusions that can be reached from our analyses.

Since the coefficient of $-0.55$ associated with $X_{t-1}$ is small compared to the coefficient of 6.6 associated with $X_t$, we will ignore the effect of $X_{t-1}$ on $Y_t$. Even though the transfer function model is obtained after an involved analysis, as discussed in the remainder of this report, the simplicity of Equation (1.1) suggests that a plot of $Y_t$ versus $X_t$, $t = 1, 2, \ldots, 28$, should be approximately linear. The actual plot confirms the reasonableness of Equation (1.1), including the values of its coefficients.

In conclusion, for the situation considered here the transfer function model is more efficient as a tool that gives us some insight into the manner in which the system operates, rather than as a tool that can give us reliable forecasts of future running times.

The remainder of this paper is devoted to a discussion of the pertinent details that lead us to our model. In Section 2, by way of presenting some aspects of transfer function model building, we also introduce some terminology and notation. In Section 3, we present an analysis of our data.

In what follows, we require the reader to have some familiarity with the material in Box and Jenkins [1] and with that in Jenkins and Watts [3].

## 2. TRANSFER FUNCTION MODELS AND THEIR ESTIMATION

Univariate transfer function models as described by Box and Jenkins [1] are models that specify the stochastic interrelationships between two time series. They are more general than regression models with lag structures on predetermined variables, in that the dependent variable can also have a lag structure. In addition, the transfer function models can have a superimposed error structure which may be of a very general nature. Engineers often refer to error with the term "noise," and "white noise" refers to errors that are independent and identically distributed.

There are two equivalent representations of a univariate transfer function model. One is the infinite or reduced form, and the other is the finite form. In the infinite form, the output series, say $Y_t$ (in our case the running time), is explicitly represented as a function of the input series $X_t$ (in our case the down time) and its lagged (previous) values; that is,

$$Y_t = v_0 + v_1 X_{t-1} + v_2 X_{t-2} + \ldots + N_t \,, \tag{2.1}$$

where the constants $v_0$, $v_1$, ... are called the *impulse response weights.* In cases where there is no immediate response, one or more of the initial v's, say $v_0$, $v_1$, ..., $v_{b-1}$, is equal to zero. The process $N_t$ represents noise, which is assumed to be independent of the level of the input series, but is additive with respect to the influence of the input; $N_t$ can have any general structure.

It can be shown that an equivalent representation of the model given by Equation (2.1) is the following finite form:

$$Y_t - \delta_1 Y_{t-1} - \ldots - \delta_r Y_{t-r} = \omega_0 X_{t-b} - \omega_1 X_{t-b-1} - \ldots - \omega_s X_{t-b-s} + N_t \,, \tag{2.2}$$

where the $\delta$'s, the $\omega$'s, and $b$ are unknown constants. The constant $b$ associated with the leading indicator series $X_t$ indicates which of the previous values of $X_t$ affect the present $Y_t$. In our application, the value of $b$ represents the number of previous maintenance times affecting the present running time.

A first step towards estimating the transfer function model is a tentative identification of the values of $r$, $b$, and $s$. This can be accomplished by an examination of the estimated impulse response weights $\hat{v}_k$, $k = 0, 1, 2, \ldots$. A plot of $\hat{v}_k$ versus $k$ is known as the *impulse response function.*

There are two general approaches for obtaining the impulse response function. The first one, outlined by [1, p. 379], is based on a "prewhitening" of the input series. Prewhitening the input series means fitting a time series model to the $X_t$ series such that the residuals from the model, say $\alpha_t$, are independent and identically distributed random variables with mean zero and a constant variance. When the $\hat{v}_k$ are estimated using the prewhitening of the input series procedure, their neighboring values tend to be correlated. Thus the graph of the impulse response function tends to be misleading. This ultimately affects our ability to obtain a realistic transfer function model. We are therefore interested in considering an alternate approach for estimating the impulse response weights.

The second approach for estimating the impulse response weights involves the use of "cross spectral analysis". Such an approach removes the difficulties associated with the problem of the correlated estimates of $v_k$, and also provides us with some additional insight into the nature of the dependencies between the input and output series. These are illustrated at the end of Section 3.

Once the impulse response function is obtained, we can isolate the noise series, $N_t$, by using Equation (2.1). Specifically, we estimate the noise series by

$$\hat{N}_t = Y_t - \hat{v}_0 - \hat{v}_1 X_{t-1} - \hat{v}_2 X_{t-2} - \ldots . \tag{2.3}$$

A knowledge of $\hat{N}_t$, plus a knowledge of the tentative values of $r$, $b$, and $s$, helps us to estimate the parameters of the transfer function model, Equation (2.2). One way of accomplishing this is by using the TIMES program package described by Willie [7].

The adequacy of the proposed model can be checked by an analysis of the residues from the model. The details of such an analysis are given in [1, p. 392].

## 3. ANALYSIS OF RUNNING TIME AND DOWN TIME DATA

Data on the operating history of nuclear reactors are generally available showing dates on which the reactors ceased operation and the duration of the stoppage. Among other facts, the reasons for the stoppage are also given. Stoppages are categorized according to whether they were scheduled or forced. In addition to this, there is a further breakdown indicating whether the stoppage was due to equipment failure, testing, refueling, regulatory reasons, operator training, administrative reasons, operational error, or other causes.

### 3.1 Screening the Data

The data that were given to us described the Robinson Power Plant's operating history from June 1974 through April 1976. These data did not contain any stoppages due to administrative reasons, operational error, or other causes; they contained one stoppage for regulatory reasons and one stoppage for refueling. In one instance, the data contained an unrealistic combination for the cause of stoppage—a scheduled failure. In this case we used our discretion to alter it to a forced failure. Whenever there were stoppages due to operator training, these were treated as running times rather than as down times. This was done for two reasons. First, the duration of each stoppage was very short (on the average about 1/2 hour); second, we would like to concentrate on those down times that pertain to the physical operation of the system rather than on those external to it.

The single stoppage due to a forced regulatory restriction was for a period of 3.67 hours, and since it immediately followed a forced equipment failure of 15.52 hours, it was combined with the equipment failure stoppage.

Refueling the reactor takes place annually and is generally of a very long duration. In our data, we had only one stoppage for scheduled refueling, and it was of 960 hours duration. Since the duration of this stoppage is out of line with the duration of the other stoppages (see Table 1.1), it was excluded from consideration.

We remark here that any time a stoppage (running time), say $X_t(Y_t)$, was excluded from consideration, its corresponding running time (stoppage) $Y_t(X_t)$ was also excluded. This is to ensure that no bias is introduced into the relationship between the two variables of interest because of the elimination of observations of either one.

Another convention followed in our analysis arises from the fact that the raw data show the dates and the duration (in hours) of each stoppage, rather than the actual time of stoppage. We assume that each down time commenced at 0000 hours (unless in some rare instances there is a second stoppage occurring during the same day). Whenever two or more breakdowns occurred during the same day, they were combined into one down time period and the intervening operating period was ignored.

Because of the paucity of data, we chose not to distinguish between stoppages due to equipment failure and those due to testing. This is reasonable because whenever there is a forced equipment faiure, maintenance and test actions on other (nonfailed) components are routine. Thus, in practice it is difficult to differentiate clearly between the consequences of equipment failure and those of testing. This strategy was suggested by some staff members of the Probabilistic Analysis Staff at NRC.

The preliminary screening and examination described above gives us a series of values of the down times $X_t$ (in days) and a series of values of the corresponding running times $Y_t$ (in days). However, as we shall soon see, some further screening is necessary.

The next step in our analysis involved prewhitening the $X_t$ series. This turned out to be quite a frustrating endeavor, since no simple univariate time series model of the Box-Jenkins type seemed to provide a reasonable fit. The difficulty turned out to have been caused by two unusually large down times due to scheduled testing and forced failure of 18.08 and 25.43 days, respectively. These were incompatible with the other down times (see Table 1.1), and thus defied the use of a simple model as a prewhitening transformation. Perhaps a time series model with an indicator variable (such as those used in the "intervention analysis" of Box and Tiao [2]) might have been adequate for these and for the refueling stoppage, but this was not attempted. In the interest of expediency, it was preferable to eliminate the two large $X_t$'s and their corresponding $Y_t$'s. Thus, in effect, some data screening was done during the prewhitening phase. Table 1.1 presents the 28 screened values of the down times $X_t$ and the corresponding running times $Y_t$. We remind the reader that the subscript $t$ is a sequential index rather than an index representing time. That is, $X_t$ and $Y_t$ are not simultaneously observed in time; $X_t$ precedes $Y_t$.

### 3.2 Transfer Function Model Development

In Figures 3.1 and 3.2 we show plots of the autocorrelation functions of the (screened) down times $X_t$ and the (screened) running times $Y_t$, respectively. Based upon these plots we are able to conclude that the two time series can be treated as stationary [1, p. 174].

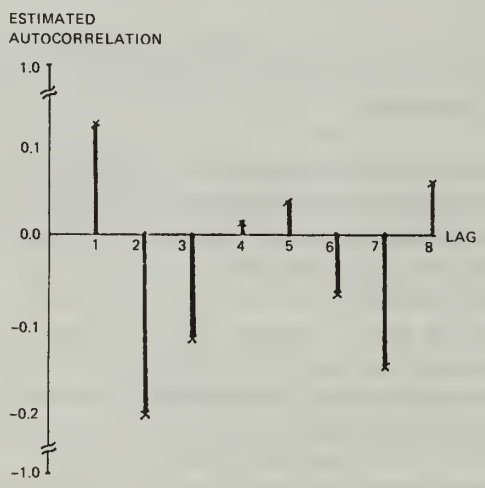


FIGURE 3.1. Estimated autocorrelations of down times $X_t$

In Figure 3.3 we show a plot of the logarithm of the smoothed estimator of the power spectrum of the screened down times $X_t$, as well as the approximate 95% confidence limits. The smoothing was performed using a rectangular window of band width .098. As a matter of fact, all the smoothing that is discussed in this paper was performed using a rectangular window of band width .098. The power spectrum curve shows us how the variance of the $X_t$ time series is distributed with frequency. For a detailed understanding of the power spectrum, its smoothing, and the band width of a smoothing window, we refer the reader to Chapter 6 of Jenkins and Watts [3].

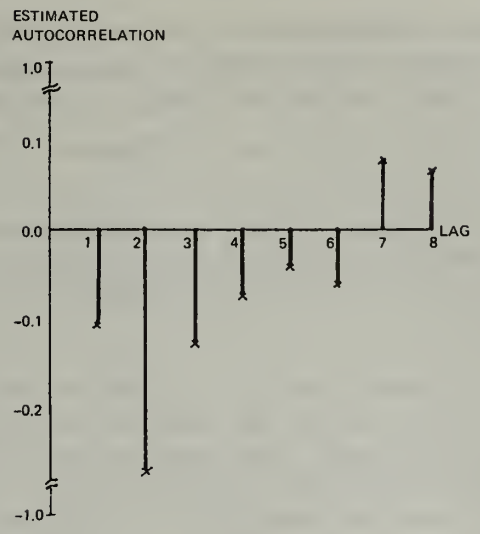


FIGURE 3.2. Estimated autocorrelations of running times $Y_t$

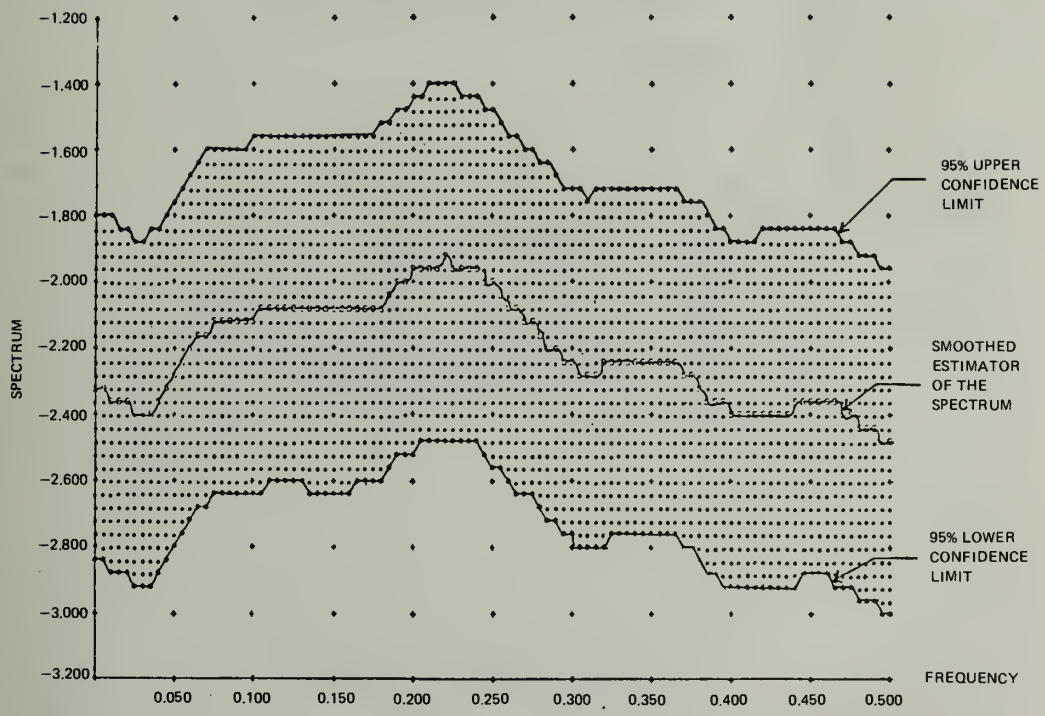


FIGURE 3.3. $\text{Log}_{10}$ of the smoothed estimator of the spectrum of down times versus frequency, using rectangular window of band width 0.98

Our next step involves the determination of a suitable prewhitening transformation for the down times $X_t$. Based upon the several models that we attempted, we conclude that a moving average process of order 3 best describes the $X_t$ series. Specifically, we find that

$$X_t - 0.7963 = \alpha_t + 0.117\alpha_{t-1} - 0.189\alpha_{t-2} - 0.133\alpha_{t-3}, \tag{3.1}$$

where .7963 is the mean of the $X_t$ series. The $\alpha_t$ represent the residuals when a moving average process of order 3 is fitted to the $X_t$ series. If the prewhitening transformation given by Equation (3.1) is correct, then the $\alpha_t$ will be independently and identically distributed with a constant mean and variance.

In Table 1.1 we give the values of the $\alpha_t$'s. In order to verify the appropriateness of the model given by Equation (3.1), we plot the estimated autocorrelation function and the estimated power spectrum of the $a_t$ series. These plots are given in Figures 3.4 and 3.5, respectively. We remark that the plot of the estimated power spectrum of the $\alpha_t$ series given in Figure 3.5 is relatively constant as compared to the plot of the estimated power spectrum of the $X_t$ series given in Figure 3.3. This is because the effect of prewhitening is to remove the dependencies among the $X_t$'s and give us a set of independent $\alpha_t$. Figures 3.4 and 3.5 confirm the appropriateness of the prewhitening transformation given by Equation (3.1).

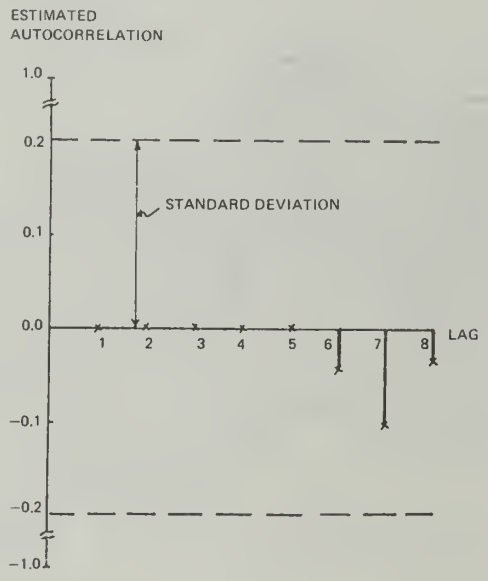


FIGURE 3.4. Estimated autocorrelations of residuals from the model used for prewhitening $X_t$

Following Box and Jenkins [1, p. 380], we next apply the same prewhitening transformation (i.e., the one given by Equation (3.1)) to the running times $Y_t$ and obtain the $\beta_t$'s as residuals. In Table 1.1 we give the values of $\beta_t$ under the heading "Prewhitened Running Times." In Figure 3.6 we show a plot of the smoothed estimator of the power spectrum of $\beta_t$. We remark that except at the very low frequencies, the plot of the estimated power spectrum of the $\beta_t$'s is fairly constant. Thus it appears that the prewhitening transformation given by Equation (3.1), when applied to the running times $Y_t$, also yields a sequence of independent and identically distributed random variables $\beta_t$.

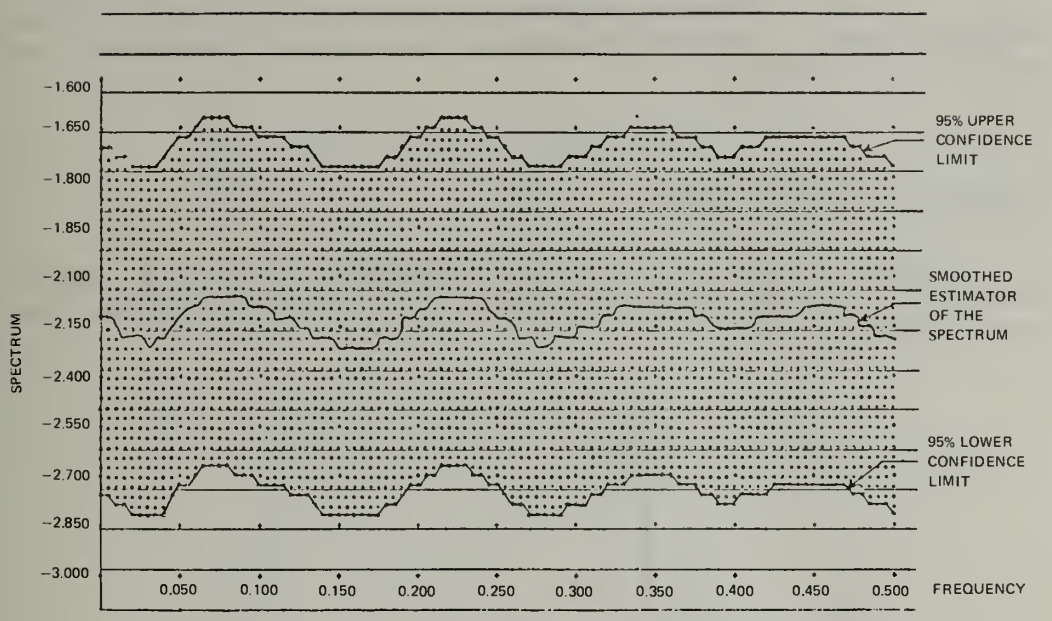


FIGURE 3.5. $\text{Log}_{10}$ of the smoothed estimator of the spectrum of the prewhitened down times versus frequency, using rectangular window of band width .098

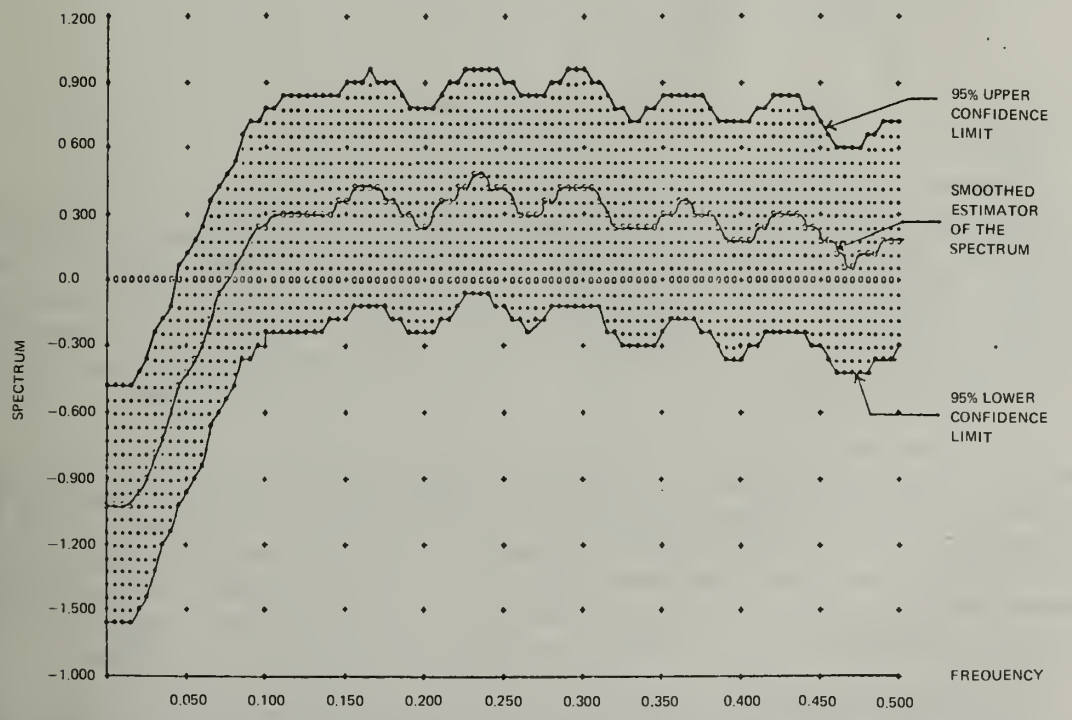


FIGURE 3.6. $\text{Log}_{10}$ of the smoothed estimator of the spectrum of $B_t$ versus frequency, using a rectangular window of band width .098

Our next step is to obtain the cross correlation between the $\alpha_t$ and $\beta_t$ at lags $k$, $k = 0, 1, 2, \ldots$. If $s_\alpha$ and $s_\beta$ denote the estimated standard deviations of the $\alpha_t$ and the $\beta_t$ series, respectively, and if $r_{\alpha\beta}(k)$ denotes the estimated cross correlation between the $\alpha_t$ and the $\beta_t$ at lag $k$, then $\hat{v}_k$, an estimate of $v_k$, is

$$\hat{v}_k = r_{\alpha\beta}(k)\,\frac{s_\beta}{s_\alpha}, \quad k = 0, 1, 2, \ldots,$$

(see Box and Jenkins [1, p. 380]).

In Figure 3.7 we show a plot of the impulse response function; that is, a plot of $\hat{v}_k$ versus $k$, $k = 0, 1, 2, \ldots$. We remark that in Figure 3.7, the value $\hat{v}_0$ is significantly larger than the other values of $\hat{v}_k$, $k = 1, 2, \ldots$.

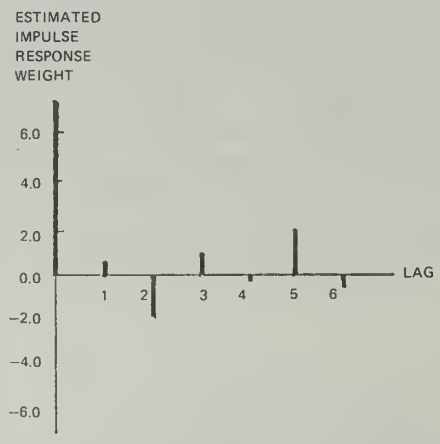


FIGURE 3.7. The impulse response function using the prewhitening of the input series

Because the neighboring values of the $\hat{v}_k$ in Figure 3.7 tend to be correlated, we also obtained the impulse response function using the "cross spectrum" between the $X_t$ and the $Y_t$ series (see Jenkins and Watts [3, p. 424]). In Figure 3.8 we show a plot of the impulse response function using the cross spectrum. Note that this plot is quite similar to the one given in Figure 3.7; that is, $\hat{v}_0$ is significantly larger than the other values of $\hat{v}_k$. Based upon Figures 3.7 and 3.8, we can conclude that *the greatest influence on the running time is exerted by the down time immediately preceding it.* This is, of course, a major point of our conclusions.

We now estimate the noise series $N_t$ using Equation (2.3). An inspection of the estimated autocorrelation function of the estimated noise series $\hat{N}_t$, and a "portmanteau lack of fit test" (1.69 with 8 degrees of freedom) [1, p. 290], lead us to conclude that the $\hat{N}_t$ are independent and identically distributed. In addition to the above, we show in Figure 3.9 a plot of the smoothed estimator of the power spectrum of the $\hat{N}_t$. Here again, except at the very low frequencies, the estimated power spectrum of the $\hat{N}_t$ series is fairly constant. Thus the estimated noise series $\hat{N}_t$ can be described reasonably well by a white noise process.

A final step in the analysis involves the fitting of a transfer function model to the running times $Y_t$. This was accomplished by using the TIMES program package. Of the several models that were attempted, the model

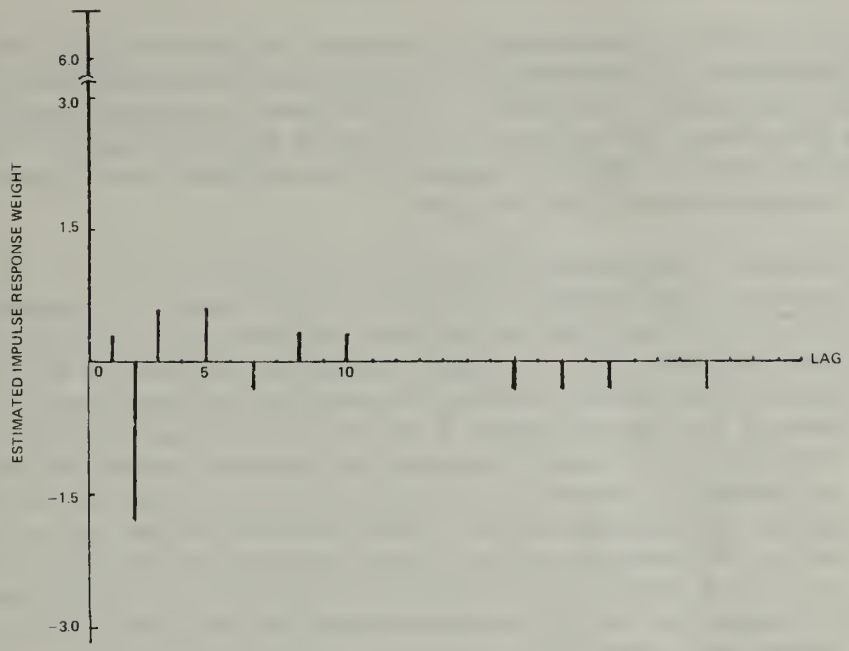


FIGURE 3.8. The impulse response function using cross spectral analysis

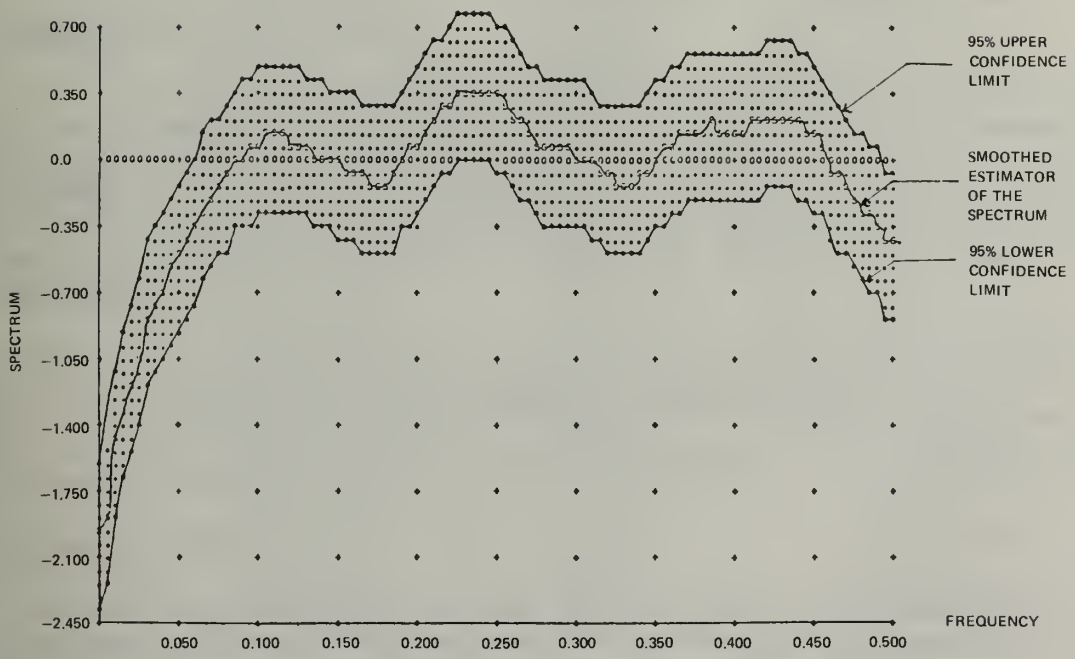


FIGURE 3.9. $\text{Log}_{10}$ of the smoothed estimator of the spectrum of the estimated noise series, using a rectangular window of band width .098

(3.2) $$Y_t - 10.630 = 6.6X_t - 0.55X_{t-1}$$

appears to be the best; 10.630 is the mean of the $Y_t$ series.

In order to verify the reasonableness of the model, two diagnostic checks were suggested in [1]. One depends on the autocorrelation of the $\hat{N}_t$ and the other depends on the estimated cross correlation between the $\hat{N}_t$ and the $\alpha_t$. For both cases a portmanteau lack of fit test was used. In the former case, the test statistic is 1.69 with 8 degrees of freedom, whereas in the latter case the test statistic is .698 with 6 degrees of freedom. These values support the reasonableness of the proposed transfer function model.

As stated earlier, a cross spectral analysis of the $X_t$ and the $Y_t$ series can give us further insight into the nature of the dependencies between the two series. For example, $G(f)$, the "gain" at frequency $f$, behaves like the regression coefficient in a linear regression model between the output and the input at frequency $f$ (see Jenkins and Watts [3, p. 352]. In Figure 3.10 we plot the gain of the running times on the down times at various frequencies. Another important function in cross spectral analysis is the "squared coherence" between the input and the output at frequency $f$. This quantity measures the correlation between the sinusoidal component of $Y_t$ and that of $X_t$ at frequency $f$. The square coherence is also in some sense a measure of the proportion of information in the $Y_t$ series that is attributable to the $X_t$ series. For more information on the coherence and the coherence spectrum, we refer the reader to Jenkins and Watts [3, p. 352]. In Figure 3.11 we show a plot of the coherence for the running time and down time data. Figures 3.10 and 3.11 give us some additional assurance on the dependence of the running times on the down times.

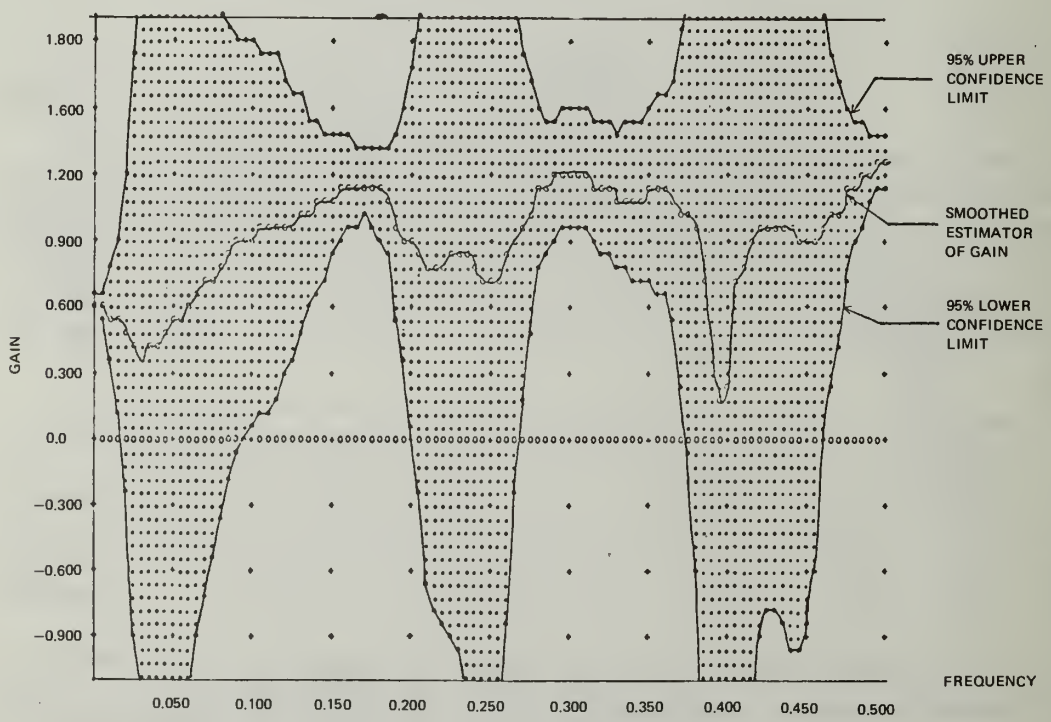


FIGURE 3.10. The gain versus frequency of running times on down times

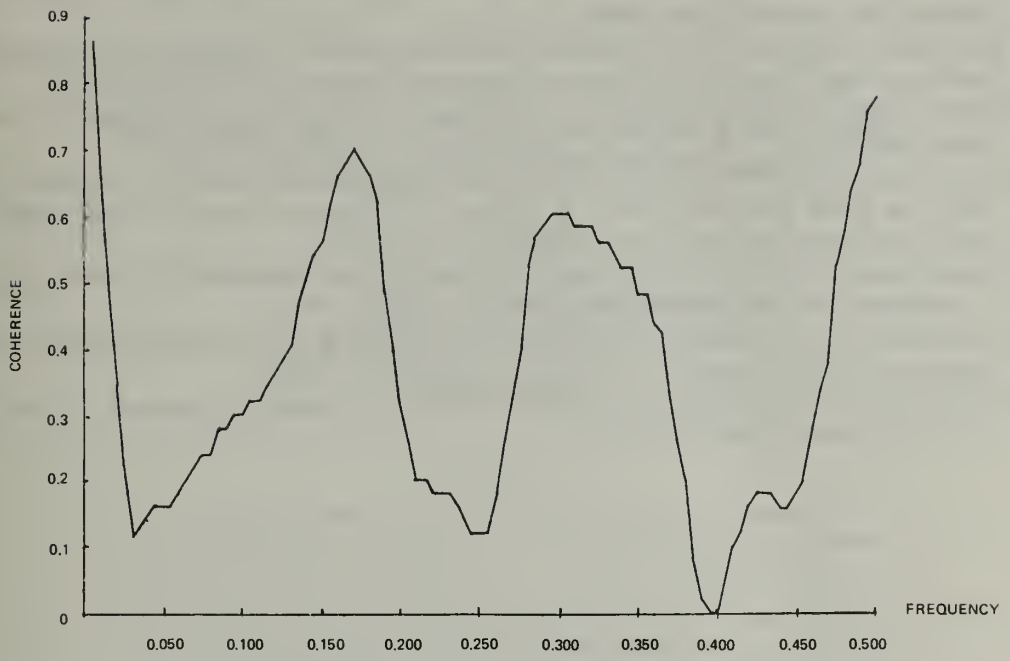


FIGURE 3.11. The coherence versus frequency of running times and down times

## 4. SUMMARY AND CONCLUSIONS

In the foregoing analysis we have demonstrated the use of time series analysis methodology for studying the interrelationships between the maintenance times and the running times of a complex system. Our analysis enables a decision maker to assess the impact of his maintenance policies on running times, or to influence the operating times by managing the maintenance times. In addition, given a down time a decision maker can, to a *limited extent*, forecast the next running time. This type of information may be very valuable, especially for large and complex systems.

Our analysis can be criticized on the grounds that it is based on an insufficient amount of data. We hope that this criticism can be overlooked in the light of the fact that our approach is to be viewed as a prototype for the analysis of reliability data involving two interrelated sources of data. For other uses of the time series analysis methodology for reliability and life data, we refer the reader to Singpurwalla [4], [5] and [6].

## ACKNOWLEDGMENTS

The written comments by Drs. Ray Waller and Gary Tietjens, and the helpful conversations with Professor Robert Shumway and Mr. Randall Willie, are much appreciated. Mr. Mahesh Chandra's efforts on behalf of the work reported here is also acknowledged. Finally, the several comments by Professor Frank Proschan with respect to the direction and the emphasis of this report are gratefully acknowledged.

## REFERENCES

[1] Box, G.E.P. and G.M. Jenkins, *Time Series Analysis, Forecasting, and Control*, revised edition (Holden-Day, San Francisco 1976).

[2] Box, G.E.P. and G.C. Tiao, "Intervention Analysis with Applications to Economic and Environmental Problems," J. American Statistical Association *70*, 70-79 (1975).

[3] Jenkins, G.M. and D.G. Watts, *Spectral Analysis and its Applications* (Holden-Day, San Francisco 1968).

[4] Singpurwalla, N.D. "Time Series Analysis and Forecasting of Failure Rate Processes," Reliability and Fault Tree Analysis (R.E. Barlow, J.B. Fussell, and N.D. Singpurwalla, Eds.) Society for Industrial and Applied Mathematics (1975).

[5] Singpurwalla, N.D. "Time Series Analysis of Failure Data. Proceedings," Annual Symposium on Reliability and Maintainability, pp. 107-112 (1978).

[6] Singpurwalla, N.D. "Estimating Reliability Growth Using Time Series Analysis," The Naval Research Logistics Quarterly, Vol. *25*, No. 1, pp. 1-14 (1978).

[7] Willie, R.R. "Everyman's guide to TIMS," ORC 77-2, Operations Research Center, University of California at Berkeley (1977).

# OPTIMAL MAINTENANCE-REPAIR POLICIES FOR THE MACHINE REPAIR PROBLEM

S. Christian Albright

*Department of Quantitative Business Analysis*
*Graduate School of Business*
*Indiana University*
*Bloomington, Indiana*

**ABSTRACT**

We consider a model with $M + N$ identical machines. As many as $N$ of these can be working at any given time and the others act as standby spares. Working machines fail at exponential rate $\lambda$, spares fail at exponential rate $\gamma$, and failed machines are repaired at exponential rate $\mu$. The control variables are $\lambda$, $\mu$, and the number of removable repairman, $S$, to be operated at any given time. Using the criterion of total expected discounted cost, we show that $\lambda$, $S$, and $\mu$ are monotonic functions of the number of failed machines $M$, $N$, the discount factor, and for the finite time horizon model, the amount of time remaining.

## 1. INTRODUCTION

In this paper we examine the structure of the optimal maintenance and repair policies for an exponential repair model. We assume that there are $M + N$ identical machines, $N$ of which can be working at any given time. The others function as standby spares. The working machines fail at rate $\lambda$, and the failed machines are repaired at rate $\mu$. We also include the possibility of spare machines failing at rate $\gamma$. Our control variables are $\lambda$, which can be thought of as a maintenance control variable, $\mu$, and the number of removable repairmen, $S$. Our principle results are that $\lambda$ decreases and $S$ and $\mu$ increase as more machines are in the failed state. (Throughout this paper, we use the terms increasing and decreasing to mean nondecreasing and nonincreasing, respectively.) We also show how these control variables behave as functions of $M$, $N$, the discount factor, and the amount of time remaining.

Our basic model has been studied by several researchers. Some of these have described the model and given its operating characteristics under exponential and nonexponential assumptions. See, for example, Barlow [3], Gnedenko, Belyayev, and Solovyev [8], and Iglehart and Lemoine [11], [12]. There has been less literature on the control of these models. Goheen [9] and Cinlar [4] study the problem of where to send failed machines when there are several repair facilities. Cinlar assumes exponential times and reduces the problem to a linear programming problem, whereas Goheen assumes Erlang times and reduces to the problem to a mathematical programming problem.

More in line with the model presented in the current paper are the papers by Crabill [5] and Winston [18]. Each considers a single-repairman system (and Winston assumes no spares)

which can be operated at a finite number of service rates, and each shows that the optimal service rate is an increasing function of the number of failed machines. Crabill does this with a long run average cost criterion by means of tedious manipulations of policy improvement equations. Winston uses a discounted cost criterion, and he builds up his infinite horizon model as a sequence of discrete time $n$-period problems. He then analyzes these quite easily be means of induction on $n$. This type of approach is also used in the present paper although our $n$-period model is defined differently than Winston's. This approach allows us to obtain more results, and to do so by simpler means, than was evidently possible for Crabill.

## 2. THE BASIC MODEL

In this section we present the basic model and main results. The system we will study is composed of $N + M$ identical machines. As many as $N$ of these can be working simultaneously in parallel, while the rest function as warm standby spares. Each of the machines which is currently working fails, independently of the others, with exponential rate $\lambda$. Whenever one of these machines fails, it is immediately replaced by a spare if any is available. We also assume each of the available spares fails, independently of the other machines, with exponential rate $\gamma$. (This is the rationale behind the term "warm standby".) There is a repair facility with as many as $N + M$ removable servers in parallel. Whenever a working machine or a spare fails, it is sent to this facility, where it is served on a first-come-first-served basis with exponential rate $\mu$.

The costs and rewards are as follows. Whenever $j$ machines, $0 \leqslant j \leqslant N$, are working, revenue is earned at a rate $h(j)$, where $h$ is an increasing function with $h(0) = 0$. The service cost for having $S$ servers each working at rate $\mu$ is $Sf(\mu)$, where $f$ is a nonnegative increasing function of $\mu$. Also, because we will wish to control $\lambda$, we assume the cost of having $j$ machines working, each subject to failure rate $\lambda$, is $jg(\lambda)$. Here $g$ is a decreasing function of $\lambda$ and can be thought of as a maintenance cost. Finally, we will continuously discount all revenues and costs by a discount rate $\alpha > 0$. Our objective is to maximize expected discounted rewards over an infinite horizon. The decision variables at any decision point are the number of servers $S$, the service rate per service $\mu$, and the failure rate per working machine $\lambda$. These must satisfy $0 \leqslant S \leqslant N + M$, $0 < \lambda_{\min} \leqslant \lambda \leqslant \lambda_{\max} < \infty$, and $0 < \mu_{\min} \leqslant \mu \leqslant \mu_{\max} < \infty$. We assume the failure rate for spares, $\gamma$, is a fixed constant.

The problem is formulated most naturally as a continuous time Markov decision model. The state space is $\{0, 1, 2, \ldots, N + M\}$, where state $i$ means that $i$ machines are currently in repair or waiting for repair. The actions, transition rates, and reward rates are easily obtained from the above description of the model. For example, in state $i$, let $N_i$ and $M_i$ be, respectively, the number of working machines and the number of available spares. Then $N_i = (N + M - i) \wedge N$ and $M_i = 0 \vee (M - i)$, where "$\wedge$" and "$\vee$" stand for "min" and "max". This implies that if we are in state $i$ and use actions $S$, $\lambda$, and $\mu$, the transition rates $q_{ij}(S, \lambda, \mu)$ are $q_{i,i+1}(S, \lambda, \mu) = N_i\lambda + M_i\gamma$, $q_{i,i-1}(S, \lambda, \mu) = (S \wedge i)\mu$, and $q_{ii}(S, \lambda, \mu) = -N_i\lambda - M_i\gamma - (S \wedge i)\mu$. The reward rate in state $i$ is $r_i(S, \lambda, \mu) = h(N_i) - N_i g(\lambda) - Sf(\mu)$. Because of the exponential assumptions and infinite time horizon, we may assume that decisions are made only at failure times and at service completion times.

Let $V(i)$ be the maximal expected discounted reward which can be obtained when starting in state $i$. Then it follows from well-known results (see, for example, Howard [10] or Ross [16]) that $V(i)$ satisfies a certain functional equation and that the solution to this equation yields the optimal actions to use in state $i$. In this paper, however, we are interested in structural properties of the optimal actions and, in particular, how they behave as functions of $i$. To discover these properties, the above functional equation is unfortunately not very useful.

Instead, it is better to build up our infinite horizon problem as a sequence of $n$-period discrete time problems. This enables us to use induction arguments (on $n$) to prove the desired structural properties of the optimal actions.

To this end, we define a finite horizon discrete time Markov decision model which in the limit (on the number of periods) is equivalent to our original problem in the sense that each gives the same optimal expected rewards and each prescribes the same optimal actions. This method of using a discrete time problem to prove structural properties of a continuous time problem was first utilized by Lippman [14], and has since been discussed and used by several other authors [1], [7], [13], [17]. An especially lucid description of the relationship between the two problems may be found in Anderson's thesis [2]. Because the method is by now quite well-known, we will only sketch the procedure. The reader is referred to the above references for further details.

Let $\Lambda = N\lambda_{\max} + M\gamma + (N + M)\mu_{\max}$, so that $\Lambda \geqslant -q_{ii}(S, \lambda, \mu)$ for all states $i$ and actions $S, \lambda, \mu$. The possible states and actions of the discrete time process are the same as for the continuous time process, but the transition probabilities, rewards, and discount factor are modified as follows. The transition probabilities are $p_{i,i+1}(S, \lambda, \mu) = q_{i,i+1}(S, \lambda, \mu)/\Lambda = (N_i\lambda + M_i\gamma)/\Lambda$, $p_{i,i-1}(S, \lambda, \mu) = q_{i,i-1}(S, \lambda, \mu)/\Lambda = (S \wedge i)\mu/\Lambda$, and $p_{ii}(S, \lambda, \mu) = (\Lambda + q_{ii}(S, \lambda, \mu))/\Lambda = (\Lambda - N_i\lambda - M_i\gamma - (S \wedge i)\mu)/\Lambda$. The one-period rewards are $\bar{r}_i(S, \lambda, \mu) = r_i(S, \lambda, \mu)/(\Lambda + \alpha) = (h(N_i) - N_i g(\lambda) - Sf(\mu))/(\Lambda + \alpha)$, and the discount factor is $\beta = \Lambda/(\Lambda + \alpha)$.

Let $V_n(i)$ be the maximal expected discounted reward over $n$ periods for the above discrete time process. Then it follows from the previously referenced results that $V_n(i)$ converges to $V(i)$ and the optimal actions for the $n$-period problem converge to the optimal actions for the continuous time infinite horizon problem. Furthermore, it is easy to see that $V_n(i)$ satisfies the following recursive equation:

$$V_n(i) = \max_{S,\ \lambda,\ \mu} \{J_n(i, S, \lambda, \mu)\}/(\Lambda + \alpha),$$

where $V_0(i) \equiv 0$, and

$$J_n(i, S, \lambda, \mu) = r_i(S, \lambda, \mu) + (N_i\lambda + M_i\gamma)V_{n-1}(i + 1) + (S \wedge i)\mu V_{n-1}(i - 1) + (\Lambda - N_i\lambda - M_i\gamma - (S \wedge i)\mu)V_{n-1}(i).$$

By letting $\Delta_n(i) = V_n(i) - V_n(i - 1)$, we may rewrite $J_n$ in the more useful form:

$$J_n(i, S, \lambda, \mu) = h(N_i) + N_i(-g(\lambda) + \lambda\Delta_{n-1}(i + 1)) + M_i\gamma\Delta_{n-1}(i + 1) - Sf(\mu) - (S \wedge i)\mu\Delta_{n-1}(i) + \Lambda V_{n-1}(i).$$

Since $f(\mu) \geqslant 0$, it makes no sense to have $S > i$. Therefore, $S \wedge i = S$ and the possible action space in state $i$ is $0 \leqslant S \leqslant i$, $\lambda_{\min} \leqslant \lambda \leqslant \lambda_{\max}$, $\mu_{\min} \leqslant \mu \leqslant \mu_{\max}$.

We will assume sufficient continuity conditions on $f$ and $g$ to insure that the above maximum is obtained. With this in mind, let $S_n(i)$ be optimal in state $i$, let $\lambda_n(i)$ be the largest optimizing action in state $i$, and let $\mu_n(i)$ be the smallest optimizing action in state $i$, all when $n$ periods remain. More concisely, we will say these actions are optimal for $(n, i)$. Our main result below is that $S_n(i)$ and $\mu_n(i)$ are increasing functions of $i$ and $\lambda_n(i)$ is a decreasing function of $i$. We also show how these optimal actions behave as functions of $n$, $\alpha$, $M$, and $N$. Our first observation follows easily (and we immediately see the benefit of this particular discrete time formulation). From the above form of $J_n$ it is clear that

(i) $\lambda_n(i)$ maximizes $-g(\lambda) + \lambda\Delta_{n-1}(i + 1)$,

(ii) $\mu_n(i)$ minimizes $f(\mu) + \mu\Delta_{n-1}(i)$,

(iii) $S_n(i) = \begin{cases} 0 & \text{if } f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i) \geqslant 0 \\ i & \text{if } f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i) < 0. \end{cases}$

We are now ready for the results of this section.

PROPOSITION 1: If $h(N_i) - h(N_{i+1}) \geqslant g(\lambda_{\min})$ for $i \geqslant M$, then $\Delta_n(i) \leqslant 0$ for $n \geqslant 0$, $1 \leqslant i \leqslant M + N$.

PROOF: We use induction on $n$. For $n = 0$, $\Delta_0(i) = 0$, so the result is trivial. Assuming $\Delta_{n-1}(i) \leqslant 0$, let $(S_1, \lambda_1, \mu_1)$ be the optimal actions for $(n, i)$ and $(S_2, \lambda_2, \mu_2)$ be the optimal actions for $(n, i-1)$, and we have

$$\begin{aligned}(\Lambda + \alpha)\Delta_n(i) &= J_n(i, S_1, \lambda_1, \mu_1) - J_n(i-1, S_2, \lambda_2, \mu_2) \\ &\leqslant J_n(i, S_1, \lambda_1, \mu_1) - J_n(i-1, S_2, \lambda_1, \mu_2) \\ &= h(N_i) - h(N_{i-1}) - (N_i - N_{i-1})g(\lambda_1) + (N_i\lambda_1 + M_i\gamma)\Delta_{n-1}(i+1) \\ &\quad - S_1 f(\mu_1) + S_2(f(\mu_2) + \mu_2\Delta_{n-1}(i-1)) \\ &\quad + (\Lambda - N_{i-1}\lambda_1 - M_{i-1}\gamma - S_1\mu_1)\Delta_{n-1}(i) \\ &\leqslant h(N_i) - h(N_{i-1}) - (N_i - N_{i-1})g(\lambda_1) + S_2(f(\mu_2) + \mu_2\Delta_{n-1}(i-1)),\end{aligned}$$

by the induction hypothesis and definition of $\Lambda$. Also, if $i \leqslant M$, $h(N_i) - h(N_{i-1}) - (N_i - N_{i-1})g(\lambda_1) = 0$, and if $i \geqslant M + 1$, $h(N_i) - h(N_{i-1}) - (N_i - N_{i-1})g(\lambda_1) = h(N_i) - h(N_{i-1}) + g(\lambda_1) \leqslant 0$, by the assumption of the proposition. Finally, either $S_2 = 0$ or $f(\mu_2) + \mu_2\Delta_{n-1}(i-1) < 0$, so that $S_2(f(\mu_2) + \mu_2\Delta_{n-1}(i-1)) \leqslant 0$. This completes the proof.

THEOREM 1: If $h(N_i) - h(N_{i+1}) \geqslant g(\lambda_{\max})$ for $i \geqslant M$, then $\Delta_{n+1}(i) \leqslant \Delta_n(i)$ for $n \geqslant 0$, $1 \leqslant i \leqslant M + N$.

PROOF: For $n = 0$, we need $\Delta_1(i) \leqslant 0$. We have

$$(\Lambda + \alpha)V_1(i) = \max_{S,\lambda,\mu} \{h(N_i) - N_i g(\lambda) - Sf(\mu)\} = h(N_i) - N_i g(\lambda_{\max}),$$

so that $\Delta_1(i) = 0$ for $i \leqslant M$ and $\Delta_1(i) = h(N_i) - h(N_{i-1}) + g(\lambda_{\max}) \leqslant 0$ for $i > M$. Now assume $\Delta_n(i) \leqslant \Delta_{n-1}(i)$ for each $i$, and let $(S_1, \lambda_1, \mu_1)$ and $(S_2, \lambda_2, \mu_2)$ be optimal for $(n+1, i)$ and $(n, i-1)$, respectively. Now we have

$$\begin{aligned}(\Lambda + \alpha)(\Delta_{n+1}(i) - \Delta_n(i)) &\leqslant J_{n+1}(i, S_1, \lambda_1, \mu_1) - J_n(i, S_1, \lambda_1, \mu_1) \\ &\quad - J_{n+1}(i-1, S_2, \lambda_2, \mu_2) + J_n(i-1, S_2, \lambda_2, \mu_2) \\ &= (N_i\lambda_1 + M_i\gamma)(\Delta_n(i+1) - \Delta_{n-1}(i+1)) \\ &\quad + S_2\mu_2(\Delta_n(i-1) - \Delta_{n-1}(i-1)) \\ &\quad + (\Lambda - N_{i-1}\lambda_2 - M_{i-1}\gamma - S_1\mu_1)(\Delta_n(i) - \Delta_{n-1}(i)) \leqslant 0,\end{aligned}$$

by the induction hypothesis.

COROLLARY 1: If $h(N_i) - h(N_{i+1}) \geqslant g(\lambda_{\max})$ for $i \geqslant M$, then $\lambda_n(i)$ is decreasing in $n$ for $0 \leqslant i \leqslant M + N - 1$, and $\mu_n(i)$ and $S_n(i)$ are increasing in $n$ for $1 \leqslant i \leqslant M + N$.

PROOF: If $\lambda_c$ is the largest value of $\lambda$ which maximizes the function $k(\lambda; c) = -g(\lambda) + \lambda c$, then it is easy to show that $\lambda_c$ is increasing in $c$. If we let $c = \Delta_n(i+1)$, the monotonicity

of $\lambda_n(i)$ in $n$ follows from Theorem 1. The proof that $\mu_n(i)$ increases in $n$ is similar. To show that $S_n(i)$ increases with $n$, it suffices to show that $f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i)$ decreases with $n$. We have

$$\begin{aligned} f(\mu_{n+1}(i)) + \mu_{n+1}(i)\Delta_n(i) &\leqslant f(\mu_n(i)) + \mu_n(i)\Delta_n(i) \\ &\leqslant f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i), \end{aligned}$$

from Theorem 1 and the definition of $\mu_{n+1}(i)$.

Corollary 1 states that when there is more time remaining, the system works harder, in the sense of more maintenance (lower $\lambda$) and more repair (higher $S$ and $\mu$). The next theorem and its corollary show that for a given $n$, the system works harder when more machines are broken down. From now on (except for Theorem 3 and Corollary 3), we will need to assume that $h$ is a linear function, so that $h(j) = jr$, $0 \leqslant j \leqslant N$.

THEOREM 2: Assume $g$ is convex, differentiable, and satisfies $-r + g(\lambda_{\min}) - g'(\lambda_{\min})\lambda_{\min} \leqslant 0$ for $i \geqslant M$. Then $\Delta_n(i+1) \leqslant \Delta_n(i)$ for $n \geqslant 0$, $1 \leqslant i \leqslant M + N - 1$.

PROOF: Before proceeding, we note that the above conditions imply $h(N_i) - h(N_{i+1}) \geqslant g(\lambda_{\min})$ for $i \geqslant M$, so that the results of Proposition 1 and Corollary 1 are in effect. This follows since $g' \leqslant 0$.

For $n = 0$, the result is trivial. Now assume inductively that $\Delta_{n-1}(i+1) \leqslant \Delta_{n-1}(i)$, $1 \leqslant i \leqslant M + N - 1$ and that $r - g(\lambda_n(i)) + \lambda_n(i)\Delta_{n-1}(i+1) \geqslant 0$ for $i \geqslant M + 1$.

First we show that $\Delta_n(i+1) \leqslant \Delta_n(i)$. The difference $(\Lambda + \alpha)(\Delta_n(i+1) - \Delta_n(i))$ is the sum of four expressions: the terms with $r$ and $\lambda$, the terms with $\gamma$, the terms with $S$ and $\mu$, and the terms with $\Lambda$. We look at each of these separately.

(i) $(r, \lambda \text{ terms}) \equiv A = N_{i+1}(r - g(\lambda_n(i+1)) + \lambda_n(i+1)\Delta_{n-1}(i+2))$
$- 2N_i(r - g(\lambda_n(i)) + \lambda_n(i)\Delta_{n-1}(i+1))$
$+ N_{i-1}(r - g(\lambda_n(i-1)) + \lambda_n(i-1)\Delta_{n-1}(i))$.

If $i \leqslant M - 1$,

$$\begin{aligned} A &\leqslant N\lambda_n(i+1)(\Delta_{n-1}(i+2) - \Delta_{n-1}(i+1)) \\ &\quad - N\lambda_n(i-1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)) \\ &\leqslant -N\lambda_n(i-1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)). \end{aligned}$$

If $i \geqslant M + 1$,

$$\begin{aligned} A &\leqslant N_{i+1}\lambda_n(i+1)(\Delta_{n-1}(i+2) - \Delta_{n-1}(i+1)) \\ &\quad - N_{i-1}\lambda_n(i-1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)) \\ &\leqslant - N_{i-1}\lambda_n(i-1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)), \end{aligned}$$

since $2N_i = N_{i+1} + N_{i-1}$.

If $i = M$,

$$\begin{aligned} A &= (N-1)(r - g(\lambda_n(M+1)) + \lambda_n(M+1)\Delta_{n-1}(M+2)) \\ &\quad - 2N(r - g(\lambda_n(M)) + \lambda_n(M)\Delta_{n-1}(M+1)) \end{aligned}$$

$$
\begin{aligned}
&+ N(r - g(\lambda_n(M-1)) + \lambda_n(M-1)\Delta_{n-1}(M)) \\
&\leqslant N\lambda_n(M+1)(\Delta_{n-1}(M+2) - \Delta_{n-1}(M+1)) \\
&- N\lambda_n(M-1)(\Delta_{n-1}(M+1) - \Delta_{n-1}(M)) \\
&\leqslant -N\lambda_n(M-1)(\Delta_{n-1}(M+1) - \Delta_{n-1}(M)),
\end{aligned}
$$

where the first inequality follows from the induction assumption.

(ii) $(\gamma \text{ terms}) = M_{i+1}\gamma\Delta_{n-1}(i+2) - 2M_i\gamma\Delta_{n-1}(i+1) + M_{i-1}\gamma\Delta_{n-1}(i)$

$$
= \begin{cases}
0 \text{ if } i \geqslant M+1 \\
\gamma\Delta_{n-1}(i) \leqslant 0 \text{ if } i = M \\
(M_i - 1)\gamma(\Delta_{n-1}(i+2) - \Delta_{n-1}(i+1)) \\
-(M_i + 1)\gamma(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)) \\
\leqslant -(M_i + 1)\gamma(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)) \text{ if } i \leqslant M-1.
\end{cases}
$$

(iii) $(S, \mu \text{ terms}) \equiv B = -S_n(i+1)(f(\mu_n(i+1)) + \mu_n(i+1)\Delta_{n-1}(i+1))$

$$
\begin{aligned}
&+ 2S_n(i)(f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i)) \\
&- S_n(i-1)(f(\mu_n(i-1)) + \mu_n(i-1)\Delta_{n-1}(i-1)).
\end{aligned}
$$

From the induction assumption, it follows as in Corollary 1 that $S_n(i)$ is increasing in $i$. Therefore, we examine expression $B$ by looking at the following four cases.

CASE 1. $S_n(i+1) = 0$. Then $B = 0$.

CASE 2. $S_n(i+1) = i+1$, $S_n(i) = 0$. Then

$$
\begin{aligned}
B &= -(i+1)(f(\mu_n(i+1)) + \mu_n(i+1)\Delta_{n-1}(i+1)) \\
&= -(i+1)\mu_n(i+1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)) \\
&\quad - (i+1)(f(\mu_n(i+1)) + \mu_n(i+1)\Delta_{n-1}(i)) \\
&\leqslant -(i+1)\mu_n(i+1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)) \\
&\quad - (i+1)(f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i)) \\
&\leqslant -(i+1)\mu_n(i+1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)),
\end{aligned}
$$

since $S_n(i) = 0$ implies $f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i) \geqslant 0$.

