A RAND NOTE

MODELS AND TECHNIQUES FOR RECOVERABLE ITEM STOCKAGE WHEN DEMAND AND THE REPAIR PROCESS ARE NONSTATIONARY--PART I: PERFORMANCE MEASUREMENT

R. J. Hillestad and M. J. Carrillo

May 1980

N-1482-AF

Prepared For

The United States Air Force

Rand
SANTA MONICA, CA. 90406

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The research reported here was sponsored by the Directorate of Operational Requirements, Deputy Chief of Staff/Research, Development, and Acquisition, Hq USAF, under Contract F49620-77-C-0023. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

The Rand Publications Series: The Report is the principal publication documenting and transmitting Rand's major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation

A RAND NOTE

MODELS AND TECHNIQUES FOR RECOVERABLE ITEM STOCKAGE WHEN DEMAND AND THE REPAIR PROCESS ARE NONSTATIONARY--PART I: PERFORMANCE MEASUREMENT

R. J. Hillestad and M. J. Carrillo

May 1980

N-1482-AF

Prepared For

The United States Air Force

Rand
SANTA MONICA, CA. 90406

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

PREFACE

The Air Force has been increasingly concerned with logistics support capability under very demanding and dynamic scenarios. Unfortunately, many measures of performance used in evaluating logistics support and methods of measurement have evolved during relative nondynamic peacetime operations and conflicts in which demand for logistics support was generally stable. These "steady state" methods have been adapted to evaluating dynamic performance by increasing demand for support to higher levels and assuming that the dynamic scenarios are adequately approximated by peacetime steady state demands at those higher levels. This assumption could lead to the overestimation (and in some cases, underestimation) of wartime spares requirements, potentially incorrect war reserve spares mixes for squadrons, and, at the least, inaccurate capability assessment of alternative logistics policies.

This report develops techniques of transient performance measurement for alternative supply and maintenance strategies. A companion document, Part II: Component and Subcomponent Spares Allocation with Dynamic Models, will describe the use of these models for determining spare stockage requirements.

Rand-sponsored research led to the dynamic modeling techniques described in this report. The techniques were used by Carrillo, Hillestad and Lippiatt in "Maintenance and Supply Performance Explorations," found in Appendix E in [11] and were further expanded to support the analysis in the forthcoming publication "Assessing the Capability of Planned Reserve Stocks and Spare Engines to Support High Sortie Rates in a Central European Contingency." Because the demands for the Central European contingency are expected to be time-varying, because squadrons are collocated with potentially shared resources, and because the availability of stockage and maintenance repair resources is time-varying, further expansion of the capability to study transitory behavior of support resources was necessary.

The models and techniques described in this report are rather technical and a background in probability theory is necessary for a full understanding of the necessary assumptions, development of the models, and some implications of their use. The report should be of greatest interest to those specialists involved in studies and analysis of dynamic logistics support requirements, those involved in determining wartime stockage (including spare engines) requirements, and those wanting an in-depth understanding of the models used by Rand in the analysis.

This work was done as part of the Project AIR FORCE project "Strategies to Improve Sortie Production in a Dynamic Wartime Environment."

CONTENTS

PREFACE iii

Section

I. INTRODUCTION AND SUMMARY 1

II. THE BASIC RESULT - TIME DEPENDENCY OF NUMBER OF COMPONENTS IN A QUEUE 5

An Aircraft Engine Example 10

III. TIME DEPENDENT MEASURES OF PERFORMANCE 16

Appendix

A. NONSTATIONARY SERVICE DISTRIBUTIONS 25

B. SOME UNRESOLVED PROBLEMS 32

REFERENCES 35

I. INTRODUCTION AND SUMMARY

Much of the past work regarding recoverable item stockage has concentrated on the development of models and policies for systems in steady state. The well known work of Feeney and Sherbrooke [3] regarding an (s-1, s) inventory policy developed distribution functions for items in service and measures of inventory performance based on the limiting behavior of an infinite server queue expressed by Palm's theorem [10]. Sherbrooke's METRIC: A Multi-Echelon Technique for Recoverable Item Control [14] used the same assumptions regarding an inventory system in steady state with constant average demand rate and service rate. More recently the MOD-METRIC model of Muckstadt [9] expanded the capability of METRIC by considering indentured components but also assumed that recoverable item stockage was to be provided for a system in steady state.

These models and variants have been widely applied to practical problems of inventory management. For example, the Air Force uses METRIC-like techniques for the management of millions of dollars of recoverable spare parts. For many inventory systems including the Air Force Supply System in peacetime the assumption of steady state behavior is both convenient and adequate. However, there are important situations in which the transient behavior is most important. A dramatic example of this is the potential dynamic behavior exhibited by demands and service in the deployment of an Air Force squadron at the onset of a conflict. Demands for components may suddenly jump very high relative to the previous peacetime activity of the squadron and then decay gradually or abruptly due to the attrition of the aircraft in the squadron. Meanwhile the initial service rate may be near zero, as the already deployed unit awaits the airlift of the specialized personnel and test equipment to repair broken components, gradually increase to its full wartime service capability, and suddenly drop to near zero again as a result of damage in an airbase attack.

Two problems regarding recoverable item inventory faced by a decisionmaker for the nonstationary case are the same as those faced in the steady state case, that is, to determine how much spare stock to provide against stockouts and to determine what level of performance can be achieved with a given level of investment in spare stock. Clearly, the limiting behavior of the dynamic case is not of interest when that limit is reached at times far beyond the period of interest. Performance measured with models which only consider limiting behavior will tend to be inaccurate and stockage provided against limiting behavior may be either inadequate or excessive.

Gaver, Lehorsky, and Perlas [5] have described certain service systems with transitory demands. Miller [11] has described a real-time management approach for the distribution of stockage among locations affected by time varying demands. The primary purpose of this Note (Part I) is to provide an integrated approach to inventory performance measurement for a given stockage allocation for systems with nonstationary demands and service rates. Part II will describe certain stock allocation algorithms for the nonstationary case.

A secondary purpose of the Note is to suggest approaches to certain aspects of recoverable item repair and supply which currently cause significant deviation between practice and theory, even in the case of stationary demands and service rates. In particular, we will address the problems of cannibalization of primary recoverable items by providing models for different degrees of cannibalization.