CASE 3. $S_n(i) = i$, $S_n(i-1) = 0$. Then

$$
\begin{aligned}
B &\leqslant -(i+1)\mu_n(i+1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)) \\
&\quad + (i-1)(f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i)) \\
&\leqslant -(i+1)\mu_n(i+1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)),
\end{aligned}
$$

since $S_n(i) = i$ implies $f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i) \leqslant 0$.

CASE 4. $S_n(i-1) = i-1$. In this case

$$
\begin{aligned}
B &\leqslant -(i+1)\mu_n(i+1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)) \\
&\quad + (i-1)\mu_n(i-1)(\Delta_{n-1}(i) - \Delta_{n-1}(i-1)) \\
&\leqslant -(i+1)\mu_n(i+1)(\Delta_{n-1}(i+1) - \Delta_{n-1}(i)).
\end{aligned}
$$

(iv) ($\Lambda$ terms) $= \Lambda(\Delta_{n-1}(i+1) - \Delta_{n-1}(i))$. The only thing left to do now is to combine the above "leftover" terms, all negative multiples of $\Delta_{n-1}(i+1) - \Delta_{n-1}(i)$, with this $\Lambda$ term to establish the result.

We finally show that $r - g(\lambda_{n+1}(i)) + \lambda_{n+1}(i)\Delta_n(i+1) \geqslant 0$ for $i \geqslant M+1$ to establish the induction. We have, for $i \geqslant M+1$,

$$\begin{aligned}\lambda_{n+1}(i)\Delta_n(i+1) = (\lambda_{n+1}(i)/(\Lambda+\alpha))\,[&N_{i+1}(r - g(\lambda_n(i+1)) \\ &+ \lambda_n(i+1)\Delta_{n-1}(i+2)) - S_n(i+1)(f(\mu_n(i+1)) + \mu_n(i+1)\Delta_{n-1}(i+1)) \\ &+ S_n(i)(f(\mu_n(i)) + \mu_n(i)\Delta_{n-1}(i)) - N_i(r - g(\lambda_n(i)) \\ &+ \lambda_n(i)\Delta_{n-1}(i+1)) + \Lambda\Delta_{n-1}(i+1)].\end{aligned}$$

By induction, $r - g(\lambda_n(i+1)) + \lambda_n(i+1)\Delta_{n-1}(i+2) \geqslant 0$. By an argument similar to above, it is easy to show that the combined terms with $S$ and $\mu$ are non-negative. Therefore, we have

$$\begin{aligned}\lambda_{n+1}(i)\Delta_n(i+1) &\geqslant (\lambda_{n+1}(i)/(\Lambda+\alpha))\,[-N_i(r - g(\lambda_n(i)) \\ &\quad + (\Lambda - N_i\lambda_n(i))\Delta_{n-1}(i+1)] \\ &\geqslant (\lambda_{n+1}(i)/(\Lambda+\alpha))[-N_i(r - g(\lambda_n(i)) \\ &\quad + (\Lambda - N_i\lambda_n(i))(-r + g(\lambda_n(i)))/\lambda_n(i)] \\ &= (\Lambda/(\Lambda+\alpha))\lambda_{n+1}(i)[(-r + g(\lambda_n(i)))/\lambda_n(i)]\end{aligned}$$

For fixed $i$, consider the graph of the function $k(\lambda) = -r + g(\lambda)$. By the condition of the theorem, the linear extension of this graph from $\lambda_{\min}$ back to $\lambda = 0$ stays nonpositive. If we recognize $k(\lambda)/\lambda$ as the slope of the line from $(0, 0)$ to $(\lambda, k(\lambda))$, then the convexity of $g$ and the fact that $\lambda_{n+1}(i) \leqslant \lambda_n(i)$ imply that $k(\lambda_{n+1}(i))/\lambda_{n+1}(i) \leqslant k(\lambda_n(i))/\lambda_n(i)$. Therefore,

$$\begin{aligned}\lambda_{n+1}(i)\Delta_n(i+1) &\geqslant (\Lambda/(\Lambda+\alpha))(-r + g(\lambda_{n+1}(i))) \\ &\geqslant -r + g(\lambda_{n+1}(i)),\end{aligned}$$

and the proof is complete.

The following corollary follows from Theorem 2 exactly as Corollary 1 follows from Theorem 1.

COROLLARY 2: Given the same conditions as in Theorem 2, $\lambda_n(i)$ is decreasing in $i$ and $S_n(i)$ and $\mu_n(i)$ are increasing in $i$.

Before proceeding, we remark that the differentiability of $g$ assumed in Theorem 2 is obviously not necessary. All we need is that a linear (or convex) extension of $g$ from $\lambda_{\min}$ back to $\lambda = 0$ remains sufficiently small, and this condition is most easily stated in terms of $g'$ when $g'$ exists.

We now examine how the optimal actions behave as functions of $\alpha$, $M$, and $N$. The additional subscripts will show the dependence of the various quantities upon these parameters.

THEOREM 3: If $h(N_i) - h(N_{i+1}) \geqslant g(\lambda_{\min})$ for $i \geqslant M$, then $\Delta_{n,\alpha_1}(i) \leqslant \Delta_{n,\alpha_2}(i)$ for discount factors $\alpha_1 \leqslant \alpha_2$, and all $n \geqslant 0$, $1 \leqslant i \leqslant M + N - 1$. (In this theorem and the following corollary, $h$ need not be linear.)

PROOF: For $n = 0$, the result is trivial. Assuming it is true for $n - 1$, let $\bar{V}_{n,\alpha}(i) = J_n(i, S_{n,\alpha}(i), \lambda_{n,\alpha}(i), \mu_{n,\alpha}(i))$ and write $\bar{\Delta}_{n,\alpha_1}(i) = \bar{V}_{n,\alpha}(i) - \bar{V}_{n,\alpha}(i-1)$. Then $\Delta_{n,\alpha_1}(i) - \Delta_{n,\alpha_2}(i) = \bar{\Delta}_{n,\alpha_1}(i)/(\Lambda + \alpha_1) - \bar{\Delta}_{n,\alpha_2}(i)/(\Lambda + \alpha_2)$. If we can show that $0 \geqslant \bar{\Delta}_{n,\alpha_2}(i) \geqslant \bar{\Delta}_{n,\alpha_1}(i)$ for each $i$, then the result will follow since $1/(\Lambda + \alpha_1) \geqslant 1/(\Lambda + \alpha_2)$. But $\bar{\Delta}_{n,\alpha}(i) \leqslant 0$ from Proposition 1, and $\bar{\Delta}_{n,\alpha_2}(i) \geqslant \bar{\Delta}_{n,\alpha_1}(i)$ by a proof almost identical to the proof of Theorem 1.

COROLLARY 3: If $h(N_i) - h(N_{i+1}) \geqslant g(\lambda_{\min})$ for $i \geqslant M$, then $\lambda_{n,\alpha}(i)$ is increasing in $\alpha$ and $S_{n,\alpha}(i)$ and $\mu_{n,\alpha}(i)$ are decreasing in $\alpha$. That is, the more the future is discounted, the less the system works.

The following theorem and its corollary say that if we compare two systems in states which have the same number of working machines, the only difference being that one system has an additional spare, then the system with less spares will work harder.

THEOREM 4: Consider two systems which are identical in every respect except that one has $M$ spares while the other has $M + 1$ spares. Assume that the conditions of Theorem 2 hold for each system. Then $\Delta_{n,M}(i) \leqslant \Delta_{n,M+1}(i+1)$ for $n \geqslant 0$, $1 \leqslant i \leqslant M + N$, if $\Lambda$ is redefined as $\Lambda = N\lambda_{\max} + (M+1)\gamma + (N+M+1)\mu_{\max}$ for each system.

PROOF: We only sketch the proof, since the details are similar to those in Theorem 2. First, notice that $N_{i,M} = N_{i+1,M+1}$ and $M_{i,M} = M_{i+1,M+1}$, since the only difference between state $i$ in the first system and state $i + 1$ in the second system is the one extra machine in repair. This observation allows us to group the $\lambda$ and $\gamma$ terms of $\Delta_{n,M}(i) - \Delta_{n,M+1}(i+1)$ in the obvious manner, while the $r$ terms cancel. Finally, for the $S, \mu$ terms, the observation that $S_{n,M}(i-1) = i - 1$ or $S_{n,M+1}(i+1) = i + 1$ implies $S_{n,M}(i) = i$ and that $S_{n,M+1}(i) = i$ implies $S_{n,M+1}(i+1) = i + 1$ defines the possible cases we need to consider. From this point, the proof is entirely analogous to the proof of Theorem 2.

COROLLARY 4: Given the same conditions as in Theorem 4, $\lambda_{n,M+1}(i+1) \geqslant \lambda_{n,M}(i)$, $\mu_{n,M}(i) \geqslant \mu_{n,M+1}(i+1)$, and if $S_{n,M+1}(i+1) = i + 1$, then $S_{n,M}(i) = i$.

We next consider two systems with the same number of spares, but where one system has one more working machine than the other. If we compare two states where each system has the same number of machines in repair, then we find the possibly surprising result that the system with more working machines works harder.

THEOREM 5: Consider two systems which are identical in every respect except that the first has $N$ working machines while the second has $N + 1$ working machines. Assume that the conditions of Theorem 2 hold for each system. Then $\Delta_{n,N+1}(i) \leqslant \Delta_{n,N}(i)$ for $n \geqslant 0$, $1 \leqslant i \leqslant M + N$, if $\Lambda$ is redefined as $\Lambda = (N+1)\lambda_{\max} + M\gamma + (M+N+1)\mu_{\max}$ for each system.

PROOF: Again we only sketch the proof. First we notice that $N_{i,N+1} = N_{i,N} + 1$ and $M_{i,N+1} = M_{i,N}$, since the extra machine in the second system will be working for any state $i$. This allows us to group the $\lambda$ and $\gamma$ terms in $\Delta_{n,N+1}(i) - \Delta_{n,N}(i)$ in an obvious way and the $r$ terms cancel. Finally we deal with the $S, \mu$ terms in the standard way by using the inequalities $S_{n,N}(i-1) \leqslant S_{n,N}(i) \leqslant S_{n,N+1}(i)$ and $S_{n,N}(i-1) \leqslant S_{n,N+1}(i-1) \leqslant S_{n,N+1}(i)$ to establish the possible cases.

COROLLARY 5: Given the same conditions as in Theorem 5, $\lambda_{n,N}(i)$ is decreasing in $N$ and $\mu_{n,N}(i)$ and $S_{n,N}(i)$ are increasing in $N$.

In our final model of this section, we examine the case where there are no spares, that is, $M = 0$. Then $N_i = N - i$ and $M_i = 0$. Here we find that if $h$ is linear, then $\lambda_n(i)$ and $\mu_n(i)$ are independent of $i$, and for a given $n$ we either never turn on servers, or we always turn on one server for each broken down machine.

THEOREM 6: Assume $h(j) = jr$, $0 \leqslant j \leqslant N$, and $M = 0$. Then $\Delta_n(i) = \Delta_n(i+1)$ for $n \geqslant 0$ and $1 \leqslant i \leqslant N-1$. Therefore, $\lambda_n(i) = \lambda_n(i+1)$ and $\mu_n(i) = \mu_n(i+1)$ for all $i$, and either $S_n(i) = i$ for all $i$ or $S_n(i) = 0$ for all $i$.

PROOF: For $n = 0$, the result is trivial. Assume $\Delta_{n-1}(i) = \Delta_{n-1}(i+1)$ for each $i$. This obviously implies that $\lambda_n(i)$ and $\mu_n(i)$ are constant functions of $i$ and that $S_n(i) = i$ for all $i$ or $S_n(i) = 0$ for all $i$. Letting $\lambda = \lambda_n(i)$, $\mu = \mu_n(i)$, and $\Delta = \Delta_{n-1}(i)$, this yields

$$\begin{aligned}(\Lambda + \alpha)\Delta_n(i) &= (N-i)(r - g(\lambda) + \lambda\Delta) - S_n(i)(g(\mu) + \mu\Delta) \\ &\quad - (N-i+1)(r - g(\lambda) + \lambda\Delta) + S_n(i-1)(g(\mu) + \mu\Delta) + \Lambda\Delta \\ &= -(r - g(\lambda) + \lambda\Delta) - (S_n(i) - S_n(i-1))(g(\mu) + \mu\Delta) + \Lambda\Delta.\end{aligned}$$

Since $S_n(i) - S_n(i-1) = 1$ for all $i$ or 0 for all $i$, we see that $\Delta_n(i)$ is constant in $i$, and the result is proved.

As a corollary to this result, consider the two systems which are identical except that their failure rates, $\lambda_1$ and $\lambda_2$, no longer considered control variables, are different. That is, suppose each working machine in system $i$ fails at rate $\lambda_i$, where $\lambda_1 < \lambda_2$. Also assume $M = 0$ and that $S$ and $\mu$ are the only control variables. Intuitively, we might expect that the system with the larger failure rate would work harder to repair to its machines. This result turns out to be false. In fact, its exact opposite is true, as exhibited in the following corollary.

COROLLARY 6: Given the above model with $g(\lambda) = 0$ and $h(j) = jr$, $\Delta_{n,\lambda_1}(i) \leqslant \Delta_{n,\lambda_2}(i)$ for each $n \geqslant 0$, $1 \leqslant i \leqslant N$, and $\lambda_{\min} \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \lambda_{\max}$. Hence $\mu_{n,\lambda_1}(i) \geqslant \mu_{n,\lambda_2}(i)$ and $S_{n,\lambda_1}(i) \geqslant S_{n,\lambda_2}(i)$ for each $n \geqslant 1$, $1 \leqslant i \leqslant N$.

PROOF: By Theorem 6, we can write $\Delta_{n,\lambda_j}(i) = \Delta_{n,j}$ and $\mu_{n,\lambda_j}(i) = \mu_{n,j}$ for each $i$. Therefore,

$$\begin{aligned}(\Lambda + \alpha)(\Delta_{n,\lambda_1}(i) - \Delta_{n,\lambda_2}(i)) &= (N-i)(\lambda_1\Delta_{n-1,1} - \lambda_2\Delta_{n-1,2}) \\ &\quad - (N-i+1)(\lambda_1\Delta_{n-1,1} - \lambda_2\Delta_{n-1,2}) - (S_{n,\lambda_1}(i) - S_{n,\lambda_1}(i-1)) \\ &\quad \cdot (f(\mu_{n,1}) + \mu_{n,1}\Delta_{n-1,1}) + (S_{n,\lambda_2}(i) - S_{n,\lambda_2}(i-1)) \\ &\quad \cdot (f(\mu_{n,2}) + \mu_{n,2}\Delta_{n-1,2}) + \Lambda(\Delta_{n-1,1} - \Delta_{n-1,2}).\end{aligned}$$

By induction, $(S_{n,\lambda_1}(i) - S_{n,\lambda_1}(i-1)) - (S_{n,\lambda_2}(i) - S_{n,\lambda_2}(i-1)) = 0$ or 1. Both cases may be handled as in Theorem 2 to dispose of the $S$, $\mu$ terms (after possibly combining part of them with the $\Lambda$ term). What is left is

$$= (\Lambda - \lambda_1)(\Delta_{n-1,1} - \Delta_{n-1,2}) + (\lambda_2 - \lambda_1)\Delta_{n-1,2} \leqslant 0.$$

In general, if $M > 0$, the result of the above corollary seems to be partly true and partly false. For most small states ($i \leqslant M$), it seems that $\mu_{n,\lambda_1}(i) \leqslant \mu_{n,\lambda_2}(i)$, but that for most large states ($i > M$), $\mu_{n,\lambda_1}(i) \geqslant \mu_{n,\lambda_2}(i)$. The following numerical example, whose results are for the original continuous time, infinite horizon problem, is typical. In this example, $N = 2$, $M = 1$, $r = 200$, $\alpha = 25$, and $f(\mu) = .05\mu^2$.

| $\lambda = 1$ | | | | $\lambda = 5$ | | | | $\lambda = 20$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $i$ | $S(i)$ | $\mu(i)$ | $V(i)$ | $i$ | $S(i)$ | $\mu(i)$ | $V(i)$ | $i$ | $S(i)$ | $\mu(i)$ | $V(i)$ |
| 0 | 0 | 0 | 15.98 | 0 | 0 | 0 | 15.67 | 0 | 0 | 0 | 13.44 |
| 1 | 1 | 2.39 | 15.74 | 1 | 1 | 8.35 | 14.83 | 1 | 1 | 16.00 | 11.84 |
| 2 | 2 | 33.65 | 12.38 | 2 | 2 | 32.73 | 11.56 | 2 | 2 | 29.20 | 8.92 |
| 3 | 3 | 37.84 | 8.59 | 3 | 3 | 36.34 | 7.92 | 3 | 3 | 31.12 | 5.81 |

We end this section by noting that all of the previous results are true for the infinite horizon continuous time problem. This follows by letting $n \to \infty$ and using the results from Lippman [14]. Furthermore, because of Theorem 1 and the results of Lippman [15], these monotonicity properties also hold for the finite horizon continuous time problem. To state these results, we will drop the subscript $n$ to denote the infinite horizon problem, and we will substitute $t$ for $n$ to denote that there is $t$ time left in the finite horizon problem.

THEOREM 7: Suppose $h(j) = jr$, $j \geq 0$, and the assumptions of Theorem 2 hold. Then for the infinite horizon problem, $V(i)$ is concave and decreasing in $i$, $\mu(i)$ and $S(i)$ are increasing in $i$ and $N$ and are decreasing in $\alpha$, and $\lambda(i)$ is decreasing in $i$ and $N$ and is increasing in $\alpha$. Also, $\lambda_{M+1}(i+1) \geq \lambda_M(i)$, $\mu_M(i) \geq \mu_{M+1}(i+1)$, and if $S_{M+1}(i+1) = i+1$, then $S_M(i) = i$. For the finite horizon problem with $t$ time remaining, these same properties hold for $V_t(i)$, $\lambda_t(i)$, $S_t(i)$, and $\mu_t(i)$. Finally, $\mu_t(i)$ and $S_t(i)$ are increasing in $t$, and $\lambda_t(i)$ is decreasing in $t$.

## 3. SUMMARY AND POSSIBLE EXTENSIONS

In this paper we have been able to extend previous results concerning the structure of optimal policies for the machine repair problem. Namely, we have let $\lambda$ and $S$, as well as $\mu$, be control variables and we have allowed spares to fail before they are put into use. There are still, however, many other complications of this model which are worthy of further study. Unfortunately, the ones we have in mind will probably present many more mathematical difficulties than the model we have examined here.

One obvious feature which would be good to include is the presence of fixed costs for turning on or turning off servers. Since there has been only limited success with this type of queuing control in other, probably less complicated, queuing models, we are not optimistic about success here. Another possible extension, or set of extensions, is to complicate the problem in such a way that there are two or more state variables. For example, we could assume that the time to install spares is an exponential random variable, or that failed spares go to their own repair facility, or that working machines may fail in more than one way and that each type of failure requires a different repair facility (see [11], [12]). In each of these models, the state space is multidimensional, which serves to complicate the mathematics immensely. In fact, we have investigated these models quite extensively. Unfortunately, aside from an analogue of Theorem 6, not any of even the most intuitive structural results could be proved by the above induction arguments. At the same time, no counterexamples to our conjectures were found. Hopefully, researchers in the future will discover how (or if) Lippman's "new device for exponential queuing systems" is able to solve these models with multidimensional state spaces.

## REFERENCES


[1] Albright, S.C., and W. Winston, "A Birth-Death Model of Advertising and Pricing," Advances in Applied Probability, *11*, 134-152 (1979).

[2] Anderson, M., "Monotone Optimal Maintenance Policies for Equipment Subject to Markovian Deterioration," Doctoral Dissertation, Indiana University, Bloomington, Ind. (1977).

[3] Barlow, R.E., "Repairman Problems," Chapter 2 in *Studies in Applied Probability and Management Science,* ed. Arrow, Karlin, and Scarf (Stanford University Press, Stanford, Ca. 1962).

[4] Cinlar, E., "Optimal Operating Policy for the Machine Repair Problem with Two Service Stations," Technical Report No. 266-3, Control Analysis Corp, Palo Alto, Ca. (1972).

[5] Crabill, T., "Optimal Control of a Maintenance System with Variable Service Rates," Operations Research, *22*, 736-745 (1975).

[6] Denardo, E., "Contraction Mappings in the Theory Underlying Dynamic Programming," SIAM Review, *9*, 165-177 (1967).

[7] Deshmukh, S.D., and W. Winston, "A Controlled Birth and Death Process Model of Optimal Product Pricing Under Stochastically Changing Demand," Journal of Applied Probability, *14*, 328-339 (1977).

[8] Gnedenko, B.V., Y.K. Belyayev, and A.D. Solovyev, *Mathematical Methods of Reliability Theory* (Academic Press, New York 1969).

[9] Goheen, L., "On the Optimal Operating Policy for the Machine Repair Problem when Failure and Repair Times Have Erlang Distribution," Operations Research, *25*, 484-492 (1977).

[10] Howard, R.A., *Dynamic Programming and Markov Processes* (M.I.T. Press, Cambridge, Mass 1960).

[11] Iglehart, D.L., and A. Lemoine, "Approximations for the Repairman Problem with Two Repair Facilities, I: No Spares," Advances in Applied Probability, *5*, 595-613 (1973).

[12] Iglehart, D.L., and A. Lemoine, "Approximations for the Repairman Problem with Two Repair Facilities, II: Spares," Advances in Applied Probability, *6*, 147-158 (1974).

[13] Kakumanu, P., "Relation Between Continuous and Discrete Time Markovian Decision Problems," Naval Research Logistics Quarterly, *24*, 431-439 (1977).

[14] Lippman, S.A., "Applying a New Device in the Optimization of Exponential Queuing Systems," Operations Research, *23*, 687-710 (1975).

[15] Lippman, S.A., "Countable-State, Continuous-Time Dynamic Programming with Structure," Operations Research, *24*, 477-490 (1976).

[16] Ross, S.M., *Applied Probability Models with Optimization Applications* (Holden-Day, San Francisco 1970).

[17] Serfozo, R., "An Equivalence Between Continuous and Discrete Time Markov Decision Processes," Tech. Report, Dept. of I.E. and O.R., Syracuse University (1976).

[18] Winston, W., "Optimal Control of Discrete and Continuous Time Maintenance Systems with Variable Service Rates," Operations Research, *25*, 259-268 (1977).

# BENDERS' PARTITIONING SCHEME APPLIED TO A NEW FORMULATION OF THE QUADRATIC ASSIGNMENT PROBLEM*

Mokhtar S. Bazaraa

*School of Industrial and Systems Engineering*
*Georgia Institute of Technology*
*Atlanta, Georgia*

Hanif D. Sherali

*School of Industrial Engineering and Operations Research*
*Virginia Polytechnic Institute and State University*
*Blacksburg, Virginia*

**ABSTRACT**

In this paper we present a new formulation of the quadratic assignment problem. This is done by transforming the quadratic objective function into a linear objective function by introducing a number of new variables and constraints. The resulting problem is a 0-1 linear integer program with a highly specialized structure. This permits the use of the partitioning scheme of Benders where only the original variables need be considered. The algorithm described thus iterates between two problems. The master problem is a pure 0-1 integer program, and the subproblem is a transportation problem whose optimal solution is shown to be readily available from the master problem in closed form. Computational experience on problems available in the literature is provided.

## 1. INTRODUCTION

In this paper, we revisit the quadratic assignment problem which was first formulated by Koopmans and Beckmann [19] for assigning $m$ indivisible entities, called facilities, to $m$ mutually exclusive locations. More specifically, we consider the following most general formulation of this problem as introduced by Graves and Whinston [14].

$$\textit{QAP 1: minimize } \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij}\, x_{ij} + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{n=1}^{p} b_{ijkl}^{n}\, x_{ij}\, x_{kl} + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{n=1}^{p} f_{ik}^{n}\, d_{jl}^{n}\, x_{ij}\, x_{kl}$$

subject to

$$x \in X_A = \left\{ (x_{11}, \ldots, x_{mm}) : \sum_{i=1}^{m} x_{ij} = 1,\ j = 1, \ldots m,\ \sum_{j=1}^{m} x_{ij} = 1,\ i = 1, \ldots, m,\ x_{ij} = 0, 1,\ i,j = 1, \ldots, m \right\} \tag{1.1}$$

Here there are $p$ products that flow among the $m$ facilities. Particularly, $f_{ik}^{n}$ is the amount of flow of product $n$ from facility $i$ to facility $k$ and $d_{jl}^{n}$ is a distance measure from location $j$ to

*This study is supported under NSF Grant # ENG77-07468

location $l$ when transporting product $n$. Further, $a_{ij}$ is the fixed cost of locating facility $i$ at location $j$, and $b^n_{ijkl}$ is a fixed cost for product $n$ dependent on a pair of assignments, viz, facility $i$ to location $j$ and facility $k$ to location $l$. Note, that without loss in generality, we can take $b^n_{ijkl} = 0$ if $i = k$ or $j = l$ and also, $f^n_{ii} = d^n_{jj} = 0$ for $i, j = 1, \ldots, m$, $n = 1, \ldots, p$.

Using a simple transformation introduced by Lawler [22], and extended by Pierce and Crowston [25], the above problem may be written as:

$$QAP\ 2\text{:}\quad \underset{x \in X_A}{\text{minimize}} \left\{ x^t\, Sx = \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} \sum_{l=1}^{m} s_{ijkl}\, x_{ij}\, x_{kl} \right\}$$

where,

$$s_{ijkl} = \begin{cases} \sum_{n=1}^{p} b^n_{ijkl} + \sum_{n=1}^{p} f^n_{ik}\, d^n_{jl} & \text{if } i \neq k \text{ or } j \neq l \\ a_{ij} & \text{otherwise} \end{cases} \tag{1.2}$$

and where a superscript $t$, will throughout this paper, denote the transpose operation. However, we will find it more convenient to use the following transformation:

$$c_{ijkl} = \frac{a_{ij} + a_{kl}}{(m-1)} + \sum_{n=1}^{p} (b^n_{ijkl} + b^n_{klij}) + \sum_{n=1}^{p} (f^n_{ik}\, d^n_{jl} + f^n_{ki}\, d^n_{lj}),$$

$$i = 1, \ldots, m-1,\ k = i+1, \ldots, m,\ l, j = 1, \ldots, m,\ l \neq j. \tag{1.3}$$

It may be easily seen that this transformation leads to the equivalent problem

$$QAP\ 3\text{:}\quad \underset{x \in X_A}{\text{minimize}} \sum_{i=1}^{m-1} \sum_{j=1}^{m} \sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} c_{ijkl}\, x_{ij}\, x_{kl}$$

The quadratic assignment problem has enjoyed great popularity mostly because of its versatile applicability, but partly because of its insurmountable resistance to efficient solution techniques. For a review on quadratic assignment problems, we refer the reader to Francis and White [6], Gaschütz and Ahrens [8], Gavett and Plyter [9], Gilmore [11], Graves and Whinston [14], Hanan and Kurtzberg [15], Heider [16], Hillier and Connors [17], Koopmans and Beckmann [19], Land [20], Lawler [22], and Pierce and Crowston [25]. For computational studies on this problem, we refer the reader to Nugent, Vollmann and Ruml [23] and, more recently, to an excellent study by Burkard and Stratmann [2].

In the following sections, we first employ a transformation which converts problem QAP 3 into a mixed integer linear program and thus makes it most amenable to the partitioning scheme of Benders [1]. This latter scheme decomposes the problem into a linear integer master program and several linear subproblems which are essentially transportation problems. The solution to the subproblems is shown to be readily available in closed form. We then develop an approach to the master problem. Finally, computational results and certain relevant concluding remarks are presented.

## 2. REFORMULATING THE QUADRATIC ASSIGNMENT PROBLEM

For the purpose of this development, let us introduce $m^2(m-1)^2/2$ new variables,

$$y_{ijkl} = x_{ij}\, x_{kl} \quad \text{for } i = 1, \ldots, m-1,\ k = i+1, \ldots, m,\ j, l = 1, \ldots, m,\ j \neq l \tag{2.1}$$

In addition, introducing $2m(m-1)$ new linear constraints, we formulate the following linear mixed integer problem which is then shown to be equivalent to problem QAP 3.

*QAP 4*: minimize $\sum_{i=1}^{m-1} \sum_{j=1}^{m} \sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} c_{ijkl}\, y_{ijkl}$

subject to

(2.2) $$\sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} y_{ijkl} - (m-i)\, x_{ij} = 0 \quad \text{for } i = 1, \ldots, m-1,\ j = 1, \ldots, m$$

(2.3) $$\sum_{i=1}^{k-1} \sum_{\substack{j=1 \\ j \neq l}}^{m} y_{ijkl} - (k-1)\, x_{kl} = 0 \quad \text{for } k = 2, \ldots, m,\ l = 1, \ldots, m$$

(2.4) $$\sum_{j=1}^{m} x_{ij} = 1 \quad \text{for } i = 1, \ldots, m$$

(2.5) $$\sum_{i=1}^{m} x_{ij} = 1 \quad \text{for } j = 1, \ldots, m$$

(2.6) $$x_{ij} \text{ binary} \quad \text{for } i,j = 1, \ldots, m$$

(2.7) $$-y_{ijkl} \geqslant -1 \quad \text{for } i = 1, \ldots, m-1,\ k = i+1, \ldots, m, j,l = 1, \ldots, m,\ j \neq l$$

(2.8) $$y_{ijkl} \geqslant 0 \quad \text{for } i = 1, \ldots, m-1,\ k = i+1, \ldots, m, j,l = 1, \ldots, m, j \neq l$$

Note that problem QAP 4 has $m^2$ integer and $m^2(m-1)^2/2$ continuous variables and $2m^2$ linear constraints. As opposed to this, Lawler's [22] linear integer formulation involves $m^2 + m^4$ integer variables and $m^4 + 2m + 1$ constraints. A more recent linearization technique due to Kaufman and Broeckx [18] results in a mixed integer program with $m^2$ zero-one and $m^2$ continuous variables and $m^2 + 2m$ constraints. In any case, we will be demonstrating how Problem QAP 4 can be solved by exploiting its structure and handling explicitly only its $m^2$ zero-one variables and its $2m$ linear assignment constraints (2.4), (2.5).

Before proceeding to propose a solution technique for problem QAP 4 we first establish that every feasible solution to Problem QAP 4 must satisfy Equation (2.1). As a consequence of this result, Theorem 1 stating that Problems QAP 3 and QAP 4 are equivalent, follows.

LEMMA 1: Let $(x,y)$ be any feasible solution to Problem QAP 4. Then, $y_{ijkl} = x_{ij}x_{kl}$ for $i = 1, \ldots, m-1,\ k = i+1, \ldots, m,\ j,l = 1, \ldots, m,\ j \neq l$.

PROOF: First of all, note that for any $p,q,r,s,\ p < r, q \neq s$, $x_{pq}x_{rs} = 0$ implies from Equation (2.2), (2.3) that $y_{pqrs} = 0$. Now consider $p,q,r,s, p < r, q \neq s$ such that $x_{pq} = x_{rs} = 1$. It suffices to show that this implies $y_{pqrs} = 1$. Since $x_{rs} = 1$, then by Equation (2.3), we have,

(2.9) $$\sum_{i=1}^{r-1} \sum_{\substack{j=1 \\ j \neq s}}^{m} y_{ijrs} = r - 1$$

But since $x$ satisfies the assignment constraints (2.4) through (2.6), and since $x_{rs} = 1$, there are precisely $r-1$ nonzero variables $x_{ij}$ for $i < r,\ j \neq s$. Since $x_{ij} = 0$ implies $y_{ijrs} = 0$ from

above, this in turn means that there are at most $r-1$ nonzero variables $y_{ijrs}$ for $i < r$, $j \neq s$. However, in view of Equations (2.7), (2.8), (2.9), we must have precisely $r-1$ unit variables $y_{ijrs}$, each corresponding to $x_{ij} = 1$ for $i < r$, $j \neq s$. In particular, since $x_{pq} = 1$, $p < r$, $q \neq s$, then $y_{pqrs} = 1$. This completes the proof.

THEOREM 1: Problems QAP 3 and QAP 4 are equivalent in the following sense. For every feasible solution to Problem QAP 3 there corresponds a feasible solution to QAP 4 with the same objective function value and vice versa.

PROOF: Let $x$ be a feasible solution to Problem QAP 3 and let us define the vector $y$ through Equation (2.1). We will show that $(x,y)$ qualifies for the corresponding solution we are seeking to Problem QAP 4. Now, Equation (2.1) implies that

$$(2.10) \qquad \sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} y_{ijkl} - \sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} x_{ij}x_{kl} = 0 \quad \text{for } i = 1, \ldots, m-1, \; j = 1, \ldots, m$$

If for any such $i,j$, we have $x_{ij} = 0$, then from Equation (2.10), Equation (2.2) must hold. On the other hand, if $x_{ij} = 1$, then since we must have $x_{kj} = 0$ for each $k > i$, we get $\sum_{k=i+1}^{m} x_{kj} = 0$. Thus, using Equation (2.4), we have,

$$(2.11) \qquad \sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} x_{ij}x_{kl} = \sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} x_{kl} + \sum_{k=i+1}^{m} x_{kj} = \sum_{k=i+1}^{m} \sum_{l=1}^{m} x_{kl} = (m-i)$$

Thus, from Equations (2.10), (2.11), again Equation (2.2) holds. Similarly, one may show that Equations (2.3) holds. Also, from the definition of $y$ through Equation (2.1), we have that Equations (2.7), (2.8) hold and that the objective function values of Problems QAP 3 and QAP 4 are equal.

Conversely, let $(x,y)$ be a feasible solution to QAP 4. We will show that $x$ qualifies for the corresponding solution we are seeking to Problem QAP 3. Clearly, $x$ is feasible to QAP 3. That the objective function values are equal follows from Lemma 1 which asserts that Equation (2.1) holds. This completes the proof.

We now proceed to develop a solution procedure for Problem QAP 4. This technique exploits the following special structure of the problem. The variables $x$ are restricted to be extreme points of the assignment polytope, or points in the set $X_A$. For each such point, the resulting problem in $j$ is a transportation problem whose solution, as seen above, is given through Equation (2.1). This structure is most efficiently exploited by Benders' Decomposition scheme [1].

## 3. APPLICATION OF THE PARTITIONING SCHEME OF BENDERS.

In this section, we will decompose Problem QAP 4 into a linear integer master problem in $m^2$ zero-one variables and a linear subproblem, using the method due to Benders (see Benders [1] and Lasdon [21]). At each iteration, the master problem generates a point of $X_A$, based on which a suitable subproblem is solved to generate a cut-constraint. This cut is then appended to the other constraints in the master problem and the solution to the latter is updated. The procedure hence iterates between the master problem and the subproblem until a suitable termination criterion is met in a finite number of steps.

To conduct such a decomposition, observe that for a fixed $x \in X_A$, problem QAP 4 is a transportation problem in the $y$-variables over a set, say, $Y(x)$ defined by constraints (2.2), (2.3), (2.7) and (2.8). Thus, we may write QAP 4 as

$$QAP\ 4\text{: minimize}_{x \in X_A} \left[ \text{minimum}_{y \in Y(x)} \sum_{i=1}^{m-1} \sum_{j=1}^{m} \sum_{k>i}^{m} \sum_{l \neq j}^{m} c_{ijkl}\, y_{ijkl} \right].$$

For the inner minimization problem above, letting $u_{ij}$, $i = 1, \ldots, m-1$, $j = 1, \ldots, m$ be the dual variables associated with constraints (2.2), $v_{kl}$, $k = 2, \ldots, m$, $l = 1, \ldots, m$ those with (2.3) and $w_{ijkl}$, $i = 1, \ldots, m-1$, $k = i+1, \ldots, m$, $l,j = 1, \ldots, m$, $l \neq j$ those with (2.7), we may use the linear programming dual to rewrite the above problem as

$$QAP\ 4\text{: minimize}_{x \in X_A} [SP(x)]$$

where, for a fixed $\bar{x} \in X_A$, we have,

$$SP(\bar{x})\text{: maximize} \left[ \sum_{i=1}^{m-1} \sum_{j=1}^{m} (m-i)\, u_{ij}\, \bar{x}_{ij} + \sum_{k=2}^{m} \sum_{l=1}^{m} (k-1)\, v_{kl}\, \bar{x}_{kl} - \sum_{i=1}^{m-1} \sum_{k=i+1}^{m} \sum_{l=1}^{m} \sum_{\substack{j=1 \\ j \neq l}}^{m} w_{ijkl} \right]$$

subject to

$$\left. \begin{aligned} u_{ij} + v_{kl} - w_{ijkl} &\leqslant c_{ijkl} \\ w_{ijkl} &\geqslant 0 \end{aligned} \right] \text{ for } i = 1, \ldots, m-1,\ k = i+1, \ldots, m,\ j,\ l \underset{j \neq l}{=} 1, \ldots, m$$

Now, using the standard transformation of replacing unrestricted variables by the difference of two nonnegative variables, one may show that $SP(\bar{x})$ attains an optimal solution at an extreme point of the resulting constraint set. Thus, letting $(u^p, v^p, w^p)$, $p \in E = \{1, \ldots, P\}$ be the finite number of points corresponding to the extreme points of this set, then $SP(\bar{x})$, and hence, QAP 4 may be rewritten with obvious notation as

$$QAP\ 4\text{: minimize}_{x \in X_A} \left[ \text{maximum}_{p \in E = \{1, \ldots, P\}} (a_p^t x - \alpha_p) \right]$$

Finally, the above problem may be represented as the master problem

$$MP(E)\text{: minimize } z$$

$$\text{subject to} \quad z \geqslant a_p^t x - \alpha_p \text{ for each } p \in E$$

$$x \in X_A$$

The scheme of Benders' asserts that one need not generate the entire set $E$, but rather, generate suitable elements of $E$ as and when needed. Thus, at any particular stage $r$, letting $E_r \subset E$, $1 \leqslant |E_r| = r \leqslant P$, and denoting the corresponding relaxed master program by $MP(E_r)$, the following scheme is validated by Benders.

Initialization: Let $z^1 = -\infty$, choose an arbitrary, preferably good, solution $x^1 \in X_A$, set $r = 1$ and $E_0 = \phi$, and go to Step 2.

Step 1: Solve the master problem $MP(E_r)$. Let $(x^{r+1}, z^{r+1})$ be an optimal solution. Increment $r$ by one and go to Step 2.

Step 2: Solve the subproblem $SP(x^r)$. Let $(u^r, v^r, w^r)$ be an optimal extreme point solution with objective function $\bar{z}$. If $\bar{z} = z^r$, terminate with $x^r$ as an optimal solution to the quadratic assignment problem. Otherwise, $\bar{z} > z^r$, and add the cut

$$z \geqslant a_r^t x - \alpha_r$$

to the current master problem. Letting $E_r = E_{r-1} \cup \{r\}$, go to Step 1.

The cardinality of $E$ being finite, the above procedure is finitely convergent since if any element of $E$ is re-generated then the termination criterion of Step 2 is necessarily met (see [1] or [21]). Over the next two subsections, we will now discuss the solution strategies for the master problem and the subproblem.

### 3.1 On Solving the Subproblem

Recall that for the sake of finiteness of Benders' scheme, we need to generate an optimal extreme point solution to problem $SP(x^r)$ where all unrestricted variables are transformed into nonnegative variables. Writing the Kuhn Tucker conditions for problem $SP(x^r)$ and using a rather elaborate proof, one can show (see [27]) that the following solution qualifies, as such an extreme point solution:

$$u_{ij}^{r+1} = \begin{cases} \underset{\substack{(k,l):\, x_{kl}^r = 1 \\ k > i}}{\text{maximum}} \, [c_{ijkl}] & \text{if } x_{ij}^r = 1 \\ & \text{for } i = 1, \ldots, m-1, j = 1, \ldots, m \\ \underset{(k,l):\, k > i}{\text{minimum}} \, [c_{ijkl} - v_{kl}^{r+1}] & \text{if } x_{ij}^r = 0 \end{cases}$$

$$v_{kl}^{r+1} = \begin{cases} 0 \quad \text{if } x_{kl}^r = 1 \\ \text{minimum} \left[ \underset{\substack{(i,j):\, x_{ij}^r = 1 \\ i < k}}{\text{minimum}} \{c_{ijkl} - u_{ij}^{r+1}\}, \; \underset{\substack{(i,j):\, x_{ij}^r = 0 \\ i < k}}{\text{minimum}} \{c_{ijkl}\} \right] & \text{for } k = 2, \ldots, m, l = 1, \ldots, m \end{cases}$$

$$w_{ijkl}^{r+1} = \begin{cases} 0 & \text{if } x_{ij}^r x_{kl}^r = 0 \\ & \text{for } i = 1, \ldots, m-1, k = i+1, \ldots, m, \; j,l = 1, \ldots, m, l \neq j. \\ u_{ij}^{r+1} - c_{ijkl} & \text{if } x_{ij}^r \, x_{kl}^r = 1 \end{cases}$$

We briefly note at this point that one may write out an alternative optimal extreme point solution by simply interchanging the roles of the variables $u$ and $v$ above. Through computational testing we found that, depending on the data structure, one of these solutions yields cuts that perform significantly better than those obtained from the other solution.

### 3.2 On Solving the Master Problem

First of all, note that the master problem need not be solved exactly at each iteration. Recall that in using Benders' decomposition, we are merely interested in generating elements of $E = \{1, \ldots, P\}$ until such time as the value of the master problem equals the current best value of Problem QAP 4. Moreover, while accomplishing this, we wish to ensure for the sake of finiteness that we do not regenerate any point of $E$ until the termination criterion is met. Now suppose we find a possibly suboptimal solution $(\hat{x}^{r+1}, \hat{z}^{r+1})$ to the master problem at stage $r$. If necessary, we then update the current best solution using $\hat{x}^{r+1}$.

Let the current best value be $\bar{z}$ and suppose that $\hat{z}^{r+1} < \bar{z}$. Then clearly, the exact solution value $z^{r+1}$ of the master problem satisfies

$$z^{r+1} \leqslant \hat{z}^{r+1} < \bar{z}$$

Moreover, we can also assert that $\hat{x}^{r+1}$ has not been previously generated. For, by contradiction, if it has been generated, at the $q^{th}$ stage, $q < r$, then let the corresponding Benders' cut be $z \geqslant a_q^t x - \alpha_q$. Noting through Lemma 1 that the subproblem yields the same value as the quadratic assignment problem for any given solution, we get $\bar{z} \leqslant a_q^t \hat{x}^{r+1} - \alpha_q \leqslant \hat{z}^{r+1}$ which contradicts $\hat{z}^{r+1} < \bar{z}$.

Thus, so long as an approximate solution yields a value for the master problem which is strictly less than that of the current best solution to the quadratic assignment problem, including the value of the new point generated, we may use this approximate solution with the assurance that it has not previously been generated. It is only when this condition does not hold that we need to solve the master problem exactly. Before discussing the proposed approximate and exact solution techniques below, we emphasize that this modification still maintains the scheme as an exact, finitely convergent procedure.

### 3.2.1 Scheme for Generating an Approximate Solution to the Master Problem

In this section, we will develop a heuristic to obtain a good quality approximate solution to the master problem $MP(E_r)$. Towards this end, suppose we compute

(3.1)
$$\beta_p = \text{maximum } \{1, \ |\text{minimum}(a_p^t x - \alpha_p)|\} \text{ for each } p \in E_r$$

and formulate the problem

$$\hat{MP}(E_r)\text{: minimize}\{\Lambda_r^t x\text{: } x \in X_A\}$$

where,

(3.2)
$$\Lambda_r = \sum_{p=1}^{r} \left(\frac{1}{\beta_p}\right) \alpha_p$$

Consider the solution obtained through the linear assignment problem $\hat{MP}(E_r)$. Since $\Lambda_r^t x$ is essentially the sum of the terms $\alpha_p^t x$, $p \in E_r$, each normalized by the magnitude $\beta_p$, Problem $\hat{MP}(E_r)$ tends not to minimize any cut expression too much at the expense of increasing the value of others. In other words, the solution to $\hat{MP}(E_r)$ tends to achieve the criterion of Problem $MP(E_r)$, viz, to minimize the maximum of several cut expressions. This is further substantiated by the fact that the quantities minimum $\{a_p^t x - \alpha_p\text{: } x \in X_A\}$ are almost always negative, with the result that (3.2) gives more weight to cut expressions which tend to have algebraically larger values. Thereby, cuts which are likely to be binding in an optimal solution to $MP(E_r)$ are given more weightage.

As an additional improvement routine for this scheme, we adopted the following strategy which enabled us to recover good quality solutions. Consider the linear bounding form $g^t x$ of Cabot and Francis [3] satisfying $g^t x \leqslant x^t Sx$ for each $x \in X_A$, with coefficients given by

$$g_{ij} = \text{minimum} \left\{ \sum_{k=1}^{m} \sum_{l=1}^{m} s_{ijkl} x_{kl}\text{: } x \in X_A, \ x_{ij} = 1 \right\}$$

Note that $g_{ij}$ is obtained by solving a linear assignment problem in $m-1$ variables after fixing the assignment $x_{ij} = 1$. Now, let

$$\nu_{\min} = \text{minimum}\{g^t x\text{: } x \in X_A\} \text{ and } \nu_{\max} = \text{maximum}\{g^t x\text{: } x \in X_A\}.$$

Working with some test problems in the literature, we found that optimal or good quality solutions often had values of $g^t x$ lying in the initial 10-15% of the range $[\nu_{\min}, \nu_{\max}]$. Hence, we found it advantageous to replace $\hat{MP}(E_r)$ by the problem

$$\text{minimize}\left\{\Lambda_r' x + \left(\frac{w}{\nu_{\min}}\right) g' x : x \in X_A\right\}, \tag{3.3}$$

where $w$ is a suitable weightage parameter. In our experience, the solution $\hat{x}^{r+1}$ to problem (3.3) almost always yielded a corresponding value of $\hat{z}^{r+1}$ satisfying $\hat{z}^{r+1} < \bar{z}$, except towards the end of the procedure when an exact solution of the master problem was frequently required. Moreover, the solution $\hat{x}^{r+1}$ tended to be of a good quality.

Finally, we note that each time an approximate solution was obtained through (3.3), we attempted to improve on this solution through pairwise interchanges, and elected to use the resulting solution in case it also qualified as an approximate solution. Thus, only when neither the solution to (3.3), nor the solution obtained through pairwise interchange improvements on it, yielded a value of $z$ less than the incumbent value, did we need to resort to an exact solution technique for the master problem.

### 3.2.2 Exact Solution Method for the Master Problem

We now consider the exact solution of Problem $MP(E_r)$. For this purpose, note that if we consider the linear relaxation $X$ of the set $X_A$, then integral solutions feasible to $X$ are in a one-to-one correspondence with the points in $X_A$. In other words, all integral solutions in $X$ are zero-one assignment solutions. Thus, we first attempted the use of Gomory's [13] dual all-integer cuts to obtain an optimal solution to $MP(E_r)$. However, in spite of incorporating all the rules for finite convergence (see [13] and [7]), we experienced the same problems as Trauth and Woolsey [29] and as formally predicted by Finkelstein [5]. Namely the procedure works very well in some instances, but is quite unpredictable in most instances. In some cases, several thousand dual-integer cuts were unable to find an integer primal feasible solution. Further, in this process, the updated coefficients in the simplex tableau tended to blow-up in magnitude.

We thus resorted to Glover's [12] pseudo-primal-dual procedure which iterates between a lexicographic dual feasible stage related to Gomory's dual all-integer algorithm [13] and a dual infeasible stage related to Young's primal all-integer algorithm [30]. As reported by Glover, this method results in a lexicographic decrease in the updated solution column of the simplex tableau on two successive visits to the dual feasible stage by an amount at least as much as the decrease which would be obtained through two successive iterations of the dual simplex method. Of course, if we have dual feasibility and the condition $\bar{z} = z'$ of Step 2 holds, then we may terminate without regard to primal feasibility. The revised-simplex method was specialized and rules for finite convergence (see [12, 13 and 7]) were also incorporated. Although this technique contained the magnitudes of the tableau coefficients, it also experienced convergence problems.

Finally, we adopted an implicit enumeration scheme which was initialized just once at the first visit of the procedure to the exact solution routine for the master problem, and was simply updated at each subsequent visit. The updating feature permitted us to delete Benders' cuts for larger sized problems since it ensured that at least fathomed solutions would not be regenerated during the exact solution of the master problem. This scheme either resulted in a solution with a value of $z$ less than the incumbent value, in which case this solution was adopted as our approximate solution, or else it verified optimality by solving the problem to termination. We avoid giving details here of this procedure in view of our experience reported in the following section. The interested reader may refer to Sherali [27], however.

## 4. COMPUTATIONAL EXPERIENCE

In this section, we report computational experience using test problems available in the literature [4, 23, 29]. We found that the procedure performed satisfactorily in detecting optimal solutions early on in the search process. Even in cases where storage limitations forced a premature termination, solutions of quality often better than any other known in the literature were obtained during the early stages of the search. In fact, recently, Burkard and Stratmann [2] have described excellent heuristics that significantly improve upon all the best known solution values of problems in the literature We were able to still improve further upon the values for Nugent et al's [23] problems of sizes $m = 20$ and $m = 30$ and perform at least as good on the other problems attempted by Burkard and Stratmann. We also obtained a significant improvement over the best known solution to Elshafei's [4] hospital layout problem.

We note that unlike the experience of Geoffrion and Graves [10] on using Benders' partitioning on a multicommodity distribution flow problem, our problem required a large number of Benders' cuts. The reason for this, as conjectured later in Section 5, may be that a solution to the linear relaxation of Problem QAP 4 is not "close" to an integral solution. In fact, when working with Nugent, Vollmann and Ruml's [23] test problems, we found that the procedure required to generate close to $m!$ cuts in order to verify optimality, even when the starting solution was optimal. Table 1 below reports this experience. We remark that the objective function value of the master problem increased very slowly except for the late stages of the procedure when a rapid increase was obtained. As a result, when the problems of sizes $m \geqslant 8$ were prematurely terminated, the master problem value was as yet negative.

TABLE 1. *Exact Implementation of Benders' Scheme*

| Problem | $m$ | $m!$ | # of cuts generated | Optimality verified? | cpu seconds execution time(1) |
|---|---|---|---|---|---|
| | 5 | 120 | 113 | YES | 6.9 |
| Nugent, Vollmann | 6 | 720 | 690 | YES | 154.4 |
| and Ruml's | 7 | 5040 | 4711(2) | NO | 770 |
| Problems [23] | 8 | 40320 | 3722(2) | NO | 770 |

(1) On a CDC Cyber 70, Model 74-28/CDC 6400 machine, with coding in Fortran IV
(2) Forced termination as time limit is reached

In view of these results, we chose to operate the procedure as a heuristic by terminating it prematurely. It may be noted that the performance of the procedure as a heuristic with regard to the quality of the best solution obtained is sensitive to both the starting solution employed, as well as the value of $w$ in Equation (3.3). In most cases, we found it appropriate to start with a value of $w = 1$ and increment it by one every three to ten cuts. For constructing a starting solution, we adopted the following strategy. For each facility, we computed the sum of flows to all the other facilities and for each location, we computed the sum of distances to all the other locations. We then arranged the facilities in nonincreasing order of their flow sums and the locations in nondecreasing order of their distance sums. The desired solution was constructed by matching the arrangement of facilities and locations element by element, and then improving upon the resulting solution through pairwise exchanges.

Table 2 reports our computational experience with some problems available in the literature [4, 23, 28]. Column 'a' gives the best values obtained on using the above starting solution and terminating the process after 25 cuts. Column 'b' gives the cut index at which this best solution was found. Through a few subsequent runs using better quality starting solutions as obtained over previous runs, we were able to improve on the initial run solutions. Column 'c' gives the best objective values we were able to obtain in this manner. The quality of these

TABLE 2. *Implementation of Benders' Scheme as a Heuristic*

| Problem | | m | a | b | c | d | Best Locations of Facilities 1, 2, ..., m Respectively for Solutions of Values in Column c | e |
|---|---|---|---|---|---|---|---|---|
| Nugent, Vollmann and Ruml's Problems [23] | | 5 | 25 | 7 | 25 | 25 | 3,4,5,1,2 | 0.69 |
| | | 6 | 43 | 1 | 43 | 43 | 3,2,1,6,5,4 | 1.30 |
| | | 7 | 74 | 3 | 74 | 74 | 1,2,5,3,4,7,6 | 2.25 |
| | | 8 | 107 | 4 | 107 | 107 | 3,4,8,2,1,5,6,7 | 3.56 |
| | | 12 | 289 | 14 | 289 | 289 | 5,1,9,8,4,3,11,7,10,2,6,12 | 13.69 |
| | | 15 | 575 | 9 | 575 | 575 | 5,4,9,10,12,13,7,2,1,15,6,11,3,8,14 | 33.93 |
| | | 20 | 1285 | 3 | 1285 | 1287 | 4,12,19,11,2,5,3,9,20,18,14,13,10,17,7,15,1, 16,6,8 | 100.61 |
| | | 30 | 3095 | 10 | 3077 | 3079 | 20,29,11,5,30,28,15,9,23,21,2,27,22,6,1,10, 19,8,16,12,24,14,7,26,13,25,3,18,17,4 | 393.96 |
| Steinberg's Problems [28] | Rectilinear Distance | 36 | 4975 | 10 | 4802 | 4802 | 14,27,35,25,34,33,15,26,7,16,32,23,24,31,6,8, 18,17,12,22,20,29,21,28,19,10,30,13,5,4,3,11, 1,2,9,36 | 829.48 |
| | Squared Euclidean Distance | 36 | 8232 | 16 | 7926 | 7926 | 31,10,2,12,3,4,22,11,29,21,5,14,13,15,30,28, 19,20,24,23,7,8,16,9,17,18,6,32,33,34,35,25, 27,26,36,1 | 921.02 |
| | Euclidean Distance | 36 | — | $1^{(1)}$ | $c_3$ | $c_3$ | Same as for the squared euclidean distance solution, but with facilities 1 and 15 interchanged in location. | 958.6 |
| Elshafei's Problem [4] | | 19 | $c_1$ | 12 | $c_1$ | $c_2$ | 17,18,19,11,12,9,3,14,1,2,10,13,7,5,16,15,8,4,6 | 96.15 |

(1) This resulted through the use of the best recorded squared euclidean solution as a starting solution.

a, b, c, d, e — See the text of Section 4 for connotation.

$c_1$ = 8, 606, 274
$c_2$ = 11, 281, 888
$c_3$ = 4125.168

solutions may be compared with the previously best known solutions reported in the literature [2] as given in column 'd'. Note that one may use such trial runs to select between the two subproblem solutions given in Section 3.1, preferring the one which yields larger values for the master problem, and one may also derive information on the manner in which $w$ may be incremented in Equation (3.3). Finally, column 'e' gives the execution time in cpu seconds for a run which generates 25 cuts on a CDC Cyber 70 Model 74-28/CDC 6400 computer, with coding in FORTRAN IV. These times do not include the effort for generating either the linear bounding form $g^t x$ of Section 3.2.1 or the starting solution.

Before concluding, we note that further improvements of the procedure as a heuristic may be possible through the implementation of more sophisticated exchange schemes [24] in lieu of the simple pairwise exchange operations we have employed at each iteration. Also, one may choose to attempt a few short trial runs with the aim of gaining the type of information alluded to above, and then execute a final run using the best solution found in the trial runs as a starting solution.

## 5. CONCLUDING REMARKS

It is our purpose in this section to discuss, at least for our problem, the relationship between the performance of Benders' partitioning scheme in regard to the number of cuts required and the closeness of the solution of the original problem to that of the linearly relaxed mixed integer program. In fact, we show below that if the linear relaxation of Problem QAP 4 has an integral optimal solution, then Benders' procedure terminates as soon as this solution is detected.

Suppose that an optimal integer solution $(\bar{x}, \bar{y})$ to the linear relaxation of Problem QAP 4 exists so that $\bar{x}$ solves the quadratic assignment problem. The Kuhn-Tucker conditions can be shown to assert the following, where $u_{ij}$, $v_{kl}$, $w_{ijkl}$ are the dual variables associated with constraints (2.2), (2.3), and (2.7) respectively:

$$\bar{y}_{ijkl} = 0 \rightarrow w_{ijkl} = 0 \text{ and } u_{ij} + v_{kl} \leqslant c_{ijkl}$$

$$\bar{y}_{ijkl} = 1 \rightarrow u_{ij} + v_{kl} - w_{ijkl} = c_{ijkl},\ w_{ijkl} \geqslant 0 \tag{5.1}$$

for $i = 1, \ldots, m-1,\ j = 1, \ldots, m,\ k = i+1, \ldots, m,\ l = 1, \ldots, m,\ l \neq j$

Suppose that $\bar{x}$ is generated by Benders' scheme, say as a solution to the Master Problem $MP(E_{r-1})$. Since $\bar{x}$ solves the quadratic assignment problem, and since by Lemma 1, $\bar{y}_{ijkl} = \bar{x}_{ij}\bar{x}_{kl}$, then the current incumbent value $\bar{z}$ is given by:

$$\bar{z} = \sum_{i=1}^{m-1} \sum_{j=1}^{m} \sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} c_{ijkl}\, \bar{y}_{ijkl}. \tag{5.2}$$

Note, however, that (5.1) implies that $(u, v, w)$ is indeed an optimal solution to problem $SP(\bar{x})$. By examining the dual of problem $SP(\bar{x})$, it immediately follows that the right hand side of (5.2) is equal to $\sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\, x_{ij} - \alpha$, where,

$$a_{ij} = (m - i)\, u_{ij} + (i - 1)\, v_{ij}, \text{ and}$$

$$\alpha = \sum_{i=1}^{m-1} \sum_{j=1}^{m} \sum_{k=i+1}^{m} \sum_{\substack{l=1 \\ l \neq j}}^{m} w_{ijkl}$$

yielding

$$\bar{z} = \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\, \bar{x}_{ij} - \alpha \tag{5.3}$$

The next Benders' cut generated is of the form $z \geqslant \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\, x_{ij} - \alpha$, and hence an optimal objective value $z^{r+1}$ to problem $MP(E_r)$ must satisfy the following inequality:

$$z^{r+1} \geqslant \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\, \bar{x}_{ij} - \alpha \tag{5.4}$$

In view of (5.3) and (5.4), we have $z^{r+1} \geqslant \bar{z}$, which is the termination criterion for Benders' scheme.

To summarize, if a solution $(\bar{x}, \bar{y})$ to the linear relaxation of problem QAP 4 is integer, and if $\bar{x}$ is generated at some iteration $r - 1$, then Benders' scheme will terminate at iteration $r$ with the conclusion that $\bar{x}$ is optimal. In particular, if $\bar{x}$ is used as a starting solution, then only a single cut is sufficient to verify optimality. We emphasize, however, according to our computational testing, that if the solution to the relaxed problem is not integer, even if we start with

the optimal solution to the quadratic assignment problem, many Benders' cuts may be needed to verify optimality. We hence conjecture that a fewer number of Benders' cuts may be required if the solution of the linear relaxation of QAP 4 is close to an integral solution. For additional insight into Benders' partitioning scheme and in particular to the strength of bounds derived therefrom, we refer the reader to Rardin and Unger [26].

## REFERENCES

[1] Benders, J.F., "Partitioning Procedures for Solving Mixed Variables Programming Problems," Numerische Mathematik, Vol. *4*, pp. 238-252 (1962).

[2] Burkard, R.E. and K.H. Stratmann, "Numerical Investigations on Quadratic Assignment Problem," Naval Research Logistics Quarterly, Vol. *25*, No. 1, pp. 129-148 (1978).

[3] Cabot, A.V. and R.L. Francis, "Solving Certain Non-Convex Quadratic Minimization Problems by Ranking the Extreme Points," Operations Research, Vol. *18*, pp. 82-86 (1970).

[4] Elshafei, A.N., "Hospital Layout as a Quadratic Assignment Problem," Operational Research Quarterly, Vol. *28*, No. 1, pp. 167-179 (1977).

[5] Finkelstein, J.J., "Estimation of the Number of Iterations for Gomory's All-Integer Algorithm," Doklady Akademii Nauk. SSSR, Vol. *193*, pp. 988-992 (1970).

[6] Francis, R.L. and J.A. White, "*Facility Layout and Location: An Analytical Approach*," (Prentice-Hall, 1974).

[7] Garfinkel, R.S. and G.L. Nemhauser, "*Integer Programming*," (John Wiley and Sons, 1972).

[8] Gaschütz, G.K. and J.J. Ahrens, "Suboptimal Algorithm for the Quadratic Assignment Problem," Naval Research Logistics Quarterly, Vol. *15*, pp. 49-62 (1968).

[9] Gavett, J.W. and N.V. Plyter, "The Optimal Assignment of Facilities to Locations by Branch and Bound," Operations Research, Vol. *14*, pp. 210-232 (1966).

[10] Geoffrion, A.M. and G.W. Graves, "Multicommodity Distribution System Design by Benders' Decomposition," Management Science, Vol. *20*, No. 5, pp. 822-844 (1974).

[11] Gilmore, P.C., "Optimal and Suboptimal Algorithms for the Quadratic Assignment Problem," SIAM Journal on Applied Mathematics, Vol. *10*, pp. 305-313 (1962).

[12] Glover, F., "A Pseudo Primal Dual Integer Programming Algorithm," *Journal of Research of the National Bureau of Standards*, 71B, pp. 187-195 (1967).

[13] Gomory, R.E., "All-Integer Programming Algorithm," *Industrial Scheduling*, J.F. Muth and G.L. Thompson, eds. (Prentice Hall, 1963).

[14] Graves, G.W. and A.B. Whinston, "An Algorithm for the Quadratic Assignment Problem," Management Science, Vol. *17*, pp. 453-471 (1970).

[15] Hanan, M. and J. Kurtzberg, "A Review of the Placement and Quadratic Assignment Problems," *SIAM Review*, Vol. *14*, pp. 324-342 (1972).

[16] Heider, C.H., "An n-Step, 2-Variable Search Algorithm for the Component Placement Problem," Naval Research Logistics Quarterly, Vol. *20*, pp. 669-724 (1973).

[17] Hillier, F.S. and M.M. Connors, "Quadratic Assignment Problem Algorithms and the Location of Indivisible Facilities," Management Science, Vol. *13*, pp. 42-57 (1966).

[18] Kaufman, L. and F. Broeckx, "An Algorithm for the Quadratic Assignment Problem Using Benders' Decomposition," European Journal of Operational Research, Vol. *2*, pp. 204-211 (1978).

[19] Koopmans, T.C. and M. Beckmann, "Assignment Problems and the Location of Economic Activities," Econometrica, Vol. *25*, pp. 53-76 (1957).

[20] Land, A.M., "A Problem of Assignment with Interrelated Costs," Operational Research Quarterly, Vol. *14*, pp. 185-198 (1963).

[21] Lasdon, L.S., *Optimization Theory for Large Systems*, (The MacMillan Company, New York, 1970).

[22] Lawler, E.L., "The Quadratic Assignment Problem," Management Science, Vol. *19*, pp. 586-590 (1963).

[23] Nugent, C.E., T.E. Vollmann and J. Ruml, "An Experimental Comparison of Techniques for the Assignment of Facilities to Locations," Operations Research, Vol. *16*, pp. 150-173 (1968).

[24] Obata, T. and P. Mirchandani, "Algorithms for a Class of Quadratic Assignment Problems," presented at the Joint National TIMS/ORSA Meeting, New Orleans (May 1979).

[25] Pierce, J.F. and W.B. Crowston, "Tree Search Algorithms for Quadratic Assignment Problems," Naval Research Logistics Quarterly, Vol. *18*, pp. 1-36 (1971).

[26] Rardin, R. and V.E. Unger, "Surrogate Constraints and the Strength of Bounds Derived from Zero-One Benders' Partitioning Processes," Operations Research, Vol. *24*, No. 6, pp. 1169-1175 (1976).

[27] Sherali, H.D., "The Quadratic Assignment Problem: Exact and Heuristic Methods," Unpublished Doctoral Dissertation, Georgia Institute of Technology, June (1979).

[28] Steinberg, L., "The Backboard Wiring Problem: A Placement Algorithm," SIAM Review, Vol. *3*, pp. 37-50 (1961).

[29] Trauth, C.A. and R.E. Woolsey, "MESA: A Heuristic Integer Linear Programming Technique," Res. Rep. SC-RR-68-299, Sandia Labs, Albuquerque, N.M. (1968).

[30] Young, R.D., "A Primal (All-Integer) Integer Programming Algorithm," Journal of Research of the National Bureau of Standards, 69B, (Math and Math Phys.) Vol. *213*, No. 3 (1965).

# TEST SELECTION FOR A MASS SCREEENING PROGRAM*

John A. Voelker

*Argonne National Laboratory*
*Argonne, Illinois*

William P. Pierskalla

*University of Pennsylvania*
*Philadelphia, Pennsylvania*

**ABSTRACT**

Periodic mass screening is the scheduled application of a test to all members of a population to provide early detection of a randomly occurring defect or disease. This paper considers periodic mass screening with particular reference to the imperfect capacity of the test to detect an existing defect and the associated problem of selecting the kind of test to use. Alternative kinds of tests differ with respect to their reliability characteristics and their cost per application.

Two kinds of imperfect test reliability are considered. In the first case, the probability that the test will detect an existing defect is constant over all values of elapsed time since the incidence of the defect. In the second case, the test will detect the defect if, and only if, the lapsed time since incidence exceeds a critical threshold $T$ which characterizes the test.

The cost of delayed detection is an arbitrary increasing function (the "disutility function") of the duration of the delay. Expressions for the long-run expected disutility per unit time are derived for the above two cases along with results concerning the best choice of type of test (where the decision rules make reference to characteristics of the disutility function).

## INTRODUCTION

Mass screening is the process of inspecting all members of a large population for defects. If the early detection of a defect provides benefits, it may be advantageous to employ a test capable of revealing the defect's existence in its earlier stages. (Throughout this paper, the words "defect" and "unit" or "individual" will refer to defect, disorder, or disease and to a member of the population, respectively).

Defects may arrive in a seemingly random fashion such as many types of machine failure, the incidence of certain types of cancer, diabetes, glaucoma, heart disease, etc.; or they may arrive as the result of some contagion such as smallpox, polio, etc. It is the former type of arrival process, random and independent arrivals, which is studied in this paper.

*This research was supported by the Office of Naval Research under Contracts N00014-67-0356-0030 and N00014-75-C-0797.

Continuous monitoring would provide the most immediate detection of a defect, but considerations of expense and practicality will frequently rule out continuous monitoring so that a schedule of periodic testing—a periodic screening program—may be the most practical means of achieving early detection of the defect. In general terms, the question then becomes one of how best to trade off the expense of testing which increases both with the frequency of test applications and with the cost of the type of test used against the benefits to be achieved from detecting the defect in an earlier state of development.

The benefits of early detection depend upon the application considered. For example, in a human population being screened for some chronic disease, the benefits of early detection might include an improved probability of ultimate cure, diminished time period of disability, discomfort, and loss of earnings; and reduced treatment costs. If the population being screened consists of machines engaged in some kind of production, the benefits of early detection might include a less costly ultimate repair and a reduction in the time period during which a faulty product is being unknowingly produced. If the population being screened consists of machines held in readiness to meet some emergency situation, an early detection of a defect would reduce the time the machine was not serving its protective function.

The expense of testing includes easily quantifiable economic costs such as those of the labor and materials needed to administer the testing. However, there can also be other important cost components which are more difficult to quantify. For example, in the case of a human population subject to medical screening, the cost of testing includes the inconvenience and possible discomfort necessitated by the test, the cost of false positives which entails both emotional distress and the need to do unnecessary follow-up testing, and even the risk of physical harm to the testee; e.g., from the cumulative effect of x-ray exposure.

The design of a mass screening program must address two important questions: How frequently to test and what kind of test to use. Optimal testing frequency has been investigated as a function of the defect incidence rate and other factors by Derman [3], Roeloffs [8,9], Barlow, et al. [1], Keller [4], Kirch and Klein [5], and Lincoln and Weiss [6].

The second question follows from the fact that more than one kind of test may be available for use in a mass screening situation. The alternate tests can be entirely different procedures; or they can be the same procedure with different criteria for what constitutes a positive outcome, e.g., alternate levels at which a recording of systolic blood pressure would induce follow-up testing.

Alternative tests will generally differ both in their reliability characteristics and in their cost of application. How to select which test to use is a question which, to our knowledge, has not previously been examined in the context of a general model. This paper will examine this question. For two different ways of modeling test reliability, we develop a framework for test selection and present certain specific decision rules.

Test reliability is assumed to be a function, $p(t)$, of elapsed time, $t$, since defect incidence. Define $\delta_S(t) = \begin{cases} 1 \text{ if } t \in S \\ 0 \text{ otherwise} \end{cases}$. Usually $S$ will be an interval, e.g., $[T, \infty)$. The two cases $p(t) = p$ and $p(t) = \delta_{[T,\infty)}(t)$ are considered. $p(t) = p$ indicates that test reliability is independent of defect age. For $p(t) = \delta_{[T,\infty)}(t)$, the test will detect a defect if, and only if, the defect has existed for at least $T$ units of time. In a sense, these two classes of $p(t)$ represent polar extremes in the responsiveness of test reliability to defect maturity. The test choice decision is posed within each of these two reliability classes.

A crucial feature in any optimization model of mass screening is the characterization of the cost due to detection delay. Detection delay is the gap between the time of detection and the time of defect incidence (or the time the defect becomes potentially detectable by a screening test). The mapping between detection delay and the resultant cost we call the disutility function, $D(\cdot)$. Obviously, the shape of the disutility depends upon the particular application considered. Section 1 gives examples.

Pierskalla and Voelker [7] demonstrated that the shape of $D(\cdot)$ impacts the optimal allocation of a screening budget among segments of the client population characterized by differing defect incidence rates. Results in Sections 3 and 4 below demonstrate the role $D(\cdot)$ plays in test choice decision rules.

Previous research in the area of optimal mass screening which utilized closed form expressions for expected cost (or disutility) placed restrictive assumptions upon the shape of $D(\cdot)$. Early work assumed $D(t) = ct$ (Barlow, et al. [1]). In Kirch and Klein [5], $D(t) = \min(t,T)$ where $t$ is detection delay and $T$ is the (possibly random) delay between defect incidence and the time when the defect would be discovered in the absence of a screening program. Keller [4] restricts the generality of $D(\cdot)$ by requiring that $D(\cdot)$ and the test frequency (density) $r(t)$ be such that $r(t)\int_0^{1/r(t)} D(s)\,ds$ be well approximated by $D(1/2r(t))$. (A density is employed to represent the schedule of test times so that the calculus of variations could serve as the optimization tool.)

Lincoln and Weiss [6] derive two kinds of optimal testing schedules. Both schedules maximize the time between successive tests subject to, respectively, a bound on the mean detection delay and a bound on the probability that detection delay will exceed a fixed threshold. Neither version is equivalent to using a general disutility function.

A few authors, Schwartz and Galliher [10], Thompson and Disney [11], and Voelker [12] let both the reliability of the test and the disutility (or utility) of detection be a function of the defect's state rather than of time since the defect's incidence. Although such models are more general and do utilize a general concept of disutility, they have not been amenable to closed form evaluation of expected disutility.

To incorporate random defect arrivals into their models, previous researchers focus upon an individual who will incur the defect. They use the density function for the age when that individual incurs the defect as a fundamental element of their model. Since the density function reflects age-specific incidence rates, a "life time" testing schedule can, thereby, be developed to tailor testing frequency at each age to the probability that the defect will occur at that age.

Our way of modeling the randomness of defect arrivals reflects a somewhat different perspective on the mass screening problem. We look through the eyes of a decision maker charged with intelligently allocating a fixed budget. The time frame over which the allocation must be made is often short compared to a typical life time of a member of the client population. Therefore, the decision maker does not plan lifetime screening schedules for particular individuals. Instead, he tries to maximize the benefit that can be derived from his available budget over a much shorter planning horizon. (For modeling purposes the objective of minimizing expected long run cost per unit time is not unreasonable with the problem viewed in this way since steady state conditions should approximately obtain after the initial screening. This is especially the case when an existing and ongoing screening program is being optimized by the decision maker. Also, lacking information to the contrary, the decision maker has no reason to anticipate abrupt changes in the screening policy at the end of the planning horizon.)

With the problem viewed in this perspective, the random nature of defect arrivals is most naturally modeled as a Poisson process with its parameter determined by the incidence rate of the defect and the size of the population. This approach has proved particularly useful in the following context: If different segments of the client population exhibit different incidence rates, subpopulations can be defined with defect incidence within each modeled as a Poisson process with its respective parameter. Then the budget can be so allocated among the subpopulation to permit appropriate relative testing frequencies (cf. Pierskalla and Voelker [7]). In this way, age-specific incidence rates can be incorporated into the notion of Poisson defect arrivals. Moreover, factors other than age which affect defect incidence rates (family history, smoking habits, work environment, etc.) can also be incorporated into the model.

Although this paper does not follow Pierskalla and Voelker [7] in studying the case of heterogeneous populations, we use the same Poisson model of random defect arrivals. Section 1 presents examples of disutility functions. Section 2 sketches the basic structure of the model and represents the expected long run disutility per unit time for general reliability and disutility functions. Section 3 considers the class of tests with $p(t) = p$. Section 4 does so for $p(t) = \delta_{[T,\infty)}(t)$. In both Sections 3 and 4, results regarding and test choice criteria are presented. Proofs are deferred to the APPENDIX.

## SECTION 1: SOME EXAMPLES OF DISUTILITY FUNCTIONS

Suppose a production process is subject to a randomly occurring defect. Although production appears to proceed normally after the incidence of the defect, the product produced is, thereafter, defective to an extent which remains constant until the production process is returned to its proper mode of operation. The only way to learn if the production process is in this degraded state is to perform a costly test. Now, if a test detects the existence of the degraded mode of production $t$ units of time after its incidence, the harm done will be proportional to the amount of defective product (unknowingly) produced which, in turn, is proportional to $t$. Hence, $D(t) = at$ for some $a > 0$.

Another example where a linear $D(\cdot)$ function may be appropriate would be for the periodic inspection of an inactive device (such as a missile) stored for possible use in an emergency. If $t$ is the time between the incidence of the disorder and its detection, the disutility incurred is proportional to the probability that the device would be needed in that time interval. If such "emergencies" arise according to a Poisson process with rate $\mu$, then the probability of an emergency in a time interval of length $t$ is $1 - e^{-\mu t}$, which, for $\mu$ small, is approximately $\mu t$. Hence, if $b$ is the cost incurred should there be an emergency while the device is defective, and if $\mu$ is the (small) arrival rate of emergencies, then $D(t) = b\mu t$.

A quadratic disutility could arise in the following situation. Suppose the magnitude of a randomly occurring defect increases linearly with time since the occurrence of the defect. For example, the magnitude of the defect might be the size of a small leak in a storage container for a fluid, and as fluid escapes, the leak gets larger. Further suppose that the harm done accumulates at a rate proportional to the magnitude of the defect. Hence, the quantity of fluid lost (at least initially) increases the longer the defect exists, and the rate of fluid loss is proportional to the size of the leak.

Let the size of the leak (as measured by rate of fluid loss), at time $s$ since the leak's incidence, be $cs$. Then, if the defect is detected at time $t$ since incidence, the disutility incurred (fluid lost) is $D(t) = \int_0^t cs\,ds = 1/2(ct)^2$.

## SECTION 2: EXPECTED DISUTILITY WITH A GENERAL RELIABILITY FUNCTION

We assume that the times of defect arrivals in the screened population form a stationary Poisson process. Since there is a certain intuitive appeal to considering the defect arrival rate proportional to both the size of the population, $N$, and the intrinsic incidence rate, $\lambda$; let $N\lambda$ designate the parameter of the above arrival process. It is not necessary to know the value of $N\lambda$ in order to apply the decision rules of test selection developed in this paper.

Let $p_l(t)$ be the probability that a test of type $l$ will detect a defect which has been present $t$ units of time. $p_l(t) = 0$ for $t < 0$. Let $\bar{S}^k_{r,l}$ be a random variable denoting the time at which the $k^{\text{th}}$ defect is detected. $\bar{S}^k_{r,l}$ depends upon the arrival time of the defect, $\delta_k$; the type of test used, $l$; and the testing frequency, $r$.

Given the application of test type $l$ at the times $\{1/r, 2/r, \ldots\}$, the disutility incurred by the $k^{\text{th}}$ defect is $D(\bar{S}^k_{r,l} - S^k)$. The total disutility incurred due to defects which occur before time $A$ is

$$\sum_{k=1}^{\infty} D(\bar{S}^k_{r,l} - S^k)\ \delta_{[O,A)}(S^k).$$

In Pierskalla and Voelker [7], the long run expected disutility per unit time under the above screening program $r,p(\ ))$ was shown to be

$$rN\lambda \sum_{n=1}^{\infty} \int_{n-1/r}^{n/r} D(u)\ p_l(u) \prod_{m-1}^{n-1} [1 - p_l(u - m/r)]\ du. \tag{2.1}$$

This result will serve as our starting piont for the technical results of this paper. Lincoln and Weiss established essentially the same result based upon a different probabilistic model of defect arrival.

## SECTION 3: CONSTANT TEST RELIABILITY

The mass screening model yields interesting results for a test which has a fixed probability $p$ of detecting the disorder if it is present in an individual. Such a model would arise if the unreliability of the test is entirely intrinsic to the test procedure rather than partially dependent upon the state or age of the defect. An example of this is the administration of a Mantoux test for tuberculosis in, say, a population of grade school children. The test has a small but relatively constant level of false negatives. There are other medical tests with similar characteristics.

### A Quality Control Example

To see how another type of situation with constant test reliability could arise, consider a production system which is subject to a randomly occurring defect which degrades the system's performance. Once the defect occurs, the level of degradation of the process remains constant until the defect is discovered. Suppose the defect is such that each item produced has probability $\delta$ of being defective and that the system without the defect never produces defective items.

The only way to discover the existence of the defect in the production system is to examine an item produced which is itself defective. Now, if the examination of an item is expensive (e.g., the item is destroyed as a result of the inspection) and if the capacity to examine a sequence of items involves a set-up cost (say $a$), the following strategy might be called for: At specified times $1/r, 2/r, \ldots$, set up the capacity to examine a sequence of items and examine, say, $l$ items at each of those times. The times $1/r, 2/r, \ldots$, are then the times of testing and sample size $l$ specifies the test type.

Assume that if a defective item is examined, the defect is always observed and the production process is, thereby, discovered to be in the degraded state. Hence, a degraded state of the production process will go undetected at the testing occasion $k/r$ if, and only if, each of the $l$ items sampled at time $k/r$ is, by chance, not defective. But the probability of that event is $(1-\delta)^l$. Note that the elapsed time $t$ between the entry of the production process into the degraded state and the test time $k/r$ does not affect this probability. Hence, $p_l = p_l(t) = 1 - (1-\delta)^l$ which represents the probability that a test (the inspection of $l$ items) will detect a degraded production process. Note that the choice of $l$ affects both the test's reliability and cost.

**Expected Disutility**

Designate the expected long run disutility per unit time per member of a population screened with frequency $r$ using a test of constant reliability $(p(t) = p,\ t > 0)$ by $C(r,p)$. From Eq. (2.1),

$$C(r,p) = N\lambda rp \sum_{n=0}^{\infty} (1-p)^n \int_{n/r}^{(n+1)/r} D(u)\,du.$$

The question examined next is how do changes in $r$ and $p$ affect $C(r,p)$. After that, explicit solutions are given for $C(r,p)$ when $D(\cdot)$ takes certain simple forms. And lastly, some general rules are indicated for selecting between a particular kind of test and a more expensive but more reliable alternative test when $D(\cdot)$ takes certain forms.

PROPOSITION 1: If $D(\cdot)$ is a strictly increasing function, $[\partial C(r,p)]/\partial p < 0$ and $[\partial C(r,p)]/\partial r < 0$.

Note that for $D(\cdot)$ nondecreasing, the above inequalities still hold, but not strictly.

From this proposition, as anticipated, when $p$ increases, the expected disutility decreases. Similarly as $r$ increases, the interval $1/r$ between tests decreases and the expected disutility decreases. Consequently, as better test types are used or the tests are more frequently applied, the value of such changes in terms of reduced disutility versus the costs of the changes can, in principle, be assessed and the tradeoffs evaluated.

It is easy to compute the Hessian for $C(r,p)$ when $D(\cdot)$ is differentiable:

$$\frac{\partial^2 C}{\partial r^2} = N\lambda p \sum_{n=0}^{\infty} q^n r^{-3}\left[(n+1)^2 D'\left(\frac{(n+1)}{r}\right) - n^2 D'\left(\frac{n}{r}\right)\right],$$

$$\frac{\partial^2 C}{\partial r^2} = N\lambda r \sum_{n=2}^{\infty} n(n-1)\, q^{n-2}\,[W(n,r) - W(n-1,r)],$$

$$\frac{\partial^2 C}{\partial p\,\partial r} = N\lambda \sum_{n=0}^{\infty} [1-(n+1)p]\, q^{n-1}\,[W(n,r) - B(n+1,r) + B(n,r)],$$

where

$$q = 1 - p,\ W(n,r) = \int_{n/r}^{(n+1)/r} D(s)\,ds, \text{ and } B(n,r) = \frac{n}{r}\, D(n/r).$$

Note that if $D'(\cdot)$ is increasing, then $(\partial^2 C/\partial r^2) \geqslant 0$ and $(\partial^2 C/\partial p^2) \geqslant 0$. Hence, along coordinate directions $C(r,p)$ is convex.

Simple expressions for $C(r,p)$ can be given when $D(\cdot)$ is specialized to a polynomial or an exponential function. Since these two types of functions are reasonably general, they can be quite useful as realistic approximations in applications.

PROPOSITION 2: If $D(t) = \sum_{i=1}^{I} a_i\, t^{m_i}$, then

$$C(r,p) = N\lambda p \sum_{i=1}^{I} [a_i / r^{m_i}(m_i + 1)] \sum_{n=1}^{\infty} n^{m_i+1} p q^{n-1}. \tag{3.1}$$

If the $m_i$; $i = 1, \ldots, N$ are positive integers, the inner summation of Eq. (3.1) is simply the $(m_i + 1)$ moment of a geometric random variable. Hence, using Laplace transforms, Eq. (3.1) becomes

$$C(r,p) = N\lambda p \sum_{i=1}^{I} [a_i / r^{m_i}(m_i + 1)]\, \psi^{m_i+1}(0{:}p) \tag{3.2}$$

where $\psi(t{:}p) = pe^t/(1 - qe^t)$ and $\psi^{(m)}(0{:}p) = d^m \psi(t{:}p)/dt^m\,|_{t=0}$. For example, when $m = 1$, $C(r,p) = (aN\lambda/2r)[(2 - p)/p]$ and for $m = 2$, $C(r,p) = [(aN\lambda/3r^2)]\,[1 + (6 - 6p)/p^2]$.

PROPOSITION 3: If $D(t) = \beta e^{at}$ for $a, \beta > 0$, then $C(r,p) = \beta N\lambda p r(e^{a/r} - 1)/a(1 - qe^{a/r})$, for $r > -a/\log(q)$.

**Test Selection**

Propositions 2 and 3 can provide a means to select between two alternative kinds of tests which differ with respect to reliability of detection and cost per application. Let test No. 1 have cost per application $c_1$ and reliability $p_1$. The corresponding parameters for test No. 2 are $c_2$ and $p_2$. If test No. 1 is administered with frequency $rc_2/c_1$ and test No. 2 administered with frequency $r$, both testing regimes will consume equal quantities of the budgeted resource; *vis.*, $Nrc_2$ per unit time. If $C(rc_2/c_1,p_1) \leqslant C(r,p_2)$ for all $r \geqslant 0$, then the expected disutility per unit time will be less with test No. 1 at all levels of budget $Nrc_2$. That is, if test No. 2 is being used with frequency $r$, the expected disutility can be decreased without any additional allotment of budget, simply by switching to test No. 1 and testing as frequently as the budget permits.

Suppose, for example, that $D(t) = at^m$ for $m$ a positive integer. Then $C(rc_2/c_1,p_1) \leqslant C(r,p_2)$ is equivalent by Eq. (3.2) to

$$(c_1/c_2)^m \leqslant \frac{p_2\, \psi^{(m+1)}(0{:}p_2)}{p_1\, \psi^{(m+1)}(0{:}p_1)}. \tag{3.3}$$

Therefore, test No. 1 is preferred over test No. 2 if, and only if, Eq. (3.3) obtains.

## SECTION 4: THRESHOLD TEST RELIABILITY

In the previous section, the reliability of the test depended only on factors intrinsic to the test itself and did not depend at all on the elapsed time since incidence at the time of the test. In this section, a special form for $p(t)$ is considered which is very different from the case of constant test reliability. Here the test reliability is zero if the elapsed time since the defect's incidence is less than $T$; otherwise, the reliability is one. That is, $p(t) = \delta_{[T,\infty)}(t)$ where the number $T$ is a characteristic of the type of test chosen.

For a screening program in which a test with the above reliability characteristics is applied with frequency $r$, let $A(r,T)$ represent the expected long run disutility per unit time.

Of course, the "blind period" of the test for 0 to $T$ does not, in a mass screening situation, delay detection of each arriving defect exactly $T$ units of time. The amount of delay depends on the interplay among the time of arrival of the defect, the testing schedules $\{1/r, 2/r, \ldots\}$, and the magnitude of $T$.

The primary results in this section are a simple characterization of $A(r,T)$ and rules which, in some cases, will permit selection betwseen two tests which differ in their reliability (i.e., in their detection threshold $T$) and in their cost per application.

### Long Run Disutility

PROPOSITION 4: If $p(\cdot) = \delta_{[T,\infty)}(\cdot)$ for some $T \geqslant 0$, then

$$A(r,T) = N\lambda r \int_0^r D(T+u)\, du.$$

Suppose, for example, that $D(t) = \exp(at)$ for $a > 0$. Then, $A(r,T) = N\lambda r \exp(aT)[\exp(a/r) - 1]/a$.

### Test Selection

Suppose the decision maker has two kinds of tests available and he must choose one of them for implementation in a mass screening program. Suppose the first kind of test—call it test No. 1— has sensitivity characterized by the "time-until-detectability" threshold $T_1$. Let $c_1 > 0$ be the cost per application (to an individual unit) of this kind of test. For the second kind of test under consideration, test No. 2, let $T_2$ and $c_2$ be the corresponding parameters.

Assume test No. 1 is better in the sense that $T_1 < T_2$. To avoid triviality, assume $c_1 > c_2$.

If the exact shape of the function $D(\cdot)$ is known, Proposition 4 can be used to decide which test to use for each possible level of budget. Let $b$ be the budget per unit time per individual in the population. Then the use of test No. 1 will permit a testing frequency of $b/c_1$ and the use of test No. 2 permits frequency $b/c_2$. To decide which test to use, compare the expected disutilities per unit time assuming a fully allocated budget, i.e., compare $A(b/c_1, T_1)$ and $A(b/c_2, T_2)$. With $D(\cdot)$ known, these quantities can be evaluated explicity by Proposition 4 and compared.

It is clear that the entire budget should be allocated because, when $D(\cdot)$ is an increasing function,

$$\frac{\partial}{\partial r} A(r,T) = N\lambda\left[\int_0^{1/r} D(T+u) - (1/r)\, D(T+\frac{1}{r})\, du\right] < 0.$$

When the exact form of the disutility function is not known, Proposition 4 does not suffice to select between tests No. 1 and 2. However, the two following theorems will permit such a determination at least for certain relative configurations of budget, relative test sensitivity $T_2 - T_1$, and cost differential $c_2 - c_1$ of the tests.

Specifically, Proposition 5 will show that for any (increasing) disutility function, test No. 1 is indicated if the budget (per unit population) exceeds $(c_1 - c_2)/(T_2 - T_1)$. On the other hand, Proposition 6 shows that for a convex increasing disutility function, test No. 2 is better if the budget is less than $(c_1 - c_2)/[2(T_2 - T_1)]$.

A decision rule for the case where the budget falls between $(c_1 - c_2)/(T_2 - T_1)$ and $(c_1 - c_2)/[2(T_2 - T_1)]$ has not been found for general disutility functions.

Just how the statement of Propositions 5 and 6 are translated into the above decision rules is explained after the statements of the respective theorems.

PROPOSITION 5: Given $D(\cdot)$ a strictly increasing function, $T_1 < T_2$ and $c_1 > c_2$, then $T_2 - T_1 > (c_1 - c_2)/rc_2$ implies

$$A(rc_2/c_1, T_1) < A(r, T_2), \tag{4.1}$$

making test No. 1 preferable at the per unit population budget level of $rc_2$.

To apply test No. 1 with frequency $rc_2/c_1$ versus test No. 2 applied with frequency $r$ (actions reflected, respectively, in the left- and right-hand sides of Eq. (4.1)) would require the same budget, $b = rc_2$, per unit population. The hypothesis of theorem 2 implies $b = rc_2 > (c_1 - c_2)/(T_2 - T_1)$. With the hypothesis in this form, the theorem provides a lower bound on the budget which is a sufficient condition for test No. 1 to entail lower expected disutility per unit time vis-a-vis test No. 2, were the two tests scheduled at their maximal (subject to budget) frequencies $rc_2/c_1$ and $r$, respectively.

The following lemma is needed for the proof of Proposition 6 and is recorded here for general interest.

LEMMA 1: If $f$ is a convex function, the $(1/y)\int_t^{t+y} f(s)\,ds \leqslant (1/2)[f(t) + f(t+y)]$.

PROPOSITION 6: If $D(\cdot)$ is convex and increasing, $T_1 < T_2$ and $c_1 > c_2$, then $T_2 - T_1 \leqslant (c_1 - c_2)/2rc_2$ implies $A(rc_2/c_1, T_1) \geqslant A(r, T_2)$, making test No. 2 preferable at the per unit population budget level of $b = rc_2$.

The hypothesis of this proposition implies $b = rc_2 \leqslant (c_1 - c_2)/2(T_2 - T_1)$. Hence, Proposition 6 indicates the superiority of test No. 2 when the budget per unit population is less than $(c_1 - c_2)/2(T_2 - T_1)$.

## REFERENCES

[1] Barlow, R., L. Hunter, and F. Proschan, "Optimum Checking Procedures," Journal SIAM, Vol. *11*, No. 4, pp. 1078-95 (1963).

[2] Butler, David A., "A Hazardous Inspection Model," Technical Report No. 187, Department of Operations Research and Department of Statistics, Stanford University (1977).
[3] Derman, C., "On Minimax Surveillance Schedules," Naval Research Logistics Quarterly, Vol. *8*, pp. 415-9 (1961).
[4] Keller, J., "Optimum Checking Schedules for Systems Subject to Random Failure," Management Science, Vol. *21*, pp. 256-60 (1974).
[5] Kirch, R. and M. Klein, "Surveillance Schedules for Medical Examinations," Management Science, Vol. *20*, pp. 1403-9 (1974).
[6] Lincoln, T. and G. H. Weiss, "A Statistical Evaluation of Recurrent Medical Examinations," Operations Research, Vol. *12*, pp. 187,205 (1964).
[7] Pierskalla, W.P. and J.A. Voelker, "A Model for Optimal Mass Screening and the Case of Perfect Test Reliability," Technical Report #3, Department of Industrial Engineering and Management Science, Northwestern University (1976).
[8] Roeloffs, R., "Minimax Surveillance Schedules with Partial Information," Naval Research Logistics Quarterly, Vol. *10*, pp. 307-22 (1963).
[9] Roeloffs, R., "Minimax Surveillance Schedules for Replacement Units," Naval Research Logistics Quarterly, Vol. *14*, pp. 461-71 (1967).
[10] Shwartz, M., and H. Galliher, "Analysis of Serial Screening in an Asymptotic Individual to Detect Breast Cancer," Tech. Report, Dept. of Ind. and Operations Engr., College of Engineering, University of Michigan (1975).
[11] Thompson, D., and R. Disney, "A Mathematical Model of Progressive Diseases and Screening," presented at the November 1976 ORSA TIMS Meeting.
[12] Voelker, J.A., "Contributions to the Theory of Mass Screening," Ph.D. Dissertation, Northwestern University (1976).

## APPENDIX

PROPOSITION 1: With $q = 1 - p$,

$$\frac{\partial}{\partial p} C(n,p) = N\lambda r \sum_{n=0}^{\infty} q^n W(n) - N\lambda rp \sum_{n=0}^{\infty} n\, q^{n-1}\, W(n)$$

where

$$W(n) = \int_{n/r}^{(n+1)/r} D(s)\,ds.$$

Now let $V_0 = W_0$; $V_n = W_n - W_{n-1}$, $n = 1, 2, \ldots$. Note that $D(\cdot)$ non-negative increasing implies $V_j > 0$ for $j = 0, 1, 2, \ldots$. Then $W_n = \sum_{j=0}^{n} V_j$ and

$$\begin{aligned}
\frac{\partial}{\partial p} C(r,p) &= N\lambda r \sum_{n=0}^{\infty} q^n \sum_{j=0}^{n} V_j - N\lambda rp \sum_{n=0}^{\infty} n\, q^{n-1} \sum_{j=0}^{n} V_j \\
&= N\lambda r \sum_{j=0}^{\infty} V_j \Big[\sum_{n=j}^{\infty} q^n - p \sum_{n=j}^{\infty} n\, q^{n-1}\Big] \\
&= -N\lambda r \sum_{j=0}^{\infty} j\, q^{j-1}\, V_j < 0.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial r} C(r,p) &= N\lambda p \sum_{n=0}^{\infty} q^n \int_{n/r}^{(n+1)/r} D(s)\,ds \\
&\quad - N\lambda p \sum_{n=0}^{\infty} q^n \left[\frac{n+1}{r} D\left(\frac{n+1}{r}\right) - \frac{n}{r} D\left(\frac{n}{r}\right)\right]
\end{aligned}$$

$$= N\lambda p \sum_{n=0}^{\infty} q^n \left[ \int_{n/r}^{(n+1)/r} D(s)\,ds - \frac{n}{r}\left[ D\left( \frac{n+1}{r} \right) - D\left( \frac{n}{r} \right) \right] \right.$$

$$\left. - \frac{1}{r} D\left( \frac{n+1}{r} \right) \right] < 0.$$

This inequality follows from $D(\cdot)$ increasing through the relations $D[(n+1)/r] - D(n/r) > 0$ and

$$\int_{n/r}^{(n+1)/r} D(s)\,ds - \frac{1}{r} D\left( \frac{n+1}{r} \right) < \int_{n/r}^{(n+1)/r} D\left( \frac{n+1}{r} \right) ds - \frac{1}{r} D\left( \frac{n+1}{r} \right) = 0.$$

Q.E.D.

PROOF OF PROPOSITION 2: $C(r,p) = N\lambda rp \sum_{n=0}^{\infty} q^n \int_{n/r}^{(n+1)/r} \sum_{i=1}^{I} a_i t^{m_i}\,dt$

$$= N\lambda rp \sum_{i=1}^{I} \sum_{n=0}^{\infty} a_i q^n/(m_i+1) \left[ \left( \frac{n+1}{r} \right)^{m_i+1} - \left( \frac{n}{r} \right)^{m_i+1} \right]$$

$$= N\lambda p \sum_{i=1}^{I} a_i/r^{m_i}(m_i+1) \left[ \sum_{n=1}^{\infty} q^{n-1} n^{m_i+1} - \sum_{n=1}^{\infty} q^n m^{m_i+1} \right]$$

$$= N\lambda p \sum_{i=1}^{I} a_i/r^{m_i}(m_i+1) \sum_{n=1}^{\infty} n^{m_i+1} p\, q^{n-1}$$

Q.E.D.

PROOF OF PROPOSITION 3: $C(r,p) = N\lambda p \sum_{k=0}^{\infty} q^k \frac{r}{a} e^{ak/r} (e^{a/r} - 1)$

$$= \frac{N\lambda pr}{a} (e^{a/r} - 1) \sum_{k=0}^{\infty} (q\, e^{a/r})^k.$$

The geometric series $\sum_{k=0}^{\infty} (q\, e^{a/r})^k$ converges if, and only if, $r > -a/\log q$. Therefore, for $r > -a/\log q$,

$$C(r,p) = \frac{N\lambda pr}{a} (e^{a/r} - 1) \frac{1}{1 - q\, e^{a/r}}.$$

Q.E.D.

PROOF OF PROPOSITION 4: Note that

$$\prod_{m=1}^{n-1} \left[ 1 - \delta_{[T,\infty)} \left( u - \frac{m}{r} \right) \right] = 1 - \delta_{[T,\infty)} \left( u - \frac{1}{r} \right).$$

By Eq. (2.1),

$$A(r,T) = N\lambda r \sum_{n=1}^{\infty} \int_{(n-1)/r}^{n/r} D(u) \delta_{[T,\infty)}(u) \prod_{m=1}^{n-1} \left[ 1 - \delta_{[T,\infty)} \left( u - \frac{m}{r} \right) \right] du$$

$$= N\, r \sum_{n=1}^{\infty} \int_{(n-1)/r}^{n/r} D(u) \delta_{[T,\infty)}(u) \left[ 1 - \delta_{[T,\infty)} \left( u - \frac{1}{r} \right) \right] du.$$

Now $u < T + (1/r) \Leftrightarrow u - (1/r) < T \Leftrightarrow 1 - \delta_{[T,\infty)}[u - (1/r)] = 1$. Therefore,

$$A(r,p) = N\lambda r \int_0^\infty D(u)\delta_{[T,\infty)}(u)\,\delta_{[-\infty,T+1/r)}(u)\,du$$
$$= N\lambda r \int_T^{T+1/r} D(u)\,du$$

Q.E.D.

PROOF OF PROPOSITION 5: By Proposition 4, Eq. (4.1) is equivalent to

$$(rc_2/c_1)\int_0^{c_1/rc_2} D(T_1 + s)\,ds < r\int_0^{1/r} D(T_2 + s)\,ds.$$

By hypothesis, $T_2 - T_1 > (c_1 - c_2)/(rc_2)$ which implies $T_2 - T_1 > u(c_1 - c_2)/c_2$ for $u \in [0, 1/r]$ or $T_2 - T_1 + u > uc_1/c_2$ for $u \in [0, 1/r]$. Since $D(\cdot)$ is increasing, this inequality implies

$$\frac{rc_2}{c_1}\int_0^{c_1/(rc_2)} D(T_1 + s)\,ds = r\int_0^{1/r} D(T_1 + uc_1/c_2)\,du$$
$$< r\int_0^{1/r} D(T_1 + (T_2 - T_1 + u))\,du = r\int_0^{1/r} D(T_2 + u)\,du.$$

Q.E.D.

PROOF OF LEMMA 1: Let $h(t + s) = f(t) + (s/y)[f(t + y) - f(t)]$, $s \in [0,y]$. Since $f$ is convex, $f(t + s) \leqslant h(t + s)$, $s \in [0,y]$. Hence, $\frac{1}{y}\int_t^{t+y} f(s)\,ds \leqslant \frac{1}{y}\int_t^{t+y} h(s)\,ds = \frac{1}{y}\int_0^y h(t + s)\,ds = \frac{1}{2}[f(t + y) + f(t)]$.

Q.E.D.

PROOF OF PROPOSITION 6: By Proposition 4, it suffices to show

$$(rc_2/c_1)\int_0^{c_1/rc_2} D(T_1 + u)\,du \geqslant r\int_0^{1/r} D(T_2 + u)\,du.$$

Letting $f(t) = D(T_1 + t)$ and $T = T_2 - T_1$, this becomes

(A.1) $$(rc_2/c_1)\int_0^{c_1/rc_2} f(s)\,ds \geqslant r\int_0^{1/r} f(T + s)\,ds$$

Note that $D(\cdot)$ being an arbitrary convex increasing function implies $f(\cdot)$ is also convex increasing.

Let $x = c_1/rc_2$ and let $a = c_2/c_1 < 1$. Then Eq. (A.1) becomes

$$(1/ax)\int_T^{T+ax} f(s)\,ds \leqslant (1/x)\int_0^x f(s)\,ds$$

which will follow from

(A.2) $$\frac{1}{ax}\int_T^{T+ax} f(s)\,ds \leqslant \frac{1}{2T + ax}\int_0^{2T+ax} f(s)\,ds \leqslant \frac{1}{x}\int_0^x f(s)\,ds$$

Note that the above expressions are all average values of $f(s)$ over their respective intervals of integration. The right-hand inequality follows from $f(\cdot)$ increasing and $2T + ax \leqslant x$. To prove the theorem, it only remains to estabilsih the left-hand inequality in Eq. (A.2).

(A.3) $$\frac{1}{2T + ax}\int_0^{2T+ax} f(s)\,ds - \frac{1}{ax}\int_T^{T+ax} f(s)\,ds$$
$$= \int_0^T \frac{d}{d\delta}\left[\frac{1}{2\delta + ax}\int_{T-\delta}^{T+ax+\delta} f(s)\,ds\right] d\delta$$

$$\frac{d}{d\delta}\left[\frac{1}{ax+2\delta}\int_{T-\delta}^{T+ax+\delta} f(s)\,ds\right]$$

$$= \frac{2}{ax+2\delta}\left[\frac{1}{2}\left(f(T-\delta)+f(T+ax+\delta)\right)\right.$$

$$\left.-\frac{1}{ax+2\delta}\int_{T-\delta}^{T+ax+\delta} f(s)\,ds\right]$$

which is non-negative for $\delta \in [0,T]$ by Lemma 1. Therefore, (A.3) is non-negative and

$$\frac{1}{2T+ax}\int_0^{2T+ax} f(s)\,ds \geqslant \frac{1}{ax}\int_T^{T+ax} f(s)\,ds.$$

Q.E.D.

# OPTIMAL ADMISSION PRICING POLICIES FOR $M/E_k/1$ QUEUES

Michael Q. Anderson

*The Robert O. Anderson Schools of Management*
*The University of New Mexico*
*Albuquerque, New Mexico*

**ABSTRACT**

This paper extends the Low-Lippman M/M/1 model to the case of Gamma service times. Specifically, we have a queue in which arrivals are Poisson, service time is Gamma-distributed, and the arrival rate to the system is subject to setting an admission fee p. The arrival rate $\lambda(p)$ is non-increasing in $p$. We prove that the optimal admission fee $p^*$ is a non-decreasing function of the customer work load on the server. The proof is for an infinite capacity queue and holds for the infinite horizon continuous time Markov decision process. In the special case of exponential service time, we extend the Low-Lippman model to include a state-dependent service rate and service cost structure (for finite or infinite time horizon and queue capacity). Relatively recent dynamic programming techniques are employed throughout the paper. Due to the large class of functions represented by the Gamma family, the extension is of interest and utility.

In this paper, we consider an $M/E_k/1$ queuing system in which the arrival rate to the system is subject to control by the service facility. Low [10] considered a similiar control model for $M/M/C$ systems and Lippman [6] has approached the Low model using new techniques. In this paper we extend Lippman's treatment (see pp 707-708 of [6]) to the case of gamma service time distribution.

The model we treat is the following. Customers arrive to a single server infinite capacity queue according to a Poisson process having rate $\lambda(p)$, where $p$ is the decision variable representing the price charged for admission to the system. We take $p \in P$ and assume $\lambda$ is a nonincreasing function of $p$. The service time is gamma with parameters $(k,\mu)$ (defined such that the mean service time is $k/\mu$). We think in terms of the customer going through $k$ serial phases of service, each having exponential service time with rate $\mu > 0$. Furthermore, we assume that these phases are observable by the decision maker (See "Applications" below, and the comment following Theorem 2). We define the state of the system to be the number of phases of service left to complete. Thus, if there are $c > 1$ customers in the system, and the customer being served has $f$ phases of service remaining, the state of the system would be $(c-1)k + f$. Denote the state by $i$ and the state space by $S = \{0, 1, 2, \ldots\}$.

The cost structure has two components. First the admission price $p$ is treated as a cost $-p$. Secondly, there is a holding cost $h(i)$ defined as a function of the state variable. The optimality criterion is the minimization of expected discounted cost. We treat the infinite time

horizon. This defines our model as a continuous time Markov decision process (see Ross [11], Bertsekas [2]).

Our aim is to prove that the optimal admission price is a nondecreasing function of the state variable. In addition, we extend the results of the Low-Lippman model [6] to include state-dependent service rates and service cost rate in the special case of exponential service distributions ($k = 1$).

In concluding this summary we remark that the techniques employed in this paper represent an effort to employ more general dynamic programming tools which have recently appeared in the literature. Lippman [6] notes the desirability of a more general class of techniques applicable to queuing optimization problems. In this paper we have formulated our problem in the general setting of Topkis' framework for analyzing monotone policies [13], Lippman's approximating techniques for finite and infinite horizon processes [6,8], and Lippman's recent result on unbounded 1-period cost functions [7].

## APPLICATIONS

Since an arrival to the system by a single customer results in the state variable increasing by $k$ units, we may view the model as applying to batch arrivals of jobs to a single facility server who services each job in series (at an exponential service rate), completing the entire batch before starting work on the next batch. This would apply to batch arrivals by identical jobs to a machine for processing. Or, taking the other view that arrivals to the system represent single jobs, the gamma service time model would apply to situations where the server must perform a series of $k$ well-defined (observable) processing operations on each job, where each of the $k$ subtasks are (essentially) identical in service time requirements.

The entrance price introduces a revenue component to the models. Holding cost might reflect costs of in-process inventories or higher machine (or service facility) maintenance cost.

Finally, the extension of the Low-Lippman model to accommodate state-dependent arrival rates is applicable to two general classes of situations: (1) Where a single service facility adjusts its service rate in response to increased work load by speeding up or slowing down, and (2) where such adjustment is made by adding or relieving additional service facility support (for example, when the service facility is a team, by adjusting the number of team members-at possibly a higher service cost rate).

## 1. EQUIVALENCE BETWEEN CTMDP AND DTMDP

In this section we summarize a new technique for transforming a continuous time Markov decision process (CTMDP) to an equivalent discrete time Markov decision process (DTMDP). Lippman appears to be the first to employ the method for queuing optimization problems [6]. See also [8], Serfozo [12], Kakumanu [4], Winston [14], and Anderson [1].

For the following discussion refer to Ross [11] or Bertsekas [2]. A DTMDP is defined by specifying four objects: a state space $S$, a collection of action spaces $\{A_i : i \in S\}$, transition probabilities $\{P_{ij}(a); i, j \in s, a \in A_i\}$, and cost functions $\{C(i,a); i \in S, a \in A_i\}$. For a CTMDP, in addition to these four objects we must specify transition time probability distributions $\{F_{ij}(a) : i, j \in S, a \in A_i\}$ where $F_{ij}(a)$ is an exponential distribution having parameter $\lambda(i,a)$. Also costs $C(i,a)$ are realized per unit time. For either process we specify a discount factor $\alpha$, $0 < \alpha < 1$ and define $V$ to be the optimal return function: $V(i)$ equals the minimum

ected discounted cost realized over an infinite time horizon starting in state $i \in S$. We may n define the following functional equations of dynamic programming.

For the DTMDP, $0 \leqslant n \leqslant +\infty$, and $i \in S$,

$$V_0(i) = 0$$

$$V_{n+1}(i) = \min_{a \in A_i}\{C(i,a) + \sum_{j=0}^{\infty} P_{ij}(a)\, V_n(j)\}/(\Lambda + \alpha)$$

e define $V_\infty = V$.)

the CTMDP,

$$V(i) = \min_{a \in A_i}\{\bar{C}(i,a) + q(i,a)\sum_{j \neq i} P_{ij}(a)\, V(j)\}$$

ere

$$q(i,a) = \lambda(i,a)/(\alpha + \lambda(i,a))$$

$$\bar{C}(i,a) = C(i,a)/(\alpha + \lambda(i,a)).$$

*olicy* is a rule for choosing actions, i.e., is a function $\delta: T \times S \to A$ where $T$ is the time axis he process ($T = \{0, 1, 2, \ldots\}$ for DTMDP and $T = [0, L]$, $L \leqslant +\infty$ for the CTMDP), and $i) \in A_i$ for each $i \in S$. Define $V_\delta$ to be the return function when using policy $\delta$. An imal policy $\delta^*$ is one satisfying $V_\delta{}^* = V$. $\delta^*$ is stationary if it is constant on the factor $T$. nditions which insure the existence of an optimal stationary policy are specified by the well-own contraction and monotonicity properties of Denardo [3]. When an optimal stationary icy exists, $V$ then satisfies equation (2) (in the case of the DTMDP) or (3) (in the case of CTMDP).

Given a CTMDP for which an optimal stationary policy exists, assume that $= \sup\{\lambda(i,a): i \in S,\ a \in A\} < +\infty$.

ine

$$P_{ij}'(a) = \begin{cases} (\Lambda - \lambda(i,a))/\Lambda & i = j \\ \lambda(i,a) P_{ij}(a)/\Lambda & i \neq j \end{cases}$$

an then be shown (see [6,8], [4], [12], [5, p. 121], or [1]) that the CTMDP is computation-equivalent to a DTMDP having the same state and action spaces, cost function $.a)/(\Lambda + \alpha)$, transition probabilities $P_{ij}'(a)$, and discount factor $\Lambda/(\Lambda + \alpha)$. Thus, the mizing actions for the CTMDP can be computed from the following functional equations, $i \in S$,

$$V(i) = \min_{a \in A_i}\{C(i,a)/(\Lambda + \alpha) + (\Lambda/(\Lambda + \alpha))\sum P_{ij}'(a)\, V(j)\}$$

(3), $\lambda(i,a) P_{ij}(a)$ is typically referred to as the *rate* of the transition $i \to j$). For a treatment CTMDP having lump sum costs see [1] or [12].

## IONOTONE OPTIMAL POLICIES

In a somewhat technical paper, Topkis [13] presents a general framework for analyzing notone optimal policies in optimization problems. For our purposes here, we specialize his

results to the problem we will be treating. Let $J$ be a function defined on $A \times B$, where $A$ $B$ are lattices. A function $J$ is said to be *submodular* on $A \times B$ if for $a_1, a_2 \in A$, $b_1$, $\in B$, $a_2 \geqslant a_1$, $b_2 \geqslant b_1$, $J(a_1, b_1) + J(a_2, b_2) \leqslant J(a_1, b_2) + J(a_2, b_1)$. Suppose we h the following optimization problem, $f(a) = \min_{b \in B} J(a, b)$. Let $b^*(a)$ be the infimum of the of points in $B$ which minimizes $J(a, b)$. Then under certain conditions on $A$, $B$, and $J$, To shows that to establish that $b^*$ is non-decreasing in $a$, it is sufficient that $J$ be submodular $A \times B$. (See Theorem 6.2 of [13] for details.)

## 3. AN $M/E_k/1$ QUEUE WITH CONTROLLED ARRIVAL RATE: THE INFINITE TIME HORIZON PROBLEM

We make the following assumptions:

(1) $p \in P$, where $P$ is a compact subspace of $[0, +\infty]$.

(2) $\lambda$ is a nonincreasing continuous function of $p \in P$. Also for $p \in P$, $0 \leqslant \lambda(p)$ $\infty$.

(3) $h$ is nondecreasing in $i \in S = \{0, 1, 2 \ldots\}$ and satisfies $h(i + k) - h(i + k$ $\geqslant h(i) - h(i - 1)$, $i \geqslant 1$.

Notation: Let $f$ be a real-valued function on $\{0, 1, 2, \ldots\}$. Define $\Delta f(i) = f(i) - f(i - 1)$, 1. Define $\Lambda = \sup_{p \in P} \lambda(p) + \mu$.

From Section 1, the equivalent discrete time recursions for our CTMDP are easily seen t the following:

For $i \in S$,

$$V_0(i) = 0 \tag{6}$$

$$V_n(i) = \min_p J(n, i, p)/(\Lambda + \alpha), \quad n > 0 \tag{7}$$

where $J$ is defined as follows:

For $i = 0$, and $0 \leqslant n \leqslant +\infty$,

$$J(n, 0, p) = -p\lambda(p) + \lambda(p) V_{n-1}(k) + V_{n-1}(0)(\Lambda - \lambda(p)), \tag{8}$$

and for $i > 0$,

$$J(n, i, p) = h(i) - p\lambda(p) + \lambda(p) V_{n-1}(i + k) + \mu V_{n-1}(i - 1)$$

$$+ V_{n-1}(i)(\Lambda - \lambda(p) - \mu) \tag{9}$$

Notation: In (7) define the infimum of the minimizing actions by $p^*(i, n)$ for $n < +\infty$ $p^*(i)$ when $n = +\infty$. Also write $V_\infty = V$ and when $n = +\infty$ we suppress $n$ in the li arguments for $J$, i.e., we will write $J(i, p)$.

THEOREM 1. There exists an optimal stationary policy to the CTMDP defined above

PROOF. The primary consideration here is the presence of the unbounded 1-perio function $h$. Although Denardo's results [3] do not directly apply, Lippman [7] has estab conditions under which Denardo's contraction and monotonicity results do apply. In our c is trivial to verify that Assumptions 1,2, and 3 of [7, p. 1227] are satisfied, and the theor established.

We now establish two supporting lemmas.

LEMMA 1. For $0 \leqslant n \leqslant +\infty$ and $i \geqslant 1$, $\Delta V_n(i) \geqslant 0$.

PROOF. We prove the lemma for finite $n$ (by finite induction); the case for infinite $n$ then follows by taking limits. From (6) the result holds trivially for $n = 0$. First assume $i > 1$ and define $p = p^*(n,i)$. Then from (7) and (9),

$$V_n(i) - V_n(i-1) \geqslant J(n,i,p) - J(n,i-,p)$$
$$= \Delta h(i) + \Delta V_{n-1}(i+k)\lambda(p) + \Delta V_{n-1}(i)(\Lambda - \lambda(p) - \mu) + \Delta V_{n-1}(i-1)\mu$$

$\geqslant 0$, where the inequality follows from Assumption (3), the definition of $\Lambda$, and the inductive hypothesis $\Delta V_{n-1} \geqslant 0$. When $i = 1$, the same computation holds except that the term $\Delta V_{n-1}(i-1)\mu$ is not present. ■

The next result establishes a condition on $V_n$ which is somewhat weaker than the convexity required in Lippman's treatment [6, p. 708, Theorem 15].

LEMMA 2. For $0 \leqslant n \leqslant +\infty$ and $i \geqslant 1$, $\Delta V_n(i+k) \geqslant \Delta V_n(i)$

PROOF. For $n = 0$, the result holds trivially. Assume it true for $n - 1 < +\infty$. Define $p_1 = p^*(n,i-1)$ and $p_2 = p^*(n,i+k)$. Then for $i > 1$, from (7), (8) and (9), $\Delta V_n(i+k) - \Delta V_n(i) \geqslant [J(n,i+k,p_2) - J(n,i+k-1,p_2)] - [J(n,i,p_1) - J(n,i-1,p_1)]$

$$= \Delta h(i+k) - \Delta h(i)$$
$$+ \Delta V_{n-1}(i+2k)\lambda(p_2)$$
$$+ \Delta V_{n-1}(i+k)(\Lambda - \mu - \lambda(p_2) - \lambda(p_1))$$
$$+ \Delta V_{n-1}(i+k-1)\mu$$
$$+ \Delta V_{n-1}(i)(-\Lambda + \mu + \lambda(p_1))$$

$\geqslant \Delta V_{n-1}(i+k-1)\mu \geqslant 0$, where the next to last inequality follows from Assumption (3), the inductive hypothesis and the definition of $\Lambda$, and the last inequality follows from Lemma 1.

For $i = 1$, a similar argument combined with the result for $i > 1$ yields the desired result. For $n = +\infty$, the result obtains by taking the limit as $n \to +\infty$ on the finite case. ■

We now establish our first main result,

THEOREM 2. The optimal entrance price is a nondecreasing function of $i \in S$, i.e., $p^*(i+1) \geqslant p^*(i)$.

PROOF. From Section 2, it suffices to show that $J$ is submodular on $S \times P$. Let $p_1 \leqslant p_2$. Define $\Delta_{ip}$ by

$$\Delta_{ip} = J(i,p_1) + J(i+1,p_2) - J(i,p_2) - J(i+1,p_1).$$

Then from (9), for $i > 0$.

$$\Delta_{ip} = (\Delta V(i+1) - \Delta V(i+k+1))(\lambda(p_1) - \lambda(p_2))$$

$\leqslant 0$, where the inequality follows from Assumption (1) and Lemma 2. when $i = 0$ we must use (8) and (9) and the same computation results. To establish the general case

$J(i,p_1) + J(i+j,p_2) - J(i,p_2) - J(i+j,p_j) \leqslant 0$. Note that this expression can be written as $\sum_{l=0}^{j-1} \Delta_{i+l,p}$ each term of which is nonpositive by the case above. ■

Thus, the greater the "load" on the server (as measured by $i \in S$), the greater is the price charged for admission to the system and thus, the lower the arrival rate. From the definition of the state variable we conclude that the optimal entrance fee $p^*$ is nondecreasing in both the number of customers in the system and the number of phases of service remaining on the customer currently being served. Along these lines, as a special case, we can take $h$ to be a function only of the number of *customers* in the system by defining $h$ to be constant on each segment. $[0k)$, $[k,2k)$, ... and requiring that the restriction of $h$ to the set $\{0,k,2k,\ldots\}$ be convex. It is easily verified that all proofs go through.

It would be of interest to extend Theorem 2 to the truly finite horizon CTMDP. However we have thus far been unable to do this. Using results of Lippman ([8]—see the connectedness condition on page 483 and also Theorem 4 of that paper) the monotonicity of $p^*$ for the finite time horizon would follow if we could prove that $J$ is submodular in $n$ and $p$, and this in turn would follow from $\Delta V_n(i) \geqslant \Delta V_{n-1}(i)$. We have been unable to establish this last inequality (or its reverse which would also be sufficient).

## 4. THE LOW-LIPPMAN MODEL WITH STATE-DEPENDENT SERVICE RATES

Once again we consider Lippman's model [6, p. 707], this time incorporating a state-dependent service rate and a state-dependent service cost. Specifically we have an $M/M/1$ queue with queue capacity $Q$, where $Q \leqslant +\infty$. An entrance price $p$ is charged and the arrival rate $\lambda$ is a function of $p$. The state of the system is the number of customers present and is denoted by $i \in S = \{0, 1, \ldots, Q\}$. Given the state is $i$, a holding cost $h(i)$ per unit time is incurred. Furthermore, we shall assume that the server works at rate $\mu_i$ and incurs a service cost $C(\mu_i)$ per unit time when the state is $i$.

The functional equations corresponding to (6), (7), (8), and (9) are, for $Q = +\infty$,

$$V_0(i) = 0$$

$$V_n(i) = \min J(n,i,p)/(\Lambda + \alpha), \quad n > 0$$

For $i = 0$, $0 \leqslant n \leqslant +\infty$

$$J(n,0,p) = -p\lambda(p) + \lambda(p)V_{n-1}(1) + V_{n-1}(0)(\Lambda - \lambda(p))$$

and for $i > 0$,

$$J(n,i,p) = h(i) + C(\mu_i) - p\lambda(p) + \lambda(p)V_{n-1}(i+1) + \mu_i V_{n-1}(i-1) + V_{n-1}(i)(\Lambda - \mu_i - \lambda(p))$$

when $Q < +\infty$, and $i = Q$, the transition $i \to i+1$ has rate zero. In addition to Assumptions (1) and (2) of Section 3 above, we make the following:

(2b) When $Q < +\infty$ we make the convention that $p^*(Q) = \sup\{p \in P\}$ and $\lambda(p^*(Q)) = 0$.