Cannibalization of components is common practice in many service systems, particularly in cases where the equipment to be repaired is highly modular and interchangeable with a minimum of effort. In effect it provides an additional source of spare components when stock is low or service times are long. For example, consider the case of an aircraft not operational due to the shortage of one item. While that aircraft is nonoperational it also provides a hedge against other aircraft becoming nonoperational due to shortages of other items. In the same fashion indentured items awaiting service provide a secondary source of supply of subcomponents. The METRIC and MOD-METRIC techniques do not consider the effect that cannibalization has on

inventory system performance and therefore tend to overestimate stockage for systems which permit cannibalization. We will describe how cannibalization can be considered in evaluating performance.

We will also describe a group of performance measures which attempt to predict the effect that inventory and service policy has on organizational performance. Some of these measures have been developed for the special case of aircraft support in which the organizational goal is to provide enough operational aircraft to perform missions. (However, it should be quite easy to adapt them to support of other equipment.) Because the measures deal with end item performance and demands they are scenario dependent.* For example, stockout of aircraft components may not be important if the aircraft are not required to fly large numbers of flights since only a few operational aircraft are needed to perform the activity. Thus, we will provide a model with the potential to evaluate inventory policy in the context of organizational goals.

Finally, in Appendix B we have indicated certain shortcomings in our models and other currently used models which have not been overcome as far as we know. We have also suggested the possible effects of these shortcomings.

Certain parts of the models in this paper have been used for evaluation of ship and shore support of the carrier aircraft in the Defense Resource Management Study by Rice [11], for analysis of centralized maintenance support of the A-10 aircraft in the same study, and for analysis of recoverable item stockage requirements for various levels of potential wartime flying demands of the F-4 aircraft.

The next section gives the derivation of a basic results regarding the time dependency of recoverable items in a queue. This

---

*See Appendix A for the methodology needed to include scenario information in the model.

is followed by a section describing a group of time dependent measures of performance.

II. THE BASIC RESULT - TIME DEPENDENCY OF NUMBER OF COMPONENTS IN A QUEUE

In this section we will obtain the probability distribution of the number of recoverable items in service when the demand rate for service and the service rate are nonstationary and there is sufficient reserve service capacity. This distribution will then be used in calculating time varying performance measures during periods of transitory demand and service. First, we review a related steady state result.

The well known queuing theorem due to Palm ([10], see also [15]) states that when demands arrive according to a Poisson process with rate λ and the service times have an arbitrary probability distribution independent of the demand process with mean T then the steady state probability of k units in service under the assumption of slack repair capacity is given by a Poisson distribution with parameter λT, that is,

$$\text{probability of k units in service} = P_k = \frac{(\lambda T)^k e^{-\lambda T}}{k!}.$$

Palm provides a similar result for the lost sales case in which demands stop when there is stockout. Feeney and Sherbrooke [4] extend these results for the case of the compound Poisson distribution. All of these results are valid only in the limit and when both the demand and service distribution are stationary. We will derive a similar result for time varying service and demand process.

We begin our derivation for the nonstationary case with the nonhomogeneous Poisson process. A demand process is said to be a nonhomogeneous Poisson process with intensity function $m(t)$, $t \geq 0$ if

i. The number of demands existing at time $t = 0$ is zero.

ii. The numbers of demands in disjoint time increments are independent of each other.

iii. The probability of more than one demand in an increment becomes infinitesimally small as the increment gets small. (A compound Poisson process relaxes this assumption.)

iv. The probability of one demand in any increment is given by the intensity function m(t) times the length of the increment as the increment gets small.

Then, if we let

$$N(t) = \text{Number of demands by time, } t$$

$$\gamma(t) = \int_0^t m(s)ds$$

it can be shown (see Ross [13], for example) that

Probability of k demands in the interval t, t + s is

$$P[N(t+s) - N(t) = k] = e^{-(\gamma(t+s)-\gamma(t))} \frac{[\gamma(t+s) - \gamma(t)]^k}{k!}$$

for $k \geq 0$

$\gamma(t)$ is called the mean value function of the process.

The following result is stated in an exercise in [12].

Theorem: Given N(t) = k, the unordered set of arrival times in a nonhomogeneous Poisson process with mean value function $\gamma(t)$ have the same distribution as k independent and identically distributed random variables having distribution function

$$G(x) = \begin{cases} \frac{\gamma(x)}{\gamma(t)} & x \leq t \\ 1 & x > t \end{cases}$$

Theorem. (Nonhomogeneous Poisson Queue)*

Let X(t) denote the number of customers in the system at time t when arrivals occur according to a nonhomogeneous Poisson process with intensity function m(t) and mean value function

$$\gamma(t) = \int_0^t m(s)ds.$$

Let the nonstationary service distribution be F(s,t) (the probability that a service started at time s will be completed by time t). Then, with the additional assumptions that the repair process is independent of the arrival process and slack repair capacity, X(t) has a Poisson distribution with mean

$$\int_0^t \bar{F}(s,t)d\gamma(s) = \int_0^t \bar{F}(s,t)m(s)ds$$

where

$$\bar{F}(s,t) = 1 - F(s,t).$$

Proof:

Conditioning on N(t) we obtain the probability that X(t) = j,

$$P(X(t) = j) = \sum_{k=0}^{\infty} P(X(t) = j \mid N(t) = k) \frac{e^{-\gamma(t)}\gamma(t)^k}{k!}$$

---

*A version of this theorem as well as a sketch of a proof (different from the one given here) was first seen by the authors in some unpublished notes by Gordon Crawford. Also, see Crawford [2] and Gaver [6].

If there have been k arrivals by time t, then the probability that an arbitrary one of these customers will still be present at t is given by

$$p = \int_0^t \overline{F}(s,t)\, d\left(\frac{\gamma(s)}{\gamma(t)}\right) .$$

This is obtained by conditioning on the time of arrival of an arbitrary customer and then unconditioning using the distribution G(x) given earlier.