(3) $h$ is nondecreasing and convex in $i \in S = \{0, 1, 2, \ldots, Q\}$

(4) $C(\mu_i)$ is convex increasing in $i \in S = \{0, 1, 2, \ldots, Q\}$. (As a special case we may take $C(\mu_i) \equiv 0$).

(5) $\mu_i$ is concave in $i$ and satisfies $2\mu_i - \mu_2 \geqslant 0$. When $Q < +\infty$ we require additionally that $2\mu_Q - \mu_Q \geqslant \sup\{\lambda(p)\}$. When $Q = +\infty$, in order that $\Lambda$ be finite we assume that $\mu_i$ is non-decreasing in $i$.

When $\mu_i$ is taken to be non-decreasing in $i$, this may be interpreted in terms of the server choosing to increase the service rate (at higher costs in case $C(\mu_i) > 0$). The case where $\mu_i$ is non-increasing in $i$ would represent a system where the effectiveness of the service facility decreases as the work load increases.

We may now prove the analogs of all the previous results. The proofs obtain by employing the same techniques as before. We remark that $\Lambda$ is defined to be the $\sup_{i,p}\{\lambda(p) + \mu_i\} < +\infty$ (by Assumptions (2) and (5)). Lemma 1 is straightforward. Lemma 2 is restated: $\Delta V_n(i) \geqslant \Delta V_n(i-1)$ i.e., $V_n$ is convex in $i$. To establish this requires Assumptions (4) and (5). In Theorem 2, we establish submodularity as before; when $Q < +\infty$ we observe Assumption (2b). Furthermore, for this model we may prove that $p^*$ is non-decreasing in $i$ for the finite horizon continuous time problem as follows. We first show that $p^*$ is non-increasing as a function of $n$ by establishing that $J$ is submodular in $n$ and $p$. This last condition will follow from $\Delta V_n(i) \geqslant \Delta V_{n-1}(i)$ which is easily proved. We then invoke Lippman's connectedness condition to establish Theorem 4 of [8] (see page 483). In every instance we observe the various special cases ($i = 0$, $i = Q < +\infty$). We summarize these remarks in the following Theorem.

THEOREM 3. In the Low-Lippman Model [6, p. 707] with state-dependent service rates and service costs, under Assumptions (1) through (5), and for finite or infinite queue capacity, in both the finite and infinite time horizon CTMDP, the optimal entrance price is a non-decreasing function of the number of customers in the system.

We note that the special case $C \equiv 0$ and $h$ linear was treated by Lippman and Stidham [9].

## 5. TOPICS FOR FURTHER RESEARCH

We have thus far been unable to entend Theorem 2 to either the finite queue capacity case or to the state-dependent service rate model. A few numerical computations have been made but no counter-examples yet found. This might be a research topic of interest to other authors working in the area.

In general the technique of transforming a CTMDP to a DTMDP appears to be a very effective tool for analyzing birth-death processes. However, when more general state transitions are allowed, complications seem to appear. The author in [1] has had some success in this respect in treating various machine maintenance models. Another difficult class of control problems (using the transformation technique CTMDP $\rightarrow$ DTMDP) are Markov decision processes involving two dimensional state spaces (such as one would have in treating tandem queues, or the operation of two machines, for example). This might be a rich field of research for authors interested in applying the new methods employed in this paper for analyzing Markov processes with structure.

## REFERENCES

[1] Anderson, M.Q., "Monotone Optimal Maintenance Policies For Equipment Subject to Markovian Deterioration," Doctoral Dissertation, Indiana University (1977).

[2] Bertsekas, D.P., *Dynamic Programming and Stochastic Control* (Academic Press, New York, 1976).

[3] Denardo, E., "Contraction Mappings in the Theory Underlying Dynamic Programming," SIAM Review, *9*, 165-177 (1967).

[4] Kakumanu, P., "Relation Between Continuous and Discrete Time Markovian Decision Problems," Naval Research Logistics Quarterly, *24*, 431-441 (1977).

[5] Howard, R., *Dynamic Programming and Markov Processes* (John Wiley, New York, 1970).

[6] Lippman, S., "Applying a New Device in the Optimization of Exponential Queuing Systems," Operations Research, Vol. *23*, 687-710 (1975).

[7] Lippman, S., "On Dynamic Programming With Unbounded Rewards," Management Science, Vol. *21*, 1225-1233 (1975).

[8] Lippman, S., "Countable-State, Continuous-Time-Dynamic Programming With Structure," Operations Research, Vol. *24*, 477-490 (1976).

[9] Lippman, S., and S. Stidham, "Individual Versus Social Optimization in Exponential Congestion Systems," Operations Research, Vol. *25*, 233-247 (1977).

[10] Low, D., "Optimal Dynamic Pricing Policies for an M/M/S Queue," Operations Research, Vol. *22*, 545-561 (1974).

[11] Ross, S., *Applied Probability Models* Helden-Day, San Francisco, CA (1970).

[12] Serfozo, R.F., "An Equivalence Between Continuous and Discrete Time Markov Decision Processes," Operations Research, Vol *27*, 616-620 (May-June 1976).

[13] Topkis, D.M., "Minimizing a Submodular Function on a Lattice," Operations Research, Vol. *28*, 305-321 (1978).

[14] Winston, W.L., Optimal Operation of Congestion Systems with Heterogeneous Arrivals and Servers, Ph.d. Dissertation, Yale University (1975).

# THE DYNAMIC TRANSPORTATION PROBLEM: A SURVEY

James H. Bookbinder and Suresh P. Sethi

*Faculty of Management Studies*
*University of Toronto*
*Toronto, Ontario*

**ABSTRACT**

The dynamic transportation problem is a transportation problem over time. That is, a problem of selecting at each instant of time $t$, the optimal flow of commodities from various sources to various sinks in a given network so as to minimize the total cost of transportation subject to some supply and demand constraints. While the earliest formulation of the problem dates back to 1958 as a problem of finding the maximal flow through a dynamic network in a given time, the problem has received wider attention only in the last ten years. During these years, the problem has been tackled by network techniques, linear programming, dynamic programming, combinational methods, nonlinear programming and finally, the optimal control theory. This paper is an up-to-date survey of the various analyses of the problem along with a critical discussion, comparison, and extensions of various formulations and techniques used. The survey concludes with a number of important suggestions for future work.

## 1. INTRODUCTION

The classical transportation problem refers to the shipment of goods from a set of sources to a collection of sinks at a minimum cost. Analysis of these problems dates back to the works of Kantorovich [28], Hitchcock [26], Koopmans [31], and Dantzig [9]. Subsequently, much work has been done. (See e.g., Holladay [27], Potts and Oliver [35], Bradley [5], Christofides [7], and Kennington [29].)

The classical problem is static in nature in the sense that shipments are instantaneous, and costs as well as supply and demand requirements are independent of time. A notational formulation of the static transportation problem appears in Section 2.

That the dynamics of the transportation problem are important was first recognized by Ford [15] in his formulation of the maximal dynamic flow problem. (See also Gale [19] and Ford and Fulkerson [16].) Since then, a number of other dynamic extensions have been studied. These include the minimum-time transportation problem (Szwarc [44,45,46], Hammer [23], Tapiero and Soliman [51], Tapiero [49], and Srinivasan and Thompson [41]); the minimum-cost transportation problem (Bellmore, Eklof, and Nemhauser [2], Szwarc [47], Tapiero [48], Tapiero and Soliman [51], and Srinivasan and Thompson [41,42]); and the maximal dynamic flow problem, both for a single commodity (Ford [15], Ford and Fulkerson [16]) and in the multi-commodity case (Bellmore and Vemuganti [3]).

It is the purpose of the present paper to survey the above work as well as some applications of the dynamic transportation problem. A schematic representation of the various treatments of this problem is given in Figure 1.

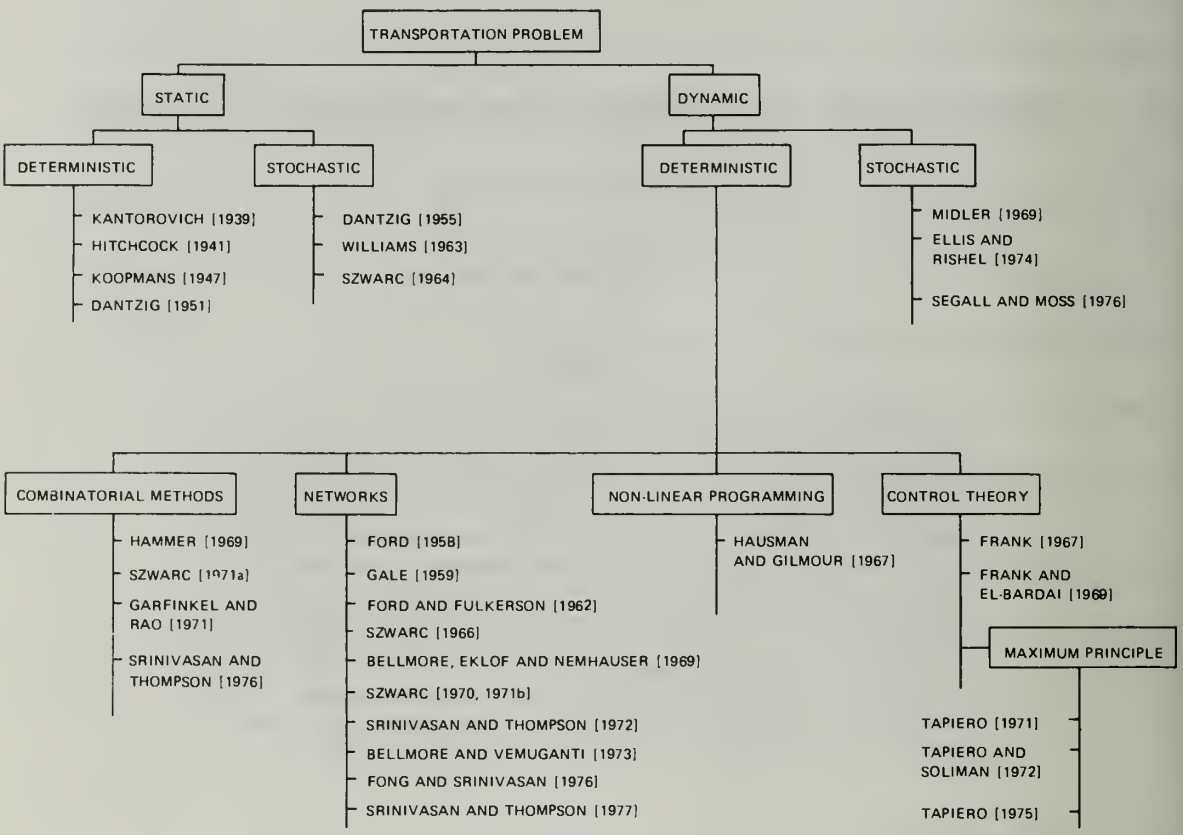


FIGURE 1

The organization of our survey is as follows. Section 2 establishes the basic notation and points out some of the possible ways in which the static problem can be generalized to the time-dependent case. Sections 3, 4, and 5 deal with the case of discrete time, and continuous time is treated in Section 6.

Section 3 is concerned with the deterministic dynamic transportation problems, i.e., the minimum-time, minimum-cost, and the maximal dynamic flow problems. We consider in Section 4 a stochastic multiperiod, multimode model (Midler [34]), a problem of discrete-time stochastic optimal control which is solved via a dynamic programming algorithm.

Section 5 deals with two illustrative applications of the discrete time dynamic transportation problem. One involves a multiperiod truck delivery problem (Hausman and Gilmour [25]), and the other studies the optimal rescheduling of air traffic in response to stochastic influences (Ellis and Rishel [11]).

Section 6 is concerned with the dynamic transportation problem over continuous time, and moreover, when explicit account is taken of the time delay between dispatch of a shipment at the sink and its receipt at the source. The appropriate framework is that of optimal control

theory, and papers surveyed include those of Frank [17], Frank and El-Bardai [18], Tapiero and Soliman [51], and Tapiero [49].

Section 7 concludes the paper with extensive outlines and suggestions for future research. Portions of this work are already under way (Sethi and Bookbinder [38]).

## 2. STATIC TRANSPORTATION PROBLEM

By the static transportation problem, we shall mean the well-known linear programming problem

$$\text{Min } J = \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij} u_{ij},$$

subject to

$$\sum_{j=1}^{n} u_{ij} = A_i \quad \text{(Supply constraints)}$$

$$\sum_{i=1}^{m} u_{ij} = B_j \quad \text{(Demand constraints)}$$

$$0 \leqslant u_{ji} \leqslant c_{ij} \quad \text{(Capacity restrictions)}.$$

In this notation (which has been chosen to conform with that of the control-theoretic treatment of transportation problems), $u_{ij}$ is the shipment from origin $i$ to destination $j$, at a cost of $q_{ij}$. Each source node $i$ has a supply $A_i$, the sink nodes $j$ require an amount $B_j$, and so a transportation schedule is to be found that minimizes the total shipping costs while not exceeding the capacity $c_{ij}$ of any arc $(i,j)$. It should be noted that the above problem refers to the shipment of a *single commodity* (or at least only a standard "product mix") and only a *single mode* of transportation is considered for each link $(i,j)$.

There are a number of possible ways in which this problem can be made time-dependent:

### 2.1. $A_i$ and/or $B_j$ Functions of Time

If the supply capabilities or the demand requirements are time-dependent, then the production schedule can be smoothed by incurring the costs of holding inventory which was delivered early. Inventory variables could be allowed at sources as well as sinks, and there may be included upperbounds, possibly time-dependent, on those inventories.

### 2.2. $q_{ij}$ Functions of Time

The remarks of the preceding paragraph are applicable here as well.

### 2.3. Associate *a* Time $\tau_{ij}$ with Each $u_{ij}$

The notation $\tau_{ij}$ refers to the time required to ship the $u_{ij}$ units over the route $(i,j)$, so that the relevant problem is the minimization of the time needed to satisfy all the demand requirements. It should be noted that the costs $q_{ij}$ would not normally enter into this problem (though they could of course be taken into account via subsidiary constraints; see Glickman and Berger [21], and also Srinivasan and Thompson [42]).

A convenient division of the papers in this survey is according to whether the analytical techniques involve *discrete time* or *continuous time.* We shall begin with the former, and in fact with a problem of type 3 above.

## 3. DETERMINISTIC DYNAMIC TRANSPORTATION PROBLEM: DISCRETE TIME

### 3.1. The Minimum-Time Transportation Problem

One is interested in minimization of the total time required to transport all the goods from the origins to their destinations, i.e.,

$$\text{Min } t^* = \{\max(\tau_{ij}) | u_{ij} > 0\} .$$

Possible applications include the transport of *perishable* goods; the movements of military units from bases to fronts; or the shipment of customers' orders which are of a priority or "rush" nature.

The minimum-time transportation problem has been considered by Szwarc [44,45,46] and Hammer [23,24]. (See also Garfinkel and Rao [20] and Srinivasan and Thompson [41].) Szwarc [45] modified his original [44] algorithm to prevent it from cycling. Szwarc [46] in a similar way refined Hammer's [23] algorithm and revised a proof concerning the equivalence of local and global optima. Hammer [24] also furnished some amendment of his own work.

Szwarc [46] has shown that his algorithm [45] is essentially equivalent to that of Hammer [23]. By this, we mean they both produce the same sequence of basic feasible solutions, given that they begin with the same initial solution. The Szwarc-Hammer algorithm is a primal algorithm, but Garfinkel and Rao [20] employ a "threshold" algorithm that yields a primal-feasible solution only upon termination. Srinivasan and Thompson [41] have studied the Bottleneck Time and Bottleneck Shipment transportation problems, and have furnished an algorithm for the former (i.e., for the minimum-time transportation problem) that is virtually identical to the Szwarc-Hammer algorithm. Srinivasan and Thompson [41] have also shown that their algorithm, and hence the Szwarc-Hammer algorithm, is computationally more efficient than that of Garfinkel and Rao [20].

The main points of the Szwarc-Hammer algorithm involve the following steps:

1. Find an initial basic feasible solution (BFS) to the problem. This is, of course, a basis of order $(m + n - 1)$. Let $t$ denote the current value of the bottleneck time.

2. Find an *adjacent* (i.e., a change of only 1 basis element) BFS which is *better* (i.e., either $t$ is lower, or if $t$ is unchanged, there is a smaller quantity shipped on the bottleneck arc). Hammer gives a 4-step procedure by which this can be done:

   (i) Among these basic arcs $(i,j)$ for which $\tau_{ij} = t$, identify the maximum quantity shipped. Call the quantity $u_{hk}$.

   (ii) Determine the set $S_{h,k}$ of all non-basic elements $(p,q)$, such that if $u_{pq}$ were introduced into the basis, the shipment $u_{hk}$ would thereby be reduced.

   (iii) Choose among the elements of $S_{h,k}$ the one, say $(p_0,q_0)$, for which $\tau_{pq}$ is the smallest.

(iv) Enter $u_{p_0 q_0}$ into the basis, as in the usual transportation problem. Update $t$. (Szwarc, at this stage, eliminates from further consideration any non-basic $u_{ij}$ for which $\tau_{ij} \geqslant t$.)

3. Check for optimality, i.e., for the existence of a still better BFS. Szwarc's procedure follows from his Theorem 2, which as above, indicates optimality when the only vacant "cells" $u_{ij}$ available for entry into the basis turn out to have $\tau_{ij} \geqslant t$.

4. Return to step 2 or terminate at optimal solution.

We remark that the preceding treatments have emphasized the time objectives with little mention of the concomitant *costs*. Hammer [23] did give an extension of his algorithm to determine, from among the alternative optimal solutions with the given $t^*$, that solution with minimum cost. An interesting extension would be the imposition of an additional constraint involving the upper bound of total cost. Alternatively, Glickman and Berger [21] have analyzed the trade-off between the cost and time of transport, by solving the usual minimum cost problem, subject to an upper bound on the time of transportation. Srinivasan and Thompson [42] have determined cost/time efficient frontiers for the multimodal problem. For every Pareto-optimal point, their algorithm furnishes the routes, modes, and quantities shipped.

### 3.2. Multiperiod Transportation Problem

Attention is now addressed to the problem of minimization of total costs when the $n$ sinks have demands for each of $T$ periods. This problem has been studied by Bellmore, Eklof, and Nemhauser [2] (B.E.N.), and by Szwarc [47]. (The latter reference first appeared in 1967 as a C.O.R.E. Discussion Paper, No. 6704.)

B.E.N. treat the general case in which $A_i$, $B_j$, and $q_{ij}$ are all functions of time, which is denoted by the presence of a superscript $t$. Shipments to the sinks from the $m$ sources can either take place in the same period as the demand requirement, corresponding to a shipping cost of

$$q_{ij}^t \, u_{ij}^t,$$

or *earlier*, in which case a linear cost for holding inventory is then also incurred. If the notation $y_j^t$ denotes the inventory carried from period $(t-1)$ to period $t$ at sink $j$, and $d_j^t$ the corresponding inventory holding cost, the problem of B.E.N. can then be formulated as:

$$\text{Min } J = \sum_{i,j} \sum_{t=1}^{T} q_{ij}^t \, u_{ij}^t + \sum_{j=1}^{n} \sum_{t=2}^{T} d_j^t \, y_j^t$$

subject to:

$$\sum_{j=1}^{n} u_{ij}^t \leqslant A_i^t, \quad \forall \, i, \; t$$

$$\sum_{i=1}^{m} u_{ij}^t + y_j^t - y_j^{t+1} \geqslant B_j^t, \quad \forall \, j, \; t$$

$$u_{ij} \geqslant 0$$

$$0 \leqslant y_j^t \leqslant N_j^t,$$

where $N_j^t$ is the maximum inventory which can be stored at sink $j$ between periods $(t-1)$ and $t$.

Were it not for the inclusion of the upperbounds $N_j^t$, the above problem could be solved by creation of a large *single-period* problem, with $(mT)$ sources and $(nT)$ sinks (Bowman [4] and Kreibel [32]). This approach, as B.E.N. point out, yields computer storage requirements that are proportional to $T^2$.

B.E.N. approach the problem through a *minimum cost flow* or *transshipment* model, i.e., $T(m$ — source, $n$ — sink) problems linked by inventory variables. (The "nodes" are the $A_i^t$ and $B_j^t$ for all values of $t$.) The computer storage requirements are then proportional to $T$. B.E.N. also point out a *decomposition* of the transshipment problem that requires rapid access computer storage of data for only $n$ more nodes and $n$ more arcs than does the standard transportation problem. That is, the storage requirements for the decomposition are then independent of $T$ (and, in fact, are at anytime comprised of the data for a single period in core).

The B.E.N. transshipment algorithm is effectively composed of the following steps:

1. Add an artificial source node $s^*$, and an artificial sink note $t^*$.

2. Introduce arc "distances", which are essentially $q_{ij}^t$ or $d_j^t$, suitably modified by the existing flow in the network. (Initial flow = 0.)

3. Find a shortest path from $s^*$ to $t^*$, and ship the maximum amount possible, consistent with the upperbounds on inventories.

4. Increase flows, revise "distances", and re-calculate shortest path.

5. Algorithm terminates when all demands have been satisfied.

The paper of Szwarc [47] is addressed to the same problem as that of B.E.N., but differs in that Szwarc does not include an inventory holding cost. B.E.N. include this cost, as well as upperbounds on inventories at each sink node. Neither Szwarc nor B.E.N. include upper bounds on the shipments $u_{ij}^t$, although as B.E.N. point out, the modification to the transshipment algorithm in this case is clear.

It may be argued that the B.E.N.'s desire to reduce the core requirements was motivated by desires more academic than practical, in that few real problems would assume as regular or straightforward a form. However, the work of B.E.N. should be quite practical for a more complicated problem with additional constraints, since the B.E.N. solution could be used for the initial iteration to the larger problem.

B.E.N. treated $A_i^t$, $B_j^t$, and $N_j^t$ as given, and determined optimal shipments $u_{ij}^t$ and inventories $y_j^t$ to minimize the sum of the latter two costs over the horizon $t = 1, 2, \ldots, T$. For intermediate values of $T$, related questions are how the prduction at each plant should be increased over time and how the shipping patterns should change, when market demands have given growth rates and there is thus an increase in the total volume handled by the system. Such optimal growth paths in logistics operations have been studied by Srinivasan and Thompson [40] when supply costs are linear or piecewise linear and convex. For still larger $T$, Fong and Srinivasan [14] have considered the multiperiod capacity expansion problem when market demands are nondecreasing over time and must be met exactly during each period. In the case of linear costs, they furnished an efficient algorithm to schedule capacity expansions and shipments to markets to minimize the discounted capacity expansion costs plus the costs of production and transportation.

## 3.3. Maximal Dynamic Flows

### 3.3.1. Single Commodity

The problem of maximal dynamic flow requires the determination of the largest commodity flow $v(T)$ which can occur between source and sink within a specified time horizon $T$. A capacity $c_{ij}$ and a minimum transit time $\tau_{ij}$ are associated with each arc $(i,j)$. Let $u_{ij}^t$ be the flow leaving node $i$ at time $t$, enroute to node $j$; for the deterministic case, arrival at node $j$ will occur at time $(t + \tau_{ij})$. Transshipment is allowed, with $u_{ii}^t$ denoting the commodity inventory at node $i$ held over from time $t$ to $(t + 1)$. (It turns out, however, that there always exists a maximal dynamic flow in which there are no holdovers at intermediate nodes.)

Letting $P$, $Q$ denote the source and sink nodes, respectively, the problem is then:

$$\text{Maximize } v(T)$$

subject to

$$\sum_{t=0}^{T} \sum_{j} [u_{Pj}^t - u_{jP}^{t-\tau_{jP}}] = v(T)$$

$$\sum_{j} [u_{ij}^t - u_{ji}^{t-\tau_{ji}}] = 0; \quad i \neq P, Q; \quad t = 0, 1, \ldots, T$$

$$\sum_{t=0}^{T} \sum_{j} [u_{Qj}^t - u_{jQ}^{t-\tau_{jQ}}] = -v(T)$$

$$0 \leqslant u_{ij}^t \leqslant c_{ij}.$$

This problem has been studied in detail by Ford [15] and by Ford and Fulkerson [16]. Their analysis is based upon a Time Expanded (T.E.) network, in which a new node is introduced for each (discrete) time period. The maximal dynamic flow is obtained by:

1. Solving a static transshipment problem in the T.E. network.

2. Repeating this flow for successive time periods.

Ford and Fulkerson showed that this maximal temporally repeated flow is the optimal dynamic flow.

### 3.3.2. Multicommodity

Ford and Fulkerson's result cannot be extended to more than one commodity, that is, to the case where the upperbounds $c_{ij}$ refer not to each commodity separately but rater to the *sum* of the $u_{ijk}$ over all commodities $k$. Bellmore and Vemuganti [3] (B.V.) have used the fact that, while the temporally repeated multicommodity flow need not be optimal, it *is* feasible. This furnishes a lower bound on the maximal dynamic flow.

B.V. showed that, as the number of time periods $T$ becomes large, an *upper* bound is furnished by

$$[\text{Multi. Comm. Max. Dyn. Flow for } T \text{ periods}]$$
$$-[\text{Max. temporally repeated flow for } T \text{ periods}] \leqslant \alpha.$$

The number $\alpha$ is independent of $T$, so that for large $T$, the percentage difference between the two terms on the left tends to zero. B.V. show how $\alpha$ can be obtained by solving the multicommodity transshipment problem on the static network.

The bounds $\alpha$ are refined by B.V. through the use of duality. As they note, their method and results generalize to the case of the *weighted*, multicommodity maximal dynamic flow problem. Nevertheless, as far as known to the present authors, no exact algorithm has been presented for the problem of maximal dynamic flow in the multicommodity case.

## 4. A STANDARD MULTIPERIOD, MULTIMODE MODEL

Our next illustrations for the case of discrete time incorporate uncertainty in the demand requirements $B_j$. For the single stage problem, treatments by Dantzig [10], Williams [52] and Szwarc [43] have employed stochastic linear programming.

In the multistage case, the probabilistic version of the transportation problem can be thought of as involving demand governed by a stochastic process. The paper by Midler [34] considers such a multiperiod problem with random demands for the multicommodity case. A dynamic programming algorithm is used to decide which mode of transport to employ, and to assign commodity classes to various modes and supply points to destinations.*

The problem is formulated as one of discrete time stochastic optimal control, in which the objective is minimization of total expected cost over the planning horizon. The costs considered are those of shipping, by commodity and route; costs of rerouting a carrier from a destination back to each origin; costs of stockouts and of carrying inventories at destinations; and costs involving deviations from port capacities. Midler assumed that the shipping costs are linear, with all other costs quadratic.

The dynamic equations were taken to be linear in the state and control variables, and the model solved in closed form. Midler found that the optimal values of the control variables could be expressed linearly in terms of the current values of the state variables, with the coefficients in this relationship depending upon the number of periods remaining in the horizon. Midler then furnished an extensive discussion of the solution to the transportation problem, including the optimal dependence upon time and upon location of the assignment of each commodity class to a mode, and the interpretation of the time-varying stochastic shadow prices.

## 5. SOME APPLICATIONS

We now discuss two problems which have been formulated as applications of one form or other of the dynamic transportation problem with discrete time.

### 5.1 A Multiperiod Truck Delivery Problem

The first application concerns the scheduling of truck deliveries. This problem, studied by Hausman and Gilmour [25], arose in the analysis of delivery of home heating oil. Deliveries are made from a single source to $n$ customers, each of which has a minimum required frequency of service $f_i$. This frequency $f_i$, which may be exceeded to take advantage of truck economies of scale, furnishes the time-dependence of the problem at hand.

*For the case of a dynamic communication network with stochastic inputs, Segall and Moss [37] have furnished an analysis (in the continuous-time framework) which uses Hamilton-Jacobi theory (see e.g., Fleming and Rishel [13]).

As an operating policy, the customers are classified into $k$ *groups*, where $k$ is treated as a decision variable. When any customer in a group needs service, the whole group is served. The objective function, then, is the assignment of each customer to a group, so that annual delivery costs are minimized.

The truck costs involve both a fixed cost per delivery, as well as a variable cost per mile travelled. The presence of this variable cost means that evaluation of the objective function requires knowledge of the *distance* travelled in the *optimal tour* of each group. This, of course, amounts to solving the *travelling salesman problem*, in general a difficult task indeed.

Hausman and Gilmour took the point of view that the customers in a group were likely to be reasonably close geographically. The truck driver would then intuitively choose a sequence of deliveries which involved a distance $D$ fairly close to that in the optimal tour, at least if there were not an "excessive" number of customers per group. (The latter is guaranteed by the limited capacity of the truck.) Accordingly, these authors estimated $\hat{D}_j$ for each group $j$ by multiple regression. This heuristic device employed as independent variables three simple statistics for the group under consideration. These variables involved the standard deviations of the latitudes and longitudes, and the number of customers in the group. Parameters were determined by comparison with some randomly generated travelling salesman problems and their optimal solutions.

Hausman and Gilmour's algorithm began with this approximation $\hat{D}_j$ for each group, followed by a search for a re-assignment of customer groupings in an attempt to lower the total costs. While the problem at hand is quite difficult, their algorithm does not appear to be very efficient, because customers are one-by-one temporarily re-assigned to each of the other groups.

Nevertheless, there are two interesting aspects to this work. The regression approach (to estimate $\hat{D}_j$) could, as Hausman and Gilmour point out, be useful in estimating bounds on the optimal solution to other problems, or in deciding whether to continue with further searches on a given problem.

Moreover, in the course of their work, Hausman and Gilmour found that they could still obtain a significant cost reduction (and, of course, considerable saving in computer time) if the *area* under consideration were partitioned, and these sub-problems treated independently. It would be worthwhile to obtain conditions on the validity of this "geographical decomposition."

### 5.2. Optimal Rescheduling of Airplanes

We next turn attention to the possibility of adjusting a previously-scheduled flow, when stochastic disturbances warrant such adjustment. Ellis and Rishel [11] have formulated a model for the one-way flow of air traffic between two airports, subject to random constraints on the takeoff and landing capacities. Their presentation was based upon the state-space approach of (stochastic) optimal control theory, but the specific example which they solved employed a dynamic programming algorithm.

Ellis and Rishel considered $(n + 1)$ discrete time intervals, by the end of which all scheduled takeoffs from airport 1 will have actually taken off and landed at airport 2. Once takeoff occurs, (a deterministic) $j$ intervals of time are required to fly from 1 to 2. During any time interval, the number of aircraft movements which a controller can permit is limited by runway or airside capacity. Each capacity was treated as a random variable to model delays due to weather or mechanical failure.

The objective function is the minimization of total waiting time, with time in the air valued higher by a factor $\alpha > 1$, compared to waiting time on the ground. The control variables are the numbers of landings and takeoffs permitted in each time interval. The state variables are the numbers of planes currently in the air; aircraft awaiting takeoff; those awaiting landing; and the current capacities of the respective airports.

To ensure with probability one that by time $n$, the schedule of movements begun at time 0 will have been completed, Ellis and Rishel assume that there exist positive upper and lower bounds on the capacities at each airport. More importantly, the completion constraints at time $n$ were shown to imply constraints on the values of variables at all intermediate times. These implied constraints define bounded convex sets giving the feasible values of the state variables. Ellis and Rishel use this convexity to prove a theorem which shows the existence of an optimal control, and which recursively exhibits an optimal solution to the dynamic programming problem. They show the optimal controls for a 5-period numerical example which, although a fairly simple problem, involved considerable computation.

The remainder of the present paper deals with the continuous-time dynamic transportation problem with delays. In particular, the following section deals with the minimum-time problems, as well as minimum cost problems with delays. Section 7 concludes with a discussion of some important extensions of the dynamic transportation problem with delays.

## 6. DYNAMIC TRANSPORTATION PROBLEM WITH TIME DELAYS

An important class of extensions of the classical transportation problem recognizes the fact that it takes time for a shipment to reach the sink after it has left the source. An appropriate framework to deal with such problems is that of optimal control theory. Frank [17] was first to recognize this as he attempted to examine a communication network problem with finite delays in the discrete-time state-space framework.*

### 6.1. Frank [17] and Frank and El-Bardai [18]

To describe Frank's formulation, we let $G$ denote a directed graph with each arc or branch $b_i$ of $G$ having an integer (or a rational) branch delay $\delta_i$. We can assume $\delta_i = 1$, for all $i$, without any loss of generality.†

Let $z_i(t)$ represent the flow in $b_i$ at time $t$ and let $z(t) = (z_1(t), z_2(t), \ldots, z_m(t))$, where $m$ denotes the number of branches, be called the flow or state vector. Further, let $u_j(t)$ be the external input and $y_j(t)$ be the output at node $j$. With $u(t) = (u_1(t), u_2(t), \ldots, u_n(t))$ and $y(t) = (y_1(t), y_2(t), \ldots, y_n(t))$, where $n$ is the number of nodes, we can write the state equations as:

$$z(t+1) = f[z(t), u(t)], \tag{1}$$

(Equation (1) continues)

*Note that the first application of optimal control theory to a static transportation problem is due to Fan and Wang [12]. They applied a discrete maximum principle to solve a nonlinear transportation problem (of Bellman and Dreyfus [1]) with two sources with unlimited supplies available at increasing (concave) cost, and multiple sinks with specified demands. Charnes and Kortanek [6] pointed out an inconsistency in the Fan-Wang procedure. The reader is referred to Halkin [22] for a precise statement of the discrete maximum principle.

†If not, we can construct a new graph $G'$ by replacing the $i$ th branch by $\delta_i/\delta$ series branches with common time delay $\delta$, where $\delta$ is the greatest common divisor of all $\delta_i$. Note that $\delta$, in turn, can be considered a unit of time.

and

$$y(t+1) = \hat{f}[z(t),\ u(t)]. \tag{1}$$

With this model, Frank relates the idea of satisfying demands at the nodes of the network to *state reachability*. He restricts the system strategy $(f,\ \hat{f})$ to be conservative linear. Note that $(f,\ \hat{f})$ is *conservative* if at each node $j$, total inflows equal total outflows, and $(f,\ \hat{f})$ is *linear* if $f = Az(t) + Bu(t)$ and $\hat{f} = Cz(t) + Du(t)$.

Frank shows that a given final state $z(T)$ is reachable from the zero state in time $T$ if and only if there exists a state $\sigma \geqslant z(T)$ in the subspace spanned by the columns of the *controllability matrix* $Q$ given by

$$Q = [B, AB, A^2B, \ldots, A^{T-1}B]\ . \tag{2}$$

Furthermore, $z(T)$ is reachable from the zero state (in a finite time) if and only if it is reachable from the zero state in time $r$, where $r$ is the degree of the minimal polynomial of $A$. Frank goes on to examine the necessary and sufficient conditions for the existence of a linear strategy under which a given set of terminal states are reachable (from the zero state).

Finally, Frank oulines a procedure to find the set of reachable states for a linear system with upperbounds on input $u(t)$, i.e., $u(t) \leqslant U$, $\forall t$. Note that it is no longer possible to set a uniform upper limit $r$ on the number of time periods necessary to reach a given state.

Frank and El-Bardai [18] also impose upperbounds on the flows in branches of the network. For this. let $c = (c_1,\ c_2,\ \ldots, c_m)$ denote the branch-capacity vector. We can now express the linear strategy with capacity constraint as:

$$z(t+1) = \min[Az(t) + Bu(t), c]$$

and (3)

$$y(t+1) = Cz(t) + Du(t).$$

Frank and El-Bardai show that with the input constraint $u(t) \leqslant U$, the set of reachable states (in time $T$) in system (3) is given by

$$\left\{d \mid d \leqslant \sum_{t=1}^{T} d(t)\right\}$$

where

$$d(1) = \min[BU, c]$$

and

$$d(t+1) = \min\left[Ad(t), c - \sum_{\tau=1}^{t} d(\tau)\right], \quad t = 1, 2, \ldots, (T-1).*$$

*Frank and El-Bardai also derive a similar result for systems with *memory* (but without input constraints). That means, real number $w_j \geqslant 0$ can be associated with node $j$, representing the maximum amount of flow that can be stored at node $j$. Note that any excess flow to be stored at node $j$ will be lost. Obviously, if a state is reachable in a memoryless system (i.e., one in which no flow can be stored), it is reachable in a system with memory. The converse is not necessarily true.

So far, we have dealt with the question of state reachability which is the same as the question of *feasibility* in operations research. The answer involved applying a large enough input in every period to saturate all branches. It may be that a given demand vector can be satisfied in time $T$ without saturating all branches. This raises the question of *optimality*. The idea is to find the 'smallest' inputs with which a demand vector may be satisfied.

The problem of finding such an input sequence can be easily formulated as a linear programming problem. That is, given $z(0)$ and parameters $A$, $B$, $z_f$, $C$ and $U$, find $u(t)$ and $z(t)$ that minimizes

$$\sum_{t=0}^{T-1} \sum_{j=1}^{n} u_j(t),$$

subject to

$$\begin{aligned} z(t+1) &\leqslant Az(t) + Bu(t) \quad t = 0,1,2,\ldots,(T-1) \\ z(t) &\geqslant z_f \\ 0 \leqslant z(t) &\leqslant c, \qquad\qquad t = 1,2,\ldots,T \\ 0 \leqslant u_j(t) &\leqslant U, \; j = 1,2,\ldots,n, \; t = 0,1,2,\ldots,(T-1)\ . \end{aligned}$$

The solution of this linear programming problem will yield a sequence of inputs that takes the state $z(0)$ of $G$ to a state $z(T) \geqslant z_f$ in a way that is optimal in *the input sense*. An important property of this sequence is that it is also optimal in *the loss sense*, i.e., it minimizes loss in the network.

It is the idea of optimality with which we are really concerned in this survey. Tapiero and Soliman [51] address themselves to this important problem of optimality in dynamic transportation networks with transport delays. They use the continuous-time state-space framework and the maximum principle for their analysis.*

### 6.2. Tapiero and Soliman [51] and Tapiero [49]

Tapiero and Soliman formulate a dynamic multi-commodity transportation problem with time delays as an optimal control problem. For simplicity in exposition, we will only develop the single-commodity case formulated in Tapiero [49] since its extension to the multi-commodities case is rather straightforward; all our discussion addressed to the single-commodity case is applicable to the multi-commodity version.

#### 6.2.1. The Model

To develop the model, let there be $m$ sources and $n$ destinations. Let $i$ denote source $i$ and $j$ denote sink $j$. Let $t$ denote the time. We can now define the following variables and parameters:

$u_{ij}(t)$ = rate of flow from source $i$ enroute to $j$ at time $t$,

$x_i(t)$ = inventory at source $i$ at time $t$,

*While we survey only his latter papers, Tapiero's 1971 paper is the first paper applying the maximum principle to the dynamic transportation problem. The paper considers the problem without delays. We also note that more recently Segall and Moss [37] have dealt with a similar problem for a communication network.

$y_j(t)$ = inventory at sink $j$ at time $t$,

$\tau_{ij}$ = transit time from $i$ to $j$,

$c_{ij}$ = maximum flow capacity of route $(i,j)$,

$K$ = total flow capacity of the transport system,

$T$ = planning horizon,

$A_i$ = initial supply at source $i$; i.e., $x_i(0) = A_i$,

$B_j$ = demand at sink $j$ at time $T$; i.e., $y_j(T) = B_j$,

$L\{x_i(t),\ y_j(t),\ u_{ij}(t),\ t\}$ = cost function expressing inventory and transportation costs.

Note that it is necessary to have the total available supplies exceed or equal the total available demand. We may assume (without any loss of generality)

(4)
$$\sum_{i=1}^{m} A_i = \sum_{j=1}^{n} B_j \ .$$

We can now state the optimal control problem as follows:

(5)
$$\min\left\{J = \int_0^T L\{x_i(t),\ y_j(t),\ u_{ij}(t),\ t\}\ dt\right\},$$

subject to

(6)
$$\dot{x}_i(t) = \sum_{j=1}^{n} u_{ij}(t),\ x_i(0) = A_i,\ x_i(T) = 0,*$$

(7)
$$\dot{y}_j(t) = \sum_{i=1}^{m} u_{ij}(t - \tau_{ij}),\ y_j(s) = 0,\ \forall s \leqslant 0;\ y_j(T) = B_j,$$

and the capacity constraints

(8)
$$0 \leqslant u_{ij}(t) \leqslant c_{ij}$$

and

(9)
$$0 \leqslant \sum_{i=1}^{m} \sum_{j=1}^{n} u_{ij}(t) \leqslant K.$$

*Note that the terminal conditions of $x_i$ can be rewritten as $x_i(T - \min_j \tau_{ij}) = 0$. We note that Tapiero's formulation has misprints in the specification of terminal conditions of $x_i$ and $y_j$. In passing, we may also remark that there are other errors in Tapiero [49] and Tapiero and Soliman [51].

Note that $x_i$ and $y_j$ are the state variables and $u_{ij}$ are the control variables in the terminology of optimal control theory. We refer to (9) as a linking constraint since it introduces interdependencies among flows on various routes.

It is convenient to transform this formulation into an equivalent formulation by defining a new variable $\eta_{ij}(t)$ denoting the cumulative flow (up to time $t$) which has left source $i$ for sink $j$. The equivalent problem is:

(10) $$\min \left\{ J = \int_0^T \Lambda \{\eta_{ij}(t),\ \eta_{ij}(t-\tau_{ij}),\ u_{ij}(t),\ t\} dt \right\},^*$$

subject to

(11) $$\dot{\eta}_{ij}(t) = u_{ij}(t),\ \eta_{ij}(s) = 0,\ \forall s \in [-\tau_{ij}, 0],$$

with constraints on terminal conditions

(12) $$\sum_{j=1}^{n} \eta_{ij}(T) = \sum_{j=1}^{n} \eta_{ij}(T-\tau_{ij}) = A_i,$$

(13) $$\sum_{i=1}^{m} \eta_{ij}(T) = \sum_{i=1}^{m} \eta_{ij}(T-\tau_{ij}) = B_j,$$

and the control constraints (8) and (9).

### 6.2.2. Application of the Maximum Principle

We are now ready to apply the maximum principle with delays in the state variables (see Kharatishvili [30]; see also Ray and Soliman [36]). The Hamiltonian is:

(14) $$H = -\Lambda\{\eta_{ij}(t), \eta_{ij}(t-\tau_{ij}),\ u_{ij}(t), t\} + \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij}(t)\ u_{ij}(t)$$

where the adjoint variables $\lambda_{ij}(t)$ satisfy

(15) $$\dot{\lambda}_{ij}(t) = \frac{\partial \Lambda\{t\}}{\partial \eta_{ij}(t)} + \frac{\partial \Lambda\{t+\tau_{ij}\}}{\partial \eta_{ij}(t)}$$

with the transversality conditions (see Tapiero and Soliman [51] and Tapiero [49])

(16) $$\lambda_{ij}(T-\tau_{ij}) + \lambda_{mn}(T-\tau_{mn}) = \lambda_{mj}(T-\tau_{mj}) + \lambda_{in}(T-\tau_{in}),\ i = 1, 2, \ldots, (m-1),$$
$$j = 1, 2, \ldots, (n-1).$$

It is noted that the adjoint variables in optimal control play a role similar to the Lagrange multipliers in nonlinear programming. For explanation of the transversality conditions, see Fleming and Rishel [13], Tapiero [50], and Sethi and Thompson [39].

---

*Note that $x_i(t) = A_i - \sum_{j=1}^{n} \eta_{ij}(t)$ and $y_j(t) = \sum_{i=1}^{m} \eta_{ij}(t-\tau_{ij})$ and, therefore, the new loss function (denoted in abbreviated form by $\Lambda\{t\}$) is obtained as $\Lambda\{t\} = L\left\{A_i - \sum_{j=1}^{n} \eta_{ij}(t),\ \sum_{i=1}^{m} \eta_{ij}(t-\tau_{ij}),\ u_{ij}(t),\ t\right\}$.

A necessary condition for optimal transportation is that the Hamiltonian be a maximum (subject to control constraints (8) and (9)) along the optimal path. This is, in general, a difficult problem to solve. However, it may be possible to resolve the problem in some special cases. Tapiero takes up the minimum time problem as the special case for consideration.

### 6.2.3. Minimum-Time Problem

The minimum-time transportation problem is concerned with determining a transport schedule which will transport supplies at $m$ sources to meet the demands at $n$ sinks in minimum time. This problem, especially its multi-commodity version, is extremely important in army logistics, in supplying help to disaster areas and, possibly, in transporting perishable goods. For the case of discrete time, we have discussed this problem in Section 3.

For minimum-time problems the objective function (10) is the total time $T$, thus

(17) $$L = \Lambda = 1.$$

That means $\dot{\lambda}_{ij} = 0$ in (15), implying that the $\lambda_{ij}$ are $mn$ (unknown) constants satisfying $(m-1)(n-1)$ linear equations (16). With this information, we can rewrite the Hamiltonian as: where

(18) $$H = -1 + \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij}\, u_{ij}(t).$$

Since it is easier to deal with the minimum-time problem *without* the transport system's capacity constraint (9), we will take this up next.

### 6.2.4. Minimum-time Problem Without (9)

Since the Hamiltonian (18) is linear in $u_{ij}$ and since the only constraints we need to worry about are the route capacity constraints (8), which are not of the linking-type, it is easy to see that

(19) $$\lambda_{ij} < 0 \rightarrow u_{ij}(t) = 0\,, \qquad \forall t,$$

$$\lambda_{ij} > 0 \rightarrow \begin{cases} u_{ij}(t) = c_{ij}, & t \leqslant t_{ij} \\ 0\,, & \text{otherwise,} \end{cases}$$

where

(20) $$t_{ij} \triangleq \min\left[ A_i \Big/ \sum_{\{l|\lambda_{il}>0\}} c_{il},\ B_j \Big/ \sum_{\{k|\lambda_{kj}>0\}} c_{kj} \right].$$

This means that along any arc $(i,j)$, either there is no shipment or there is a shipment at full capacity from time zero to some time $t_{ij} \leqslant T$. This observation allows us to set up a linear program which will solve the minimum-time problem without (9):

$$\text{Min } T$$

subject to

$$t_{ij} \leqslant T - \tau_{ij}, \quad \forall i,j,$$

$$\sum_{j=1}^{n} c_{ij}\, t_{ij} = A_i, \quad \forall i,$$

(21)

$$\sum_{i=1}^{m} c_{ij}\, t_{ij} = B_j, \quad \forall j,$$

$$t_{ij} \geqslant 0, \quad \forall i,j.$$

This L.P. problem will give the optimal $t_{ij}$ and $T$ provided there is a feasible solution to the given problem. We remark that in the optimal solution, $T = \min_{i,j} (t_{ij} + \tau_{ij})$.

Note that it may not be easy or even possible to modify this program to take into account (9) and still retain an L.P. problem. The existence of other ways to find an optimal solution with (9) is our next concern.

### 6.2.5. Minimum-Time Problem with (9)

If the linking constraint (9) is binding, there will be some kind of priority assignment in the set of routes $(i,j)$. Obviously, the Hamiltonian-maximizing condition will imply that this priority assignment be reflected in the values of $\lambda_{ij}$.

Suppose that we have a given set of values for $\lambda_{ij}$ which satisfy (16). There are many possible sets that may do this. Note that $\lambda_{ij} \equiv 1$, $\forall i,j$ always satisifies (16). Given the set of values for $\lambda_{ij}$, the Hamiltonian maximization is a knapsack problem, in fact a parametric knapsack problem. A tentative outline of this procedure is suggested below; see Sethi and Bookbinder [38] for more details.

We begin by listing positive $\lambda_{ij}$ in decreasing order. Starting with the route associated with the largest $\lambda_{ij}$, we go down the list by shipping the maximum flow $c_{ij}$ in the associated route $(i,j)$ until all routes with positive $\lambda_{ij}$ are full or the sum of the flows equals or just exceeds the total system capacity $K$. If it exceeds $K$, we cut back the flow in the last route just enough so that the sum of the flows equals $K$.

The flows defined above commence at time $t = 0$ and continue until the first $t_{ij}$ defined in (20) is encountered. Let this be $t_{i_1 j_1}$. This $t_{i_1 j_1}$ can be of one of two types depending upon whether the first or the second argument in (20) is smaller.

a) If $t_{i_1 j_1}$ is of the first type, set all $u_{i_1,j} = 0$, $\forall t > t_{i_1 j_1}$.

b) If $t_{i_1 j_1}$ is of the second type, set all $u_{ij_1} = 0$, $\forall t > t_{i_1 j_1}$.

We then go back to assigning the flows according to the knapsack procedure by making a new list.

These iterations continue until no further flow can be assigned. At this point, there are two possible situations. Either the terminal conditions (12) and (13) (equivalent to terminal

conditions in (6) and (7)) are met in which case we term the set of $\lambda_{ij}$ a feasible set, or they are not met, in which case the set of $\lambda_{ij}$ is clearly infeasible.

If infeasible, we have to get another set of $\lambda_{ij}$ satisfying (16) and start over again. If feasible, we have obtained a local minimum. It may not be possible to obtain the global minimum since this may require obtaining all the local minima, of which there could be a large number.

To conclude this section, we note that it would be more realistic to replace the total system constraint (9) by loading and unloading constraints,

(22)
$$0 \leqslant \sum_{j=1}^{m} u_{ij}(t) \leqslant b_i, \quad \forall t,$$

and

(23)
$$0 \leqslant \sum_{i=1}^{m} u_{ij}(t) \leqslant c_j, \quad \forall t.$$

The algorithm suggested in this section can be easily modified to deal with (22) and (23).

### 6.2.6. The Multicommodity Case

Tapiero and Soliman [51] have treated the multicommodity version of the minimum-time problem. Their paper does not contain a proof of their algorithm. It should, however, be noted that in the multicommodity case, an essential feature of the problem is the linking constraint. To state the linking constraint (the analogue of (9)), let there be $r$ commodities and let subscript $k$ denote the $k^{\text{th}}$ commodity. The linking constraint can be stated as

(24)
$$0 \leqslant \sum_{k=1}^{r} \alpha_k u_{ijk}(t) \leqslant C_{ij}, \quad \forall t,$$

where $\alpha_k$ is the capacity required per unit of commodity $k$. These constraints are analogous to (9) in so far as their effect in developing an algorithm is concerned.

We offer the following remarks concerning (24). The constraint states that the total shipment leaving source $i$ at time $t$ enroute to sink $j$ should not exceed $C_{ij}$. However, the loads-in-transit at any point along arc $(i,j)$ need not be less than $C_{ij}$. An example will further clarify this issue. Suppose there is a bridge along arc $(i,j)$ which is a bottleneck and whose capacity is $C_{ij}$. Imposition of constraint (24) does not guarantee that at some time later than $t$, there will not be a load on the bridge in excess of $C_{ij}$. Different commodities $k$ may require different times $\tau_{ijk}$ to travel from source $i$ to sink $j$, and it is possible for them to leave the source $i$ satisfying (24) and yet exceed $C_{ij}$ on the bridge. Tapiero and Soliman have not recognized this in stating the arc-capacity constraints (24). Once again, constraints (24) are more like loading constraints associated with arc $(i,j)$ rather than the arc-capacity constraints. Note also that the arc-capacity constraints, such as in the case of the bridge situation above, will be extremely difficult to handle.

### 6.2.7. Linear Inventory and Transportation Costs

Tapiero and Soliman [51] have also considered the problem wherein the inventory and transportation costs are assumed linear. In this case, the loss function

$$L = \sum_{i=1}^{m} a_i x_i + \sum_{j=1}^{n} d_j y_j + \sum_{j=1}^{n} \sum_{i=1}^{m} q_{ij}\, u_{ij}, \tag{25}$$

where $a_i$ and $d_j$ are the costs per unit time of holding a unit of inventory at source $i$ and at sink $j$, respectively, and $q_{ij}$ is the cost per unit time of transporting the commodity at a unit rate. In terms of $\eta_{ij}$, we can write the loss function

$$\begin{aligned}\Lambda\{t\} = &\sum_{i=1}^{m} a_i[A_i - \sum_{j=1}^{n} \eta_{ij}(t)] + \sum_{j=1}^{n} \sum_{i=1}^{m} d_j \eta_{ij}(t - \tau_{ij}) \\ &+ \sum_{j=1}^{n} \sum_{i=1}^{m} q_{ij}\, u_{ij}(t).\end{aligned} \tag{26}$$

The Hamiltonian can be written as

$$\begin{aligned}H = &- \sum_{i=1}^{m} a_i[A_i - \sum_{j=1}^{n} \eta_{ij}(t)] - \sum_{j=1}^{n} \sum_{i=1}^{m} d_j \eta_{ij}(t - \tau_{ij}) \\ &+ \sum_{j=1}^{n} \sum_{i=1}^{m} [-q_{ij} + \lambda_{ij}(t)] u_{ij}(t),\end{aligned} \tag{27}$$

where the adjoint variables $\lambda_{ij}(t)$ can be expressed as

$$\lambda_{ij}(t) = \lambda_{ij}(T - \tau_{ij}) + (a_i - d_j)(T - \tau_{ij} - t), \tag{28}$$

with $\lambda_{ij}(T - \tau_{ij})$ satisfying (16). Furthermore, it is obvious that

$$\begin{aligned}&a_i > d_j \rightarrow \lambda_{ij}(t) \text{ decreases with } t, \\ &a_i < d_j \rightarrow \lambda_{ij}(t) \text{ increases with } t.\end{aligned} \tag{29}$$

But the Hamiltonian (27) is linear in $u_{ij}$, implying that the optimal control is bang-bang. From (29), we can conclude that for the problem without (9), the form of the optimal policy for each are $(i,j)$ can be characterized by one number $t_{ij}$ if $a_i \geqslant d_j$ and $\bar{t}_{ij}$ if $a_i < d_j$:

$$\begin{aligned}a_i > d_j \rightarrow u_{ij}(t) &= \begin{cases} c_{ij}, & 0 \leqslant t \leqslant t_{ij} \\ 0, & \text{otherwise,} \end{cases} \\ a_i < d_j \rightarrow u_{ij}(t) &= \begin{cases} 0, & 0 \leqslant t \leqslant \bar{t}_{ij} \\ c_{ij}, & \bar{t}_{ij} < t \leqslant T - \tau_{ij}. \end{cases}\end{aligned} \tag{30}$$

This policy is consistent with our intuition. It states that if the inventory cost at source $i$ is higher than that at sink $j$, then shipment along arc $(i,j)$, if any, must commence as early as possible. On the other hand, if $a_i < d_j$ then delay the shipment along arc $(i,j)$, if any, as long as possible.

Having characterized the form of the optimal policy, it is possible to formulate a quadratic program to solve the linear-costs transportation problem without (9). This is done in a manner similar to that in Section 6.2.4. We note that Tapiero and Soliman [51] had earlier made this observation. For the actual formulation of the quadratic program, see Sethi and Bookbinder [38]. Finally, Sethi and Bookbinder also attempt to obtain an algorithm for the problem with (9). This is similar but far more difficult than the procedure in Section 6.2.5.

### 6.2.8. Linear Inventory and Quadratic Transportation Costs

Linear-quadratic control problems have a special place in the optimal control theory. Usually these problems yield closed-form solutions. Midler [34] treated a transportation problem with linear transportation costs and quadratic inventory costs in the stochastic dynamic programming framework. Tapiero and Soliman [51] treat a similar problem with delays in the deterministic framework, but with linear inventory costs and quadratic transportation costs. To state the latter problem, we only need to replace the $q_{ij}\ u_{ij}$ term in (25) by $q_{ij}(u_{ij} - \hat{u}_{ij})^2$, where $\hat{u}_{ij}$ is the most desirable shipment rate from the point-of-view of transportation. *Furthermore, constraints (8) and (9) are assumed to be no longer acting for this problem.*

The Hamiltonian of this problem can be written as

$$H = -\sum_{i=1}^{m} a_i\Big[A_i - \sum_{j=1}^{n} \eta_{ij}(t)\Big] - \sum_{j=1}^{n}\sum_{i=1}^{m} d_j \eta_{ij}(t - \tau_{ij}) - \sum_{j=1}^{n}\sum_{i=1}^{m} q_{ij}(u_{ij}(t) - \hat{u}_{ij})^2 + \sum_{j=1}^{n}\sum_{i=1}^{m} \lambda_{ij} u_{ij}(t), \tag{31}$$

where the adjoint variable $\lambda_{ij}(t)$ satisfies (28) and (16). Furthermore, the Hamiltonian maximizing condition yields

$$u^*_{ij}(t) = \hat{u}_{ij} + \lambda_{ij}(t)/2q_{ij}, \tag{32}$$

where $u^*_{ij}(t)$ is the optimal control if $u^*_{ij}(t) \geqslant 0$, $\forall t$. We will assume this to be the case; otherwise the problem becomes more difficult.

We can now use (32) in (11) to obtain $\eta^*_{ij}(T - \tau_{ij})$ which must satisfy (12) and (13). These conditions can be simplified as:

$$\sum_{j=1}^{n} \frac{T - \tau_{ij}}{2q_{ij}} [\lambda_{ij}(T - \tau_{ij}) + (a_i - d_j)(T - \tau_{ij})/2] = A_i, \quad i = 1, 2, \ldots, m, \tag{33}$$

$$\sum_{i=1}^{m} \frac{T - \tau_{ij}}{2q_{ij}} [\lambda_{ij}(T - \tau_{ij}) + (a_i - d_j)(T - \tau_{ij})/2] = B_j, \quad j = 1, 2, \ldots, (n-1). \tag{34}$$

We have not written (34) for $j = n$ since it is redundant on account of condition (4), which states that total demand equals total supply.

The system of Equations (33) and (34) contains $m + n - 1$ equations. Along with (16) which contains $(m - 1)\ (n - 1)$ equations, we have a total of $mn$ equations which must be satisfied by $mn$ variables $\lambda_{ij}(T - \tau_{ij})$. The solution of this linear system of equations should provide the optimal control when substituted in (32).

## 7. SUGGESTED IMPORTANT EXTENSIONS

A natural extension of the dynamic transportation problem with delays involves time-dependent demands at various sinks. If we let $\xi_j(t)$ represent the demand rate at sink $j$, then (7) becomes

(35) $$\dot{y}_j = \sum_{i=1}^{n} u_{ij}(t - \tau_{ij}) - \xi_j(t),$$

with an additional constraint that

(36) $$y_j(t) \geqslant 0, \quad \forall t.$$

For this extension, it may be necessary to either assume $y_j(0) > 0$ or $\xi_j(t) = 0$ for $0 \leqslant t \leqslant \min_i \tau_{ij}$ to have feasibility. Another way to handle this problem is to allow shortages and incorporate shortage costs in the objective function. See also Midler [34] in this connection.

A simple and interesting extension is the case of perishable goods. If we assume a constant spoilage rate $\gamma$ in transit, it amounts to replacing (7) by

(37) $$\dot{y}(t) = \sum_{i=1}^{n} u_{ij}(t-\tau_{ij}) e^{-\gamma \tau_{ij}} .$$

In this case we cannot require condition (4), since that would imply infeasibility. Consequently, feasibility also becomes an issue in this case. Of course, the cost of spoilage must be incorporated in the objective function by modifying unit transportation costs on the arcs of the transportation network.

Another extension is the case in which there are unloading delays at various sinks. Typically, the unloading delay at a sink will depend on the inventory at the sink and (possibly) the time of unloading. The latter dependence may derive from an expansion program which is in progress at the sink under consideration. For the single commodity case, this amounts to replacing (7) by

(38) $$\dot{y}_j(t) = \sum_{i=1}^{m} u_{ij}(t - \tau_j(y_j(t),t))$$

where the delay $\tau_j$ may be assumed to satisfy $\partial \tau_j / \partial y_j \geqslant 0$ and $\partial \tau_j / \partial t \leqslant 0$. Ray and Soliman [36] have a weak maximum principle for dealing with systems subject to equations of the form (38).