Now,

$$P(X(t) = j \,|\, N(t) = k) = \begin{cases} \binom{k}{j} p^j (1-p)^{k-j} & j = 0, 1, \ldots, k \\ 0 & j > k \end{cases}$$

Unconditioning on N(t), we get

$$P(X(t) = j) = \sum_{k=j}^{\infty} \binom{k}{j} p^j (1-p)^{k-j} \frac{e^{-\gamma(t)} \gamma(t)^k}{k!}$$

$$= e^{-\gamma(t)} \frac{(p \cdot \gamma(t))^j}{j!} \sum_{k=j}^{\infty} \frac{[(1-p)\gamma(t)]^{k-j}}{(k-j)!}$$

$$= e^{-\gamma(t)} \frac{(p \cdot \gamma(t))^j}{} \; e^{(1-p)\gamma(t)}$$

$$= e^{-p\gamma(t)} \frac{[p \cdot \gamma(t)]^j}{j!}$$

The proof is then complete since

$$p\gamma(t) = \left[\int_o^t \bar{F}(s,t) d \frac{\gamma(s)}{\gamma(t)} \right] \gamma(t)$$

$$= \int_o^t F(s,t)m(s)ds.$$

In the work that follows we define

$$\lambda(t) = \int_o^t \bar{F}(s,t)m(s)ds.$$

The importance of the result lies in the fact that nonstationary demands, described by the parameter m(t), and nonstationary service time, given by $\bar{F}(s,t)$, are both represented in the model.

If demands in the Poisson process have a probability f of being registered independent of the process, the number $\tilde{X}$ of registered demands in service is given by

$$P(\tilde{X}(t) = k) = \frac{[f \cdot \lambda(t)]^k e^{-[f \cdot \lambda(t)]}}{k!}.$$

In fact, f can also be nonstationary so that f = f(t) and we compute

$$\tilde{\lambda}(t) = \int_o^t \bar{F}(s,t)f(s)m(s)ds$$

and then use

$$P(\tilde{X}(t) = k) = \frac{\tilde{\lambda}(t)^k e^{-\tilde{\lambda}(t)}}{k!}.$$

Furthermore, if we let X(t) and Y(t) be the number of customers in service in two independent processes it can be shown that the total number in service is given by the Poisson distribution,

$$P(X(t) + Y(t) = k) = \frac{[\lambda_X(t)+\lambda_Y(t)]^k \, e^{-[\lambda_X(t)+\lambda_Y(t)]}}{k!}.$$

We will use the following example to give some perspective on these results.

AN AIRCRAFT ENGINE EXAMPLE

Assume that aircraft engines at an airbase demand service at an average rate m(t) where

$$m(t) = \begin{cases} m_1 & t < \tau \\ m_2 & t \geq \tau \end{cases}$$

and that τ reflects a change in the base flying program causing a change in the rate of demand of service. Assume that an arbitrary service demand has a probability f of being satisfied at the base and must be satisfied at a remote location with probability 1-f. Assume further that the service time distribution functions for base and remote locations are exponential with average rates μ_B and μ_R. Finally, assume that engines demanding service at the remote location cannot leave the base for the first T (assume $T < \tau$) days of the period we are considering. Then we have, for base service

$$\overline{F}_B(s,t) = e^{-\mu_B(t-s)}$$

and for remote service

$$\bar{F}_R(s,t) = \begin{cases} 1 & s \leq t < T \\ e^{-\mu_R(t-T)} & s < T \leq t \\ e^{-\mu_R(t-s)} & T \leq s \leq t \end{cases}$$

The average number of engines in the base repair process is then given by

$$\lambda_B(t) = \int_o^t \bar{F}_B(s,t)m(s)ds$$

$$\lambda_B(t) = \int_o^t e^{-\mu_B(t-s)} m_1 f ds = f \cdot \frac{m_1}{\mu_B}(1-e^{-\mu_B t}) \qquad \text{for } t < \tau$$

$$\lambda_B(t) = \int_o^t e^{-\mu_B(t-s)} m_1 f ds + \int_\tau^t e^{-\mu_B(t-s)} m_2 f ds \qquad \text{for } t \geq \tau$$

$$= f \cdot \frac{m_1}{\mu_B}(1-e^{-\mu_B \tau})\, e^{-\mu_B(t-\tau)}$$

$$+ f \cdot \frac{m_2}{\mu_B}(1-e^{-\mu_B(t-\tau)}) \qquad \text{for } t \geq \tau$$

and the average number of engines in repair at the remote location or awaiting repair at the remote location is

$$\lambda_R(t) = \int_o^t \bar{F}_R(s,t)m(s)ds$$

$$\lambda_R(t) = \int_o^t 1 \cdot m_1(1-f)ds = (1-f)m_1 t \qquad \text{for } t < T$$

$$\lambda_R(t) = \int_o^T e^{-\mu_R(t-T)} m_1(1-f)ds + \int_T^t e^{-\mu_R(t-s)} m_1(1-f)ds$$

$$= (1-f)m_1 T\, e^{-\mu_R(t-T)} + (1-f)\frac{m_1}{\mu_R}(1-e^{-\mu_R(t-T)}) \qquad \text{for } T \leq t < \tau$$

$$\lambda_R(t) = \int_o^T e^{-\mu_R(t-T)} m_1(1-f)ds + \int_T^\tau e^{-\mu_R(t-s)} m_1(1-f)ds$$

$$+ \int_\tau^t e^{-\mu_R(t-s)} m_2(1-f)ds$$

$$= (1-f)m_1 T\, e^{-\mu_R(t-T)}$$

$$+ (1-f)\frac{m_1}{\mu_R}(1-e^{-\mu_R(\tau-T)})\, e^{-\mu_R(t-\tau)}$$

$$+ (1-f)\frac{m_2}{\mu_R}(1-e^{-\mu_R(t-\tau)}) \qquad \text{for } \tau \leq t$$

The average total engines in service is then given by

$$\lambda_T(t) = \lambda_B(t) + \lambda_R(t)$$

and the distribution of engines in or awaiting service as a function of time is

$$P(X_B(t) + X_R(t) = k) = \frac{\lambda_T(t)^k\, e^{-\lambda_T(t)}}{k!}$$

Figure 1 shows the mean number being serviced as a function of time. Note that as t gets larger this result converges to the steady state result predicted by Palm's theorem,

$$\lambda_T(\infty) = f \frac{m_2}{\mu_B} + (1-f) \frac{m_2}{\mu_R}.$$

Note also that the dynamic result gives considerably more information regarding the distribution of engines in service during the period of transient behavior. It can be used, for example, to determine the sensitivity to base and remote repair rates, the effect of transition to a different flying rate (with subsequently different demand rate) and the effect of various delays in shipping engines off base. Furthermore, it can be used to determine the time dependency of the probability of stockout based on a time varying supply of spare engines.