Perhaps the most important extension is the case in which the time required for a shipment to reach from source $i$ to sink $j$ is a function of the amount shipped. That is

(39) $$\tau_{ij} = \tau_{ij}(u), \quad d\tau_{ij}/du_{ij} \geqslant 0.$$

In this case the transit time increases with the amount shipped. For the single commodity case, this amounts to replacing (7) by

(40) $$\dot{y}_j(t) = \sum_{i=1}^{m} \sum_{\{\theta \mid \theta + \tau_{ij}[u_{ij}(\theta)] = t\}} u_{ij}(\theta).$$

We note that this equation is an essential feature of any dynamic transportation problem in which the transit time is a function of the load-in-transit.

Equation (40) is a state equation in which not only the delays depend on controls at previous times but also these delays are defined recursively.* We do not know of any maximum principle allowing even for delays depending upon controls at an earlier time, let alone where that time is defined recursively. Ray and Soliman's [36] paper contains only a weak maximum principle for systems with delays depending on *state and control at time t* and time *t*.

Treatment of state equations of the type (40) will obviously require some theoretical developments, including a suitable maximum principle. For the time being, we are working with a discrete-time version of (40), employing dynamic programming and/or some theory for optimization systems developed by Clarke [8].

Finally, the problem could be extended to take into account the stochastic nature of transit delays. This would require replacing (7) by a stochastic differential equation. One could also consider demands at sinks to be stochastic (see Midler [34] and Segall and Moss [37]). Either case results in difficult stochastic optimal control problems.†

## 8. ACKNOWLEDGMENT

This paper is dedicated to the memory of Ray Fulkerson. This research is supported by Grant 3-214-385-80 from the University of Toronto/York University Joint Program in Transportation. The authors are grateful to the referees for their instructive comments concerning an earlier draft of this paper.

## REFERENCES

[1] Bellman, R.E. and S.E. Dreyfus, *Applied Dynamic Programming* (Princeton University Press, Princeton, New Jersey, 1962).

[2] Bellmore, M., W.D. Eklof, and G.L. Nemhauser, "A Decomposable Transshipment Algorithm for a Multiperiod Transportation Problem," Naval Research Logistics Quarterly *16*, 517-524 (1969).

[3] Bellmore, M. and R.R. Vemuganti, "On Multicommodity Maximal Dynamic Flows," Operations Research *21*, 10-21 (1973).

[4] Bowman, E.R., "Production Scheduling by the Transportation Method of Linear Programming," Operations Research *4*, 100-103 (1956).

[5] Bradley, G.H., "Survey of Deterministic Networks," Association of Industrial Engineers Transactions *7*, No. 3, 222-234 (1975).

[6] Charnes, A. and K. Kortanek, "A Note on the Discrete Maximum Principle and Distribution Problems," Journal of Mathematics and Physics, Vol. *45*, No. 1, 121-126 (1966).

[7] Christofides, N., *Graph Theory: An Algorithmic Approach* (Academic Press Inc., New York, 1976).

[8] Clarke, F., "Necessary Conditions for a General Control Problem," in *Calculus of Variations and Control Theory* (Ed. David Russell, Academic Press, N.Y., 1976).

[9] Dantzig, G.B., "Application of the Simplex Method to a Transportation Problem," in *Activity Analysis of Production and Allocation*, 330-335 (T.C. Koopmans, ed., John Wiley & Sons, Inc., New York, 1951).

*Here the delays are in control variables. For delays in the state variables, see Manitius [33].

†See Fleming and Rishel [13] for the theory of stochastic optimal control without delays.

[10] Dantzig, G.B., "Linear Programming under Uncertainty," Management Science *1*, 301-312 (1955).

[11] Ellis, R.S. and R.W. Rishel, "An Application of Stochastic Optimal Control Theory to the Optimal Rescheduling of Airplanes," IEEE Transactions on Automatic Control, 139-142 (1974).

[12] Fan, L. and C. Wang, "The Application of the Discrete Maximum Principle to a Transportation Problem," Journal Mathematics and Physics, Vol. *43*, 255-260 (1964).

[13] Fleming, W.H. and R.W. Rishel, *Deterministic and Stochastic Optimal Control* (Springer-Verlag, New York, 1975).

[14] Fong, C.O. and V. Srinivasan, "Multiperiod Capacity Expansion and Shipment Planning with Linear Costs," Naval Research Logistics Quarterly *23*, No. 1, 37-52 (1976).

[15] Ford, L.R., "Constructing Maximal Dynamic Flows from Static Flows," Operations Research *6*, 419-433 (1958).

[16] Ford, L.R. and D.R. Fulkerson, *Flows in Networks* (Princeton University Press, Princeton, New Jersey, 1962).

[17] Frank, H., "Dynamic Communication Networks," IEEE Transactions on Communication Technology, Vol. *COM-15*, No. 2, 156-163 (1967).

[18] Frank, H. and M.T. El-Bardai, "Dynamic Communication Networks with Capacity Constraints," IEEE Transactions on Communication Technology, Vol. *COM-17*, No. 4, 432-437 (1969).

[19] Gale, D., "Transient Flows in Networks," Michigan Mathematical J. *6*, 59-63 (1959).

[20] Garfinkel, R.S. and M.R. Rao, "The Bottleneck Transportation Problem," Naval Research Logistics Quarterly, Vol. *18*, 465-472 (1971).

[21] Glickman, T.S. and P.D. Berger, "Cost-Completion-Date Tradeoffs in the Transportation Problem," Operations Research, Vol. *25*, No. 1, 163-168 (1977).

[22] Halkin, H., "A Maximum Principle of the Pontryagain Type for Systems Described by Nonlinear Difference Equations," SIAM J. Control, Vol. 4, 90-111 (1966).

[23] Hammer, P.L. "Time-Minimizing Transportation Problems," Naval Research Logistics Quarterly *16*, 345-357 (1969).

[24] Hammer, P.L., "Communication on 'The Bottleneck Transportation Problem' and 'Some Remarks on the Time Transportation Problem'," Naval Research Logistics Quarterly *18*, 487-490 (1971).

[25] Hausman, W.H. and P. Gilmour, "A Multiperiod Truck Delivery Problem," Transportation Research *1*, 349-357 (1967).

[26] Hitchcock, F.L., "The Distribution of a Product from Several Sources to Numerous Localities," Journal Mathematics and Physics, *20*, 224-230 (1941).

[27] Holladay, J., "Some Transportation Problems and Techniques for Solving Them," Naval Research Logistics Quarterly *11*, 15-42 (1964).

[28] Kantorovich, L.V., "Mathematical Methods in the Organization and Planning of Production," Publication House of the Leningrad State University (1939). Translated in Management Science *6*, 366-422 (1960).

[29] Kennington, J.L., "A Survey of Linear Cost Multicommodity Network Flows," Operations Research *26*, No. 2, 209-236 (1978).

[30] Kharatishvili, G.L., "A Maximum Principle in Extremal Problems with Delays," in *Mathematical Theory of Control* (edited by A.V. Balakrishnan and L.W. Neustadt, Academic Press, New York, 26-34 1966).

[31] Koopmans, T.C., "Optimum Utilization of the Transportation System," Vol. 5 in Proceedings of the International Statistical Conferences, Washington, D.C. (1947). Reprinted as Supplement to Econometrica, Vol. *17* (1949).

[32] Kreibel, C.H., "Warehousing and Transshipment under Seasonal Demand," Journal of Regional Science *3*, 57-71 (1961).

33] Manitius, A., "On the Optimal Control of Systems with a Delay Depending on State, Control and Time," Séminaires IRIA, Analyse et Contrôle de Systémes, IRIA, Rocquencourt, France (1975).

34] Midler, J.L. "A Stochastic Multiperiod Multimode Transportation Model," Transportation Science *3*, 8-29 (1969).

35] Potts, R.B. and R.M. Oliver, *Flows in Transportation Networks* (Academic Press, New York, 1972).

36] Ray, W.H. and M.A. Soliman, "The Optimal Control Processes Containing Pure Time Delays — I: Necessary Conditions for an Optimum," Chemical Engineering Science, Vol. *25*, 1911-1925 (1970).

37] Segall, A. and F.H. Moss, "Application of Optimal Control Theory to Dynamic Routing for Computer Networks," Proceedings of the Joint Automatic Control Conference, Purdue University, Lafayette, Indiana (July 1976).

38] Sethi, S.P. and J.H. Bookbinder, "Solution of Dynamic Transportation Problems Via Maximum Principle and Mathematical Programming," presented at the Joint ORSA/TIMS National Meeting, Atlanta, Georgia in November, 1977.

39] Sethi, S.P. and G.L. Thompson, *Applied Optimal Control and Management Science Applications*, (forthcoming, 1980).

40] Srinivasan, V. and G.L. Thompson, "Determining Optimal Growth Paths in Logistics Operations," Naval Research Logistics Quarterly *19*, No. 4, 575-599 (1972).

41] Srinivasan, V. and G.L. Thompson, "Algorithms for Minimizing Cost, Bottleneck Time and Bottleneck Shipment in Transportation Problems," Naval Research Logistics Quarterly *23*, No. 4, 567-595 (1976).

42] Srinivasan, V. and G.L. Thompson, "Determining Cost vs. Time Pareto-Optimal Frontiers in Multimodal Transportation Problems," Transportation Science *11*, No. 1, 1-19 (1977).

43] Szwarc, W., "The Transportation Problem with Stochastic Demand," Management Science *11*, 35-50 (1964).

44] Szwarc, W., "The Time Transportation Problem," Zastowania Matematyki, (Warsaw) *8*, 231-242 (1966).

45] Szwarc, W., "The Time Transportation Problem," Technical Report No. 200, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, Pennsylvania (1970).

46] Szwarc, W., "Some Remarks on the Time Transportation Problem," Naval Research Logistics Quarterly *18*, 473-485 (1971).

47] Szwarc, W., "The Dynamic Transportation Problem," Mathematica *13*, 335-345 (1971).

48] Tapiero, C.S. "Transportation-Location-Allocation Problems over Time," Journal of Regional Science, Vol. *11*, 377-384 (1971).

49] Tapiero, C.S., "Transportation over Time: A System Theory Approach," *Proceedings — XX International TIMS Meeting held in Tel-Aviv, Israel on June 24-29, 1973*, Vol. *1* (Jerusalem Academic Press, Jerusalem, 239-243, 1975).

50] Tapiero, C.S., *Managerial Planning: An Optimum and Stochastic Control Approach* (Gordon & Breach, New York, 1978).

51] Tapiero, C.S. and M.A. Soliman, "Multicommodites Transportation Schedules over Time," Networks, Vol. *2*, 311-327 (1972).

52] Williams, A.C., "A Stochastic Transportation Model," Operations Research *11*, 759-770 (1963).

# THE USE OF DYNAMIC PROGRAMMING METHODOLOGY FOR THE SOLUTION OF A CLASS OF NONLINEAR PROGRAMMING PROBLEMS

Mary W. Cooper

*Department of Operations Research*
*and*
*Engineering Management*
*Southern Methodist University*
*Dallas, Texas*

**ABSTRACT**

This paper presents an application of a method for finding the global solution to a problem in integers with a separable objective function of a very general form. This report shows that there is a relationship between an integer problem with a separable nonlinear objective function and many constraints and a series of nonlinear problems with only a single constraint, each of which can be solved sequentially using dynamic programming. The first solution to any of the individual smaller problems that satisfies the original constraints in addition, will be the optimal solution to the multiply-constrained problem.

## INTRODUCTION

Let us define a nonlinear integer programming problem using the following notation:

(1) $\text{Max } z = \sum_{j=1}^{n} f_j(x_j)$

(2) such that $h_i(x) \leqslant 0 \quad i = 1, 2, \ldots, m$

(3) $x_j \in I_p \qquad j = 1, 2, \ldots, n$

where $f_j: I_p \rightarrow R_p \quad j = 1, 2, \ldots, n,$

and $h_i: I_p \rightarrow R_p \quad i = 1, 2, \ldots, m$ and $x = (x_1, x_2, \ldots, x_n)$.

Additional assumptions are as follows:

1. The functions $f_j(\cdot)$ satisfy a necessary conditon for dynamic programming.

2. The region defined by (2) and (3) is non-empty, with at least one lattice point in the region and the region is bounded.

In [1] and [2] algorithms are developed to find candidate lattice points on a series of hypersurfaces. A dynamic programming technique is used to solve a sequence of problems

with only a single constraint. Each of these problems corresponds to a value of the objective function, and this value is reduced from an upper bound until a feasible lattice point is found. In [1] the value of the objective function is reduced by one unit for each successive single constraint problem. In [2] an argument is made for allowing reduction of the trial value by more than a unit amount using numbers corresponding to hyperplanes with integer solutions in the dynamic programming return function tables. Therefore, only hypersurfaces which contain lattice points are investigated and the speed of the algorithm is greatly improved. The effort contained in this paper is an explication of the algorithm developed in [2] and its application to problems with a different type of objective function: namely, those for which the separable functions are such that

$$f_j(\cdot): I_p \rightarrow R_p \qquad j = 1, 2, \ldots n.$$

(It was previously assumed that $f_j(\cdot)$: $(I_p \rightarrow I_p, j = 1, 2, \ldots, n)$. Such problems would typically have an objective function like $z = x_1^{1/2} + 3x_2^{1/4} + 4x_3^2$. It will be shown that the algorithm of [2] can be used on problems where the objective function is not restricted to integers. The algorithm will find a global optimum, even for these less restricted problems. Therefore, this method can be used for nonlinear problems such as the concave transportation problem or for discrete problems in which the objective function is neither concave nor convex.

**Summary of the Algorithm**

Let us formulate the following approach for solving the problem described in (1)-(3). In this development, we will consider a sequence of hypersurfaces of the following form:

$$\sum_{j=1}^{n} f_j(x_j) = z_k, \qquad k = 0, 1, \ldots \tag{4}$$

If we can find an upper bound $z_0$ on the optimal solution, then we search the hypersurfaces (4) corresponding to a sequence of values of $z_k$ for lattice points. We choose $z_0$ so that $z_0$ corresponds to an upper bound on the value of the objective function, and every successive value of $z_k$ is such that $z_k < z_0$, and the hyperplane (4) corresponding to $z_k$ contains an integer point. We will show how this is done in the following sections. The search for lattice points on these hyperplanes is performed by a dynamic programming algorithm. Any lattice points so found are only candidates for the optimal solution to (1)-(3). They must be tested finally for feasibility with the constraints given in (2)-(3). Using this method of examining all lattice points which correspond to some hyperplane (4), then testing for feasibility with the original constraints (2)-(3), and terminating as soon as a feasible lattice point is found is the sequence of steps of the algorithm given in [1] and improved in [2]. The method in both references looks at all the lattice points corresponding to one hypersurface at a time and checks any points found for feasibility. The improvement given in [2] is that the sequence of values $z_k$ can be calculated so that *only* those hypersurfaces are examined that are known to contain lattice points. Since the sequence of values $z_k$, $k=0, 1, \ldots$, is strictly decreasing, and if we have examined all possible surfaces that contain lattice points and check these points for feasibility with the original constraints (2)-(3), then we can guarantee that the first feasible lattice point will be optimal. In the special case in which some of the constraints are separable it is possible to discard or fathom partial solutions thereby improving the speed of the algorithm. This is done by calculating

$$s_i = b_i - \sum_{j=n-l}^{n} a_{ij} x_j \tag{5}$$

for each partial solution $x_n, x_{n-1}, \ldots, x_l$. If $s_i < 0$ for any $i$, $i = 1, \ldots, m$, the partial solution will only lead to infeasible solutions and can be dropped.

This procedure is able to use several advantageous properties of dynamic programming. First we may find global optima for nonlinear functions and identify integer points efficiently. Second, it is possible to extend the algorithm to handle even nonseparable constraints (as shown in the numerical example (10)). This is done by dropping the feasibility test for constraints that are nonseparable. So the calculations of $s_i$ is made only for constraints that are separable. If all constraints are nonseparable, then feasibility is checked after integer points are identified and there is no possibility of fathoming at the partial solution stage. So it is possible to handle highly nonlinear nonseparable constraints. Hence the emphasis in the development of this algorithm has been to utilize the advantages of a dynamic programming approach while avoiding, as much as possible, the "curse of dimensionality" which has doomed such efforts in the past.

We now present a stepwise description of the algorithm.

## MODIFIED HYPERSURFACE SEARCH ALGORITHM

. Determine upper bounds $u_j$ for each variable. We then have:

$$0 \leqslant x_j \leqslant u_j \qquad \text{for } j = 1, 2, \ldots, n$$

$$x_j \in I_p$$

. Compute $z_0$ if one is not known *a priori* by

$$Z_0 = \sum_{j=1}^{n} f_j(u_j) = z_k$$

. Find all combinations of $x_j$, $j = 1, 2, \ldots, n$ which satisfy:

$$\sum_{j=1}^{n} f_j(x_j) = z_k \qquad k = 0, 1, \ldots$$

$$0 \leqslant x_j \leqslant u_j \qquad j = 1, 2, \ldots, n$$

$$x_j \in I_p \qquad j = 1, 2, \ldots, n.$$

. Test any integer point found on the hypersurface of step 3 at the partial solution stage (for separable constraints) for feasibility with the original constraints (2)-(3). If a feasible point is found, then the value of $z_k$ is tested against the greatest feasible lower bound on $z$. If it is less than this bound, then the point corresponding to that bound is optimal. Otherwise go to step 5.

. Calculate a new $z_{k+1}$ according to logic and notation developed in succeeding sections.

. Return to step 3.

The process of finding lattice points in step 3 is achieved by using a dynamic programming formulation of the following problem: Find all $x_j$, $j=1, \ldots, n$ such that:

$$\sum_{j=1}^{n} f_j(x_j) = z_k$$

6) $$0 \leqslant x_j \leqslant u_j \qquad j = 1, \ldots, n$$

$$x_j \in I_p \qquad j = 1, \ldots, n$$

In [1] and [2] an equivalent formulation of (5) is given:

$$\text{Max } z = \sum_{j=1}^{n} f_j(x_j)$$

such that

$$\sum_{j=1}^{n} f_j(x_j) \leqslant z_k$$

$$0 \leqslant x_j \leqslant u_j \quad j = 1, \ldots, n \tag{7}$$

$$x_j \in I_p \quad j = 1, \ldots, n.$$

For every value of $z_k$ this gives us the values $x_j$, $j = 1, \ldots, n$ of all lattice points $x$ on the objective function hypersurface. Therefore, this problem can be used to identify integer points $x$ at any value $z_k$ of the objective function. The optimal return functions are given by the following:

$$g_1(\lambda) = \operatorname*{Max}_{x_1=\delta_1} f_1(x_1) = \begin{cases} \lambda, & \lambda = f_1(\delta_1),\ \delta_1 = 0, 1, \ldots, u_1 \\ -\infty, & \text{otherwise} \end{cases} \tag{8}$$

$$g_s(\lambda) = \operatorname*{Max}_{0 \leqslant x_s = \delta_s} [f_s(x_s) + g_{s-1}(\lambda - f_s(x_s))] \quad \begin{matrix} s = 2, 3, \ldots, n \\ \lambda = 0, 1, \ldots, \Lambda_s \end{matrix} \tag{9}$$

where $\delta_s = \min(u_s, [\xi_s])$

$$f_s(\xi_s) = \lambda$$

$$\Lambda_s = \sum_{j=1}^{s} (u_j)$$

and the notation [b] indicates the largest integer $\leqslant$ (b).

Equations (7) and (8) give the dynamic programming recursion formulas for calculating integer points in step 3.

In step 5 a new objective function value $z_k$ is calculated. It is not necessary to consider every integer value between an upper bound on the problem $z_o$ and the optimal objective function value. In the case in which the separable functions are required to map onto the range $I_p$, this would have been a feasible strategy. However, for functions $f_j$: $I_p \rightarrow R_p$ this method would skip any noninteger $z$. Let us consider a different method of calculating the sequence $z_0$, $z_1, \ldots, z_k$ in step 5. We know that only values corresponding to integer points will be finite entries in the tabulation of the return function $g_s(\lambda)$. (Values for noninteger $x_s(\lambda)$ are associated with $g_s(\lambda) = -\infty$.) Therefore, let

$$z_k = \lambda^{(n-1)} + f_n(x_n) \quad x_n = u_n, u_n - 1, \ldots, 0,$$

where $\lambda^{(n-1)}$ are values of $\lambda$ entered in the $g_{n-1}(\cdot)$ function tabulation. This formula for calculating $z_k$ will give all hypersurface level sets which contain lattice points. The interesting fact for the current application is that this calculation will give all $z_k$ corresponding to hypersurfaces containing lattice points even for noninteger values of $\lambda^{(n-1)}$ and $f_n(x_n)$.

## A NUMERICAL EXAMPLE

Let us consider the solution of the following example:

Maximize

$$z = 6x_1^2 + 3x_2^{1/3} + 2x_3^{1/2}$$

such that

$$
\begin{aligned}
(10) \qquad & 3x_1 + 4x_2 + 3x_3 \leqslant 10 \\
& 2x_1 + 3x_2 + 3x_3 \leqslant 10 \\
& x_1 x_2^2, \leqslant 7 \\
& x_1,\ x_2,\ x_3 \geqslant 0,\ \text{integer}
\end{aligned}
$$

then

$$
\begin{aligned}
f_1(x_1) &= 6x_1^2, \\
f_2(x_2) &= 3x_2^{1/3} \\
f_3(x_3) &= 2x_3^{1/2}.
\end{aligned}
$$

from the constraints of (10) we see that:

$$
\begin{aligned}
0 \leqslant x_1 \leqslant 3 &= u_1 \\
0 \leqslant x_2 \leqslant 2 &= u_2 \\
0 \leqslant x_3 \leqslant 3 &= u_3.
\end{aligned}
$$

Therefore,

$$
g_1(\lambda) = \begin{cases} \lambda, & \lambda = 6\,\delta_1^2,\ \delta_1 = 0, 1, 2, 3 \\ -\infty, & \text{otherwise.} \end{cases}
$$

Let us tabulate the return function for the first variable

TABLE 1. *Optimal Return and Policy*

| $\lambda$ | $g_1(\lambda)$ | $x_1^*(\lambda)$ |
|---|---|---|
| 0 | 0 | 0 |
| 6 | 6 | 1 |
| 24 | 24 | 2 |
| 54 | 54 | 3 |

For the second stage the dynamic programming recursion is

$$
g_2(\lambda) = \max_{0 \leqslant x_2 \leqslant \delta_2} [3x_2^{1/3} + g_1(\lambda - 3x_2^{1/3})]
$$

for $\delta_2 = \min(2, [\xi_2])$ and $\lambda = 3\xi^{1/3}$.

We can thus compile the following table:

TABLE 2. *Optimal Return and Policy*

| $\lambda$ | $g_2(\lambda)$ | $x_2^*(\lambda)$ | $\lambda$ | $g_2(\lambda)$ | $x_2^*(\lambda)$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 24 | 24 | 0 |
| 3 | 3 | 1 | 27 | 27 | 1 |
| 3.78 | 3.78 | 2 | 27.78 | 27.78 | 2 |
| 6 | 6 | 0 | 54 | 54 | 0 |
| 9 | 9 | 1 | 57 | 57 | 1 |
| 9.78 | 9.78 | 2 | 57.78 | 57.78 | 2 |

Now we may choose the sequence of values $z_0, z_1, z_2, \ldots$ by using the following rule:

$$
z_k = \lambda^{(n-1)} + f_3(x_3),
$$

where $\lambda^{n-1}$ takes on all values entered for the state variable in Table 2, and $x_3 = u_3, u_3 - 1, \ldots, 0$. It is shown in reference [2], if it is not immediately apparent to the reader, that lattice points may lie only on hypersurfaces corresponding to

$$\sum_{j=1}^{n} f_j(x_j) = z_k = \lambda^{(n-1)} + f_n(x_n);$$

therefore,

$$z_0 = \lambda^{(2)} + f_3(u_3) = 57.78 + 2 \cdot 3^{1/2} = 61.24.$$

The corresponding integer solution is $x_3^* = 3$, $x_2^* = 2$, $x_1^* = 3$, which is not feasible for the constraints given in (9). Now we choose

$$z_1 = \lambda^{(2)} + f_3(u_3 - 1) = 57.78 + 2 \cdot 2^{1/2} = 60.61$$

corresponding to $x_1^* = 3$, $x_2^* = 2$, $x_3^* = 2$. Again the constraints given in (9) are violated. We continue the steps of the algorithm until we generate a feasible lattice point:

| | | |
|---|---|---|
| $z_2 = 57.78 + 2 = 59.78;$ | $x_1^*=3,\ x_2^*=2,\ x_3^*=1$ | (infeasible) |
| $z_3 = 57.78 + 0 = 57.78;$ | $x_1^*=3,\ x_2^*=2,\ x_3^*=0$ | (infeasible) |
| $z_4 = 57 + 3.465 = 60.46;$ | $x_1^*=3,\ x_2^*=1,\ x_3^*=3$ | (infeasible) |
| $z_5 = 57 + 2.828 = 59.828;$ | $x_1^*=3,\ x_2^*=1,\ x_3^*=2$ | (infeasible) |
| $z_6 = 57 + 2 = 59;$ | $x_1^*=3,\ x_2^*=1,\ x_3^*=1$ | (infeasible) |
| $z_7 = 57 + 0 = 57;$ | $x_1^*=3,\ x_2^*=1,\ x_3^*=0$ | (infeasible) |
| $z_8 = 54 + 3.464 = 57.464;$ | $x_1^*=3,\ x_2^*=0,\ x_3^*=3$ | (infeasible) |
| $z_9 = 54 + 2.828 = 56.828;$ | $x_1^*=3,\ x_2^*=0,\ x_3^*=2$ | (infeasible) |
| $z_{10} = 54 + 0 = 56$ | $x_1^*=3,\ x_2^*=0,\ x_3^*=1$ | (infeasible) |
| $z_{11} = 54 + 0 = 54;$ | $x_1^*=3,\ x_2^*=0,\ x_3^*=0$ | (feasible). |

Now we have a bound which is feasible for the integer problem. However, we must continue the calculations until no hypersurface with a greater $z$ is unexamined. If we check $z_{12}$, we find that

$$z_{12} = 27.78 + 3.465 < 54$$

and all subsequent calculations also have $z$ values less than this bound. Hence 54 is the optimal value of the objective function and this value corresponds to a global maximum at $x_1 = 3$, $x_2 = 0$, $x_3 = 0$.

## COMPUTATIONAL RESULTS

Results from 51 randomly generated problems are reported. The problems were of the following form:

$$\text{Max } z = \sum_{j=1}^{n} f_j(x_j)$$

$$\sum a_{ij} x_j \leqslant b_i \qquad i = 1, 2, \ldots, m \tag{11}$$

where

$$f_j(x_j) = \alpha_j x_j + \beta_j x_j^2 + \gamma_j x_j^3 \tag{12}$$

and $\alpha_j$, $\beta_j$, $\gamma_j$ were non-negative integers. Results are given in Table 3. The number of terms in the objective function for $n = 20$ could be as large as $3n = 60$. In addition, the problems were generated so that the upper bounds $u_j$ for each $x_j$ determined from the constraints (15) varied among the values 0, 1, 2, 3 so that a maximum of 4 integer values need be considered

TABLE 3

| $z_0 - z^*$ | $m \times n$ 4 × 10 | $m \times n$ 4 × 15 | $m \times n$ 4 × 20 | $m \times n$ 4 × 25 | $m \times n$ 4 × 30 |
|---|---|---|---|---|---|
| 79-38 | .92 | 1.35 | 32.38 | 16.26 | 69.86 |
| 101-71 | 1.16 | 14.83 | 45.66 | 68.61 | 52.28 |
| 129-37 | 2.38 | 13.19 | 16.45 | 26.18 | 63.77 |
| 102-35 | 2.45 | 4.84 | 25.45 | 94.65 | 229.74 |
| 75-34 | 5.28 | 5.39 | 47.01 | 57.32 | |
| 35-16 | 1.30 | 9.74 | 21.06 | 27.35 | |
| 51-30 | .49 | 26.53 | 77.69 | 35.65 | |
| 95-84 | .58 | 3.47 | 18.69 | | |
| 62-23 | 1.09 | 10.20 | 22.88 | | |
| 53-28 | 1.99 | 3.09 | 48.19 | | |
| 58-29 | .38 | 11.14 | 68.62 | | |
| 85-57 | .53 | 6.82 | | | |
| 59-36 | .49 | 9.27 | | | |
| 169-94 | 11.71 | 4.51 | | | |
| 61-28 | .76 | 7.92 | | | |
| 80-28 | 1.00 | 18.12 | | | |
| 125-34 | 4.52 | 7.64 | | | |
| 53-25 | .32 | 12.55 | | | |
| 77-23 | 1.16 | 17.07 | | | |
| 64-30 | .84 | 23.81 | | | |
| Total (sec) | 39.35 | 211.48 | 424.08 | 326.02 | 415.65 |
| Avg. (sec) | 1.97 | 10.57 | 38.55 | 37.26 | 103.91 |

or each variable. All computations were carried out on a CDC Cyber 70, Model 72, a medium peed computer. An important feature of the program when the constraints are separable is hat partial solutions may be discarded when they become infeasible by calculating the slack in ach constraint as soon as $x_j$ is determined. This procedure is not carried out in the detailed xample of the previous section because of the presence of a single nonseparable constraint. Even in this case, when some constraints are separable, partial solutions may still be discarded sing only the separable constraints.

## CONCLUSION

This paper presents an exact solution method for an extended class of problems with ighly nonlinear objective functions, $\sum_{i=j}^{n} f_j\ (x_j)$, that have the property that $f_j$: $I_P \rightarrow R_P$. Hence problems with an extended class of objective functions—namely those that have rational alues—may be solved exactly for the global optimum. The method is one example of an ffort to utilize the favorable features of dynamic programming while avoiding the storage and omputational difficulties associated with multiply-constrained problems.

## REFERENCES

1] Cooper, L., and M.W. Cooper, "Non-Linear Integer Programming," Computers and Mathematics With Applications, *1*, 215-222 (1975).

2] Cooper, M.W., "An Improved Algorithm for Non-Linear Integer Programming," Report IEOR 77005, Southern Methodist University (February 1977).

# THE U.S. VERSUS THE SOVIET INCENTIVE MODELS

Gregory G. Hildebrandt

*Department of Economics, Geography and Management*
*United States Air Force Academy*
*Colorado Springs, Colorado*

**ABSTRACT**

This paper is concerned with models of the use of performance incentives in the Soviet Union and United States. The principal analytical result is an extension of an analysis of the methods whereby Soviet planners make the decision about production targets a variable under control of the producer, who is the only one possessing a knowledge of the uncertain condition of production. It is shown that this device can be viewed as a classical inventory problem. There is also an examination of the "U.S. incentive program" referring to multi-incentive contracts in which the profits received by the private producers are related to performance, outcome and cost. The analysis describes how this device can be extended to solve the target output selection problem of the Soviet planning system.

## INTRODUCTION

The central planning organizations of both the Soviet Union and the United States have been concerned with influencing the behavior of enterprises in order to achieve improved allocation of resources. Although one is more likely to associate the central planning task with an economy such as the Soviet Union, the provision of many goods is centrally planned in the United States. For example, the Planning-Programming-Budgeting System of the United States government can be viewed as part of a central planning process. In the analysis I compare the incentive systems of the Soviet Union and the United States, thereby clarifying the similarities that exist between the two economic systems. A suggestion for improvement to the U.S. incentive system is also made.

The key similarity between planning in the Soviet Union and the United States is that the government does not know as much about individual values and technological opportunities as do the relevant producers or consumers. For example, in the United States, the Planning-Programming-Budgeting System frequently culminates in government acquisition of goods such as military hardware and space systems from private enterprises. The production technolgies associated with many of these goods are not only highly uncertain, but are also more accurately known by the producers than by the government. Such goods as intercontinental ballistic missiles and manned space vehicles have embodied in them advanced technology and associated uncertainty about the conditions of production. This uncertainty is probably most pervasive during the engineering development phase of the "production" process when the performance characteristics of these goods are determined. The producers of these goods, however, will

typically have a greater knowledge of this advanced technology and its impact on the conditions of production than does the government. Thus, the conditions of production are more uncertain for the government than for the producer at that time. The uneven impact of uncertainty implies that it is not possible for the government to specify the best output level.

A similar situation exists in the Soviet Union. The state enterprise may have better knowledge of its production technology than the planners, and the Soviet planners, therefore, may be unable to specify the optimal output level. Yet both the U.S. and Soviet decision makers attempt to provide appropriate incentives to motivate producers to select the "right" output level.

In the United States this system employs the so-called "contractual incentive function" which specifies a mutually acceptable rule connecting the monetary rewards of one decision maker to the subsequent performance of another. Numerous enterprises have devised profit-sharing formulae to motivate supervisory and managerial personnel, and the Department of Defense and NASA have relied on the use of performance incentives to monitor the work of major contractors. For example, performance incentives were included in contracts with a total value of several billion dollars during the U.S. moon program.* Recent innovations in the use of performance incentives have appeared in the new Amtrak contract which provides payments to the railroads according to the quality of services they provide [2]. A similar contractual arrangement guarantees a one percent increase in the salaries of the policemen of Orange, California for every three percent decline in rape, robbery, burglary and auto theft [15].

Although the existing literature on economic planning does not specifically mention the use of contractual incentive functions, a related concept has arisen in discussions of "success indicators" in Soviet planning. The Soviet planning system often rewards enterprise agents according to the degree to which producers reach certain planned targets.† Thus, Soviet planners have implicitly defined a performance incentive system. In contrast to similar systems employed in the West, the Soviet system has not been "contractual" in the sense that it has been agreed upon by the planners and the enterprise managers. Instead, the state has unilaterally chosen the targets and rewards, and the enterprise managers have been expected to comply in order to attain their own maximum reward within the confines of the rules laid down by the planners. This "non-contractual" incentive system has clearly been an example of the use of performance incentives in the implementation of economic planning.

Recently, the Soviets have experimented with an extended incentive system which provides motivation for the state enterprises to select the optimal target output level before the determination of the actual output. The importance of this additional incentive stems from the fact that if the central planners have a good estimate of the amount of the good which *will be*

*Using $z$, $x$, and $r$ to represent measures of relative profit, cost, and performance, the functional form which applied to several of the large dollar value incentive contracts used during the U.S. moon program is

$$z = f(x) + g(r) + \alpha f(x)g(r) + \beta$$

where

$$f(x) = a_1 x e^{a_2 x^{a_3}} \quad g(r) = b_1 r e^{b_2 r^{b_3}}$$

and $a_1$, $a_2$, $a_3$, $b_1$, $b_2$, $b_3$, $\alpha$, and $\beta$ are constants. In addition, between 1967 and 1970, there were approximately $27 billion of multiple incentive contracts evaluated by a Department of Defense analysis group.

†It is widely accepted that this is the Soviets' most famous planning problem. In addition to monetary incentives, the Soviets have also tried to solve this problem using informational exchange during bargaining with the enterprise.

oduced before it is *actually* produced, then a better coordinated plan can be achieved. There also a need for planners to coordinate outputs that are jointly used. For example, some intermediate goods are used jointly in the production of final goods, and final goods may be jointly nsumed. The reason why an incentive is required to motivate the managers of the state terprises to reveal the optimal target is that the enterprise may also receive a reward based on e actual output achieved in relation to the target output level. The existence of this reward ay motivate the managers to understate the target output level if they are simply asked its lue.

Martin Weitzman [16] has analyzed this new incentive system using a model whose properties have been investigated by Fan [6] and Bonin [3]. In this model, the planners fix the sources or inputs available to the enterprise, but there is uncertainty associated with the output that can be produced with these fixed inputs. The uncertainty rests with the planners, thus stifying the selection of the target output level by the enterprise. Although the output actuly achieved is not selected by the enterprise, Weitzman shows how the enterprise can use its nowledge of the uncertain conditions of production in conjunction with a specified performance incentive to select the best target output.

This report will first review the Weitzman analysis and then show that the new Soviet centive program can be viewed as a classical inventory problem, which is a problem of determining how much of product to keep in storage. This interpretation of the incentive program is nportant because inventory theory is a well-developed analytical framework and general associons between inventory theory and planning may prove fruitful.

Recently, the state enterprises in the Soviet Union have been given greater flexibility in eir use of inputs. In view of this change, I next show how the new Soviet incentive program n be extended to deal with a situation in which the inputs used by the enterprise are choice riables with associated cost. This extension has some similarities to the analyses of Bonin nd Marcus [4], Snowberger [14], and Miller and Thornton [11] who have investigated the nplications of a relationship between output and variations in the producer's effort.

An analysis of the U.S. incentive program will follow the discussion of the Soviet incenve system. To ease comparison between the two systems, a cost-effectiveness model will be sed. Thus, I assume that the objective of the goverment is the achievement of some specified vel of performance at minimum cost. The large degree of uncertainty that exists during ngineering development prevents the government from knowing in advance what performance vel will be achieved for any level of expenditure. Furthermore, both the estimate of the tual performance level, the target, and the performance level actually achieved have associed costs that must be borne by the government rather than the producer. Therefore, the perrmance incentive can be viewed as a method of motivating the producer to take appropriate count of these costs during engineering development. This internalization of social costs by e producer suggests a relationship not only to the control of environmental externalities, but so to the design of a transfer price by one division of an enterprise that is providing some termediate good to a second division. The transfer pricing problem has been recently alyzed by Ronen [13], and Groves and Loeb [7].

The existing incentive system motivates the producer to economize on the costs associed with the output level actually achieved. This system can be expanded to solve the target tput selection problem. This expansion would lead to a greater compatibility of the various terrelated output decisions, thus making the target output itself a product worth paying for in e U.S.

## I. THE NEW SOVIET INCENTIVE MODEL

In analyzing the new Soviet incentive system, Weitzman uses a model in which the factors of production used by the enterprise are set by the planners, an assumption which is realistic in the U.S.S.R. where inputs have typically been rationed by the state. A tentative target $\bar{y}$ and a tentative bonus fund $\bar{B}$ are assigned to the enterprise during the first or *preliminary* phase. The tentative target is the planner's best estimate of the target output level at that time. During the second, or *planning* phase, the enterprise has the option of revising the tentative target to $\hat{y}$ which has associated with it a revised bonus fund $\hat{B}$ computed in accordance with the formula,*

$$\hat{B} = \bar{B} + \beta(\hat{y} - \bar{y}),$$

where the constant $\beta$ is proportional to the "real social value of having an extra unit which has been pre-planned" [16].

In the third or *implementation* phase, when the enterprise ends up producing amount $y$, it actually receives the bonus fund

$$B = \begin{cases} \hat{B} + \alpha(y - \hat{y}) : y \geqslant \hat{y} & \text{(overfulfillment)} \\ \hat{B} - \gamma(\hat{y} - y) : y < \hat{y} & \text{(underfulfillment)} \end{cases}$$

where $\alpha$ is proportional to the "real social value of having an extra unit unexpectedly delivered," and $\gamma$ is proportional to the "real social cost of being unexpectedly caught short by one unit" [16]. Subsequent analysis reveals that correct decision making by the enterprise requires that the constant $\beta$, $\alpha$, and $\gamma$ be in the same proportion to their respective value coefficients. Under the "old" Soviet incentive system, $\hat{B}$ and $\hat{y}$ were fixed by the planners. Under the new system, they are set by the enterprise.

In the model developed by Weitzman, there is uncertainty during the planning phase as to the amount of output that will actually be produced with the fixed inputs. Only the producer knows the probability density function $f(y)$. Thus, we have an example of the informational asymmetry which is so prevalent during the planning process and a justification for the producer to select the target output level. This uncertainty might in fact persist during the implementation phase, but with fixed inputs, actual output $y$ is not a choice variable, and thus, the character of the uncertainty that applies then is not relevant to this analysis.

During the planning phase, when $\hat{y}$ is selected, the problem faced by the enterprise (assumed risk neutral) is to choose $\hat{y}$ to maximize

$$\int_{-\infty}^{\hat{y}} [\bar{B} + \beta(\hat{y} - \bar{y}) + \gamma(y - \hat{y})]\, f(y)\, dy + \int_{\hat{y}}^{\infty} [\bar{B} + \beta(\hat{y} - \bar{y}) + \alpha(y - \hat{y})]\, f(y)\, dy. \tag{1}$$

By differentiating with respect to $\hat{y}$, Weitzman shows that the optimal solution to this maximization problem is to select $\hat{y}$ such that

$$P(y \geqslant \hat{y}) = \frac{\gamma - \beta}{\gamma - \alpha}, \tag{2}$$

where

*If the selection of the target $\hat{y}$ influences the allocation of inputs to the enterprise, then the producer will take this into account in the selection of $\hat{y}$. This raises a serious incentive compatability problem which has been discussed by Magat and Loeb [10]. We assume that such a relationship does not exist.

$$P(y \geqslant \hat{y}) = \int_{\hat{y}}^{\infty} f(y)dy.$$

Because it is possible to multiply all of the coefficients by a constant without changing (2), only the relative magnitudes of the coefficients matter in determining the optimal $\hat{y}$. The appropriate relative magnitudes are achieved when these coefficients are in the same proportion to their respective value coefficients.* Furthermore, in view of the fact that (2) must be positive, this incentive system is meaningful only when the coefficients are set such that $\alpha < \beta < \gamma$.

**An Inventory Theory Interpretation**

The fact that there are costs borne by the center when the actual outcome is both below and above target suggests that an inventory theoretic interpretation can be given to Weitzman's analysis. To see the classical inventory structure of this problem, rewrite (1) as

$$\overline{B} + \beta(\hat{y} - \overline{y}) + \int_{-\infty}^{\hat{y}} \gamma(y - \hat{y}) f(y)dy + \int_{\hat{y}}^{\infty} \alpha(y - \hat{y}) f(y)dy. \tag{3}$$

The difference, $\overline{B} - \beta\overline{y}$, is fixed and therefore not relevant when choosing the target output level, but we must concern ourselves with the term $\beta\hat{y}$ which can be written as

$$\beta\hat{y} \equiv \beta \int_{-\infty}^{+\infty} yf(y)dy + \beta \int_{-\infty}^{\hat{y}} (\hat{y} - y)f(y)dy - \beta \int_{\hat{y}}^{\infty} (y - \hat{y})f(y)dy.$$

The expression,

$$-\beta \int_{\hat{y}}^{\infty} (y - \hat{y})f(y)dy, \tag{4}$$

is proportional to the *benefits foregone*, weighted by the probabilities, as a result of the economic system not being geared to a higher $\hat{y}$ when the actual output is larger than the target.

If the target is not achieved, the actual $y$ is less than $\hat{y}$, and

$$\beta \int_{-\infty}^{\hat{y}} (\hat{y} - y)\, f(y)dy \tag{5}$$

is applicable. This expression can be viewed as (proportional to) the *benefits still received* (weighted by probabilities) from having the system geared to $\hat{y}$. Now insert (4) and (5) into the last two parts of (3), and obtain as the producer's problem the maximization of

$$(\beta - \gamma) \int_{-\infty}^{\hat{y}} (\hat{y} - y)f(y)dy + (\alpha - \beta) \int_{\hat{y}}^{\infty} (y - \hat{y})f(y)dy$$

which is equivalent to the minimization of

$$(\gamma - \beta) \int_{-\infty}^{\hat{y}} (\hat{y} - y)f(y)dy + (\beta - \alpha) \int_{\hat{y}}^{\infty} (y - \hat{y})f(y)dy.$$

The coefficient $\gamma - \beta$ can now be identified as (proportional to) the net social cost per unit of output actually achieved below the target and $\beta - \alpha$ as (proportional to) the net social cost per unit of output above the target.

---

*Each coefficient of the right hand side of (2) can be multiplied by a constant $k$ yielding

$$\frac{k\gamma - k\beta}{k\gamma - k\alpha} = \frac{k(\gamma - \beta)}{k(\gamma - \alpha)} = \frac{\gamma - \beta}{\gamma - \alpha}$$

In that the value coefficients are measured in rubles per extra output, multiplying each coefficient by a constant can be viewed as a change in the monetary unity which could never affect the selection of $\hat{y}$. Also, note that the units associated with each coefficient of the right hand side of (2) cancel. As the left hand side of (2) is a probability (a pure number), such a cancellation is required to equate both sides of (2).

Let us use the notational convention

$$c_1 = \gamma - \beta,$$

$$c_2 = \beta - \alpha.$$

The problem facing the enterprise can therefore be written

$$\underset{\hat{y}}{\text{Min}}\; c_1 \int_{-\infty}^{\hat{y}} (\hat{y} - y) f(y)\, dy + c_2 \int_{\hat{y}}^{\infty} (y - \hat{y}) f(y)\, dy. \tag{6}$$

When the producer's maximization problem (1) is rewritten as the minimization problem (6), it is possible to view the selection of $\hat{y}$ as the selection of the amount of a good (the target) to be placed in inventory. The coefficient $c_1$ can be viewed as the carrying cost per unit of unsold inventory and $c_2$ the per unit shortage cost. Taking the derivative of (6) with respect to $\hat{y}$ (and equating it to zero) we see that for the optimal solution value $\hat{y}$,

$$P(y < \hat{y}) = \frac{c_2}{c_1 + c_2} = \frac{\beta - \alpha}{\gamma - \alpha}. \tag{7}$$

Equation (7) is a well-known formula from inventory theory [5]. Therefore,

$$P(y \geqslant \hat{y}) = \frac{c_1}{c_1 + c_2} = \frac{\gamma - \beta}{\gamma - \alpha},$$

which is the solution obtained by Weitzman.

The inventory formula (7) has a simple economic interpretation. Letting $P = P(y < \hat{y})$, this formula can be rewritten as

$$Pc_1 = (1 - P)\, c_2 \tag{8}$$

and indicates that $P$ should be selected through the selection of $\hat{y}$, so that the expected *net* social cost of the output produced less than the target and the output produced at least as great as the target are equal. The reason an inventory theoretical interpretation is interesting is that inventory theory is a well-developed framework and analogies that can be found with the planning process might prove fruitful in the development of a theory of economic planning.

**Production Inputs Variable**

Although Weitzman has chosen to view production inputs as fixed, largely because this assumption reflects the Soviet planning environment, it is possible to extend his analysis by allowing the production inputs used during the implementation phase to be choice variables of the enterprise. This extension may have relevance to the Soviet planning problem now that the managers of state enterprises are being given greater flexibility in the use of inputs. In order to simplify the analysis, I assume that there is the *same degree of uncertainty* about the conditions of production during the planning phase when the enterprise selects the target output level and during the implementation phase when a level of cost expenditure is selected. In the view of the enterprise, the conditions of production during both phases can be represented by

$$y = h(c, \theta),$$

where $\theta$ is a random variable with density function $f(\theta)$ applicable for both the planning phase and the implementation phase, and $c$ represents production costs. Although identical uncertainty permits one to view the enterprise as selecting $\hat{y}$ and $c$ simultaneously, a meaningful economic interpretation can still be given for the need to select $\hat{y}$ beforehand by assuming that

the actual output $y$ is not revealed at the time $c$ is selected but rather at some later time which can be called the implementation phase. It will also be assumed in this extension that the share of cost expenditure borne by the enterprise is equal to $s$.

Assuming that the coefficients $\beta$, $\gamma$, and $\alpha$ continue to apply, and the profits are determined by $T(\hat{y},c)$, the producer must solve

$$\text{(9)} \quad \max_{\hat{y},c} T(\hat{y},c) = \int_{-\infty}^{\theta^{-1}(\hat{y},c)} [\bar{B} + \beta(\hat{y} - \bar{y}) + \gamma(h(c,\theta) - \hat{y}) - sc] f(\theta)\, d\theta + \int_{\theta^{-1}(\hat{y},c)}^{\infty} [\bar{B} + \beta(\hat{y} - \bar{y}) + \alpha(h(c,\theta) - \hat{y}) - sc] f(\theta)\, d\theta$$

where the inverse function $\theta^{-1}(\hat{y},c)$ determines the value of $\theta$ which achieves $y = \hat{y}$ when the production costs are $c$. The enterprise must set the derivative of this with respect to $\hat{y}$ equal to zero obtaining

$$\partial T/\partial \hat{y} = \int_{-\infty}^{\theta^{-1}(\hat{y},c)} (\beta - \gamma) f(\theta)\, d\theta + \int_{\theta^{-1}(\hat{y},c)}^{\infty} (\beta - \alpha) f(\theta)\, d\theta = 0.$$

It is easy to verify that this equality implies that

$$\text{(10)} \quad P(\theta \geqslant \theta^{-1}(\hat{y},c)) = \frac{\gamma - \beta}{\gamma - \alpha}.$$

Similarly, the derivative of (9) with respect to $c$ set equal to zero yields*

$$\text{(11)} \quad \frac{\partial T}{\partial c} = \int_{-\infty}^{\theta^{-1}(\hat{y},c)} [\gamma h_c(c,\theta) - s] f(\theta)\, d\theta + \int_{\theta^{-1}(\hat{y},c)}^{\infty} [\alpha h_c(c,\theta) - s] f(\theta)\, d\theta = 0.$$

To obtain qualitative results we require knowledge of the function $h(c,\theta)$. Assume that the uncertainty is additive and that $h_c$ depends only on $c$ ($h_{c\theta} = 0$).† Then it can be shown that (11) implies that

$$\text{(12)} \quad P(\theta \geqslant \theta^{-1}(\hat{y},c)) = \frac{\gamma h_c - s}{(\gamma - \alpha) h_c}.$$

For both (10) and (12) to be satisfied simultaneously, it must be true that

$$\frac{\gamma - \beta}{\gamma - \alpha} = \frac{\gamma h_c - s}{(\gamma - \alpha) h_c}$$

which implies that the producer must set

$$\text{(13)} \quad \beta h_c(c) = s.$$

This condition implies that when $\hat{y}$ is optimal, the selection of the level of cost by the enterprise can be determined by evaluating the profit from a small adjustment in $\hat{y}$. The effect on the profit obtained from $y$ captured when the optimal value of $\hat{y}$ is selected. Thus, the producer should vary $c$ until the extra profit associated with a small increase in target output ($\beta h_c$) just equals the reduced profits from increasing $c$ by one unit ($s$).

One can rewrite (13) as

$$c = h_c^{-1}(s/\beta),$$

---

*Notation such as $h_c$ represents the partial derivative of the function $h$ with respect to the variable $c$.

†Although this assumption is strong, it is frequently interesting to know what assumptions are required to obtain a sharp characterization of an optimal policy. It is not difficult, however, to imagine an interaction between the level of cost expenditure and the random variable $\theta$. For example, high levels of cost expenditure might be associated with greater uncertainty. Such interactions have been excluded from the analysis.

thereby permitting (10) to be written as

$$P(\theta > \theta^{-1}(\hat{y}, h_c^{-1}(s/\beta)) = \frac{\gamma - \beta}{\gamma - \alpha}. \tag{14}$$

The enterprise must satisfy this condition during the planning phase when selecting $\hat{y}$. This condition recognizes that during the implementation phase the producer selects the optimal cost expenditure. Comparing (14) with (2) shows that the producer must simply account for the impact of the additional choice variable (cost) on the likelihood of being over target. However, once this adjustment is made, the economic interpretation described by (8) continues to apply.

## II. THE U.S. INCENTIVE MODEL

The purpose of the U.S. incentive model is to motivate producers to select a performance level that is socially optimal. The DOD and NASA Guide states that

> the concept of multiple incentive contracting must quantitatively relate profit motivation directly and in accordance with the Government's objectives. . . . it establishes the contractor's profit in direct relationship to the value of the combined level of performance in all areas [5].

Furthermore,

> the process of including performance in an incentive structure must logically begin with the determination of the "value" of the characteristics which will be incentivized. The multiple incentive contract should reflect the importance to the government of various cost, schedule, and performance outcomes, through the profits assigned to each part of the multiple incentive structure [5].

### Cost-Effectiveness Analysis

One method of describing the U.S. incentive model is to use a cost-effectiveness analysis approach.* This approach applies when the government's objective is the achievement of some specified level of *system* performance at minimum cost and it simplifies comparison of the U.S. incentive model with the Soviet model. It is assumed that increasing performance level $p$ of some component of the system during the engineering developing phase of procurement leads to future, or "downstream" cost savings for the government because of reduced acquisition cost, maintenance costs, etc. The basic structure of the U.S. incentive model can be most easily illustrated if it is assumed that the producer is given a performance reward based on the level of $p$ actually achieved and on development costs. Later, a more complicated model will show how the U.S. incentive program can be expanded to incorporate the target selection features of the new Soviet incentive program.

I assume that the cost of development function, $C(p)$, is deterministic during the implementation phase when the producer actually selects $p$. This function may, however, be known only to the producer. Indeed, in order to justify using a performance incentive in the first place, there must be some uncertainty in the government's mind about the cost of development at the time the incentive is specified. Otherwise, the government would simply specify $p$. The

*As far as I am aware, the first mathematical treatment of multiple incentive contracting using a cost-effectiveness approach similar to the one presented here was by Ackerman and Krutz [1].

downstream cost function, $D(p)$, determines the costs borne by the *government* through the dependence on the output level selected by the *producer*. This function is assumed to be known by the government. Total cost is the sum of the development cost and the downstream cost and is designated $T(p)$. At the time of *producer* decision making, the *objective* of the government is to solve

$$\underset{p}{\text{Min}}\ T(p) = C(p) + D(p).$$

The first order condition for this problem is

$$C' = -D', \tag{15}$$

which simply says that the performance level should be increased until the producer's marginal development cost expenditure just equals the government's marginal downstream cost reduction.

The profit, or performance incentive function given to the producer under the U.S. incentive system is typically of the form

$$\pi = G(p) - sC \tag{16}$$

where $G(p)$ represents dollars of profit earned as a function of the performance level $p$, and $s$ equals the share of the development cost borne by the producer. The relevant first order, or profit maximizing condition for the *producer* is

$$G'(p) = sC'(p).$$

In view of the government's optimization condition (15), the optimal incentive structure is obtained when $-D'$ is substituted for $C'$, and the government constructs the performance incentive function such that

$$G'(p) = -sD'(p).$$

Therefore, with the inclusion of a constant $A$, a performance incentive function of the form

$$\pi = -sD(p) - sC + A$$

will motivate the producer to satisfy (15), thereby satisfying the objectives of the government.*

**Extending the U.S. Incentive Model**

For selected U.S. procurements in which the producer is the only supplier of a good whose performance is rewarded in relation to some target (thereby creating an incentive for the producer to understate the target if simply asked its level), there is value in extending the U.S. incentive program to include producer target specification. In addition to depending on the actual performance level, downstream costs also depend on the target performance level because of the time needed to prepare the operational environment (e.g., train maintenance people, etc.) for the actual performance level.

In extending the U.S. incentive model, I assume, for the purpose of comparison with the Soviet incentive model, that during the planning phase an incentive function is specified and the producer selects a *target* performance level $\hat{p}$. The actual performance level $p$ is not achieved until an implementation phase.

*N. J. Ireland [9] has shown that the "ideal price" is the benefit function of the central planning organization for certain types of planning environments. The function $D(p)$ can be interpreted as the benefit function of the planners for the problem at hand. Note, however, that in this analysis, the cost sharing ratio is also applied to the function $D(p)$.

During the planning phase, the downstream cost function will be of the form $D(\hat{p},p)$. Although all costs are variable at that time, certain downstream costs are fixed at the time the actual performance is achieved. During the planning phase the dependence of these costs on the target performance level can be represented by $F(\hat{p})$. Those costs which remain variable when the actual performance level is determined can be represented by $Dv(\hat{p},p)$.

In order to parallel the extension of the new Soviet incentive model to the situation where inputs are variable, I now assume that the choice variable of the producer is a level of development cost expenditure $c$. In the cost-effectiveness analysis section above, the performance level $p$ was selected as the producer's choice variable. At both the time the producer selects the target performance level and the time that a cost expenditure level is selected, the producer's view of the conditions of production is represented by

$$p = g(c,\theta),$$

where $\theta$ is a random variable which has the same density function at both of the times of producer decision making. It is assumed that the government does not know $g$ during the planning phase, thus justifying the selection of $\hat{p}$ by the producer. As we shall see, the government's information about downstream cost is transmitted to the producer in the incentive function. This information combined with the producer's information about the conditions of production yields, via profit maximization, the best solution to the target selection problem.

The government in taking a social view is interested in the minimization of

$$E(c + Dv(\hat{p},p) + F(p)).$$

The first order conditions associated with this minimization are

$$1 + E(\partial Dv/\partial p)\ (\partial p/\partial c) = 0 \tag{17}$$

$$E(\partial Dv/\partial \hat{p}) + \partial F(\hat{p})/\partial \hat{p} = 0. \tag{18}$$

The profit function given to the producer is of the form

$$\pi = G(\hat{p},p) - sc, \tag{19}$$

where $s$ again represents the share of the development cost borne by the producer. This function has the same basic form as (16) to retain compatibility with what has typically been used for the existing U.S. incentive system. The first order conditions which apply for the *producer* are

$$E(\partial G/\partial \hat{p}) = 0, \tag{20}$$

$$E(\partial G/\partial p)\ (\partial p/\partial c) = s. \tag{21}$$

Comparing (17) and (18) with (19) and (20) shows that the government can achieve its objective if it constructs an incentive function such that

$$G_{\hat{p}} = -\ s(Dv_{\hat{p}} + F_{\hat{p}}), \tag{22}$$

$$G_p = -\ s(D_p). \tag{23}$$

When condition (22) is satisfied, the incentive profit received by the producer from a change in the target performance level is just equated to a proportion of the incremental downstream cost savings. A similar interpretation applies to (23).*

*The expectation operators are not required in (22) or (23) because the terms inside the expectation operators of (20) and (21) are substituted for the terms of (17) and (18) inside these operators.

Thus,

$$G(\hat{p},p) = -sD(\hat{p},p) + A,$$

where $A$ is a constant.

By linearizing the function $Dv(\hat{p},p)$ about $\hat{p}$, and $F(\hat{p})$ about $\bar{p}$, where $\bar{p}$ is some specified performance level, e.g., a government estimate of the target performance level, one can obtain a formal equivalence of the U.S. and the Soviet incentive models. Thus, if one approximates $Dv(\hat{p},p)$ by

$$K/s + \alpha/s(\hat{p} - p) \text{ when } p \geqslant \hat{p}$$
$$K/s + \gamma/s(\hat{p} - p) \text{ when } p < \hat{p},$$

and $F(\hat{p})$ by

$$M/s + \beta/s(\bar{p} - \hat{p}),$$

then

$$-sD(\hat{p},p) = \begin{bmatrix} -(K+M) + \alpha(p-\hat{p}) + \beta(\hat{p}-\bar{p}) : p \geqslant \hat{p}) \\ -(K+M) + \gamma(p-\hat{p}) + \beta(\hat{p}-\bar{p}) : p < \hat{p}). \end{bmatrix}$$

The parameters $\alpha$, $\gamma$, and $\beta$ have the same interpretation as in the new Soviet incentive model. For example, in that $\hat{p}$ is a preplanned performance outcome, $\beta$ is simply proportional to the social value of having an extra unit which has been preplanned and can be identified as a proportion of the cost savings achieved when $\hat{p}$ is varied during the *planning* phase.

If inputs are fixed as assumed by Weitzman, the term $sc$ vanishes from the profit function (19), and one obtains an equivalence to the new Soviet incentive model. If the inputs are variable and $g(c,\theta)$ applies, then one obtains an equivalence to the extended Soviet incentive model developed above. The relevant maximization problem that must be solved by the producer is analogous to (9).

Note that the cost share $s$ is the factor of proportionality that applies to the parameters $\alpha$, $\gamma$, and $\beta$. If the cost share changes, then so too will the parameters. Thus, there appears to be a degree of freedom in the selection of these parameters. However, this factor of proportionality has distributional significance and, in fact, is related to the distribution of societal profits between the center and the enterprise. It is subject to optimization in an analysis of risk sharing between the center and the producer, and has been discussed by Hildebrandt and Tyson [8] who show that under certain types of differentiated information structures, it is still possible to achieve both the distributional and allocational objectives of the planner.

## CONCLUSIONS

Decision makers in both the United States and the Soviet Union face similar problems of correctly guiding production at the enterprise level. To achieve certain social objectives, the United States government has employed the contractual incentive function whereas the Soviet planners have used the non-contractual or unilarteral incentive function.

The new Soviet incentive system provides an incentive for the enterprise to reveal the socially optimal target output level. My analysis has shown that this system can be expanded to deal with the situation when the enterprise controls the amount of resources utilized, a situation which is becoming increasingly typical in the Soviet Union and which continues to be the

norm in the United States. Although the option of placing an incentive on the target output level has not yet been used in the United States, the existing U.S. incentive system can be expanded to permit that possibility.

## BIBLIOGRAPHY

[1] Ackerman, D. and Krutz, R. "Structuring of Multiple Incentive Contracts," United States Air Force Academy Working Draft (May 1966).
[2] Baumol, W.J. "Payment by Performance in Rail Passenger Transportation: An Innovation in Amtrak's Operations," Bell Journal of Economics, 281-99 (Spring 1975).
[3] Bonin, J.P. "On the Design of Managerial Incentive Structures in a Decentralized Planning Environment," American Economic Review, 682-87 (September 1976).
[4] Bonin, J.P. and A. Marcus, "Information, Motivation and Control in Decentralized Planning: The Case of Discretionary Managerial Behavior," Unpublished Manuscript, 1-27 (June 1977).
[5] DOD and NASA Guide, "Incentive Contracting Guide" (October 1969).
[6] Fan, Liang-Shin. "On the Reward System," American Economic Review, 226-29 (March 1975).
[7] Groves, T. and M. Loeb, "Reflections in "Social Costs and Benefits and the Transfer Pricing Problem," Journal of Public Economics, *5*: 353-59 (1976).
[8] Hildebrandt, G. and L. Tyson, "Performance Incentives and Planning Under Uncertainty," Econometric Research Program Research Memorandum No. 201, Princeton University (July 1976).
[9] Ireland, N.J. "Ideal Prices vs. Prices vs. Quantities," Review of Economic Studies, *44*: 183-86 (1977).
[10] Loeb, M. and W. Magat, "Success Indicators in the Soviet Union: The Problem of Incentives and Efficient Allocations," American Economic Review, 173-81 (March 1978).
[11] Miller, J. and J. Thornton, "Effort, Uncertainty, and the New Soviet Incentive System," Southern Economic Journal, 432-60 (October 1978).
[12] Naddor, E. *Inventory Systems*, (John Wiley and Sons, Inc., 1966).
[13] Ronen, J. "Social Costs and Benefits and the Transfer Pricing Problem," Journal of Public Economics, *3*: 71-82 (1974).
[14] Snowberger, V. "The New Soviet Incentive Model: Comment", Bell Journal of Economics, 591-600 (Autumn 1977).
[15] *The Trenton Times*, p. 16 (Sunday, December 15, 1974).
[16] Weitzman, M. "The New Soviet Incentive Model," Bell Journal of Economics, 251-57 (Spring 1976).

# THEORETICAL ANALYSIS OF LANCHESTER-TYPE COMBAT BETWEEN TWO HOMOGENEOUS FORCES WITH SUPPORTING FIRES*

James G. Taylor

*Department of Operations Research*
*Naval Postgraduate School*
*Monterey, California*

**ABSTRACT**

This paper studies combat between two homogeneous forces modelled with variable-coefficient Lanchester-type equations of modern warfare with supporting fires not subject to attrition. It shows that this linear differential-equation model for combat with supporting fires may be transformed into one without the supporting fires so that all the previous results for variable-coefficient Lanchester-type equations of modern warfare (without supporting fires) may be invoked. Consequently, new important results for representing the solution (i.e. force levels as functions of time) in terms of canonical Lanchester functions and also for predicting force annihilation are developed for this model with supporting fires. Important insights into the dynamics of combat between two homogeneous forces with such supporting fires are discussed.

## 1. INTRODUCTION

Today military operations analysts commonly use deterministic Lanchester-type† models for developing insights into the dynamics of combat. Militarily realistic computer-based Lanchester-type models of quite complex military systems have been developed for almost the entire spectrum of combat operations, from combat between battalion-sized units [3], [7] to theater-level operations [5], [6]. Nevertheless, a simple combat model may yield a clearer understanding of significant interrelationships that are difficult to perceive in a more complex model, and such insights can subsequently provide valuable guidance for more detailed computerized investigations (see [2], [18]). In this paper we consider such a simplified variable-coefficient Lanchester-type model of combat between two homogeneous forces with supporting fires not subject to attrition, and develop important results concerning the representation and behavior of its solution. These theoretical results are shown to provide important insights into the dynamics of this combat situation.

Thus, the model that we study is important because it yields some important insights into the effects of supporting fires on the dynamics of combat (see [11], [15], and [17]). Our work here extends and unifies previous results of a number of authors [1], [11], and [13] through

*This research was partially supported by the Office of Naval Research (both through direct funding and also through the Foundation Research Program at the Naval Postgraduate School) and partially by the U.S. Army Research Office, Durham, North Carolina, under R & D Project No. 1L161102H57-05 Math (funded with MIPR No. ARO 7-79).

† So-called after pioneering work by F. W. Lanchester [10].

[17]. Taylor and Parry [17] have considered the same model and developed more-restrictive victory-prediction conditions for fixed-force-ratio-breakpoint battles by considering the force-ratio equation (see also [11]). Taylor and Brown [14] have developed a mathematical theory for solving variable-coefficient Lanchester-type equations of modern warfare (without supporting fires) and introduced canonical hyperbolic-like Lanchester functions for constructing their solution. Taylor and Comstock [16] extended this work by developing theoretical results for predicting force annihilation from initial conditions without having to spend the time and effort to explicitly compute force-level trajectories for the model without supporting fires.

In the paper at hand we show that the variable-coefficient model with supporting fires may be transformed into the one without them so that all the known results about the latter may be applied: representation of solution in terms of canonical Lanchester functions, nonoscillation of the solution, force-annihilation-prediction conditions, explicit calculation of annihilation time, etc. We consequently can translate all these results to the case of combat with supporting fires. As a result of our work here, the theory of the model with supporting fires may be considered to be as complete as that of the model without them. Consequently, one can now study this variable-coefficient model almost as easily and thoroughly as Lanchester's original simple constant-coefficient model without supporting fires.

## 2. NOTATION

The symbols that are used in this paper are defined as follows:

$a,\ b,\ \alpha,\ \beta$ = constant attrition-rate coefficients,

$a(t), b(t),\ \alpha(t),\ \beta(t)$ = time-dependent attrition-rate coefficients, the first two are taken to be given in the form $a(t) = k_a g(t)$ and $b(t) = k_b h(t)$,

$A(t),\ B(t)$ = time-dependent attrition-rate coefficients in the transformed model (5.2); given by (5.3),

$C_P(t),\ S_P(t)$ = hyperbolic-like general Lanchester functions (*GLF*) which are linearly-independent solutions to the transformed *P* force-level equation (5.5); they are analogous to the hyperbolic cosine and hyperbolic sine respectively, and their quotient is denoted as $T_P(t) = S_P(t)/C_P(t)$; $C_Q(t)$ and $S_Q(t)$ are similarly defined,

$C_X(t),\ S_X(t)$ = hyperbolic-like *GLF* which are linearly-independent solutions to the *X* force-level equation (4.3); they are analogous to the hyperbolic cosine and hyperbolic sine respectively, and their quotient is denoted as $T_X(t) = S_X(t)/C_X(t)$; $C_Y(t)$ and $S_Y(t)$ are similarly defined,

$k_a,\ k_b$ = positive constants ("scale" factors) used for the representation of $a(t)$ and $b(t)$,

$p(t),\ q(t)$ = transformed "force-level" variables corresponding to $x(t)$ and $y(t)$, respectively and related to them by (5.1); with initial values $x_o$ and $y_o$,

$\dot{p} = dp/dt$ = $dp/dt$,

$Q^*$ = parity-condition parameter for the model without supporting fires (4.1); defined by (4.7),

$t$ = battle time, with $t = 0$ denoting the beginning of battle,

$t_a^X$ = time at which the $X$ force is annihilated, i.e. $x(t_a^X) = 0$,

$t_0$ = largest finite time at which $a(t)$ or $b(t)$ ceases to be defined, positive, or continuous; we take $t_0 = 0$ when no such finite time exists,

$x(t)$, $y(t)$ = force levels of $X$ and $Y$ at time $t$ for the model with supporting fires (3.1); with initial values $x_0$ and $y_0$,

$X(t)$, $Y(t)$ = force levels of $X$ and $Y$ at time $t$ for the model without supporting fires (4.1); with initial values $x_0$ and $y_0$,

$z(t)$ = $x(t)/y(t)$,

$\lambda_I$ = $\sqrt{k_a k_b}$ and is called the combat-intensity parameter,

$\lambda_R$ = $k_a/k_b$ and is called the relative-fire-effectiveness parameter,

$\Lambda^*$ = parity-condition parameter for the model with supporting fires (3.1); defined by (6.4),

$\theta$ = $\sqrt{ab + [(\alpha - \beta)/2]^2}$.

## 3. COMBAT MODELLED BY VARIABLE-COEFFICIENT LANCHESTER-TYPE EQUATIONS OF MODERN WARFARE WITH SUPPORTING FIRES

We consider combat modelled by the following Lanchester-type equations

$$(3.1)\qquad \begin{cases} \dfrac{dx}{dt} = -a(t)y - \beta(t)x & \text{with } x(0) = x_0, \\ \dfrac{dy}{dt} = -b(t)x - \alpha(t)y & \text{with } y(0) = y_0, \end{cases}$$

where $t = 0$ denotes the time at which the battle begins, $x(t)$ and $y(t)$ the numbers of $X$ and $Y$ at time $t$, and $a(t)$, $b(t)$, $\alpha(t)$, and $\beta(t)$ denote time-dependent Lanchester attrition-rate coefficients, which represent the effectiveness of each side's fire. In any analysis of combat, we should use the above equations only for $x$ and $y > 0$ and, for example, set $dx/dt = 0$ when $x = 0$, since negative force levels have no physical meaning. However, for studying the mathematical properties of the functions defined by these differential equations, we will find it more convenient to ignore this restriction and assume that (3.1) holds for all values of $x$ and $y$.

Two situations that have been hypothesized to yield the above equations are:

(S1) "aimed-fire" combat between two homogeneous forces with "operational" losses [1], [9],

(S2) "aimed-fire" combat between two homogeneous (primary) forces with superimposed effects of supporting fires not subject to attrition [17] (see Figure 1).

The practical use of such equations in analysis depends on one's ability to obtain realistic values for the coefficients: the prediction of the attrition-rate coefficients from weapon-system-performance data has been discussed by Bonder and Farrell [2], (see also [14], and [17]).

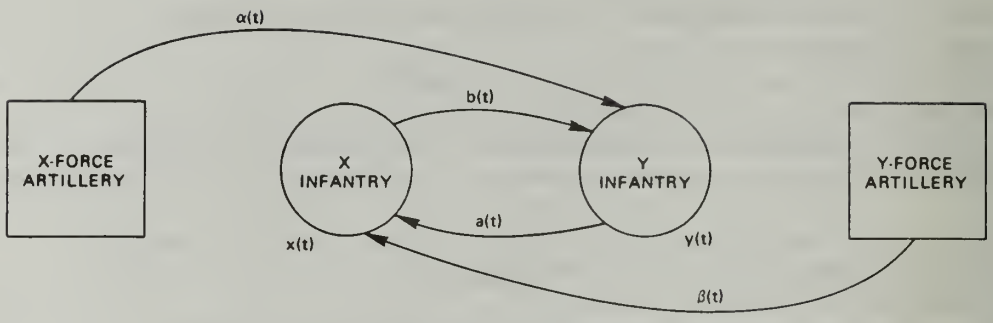


Figure 1. Combat between two homogeneous forces (infantry) with supporting weapons (artillery) not subject to attrition.

In our study here of the behavior of solutions to (3.1), we make the following assumptions about the attrition-rate coefficients:

(A1) $a(t)$ and $b(t)$ are defined, positive, and continuous for $t_0 < t < +\infty$ with $t_0 \leqslant 0$,

(A2) $\alpha(t)$ and $\beta(t) \geqslant 0$ for $t_0 \leqslant t < +\infty$,

(A3) $a(t)$, $b(t)$, $\alpha(t)$, and $\beta(t) \in L(t_0, T)$ for any finite $T$.

Here we use the notation $a(t) \in L(t_0, T)$ to mean that $\int_{t_0}^{T} a(t)\,dt$ exists (and is given by a finite quantity). It follows that (and we will assume so below) $a(t) \notin L(0, +\infty)$ means that $\lim_{T \to +\infty} \int_0^T a(t)dt = +\infty$.

We further take $a(t)$ and $b(t)$ to be given in the form $a(t) = k_a g(t)$ and $b(t) = k_b h(t)$, where $k_a$ and $k_b$ are positive constants ("scale" factors) chosen so that $a(t)/b(t) = k_a/k_b$ when $g(t) = h(t)$ for all $t$. We introduce the primary weapon systems' *combat-intensity parameter* $\lambda_I$ and the *relative-fire-effectiveness parameter* $\lambda_R$ (for the primary weapon systems) defined by

$$\lambda_I = \sqrt{k_a k_b}, \text{ and } \lambda_R = k_a/k_b. \tag{3.2}$$

Taylor and Parry† [17] noted that the force ratio, $z = x/y$, satisfies the Riccati equation

$$\frac{dz}{dt} = b(t)z^2 + \{\alpha(t) - \beta(t)\}z - a(t) \text{ with } z(0) = z_0 = \frac{x_0}{y_0}, \tag{3.3}$$

and used this fact to develop much useful information about the behavior and implications of the model (3.1). For the model (3.1) with $\alpha(t) = \beta(t) = 0$ for all $t \geqslant 0$ [i.e. the model without supporting fires (4.1) below], Taylor and Comstock [16] have developed theoretical conditions that predict force annihilation without having to spend the time and effort of computing the entire force-level trajectories. They also briefly considered (3.1) and (3.3) and observed that for identically equal fire effectivenesses of the supporting weapons [i.e. $\alpha(t) = \beta(t)$ for all $t \geqslant 0$] the same Riccati equation is satisfied by the force ratio for both (3.1) and (4.1). Consequently, in terms of the force ratio, a battle's evolution is the same for the two models (3.1) and (4.1), although the force levels initially decay more quickly for (3.1). Thus, we are led to conjecture that there must be some kind of intimate relationship between the two models with and without supporting fires.

†It was the author's good fortune to be awarded (jointly with S. Parry) the 1975 MAS Prize by the Military Applications Section of the Operations Research Society of America for the three papers Taylor and Parry [17] and Taylor [11], [12].

In this paper, we will show that the model with supporting fires (3.1) may be transformed into the one without supporting fires [i.e. (4.1)] and use this fact to develop a fairly thorough characterization of the mathematical nature of solutions to (3.1). Our results allow one, in theory, to study this particular variable-coefficient model (3.1) almost as easily and thoroughly as Lanchester's simple constant-coefficient one.

## 4. SUMMARY OF RESULTS FOR VARIABLE-COEFFICIENT LANCHESTER-TYPE EQUATIONS OF MODERN WARFARE WITHOUT SUPPORTING FIRES

In this section we summarize results about the model without supporting fires that we will use in our study of the model with supporting fires. These results have appeared in a scattered fashion in the literature (see [13], [14], and [16]), and we summarize them here in unified form for the reader's convenience. Accordingly, we consider the following variable-coefficient Lanchester-type equations of modern warfare without supporting fires

$$\begin{aligned} \frac{dx}{dt} &= -a(t)Y \quad \text{with} \quad X(0) = x_0, \\ \frac{dY}{dt} &= -b(t)X \quad \text{with} \quad Y(0) = y_0, \end{aligned} \tag{4.1}$$

Here we assume that the attrition-rate coefficients $a(t)$ and $b(t)$ satisfy assumptions (A1) and (A3) of Section 3.

Taylor and Brown [14] have shown that the $X$ (without supporting fires) force level as a function of time may be written as

$$X(t) = x_0\{C_Y(0)C_X(t) - S_Y(0)S_X(t)\} - y_0\sqrt{\lambda_R}\,\{C_X(0)S_X(t) - S_X(0)C_X(t)\}, \tag{4.2}$$

where the hyperbolic-like general Lanchester functions (GLF) $C_X(t)$ and $S_X(t)$ satisfy

$$\frac{d^2X}{dt^2} - \left\{\frac{1}{a(t)}\,\frac{da}{dt}\right\}\frac{dX}{dt} - a(t)b(x)X = 0, \tag{4.3}$$

with initial conditions

$$\begin{aligned} &C_X(t_0) = 1, \quad S_X(t_0) = 0, \\ &\frac{1}{a(t_0)}\,\frac{dC_X}{dt}(t_0) = 0, \quad \frac{1}{a(t_0)}\,\frac{dS_X}{dt}(t_0) = \frac{1}{\sqrt{\lambda_R}}. \end{aligned} \tag{4.4}$$

In other words, these functions are linearly independent solutions to the $X$ force-level equation (4.3) and are analogous to the ordinary hyperbolic functions (see [13] and [14] for further details). We will refer to any such basic pair of linearly-independent solutions (normalized in an appropriate manner) to a force-level equation like (4.3) as *general Lanchester functions* (GLF). Thus, for such linear Lanchester-type differential combat equations, the hyperbolic-like GLF play in variable-coefficient combat a role analogous to that which the ordinary hyperbolic functions play in constant-coefficient combat. The GLF $C_Y(t)$ and $S_Y(t)$ are similarly defined. The reader should recall (see Section 3 above) that $t_0$ denotes the largest finite time at which $a(t)$ and $b(t)$ ceases to be defined, positive, or continuous (see [14] and [15] for further discussion). Also, we will set $t_0 = 0$ if no such finite time exists.

Thus, the hyperbolic-like GLF $C_X$, $S_X$, $C_Y$, and $S_Y$ are basic "building blocks" for constructing the solution to (4.1). They depend on only the attrition-rate coefficients $a(t)$ and $b(t)$ and do not depend on the initial force levels $x_0$ and $y_0$ at all. Taylor and Brown [14] have given examples of various such GLF for different attrition-rate coefficients of tactical interest. Various numerical examples are also given in their work (see [14] and [15]). We observe that (4.2) simplifies to

$$X(t) = x_0 C_X(t) - y_0\sqrt{\lambda_R}\, S_X(t), \tag{4.5}$$

when $t_0 = 0$. However, Taylor and Brown [14] have shown that (4.2) simplifies to (4.5) for $t_0 \neq 0$ if and only if $a(t)/b(t) = \text{constant} = k_a/k_b$ for all $t$ (constant ratio of attrition-rate coefficients). In other words, the hyperbolic-like GLF only possess so-called algebraic addition theorems like the hyperbolic functions do only for a constant ratio of attrition-rate coefficients.

The following results have been developed elsewhere in the literature. They form the basis for our subsequent theoretical analysis of the model (3.1).

RESULT 4.1: *At most one of the two force levels $X(t)$ and $Y(t)$ can ever vanish in finite time*

RESULTS 4.2: *If either $a(t) \notin L(0, +\infty)$ or $b(t) \notin L(0, +\infty)$, then the X force (without supporting fires) will be annihilated in finite time if and only if*

$$\frac{x_0}{y} < \sqrt{\lambda_R}\left\{\frac{C_X(0) - Q^* S_X(0)}{Q^* C_Y(0) - S_Y(0)}\right\}, \tag{4.6}$$

*where*

$$\lim_{t\to+\infty}\{S_X(t)/C_X(t)\} = 1/Q^*. \tag{4.7}$$

RESULT 4.3: *The time of annihilation of the X force (without supporting fires), $t_a^X$, is given by*

$$t_a^X = T_X^{-1}(\{x_0 C_Y(0) + y_0\sqrt{\lambda_R} S_X(0)\}/\{x_0 S_Y(0) + y_0\sqrt{\lambda_R}\, C_X(0)\}), \tag{4.8}$$

*where $T_X(t)$ denotes a hyperbolic-like GLF that is analogous to the hyperbolic tangent and is defined by $T_X(t) = S_X(t)/C_X(t)$.*

Results 4.2 (see [13], however, for the proof of a more general result) and 4.3 were developed by Taylor and Comstock [16], although Result 4.2 was given in a slightly more restrictive form (namely, both $a(t)$ and $b(t) \notin L(0, +\infty)$). Let us now sketch the proof of Result 4.1 (see also [13] for a less detailed proof), which says that all solutions to (4.3) are nonoscillatory (see, for example, Ince [8, p. 224]). Multiplying the first equation of (4.1) by $Y$, the second by $X$, adding, and integrating the result between 0 and $t$, we obtain

$$X(t)\,Y(t) = x_0 y_0 - \int_0^t \{a(s)\,Y^2(s) + b(s)X^2(s)\}\,ds. \tag{4.9}$$

Considering the standard uniqueness theorem for linear differential equations (see, for example, Coddington and Levinson [4, p. 67]), we see that it is impossible for both $X(t)$ and $Y(t)$ to be equal to zero at any finite time, since they then would have to be equal to zero for all time. Thus, since $x_0 y_0 > 0$ and the integral in (4.9) is strictly increasing in $t$ and positive for $t > 0$, it follows that $X(t)\,Y(t)$ has at most one finite zero for $t \geqslant 0$, and Result 4.1 has been proven.

Result 4.1 is very useful in the theoretical study of the behavior of solutions to (4.1). It shows that if there exists a finite $t_a^X$ such that $X(t_a^X) = 0$, then $Y(t) > 0$ for all $t \geqslant 0$. Thus, if we can find a zero for $X(t)$, we know that $Y(t)$ is guaranteed to be positive for all $t$. This result is useful in proving Result 4.2. Furthermore, when continuous withdrawals for both sides (denoted as $r(t)$ and $s(t)$) are added to the model (4.1) (e.g. $dX/dt = -a(t)Y - r(t)$ where $r(t) > 0$), then Result 4.1 no longer holds (at least when $a(t), b(t), r(t)$, and $s(t)$ are constants).

## 5. TRANSFORMATION OF MODEL WITH SUPPORTING FIRES TO ONE WITHOUT THEM

In this section we show that the model with supporting fires (3.1) may be transformed into one without them (4.1). This transformation is the key result upon which this paper is essentially based. Thus, the substitution

$$\begin{cases} p(t) = x(t) \exp\left\{\int_0^t \beta(s)\,ds\right\}, \\ q(t) = y(t) \exp\left\{\int_0^t \alpha(s)\,ds\right\}, \end{cases} \tag{5.1}$$

transforms (3.1) into

$$\begin{cases} \dfrac{dp}{dt} = -A(t)q \quad \text{with } p(0) = x_0, \\ \dfrac{dq}{dt} = -B(t)p \quad \text{with } q(0) = x_0, \end{cases} \tag{5.2}$$

where

$$A(t) = a(t) \exp\left\{\int_0^t \left[\beta(s) - \alpha(s)\right] ds\right\}, \tag{5.3}$$

and

$$B(t) = b(t) \exp\left\{-\int_0^t \left[\beta(s) - \alpha(s)\right] ds\right\}.$$

The above assumptions (A1) through (A3) of Section 3 about $a(t)$, $b(t)$, $\alpha(t)$, and $\beta(t)$ imply that $A(t)$ and $B(t)$ have the same properties as $a(t)$ and $b(t)$, i.e. they $\in L(t_0, T)$ for any finite $T$ and are defined, positive, and continuous for $t_0 < t < +\infty$. Thus, we can invoke all the results for the model without supporting fires (4.1) and obtain results for the model with supporting fires (3.1) by inversion of the transformation (5.1). Before we do this, however, let us consider the representation of the solution to (5.2).

The transformed "force-level" variable $p(t)$ satisfies

$$\frac{d^2p}{dt^2} - \left\{\frac{1}{A(t)}\frac{dA}{dt}\right\}\frac{dp}{dt} - A(t)B(t)p = 0, \tag{5.4}$$

which may be written in the equivalent form

$$\frac{d^2p}{dt^2} - \left\{\beta(t) - \alpha(t) + \frac{1}{a(t)}\frac{da}{dt}\right\}\frac{dp}{dt} - a(t)b(t)p = 0, \tag{5.5}$$

with initial conditions

$$p(0) = x_0, \quad \text{and} \quad \frac{1}{a(0)}\frac{dp}{dt}(0) = -y_0, \tag{5.6}$$

Hence, by the results of Taylor and Brown [14] reported in the previous section (i.e. (4.2)), the solution to (5.4) that satisfies the initial conditions (5.6) may be written as

$$p(t) = x_0\{C_Q(0)C_P(t) - S_Q(0)S_P(t)\} - y_0\sqrt{\lambda_R}\{C_P(0)S_P(t) - S_P(0)C_P(t)\}, \tag{5.7}$$

where the hyperbolic-like GLF $C_P(t)$ and $S_P(t)$ are linearly independent solutions to the $P$ force-level equation (5.5) that satisfy the initial conditions

$$C_P(t_0) = 1, \qquad S_p(t_0) = 0,$$
$$\frac{1}{a(t_0)}\frac{dC_P}{dt}(t_0) = 0, \qquad \frac{1}{a(t_0)}\frac{dS_P}{dt}(t_0) = \frac{1}{\sqrt{\lambda_R}}. \tag{5.8}$$

In other words, $C_P(t)$ and $S_P(t)$ play exactly the same role for the equation (5.2) that $C_X(t)$ and $S_X(t)$ do for (4.1). We have, however, expressed the $P$ force-level equation in terms of the attrition-rate coefficients of the model with supporting fires (3.1). We note again that $t_0$ denotes the largest finite time at which $a(t)$ or $b(t)$ ceases to be defined, positive, or continuous. Also, we will set $t_0 = 0$ if no such finite time exists. The GLF $C_Q(t)$ and $S_Q(t)$ are similarly defined, with the following initial condition worthy of note:

$$\frac{1}{b(t_0)}\frac{dS_Q}{dt}(t_0) = \sqrt{\lambda_R}. \tag{5.9}$$

## 6. BEHAVIOR OF THE MODEL WITH SUPPORTING FIRES

It follows from (5.1), (5.2), and the results of Section 4 that when the supporting fires are present (i.e. for our model (3.1) under study), the $X$ force level as a function of time is given by

$$x(t) = [x_0\{C_Q(0)C_P(t) - S_Q(0)S_P(t)\} - y_0\sqrt{\lambda_R}\{C_P(0)S_P(t) - S_P(0)C_P(t)\}] \exp\left\{-\int_0^t \beta(s)\,ds\right\}. \tag{6.1}$$

From (5.1) it follows that for any finite $t$

$$p(t) = 0 \textit{ if and only if } x(t) = 0, \text{ and } q(t) = 0 \textit{ if and only if } y(t) = 0, \tag{6.2}$$

whence Result 4.1 yields Result 6.1.

The behavior of the model with supporting fires is then largely described by the following results.

RESULT 6.1: *At most one of the two force levels $x(t)$ and $y(t)$ can even vanish in finite time.*

RESULT 6.2: *If either $A(t) \notin L(0, +\infty)$ or $B(t) \notin L(0, +\infty)$, then the $X$ force (with $X$ force (with supporting fires) will be annihilated in finite time if and only if*

$$\frac{x_0}{y_0} < \left\{ \frac{C_P(0) - \Lambda^* S_P(0)}{\Lambda^* C_Q(0) - S_Q(0)} \right\}, \tag{6.3}$$

*where*

$$\lim_{t \to +\infty} \{S_P(t)/C_P(t)\} = 1/\Lambda^*. \tag{6.4}$$

RESULT 6.3: *The time of annihilation of the X force ( with supporting fires), $t_a^X$ , is given by*

$$t_a^X = T_P^{-1}(\{x_0 C_Q(0) + y_0\sqrt{\lambda_R} S_P(0)\}/\{x_0 S_Q(0) + y_0\sqrt{\lambda_R} C_P(0)\}), \tag{6.5}$$

*where $T_P(t)$ denotes a hyperbolic-like GLF that is analogous to the hyperbolic tangent and is defined by $T_P(t) = S_P(t)/C_P(t)$.*

RESULT 6.4: *If both $a(t)$ and $b(t) \notin L(0, +\infty)$, then either $A(t) \notin L(0, +\infty)$ or $B(t) \notin L(0, +\infty)$.*

RESULT 6.5: *If $\alpha(t) = \beta(t)$ for all $t \in [0, +\infty)$, then*

$$\begin{aligned} x(t) = [x_0\{C_Y(0)C_X(t) - S_Y(0)S_X(t)\} - y_0\sqrt{\lambda_R}\{C_X(0)S_X(t) \\ - S_X(0)C_X(t)\}] \exp\left\{ - \int_0^t \beta(s)\,ds \right\}, \end{aligned} \tag{6.6}$$

*and when either $a(t) \notin L(0, +\infty)$ or $b(t) \notin L(0, +\infty)$, then the X force (with supporting fires) will be annihilated in finite time if and only if (4.6) holds.*

RESULT 6.6: *If $\alpha(s) \geqslant \beta(s)$ for all $s \in [0,t]$ with inequality holding for a subinterval of positive length, then*

$$\frac{x(t)}{y(t)} > \frac{X(t)}{Y(t)} \quad \text{for} \quad t > 0. \tag{6.7}$$

Results 6.1 through 6.5 follow from the material in Sections 4 and 5. Let us now sketch the proof of Result 6.6. Equations (4.3) and (5.4) may be combined to yield

$$\begin{aligned} \left[\frac{p}{X}\left(\frac{\dot{p}X}{A} - \frac{p\dot{X}}{a}\right)\right]_0^t &= \int_0^t \{B(s) - b(s)\}p^2\,ds + \int_0^t \{1/A(s) - 1/a(s)\}\dot{p}^2\,ds \\ &\quad + \int_0^t \frac{1}{a(s)} \frac{(\dot{p}X - p\dot{X})^2}{X^2}\,ds, \end{aligned} \tag{6.8}$$

which is known as the Picone formula (e.g. *see* Ince [8, p. 226]). Next, we observe that the initial conditions to (4.3) and (5.4) are given by

$$p(0) = X(0) = x_0, \quad \text{and} \quad \frac{1}{a(0)}\frac{dp}{dt}(0) = \frac{1}{a(0)}\frac{dX}{dt}(0) = -y_0. \tag{6.9}$$

Substituting (6.9) into the left-hand side of (6.8) and observing that the right-hand side of (6.8) is positive for $t > 0$ because $\alpha(s) \geqslant \beta(s)$ for all $s \in [0,t]$ with strict inequality holding on a subinterval of positive length implies that $0 < A(t) < a(t)$ and $B(t) > b(t)$ for $t > 0$, we find that

$$\frac{X(t)}{Y(t)} < \frac{p(t)}{q(t)} = \frac{x(t)}{y(t)} \exp\left\{\int_0^t [\beta(s) - \alpha(s)]\,ds\right\} < \frac{x(t)}{y(t)},$$

and Result 6.6 has been proved.

## 7. ILLUSTRATION OF THEORY WITH RESULTS FOR CONSTANT-COEFFICIENT MODEL

In this section we illustrate the above general mathematical theory of the Lanchester-type equations (3.1) for the special case in which all the attrition-rate coefficients are constants. Hence, we consider

$$\frac{dx}{dt} = -ay - \beta x, \quad \text{and} \quad \frac{dy}{dt} = -bx - \alpha y, \tag{7.1}$$

where $a$, $b$, $\alpha$, and $\beta$ denote constants, and in this case $t_0 = 0$.

Applying the key transformation (5.1) of Section 5, we find that the transformed force levels $p(t)$ and $q(t)$ satisfy (5.2) with

$$A(t) = ae^{(\beta - \alpha)t}, \quad \text{and} \quad B(t) = be^{(\alpha - \beta)t}, \tag{7.2}$$

and consequently (5.4) reads

$$\frac{d^2p}{dt^2} - (\beta - \alpha)\frac{dp}{dt} - ab\ p = 0.$$

We then find that

$$C_P(t) = e^{t(\beta - \alpha)/2}\left\{\cosh \theta t + \left(\frac{\alpha - \beta}{2\theta}\right) \sinh \theta t\right\}, \tag{7.3}$$

$$S_P(t) = \frac{\sqrt{ab}}{\theta}\ e^{t(\beta - \alpha)/2} \sinh \theta t, \tag{7.4}$$

where $\theta = \sqrt{ab + [(\alpha - \beta)/2]^2}$. It follows that

$$\frac{1}{\Lambda^*} = \frac{\theta - [(\alpha - \beta)/2]}{\sqrt{ab}}$$

Hence, (6.3) yields the known constant-coefficient result (see Bach, Dolansky, and Stubbs [1] or Taylor and Parry [17]) that the $X$ force will be annihilated in finite time if and only if

$$\frac{x_0}{y_0} < \frac{1}{b}\left\{\theta - \left(\frac{\alpha - \beta}{2}\right)\right\}, \tag{7.5}$$

and also (6.1) reads (since $t_0 = 0$)

$$x(t) = \left\{x_0 C_P(t) - y_0\sqrt{\frac{a}{b}}\ S_P(t)\right\} e^{-\beta t}, \tag{7.6}$$

which may be written in a more recognizable form as

$$x(t) = \left\{x_0 \cosh \theta t - \frac{1}{\theta}\left[ay_0 + \left(\frac{\beta - \alpha}{2}\right) x_0\right] \sinh \theta t\right\} e^{-t(\alpha + \beta)/2}. \tag{7.7}$$

## 8. DISCUSSION

The above theoretical results provide many important insights into the dynamics of combat between two homogeneous forces with supporting fires not subject to attrition, and also they are quite useful for facilitating parametric analyses (see [14] and [15]). For example, we observe that when each side's supporting fires are always equally effective (i.e. $\alpha(t) \equiv \beta(t)$;

see Result 6.5), they cancel out and the battle's outcome in a fight-to-the-finish is the same (although the victor suffers greater losses) as when they are not present. Strangely enough, although both force levels decay more rapidly in the early stages of battle (see equation (6.6)), both the length of battle and the time history of the force ratio are the same both with and without the supporting fires.

Moreover, when the effectiveness of one side's supporting fires always dominates that of the other side (e.g. $\alpha(t) > \beta(t)$ always; see Result 6.6), then the force ratio is always more favorable in this model to the side with superiority in fire support than it is in the corresponding model without the supporting fires. In other words, gaining superiority in fire support is in some sense equivalent to an improvement in the force ratio of primary fighting systems (e.g. infantries). Result 6.2 shows, however, that supporting fires alone cannot win a battle of attrition but that the force ratio of the primary fighting systems must be above a threshold value (which, however, does depend on the net effectiveness of the supporting systems; see equations (5.5) and (6.4)) in order for a side to win. The main point is that it is the (cumulative) *net* effectiveness of the *supporting weapon systems* that modifies the *relative* effectiveness of the *primary systems* in order to determine the outcome of battle (see (5.2), (5.3), (6.3), and (6.4)).

In this paper we have generalized the results of Taylor and Brown [14] on representing the solution in terms of GLF and those of Taylor and Comstock [16] on predicting force annihilation for the model (3.1) by means of the key transformation (5.1). Such a close relationship between the two models (3.1) and (4.1) was suggested by Taylor and Comstock's observation that the Riccati equation (e.g. (3.3)) satisfied by the force ratio is the same for both models when each side's supporting fires are always equally effective, i.e. $\alpha(t) \equiv \beta(t)$. Moreover, Result 6.5 means that all the series solutions and GLF developed [14], [15] and [16] for the model (4.1) without supporting fires may be used for the model (3.1) in this case. Furthermore (as observed by Taylor and Comstock's [16]), the force-annihilation-prediction conditions are the same for these two models in this special case. However, we have given more general force-annihilation-prediction conditions (expressed in terms of special transcendental functions, i.e. the GLF), which apply when $\alpha(t) \not\equiv \beta(t)$ and are complementary to Taylor and Parry's [17] conditions (expressed in terms of elementary functions under more restrictive conditions).

In summary, we have presented new important results about the variable-coefficient model (3.1) for Lanchester-type combat between two homogeneous forces with supporting fires not subject to attrition (see Figure 1). These results make the mathematical theory about representing and analyzing the model's behavior just as complete as that for variable-coefficient Lanchester-type equations of modern warfare without supporting fires. In many cases of tactical interest (see the time-dependent attrition-rate coefficients considered in [13 through 16]), we can now study this variable-coefficient model (even with supporting fires) almost as easily and thoroughly as Lanchester's classic constant-coefficient one. Also, our results may be viewed as results in the theory of differential equations, and then, of course, they apply to any other system that can be modelled with (3.1) (e.g. reaction kinetics for chemical or biological processes).

## 9. CONCLUSIONS

(I) The net effectiveness of supporting weapons is the parameter that modifies the relative effectiveness of primary systems and determines (along with the relative effectiveness and the initial force ratio of the primary systems) the outcome of combat between two homogeneous forces with supporting fires not subject to attrition.

(II) Such combat with supporting fires is equivalent to combat without the supporting fires, only the relative effectiveness of the primary systems is modified by the cumulative net effectiveness of the supporting systems.

(III) Supporting weapons augment, but do not replace, primary weapon systems.

## REFERENCES

[1] Bach, R.E., L. Dolansky, and H.L. Stubbs, "Some Recent Contributions to the Lanchester Theory of Combat," Operations Research *10*, 314-326 (1962).

[2] Bonder, S. and R.L. Farrell (Editors), "Development of Models for Defense Systems Planning," Report No. SRL 2147 TR 70-2 (U), Systems Research Laboratory, The University of Michigan, Ann Arbor, Michigan (September 1970).

[3] Bonder, S. and J. Honig, "An Analytic Model of Ground Combat: Design and Application," Proceedings of U. S. Army Operations Research Symposium *10*, 319-394 (1971).

[4] Coddington, E.A. and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955).

[5] Cordesman, A. (Editor), "Developments in Theater Level War Games," Unpublished materials for C-5 Working Group of 35th Military Operations Research Symposium (1975).

[6] Farrell, R.L. "VECTOR 1 and BATTLE: Two Versions of a High-Resolution Ground and Air Theater Campaign Model," pp. 233-241 in *Military Strategy and Tactics,* R. Huber, L. Jones, and E. Reine (Eds.), (Plenum Press, New York, 1975).

[7] Hawkins, J. "The AMSAA War Game (AMSWAG) Computer Combat Simulation," AMSAA Tech. Report No. 169, U. S. Army Material Systems Analysis Activity, Aberdeen Proving Ground, Maryland (July 1976).

[8] Ince, E.L. *Ordinary Differential Equations*, Longmans, Green and Co., London, 1927 (reprinted by Dover Publications, Inc., New York, 1956).

[9] Isbell, J.R. and W.H. Marlow, "Attrition Games," Naval Research Logistics Quarterly *3* 71-94 (1956).

[10] Lanchester, F.W. "Aircraft in Warfare: The Dawn of the Fourth Air—No. V., The Principle of Concentration," Engineering 98, 422-423 (1914) reprinted on pp. 2138-2148 of *The World of Mathematics, Vol. IV,* J. Newman (Ed.), (Simon and Schuster, New York, 1956).

[11] Taylor, J.G., "On the Relationship Between the Force Ratio and the Instantaneous Casualty—Exchange Ratio for Some Lanchester-Type Models of Warfare," Naval Research Logistics Quarterly *23*, 345-352 (1976).

[12] Taylor, J.G., "Optimal Commitment of Forces in Some Lanchester-Type Combat Models," Operations Research *27*, 96-114 (1979).

[13] Taylor, J.G., "Prediction of Zero Points of Solutions to Lanchester-Type Differential Combat Equations for Modern Warfare," SIAM J. Applied Mathematics *36*, 438-456 (1979).

[14] Taylor, J.G. and G.G. Brown, "Canonical Methods in the Solution of Variable-Coefficient Lanchester-Type Equations of Modern Warfare," Operations Research *24*, 44-69 (1976).

[15] Taylor, J.G. and G.G. Brown, "Numerical Determination of the Parity-Condition Parameter for Lanchester-Type Equations of Modern Warfare," Computers and Operations Research *5*, 227-242 (1978).

[16] Taylor, J.G. and C. Comstock, "Force-Annihilation Conditions for Variable-Coefficient Lanchester-Type Equations of Modern Warfare," Naval Research Logistics Quarterly *24*, 349-371 (1977).

[17] Taylor, J.G. and S.H. Parry, "Force-Ratio Considerations for Some Lanchester-Type Models of Warfare," Operations Research *23*, 522-533 (1975).

[18] Weiss, H.K. "Some Differential Games of Tactical Interest and the Value of a Supporting Weapon System," Operations Research *7*, 180-196 (1959).

# ON SINGLE-MACHINE SEQUENCING WITH ORDER CONSTRAINTS

K. D. Glazebrook

*University of Newcastle upon Tyne, England*

**ABSTRACT**

A collection of jobs is to be processed by a single machine. Each job has a cost function associated with it which may be either linear or exponential, costs accruing when a job is completed. The machine may be allocated to the jobs according to a precedence relation. The problem is to find a strategy for allocating the machine which minimizes the total cost and which is consistent with the precedence relation. The paper extends and simplifies some previous work done by Sidney.

## 1. INTRODUCTION

A job-shop consists of one machine and a set $J = \{1, \ldots, n\}$ of jobs to be processed on it. Associated with job i is a deterministic *processing time* $X_i$ and a *cost function* $C_i$ (.) such that if job $i$ is completed at time $F_i$ (*flow time*) its cost is $C_i(F_i)$. We usually either assume that

$$C_i(t) = -C_i a^t, \ 1 \leqslant i \leqslant n, \tag{1}$$

for positive constants $C_i$ and discount rate $a, O \leqslant a < 1$, in which case the problem is said to have *discounted costs* or we assume that

$$C_i(t) = C_i t, \ 1 \leqslant i \leqslant n, \tag{2}$$

for positive constants $C_i$ in which case the costs are said to be *linear.*

A *precedence relation* $R$ exists on $J$ such that if $(i \rightarrow j) \in R$ then job i must precede job j in any *feasible permutation* of $J$. The objective is to find those permutations $\alpha$ of J which are consistent with $R$ and which minimize the *total cost* $TC(\alpha) = \sum_{i=1}^{n} C_i \{F_i(\alpha)\}$, $F_i(\alpha)$ being the completion time of job i under $\alpha$.

Sidney [5] considered the above problem with linear costs. He proposed an algorithm that yields all the optimal permutations. This paper modifies and extends Sidney's in the following ways:

(i) Some new results for stochastic versions of this problem are reported in Section 2.

(ii) Sections 3 to 6 contain a discussion of the difficult deterministic problem with discounted costs as well as the rather easier problem with linear costs. An algorithm is presented in Section 5 which yields optimal permutations for all of these problems. The proofs used to establish the validity of this algorithm are, in the author's opinion, simpler than those used by Sidney.

## 2. STOCHASTIC SCHEDULING

As Banerjee [1] points out, 'in real life the time taken to complete a job on a machine is invariably random'. In the light of this assertion, consider the following stochastic scheduling problem: Let job set $J$, precedence relation $R$ and cost functions $C_i$, $1 \leqslant i \leqslant n$, be as in Section 1. The processing times $\{X_1, \ldots, X_n\}$ are independent integer-valued random variables with honest distributions. During each time interval $[t, t+1)$, $t$ a nonnegative integer, just one of the unfinished jobs is processed on the machine.

A *feasible strategy* $\pi$ is any rule for allocating the machine to the jobs which is consistent with $R$. Under strategy $\pi$, job $i$ is completed at random time $F_i(\pi)$. The objective is to find those feasible strategies $\pi$ that minimize the *total expected cost* $TC\ (\pi) = E\left\{\sum_{i=1}^{n} C_i[F_i(\pi)]\right\}$.

To date, progress has only been made in solving such stochastic problems where the constraint set $R$ satisfies fairly stringent requirements; see, for example, Glazebrook [2] and Meilijson and Weiss [4]. The problem with general $R$ seems very complex. One possible way forward is to find a large class of problems for which the optimal strategy can be shown to be *nonpreemptive* (that is, given by a permutation of the jobs).

The following result holds whether the cost structure is discounted or linear.

THEOREM 1: If the distributions of the processing times $X_i$, $1 \leqslant i \leqslant n$, have nondecreasing hazard rate there is a nonpreemptive strategy which is optimal among all the feasible strategies.

This result is a consequence of Theorems 3 and 4 in Glazebrook and Gittins [3]. Note that the class of probability distributions covered by Theorem 1, namely those with nondecreasing hazard rates, contains a large number of examples of interest including the geometric distribution. Note also that a similar result to Theorem 1 can be obtained for continuous-time problems by considering appropriately chosen sequences of discrete-time problems.

It follows from Theorem 1 that many stochastic scheduling problems are formally equivalent to deterministic ones in that they have nonpreemptive optimal strategies. Hence the algorithm in Section 5, although given in the context of deterministic problems, also yields optimal strategies for a large class of stochastic problems.

For the rest of the paper we concentrate exclusively on the deterministic case.

## 3. ORDERING JOB MODULES: DISCOUNTED COSTS

Throughout this section we are concerned with the deterministic problem with costs given by (1). However, before we can state and prove our main result we require some terminology and notation.

A set of jobs $M \subset J$, to be denoted by $\{(i,j), 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n_i\}$, forms a *job module* if and only if

(i) $[(i, j) \rightarrow (i, l)] \in R$, $1 \leqslant i \leqslant m$, $1 \leqslant j < l \leqslant n_i$.

(ii) $[(i, j) \rightarrow (k, l)] \notin R$, $1 \leqslant i \neq k \leqslant m$, $1 \leqslant j \leqslant n_i$, $1 \leqslant l \leqslant n_k$.

(iii) If $\alpha \in J\text{-}M$ then precisely one of (a), (b) or (c) holds:

(a) $[\alpha \rightarrow (i, j)] \in R$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n_i$;

(b) $[(i, j) \rightarrow \alpha] \in R$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n_i$;

(c) $\{[\alpha \rightarrow (i, j)], [(i, j) \rightarrow \alpha]\} \cap R = \phi$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n_i$.

From (i) and (ii), the jobs in $M$ form disjoint chains $\{(i, j), 1 \leqslant j \leqslant n_i\}$, $1 \leqslant i \leqslant m$. Condition (iii) indicates that a job in $J\text{-}M$ has the same precedence relation to every job in $M$. Hence a job in $J\text{-}M$ will either precede every job in $M$ {iii(a)} or be preceded by every job in $M$ {iii(b)} or have no relation to any job in $M$ {iii(c)}.

To simplify the algebraic expressions in what follows we introduce the notation $M_j = a^{X_j}$ and

$$CM[\alpha_1, \alpha_2, \ldots, \alpha_p] = \left\{\sum_{r=1}^{m} C_{\alpha_r} \left[\prod_{l=1}^{r} M_{\alpha_l}\right]\right\} \left\{1 - \prod_{r=1}^{m} M_{\alpha_r}\right\}^{-1} \tag{3}$$

where $\{\alpha_1, \alpha_2, \ldots, \alpha_p\}$ is an ordered subset of $J$ consistent with $R$. The function $CM$ will play an important part in the analysis.

The following two Lemmas which are both easy to establish are used in the proof of Theorem 2.

LEMMA 1: Let $\max_{j=1(1)p} \{CM[\alpha_1, \alpha_2, \ldots, \alpha_j]\} = CM[\alpha_1, \alpha_2, \ldots, \alpha_{j^*}]$ where $j^*$ is the largest such maximizing index, then

$$\max_{j=r(1)p} \{CM[\alpha_r, \alpha_{r+1}, \ldots, \alpha_j]\} \geqslant \max_{j=1(1)p} \{CM[\alpha_1, \alpha_2, \ldots, \alpha_j]\}, \quad 1 \leqslant r \leqslant j^*, \tag{4}$$

and

$$\max_{j=r(1)p} \{CM[\alpha_r, \alpha_{r+1}, \ldots, \alpha_j]\} < \max_{j=1(1)p} \{CM[\alpha_1, \alpha_2, \ldots, \alpha_j]\}, \quad j^* < r \leqslant p. \tag{5}$$

LEMMA 2: If $S_1 = \{\alpha_1, \ldots, \alpha_p\}$, $S_2 = \{\beta_1, \ldots, \beta_q\}$ and $S_3 = \{\gamma_1, \ldots, \gamma_r\}$ are mutually disjoint ordered subsets of $J$ consistent with $R$ such that

(i) $\{S_2, S_3\} = \{\beta_1, \ldots, \beta_q, \gamma_1, \ldots, \gamma_r\}$ is also an ordered subset of $J$ consistent with $R$;

(ii) $CM[\alpha_1, \ldots, \alpha_p] \geqslant CM[\beta_1, \ldots, \beta_q]$, $CM[\alpha_1, \ldots, \alpha_p] \geqslant CM[\gamma_1, \ldots, \gamma_r]$ and one of these two inequalities is strict, then

$$CM[\alpha_1, \ldots, \alpha_p] > CM[\beta_1, \ldots, \beta_q, \gamma_1, \ldots, \gamma_r].$$

Theorem 2 tells us in which order the jobs in job module $M$ will be processed in an optimal permutation for the jobs in $J$.

THEOREM 2: There is an optimal permutation for the jobs in $J$ such that the following rule is adopted for processing the jobs $\{(i, j), 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n_i\}$ in $M$: If at some time only jobs $\{(i, j), 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant s_i\}$ among the jobs in $M$ have been completed, then the next job in $M$ to be processed will be a job $(k, s_k + 1)$ where $k$ is chosen such that $s_k < n_k$ and

$$\max_{j=(s_k+1)(1)n_k} \{CM[(k, s_k+1), \ldots, (k, j)]\} = \max_{[i:s_i<n_i]} \left\{\max_{j=(s_i+1)(1)n_i} \{CM[(i, s_i+1), \ldots, (i, j)]\}\right\} \tag{6}$$

If more than one job satisfies these criteria it is immaterial which of them is chosen next.

PROOF: Those jobs from $\{(i, j),\ 1 \leqslant j \leqslant n_i\}$ still remaining to be processed at any time will be referred to as chain i. If at some time chain i consists of the jobs $\{(i, j),\ s_i + 1 \leqslant j \leqslant n_i\}$ it has associated with it (see (6)) the quantity

$$\max_{j=(s_i+1)(1)n_i} \{CM[(i, s_i + 1), \ldots, (i, j)]\} \tag{7}$$

which will be called the *allocation index* for chain i.

The theorem is proved by using an induction on the number of jobs in $M$. If $|M| = 1$ the result is trivially true. We suppose that the theorem is true for $1 \leqslant |M| < n^*$ but not for $|M| = n^*$ and show that this leads to a contradiction. Consider a problem in which $|M| = n^*$.

We suppose that in an optimal permutation for $J$ the first job to be chosen from $M$ is $(s, 1)$ where chain 1, say, has the largest allocation index and chain s does not. Hence, from (7)

$$\max_{j=1(1)n_s} \{CM[(s, 1), \ldots, (s, j)]\} = CM[(s, 1), \ldots, (s, j\{s\})] < CM[(1, 1), \ldots, (1, j\{1\})]$$

$$= \max_{j=1(1)n_1} \{CM[(1, 1), \ldots, (1, j)]\} \tag{8}$$

where $j\{1\}$ and $j\{s\}$ are the largest such maximizing indices. From the inductive hypothesis, when job $(s, 1)$ has been completed the remaining jobs in $M$ will be processed in the order described in the statement of the theorem. It follows from Lemma 1 that the jobs in $M$ will be processed in the order $\{(s, 1), \ldots, (s, k\{s\}), (1, 1), \ldots, (1, j\{1\});\ T\}$ for some $k\{s\}$, $1 \leqslant k\{s\} \leqslant n_s$, where $T$ is an ordering of the job set $J - \{(1, l),\ 1 \leqslant l \leqslant j\{1\};\ (s, l),\ 1 \leqslant l \leqslant k\{s\}\}$. Consider now the following three cases which are mutually exclusive and exhaustive.

The first set of jobs in $M$ to be allocated service before being interrupted by the allocation of the machine to a job in $J$-$M$:

CASE 1 contains $\{(s, 1), \ldots, (s, k\{s\}), (1, 1), \ldots, (1, j\{1\})\}$;

CASE 2 is $\{(s, 1), \ldots, (s, k\{s\}), (1, 1), \ldots, (1, p)\}$ for some $p$, $1 \leqslant p \leqslant j\{1\}$;

CASE 3 is $\{(s, 1), \ldots, (s, q)\}$ for some $q$, $1 \leqslant q \leqslant k\{s\}$.

We now consider case 2 in detail.

**CASE 2:** Following the processing of jobs $\{(s, 1), \ldots, (s, k\{s\}), (1, 1), \ldots, (1, p)\}$ from $M$, the first "visit" is made to $J$-$M$ (that is, the first visit after service is begun on $M$). During this first visit, jobs $\{(\alpha, \mathrm{v}),\ 1 \leqslant \mathrm{v} \leqslant t_1\}$, say, are processed in order of increasing v. After this visit, jobs $\{(1, p + 1), \ldots, (1, p\{1\})\}$ from $M$ are processed where $p < p\{1\} \leqslant j\{1\}$. During the $r$th visit to $J$-$M$ jobs $\{(\alpha, \mathrm{v}),\ t_{r-1} < \mathrm{v} \leqslant t_r\}$ are processed in order of increasing v. Assume that in general, after the $r$th visit to $J$-$M$, module jobs $\{(1, p\{r - 1\} + 1), \ldots, (1, p\{r\})\}$ are processed and that $p\{r^*\} = j\{1\}$. The fact that after $(1, j\{1\})$ has been processed other jobs from $M$ may also be processed before the $(r^* + 1)^{st}$ visit to $J$-$M$ causes no problems.

Depict the permutation described above as follows:

$$(\{(s, j),\ 1 \leqslant j \leqslant k\{s\};\ (1, j),\ 1 \leqslant j \leqslant p\},\ \{(\alpha, \mathrm{v}),\ 1 \leqslant \mathrm{v} \leqslant t_1\}, \{(1, j),\ p+1 \leqslant j \leqslant p\{1\}\},\ U)$$

where $U$ is an appropriate job ordering. Call this permutation 1.

Consider now the following two permutations which, from the definition of a job module, are also consistent with $R$. Permutation 2 is

$$(\{(\alpha, v), 1 \leqslant v \leqslant t_1\}, \{(s, j), 1 \leqslant j \leqslant k\{s\}; (1, j), 1 \leqslant j \leqslant p\{1\}\}, U).$$

Permutation 3 is

$$(\{(s, j), 1 \leqslant j \leqslant k\{s\}; (1, j), 1 \leqslant j \leqslant p\{1\}\}, \{(\alpha, v), 1 \leqslant v \leqslant t_1\}, U).$$

It is easy to show that the fact that permutation 2 gives rise to at least as high a cost as does permutation 1 implies that

$$CM[(s, 1), \ldots, (s, k\{s\}), (1, 1), \ldots, (1, p)] \geqslant CM[(\alpha, 1), \ldots, (\alpha, t_1)]. \tag{9}$$

Similarly, the fact that permutation 3 gives rise to at least as high a cost as does permutation 1 implies that

$$CM[(\alpha, 1), \ldots, (\alpha, t_1)] \geqslant CM[(1, p+1), \ldots, (1, p\{1\})]. \tag{10}$$

From (9) and (10) it follows that

$$CM[(s, 1), \ldots, (s, k\{s\}), (1, 1), \ldots, (1, p)] \geqslant CM[(1, p+1), \ldots, (1, p\{1\})]. \tag{11}$$

Repeating the above argument for all visits to $J$-$M$ up to and including the $r^*$th one, the following inequality is obtained:

$$CM[(s, 1), \ldots, (s, k\{s\}), (1, 1), \ldots, (1, p)] \geqslant CM[(1, p(r^*-1)+1), \ldots, (1, j\{1\})]. \tag{12}$$

However, it follows from Lemma 1 that

$$CM[(1, p(r^*-1)+1), \ldots, (1, j\{1\})] \geqslant CM[(1, 1), \ldots, (1, j\{1\})]$$
$$\geqslant CM[(1, 1), \ldots, (1, p)]. \tag{13}$$

It follows from Lemma 1 and the fact that chain 1 has a larger allocation index than chain $s$ that

$$CM[(1, p(r^*-1)+1), \ldots, (1, j\{1\})] \geqslant CM[(1, 1), \ldots, (1, j\{1\})]$$
$$> CM[(s, 1), \ldots, (s, j\{s\})] \geqslant CM[(s, 1), \ldots, (s, k\{s\})]. \tag{14}$$

It now follows from (13), (14) and Lemma 2 that

$$CM[(1, p(r^*-1)+1), \ldots, (1, j\{1\})] > CM[(s, 1), \ldots, (s, k\{s\}), (1, 1), \ldots, (1, p)],$$

which contradicts (12).

We obtain similar contradictions for cases 1 and 3. It may further be shown, using similar techniques, that in the event of more than one unfinished chain satisfying (6) it is immaterial which of the appropriate jobs is chosen next. Hence the induction proceeds and the theorem is proved.

## I. ORDERING JOB MODULES: LINEAR COSTS

Theorem 3 tells us in which order the jobs in job module $M$ will be processed in an optimal permutation for the jobs in $J$ when the costs are linear as in (2).

THEOREM 3: There is an optimal permutation for the jobs in $J$ such that the following rule is adopted for processing the jobs $\{(i, j), 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n_i\}$ in $M$: If at some time only jobs $\{(i, j), 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant s_i\}$ among the jobs in $M$ have been completed, then the next job in $M$ to be processed will be a job $(k, s_k + 1)$ where $k$ is chosen such that $s_k < n_k$ and

$$\max_{j=(s_k+1)(1)n_k}\left\{\left(\sum_{r=s_k+1}^{j} C_{kr}\right)\left(\sum_{r=s_k+1}^{j} X_{kr}\right)^{-1}\right\} = \max_{[i:s_i<n_i]}\left\{\max_{j=(s_i+1)(1)n_i}\left\{\left(\sum_{r=s_i+1}^{j} C_{ir}\right)\left(\sum_{r=s_i+1}^{j} X_{ir}\right)^{-1}\right\}\right\}$$

If more than one job satisfies these criteria it is immaterial which of them is chosen next. This result, previously established by Sidney [5], may be deduced from Theorem 2 by considering the limit as the discount rate a tends to one.

## 5. THE ALGORITHM

Suppose the jobs $\{(i, j),\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n_i\}$ form a job module $M$ within the set of jobs $J$ with precedence relation $R$. A *chain transformation* $C_M$ acts on the pair $(J, R)$ transforming it to the pair $(J, R \cup S)$ where

(i) $R \cap S = \phi$;

(ii) with respect to $R \cup S$ the jobs in $M$ form a single chain;

(iii) $S$ contains only elements of the form $(\alpha \rightarrow \beta)$ $\alpha \in M$, $\beta \in M$. A chain transformation, then, is a device for changing job modules into chains. $(J, R)$ is *chain reducible* if it can be transformed to a pair $(J, R \cup T)$ by successive chain transformations, where with respect to precedence relation $R \cup T$ the jobs in form a single chain.