In this example we assumed that there were no engines requiring service initially. However, by superposition* it is possible to show that if there was an average of Γ_0 engines requiring service initially, the mean number in service at time t is

$$\lambda(t) = \int_0^t \bar{F}(s,t)m(s)ds + \bar{F}(0,t)\Gamma_0$$

provided that service times have an exponential distribution.

Appendix A gives a number of closed form solutions for $\lambda(t)$ given various assumptions about $\bar{F}(s,t)$ and $m(s)$.

---

*Consider two separate repair processes, one which operates prior to t = 0, and another which operates after t = 0. The second term in the equation above represents the servicing of all demands which have occurred and not been serviced by time t = 0.

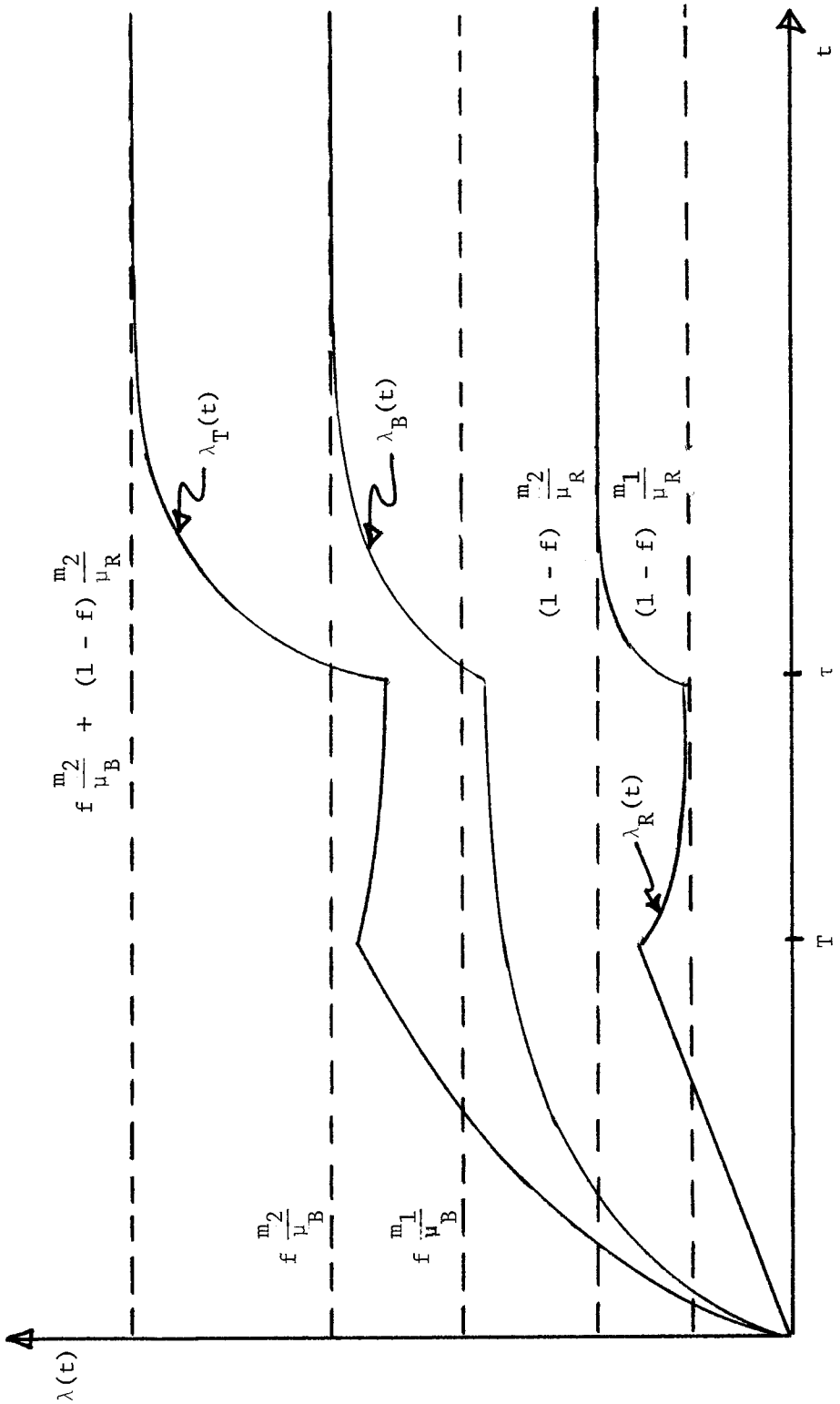

Fig. 1--Average number of engines being serviced as a function of time

Once an expression or value of $\lambda(t)$ is determined it is possible to compute other measures of performance such as the probability of stockout given a spare stock level, $S(t)$, at time t. The next section derives a number of performance measures for the nonstationary inventory system.

III. TIME DEPENDENT MEASURES OF PERFORMANCE

In this section we describe a number of performance measures of a nonstationary inventory system*, some of which are the well known measures such as average backorders, while others are more specifically related to system supporting activity of vehicles such as aircraft. Let $\lambda_i(t)$ represent the mean number of a single recoverable item, i, undergoing service as given in the previous section and let $S_i(t)$ be the amount of spare stock provided for that item at time t. (We assume here that $S_i(t)$ is a deterministic quantity.) Then several inventory system measures are:

a) Probability of stockout at time t, $PO_i(t)$

This is the probability that demands exceed supply and is given by

$$PO_i(t) = \sum_{k=S_i(t)+1}^{\infty} \lambda_i(t)^k \frac{e^{-\lambda_i(t)}}{k!}$$

$$= 1 - \sum_{k=0}^{S_i(t)} \lambda_i(t)^k \frac{e^{-\lambda_i(t)}}{k!}$$

b) Fill rate at time t, $FR_i(t)$

This is the probability that a demand that has occurred can be filled

$$FR_i(t) = \sum_{k=0}^{S_i(t)-1} \lambda_i(t)^k \frac{e^{-\lambda_i(t)}}{k!}$$

---

*See [1] for steady state supply performance measures.