An optimal permutation for any chain reducible pair $(J, R)$ with discounted or linear costs may be obtained by successive applications of whichever is appropriate of Theorem 2 or Theorem 3. This is easily seen since those results demonstrate how to process jobs within a job module in an optimal fashion. Such a module, once ordered in this way, effectively becomes a chain. In this way, using these theorems is seen to be equivalent to applying a chain transformation. If by successive such applications of the theorems, the pair $(J, R)$ is reduced to a single chain of jobs, then this chain must evidently correspond to an optimal permutation.

However, as can easily be shown, not all pairs $(J, R)$ are chain reducible. The procedure for solving general problems that is described in this section utilizes in a helpful way what we know about ordering chain reducible pairs. Before presenting this procedure some extra terminology is required.

Suppose that $U$ is a subset of $J$ then $R|U$, the *restriction* of $R$ to $U$, is the set $R \cap \{(i \rightarrow j),\ i \in U,\ j \in U\}$. If $\alpha$ is a permutation of $U$ consistent with $R|U$ then $\alpha(U)$ is the ordered set obtained by the application of $\alpha$ to $U$. Further, $U$ is said to be *initial* if there does not exist $i \in J - U$, $j \in U$ such that $(i \rightarrow j) \in R$. An *initial permutation* $\alpha$ is a permutation for an initial subset $U$.

If $\alpha$ is a permutation for the subset $U$, then $\beta$ is a *sub-ordering* of $\alpha$ if and only if $\beta$ is a permutation for some subset $V \subseteq U$, and there exists a permutation for $U - V$, $\gamma$ say, such that $\alpha(U) = \{(\beta(V), \gamma(U - V)\}$. When this happens we write $\alpha = \{\beta, \gamma\}$.

The algorithm for solving the problem associated with a general pair $(J, R)$ with either discounted or linear costs is as follows:

STEP 1: Compute all the initial permutations $\alpha_i$, $1 \leqslant i \leqslant q$, say, with the following properties:

(i) If $\alpha_i$ is a permutation for initial subset $U_i$, then $(J - U_i,\ R|J - U_i)$ is chain reducible, $1 \leqslant i \leqslant q$;

(ii) $\alpha_i$ is *minimal* in the sense that it has no sub-ordering which satisfies (i), $1 \leqslant i \leqslant q$.

STEP 2: Compute an optimal permutation $\beta_i$ for the pair $(J - U_i R|J - U_i)$, $1 \leqslant i \leqslant q$. Since these pairs are all chain reducible such permutations may be obtained by the successive use of whichever is appropriate of Theorem 2 or Theorem 3.

STEP 3: Select that permutation for $J$ from among the $\{\alpha_i, \beta_i\}$, $1 \leqslant i \leqslant q$, which gives the smallest total cost. This is an optimal permutation.

It is clear that this procedure does yield an optimal permutation. Its computational efficiency in any problem will be dependent upon the nature of the pair $(J, R)$. In any problem with linear deferral costs the option of using Sidney's algorithm is also open to us. Sidney's algorithm successively chooses from $J$ initial subsets of a particular kind and operates as follows: In $r$ steps of the algorithm we choose subsets $U_i$, $1 \leqslant i \leqslant r$, subject to specified criteria where $U_j$ is initial in $\left\{J - \bigcup_{i=1}^{j-1} U_i,\ R|J - \bigcup_{i=1}^{j-1} U_i\right\}$ and where $\bigcup_{i=1}^{r} U_i = J$. If $\alpha_j$ is an optimal permutation for $U_j$, then $\{\alpha_1, \ldots, \alpha_r\}$ is optimal for $\{J, R\}$.

Note that it is possible to use Theorem 3 in conjunction with Sidney's algorithm in a problem with linear costs. Such an approach is exemplified in the next section.

## 6. AN EXAMPLE

The example discussed here is the one given by Sidney [5]. We seek an optimal permutation of the job set $J = \{1, 2, 3, 4, 5, 6, 7\}$. Each job $i$ has a linear cost function as in (2) with $C_i = 1$. The processing times are given by $\{X_1, X_2, X_3, X_4, X_5, X_6, X_7\} = \{5, 8, 3, 5, 3, 7, 6\}$ and the precedence relation $R$ is $\{(1 \to 3), (1 \to 4), (1 \to 6), (1 \to 7), (2 \to 4), (2 \to 5), (2 \to 6), (2 \to 7), (3 \to 7), (4 \to 6), (4 \to 7), (5 \to 6), (5 \to 7), (6 \to 7)\}$.

METHOD 1 (Glazebrook's algorithm): Adopt the procedure outlined in the previous section.

STEP 1: Initial permutations $\alpha_1$, $\alpha_2$ and $\alpha_3$ are chosen together with initial subsets $U_1$, $U_2$ and $U_3$, where $\alpha_1(U_1) = \{1\}$, $\alpha_2(U_2) = \{2, 1\}$ and $\alpha_3(U_3) = \{2, 5\}$.

STEP 2: Use Theorem 3 to compute $\beta_1$, $\beta_2$ and $\beta_3$. Consider, for example, the computation of $\beta_3$, an optimal permutation for $(J - \{2, 5\};\ R|J - \{2, 5\}) = (\{1, 3, 4, 6, 7\};\ \{(1 \to 3), (1 \to 4), (1 \to 6), (1 \to 7), (3 \to 7), (4 \to 6), (4 \to 7), (6 \to 7)\})$. Job set $\{3, 4, 6\}$ is a job module in this new problem since job 1 precedes all of jobs 3, 4 and 6, job 7 is preceded by them all and $\{3\}$ and $\{4, 6\}$ are disjoint chains. Theorem 3 applied to this new problem indicates that in an optimal permutation the jobs in the module will be processed in the order $\{3, 4, 6\}$ since

$$\frac{1}{3} = \frac{C_3}{X_3} > \max\left\{\frac{C_4}{X_4}, \frac{C_4 + C_6}{X_4 + X_6}\right\} = \frac{1}{5}.$$

Hence $\beta_3(J - U_3) = \{1, 3, 4, 6, 7\}$. Similarly $\beta_1(J - U_1) = \{3, 2, 5, 4, 6, 7\}$ and $\beta_2(J - U_2) = \{3, 5, 4, 6, 7\}$.

STEP 3: Of the permutations $\{\alpha_1, \beta_1\}$, $\{\alpha_2, \beta_2\}$ and $\{\alpha_3, \beta_3\}$ the first gives rise to the smallest total cost. Hence an optimal permutation for $(J,R)$ is given by $\{\alpha_1(U_1), \beta_1(J - U_1)\} = \{1, 3, 2, 5, 4, 6, 7\}$.

METHOD 2 (Sidney's algorithm): According to the criteria which Sidney sets out in [5], the subsets $U_1 = \{1, 3\}$, then $U_2 = \{2, 4, 5\}$ and then $U_3 = \{6, 7\}$ are chosen. $\gamma_1$, $\gamma_2$ and $\gamma_3$ are optimal permutations for $U_1$, $U_2$, and $U_3$ respectively. It can be shown that $\gamma_1(U_1) = \{1, 3\}$, $\gamma_2(U_2) = \{2, 5, 4\}$ and $\gamma_3(U_3) = \{6, 7\}$ so that an optimal permutation for $(J, R)$ is given by $\{1, 3, 2, 5, 4, 6, 7\}$.

METHOD 3 (Combination of Methods 1 and 2): Use Sidney's algorithm to choose the initial subset $U_1 = \{1, 3\}$ and note that its optimal permutation $\gamma_1$ is such that $\gamma_1(U_1) = \{1, 3\}$. $(J - \{1, 3\}, R|J - \{1, 3\})$ or $(\{2, 4, 5, 6, 7\}, \{(2 \rightarrow 4), (2 \rightarrow 5), (2 \rightarrow 6), (2 \rightarrow 7), (4 \rightarrow 6), (4 \rightarrow 7), (5 \rightarrow 6), (5 \rightarrow 7), (6 \rightarrow 7)\})$ is chain reducible with an optimal permutation $\delta$ which may be computed by using Theorem 3. It is found that $\delta(J - \{1, 3\}) = \{2, 5, 4, 6, 7\}$ and so an optimal permutation for $(J, R)$ is that implied by the ordered set $\{1, 3, 2, 5, 4, 6, 7\}$.

## ACKNOWLEDGMENTS

I should like to thank Drs. J.C. Gittins and P. Nash for many enjoyable and stimulating discussions on these and other matters.

## REFERENCES

[1] Banerjee, B.P., "Single Facility Sequencing with Random Execution Times," Operations Research *13*, 358-364 (1965).

[2] Glazebrook, K.D., "Stochastic Scheduling with Order Constraints," International Journal of System Science *7*, 657-666 (1976).

[3] Glazebrook, K.D. and J.C. Gittins "On Single-Machine Scheduling with Precedence Relations and Linear or Discounted Costs," (submitted) (1979).

[4] Meilijson, I. and G. Weiss, "Multiple Feedback at a Single Server Station." Stochastic Processes and Their Applications *5*, 195-205, (1977).

[5] Sidney, J.B., "Decomposition Algorithms for Single-Machine Sequencing with Precedence Relations and Deferral Costs," Operations Research *23*, 283-298 (1975).

# STATISTICAL ANALYSIS OF THE OUTPUT DATA FROM TERMINATING SIMULATIONS*

Averill M. Law

*University of Wisconsin*
*Madison, Wisconsin*

**ABSTRACT**

In this paper we precisely define the two types of simulations (terminating and steady-state) with regard to analysis of simulation output and discuss some common measures of performance for each type. In addition, we conclude, on the basis of discussions with many simulation practitioners, that both types of simulations are important in practice. This is contrary to the impression one gets from reading the simulation literature, where the steady-state case is almost exclusively considered.

Although analyses of terminating simulations are considerably easier than are those of steady-state simulations, they have not received a careful treatment in the literature. We discuss and give empirical results for fixed sample size, relative width, and absolute width procedures that can be used for constructing confidence intervals for measures of performance in the terminating case.

## . TYPES OF SIMULATIONS WITH REGARD TO ANALYSIS OF THE OUTPUT

We begin by giving a precise definition of the two types of simulations with regard to nalysis of the output data (cf. Gafarian and Ancker [4] and Kleijnen [6]). A *terminating simu-ation* is one for which the quantities of interest are defined relative to the interval of simulated ime $[0, T_E]$, where $T_E$, a possibly degenerate random variable (r.v.), is the time that a specified vent $E$ occurs. The following are some examples of terminating simulations:

) Consider a retail establishment (e.g., a bank) which closes each evening (physically terminating). If the establishment is open from 9 to 5, then the objective of a simulation might be to estimate some measure of the quality of customer service over the period beginning at 9 and ending when the last customer who entered before the doors closed at 5 has been served. In this case $E$ ={at least 8 hours of simulated time have elapsed and the system is empty}, and reasonable initial conditions for the simulation might be that no customers are present at time 0.

) Consider a telephone exchange which is always open (physically nonterminating). The objective of a simulation might be to determine the number of (permanent) telephone lines needed to service adequately incoming calls. Since the arrival rate of calls changes with the time of day, day of the week, etc., it is unlikely that a steady-state measure of

This research was supported by the Office of Naval Research under contract N00014-76-C-0403 (NR 047-145) and the rmy Research Office under contract DAAG29-75-C-0024.

performance (see Section 2), which is defined as a limit as time goes to infinity, will exist. A common objective in this case is to study the system during the period of peak loading, say, of length $t$ hours, since the number of lines sufficient for this period will also do for the rest of the day. In this case, $E = \{t$ hours of simulated time have elapsed$\}$. However, care must be taken in choosing the number of waiting calls at time 0, since the actual system will probably be quite congested at the beginning of the period of peak loading. One approach would be to collect data from the actual system on the distribution of the number of waiting calls at the beginning of the peak period. Then for each simulation run of length $t$ hours, a sample is generated from this distribution and used as the number of waiting calls at time 0.

c) Consider a military confrontation between a defensive (fixed position) blue force and an offensive (attacking) red force. Relative to some initial force strengths, the objective of a simulation might be to estimate some function of the (final) force strengths at the time that the red force moves to within a certain specified distance from the blue force. In this case, $E =$ {red force has moved to within a certain specified distance from the blue force}.

A *steady-state simulation* is one for which the quantities of interest are defined as limits as the length of the simulation goes to infinity. Since there is no natural event to terminate the simulation, the length of the simulation is made large enough to get "good" estimates of the quantities of interest. Alternatively, the length of the simulation could be determined by cost considerations; however, this may not produce acceptable results. The following are some examples of steady-state simulations:

a) Consider a computer manufacturer who constructs a simulation model of a proposed computer system. Rather than use data from the arrival process of an existing computer system as input to the model, he typically assumes that jobs arrive in accordance with a Poisson process with rate equal to the predicted arrival rate of jobs during the period of peak loading. He is interested in estimating the response time of a job after the system has been running long enough so that initial conditions (e.g., the number of jobs in the system at time 0) no longer have any effect. (Assuming that the arrival rate is constant over time allows steady-state measures to exist.)

b) A chemical manufacturer constructs a simulation model of a proposed chemical process operation. The process, when in operation, will be subject to randomly occurring breakdowns. The input rate of raw materials to the process and the controllable parameters of the process are both assumed to be stationary with respect to time. The company would like to estimate the production rate after the process has been running long enough so that initial conditions no longer have any effect.

The remainder of this paper is organized as follows. In Section 2 we discuss some common measures of performance for both types of simulations and in Section 3 we present our findings on the relative occurrence of each type in the real world. A number of procedures which can be used to construct confidence intervals for terminating simulations are discussed in Section 4 and, finally in Section 5 we summarize our findings.

## 2. MEASURES OF SYSTEM PERFORMANCE

To the best of our knowledge, nowhere in the simulation literature are measures of performance for terminating simulations explicitly defined. In this section we define and contrast several common measures of performance for terminating and steady-state simulations by

means of examples. (Because of the diversity of terminating simulations, it is not possible to give one definition that fits all cases.) *For the simple examples that we consider, it is possible to compute analytically measures of performance. This fact will be useful in Section 4 where we study stopping rules for terminating simulations.*

### A. Averages

Consider first the stochastic process $\{D_i, i \geq 1\}$ for the M/M/1 queue with $\rho < 1$, where $D_i$ is the delay in queue of the $i$ th customer. The objective of a terminating simulation of the M/M/1 queue might be to estimate the *expected average delay of the first m customers* given some initial conditions, say, that the number of customers in the system at time 0, $N(0)$, is zero. The desired quantity, which we denote by $d(m|N(0) = 0)$, is then given by

$$d(m|N(0) = 0) = E\left[\sum_{i=1}^{m} D_i/m \,|N(0) = 0\right]$$

$$= \sum_{i=1}^{m} E[D_i|N(0) = 0]/m.$$

(Alternatively, if one is interested in estimating the expected average delay of all customers who arrive and complete their delay in the time interval $[0, t]$, then the desired quantity is given by

$$d(t|N(0) = 0) = E\left[\sum_{i=1}^{M(t)} D_i/M(t)|N(0)=0\right],$$

where $M(t)$ (a r.v.) is the number of customers who arrive and complete their delay in the interval $[0, t]$. Note that in this case the expectation and summation are not interchangeable. Thus, the label "expected average delay" is more general than "average expected delay.") Note also that the quantity $d(m|N(0) = 0)$, which is often called a transient characteristic of the stochastic process $\{D_i, i \geq 1\}$, explicitly depends on the state of the system at time 0; i.e., $d(m|N(0) = i) \neq d(m|N(0) = j)$ for $i \neq j$.

The objective of a steady-state simulation of $\{D_i, i \geq 1\}$ for the M/M/1 queue would be to estimate the *steady-state expected average delay d*, which is given by

$$d = \lim_{m\to\infty} d(m|N(0) = i) \quad \text{for any } i = 0, 1, \ldots .$$

(Under our assumption that $\rho < 1$, it can be shown that $d$ exists.) Observe that $d$ is independent of $N(0)$. In Figure 1 we plot $d(m|N(0) = 0)$ as a function of $m$. (The arrival rate $\lambda = 1$ and the service rate $\mu = 10/9$, so $\rho = .9$.) The horizontal line that $d(m|N(0)=0)$ asymptotically approaches is at height $d = 8.1$.

As a second example consider the stochastic process $\{E_i, i \geq 1\}$ for an $(s, S)$ inventory system with zero delivery lag and backlogging, where $E_i$ is the expenditure in the $i$ th period. This system is described in detail in Law [8]. A possible objective of a terminating simulation would be to estimate the *expected average cost for the first m periods* given that the inventory level at the beginning of period 1, $I_1$, is $S$:

$$e(m|I_1 = S) = E\left[\sum_{i=1}^{m} E_i/m \,|I_1 = S\right].$$

The objective of a steady-state simulation of $\{E_i, i \geq 1\}$ would be to estimate the *steady-state expected average cost:*

$$e = \lim_{m\to\infty} e(m|I_1 = i) \quad \text{for any } i = 0, \pm 1, \pm 2, \ldots .$$

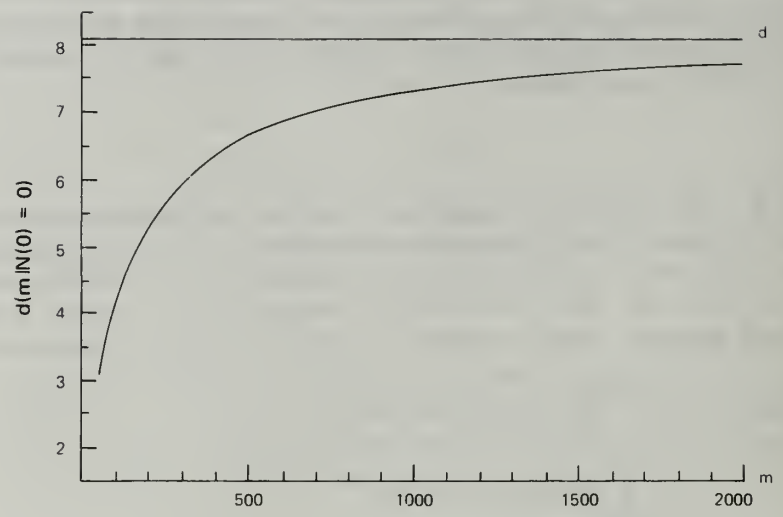


FIGURE 1. $d(m|N(0)=0)$ as a function of $m$ for the M/M/1 queue with $\rho = 0.9$

(It can be shown that $e$ exists.) We plot $e(m|I_1 = S)$ as a function of $m$ and also $e = 112.108$ in Figure 2.

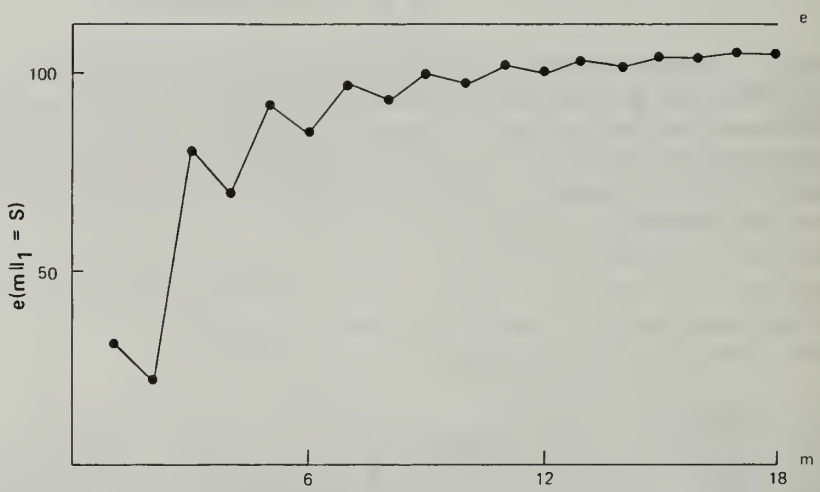


FIGURE 2. $e(m|I_1=S)$ as a function of $m$ for the $(s,S)$ inventory system

Our third example is quite different from the first two. A reliability model consisting of three components will function as long as component 1 works and either component 2 or 3 works. If $T$ is the time to failure of the whole system and $T_i$ the time to failure of component $i (i = 1, 2, 3)$, then

$$T = \min[T_1, \max(T_2, T_3)].$$

We further assume that the $T_i$'s are independent r.v.'s and each $T_i$ has a Weibull distribution with shape parameter $\alpha = .5$ and scale parameter $\beta = 1$. (A distributional assumption for the $T_i$'s is needed in Section 4 where we present simulation results for this model.) The objective of a terminating simulation might be to estimate the *expected time to failure of the system* given that all components are new, $E(T|\text{all components are new})$. If we assume that the system is not repaired when it fails, then a steady-state simulation makes no sense for this system. Such could be the case if this system were part of a space probe.

### B. Proportions

The usual criterion for comparing two or more systems is some sort of average behavior. However, different kinds of information may be of more value in some situations. For example, a bank manager may be concerned with estimating the proportion of customers who experience a delay in excess of 5 minutes. Since proportions are really just a special case of averages, we will illustrate them by means of the M/M/1 example.

In a terminating simulation of the M/M/1 queue the objective might be to estimate, instead of an expected average delay, the *expected proportion of the first m customers whose delay is less than or equal to x* (a specified number) given that $N(0) = 0$. Denote the desired quantity by $P(m, x|N(0) = 0)$ and let

$$Y_i(x) = \begin{cases} 1 & \text{if } D_i \leqslant x \\ 0 & \text{if } D_i > x \end{cases} \text{ for } i = 1, 2, \ldots .$$

Then $P(m, x|N(0) = 0)$ is given by

$$P(m, x|N(0) = 0) = E\left[\sum_{i=1}^{m} Y_i(x)/m \,|N(0) = 0\right].$$

For a steady-state simulation, the objective would be to estimate the *steady-state expected proportion of customers whose delay is less than or equal to x :*

$$P(x) = \lim_{m \to \infty} P(m, x|N(0) = i) \quad \text{for any } i = 0, 1, \ldots .$$

## . RELATIVE IMPORTANCE OF EACH TYPE OF SIMULATION

Reading the simulation literature leads one to think that only steady-state simulations are important; almost every paper written on the analysis of simulation output data deals with the steady-state case. This may be a carry-over from mathematical queueing theory where only a steady-state analysis is generally possible. However, we believe that terminating simulations are also important. We have discovered, by talking to a large number of simulation practitioners, that a significant proportion of simulations in the real world are actually of the terminating type. The following are some reasons why a steady-state analysis may not be appropriate:

- The system under consideration is physically terminating. In this case, letting the length of a simulation be arbitrarily large makes no sense.
- The input distributions for the system change over time. In this case, steady-state measures of performance will probably not exist.
- One is often interested in studying the transient behavior of a system even if steady-state measures of performance exist.

## STOPPING RULES FOR TERMINATING SIMULATIONS

In the following three subsections we consider procedures that can be used to construct confidence intervals (c.i.'s) for measures of performance for terminating simulations. We will not consider the steady-state case since it has been widely discussed in the simulation literature. For surveys of fixed sample size and sequential procedures that can be used to construct c.i.'s for steady-state measures of performance, see Law and Kelton [9, 10]. The random numbers used in the remainder of this paper were generated from the generator discussed in [8].

### A. Fixed Sample Size Procedures

Suppose we make $n$ independent replications of a terminating simulation. The independence among replications is accomplished by using different random numbers for each replication and by starting each one with the same initial conditions. If $X_i$ is the estimator of interest from the $i$ th replication ($i = 1, 2, \ldots, n$), then the $X_i$'s are independent identically distributed (i.i.d.) r.v's. (For the M/M/1 queue, $X_i$ might be the average $\sum_{j=1}^{m} D_j/m$ or the proportion $\sum_{j=1}^{m} Y_j(x)/m$.) If, in addition, the $X_i$'s are normally distributed, then a $100(1-\alpha)\%$ $(0 < \alpha < 1)$ c.i. for $\mu = E(X)$ is given by

$$\bar{X}(n) \pm t_{n-1,\,1-\alpha/2}\sqrt{s^2(n)/n}, \tag{1}$$

where $\bar{X}(n)$ and $s^2(n)$ are the usual sample mean and variance, respectively, and $t_{n-1,\,1-\alpha/2}$ is the $1 - \alpha/2$ point for a $t$ distribution with $n$-1 degrees of freedom.

In practice the $X_i$'s will not be normally distributed and the c.i. given by (1) will be only approximate. To investigate the effect of nonnormality, we simulated the three stochastic models of Section 2. For the M/M/1 queue with $\rho = .9$, the $(s,S)$ inventory system, and the reliability model, respectively, the quantities of interest were $d(25|N(0) = 0) = 2.12$, $e(12|I_1 = S) = 99.52$, and $E(T|\text{all components are new}) = .778$. (See [8] for a discussion of how to compute the first two quantities.) For each model we performed 500 independent simulation experiments, for each experiment we considered $n = 5, 10, 20, 40$, and for each $n$ we used (1) to construct a 90% c.i. for the desired quantity. In Tables 1, 2, and 3 we give the proportion, $\hat{p}$, of the 500 c.i.'s that covered the desired quantity, a 90% c.i. for the true coverage and the average value of the c.i. half length divided by the point estimate over the 500 experiments for the three models. The 90% c.i. for the true coverage was computed from

$$\hat{p} \pm 1.645\sqrt{\hat{p}(1-\hat{p})/500}.$$

Observe that for the M/M/1 queue and the $(s,S)$ inventory system the coverages are quite close to 90%, but for the reliability model there is a significant degradation in coverage apparently caused by a severe departure from normality. To see if this is indeed the case, we generated 1000 $X_i$'s for each stochastic model and estimated the skewness and kurtosis. These estimates, which are presented in Table 4, indicate that the $X_i$'s for the reliability model are considerably more nonnormal than are those for the other two models. This conclusion was reinforced by plotting histograms for the three sets of data.

TABLE 1. *Fixed Sample Size Results for $d(25)|N(0)=0) = 2.12$, M/M/1 Queue with $\rho = .9$*

| $n$ | coverage | average of $\frac{\text{c.i. half length}}{\bar{X}(n)}$ |
|---|---|---|
| 5 | .880 ± .024 | .672 |
| 10 | .864 ± .025 | .436 |
| 20 | .886 ± .023 | .301 |
| 40 | .914 ± .021 | .212 |

TABLE 2. *Fixed Sample Size Results for* $e(12|I_1 = S) = 99.52$, *(s,S) Inventory System*

| $n$ | coverage | average of $\frac{\text{c.i. half length}}{\bar{X}(n)}$ |
|---|---|---|
| 5 | .908 ± .021 | .048 |
| 10 | .904 ± .022 | .031 |
| 20 | .880 ± .024 | .021 |
| 40 | .894 ± .023 | .014 |

TABLE 3. *Fixed Sample Size Results for* $E(T|$*all components new*$) = .778$, *Reliability Model*

| $n$ | coverage | average of $\frac{\text{c.i. half length}}{\bar{X}(n)}$ |
|---|---|---|
| 5 | .708 ± .033 | 1.163 |
| 10 | .750 ± .032 | .820 |
| 20 | .800 ± .029 | .600 |
| 40 | .840 ± .027 | .444 |

TABLE 4. *Skewness and Kurtosis for the Three Stochastic Models and the Normal Distribution*

| Stochastic Model or Distribution | Skewness | Kurtosis |
|---|---|---|
| Normal Distribution | 0* | 3* |
| M/M/1 Queue | 1.66 | 6.43 |
| (*s,S*) Inventory System | .45 | 3.76 |
| Reliability Model | 5.18 | 54.39 |

*Theoretical Values

**B. Relative Width Procedures**

One disadvantage of the fixed sample size approach to constructing a c.i. is that the simulator has no control over the c.i. half length; for fixed $n$, the half length will depend on the population variance $\sigma^2 = Var(X)$. In this subsection we consider two sequential procedures which allow one to specify the "relative precision" of a c.i. Both assume that $X_1, X_2, \ldots$ is a sequence of i.i.d. r.v.'s which need not be normal.

The first procedure has been suggested for use in several different contexts; see Iglehart [5], Lavenberg and Sauer [7], and Thomas [13]. The objective of the procedure is to construct a $100(1-\alpha)\%$ c.i. for $\mu$ such that the difference between the point estimator $\bar{X}(n)$ and $\mu$ is no more than $100\,\gamma\%$ of $\bar{X}(n)$, that is,

$$|\bar{X}(n) - \mu| \leq \gamma|\bar{X}(n)| \quad \text{for } 0 < \gamma < 1. \tag{2}$$

Choose an initial sample size $n_0 \geq 2$, let

$$\delta_{r,1}(n,\alpha) = t_{n-1,1-\alpha/2}\sqrt{s^2(n)/n},$$

and let

$$N_{r,1}(\gamma,\alpha) = \min\left\{n: n \geq n_0,\ s^2(n) > 0,\ \frac{\delta_{r,1}(n,\alpha)}{|\bar{X}(n)|} \leq \gamma\right\}. \tag{3}$$

(Note that $N_{r,1}(\gamma, \alpha)$, which is the required number of replications, is a r.v.). Then use

$$I_{r,1}(\gamma,\alpha) = [\bar{X}(N_{r,1}(\gamma,\alpha)) - \delta_{r,1}(N_{r,1}(\gamma,\alpha),\alpha), \bar{X}(N_{r,1}(\gamma,\alpha)) + \delta_{r,1}(N_{r,1}(\gamma,\alpha),\alpha)] \tag{4}$$

as an approximate 100(1-$\alpha$)% c.i. for $\mu$. It easily follows from (3) and (4) that $I_{r,1}(\gamma,\alpha)$ satisfies the criterion given by (2). Furthermore, using an argument similar to the one employed by Lavenberg and Sauer in the context of the regenerative method for steady-state simulations, we have been able to prove the following theorem.

THEOREM 1. If $\mu \neq 0$ and $0 < \sigma^2 < \infty$, then $\lim_{\gamma \to 0^+} P\{\mu \in I_{r,1}(\gamma,\alpha)\} = 1 - \alpha$.

The objective of the second procedure, which is due to Nadas [11], is to construct a c.i. such that

$$|\bar{X}(n)-\mu| \leq \gamma\,|\mu| \quad \text{for } 0 < \gamma < 1. \tag{5}$$

Let

$$v^2(n) = \left\{1 + \sum_{i=1}^{n} [X_i - \bar{X}(n)]^2\right\}/n = (1/n) + (n-1)s^2(n)/n,$$

$$\delta_{r,2}(n,\alpha) = t_{n-1,1-\alpha/2}\sqrt{v^2(n)/n},$$

and

$$N_{r,2}(\gamma,\alpha) = \min\left\{n: n \geq n_0,\ \frac{\delta_{r,2}(n,\alpha)}{|\bar{X}(n)|} \leq \gamma\right\}.$$

Then use

$$I_{r,2}(\gamma,\alpha) = \left[\frac{\bar{X}(N_{r,2}(\gamma,\alpha))}{1+\gamma},\ \frac{\bar{X}(N_{r,2}(\gamma,\alpha))}{1-\gamma}\right] \tag{6}$$

as an approximate 100(1-$\alpha$)% c.i. for $\mu$. From (6) it is easy to show that $I_{r,2}(\gamma,\alpha)$ satisfies the criterion given by (5). Furthermore, the following theorem was proved by Nadas.

THEOREM 2. If $\mu \neq 0$ and $0 < \sigma^2 < \infty$, then $\lim_{\gamma \to 0^+} P\{u \in I_{r,2}(\gamma,\alpha)\} = 1 - \alpha$.

In order to compare the two procedures and to determine the effect of non-infinitesimal $\gamma$ on coverage, we once again simulated the three stochastic models. For each model we performed 500 independent experiments, for the M/M/1 queue and the reliability model we considered $\gamma$ = .2, .1, .05 for each experiment, and for the inventory system we considered $\gamma$ = .2, .1, .05, .025, .0125, .00625 for each experiment. In all cases, $n_0 = 5$. In Tables 5, 6, and we give point estimates and 90% c.i.'s for the true coverages, point estimates and 90% c.i.'s for $E\{N_{r,i}(\gamma,\alpha)\}$ ($i$ = 1, 2), and the average c.i. half lengths over the 500 experiments. We considered more values of $\gamma$ for the inventory system because it appeared from our empirical

TABLE 5. *Relative Width Results for $d(25|N(0) = 0) = 2.12$, M/M/1 Queue with $\rho = .9$*

| | Procedure 1 | | | Procedure 2 | | |
|---|---|---|---|---|---|---|
| $\gamma$ | $E\{N_{r,1}(\gamma,\alpha)\}$ | coverage | average c.i. half length | $E\{N_{r,2}(\gamma,\alpha)\}$ | coverage | average c.i. half length |
| .2 | 42.3 ± 0.9 | .842 ± .027 | .414 | 41.9 ± 0.8 | .862 ± .025 | .437 |
| .1 | 175.2 ± 1.7 | .860 ± .026 | .211 | 174.5 ± 1.7 | .868 ± .025 | .213 |
| .05 | 704.4 ± 3.5 | .884 ± .024 | .106 | 703.7 ± 3.5 | .882 ± .024 | .106 |

TABLE 6. *Relative Width Results for $e(12|I_1 = S) = 99.52$, $(s, S)$ Inventory System*

| | Procedure 1 | | | Procedure 2 | | |
|---|---|---|---|---|---|---|
| $\gamma$ | $E\{N_{r,1}(\gamma,\alpha)\}$ | coverage | average c.i. half length | $E\{N_{r,2}(\gamma,\alpha)\}$ | coverage | average c.i. half length |
| .2 | 5.0 ± 0.0 | .902 ± .022 | 4.89 | 5.0 ± 0.0 | 1.0 | 20.74 |
| .1 | 5.0 ± 0.0 | .902 ± .022 | 4.86 | 5.0 ± 0.0 | 1.0 | 10.06 |
| .05 | 5.9 ± 0.1 | .892 ± .023 | 3.97 | 5.7 ± 0.1 | .962 ± .014 | 4.99 |
| .025 | 13.3 ± 0.4 | .834 ± .027 | 2.35 | 12.3 ± 0.4 | .858 ± .026 | 2.48 |
| .0125 | 51.0 ± 1.0 | .856 ± .026 | 1.23 | 49.8 ± 1.0 | .862 ± .025 | 1.24 |
| .00625 | 206.3 ± 1.8 | .872 ± .025 | 0.62 | 205.4 ± 1.8 | .876 ± .024 | 0.62 |

TABLE 7. *Relative Width Results for $E(T|$all components new$) = .778$, Reliability Model*

| | Procedure 1 | | | Procedure 2 | | |
|---|---|---|---|---|---|---|
| $\gamma$ | $E\{N_{r,1}(\gamma,\alpha)\}$ | coverage | average c.i. half length | $E\{N_{r,2}(\gamma,\alpha)\}$ | coverage | average c.i. half length |
| .2 | 213.7 ± 4.5 | .876 ± .024 | .152 | 214.1 ± 4.5 | .908 ± .021 | .160 |
| .1 | 907.4 ± 11.2 | .898 ± .022 | .077 | 908.6 ± 10.8 | .902 ± .022 | .078 |
| .05 | 3720.5 ± 23.7 | .882 ± .024 | .039 | 3720.0 ± 23.7 | .884 ± .024 | .039 |

sults that a smaller $\gamma$ is required for the coverage ultimately to converge to the desired level. A smaller $\gamma$ is required for this model to get a large value of $N_{r,i}(\gamma,\alpha)$.) Note also that convergence of coverage does not appear to be monotone.

We repeated the above 500 experiments using the same random numbers and $n_0 = 2$. For procedure 2 the results were identical; however, for procedure 1 there was a significant degradation in coverage due to premature stopping on replications 2, 3, or 4. For example, the coverage for the M/M/1 queue with $\gamma = .2$ was .798.

**Absolute Width Procedures**

In this subsection we present two procedures which allow one to construct a $100(1-\alpha)\%$ for $\mu$ such that

$$|\bar{X}(n) - \mu| \leq c,$$

where $c$ is a specified positive number.

The first procedure, which is due to Chow and Robbins [1], assumes that $X_1, X_2, \ldots$ is a sequence of i.i.d. r.v's. Choose $n_0 \geqslant 2$. Let $v^2(n)$ be defined as in Subsection 4.B, let

$$N_{a,1}(c, \alpha) = \min\left\{n : n \geqslant n_0,\ v^2(n) \leqslant \frac{c^2 n}{(t_{n-1,1-\alpha/2})^2}\right\},$$

and then use

$$I_{a,1}(c, \alpha) = [\bar{X}(N_{a,1}(c, \alpha)) - c,\ \bar{X}(N_{a,1}(c, \alpha)) + c]$$

as an approximate $100(1-\alpha)\%$ c.i. for $\mu$. It is clear that $I_{a,1}(c,\alpha)$ satisfies the criterion given by (7). The following theorem was proved by Chow and Robbins.

THEOREM 3. If $0 < \sigma^2 < \infty$, then $\lim_{c \to 0^+} P\{\mu \in I_{a,1}(c,\alpha)\} = 1 - \alpha$. For an empirical evaluation of the above procedure under the assumption that the $X_i$'s are normal, see Starr [12].

The second procedure, which is due to Dudewicz [2], assumes that the $X_i$'s are i.i.d *normal* r.v.'s. Initially make $n_0 (n_0 \geqslant 2)$ replications of the simulation and compute $\bar{X}(n_0)$ and $s^2(n_0)$. Let

$$N_{a,2}(c,\alpha) = \max\{n_0 + 1, \lceil w^2 s^2(n_0)\rceil\},$$

where $w = t_{n_0-1,1-\alpha/2}/c$ and $\lceil z\rceil$ is the smallest integer $\geqslant z$. Make $N_{a,2}(c,\alpha) - n_0$ additional replications of the simulation, let

$$\bar{Y}(N_{a,2}(c,\alpha) - n_0) = \sum_{i=n_0+1}^{N_{a,2}(c,\alpha)} X_i/(N_{a,2}(c,\alpha) - n_0),$$

and let $\tilde{\bar{X}}(N_{a,2}(c,\alpha)) = a_1\bar{X}(n_0) + a_2\bar{Y}(N_{a,2}(c,\alpha) - n_0)$, where

$$a_1 = \frac{n_0}{N_{a,2}(c,\alpha)}\left[1 + \sqrt{1 - \frac{N_{a,2}(c,\alpha)}{n_0}\left[1 - \frac{N_{a,2}(c,\alpha) - n_0}{w^2 s^2(n_0)}\right]}\right]$$

and $a_2 = 1 - a_1$. Then use

$$I_{a,2}(c,\alpha) = [\tilde{\bar{X}}(N_{a,2}(c,\alpha)) - c,\ \tilde{\bar{X}}(N_{a,2}(c,\alpha)) + c]$$

as an approximate $100(1-\alpha)\%$ c.i. for $\mu$. Dudewicz has proved the following theorem.

THEOREM 4. $P\{\mu \in I_{a,2}(c,\alpha)\} = 1 - \alpha$ for all $c > 0$.

To compare the sequential procedure of Chow and Robbins and the two-stage procedure of Dudewicz, we performed 500 independent experiments for each model. To make the absolute width results somewhat comparable to the relative width results, we chose the values of $c$ to correspond to the values of $\gamma$; that is, for each $\gamma$ we chose $c = \gamma\mu$. For the Chow and Robbins procedure we chose $n_0 = 5$ and for the Dudewicz procedure we considered $n_0 = 15$, 30 and 60. (Dudewicz [3] recommended that $n_0$ be at least 12.) The results of the simulation experiments for the three models are given in Tables 8, 9, 10.

TABLE 8. *Absolute Width Results for $d(25|N(0) = 0) = 2.12$, M/M/1 Queue with $\rho = .9$*

| | Chow and Robbins | | Dudewicz | | |
|---|---|---|---|---|---|
| $c$ | $E\{N_{a,1}(c,\alpha)\}$ | coverage | $n_0$ | $E\{N_{a,2}(c,\alpha)\}$ | coverage |
| .425 | 38.0 ± 1.2 | .800 ± .029 | 15 | 49.9 ± 2.1 | .850 ± .026 |
| | | | 30 | 48.2 ± 1.3 | .912 ± .020 |
| | | | 60 | 62.1 ± 0.4 | .926 ± .019 |
| .212 | 173.5 ± 2.5 | .898 ± .022 | 15 | 196.9 ± 8.5 | .854 ± .026 |
| | | | 30 | 185.7 ± 5.6 | .888 ± .023 |
| | | | 60 | 182.9 ± 4.0 | .894 ± .023 |
| .106 | 706.8 ± 4.8 | .906 ± .021 | 15 | 786.1 ± 34.2 | .868 ± .025 |
| | | | 30 | 741.1 ± 22.6 | .878 ± .024 |
| | | | 60 | 730.2 ± 15.7 | .898 ± .022 |

TABLE 9. *Absolute Width Results for $e(12|I_1 = S) = 99.52$, (s,S) Inventory System*

| | Chow and Robbins | | Dudewicz | | |
|---|---|---|---|---|---|
| $c$ | $E\{N_{a,1}(c,\alpha)\}$ | coverage | $n_0$ | $E\{N_{a,2}(c,\alpha)\}$ | coverage |
| 19.90 | 5.0 ± 0.0 | 1.0 | 15 | 16.0 ± 0.0 | .936 ± .018 |
| | | | 30 | 31.0 ± 0.0 | .878 ± .024 |
| | | | 60 | 61.0 ± 0.0 | .888 ± .023 |
| 9.95 | 5.0 ± 0.0 | 1.0 | 15 | 16.0 ± 0.0 | .936 ± .018 |
| | | | 30 | 31.0 ± 0.0 | .880 ± .024 |
| | | | 60 | 61.0 ± 0.0 | .890 ± .023 |
| 4.98 | 5.7 ± 0.1 | .976 ± .011 | 15 | 16.0 ± 0.0 | .922 ± .020 |
| | | | 30 | 31.0 ± 0.0 | .882 ± .024 |
| | | | 60 | 61.0 ± 0.0 | .886 ± .023 |
| 2.49 | 12.3 ± 0.4 | .880 ± .024 | 15 | 18.5 ± 0.3 | .908 ± .021 |
| | | | 30 | 31.0 ± 0.0 | .894 ± .023 |
| | | | 60 | 61.0 ± 0.0 | .882 ± .024 |
| 1.24 | 48.3 ± 1.1 | .872 ± .025 | 15 | 60.8 ± 2.0 | .904 ± .022 |
| | | | 30 | 55.0 ± 1.3 | .912 ± .020 |
| | | | 60 | 62.9 ± 0.4 | .902 ± .022 |
| 0.62 | 204.4 ± 1.8 | .896 ± .022 | 15 | 241.7 ± 8.1 | .912 ± .020 |
| | | | 30 | 217.9 ± 5.1 | .898 ± .022 |
| | | | 60 | 211.5 ± 3.4 | .912 ± .020 |

## 5. SUMMARY AND CONCLUSIONS

We have defined terminating and steady-state simulations and have discussed some common measures of performance for each type. In addition, we have concluded from talking with simulation practitioners that a significant proportion of real-world simulations are of the terminating type. This is fortunate because it means that classical statistical analysis for i.i.d. observations (e.g., confidence intervals, hypothesis testing, ranking and selection, etc.) is applicable to analyzing many simulations. On the other hand, in the steady-state case there is still not a totally acceptable procedure even for the relatively simple problem of constructing a c.i. for a steady-state expected average.

TABLE 10. *Absolute Width Results for $E(T|$all components new$) = .778$, Reliability Model*

| | Chow and Robbins | | Dudewicz | | |
|---|---|---|---|---|---|
| $c$ | $E\{N_{a,1}(c,\alpha)\}$ | coverage | $n_0$ | $E\{N_{a,2}(c,\alpha)\}$ | coverage |
| .156 | 179.5 ± 7.0 | .774 ± .031 | 15 | 246.0 ± 27.2 | .704 ± .034 |
| | | | 30 | 220.8 ± 17.0 | .772 ± .031 |
| | | | 60 | 231.7 ± 14.9 | .812 ± .029 |
| .078 | 888.0 ± 14.5 | .900 ± .022 | 15 | 981.8 ± 109.0 | .728 ± .033 |
| | | | 30 | 880.6 ± 68.0 | .794 ± .030 |
| | | | 60 | 922.2 ± 59.6 | .838 ± .027 |
| .039 | 3672.1 ± 32.9 | .884 ± .024 | 15 | 3925.6 ± 435.8 | .772 ± .031 |
| | | | 30 | 3520.9 ± 272.0 | .788 ± .030 |
| | | | 60 | 3687.2 ± 238.5 | .832 ± .028 |

We have also considered procedures for constructing c.i.'s for terminating simulations. If one is performing an exploratory experiment where precision of the c.i. may not be overwhelmingly important, then we recommend using a fixed sample size procedure. However, if the $X_i$'s are highly nonnormal and if the number of replications $n$ is too small, then the actual coverage of the constructed c.i. may be considerably lower than that desired (see Table 3).

If one wants a c.i. having half length that is small relative to the point estimate, then a relative width procedure may be used. We recommend using Procedure 2 (due to Nadas) with $n_0 \geqslant 5$. Procedure 2 appears to give slightly better coverage; its criterion (see (5)) is more intuitive than the criterion of Procedure 1 (see (2)), and Procedure 2 does not seem subject to premature stopping even for $n_0 = 2$. (On the other hand, Procedure 1 uses a more intuitive expression to construct a c.i.)

If one wants a c.i. for which the half length is a specified number, then an absolute width procedure may be used. We recommend using the Chow and Robbins procedure with $n_0 \geqslant 5$. Their procedure generally requires a smaller average sample size, the variance of the sample size is smaller, and its coverage seems to be less affected by departures from normality (see Table 10).

In general, we believe that relative width procedures are more useful than absolute width procedures due to the difficulty in specifying the absolute width $c$ for most simulation experiments. When using either the Nadas procedure or the Chow and Robbins procedure, we believe that it is advisable to choose a $\gamma$ or $c$ which will cause the procedure to run until the sample is at least of moderate size; perhaps, at least 30. (Since both procedures are based on the central limit theorem, it is unreasonable to think that they will work well in general for a small sample size; see the results for $\gamma = .025$ in Table 6.) Note that precise c.i.'s may be unaffordable in the real world due to the high cost of making a single replication.

Our conclusions on the efficacy of the procedures are based on only three models and, thus, should be considered tentative. However, since the performance of a procedure depends only on the distribution of an $X_i$ and not on the complexity of a model, we feel that there is no particular reason to suspect that the results for real-world models should differ significantly from those for the simple models presented here.

## ACKNOWLEDGMENTS

The author would sincerely like to thank David Kelton for his programming assistance and for his reading of the manuscript. The author is also grateful to Stephen Lavenberg (IBM

/atson Research Center), Richard Smith (IBM Boulder), Charles White (Dupont), U. A. /eber (British Columbia Telephone Company), and Barney Watson (U. S. Army TRASANA) )r their discussions of simulation at their organizations.

## REFERENCES

1] Chow, Y.S. and H. Robbins, "On the Asymptotic Theory of Fixed-Width Sequential Confidence Intervals for the Mean," Annals of Mathematical Statistics *36*, 457-462 (1965).

2] Dudewicz, E.J., "Statistical Inference with Unknown and Unequal Variances," Transactions of the Annual Quality Control Conference of the Rochester Society for Quality Control *28*, 71-85 (1972).

3] Dudewicz, E.J., Personal Communication (1977).

4] Gafarian, A.V. and C.J. Ancker, Jr., "Mean Value Estimation from Digital Computer Simulation," Operations Research *14*, 25-44 (1966).

5] Iglehart, D.L., "The Regenerative Method for Simulation Analysis," Technical Report No. 86-20, Control Analysis Corporation, Palo Alto, California (1975).

6] Kleijnen, J.P.C., "The Statistical Design and Analysis of Digital Simulation: A Survey," Management Informatics *1*, 57-66 (1972).

7] Lavenberg, S.S. and C.H. Sauer, "Sequential Stopping Rules for the Regenerative Method of Simulation," IBM Journal of Research and Development *21*, 545-558 (1977).

8] Law, A.M., "Confidence Intervals in Discrete Event Simulation: A Comparison of Replication and Batch Means," Naval Research Logistics Quarterly *24*, 667-678 (1977).

9] Law, A.M. and W.D. Kelton, "Confidence Intervals for Steady-State Simulations, I: A Survey of Fixed Sample Size Procedures," Technical Report 78-5, Dept. of Industrial Engineering, University of Wisconsin (1978).

0] Law, A.M. and W.D. Kelton, "Confidence Intervals for Steady-State Simulations, II: A Survey of Sequential Procedures," Technical Report 78-6, Dept. of Industrial Engineering, University of Wisconsin (1978).

1] Nadas, A., "An Extension of a Theorem of Chow and Robbins on Sequential Confidence Intervals for the Mean," Annals of Mathematic Statistics *40*, 667-671 (1969).

2] Starr, N., "The Performance of a Sequential Procedure for the Fixed-Width Interval Estimation of the Mean," Annals of Mathematical Statistics *37*, 36-50 (1966).

3] Thomas, M.A., "A Simple Sequential Procedure for Sampling Termination in Simulation Investigations," *Journal of Statist. Comput. Simul. 3*, 161-164 (1974).

# COMPUTATIONAL EXPERIENCE ON AN ALGORITHM FOR THE TRANSPORTATION PROBLEM WITH NONLINEAR OBJECTIVE FUNCTIONS*

Ram C. Rao

*Purdue University*
*West Lafayette, Indiana*

Timothy L. Shaftel

*University of Arizona*
*Tucson, Arizona*

ABSTRACT

This paper explores computational implications of allowing nonlinear objective functions in the transportation problem. Two types of nonlinearities, including polynomials, are studied. The choice of these functions resulted from our interest in models of integrated water management. Zangwill's convex simplex method and the primal method of transportation problem form the basis of our algorithm. Innovative features of our work are compact storage and efficient computation procedures. We study the effects on computation time of problem size; the density of nonlinear terms; the size of tolerances for stopping rules; and rules for choice of new variables to enter the solution. We find that problems up to 95 × 95 in size are capable of reasonably fast solution. A particularly surprising finding is that one-dimensional search for improving solutions performs adequately, at least for the kinds of problems posed in this paper. We are encouraged by our results and believe that models involving nonlinear objective functions may be tractable even for relatively large problems, thus making possible more accurate descriptions of real situations.

## 1. INTRODUCTION

The recent advent of sophisticated codes [3], [7] for solving the transportation problem makes it possible to solve extremely large problems of this type. This makes the transportation model a powerful modeling device. In some instances, however, a real situation demands augmentation of the transportation problem. In particular, such effects as economies or diseconomies of scale and other cost interactions among variables cannot be modeled in the transportation framework. We feel that the ability to solve large transportation problems that incorporate nonlinear cost functions will represent an important addition to the usefulness of transportation models in many situations. Our own motivation for solving problems such as

*We would like to thank Professors V. Srinivasan of Stanford University and G.L. Thompson of Carnegie-Mellon University for making available their code for the Transportation problem.

this stems from work in water recycling models [4]. As will be seen later, we attempt to handle two types of nonlinearities in the objective function. One of these types consists of polynomials and is thus quite general.

Our algorithm represents a combination of the primal transportation method [1] and the convex simplex method of Zangwill [8]. In developing a computational procedure we use as the basic building block the methods of Srinivasan and Thompson [6], [7]. Where possible we use their efficient coding procedures for the linear problem. Our innovations include the way we store the information required for the nonlinear objective function in an efficient manner, and in performing the nonlinear calculations in ways we think will minimize their impact on total computation time. In developing these techniques we have drawn from some of the ideas in Shaftel and Thompson [5]. We report computation times for problems with up to 95 rows and 95 columns. Though computation times are larger than those for similar sized linear problems, they compare favorably with what was considered acceptable for the linear transportation problem only a few years ago. To our knowledge computational experience on transportation problems with general nonlinear objective functions is not heretofore available. In this light our study is significant. We believe that it will prove useful to those engaged in developing models with a view to computational viability.

We report on computation times for some of the subroutines used in our code in addition to overall solution time so as to isolate areas for further improvements. We will also discuss our experience in using different parameters, such as stopping rule tolerance, in the code. We believe that many improvements to this code can be made, and that even faster times than those reported here are possible.

## 2. PROBLEM STATEMENT

The problem that we are interested in can be stated as follows: find a schedule of shipments or flows of a homogeneous good from a set of $m$ source nodes (often thought of as warehouses) to a set of $n$ sink nodes (often thought of as markets) to minimize the total cost of effecting the shipments. We denote $x_{ij}$ to be the flow from $i$ to $j$, and $a_i$ and $b_j$ to be the availability and demand at nodes $i$ and $j$, respectively. The objective function to be minimized consists of a linear component $c_{ij}$, the cost of unit shipment from $i$ to $j$, and some other costs that are nonlinear in the decision variables $x_{ij}$. We note that the usual transportation problem has only linear costs. See Dantzig [2], for example. Mathematically, we can write our problem as:

Minimize

$$(1) \qquad \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\,x_{ij} \qquad \text{Linear Terms}$$

$$+ \sum_{t=1}^{T} d_t \prod_{i=1}^{m}\prod_{j=1}^{n} (x_{ij})^{h_{tij}} \qquad \text{Polynomial Terms (a total of } T \text{ terms)}$$

$$+ \sum_{i=1}^{m} D_i \left(\sum_{j=1}^{n} k_{ij}\,x_{ij}\right)^{\alpha_i} \qquad \text{Demand Interdependencies}$$

$$+ \sum_{j=1}^{n} S_j \left(\sum_{i=1}^{m} l_{ij}\,x_{ij}\right)^{\beta_j} \qquad \text{Supply Interdependencies.}$$

Subject to

2) $$\sum_{j=1}^{n} x_{ij} = a_i \quad \text{for } i = 1, 2, \ldots, m$$

3) $$\sum_{i=1}^{m} x_{ij} = b_j \quad \text{for } j = 1, 2, \ldots, n$$

4) $$x_{ij} \geqslant 0 \quad \text{for all } i \text{ and } j.$$

Ve assume that $\sum_i a_i = \sum_j b_j$. In addition we require the following:

$$\alpha_i \geqslant 0 \quad \text{for all } i$$

$$\beta_j \geqslant 0 \quad \text{for all } j$$

$$k_{ij},\ l_{ij} \geqslant 0 \quad \text{for all } i \text{ and } j$$

$$h_{tij} \geqslant 0 \quad \text{for all } t,\ i \text{ and } j.$$

ctually the nonnegativity of $k_{ij}$ and $l_{ij}$ may be relaxed if $a_i$ and $\beta_j$ are restricted to integers. his would be the case if we are dealing with quadratics, for example. For convenience we troduce the following additional notation:

$$R_i = \sum_{j=1}^{n} k_{ij}\, x_{ij} \quad \text{for } i = 1, 2, \ldots, m$$

$$C_j = \sum_{i=1}^{m} l_{ij}\, x_{ij} \quad \text{for } j = 1, 2, \ldots, n$$

Although our purpose in this paper is to examine the computational implications of allowg nonlinear objective functions such as (1) in the framework of a transportation problem, it is eful to motivate the choice of such a function. As noted earlier, we first encountered the ed for it in modeling integrated water management. The polynomials which we have incorrated reflect the fact that the use of some arcs (shipment routes) affects the cost of use of hers. The transshipment type of model used in water management makes this a naturally countered type of objective function. The supply interdependencies involve functions of $C_j$, hich is a linear combination of the shipments into node $j$. As an illustration these can be used model a demand curve at each of the markets. The demand interdependencies involve funcons of $R_i$, which is a linear combination of the shipments emanating from source $i$. To illustte the use of this, suppose sink node $n$ is a "dummy" market. Shipments from node $i$ to $j$, $= 1, 2, \ldots, n-1$, can be thought of as the total production at node $i$. The demand interpendencies term can then be used to model the total cost of production at node $i$.

We now proceed to discuss the different steps in the algorithm and some features of the mputer code we have written for the algorithm.

## 2 DESCRIPTION OF THE ALGORITHM AND THE COMPUTER CODE

The algorithm for solving the proposed problem involves a marriage of Zangwill's convex iplex method [8] and the primal transportation method [1]. We specialize this algorithm to t transportation problem. Our contribution is in developing procedures that render the algorm computationally viable both in respect of demands on storage as well as computation me. We will therefore discuss the details of our code at length, but before that we briefly outlie the various steps of the convex simplex algorithm.

The convex simplex method starts with an initial basic feasible solution which is obtained as for a linear problem. Let $f(x)$ be the objective function to be minimized over $x$ in some convex polyhedral set. A straight forward way to understand the nonlinear approach would be to view the objective function coefficient vector $c$ of the linear problem (clearly $\nabla f(x) = c$) as being replaced by the gradient of the objective function $\nabla f(x)$ at any iteration of the simplex tableau. If this is done then the reduced cost calculation yields the rate of change in the objective function with respect to $x$. A variable which has a negative reduced cost may be raised causing the objective function to decrease.

In a linear problem, the identification of a candidate variable is followed by a pivot to a new basis. For a nonlinear objective function, the optimum does not necessarily lie on an extreme point of the feasible set. Thus the candidate variable provides a locally useful direction along which we search for a step length. Having found the extent of a desirable movement, the candidate variable is raised to the appropriate value, and a new solution point has been found. In the nonlinear problem $\nabla f(x)$ must be calculated at each new solution point. In addition to increasing (raising) variables with negative reduced costs, nonzero (nonbasic) variables with positive reduced costs may be decreased in value in order to reduce the objective function. At any iteration, a new solution point may be the result of a basis change or a modification of the values of the basic variables without a change of basis. The convex simplex method then, can be viewed as the following:

At any iteration,

(i) Calculate $\nabla f(x)$. Using these, calculate the reduced costs associated with the nonbasic variables.

(ii) Find either a raised (nonzero) nonbasic variable which does not have a zero reduced cost or a nonbasic variable at zero with a negative reduced cost. If none stop, the optimum is obtained.

(iii) Raise or lower the value of the chosen variable in the direction which reduces the objective function. Find out how far this variable should be adjusted using a one-dimensional unconstrained search. (Of course, the variable is restricted to be nonnegative). If a basic variable goes to zero prior to reaching the unconstrained minimum go to (iv). Otherwise go to (v).

(iv) The candidate variable enters the basis, and the basic variable driven to zero is removed from the basis. Go to (i).

(v) Modify the values of the basic variables and the nonbasic variable being raised or lowered. The basis remains the same. Go to (i).

The stopping rule in (ii) will lead to a global minimum if the objective function is convex. Otherwise, a local minimum or a stationary point is arrived at. In effect, a Kuhn-Tucker point is obtained. Although we will refer to the optimum solution in the remainder of the paper, the above qualifications are to be borne in mind.

An important aspect of the work we have accomplished is the modification of the primal transportation solution technique to solve a family of nonlinear problems. We now discuss the computational procedures in detail.

The discussion in this section will make use of Figure 1 to a great extent. This figure develops a flow chart which shows the modifications necessary for accommodating nonlinear objective functions in the transportation problem. The figure is divided into three parts each containing different blocks of the flow chart. The first part represents calculations for the linear problem which are not used in the nonlinear case. The second third of the figure represents those blocks which would be used by both linear and nonlinear problems. The first two parts together represent a flow chart for the linear transportation problem. The last part of the figure shows the added computation necessary for nonlinear problems. Combining the last two parts of Figure 1 yields the flow chart for solving the nonlinear transportation problem.

The flow chart in Figure 1 closely follows the FORTRAN computer code which we have written. Thus, each section below from 3.1 to 3.9 represents a subroutine in our code. Wherever appropriate, the block number from Figure 1 will be indicated in the body of the text.

### 3.1 Input

The input for the problem consists of (a) the linear cost coefficients, (b) the rim conditions, i.e., capacities of sources (plants) and demands at sinks (markets) (block A.1) and (c) the nonlinear part of the objective function (block B.1). The linear cost coefficients and rim conditions for a problem of size $m \times n$ are stored in a $(m + 1) \times (n + 1)$ matrix in the usual way. The storage of the nonlinear part however needs more elaboration. We will address the storage of polynomial terms (type 1) and demand and supply interdependencies (type 2) in the objective function separately.