c) Average backorders at time t, $EB_i(t)$

This gives the average level of shortages of item i at time t

$$EB_i(t) = \sum_{k=S_i(t)+1}^{\infty} [k - S_i(t)] \frac{\lambda_i(t)^k e^{-\lambda_i(t)}}{k!}$$

$$= \lambda_i(t) - S_i(t) + \sum_{k=0}^{S_i(t)} (S_i(t) - k) \frac{\lambda_i(t)^k e^{-\lambda_i(t)}}{k!}$$

d) Variance in backorders at time t

$$VB_i(t) = \sum_{k=S_i(t)+1}^{\infty} [k - S_i(t)]^2 \frac{\lambda_i(t)^k e^{-\lambda_i(t)}}{k!} - EB_i(t)^2$$

$$= \lambda_i(t) + [\lambda_i(t) - S_i(t)]^2$$

$$- \sum_{k=0}^{S_i(t)} (k - S_i(t))^2 \frac{\lambda_i(t)^k e^{-\lambda_i(t)}}{k!} - EB_i(t)^2$$

e) Average number of systems NMCS (Not Mission Capable for Supply reasons) without cannibalization, EN(t)

This measure gives the average combined effect of recoverable item shortages on the major system (such as aircraft) those items support. Here we assume that NA(t) represents the number of such major systems which are supported at time t. We further assume that there are N types of recoverable items required on each system and that the shortages of any one of these items will make the system nonoperational. In the noncannibalization case we assume that shortages of components cannot be consolidated among the systems.

The probability that an arbitrary system has a shortage of item i when there are k shortages of i is

$$\frac{k}{NA(t)}$$

and therefore, the probability that an arbitrary system has a shortage of item i is

$$\frac{EB_i(t)}{NA(t)}$$

The probability that a system is down due to shortages of some item is

$$1 - \prod_{i=1}^{N}\left(1 - \frac{EB_i(t)}{NA(t)}\right)$$

and, finally, the expected number of systems not operational at time t is

$$EN(t) = NA(t)\left[1 - \prod_{i=1}^{N}\left(1 - \frac{EB_i(t)}{NA(t)}\right)\right]$$

This derivation required the assumption that only one component of given type was on each system. If this is not the case let Q_i be the quantity of item i per system. If no consolidation of shortages occurs then

$$EN(t) = NA(t)\left[1 - \prod_{i=1}^{N} \frac{\sum_{j=0}^{\infty}\binom{Q_i NA(t) - y}{Q_i} PB_i(y)}{\binom{Q_i NA(t)}{Q_i}}\right]$$

where $PB_i(y)$ is the probability that item i has y shortages at time t

$$PB_i(y) = \begin{cases} \sum_{k=0}^{S_i(t)} \frac{e^{-\lambda_i(t)} \lambda_i^k(t)}{k!} & y = 0 \\ \frac{e^{-\lambda_i(t)} \lambda_i^{S_i(t)+y}(t)}{(S_i(t)+y)!} & y > 0 \end{cases}$$

f) NORS with complete cannibalization

Here we assume that shortages of all components are consolidated to make the smallest number of nonoperational systems. Let $P^i(j)$ be the probability that shortages of the i^{th} item are less than or equal to j. Then

$$P^i(j) = \sum_{k=0}^{S_i(t)+j} \frac{\lambda_i(t)^k e^{-\lambda_i(t)}}{k!}$$

Let P(j) be the probability that the number of nonoperational (NORS) systems is less than or equal to j. Then

$$P(j) = \prod_{i=1}^{N} P^i(j)$$

If there is more than one item of a type on each system we again employ Q_i as the number of item i per system and obtain

$$P(j) = \prod_{i=1}^{N} P^{i}(Q_i \cdot j).$$

The average number of NORS systems with full cannibalization is then

$$EN_c(t) = \sum_{j=0}^{NA(t)-1} [1 - P(j)].$$

The NORS distribution function with full cannibalization is

$$PN_j(t) = P(j) - P(j - 1)$$

which then can be used to give the variance in the number of NORS systems.

$$VN(t) = \left[\sum_{j=1}^{NA(t)} j^2 PN_j(t) \right] - EN_c^2(t)$$

g) NORS with partial cannibalization

Here we assume that some items are relatively easy to cannibalize and that some are so difficult to remove or install that it is not desirable to cannibalize them. Let

$$I_c = \{ i \mid \text{item } i \text{ is cannibalizable} \}$$

$$I_n = \{ i \mid \text{item } i \text{ is not cannibalizable} \}$$

We first compute the probability of exactly k NORS systems due to items in the set I_c. This is just the $PN_j(t)$ derived earlier using P(j) and where P(j) is computed only using the i in I_c. We will denote this probability $PN_j^c(t)$. The probability that an arbitrary system is operational considering only shortages of cannibalizable items is then

$$\sum_{j=1}^{NA(t)} \frac{j}{NA(t)} PN_j^c(t) = \frac{EN^c(t)}{NA(t)}.$$

The probability that an arbitrary system is operational after shortages of noncannibalizable components only is

$$\prod_{i \varepsilon I_n} \left(1 - \frac{EB_i(t)}{NA(t)}\right)$$

assuming Q_i, the quantity of item i per system, is one for each part, i, in the set I_n. The probability that an arbitrary system is not operational, assuming independence of demands and that the cannibalization of the items in I_c takes place with no information of the failed items belonging to I_n, is

$$1 - \left[1 - \frac{EN^c(t)}{NA(t)}\right]\left[\prod_{i \varepsilon I_n} \left(1 - \frac{EB_i(t)}{NA(t)}\right)\right]$$

and the expected number of nonoperational systems given partial cannibalization and given NA(t) total systems is

$$EN^p(t) = NA(t)\left(1 - \left[1 - \frac{EN^c(t)}{NA(t)}\right]\left[\prod_{i \varepsilon I_n} \left(1 - \frac{EB_i(t)}{NA(t)}\right)\right]\right).$$

$EN^P(t)$ is an overestimation of the expected NORS if the information describing which noncannibalizable items had failed is used (since, in this case, the strategy would be to move "holes" due to cannibalizable items to those aircraft with holes due to noncannibalizable items. We did not do this in our derivation).