#### 3.1.1 Type 1 — Objective Function

For each term $t$, in this part of the objective function we must keep the following pieces of information: (i) $d_t$ the constant and (ii) $h_{tij}$ the exponent of each variable $x_{ij}$ in term $t$. If a particular variable does not occur in a term, clearly the corresponding exponent would be zero. Since in most problems not all variables will occur in each term with a nonzero exponent, we can take advantage of this to significantly reduce the storage requirements for this part of the objective function. To do this, we assign an index, v, to variable $x_{ij}$, where $v = n \times (i - 1) + j$. Then, we create a list, denoted HL, that contains the term in which a variable occurs and its nonzero exponent. If it does not occur in a term, it is *not stored.* The HL list is ordered by the index v. We then use a pointer list, HTL, in order to extract information about each variable as it is needed. In order to accomplish this information extraction, the HTL list keeps the beginning and ending positions for information regarding each variable. In this fashion we need only store the term numbers in which a variable occurs with a nonzero exponent, the value of that exponent, and a pointer list of length $m \times n$. This as opposed to an $m \times n$ matrix for each term of the polynomial.

#### 3.1.2 Type 2 — Objective Function

For each row nonlinear term, $i = 1, \ldots, m$, the coefficients $k_{ij}$, $j = 1, \ldots, n$ and the coefficients $\alpha_i$ are stored. The data for the $m$ rows along with the exponent are stored in a $n \times (n + 1)$ matrix. The data for the column nonlinear terms $l_{ij}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. and $\beta_j$, are similarly stored in a $(m + 1) \times n$ matrix.

### 3.2 INSOL

This step finds an initial feasible solution (Block A.2). We do this by using the linear cost coefficients only. An initial solution is found using the Modified Row Minimum Rule of Srinivasan and Thompson [7].

*The figure shows necessary modifications to a linear transportation problem in order to accommodate non-linear objective functions. Dashed lines are unique to nonlinear problem. Crossed lines are unique to linear problem. Solid lines are mutual. Thick-lined boxes must be modified for different types of non-linear objective functions.*

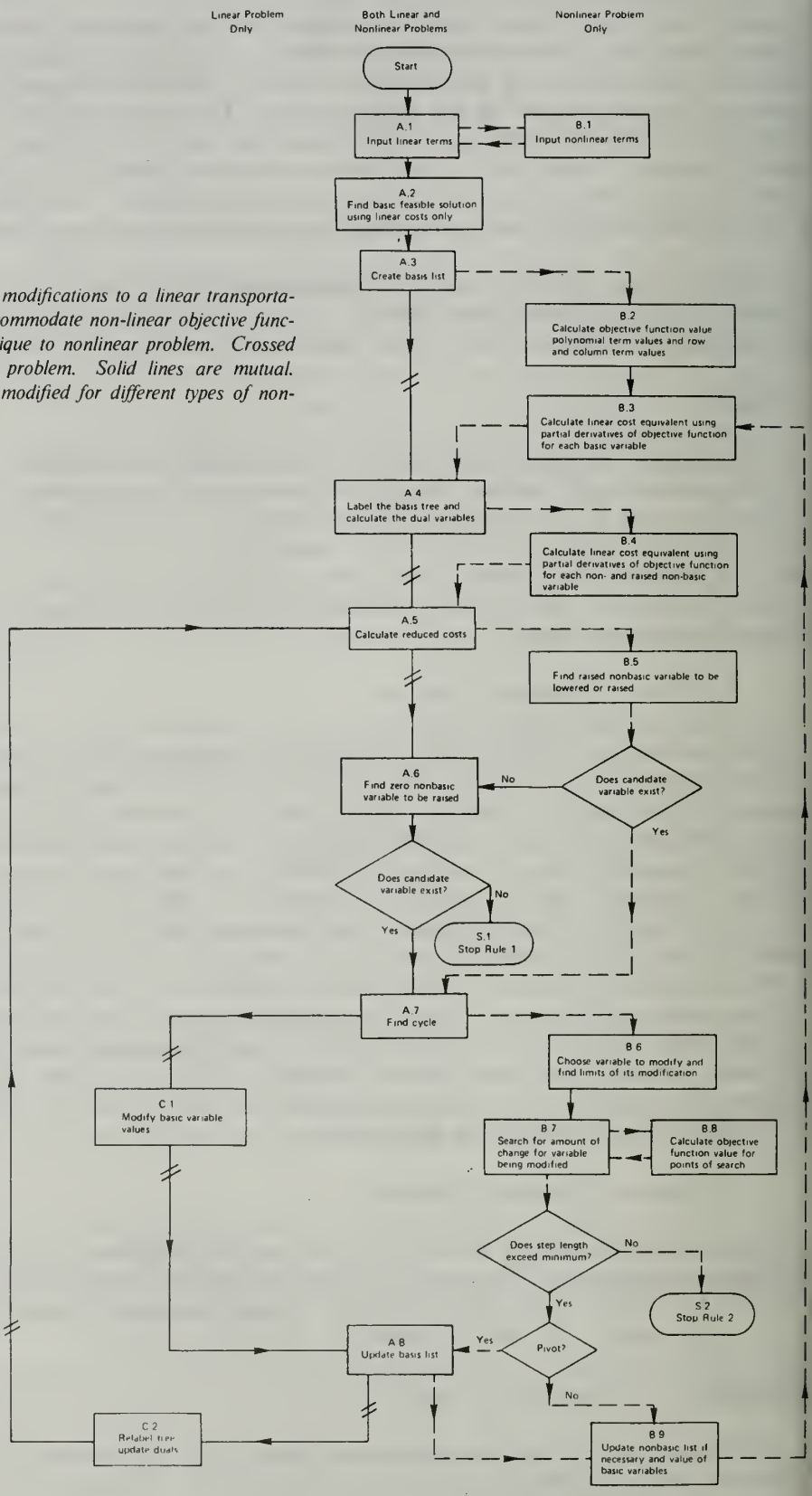


FIGURE 1. Flow chart of computer code

### 3 LABEL 1

This step creates the basis list in the manner of a "pointed list" (Block A.3). The basis list ıows the information on current basic cells, i.e., their row $i$, column $j$ and current value. See rinivasan and Thompson [7] for more details.

### 4 OBJFUN 2

This step calculates the value of the nonlinear part of the objective function for each of e two nonlinear functions (Block B.2). For the polynomials each term value is calculated parately. These values will be needed later in the program for finding the partial derivatives the objective function with respect to each of the variables. Actually two vectors store these rm values for the polynomial objective function. The first vector, $Z$, stores the term values nich when added will give the current value of the nonlinear objective function. The second ctor, $Z1$, stores the term values in a form that enables the calculation of partial derivatives. nis calculation will be discussed in the next section. The objective function value for row and lumn nonlinear terms are calculated in two steps. First the row terms $\sum_j k_{ij}\, x_{ij}$ for = 1, ..., $m$ and the column terms $\sum_i l_{ij}\, x_{ij}$ for $j = 1, \ldots, n$ are calculated and stored in ctors RSUM and CSUM. Then these sums are raised to the appropriate exponent values and ored in RTERM and CTERM. Once again these stored values will be extremely useful for lculation of partial derivatives discussed in the next section.

### 5 LABEL 2

This subroutine of the code accomplishes three things. First, it calculates the partial rivatives of the objective function with respect to each of the basic variables (Block B.3). Scond, it labels the basis tree (Block A.4) using the predecessor successor index method of nivasan and Thompson [7]. This tree is used to determine the cycle created by the addition a new cell to the solution, i.e., a new variable entering the basis. Finally, it calculates the rrent value of the dual variables (Block A.4). In the linear primal transportation code only a rt of the tree basis must be relabled and only a subset of the dual variables needs to be vised. For the nonlinear problem, however, we must use completely revised partial deriva- es at each iteration in order to calculate the duals—in doing so we also relabel the entire tree.

We now turn to the technique we use for calculating the partial derivatives. We attempt t determine the partial derivatives in an efficient way. For each variable, we calculate the rivative with respect to the linear, polynomial and row and column sum terms separately and th n add them together. The derivatives associated with the linear costs are of course merely th initial coefficients. The derivatives for the two types of nonlinear functions are then calcu- la d as follows:

#### 3.5.1 Derivatives for Polynomial Terms

We would like to avoid recalculating each polynomial term for each variable whose partial rivative is calculated. To do this we take advantage of the fact that at each iteration the value o each polynomial term is known to be some value $Z1_t$, and that taking the derivative with rpect to a variable $x_{ij}$ will yield the result, $\dfrac{h_{tij} \cdot Z1_t}{x_{ij}}$. This provides us with a very fast tech- n ue for finding the derivative of each term since we perform just one multiplication and divi- sn.

By using the HL and HTL vectors we can quickly find only those polynomial terms which contain $x_{ij}$. In the use of this technique, problems arise whenever the value of $x_{ij}$ is zero and the exponent $h_{tij}$ is less than or equal to one. When the exponent is greater than one or when more than one zero valued variable occurs in a term the partials are zero and offer no computational difficulty. In the case where a single $x_{ij}$ with an exponent less than or equal to one is at the value zero we act as though this variable is bounded to a prespecified small value (EPS1) until the partials are found. It is then returned to its original value. (The storage of the extra vector $Z$ allows us to perform this modification without losing the true value of the polynomial terms.)

### 3.5.2 Derivatives for Row and Column Nonlinear Terms

The derivatives for these terms are found in a similar fashion to the polynomial terms. Since the row and column sums are known, the derivative with respect to any variable $x_{ij}$ will be $\frac{k_{ij} \cdot \alpha_i \cdot \text{RTERM}}{\text{RSUM}}$, and similarly for the column terms. (RTERM and RSUM were calculated and stored earlier in OBJFUN 2.) In this case, if RSUM is zero the derivative is also zero.

## 3.6 MAIN

This subroutine of the code first calculates partial derivatives with respect to the nonbasic variables (Block B.4) in the same fashion as described in the last section for basic variables. In fact there will be two types of nonbasic variables. (i) Raised nonbasic variables whose value are not zero that will be stored in an array similar to that for basic variables, and (ii) nonbasic variables that are at the value zero (zero nonbasic variables). Once the partial derivative of a nonbasic variable is known, its reduced cost may be found using the standard linear transportation dual equations and the dual variables calculated in LABEL 2 above (Block A.5). We now attempt to find a raised nonbasic variable to be raised or lowered (Block B.5) or a zero nonbasic variable to be raised (Block A.6). Any nonbasic variable with a negative reduced cost is a candidate for being raised above its current value. A raised, nonbasic variable with a positive reduced cost is also a candidate for being modified. In this case the variable must be reduced below its current value in order to improve the objective function. Of these candidates we choose a cell according to one of the following rules:

(i) Matrix-minimum: here the cell that has the smallest reduced cost among all cells is chosen,

(ii) Row-minimum: here cells are examined row by row. As soon as a candidate is encountered, the cell with the lowest reduced cost in that row is chosen,

(iii) Lot-minimum: here cells are examined in specified lots of NP variables. As soon as a lot contains a candidate cell, the cell with the lowest reduced cost in that lot is chosen. Of course the row-minimum rule is equivalent to the lot minimum rule with NP = $n$ where $n$ is the number of columns. The lowest reduced cost should be interpreted as largest in absolute value.

More will be said later about the effects of the choice of rule for picking candidate cells.

## 3.7 CYCLE

In this step we find the cycle that the candidate cell forms with the current basis using Srinivasan's and Thompson's procedure (Block A.7). We isolate the cells whose shipments will

be altered by modifying the candidate cell. We also find the maximum extent that the candidate cell can be modified. This provides an upper bound for a cell if it is to be raised, or a lower bound if it is to be lowered (Block B.6). Of course in the linear transportation problem there is always a pivot to the new basis which permits one to alter shipments at this stage. In the nonlinear case however, before we can alter shipments we must determine the extent to which we wish to modify the candidate cell. We do this in the next step.

### 3.8 SEARCH

In this step we search for a value of the candidate variable that lies between its current value and the bound obtained from the previous step. We do this with a one-dimensional golden section search (Block B.7). Although many possible one-dimensional searches could have been chosen, we preferred the robust characteristics of golden sections. For this search we have to calculate the objective function for each trial value of the candidate variable (Block B.8). This is done in two steps. First we calculate the value of the objective function that is not affected by changes in the shipments of the cells in the cycle. This fixed part is calculated for the linear and the row and column nonlinear portions by setting all variables in the cycle to be zero and then calculating the objective function. For the polynomial portion the same procedure is used after setting all variables in the cycle to be one. Then, for each trial value we determine the shipments of the cells in the cycle and augment the above fixed part of the objective function. This process greatly reduces the number of computations needed for each trial value. The exit possibilities from SEARCH are:

(i) we pivot to a new extreme point,

(ii) we move to an interior point,

(iii) we are not able to move more than a pre-specified small value (EPS2), i.e., the candidate cell does not lead to a better solution.

Actually, before we use the method of golden sections we check to see if a pivot will result in a lower objective function. If such is the case, we pivot rather than look for an interior point with a lower objective. This takes advantage of the speed of the transportation algorithm, ignoring the difficult nonlinear search as often as possible.

### 3.9 UPDATE

This step updates the solution. For all exits for SEARCH we update the nonbasic list as necessary (Block B.9). If we pivot to a new basis then we also update the basis list (Block A.8). We then return to LABEL 2 and begin a new iteration.

### 3.10 Stopping Rules

There are two possible criteria that will terminate the algorithm. (i) If no candidate variable with a reduced cost above a certain prespecified amount exists, the program stops (Block S.1). This stopping rule is the normal rule for linear programs and is invoked most frequently in linear or nonlinear problems. (ii) When performing the nonlinear search (Block B.7), it is possible that the candidate variable that is not at its lower bound cannot be moved a greater amount than a prespecified tolerance with a resulting improved objective value. When this occurs the program searches for another candidate cell among another subset of NP variables. The program stops whenever no candidate cell from these subsets leads to an improved objective function value (Block S.2).

## 4. COMPUTATIONAL EXPERIENCE

Computational experience is based on random problems. Input for each problem includes:

| | |
|---|---|
| $m$, | the number or rows |
| $n$, | the number of columns |
| $T$, | the desired number of terms in the polynomial |
| AXT, | the average number of variables per term |
| ADEN, | the average number of nonzero row and column coefficients, $k_{ij}$ and $l_{ij}$ |
| EPS2, | the tolerance for stopping rule (ii) |
| EPS3, | the tolerance for stopping rule (i) (set at .001 throughout these tests) |
| NP, | the number of variables whose reduced cost is calculated during any pass of MAIN |

With these inputs, test problems are generated with the following specifications:

| | |
|---|---|
| $a_i, b_j$ | : random integer numbers between 0 and 100 |
| $c_{ij}$ | : random numbers between 0 and 100 plus the row number of that variable |
| $d_t$ | : random numbers between 0 and 30 |
| $h_{tij}$ | : random numbers between 0 and 3 for variables chosen at random to be in each t |
| $S_j, D_i, k_{ij}, l_{ij}$ | : random numbers between 0 and 10 |
| $\alpha_i, \beta_j$ | : random numbers between 0 and 3 |

Figures 2 through 7 summarize the results of the computation runs. All numbers except those with * are based on 3 runs; * indicate a single run only. For each problem type with multiple runs, the computation times are the average of total execution times including time for generating input and are in seconds. The code itself is written in FORTRAN. All computation was performed on CYBER 70 at the University of Arizona computation center. Inspection of Figure 7 indicates that CPU time is improved if only a subset of reduced costs are calculated. Generally 10 to 20% of the variables seem about optimal in this case. A number in this range was chosen for NP in all the computational results presented. As expected, times increased with the finer tolerance in the search to a point where too fine a search would not yield any results whatever—see Figure 3. (These were the only cases of no consequence.) Times increase dramatically as the row and column nonlinear terms have more nonzero coefficients and thus become more active—see Figure 4. As the number of variables occurring in each polynomial term increased, the computation time is not greatly affected. This is seen in Figure 5. This situation is not unexpected since a single zero variable in a term gives it a zero value (and two zero variables make the derivative of that term zero). Times decreased as the number of terms in the polynomial rose—see Figure 6. We suspect that the multiple random terms in the polynomial somehow balanced each other so that the cumulative effect was to offset each other, thus leading to a solution closer to an extreme point.

Due to the realities of computer budgets, more extensive tests were not possible. The basic results, however, can be found in Figure 2. Total time even for the 95 × 95 problem are well within acceptable computation times. Near optimal results (10% from optimum) were obtained very quickly. Also obvious is that the bulk of the time spent in solving the nonlinear

| oblem Size | Average Total Time[a] | Average Time to Within 10% of OPT | Average Time in Search | Average Time in Main | Average No. of Basic Variables | Average No. of Raised Non-Basic Variables (% increase in non-zero variables) |
|---|---|---|---|---|---|---|
| × 10 | 1.52 0(1.02) | .42 | 1.16 | .10 | 19 | 3.00 (15.7) |
| × 20 | 4.64 0(4.20) | 2.00 | 3.31 | .53 | 39 | 6.33 (16.2) |
| × 30 | 21.02 (18.81) | 10.85 | 15.03 | 3.00 | 59 | 16.67 (28.3) |
| × 50 | 56.71 (56.91) | 34.55 | 26.74 | 25.85 | 99 | 28.67 (29.0) |
| × 70* | 198.74 (00.00) | 56.13 | 143.24 | 41.72 | 139 | 42.00 (30.2) |
| × 95* | 354.38 (00.00) | 227.73 | 255.19 | 76.02 | 189 | 63.00 (33.3) |

FIGURE 2. Results for a Range of Problem Sizes
EPS2 = .05; T = 20; AXT = 20; ADEN = .9
NP = 7 + 20 percent of the total number of variables
[a]Number in parenthesis are median times

| EPS2 | Time | Value of OBJ Function |
|---|---|---|
| .10 | 17.07 | $3.562 \times 10^5$ |
| .05 | 21.30 | $3.533 \times 10^5$ |
| .01 | Solution not found | |

FIGURE 3. EPS2 vs. Time
T = 20, AXT = 2, ADEN = .9
m = n = 30, NP = 150

| ADEN | Time |
|---|---|
| .2 | 4.30 |
| .5 | 14.68 |
| .9 | 21.30 |

FIGURE 4. ADEN vs. Time
T = 20, EPS2 = .05, AXT = 2,
m = n = 30, NP = 150

| AXT | Time |
|---|---|
| 2.5 | 20.49 |
| 5.0 | 20.89 |
| 20.0 | 21.30 |

FIGURE 5. AXT vs. Time
T = 20, ESP2 = .05, ADEN = .9,
m = n = 30, NP = 150

| T | Time |
|---|---|
| 10 | 23.51 |
| 20 | 20.89 |
| 50 | 17.89 |

FIGURE 6. T vs. Time
AXT = 5.0, ESP2 = .05, ADEN = .9,
m = n = 30, NP = 150

| Problem Size | 30 × 30 | | 20 × 20 | | 10 × 10 | |
|---|---|---|---|---|---|---|
| | NP | Time | NP | Time | NP | Time |
| | 30 | 20.90 | 20 | 6.93 | 10 | 1.70 |
| | 60 | 21.02 | 40 | 4.64 | 20 | 1.52 |
| | 150 | 17.07 | 100 | 4.80 | 40 | 1.53 |
| | 300 | 17.23 | 200 | 7.39 | 50 | 2.42 |
| | 450 | 31.88 | | | | |

FIGURE 7. NP vs. Time
T = 20, AXT = 20, ADEN = .9,
EPS2 = .05 (= .10 for 30 × 30 case)

problem is in MAIN where the reduced costs are calculated, and in SEARCH. One obvious area of further research will be to replace the relatively slow (but robust) golden sections search with a faster type of linear search. The final column in figure 2 is the number of nonbasic variables in the optimum solution at a value other than zero. We use this as one measure of the nonlinearity of the problems we have chosen to solve. We also indicate the percent increase in nonzero variables caused by raised nonbasic variables.

## 5. CONCLUSIONS

In this paper we have shown that transportation problems with a class of nonlinear objective functions can be solved efficiently for problems large enough to make them useful. We believe in fact that the results are startling. These results are even more impressive given the common belief that one variable at a time searches cannot be useful in solving nonlinear problems. In the cases we studied, only a small percentage of nonbasic variables became active so that our problems were in fact smallish nonlinear problems—nevertheless there is no way of telling a priori which variables will be inactive. Also, at the optimum the number of raised nonbasic variables is significant in relation to the number of basic variables.

Obvious areas of further research remain. In particular, replacing the slow golden section search with a more efficient routine will lead to some improvement. It is also very simple in the present code to modify the calculations to solve nonlinear problems with any differentiable objective function. Finally, attempting to modify more than one variable during the nonlinear search presents an intriguing option.

We hope that this paper will stimulate others to explore using, where appropriate, more extensive nonlinear models in their own problem solving studies. Nonlinear problems pose interesting challenges. Computational viability uiltimately depends on exploiting special problem structures as well as experimentation with computational procedures. This has been the primary aim of our research.

## REFERENCES

[1] Charnes, A. and W.W. Cooper, *Management Models and Industrial Applications of Linear Programming*, Vol. I (Wiley, New York, 1961).

[2] Dantzig, G.B., *Linear Programming and Extensions* (Princeton U. Press, Princeton, 1963).

[3] Glover, F., D. Karney, D. Klingman and A. Napier, "Computational Study on Basis Change Criteria, and Solution Algorithms for Transportation Problems," Management Science, pp. 793-814 (1974).

[4]. Pingry, D. and T. Shaftel, "Integrated Water Management: Treatment, Reuse and Disposal," Water Resources Research, to appear.

[5] Shaftel, T. and G.L. Thompson, "A Simplex-Like Algorithm for the Continuous Modular Design Problem," Operations Research *25*, pp. 788-807 (1977).

[6] Srinivasan, V. and G.L. Thompson, "Accelerated Algorithms for Labelling and Relabelling of Trees, with Applications to Distribution Problems," Journal of the Association for Computing Machines *19*, pp. 712-726 (Oct. 1972).

[7] Srinivasan, V. and G.L. Thompson, "Benefit-Cost Analysis of Coding Techniques for the Primal Transportation Algorithm," Journal of the Association for Computing Machines *20*, pp.194-213, (April 1973).

[8] Zangwill, W., *Nonlinear Programming: A Unified Approach* (Prentice-Hall, New Jersey, 1969).

# A NOTE ON DETERMINING OPERATING STRATEGIES FOR PROBABILISTIC VEHICLE ROUTING

James R. Yee

*Department of Electrical Engineering*
*University of Maryland*
*College Park, Maryland*

Bruce L. Golden

*College of Business and Management*
*University of Maryland*
*College Park, Maryland*

**ABSTRACT**

The stochastic vehicle routing problem is a problem of current importance and research interest. Applications include schoolbus routing, municipal waste collection, subscription bus scheduling, daily delivery of dairy goods, and a host of related transportation and distribution activities. In this paper, we assume that routes for vehicles have already been generated and we focus on determining operating strategies. That is, under what conditions should a driver return to the central depot in order to replenish his supply? We present a dynamic programming recursion which addresses this question and we show that the optimal policy is of a rather simple form. Finally, an algorithm and example illustrate the policy.

## ITRODUCTION

The stochastic vehicle routing problem is a problem of current importance and considerle research interest. Applications include schoolbus routing, municipal waste collection, subiption bus scheduling, daily delivery of dairy goods, and a host of related transportation and stribution activities. The problem that has been considered to date is to determine a fixed set vehicle routes of minimal expected total distance. We assume that all vehicles leave from d eventually return to a central depot, and that vehicle capacity constraints and probabilistic stomer demands must be satisfied (for deterministic demands see Golden, Magnanti and guyen [1]).

In early work on stochastic vehicle routing, Tillman [4] introduced a heuristic approach to termine, under some rather strong assumptions, a fixed set of routes of minimal expected al cost. The total cost of traveling a route included a cost associated with not hauling enough ods to satisfy customer demands and a cost associated with carrying unnecessary excess goods the route.

Golden and Stewart [2] address the problem of determining a fixed set of routes where the demand of customer $i$ is modeled by a Poisson distribution with mean $\lambda_i$. The objective is the minimization of the expected total distance traveled. In Golden and Yee [3], this work is extended and a general framework for solving the stochastic vehicle routing problem is provided. The authors discuss a solution procedure for correlated demands and develop a general solution technique for the situation where demands are Poisson, negative binomial, binomial or gamma distributed. In addition, they derive analytical results which describe the relationships between design parameters and offer a scheme for performing perturbation analysis.

In previous work, whenever a vehicle does not have enough goods to satisfy a customer's demand, it must immediately return to the central depot for reloading before supplying goods to the remaining customers on the route. This, however, is not always the most cost-effective strategy. It is sometimes judicious not to wait until a vehicle is empty in order to return to the central depot for reloading. In this paper, we determine conditions under which a vehicle *should* return to the central depot. These operating strategies lead to the minimization of expected distance traveled.

## DISCUSSION

Suppose that the routes for vehicles have already been determined by a solution procedure such as the Clarke-Wright algorithm (see [1] for details). Each vehicle leaves the central depot and proceeds to deliver goods to a number of demand points. After delivering goods to a demand point on a route, the driver is faced with the decision of whether or not to return to the depot in order to replenish his supply. We shall prove that the optimal decision is based upon whether or not the remaining supply of goods in the vehicle is greater or less than some critical value $\hat{y}$ which must take into account

(i) the probabilistic demands on the remaining portion of the route, and
(ii) the distances between customers.

Let the demand of customer $i (i = 1, 2, \ldots, N)$ be given by the discrete random variable $X_i$ and have some arbitrary mass function $P(X_i = k)$; we remark that continuous probability distributions can be handled analogously. Each vehicle has a capacity of $c$ and starts out on a route with a supply of $c$ units. Without loss of generality, we number the customers in reverse order from which deliveries are made, i.e., deliveries are made to customer $N$ first, customer $N - 1$ second, and so on (customer 0 denotes the central depot). The matrix of distances $D = [d_{ij}]$ between customers $i$ and $j$ is known in advance and we will assume that it is symmetric in order to simplify computations.

In general, after delivering goods to customer $n + 1$, let $y_n$ represent the remaining supply of goods on the vehicle with $n$ customers still to be visited and let $V_n(y_n)$ for $n = 0, 1, \ldots, N$ be the expected distance to be traveled in supplying the remaining $n$ customers given that there are $y_n (0 \leqslant y_n \leqslant c)$ remaining units of supply in the vehicle and an optimal policy is followed; for $n = 0$, $V_n(y_n)$ is defined to be $d_{1,0}$.

We now present a dynamic programming approach for determining the optimal operating policy for a vehicle. After delivering to customer $n + 1$, we have

$$V_n(y_n) = \min \left\{ d_{n+1,0} + d_{0,n} + \sum_{k=0}^{\infty} P(X_n = k)\, V_{n-1}(c - k)^+, \right. \tag{1}$$

$$d_{n+1,n} + \sum_{k=0}^{y_n} P(X_n = k)\, V_{n-1}(y_n - k) + \sum_{k=y_n+1}^{\infty} P(X_n = k)[2d_{0,n} + V_{n-1}(c - k + y_n)^+]\Bigg\}$$

where $V(\omega)^+ = V(\max(0,\omega))$.

The recursive equation in (1) may be used to determine $V_n(y_n)$, the minimal expected distance for completing the route with $y_n$ remaining units of supply in the vehicle and $n$ demand points still to be serviced. The two options available to the driver are to go back to the depot and replenish his supply or to proceed to the next customer on the route. The expression representing the expected distance of first returning to the depot includes the distance of traveling from customer $n + 1$ to the depot, then from the depot to customer $n$, and finally the minimal expected distance with $n - 1$ customers remaining. On the other hand, if the driver advances to the next node directly, the expected distance includes the distance between customers $n + 1$ and $n$, the distance from customer $n$ to the central depot which might be incurred twice as a penalty, and the minimal expected distance with $n - 1$ customers remaining. It is assumed that the probability of any customer's demand exceeding vehicle capacity is infinitesimal.

In the remainder of this section we study properties of the optimal solution to our problem. First, we state the following lemma and definition.

LEMMA 1: $2d_{0,n} + V_{n-1}(a + b) \geqslant V_{n-1}(b)$, where $a$ and $b$ are positive integers and $a + b \leqslant c$.

PROOF: After supplying customer $n$ and with $b$ units of supply remaining in the vehicle, the driver may decide to return to the depot, pick up an additional $a$ units of supply and return to customer $n$. Since this is not necessarily optimal, the result follows.

DEFINITION: $\hat{y}_n = \min\Bigg\{\text{integer } y\,(0 \leqslant y \leqslant c) \mid d_{n+1,0} + d_{0,n} + \sum_{k=0}^{\infty} P(X_n = k)\, V_{n-1}$

$(c - k)^+ \geqslant d_{n+1,n} + \sum_{k=0}^{y} P(X_n = k)\, V_{n-1}(y - k) + \sum_{k=y+1}^{\infty} P(X_n = k)[2d_{0,n} + V_{n-1}(c - k + y)^+]\Bigg\}$

or $c + 1$ if no such $y$ exists.

We can now state and prove the general structure of the optimal policy.

THEOREM 1: After supplying customer $n + 1$, if the remaining supply $y_n \geqslant \hat{y}_n$, then the optimal policy is to proceed to the next demand point. Otherwise return to the depot.

PROOF: We prove the first part of the theorem here; the second part is proved in an analogous fashion. Suppose $\hat{y}_n \leqslant c$. Then, by definition

$$d_{n+1,0} + d_{0,n} + \sum_{k=0}^{\infty} P(X_n = k)\, V_{n-1}(c - k)^+ \geqslant d_{n+1,n} + \sum_{k=0}^{\hat{y}_n} P(X_n = k)\, V_{n-1}(\hat{y}_n - k)$$

$$+ \sum_{k=\hat{y}_n+1}^{\infty} P(X_n = k)[2d_{0,n} + V_{n-1}(c - k + \hat{y}_n)^+]. \tag{2}$$

Let $y_n = \hat{y}_n + m$, where $m$ is a positive integer. For the decision of proceeding to the next customer, the expected cost of completing the route is

(3)
$$d_{n+1,n} + \sum_{k=0}^{\hat{y}_n+m} P(X_n = k)\, V_{n-1}(\hat{y}_n + m - k) + \sum_{k=\hat{y}_n+m+1}^{\infty} P(X_n = k)[2\, d_{0,n} + V_{n-1}(c - k + \hat{y}_n + m)^+].$$

Subtracting (3) from the right-hand side of (2) results in

$$\sum_{k=0}^{\hat{y}_n} P(X_n = k)[V_{n-1}(\hat{y}_n - k) - V_{n-1}(\hat{y}_n + m - k)] + \sum_{k=\hat{y}_n+1}^{\hat{y}_n+m} P(X_n = k)[2\, d_{o,n} + V_{n-1}(c - k + \hat{y}_n)^+ - V_{n-1}(\hat{y}_n + m - k)] + \sum_{k=y_n+m+1}^{\infty} P(X_n = k)[V_{n-1}(c - k + \hat{y}_n)^+ - V_{n-1}(c - k + \hat{y}_n + m)^+] \geqslant 0.$$

The last inequality is obtained from Lemma 1 and the fact that for all $n$

$$V_n(a + m) \leqslant V_n(a), \text{ where } a,m \geqslant 0 \text{ and } a + m \leqslant c.$$

If $\hat{y}_n = c + 1$, then the triangle inequality is violated since $d_{n+1,n} > d_{n+1,0} + d_{0,n}$ and the optimal decision is to return to the central depot.

From the previous theorem, the following algorithm becomes apparent.

*OPERATING STRATEGY ALGORITHM:*

Step 0. $n \leftarrow 0$. $V_0(y) = d_{1,0}(y = 0, 1, \ldots, c)$.
Step 1. $n \leftarrow n + 1$; determine $V_n(y)(y = 0, 1, \ldots, c)$ and $\hat{y}_n$.
Step 2. If $n < N - 1$, go to step 1; otherwise compute $V_N(c)$ and stop.

*EXAMPLE*:

In this section, we illustrate the algorithm by presenting data and results for a small sample problem in order to gain more insight into the performance of this simple dynamic programming algorithm. We consider a vehicle routing problem with four demand points and the distance matrix shown in Figure 1. Suppose that customer demands are Poisson distributed with average demands of $\lambda_1 = 30$, $\lambda_2 = 10$, $\lambda_3 = 30$, $\lambda_4 = 25$ and that vehicle capacity is 100 units. After some preliminary analysis, the following fixed route emerges: $0 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow 0$. We seek to determine an effective operating strategy. In this case, the dynamic programming recursion yields

$\hat{y}_1 = 33$

$\hat{y}_2 = 0$

$\hat{y}_3 = 38$

and the expected distance traveled under this optimal policy is 30.59 units. If this operating policy were not followed and a vehicle returned to the depot only after it discovered that customer demand could not be satisfied, then the expected distance would be 32.04 units. Thus, the operating policy results in a 4.53% savings in expected distance traveled. Furthermore, we note that if vehicle capacity were infinite, distance traveled would be 29 units. The operating policy is straightforward and easy to calculate, intuitively appealing, and very effective in minimizing distance covered.

| | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 | - | 5 | 2 | 11 | 6 |
| 1 | 5 | - | 4 | 12 | 8 |
| 2 | 2 | 4 | - | 2 | 10 |
| 3 | 11 | 12 | 2 | - | 12 |
| 4 | 6 | 8 | 10 | 12 | - |

FIGURE 1. Distance Matrix

## CKNOWLEDGMENT

A preliminary version of this paper was presented at the 1978 Northeast AIDS Conference in Washington, D.C.

## REFERENCES

Golden, B., T. Magnanti and H. Nguyen, "Implementing Vehicle Routing Algorithms," Networks, *7* (2), 113-148 (1977).

Golden, B. and W. Stewart, "Vehicle Routing with Probabilistic Demands," Computer Science and Statistics: Tenth Annual Symposium on the Interface (D. Hogben, D. Fife, eds.), National Bureau of Standards Special Publication 503, Washington, D.C., 252-259 (1978).

Golden B. and J. Yee, "A Framework for Probabilistic Vehicle Routing," American Institute of Industrial Engineers Transactions, *11*(2), 109-112 (1979).

Tillman, F., "The Multiple Terminal Delivery Problem with Probabilistic Demands," Transportation Science, *3*(3), 192-204 (1969).

# A NOTE ON THE "VALUE" OF BOUNDS ON EVPI IN STOCHASTIC PROGRAMMING*

J. G. Morris and H. E. Thompson

*Graduate School of Business*
*University of Wisconsin*
*Madison, Wisconsin*

**ABSTRACT**

The existing literature concentrates on determining sharp upper bounds for EVPI in stochastic programming problems. This seems to be a problem without an application. Lower bounds, which we view as having an important application, are only the incidental subject of study and in the few instances that are available are obtained at an extremely high cost.

In order to suggest a rethinking of the course of this research, we analyze the need for bounds on EVPI in the context of its significance in decision problems.

## 1. INTRODUCTION

The standard definition of the expected value of perfect information, EVPI, is the maximum amount that the decision maker would pay for perfect prior information as to the realization of the random variable in a decision problem. This definition is contained in much of the recent literature that concentrates on determining upper bounds for EVPI in stochastic programming problems. For example, Ziemba and Butterworth [6, pp. 365-66] note that EVPI is the upper bound on the "amount that the decision maker would pay a clairvoyant" to provide the information but go on to suggest that since the calculations are formidable, "it's of interest to find an upper bound on EVPI that requires less computational effort." No attention is paid to developing a lower bound other than zero. (See also [1], [2], [4].) That upper bounds on EVPI and this standard definition should be juxtaposed is peculiar. There is little managerial "value" in providing upper bounds on EVPI within the context of its definition. In contrast, a lower bound on EVPI, which has managerial significance, is typically left at zero by authors writing on the subject. To illustrate, suppose that we determined an upper bound on EVPI to be $5. Armed with this information, how much will we pay a clairvoyant to predict realizations of the random variable? Since $5 in an *upper bound*, it may be unjustifiable to offer $5. To be safe, we should pay close attention to the *lower bound.* And it seems reasonable to offer to pay the clairvoyant only up to the lower bound for this information since it is the *maximum* amount that we could pay and still be assured of justifying the expense.

In the next section, we discuss a stochastic programming problem from an applications point of view.

Support from the Rennebohm Foundation is acknowledged.

## 2. STOCHASTIC PROGRAMMING AND EVPI

Consider a decision problem given by

$$\text{Max } \{E_z u(f(x, z)) | x \in K\} \tag{1}$$

where the decision, $x$, must be made prior to the realization of the random variable, $z$. In (1), $x$ and $z$ may be vectors and $f(x, z)$ is the accumulated monetary wealth of the decision maker subsequent to the random event and $u$ is a utility function; $K$ is the set of feasible decision alternatives.

Problem (1) can be very difficult computationally since it may involve maximizing an expected value of a nonlinear function of both $x$ and the random variable $z$. One way to ease this computational burden is to find and solve a computationally easier problem which gives approximately the same solution as (1). Another approach would be to expend resources to obtain aforehand knowledge of $z$, thus at least eliminating the expectation from the problem. This latter approach also provides an opportunity to produce a greater expected utility.

### SOLVING AN EASIER PROBLEM

Suppose we solve an easier problem and obtain the solution $x_e$. Suppose also that the cost to solve (1) is $C_1$. Then it would be worthwhile adopting $x_e$, thus avoiding $C_1$, as long as

$$E_z u(f(x_e, z)) > \text{Max } \{E_z u(f(x, z) - C_1) | x \in K\}. \tag{2}$$

Now if an easily computed upper bound on the RHS of (2), call it $U_e$, can be found, then a decision rule is available. It is to adopt $x_e$ as the problem solution if $E_z u(f(x_e, z)) > U_e$.

To see how this might work, consider an example where $u$ is *linear* and $f$ is concave in $z$. Then (2) becomes

$$C_1 > \text{Max } \{E_z f(x, z) | x \in K\} - E_z f(x_e, z). \tag{2a}$$

Let the easier problem be

$$\text{Max } \{u(f(x, \mu)) | x \in K\} = u(\text{Max } \{f(x, \mu) | x \in K\}) \tag{3}$$

where $\mu = E(z)$. Avriel and Williams [1] used the following inequalities for determining an upper bound on EVPI:

$$E_z f(\bar{x}, z) \leqslant \text{Max } \{E_z f(x, z) | x \in K\} \leqslant E_z \text{Max} \{f(x, z) | x \in K\} \leqslant \text{Max } \{f(x, \mu) | x \in K\} \tag{4}$$

where $\bar{x}$ solves (3). For $u$ linear,

$$\text{EVPI} = E_z \text{ Max } \{f(x, z) | x \in K\} - \text{Max } \{E_z f(x, z) | x \in K\} \tag{5}$$

and

$$\begin{aligned} \text{Max } \{f(x, \mu) | x \in K\} - E_z f(\bar{x}, z) &\geqslant E_z \text{ Max } \{f(x, z) | x \in K\} \\ &\quad - \text{Max } \{E_z f(x, z) | x \in K\} = \text{EVPI}. \end{aligned} \tag{6}$$

If we define $x_e = \bar{x}$, then the LHS of (6) is an upper bound on the RHS of (2a) and $x_e$ would be adopted if

$$C_1 > f(x_e, \mu) - E_z f(x_e, z). \tag{7}$$

e RHS of (7) is relatively easy to compute since it involves solving a deterministic program-ıg problem and taking an expectation.

That the RHS of (7) is an upper bound on EVPI is incidental. The essential elements are: ling an easier problem to solve and finding an upper bound on (1) which provides the deci-ı rule (7). Research directed toward finding an upper bound on EVPI such as found in riel and Williams [1], Ziemba and Butterworth [6], and Huang, Vertinsky, and Ziemba [4] is ırly applicable to the decision problem posed. However, if we were able to discover an ier problem for which an upper bound on the RHS of (2a) is less than EVPI, we would be eedingly happy.

## QUIRING AFOREHAND PERFECT INFORMATION

Assume perfect aforehand information can be acquired for a cost of $C_I < \infty$. Given this ɔrmation, one can solve the wait-and-see problem

$$\text{Max } \{u(f(x, z))|x \in K\}$$

that value of $z$ which will be realized. Solving (8) is ordinarily much easier than solving Suppose that the cost to do so is $C_W < C_1$. Then it is desirable to acquire the informa-if

$$E_z \text{ Max } \{u(f(x, z) - C_W - C_I)|x \in K\} \geqslant \text{Max } \{E_z u(f(x, z) - C_1)|x \in K\}.$$

define an EVPI′ implicity in

$$E_z \text{ Max } \{u(f(x, z) - C_W - \text{EVPI}')|x \in K\} = \text{Max } \{E_z u(f(x, z) - C_1)|x \in K\}.$$

Equation (10) is in contrast with the implicit definition of EVPI usually stated as

$$E_z \text{ Max } \{u(f(x, z) - \text{EVPI})|x \in K\} = \text{Max } \{E_z u(f(x, z))|x \in K\}.$$

is *linear*, then (10) yields

$$\text{EVPI}' = E_z \text{ Max } \{f(x, z)|x \in K\} - \text{Max } \{E_z f(x, z)|x \in K\} + C_1 - C_W$$

(11) yields the definition of Equation (5). Therefore,

$$\text{EVPI}' = \text{EVPI} + C_1 - C_W.$$

Now if the acquisition cost of the information $C_I$ is less than EVPI′, then it is worth ining, i.e., if

$$C_I \leqslant \text{EVPI} + C_1 - C_W.$$

re $C_1$ and $C_W$ are assumed known, lower bounds on EVPI are significant to the decision clem. Upper bounds are not.

One exception to the concentration on upper bounds appears in Huang, Vertinsky, and eba [4]. They show that for a *linear* utility function, "sharp" upper and lower bounds can be ıned for both $z_p \equiv E_z \text{ Max } \{f(x, z)|x \in K\}$ and $z_n \equiv \text{Max } \{E_z f(x, z)|x \in K\}$. Then if $z_p \leqslant z_p^U$ and $z_n^L \leqslant z_n \leqslant z_n^U$, we have

$$\text{Max } [0, z_p^L - z_n^U] \leqslant \text{EVPI} \leqslant z_p^U - z_n^L$$

since EVPI $= z_p - z_n$, and the problem becomes one of finding sharp bounds $z_p^L$, $z_p^U$, $z_n^L$, $z_n^U$. Considerable computing effort renders these bounds as sharp as desired by the decision maker. Therefore, we could reconstitute the decision inequality of (14) as

$$C_I \leqslant Max\ [0,\ z_p^L - z_n^U] + C_1 - C_W. \tag{16}$$

When $u$ is strictly increasing and concave, Ziemba and Butterworth [6, p. 375] show that

$$0 \leqslant \text{EVPI} \leqslant f(\bar{x},\ \mu) - u^{-1}\{E_z u(f(\bar{x},\ z))\}, \tag{17}$$

where $u^{-1}$ is the inverse function of $u$. The lower bound which is essential for decision making is left at zero.

Also for the case of a concave utility function, Huang, Vertinsky and Ziemba [4, pp. 136-137] show that with additional assumptions about the probability distribution of $z$, one can obtain *either* an upper or a lower bound for EVPI, but not both.

Although Huang, Vertinsky, and Ziemba make some inroads into determining lower bounds on EVPI, it appears that such a determination is entirely accidental rather than by design. They appear to ignore the decision maker's real problem and concentrate on the upper bound for EVPI. The research direction in this area would benefit from returning to its applications base initiated by Dantzig [3] and Madansky [5].

## 3. ADDITIONAL COMMENTS ON LOWER BOUNDS

For the *linear* utility case, by Equation (5) $z_p - z_n$ defines EVPI. In general, $z_n$ is difficult to calculate as has been the assumption in motivating discussion of the decision problem. If the domain of $z$ is $R \subset E^S$ (Euclidean S-space), the evaluation of $z_p$ involves maximizing a function involving multiple integrals. This may be particularly undesirable and resort to numerical methods may be required. Likewise, $z_p$ would also involve evaluating a multiple integral and again resort to numerical methods may be required.

It is instructive to return to the paper by Huang, Vertinsky and Ziemba to discuss the nature of the lower bounds on EVPI which they determine. The methods suggested consist of approximating the integrals imbedded in $z_n$ and $z_p$ with summations by partitioning. For example, consider $R \equiv [a,b] \subset E^1$. The upper bound on $z_n$ would be found by evaluating

$$z_n^U = \text{Max}\left\{\sum_{i=1}^{I} \alpha_i f(x,\ \beta_i) \,|\, x \in K\right\} \tag{18}$$

where $\alpha_i$ is the probability of $z$ belonging to an interval $i$ and $\beta_i$ is the mean $z$ conditional on $z$ belonging to interval $i$ [4, pp. 130-131]. As $I \to \infty$ finer partitions are made and $z_n^U \to z_n$. However, (18) is merely a discrete version of (1) and solving it approximates the solution of (1). Thus, we solve a problem *almost* as computationally undesirable as (1) save for the elimination of an integral.

The lower bounds on $z_p$ would be found by evaluating

$$z_p^L = \sum_{i=0}^{I} \delta_i \text{ Max } \{f(x,\ d_i) \,|\, x \in K\} \tag{19}$$

where $\delta_i$ is an appropriate probability measure and $d_i$ is an appropriate point in interval $i$ [4, pp. 130-131]. Again, as finer partitions are made, $z_p^L \to z_p$. But calculating $z_p^L$ involves solving

$l + 1$ programming problems (whose values are then averaged). Equation (19) consists of a discrete version of $z_p$ and has as its main computational virtue the elimination of an integral. Similar results hold for the case of several random variables.

Since

$$z_p^L - z_n^U \leqslant \text{EVPI} = z_p - z_n, \tag{20}$$

we have a lower bound on EVPI. If we now ask the question of how much we would be willing to pay a clairvoyant for the benefits of perfect aforehand knowledge of $z$ *ignoring* the cost of calculating $z_p^L$ and $z_n^U$ and *ignoring* elements $C_1$ and $C_W$ as factors in the decision, then the lower bound from (20) would help to answer the question.

If, however, the computation costs are not negligible and if the reason for seeking a clairvoyant was to ease the computational burden, then we have an interesting situation. It is worthwhile employing the clairvoyant if (14) holds. But (14) is predicated on the assumption that $C_1$ will be avoided if the clairvoyant is hired. But in finding the lower bound on EVPI, we solved a problem which was an approximation to (1) and as that approximation got better, the cost of doing it would approach $C_1$. That is, $C_1$ was not entirely avoided. Now if the decision rule (14) tells us to avoid solving (1), we have a dilemma since it suggests that we should not do what we just (approximately) did!

## REFERENCES

[1] M. Avriel and A.C. Williams, "The Value of Information and Stochastic Programming," Operations Research *18*, 947-954 (1970).

[2] D.P. Baron, "Information in Two-Stage Programming Under Uncertainty," Naval Research Logistics Quarterly *18*, 169-176 (1971).

[3] G.B. Dantzig, "Linear Programming Under Uncertainty," Management Science *1*, 197-206 (1955).

[4] C.C. Huang, I. Vertinsky, and W.T. Ziemba, "Sharp Bounds on the Value of Perfect Information," Operations Research *25*, 128-139 (1977).

[5] A. Madansky, "Inequalities for Stochastic Linear Programming Problems," Management Science *6*, 197-204 (1960).

[6] W.T. Ziemba and J.E. Butterworth, "Bounds on the Value of Information in Uncertain Decision Problems," Stochastics *1*, 361-378 (1975).

# A NOTE ON INTEGER LINEAR FRACTIONAL PROGRAMMING

Suresh Chandra and M. Chandramohan

*Indian Institute of Technology*
*Delhi, India*

**ABSTRACT**

This note consists of developing a method for enforcing additional constraints to linear fractional programs and showing its usefulness in solving integer linear fractional programs.

## 1. INTRODUCTION

Fractional cutting plane methods for solving integer linear fractional programs have been proposed, for example, by Swarup [6], Grunspan and Thomas [3] and Granot and Granot [2]. In contrast with these, the method proposed in the sequel does not impose any severe restriction on the problem as in Swarup [6], does not consist of solving many integer programs as in Grunspan and Thomas [3] or does not involve complicated computations in deriving the cuts as in Granot and Granot [2]. The results to be followed consist of developing a method for enforcing additional constraints to a linear fractional program and then using it for solving pure and mixed integer linear fractional programs by cutting plane methods.

## 2. METHOD OF ENFORCING ADDITIONAL CONSTRAINTS

Consider the following linear fractional program:

$$(P): \quad \text{maximize} \quad \frac{\sum_{j=1}^{n} c_j x_j + \alpha}{\sum_{j=1}^{n} d_j x_j + \beta}$$

$$\text{subject to} \quad \sum_{j=1}^{n} a_{ij} x_j = a_{io}, \qquad i = 1, 2, \ldots, m,$$

$$x_j \geqslant 0, \qquad j = 1, 2, \ldots, n,$$

where it is assumed that $\sum_{j=1}^{n} d_j x_j + \beta > 0$ for all feasible solutions.

Suppose that a basic feasible solution to $(P)$ is known and with reference to this, let the constraints be

$$x_{B_i} = b_i + \sum_{j \in N} y_{ij}(-x_j), \quad i = 1, 2, \ldots, m, \tag{1}$$

$$x_j \geqslant 0, \quad j = 1, 2, \ldots, n,$$

where $x_{B_i}$, i = 1, 2, ..., $m$ are the basic variables, $N$ is the index set of nonbasic variables an the basic solution is obtained by putting $x_j = 0$ for every $j \in N$ in (1). Let $z^1$, $z^2$ be respectiv ly $\sum_{i=1}^{m} c_{B_i} b_i + \alpha$, $\sum_{i=1}^{m} d_{B_i} b_i + \beta$, $z_j^1 = \sum_{i=1}^{m} c_{B_i} y_{ij}$, $z_j^2 = \sum_{i=1}^{m} d_{B_i} y_{ij}$ and $\Delta_j = \{z^2(z_j^1 - c_j)\ z^1(z_j^2 - d_j)\}/z^2$ be calculated for all $j \in N$. Let the additional constraint to be appended be

$$\sum_{j=1}^{n} p_j x_j \geqslant q. \tag{2}$$

Substituting for $x_{B_i}$, $i = 1, 2, \ldots, m$ from (1) in (2) let (2) take the form

$$x_{n+1} = b_{m+1} + \sum_{j \in N} y_{m+1}(-x_j) \geqslant 0. \tag{3}$$

Let us denote by $(P_1)$ the new problem obtained from $(P)$ by appending (3). Then

$$\begin{cases} x_{B_i} = b_i, & i = 1, 2, \ldots, m+1 \\ x_j = 0, & j \in N \end{cases} \tag{4}$$

is a basic solution to $(P_1)$ where we designate $n + 1$ for $B_{m+1}$. Let $c_{n+1}$ and $d_{n+1}$ be assign the values zero.

The problem of interest here occurs when $b_{m+1} < 0$. To establish the validity of the su cessive steps of the method to be followed in reoptimizing the problem $(P_1)$, we require t following lemmas, the proofs of which can be constructed with the help of Hadley [4].

LEMMA 1. If $b_{m+1} < 0$ and if the set of feasible solutions to $(P_1)$ is nonempty, th there exists a $j \in N$ such that $y_{m+1,j} < 0$.

LEMMA 2. For every $j \in N$, there exists an i $\in \{1, 2, \ldots, m\}$ such that $y_{ij} > 0$.

LEMMA 3. If $b_{m+1} < 0$ and if there exists a $k \in N$ satisfying $y_{m+1,k} < 0$ and

$$\frac{b_{m+1}}{y_{m+1,k}} \leqslant \frac{b_r}{y_{rk}} = \operatorname*{Min}_{i \in \{1,2,\ldots,m\}} \left\{ \frac{b_i}{y_{ik}} \,\middle|\, y_{ik} \geqslant 0 \right\},$$

then by a change of basis, i.e., by replacing $x_k$ by $x_{B_{m+1}}$ in the set of basic variables a ba feasible solution to $(P_1)$ is obtained and

$$y_{m+1,k} - \frac{b_{m+1}}{z^2}(z_k^2 - d_k) < 0.$$

The method of reoptimizing $(P_1)$ can now be given as follows:

ALGORITHM 1.

STEP 1. Set up a tableau giving $x_{B_i}$, $b_i$, $y_{ij}$, $z_j^1 - c_j$, $z_j^2 - d_j \Delta_j$ for $i = 1, 2, \ldots, m$ and $j \in$ Compute $z^1$, $z^2$ and $z$. Append the additional constraint in the form (3).

STEP 2. If $b_{m+1} < 0$, set $J = 0$ and go to step 3. Otherwise, set $J = 1$ and go to step 8.

STEP 3. Set $M = \{j | j \in N,\ y_{m+1,j} < 0\}$. If $M$ is empty, no feasible solution exists, stop.

STEP 4. Compute $\dfrac{b_{r(j)}}{y_{r(j),j}} = \underset{i \in \{1,2,\ldots,m\}}{\text{Min}} \left\{ \dfrac{b_i}{y_{ij}},\ y_{ij} > 0 \right\}$ for every $j \in N$. Set $M_1 = \left\{ j | j \in M,\ \dfrac{b_{m+1}}{y_{m+1,j}} \leqslant \dfrac{b_{r(j)}}{y_{r(j),j}} \right\}$. If $M_1$ is empty go to step 6. Otherwise, set $J = 1$, $r = m + 1$ and go to step 5.

STEP 5. Find $k$ such that $\dfrac{\Delta_k}{|l_{m+1,k}|} = \underset{j \in M_1}{\text{Min}} \left\{ \dfrac{\Delta_j}{|l_{m+1,j}|},\ l_{m+1,j} < 0 \right\}$ where $l_{m+1,j} = y_{m+1,j} - \dfrac{b_{m+1}}{z^2}(z_j^2 - d_j)$ and go to step 7.

STEP 6. Find $k$ such that $y_{m+1,k} = \underset{j \in N}{\text{Min}}\ y_{m+1,j}$. If $y_{m+1,k} \geqslant 0$ no feasible solution exists; stop. Otherwise, find $r$ such that

$$\frac{b_r}{y_{rk}} = \underset{i \in \{1,2,\ldots,m\}}{\text{Min}} \left\{ \frac{b_i}{y_{ik}},\ y_{ik} > 0 \right\}.$$

STEP 7. Do a simplex pivoting to obtain a new basic solution by replacing $x_{B_r}$ by $x_k$ in the set of basic variables. Modify $N$, calculate $\Delta_j$, $z_j^1 - c_j$, $z_j^2 - d_j$ for all $j \in N$ by modifying their definitions to take the summation from 1 to $m + 1$. Also calculate $z^1$, $z^2$ and $z$. If $J = 0$ return to step 3.

STEP 8. If $\Delta_j \geqslant 0$ for all $j$, stop; the optimal solution is $x_{B_i} = b_i$ for $i = 1, 2, \ldots, m + 1$, and $x_j = 0$ for $j \in N$. Otherwise let $\Delta_k = \underset{j \in N}{\text{Min}}\ \Delta_j$, $\dfrac{b_r}{y_{rk}} = \underset{i \in \{1,2,\ldots,m+1\}}{\text{Min}} \left\{ \dfrac{b_i}{y_{ik}},\ y_{ik} > 0 \right\}$ and return to step 7.

## 3. FRACTIONAL CUTTING PLANE METHOD FOR INTEGER LINEAR FRACTIONAL PROGRAMS

As an application of Algorithm 1 we present in this section a cutting plane method for solving integer linear fractional programs. The method for (mixed) integer linear fractional programs follows closely Gomory's fractional cutting plane method for (mixed) integer linear programs and can briefly be described as follows.

Solve the problem obtained by omitting the integer restrictions. If the solution satisfies the integer restrictions, then it is optimal; otherwise, introduce a Gomory's (mixed integer) fractional cut and reoptimize by using Algorithm 1 and repeat the process.

REMARK. Enforcing a single additional constraint to a linear fractional program can be done easily by using the dual simplex method after employing the Charnes and Cooper [1] transformation. But in the case of integer linear fractional programs, the transformation destroys the integer nature of the variables and Gomory's fractional cuts in their usual forms cannot be used. However, cuts can be obtained as given by Granot and Granot [2], but it can be observed that such derivations of cuts involve more complicated computations than those in our case.

Since the objective function is not integer constrained, the proof of finiteness of the above method cannot possibly be given in a manner similar to that of Gomory's methods.

## ACKNOWLEDGMENT

The authors wish to thank the referee for his suggestions for improvement of an earlier version of this paper.

## REFERENCES

[1] Charnes, A., and W.W. Cooper, "Programming with Linear Fractional Functionals," Naval Research Logistics Quarterly, Vol. *9*, pp. 181-186 (1962).

[2] Granot, D., and F. Granot, "On Integer and Mixed Integer Fractional Programming Problems," *Annals of Discrete Mathematics 1, Studies in Integer Programming*, pp. 221-231, eds. P.L. Hammer et al. (North Holland Publishing Company, 1977).

[3] Granspan, M., and M.E. Thomas, "Hyperbolic Integer Programming," Naval Research Logistics Quarterly," Vol. *20* (2), pp. 341-356 (1973).

[4] Hadley, G., "Linear Programming," (Addison Wesley, Reading, Massachusetts, 1962).

[5] Swarup, K., "Linear Fractional Functional Programming," Operations Research, *13*, pp. 1029-1036 (1965).

[6] Swarup, K., "Some Aspects of Linear Fractional Functional Programming," Australian Journal of Statistics, Vol. *7*, pp. 90-104 (1965).

# NEWS AND MEMORANDA

## THE 1979 LANCHESTER PRIZE

### Call for Nominations

Each year since 1954 the Council of the Operations Research Society of America has offered the Lanchester Prize for the best English-language published contribution in operations research. The Prize for 1979 consists of $2,000 and a commemorative medallion.

The screening of books and papers for the 1979 Prize will be carried out by a committee appointed by the Council of the Society. To be eligible for consideration, the book or paper must be nominated to the Committee. Nominations may be made by anyone; this notice constitutes a call for nominations.

To be eligible for the Lanchester Prize, a book, a paper or a group of books or papers must meet the following requirements:

(1) It must be on an operations research subject,

(2) It must carry a current award year publication date, or, if a group, or at least one member of the group must carry a current award year publication date,

(3) It must be written in the English language, and

(4) It must have appeared in the open literature.

The *book(s)* or *paper(s)* may be a case history, a report of research representing new results, or primarily expository.

For any nominated set (e.g., article and/or book) covering more than the most recent year, it is expected that each element in the set represents work from one continuous effort, such as a multi-year project or a continuously written, multi-volume book.

Judgments will be made by the Committee using the following criteria:

(1) The magnitude of the contribution to the advancement of the state of the art of operations research,

(2) The originality of the ideas or methods,

(3) New vistas of application opened up,

(4) The degree to which unification or simplification of existing theory or method is achieved, and

(5) Expository clarity and excellence.

Nominations should be sent to:

Marshall L. Fisher, Chairman
1979 Lanchester Prize Committee
Department of Decision Sciences
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104

Nominations may be in any form, but must include as a minimum the title(s) of the paper(s) or book, author(s), place and date of publication, and six copies of the material. Supporting statements bearing on the worthwhileness of the publication in terms of the five criteria will be helpful; but are not required. Each nomination will be carefully screened by the Committee; nominations must be received by May 30, 1980, to allow time for adequate review.

Announcement of the results of the Committee and ORSA Council action, as well as award of any prize(s) approved, will be made at the 58th National Meeting of the Society, November 10-12, 1980 in Colorado Springs, Colorado.

## INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.

NAVAL RESEARCH
LOGISTICS
QUARTERLY

MARCH 1980
VOL. 27, NO. 1
NAVSO P-1278

CONTENTS



OFFICE OF NAVAL RESEARCH
Arlington, Va. 22217