In the NORS expressions found in sections e, f, and g, the reader may observe that backorders and shortages probabilities have been computed as though demands were independent of the number of systems nonoperational. In most systems the demand rate lowers as the number of systems nonoperational increases due to lower activity of the systems. Unfortunately, the independent increment assumption of the Poisson process is then violated. Usually, with spare stock provided for the appropriate level of activity, the percentage of systems nonoperational is low so that there is a very low probability that failures exceed spare stock plus the number of systems. For cases in which NORS percentages get high, it is possible to use the lost sales distribution with the demand turned away when the number of demands exceeds spare stock plus the number of systems.

h) Probability of meeting system demands and number of system demands met

This group of performance measures is for the situation in which it is desirable to have enough systems operational to perform a certain demanded level of activity. In the case of aircraft systems this might be the number of operational airframes which permits a desired level of flying activity. Let $B(t)$ be the maximum number of allowable NORS systems which still permits the system demands to be met and let $D(t)$ be the level of system demands at time t. Let $r(t)$ be the amount of system demand which can be satisfied by a single system at time t. In the aircraft case $D(t)$ would be the total desired sorties per unit time while $r(t)$ would be the maximum number of sorties per unit time achievable by a single aircraft.

Note that

$$B(t) = NA(t) - \left[\frac{D(t)}{r(t)}\right]$$

where [x] is the ceiling of x, that is the smallest integer y so that $y \geq x$, and $D(t) \leq r(t)$. NA(t) is assumed.

i. The probability of meeting system demands is then

$$PD(t) = P(B(t))$$

That is, the probability that system demands are met is just the probability that the number of nonoperational systems is less than or equal to B(t).

ii. The expected number of demands met given k nonoperational system is

$$ES(t)/k = \begin{cases} D(t) & \text{if } k \leq B(t) \\ r(t)(NA(t) - k) & \text{if } k > B(t) \end{cases}$$

Unconditioning

$$ES(t) = D(t)P(B(t)) + \sum_{k=B(t)+1}^{NA(t)} r(t)(NA(t) - k)PN_k(t)$$

where P(j) and $PN_k(t)$ were described earlier.

iii. The distribution of the number of demands met is given by

$$PS_k(t) = \begin{cases} P(B(t)) & \text{if } k = D(t) \\ PN_j(t) & \text{if } k = r(t)(NA(t) - j); \\ & j = B(t) + 1, \ldots, NA(t) \\ 0 & \text{otherwise} \end{cases}$$

iv. The variance in demands met is given by

$$VS(t) = D^2(t)\ P(B(t)) + \sum_{k=B(t)+1}^{NA(t)} r^2(t)\ (NA(t) - k)^2\ PN_k(t) - ES^2(t)$$

Note that the latter group of measures bring into play, in a deterministic sense, additional time varying parameters such as the number of systems, NA(t), the demands on all systems, D(t), and the portion of demand a single system can satisfy, r(t). In the deployment of an aircraft squadron, NA(t) might represent a time phased deployment of the aircraft as well as attrition due to losses in a conflict. The demand D(t) would be the time dependent demand for sorties as an engagement proceeds and r(t) would be time varying because of a short-term capability to provide surges of flying activity.

APPENDIX A

NONSTATIONARY SERVICE DISTRIBUTIONS

In this section, we will present various non-stationary service distributions which are of interest in evaluating the sensitivity of a recoverable item inventory system where the demands for items arrive according to a nonhomogeneous Poisson process, implying changing rates of demand for service. Under the assumption of sufficient slack service capacity and the independence of the demand and service processes, it was shown, in the main sections of this report, that for such systems the number of items in service at time t is given by the integral

$$\lambda(t) = \int_0^t \bar{F}(s,t)m(s)ds$$

where $\bar{F}(s,t)$ is the probability that a service started at time s will not be completed before time t and m(s) is the intensity function of the Poisson process.

The service distributions to be studied have one main property in common: the service is either nonexistent for all or some period of time or a service process is available with service times that have an exponential distribution with rate μ. This property will allow us to write $\lambda(t)$ in terms of two functions:

$$\gamma(t) = \int_0^t m(s)ds,$$

the mean value function at time t of the nonhomogeneous Poisson process, and

$$\Phi_\mu(t) = \int_0^t e^{-\mu(t-s)}m(s)ds$$

which is discussed next.

A) Basic Case:

A1. Uninterrupted service process with exponential service times, i.e., $F(s,t) = 1 - e^{-\mu(t-s)}$ for $0 \leq s \leq t$.

A2. Constant demand rate m, that is

$$m(s) = \begin{cases} 0 & s \leq 0 \\ m & s > 0 \end{cases}$$

Clearly, in this case

$$\gamma(t) = \int_0^t m(s)ds = m\ t.$$

The number of items in service is a Poisson random variable with mean

$$\lambda(t) = \int_0^t e^{-\mu(t-s)} m(s)ds = \Phi_\mu(t)$$

For the assumed m(t), we get

$$\Phi_\mu(t) = \frac{m}{\mu}\ (1-e^{-\mu t})$$

Note that for large t, when the system has been operating for a long time, we have that the number of items in service approximates $\frac{m}{\mu}$, the steady state result.

B) A More General Demand Rate for Service:

B1. Same as A1.

B2. Piecewise constant m(t), i.e.,

$$m(s) = \begin{cases} 0 & s \leq 0 \\ m_i & i-1 < t \leq i, \text{ for } i=1, \ldots, t, \text{ t assumed integral} \end{cases}$$

so that the mean value function

$$\gamma(t) = \int_0^t m(s)ds = \sum_{i=1}^{t} m_i \quad \text{for integral } t > 0.$$

Here, the mean number in service, $\lambda(t)$, becomes

$$\lambda(t) = \Phi_\mu(t)$$

for t = 2, we have

$$\Phi_\mu(2) = \int_0^1 e^{-\mu(2-s)} m_1 ds + \int_1^2 e^{-\mu(2-s)} m_2 ds$$

$$= \frac{m_1}{\mu}(1-e^{-\mu})e^{-\mu} + \frac{m_2}{\mu}(1-e^{-\mu})$$

$$= \frac{(1-e^{-\mu})}{\mu}(m_1 e^{-\mu} + m_2)$$

For integral $t \geq 1$

$$\lambda(t) = \Phi_\mu(t) = \frac{1-e^{-\mu}}{\mu} \sum_{i=1}^{t} m_i e^{-\mu(t-i)}$$

C) Service Process Unavailable for Some Initial Time Period:

C1. No service available on $[0, \tau]$, i.e., up to time τ; thereafter, have exponential service times with mean $\frac{1}{\mu}$.

C2. General intensity function $m(s)$ for $0 \leq s \leq t$.

In this case,

$$\lambda(t) = \begin{cases} \gamma(t) & \text{for } t \leq \tau \\ \gamma(\tau)e^{-\mu(t-\tau)} + \Phi_{\mu}(t) - e^{-\mu(t-\tau)} \cdot \Phi_{\mu}(\tau) & \text{for } t > \tau \end{cases}$$

D) Subsequent Loss of Repair Process:

D1. Exponential service time (rate μ) on $[0, \tau]$, i.e., before time τ, after which the system completely loses its service capability.

D2. Same as C2.

Here

$$\lambda(t) = \begin{cases} \Phi_{\mu}(t) & \text{for } t \leq \tau \\ \Phi_{\mu}(\tau) + \gamma(t) - \gamma(\tau) & \text{for } t > \tau \end{cases}$$

E) Service Times with a Fixed Time Component:

E1. Service times made up of a fixed component d plus an exponential random component with mean $\frac{1}{\mu}$.

E2. Same as C2.

The result is

$$\lambda(t) = \begin{cases} \gamma(t) & t \leq d \\ \Phi_{\mu}(t-d) + \gamma(t) - \gamma(t-d) & d < t \end{cases}$$

F) Service Times with Fixed Component Plus Loss of Service:

F1. Same as E1 except that after time τ the system permanently loses its service capability.

F2. Same as E2.

Get

$$\lambda(t) = \begin{cases} \gamma(t) & t \leq d \\ \Phi_{\mu}(t-d) + \gamma(t) - \gamma(t-d) & d < t \leq \tau \\ \Phi_{\mu}(\max(0, \tau-d)) + \gamma(t) - \gamma(\max(0, \tau-d)) & \tau < t;\ t > d \end{cases}$$

G) Discrete Change in Service Rate:

G1. Assume exponential service times with rate μ_1 on $[0, \bar{t}_1]$, i.e., prior to time $\bar{t}_1$, after which the rate permanently changes to μ_2.

G2. Same as C2.

We obtain

$$\lambda(t) = \begin{cases} \Phi_{\mu_1}(t) & t \leq \bar{t}_1 \\ e^{-\mu_2(t-\bar{t}_1)} \Phi_{\mu_1}(\bar{t}_1) + \Phi_{\mu_2}(t) - e^{-\mu_2(t-\bar{t}_1)} \cdot \Phi_{\mu_2}(\bar{t}_1) & \bar{t}_1 < t \end{cases}$$

H) Two Discrete Changes in Service Rate

H1. Assume exponential service times with rates μ_1, μ_2, μ_3, in the respective time intervals $[0, \bar{t}_1]$, $(\bar{t}_1, \bar{t}_2]$, $(\bar{t}_2, \infty]$ where $0 < \bar{t}_1 < \bar{t}_2$.

H2. Same as C2.

The result is

$$\lambda(t) = \begin{cases} \Phi_{\mu_1}(t) & t \leq \bar{t}_1 \\ e^{-\mu_2(t-\bar{t}_1)} \Phi_{\mu_1}(\bar{t}_1) + \Phi_{\mu_2}(t) - e^{-\mu_2(t-\bar{t}_1)} \cdot \Phi_{\mu_2}(\bar{t}_1) & \bar{t}_1 < t \leq \bar{t}_2 \end{cases}$$

$$\left\{ \begin{array}{l} e^{-\mu_2(\bar{t}_2-\bar{t}_1)} e^{-\mu_3(t-\bar{t}_2)} \Phi_{\mu_1}(\bar{t}_1) \\ + e^{-\mu_3(t-\bar{t}_2)} \left[\Phi_{\mu_2}(\bar{t}_2) - e^{-\mu_2(\bar{t}_2-\bar{t}_1)} \Phi_{\mu_2}(\bar{t}_1) \right] \\ + \Phi_{\mu_3}(t) - e^{-\mu_3(t-\bar{t}_2)} \Phi_{\mu_3}(\bar{t}_2) \end{array} \right\} \quad \bar{t}_2 < t$$

I) Two Discrete Changes in Service Rate and Service Times with Fixed Component

I1. Same as H1 but in addition, there is a delay of d units of times prior to an item starting its service.

I2. Same as C2.

$$\lambda(t) = \left\{ \begin{array}{ll} \gamma(t) & t \leq d \\ \Phi_{\mu_1}(t-d) + \gamma(t) - \gamma(t-d) & d < t \leq \bar{t}_1 \\ \left. \begin{array}{l} e^{-\mu_2(t-\bar{t}_1)} \Phi_{\mu_1}(\max(0, \bar{t}_1-d)) \\ + \Phi_{\mu_2}(t-d) - e^{-\mu_2(t-\bar{t}_1)} \Phi_{\mu_2}(\max(0, \bar{t}_1-d)) \\ + \gamma(t) - \gamma(t-d) \end{array} \right\} & \text{for } \bar{t}_1 < t \leq \bar{t}_2;\ t > d \\ \left. \begin{array}{l} e^{-\mu_2(\bar{t}_2-\bar{t}_1)} e^{-\mu_3(t-\bar{t}_2)} \Phi_{\mu_1}(\max(0, \bar{t}_1-d)) \\ + e^{-\mu_3(t-\bar{t}_2)} \left[\Phi_{\mu_2}(\max(0, \bar{t}_2-d)) - e^{-\mu_2(\bar{t}_2-\bar{t}_1)} \Phi_{\mu_2}(\max(0, \bar{t}_1-d)) \right] \\ + \Phi_{\mu_3}(t-d) - e^{-\mu_3(t-\bar{t}_2)} \Phi_{\mu_3}(\max(0, \bar{t}_2-d)) + \gamma(t) - \gamma(t-d) \end{array} \right\} & \text{for } \bar{t}_2 < t;\ t > d \end{array} \right.$$

As it becomes clear, further assumptions would imply more complicated expressions which may lead to numeric integration. However, the various service distributions presented allow a quite general sensitivity testing of an inventory system.

APPENDIX B

SOME UNRESOLVED PROBLEMS

The development of models and techniques for transitory behavior in the body of this Note required certain assumptions that are sometimes violated to the extent that the models are no longer good approximations of the situation modeled. The most important assumptions are:

1. Sufficient slack service capacity
2. Independence of the service and demand process
3. Poisson arrivals

The first assumption implies that each recoverable item demanding service immediately receives service with average service time based on the function F(s,t). In general, the service time is dependent on the number of servers, the scheduling of repair, and the number of components already placing demands on servers. With respect to aircraft recoverable components the number of servers may be limited by the availability of specialized test equipment or specialized personnel. Furthermore, the test equipment and personnel are generally used to service several or many different kinds of recoverable items. Queuing results for finite numbers of servers are available but have the drawback that they do not account for the cross training and sharing of servers for different types of repair. Furthermore, they do not lead to the Poisson result described in this Note and consequently do not allow the separation and combination of different repair processes or "pipeline."

The "steady state" stockage calculations in use by the Air Force today use an average resupply time as the average service time. This average resupply time includes waiting time in service queues and is a historical average of the length of time elapsed between when it is checked in for repair and checked out of repair. An infinite server model using this time as the average service time is then used for

recoverable item stockage. This approximation is reasonable as long as the system remains strictly in steady state or when there is enough slack capacity to maintain the same resupply time regardless of demands on the system. An advantage of this approach is the ability to decouple maintenance policies such as service scheduling, number of specialists, and amount of test equipment, from the stockage policies involving the amount of spare stock provided. At the same time, it is difficult to identify the end benefits of certain maintenance policies when such loose coupling to inventory calculations exists. (Clearly, there must be a tradeoff in amount of spare stock provided and the service capability provided.)

The second assumption is generally violated because the scheduling and priority for service are typically based on the current set of inventory shortages (or at least those shortages affecting system performance). Furthermore, when system demands for service are affected by the past history of demands (such as when so many aircraft are nonoperational that flying demands are reduced and there is a consequent reduction in demands for service) then the third assumption regarding Poisson demands is also violated. The latter assumption may also be violated when demands occur with a higher variance to mean ratio than expressed by the Poisson process. The violation of the latter two assumptions generally means that the models lead to conservative predictions of behavior in that they assume more demands than might occur and they assume longer service times than might be achieved given efficient scheduling and expediting of certain types of service. In stockage calculations, this will lead to a higher demand for spare stock and in performance measurement with a given level of stock the models will consequently estimate performance lower than might actually be achieved.

We have examined an alternative modeling approach which involves direct integration of the differential equations describing the dynamic queue while assuming only the Markov property that the next state is determine by only the transition probability from the current state. The advantage of this approach is that the dynamic behavior of queues as they respond to server limitation, service rate changes

as a function of both state (number in the queue) and time, and demand changes as a function of state and demand can be considered. This method is also appealing relative to Monte Carlo type simulations in that only one integration is required to obtain the complete time dependent probability distribution of number of items in a queue with this approach while the Monte Carlo simulation may require a large number of trials to obtain the distribution with a sufficient degree of confidence. Although the results of this approach appear promising there are certain drawbacks still to be overcome. First, all "pipeline" or different service processes must be included in the set of state equations because of the loss of the Poisson properties and, secondly, the states of all components sharing the same set of servers must be considered at one time. This can lead to a potentially large number of state equations to be solved because of geometric growth. We are currently examining methods of reducing the number of equations which must be considered at one time.

REFERENCES

1. Brooks, R., C. A. Gillen, and J. Y. Lu, "Alternative Measures of Supply Performance: Fills, Backorders, Operational Rate, and NORS," RM-6094-PR, The Rand Corporation, August 1969,

2. Crawford, G. B., "WRSK/BLSS Analysis, I. Generalization of Palm's Theorem," Operations Analysis Office, Hq. Pacific Air Forces, March 1977.

3. Feeney, G. F., and C. C. Sherbrooke, "The (s-1, s) Inventory Policy Under Compound Poisson Demand," Management Science, Vol. 12, No. 5, January 1976. Also published by The Rand Corporation, RM-4176-PR, March 1976.

4. Fox, B. L., and D. M. Landi, "Optimization Problems with One Constraint," RM-5791-PR, The Rand Corporation, October 1968.

5. Gaver, D. P., J. P. Lehorsky, and M. Perlas, "Service Systems with Transitory Demand," Chapter in Logistics, Vol. 1 of North Holland/TIMS Studies in Management Science, North Holland/ American Elsevier, New York and Amsterdam, 1975.

6. Gaver, D. P., "Dynamic Queuing Expressions," Unpublished Communication.

7. Kotkin, M., "An Heuristic for Multi-Echelon Multi-Indentured Inventory Problems," AD-Technical Report TR 79-1, US Army Inventory Research Office, December 1978.

8. Miller, B. L., "A Real Time Metric for the Distribution of Servicable Assets," RM-5697-PR, The Rand Corporation, October 1968.

9. Muckstadt, J. A., "A Model for a Multi-Item Multi-Echelon Inventory System." Management Science, Vol. 20, No. 4 December 1973, pages 472-481.

10. Palm, C., "Analysis of the Erlang Traffic Formula for Busy-Signal Arrangements," Ericsson Techniques, No. 5, 1938, pages 39-58.

11. Rice, D. B., "Defense Resource Management Study - Case Studies of Logistics Support Alternatives," US Government Printing Office, February 1979, pages 85-108.

12. Ross, S. M., "Applied Probability Models with Optimization Applications," Holden Day, San Francisco, 1970, page 29.

13. Ross, S. M., "Introduction to Probability Models," Academic Press, New York, 1972.

14. Sherbrooke, C. C., "METRIC - Multi Echelon Techniques for Recoverable Item Control," Operations Research, Vol. 16, 1968, pages 122-141.

15. Takacs, Y., "Introduction to the Theory of Queues," Oxford University Press, New York, 1962.