63-2-1

MEMORANDUM
RM-3358-PR
JANUARY 1963

PROCEEDINGS OF RAND'S DEMAND PREDICTION CONFERENCE JANUARY 25-26, 1962

Edited by Max Astrachan and Albert S. Cahn

PREPARED FOR:
UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM
RM-3358-PR
JANUARY 1963

PROCEEDINGS OF RAND'S DEMAND PREDICTION CONFERENCE JANUARY 25-26, 1962

Edited by Max Astrachan and Albert S. Cahn

This research is sponsored by the United States Air Force under Project RAND — Contract No. AF 49(638)-700 — monitored by the Directorate of Development Planning, Deputy Chief of Staff, Research and Technology, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

EDITORS' PREFACE

Spare parts demand prediction is a basic logistics problem, and one that has many unresolved facets. The problem's various aspects have been under study at RAND and elsewhere, in both military and industrial environments. But, until recently, researchers in this field had never come together to discuss mutual problems and exchange ideas. To facilitate this exchange, RAND sponsored a Demand Prediction Conference at Stanford University on January 25 and 26, 1962. The main purpose was to review the present state-of-the-art in demand prediction, and delineate possible fruitful areas of research.

Eighteen people, representing nine military and civilian organizations in all, attended the Conference. One participant from each group presented some results of their work, centering the discussion around the following topics and questions:

(1) Demand patterns -- what underlying statistical patterns seem to be present in demand histories?

(2) Prediction techniques -- what methods or models seem to yield the best predictions from historical data by any appropriate criteria?

(3) Measures of prediction accuracy -- what criteria should be used to compare techniques? What arguments should be advanced for and against the introduction of explicit models of the costs of over- or underestimates?

(4) Data availability -- to what extent is the prediction problem one of completeness and accuracy of data, as opposed, for example, to choice of technique and measure of accuracy?

(5) Time period -- to what extent, if any, are the predictions affected by the period of time used as the experience period or as the prediction period? Alternatively, do we predict better for those items which experience a large or a small number of demands.

(6) System-wide versus base-level demand prediction -- are any of the above affected by this distinction? If so, how?

A general discussion of the areas in which further research might be undertaken followed the presentations. To focus the discussion, the following questions were advanced to the group: what data are readily available for different classes of items? What data would we like to have for prediction purposes? Can the data processing be improved? Should demand patterns be investigated in greater detail applying, for example, the techniques used in analyzing time series such as trend, etc.? Can we find better program elements on which to base predictions? Which additional prediction techniques are likely to prove fruitful? What is the best way of evaluating the accuracy of predictions? How can we take account of dynamic factors such as changes in policy, item design, T. O., or effects of inspections, repair, replacement procedures, transportation and environmental usage, etc.? What theoretical studies should be undertaken? Have we reached the limits of prediction accuracy with which logistics systems must be content and with which they must live?

Although the Conference represented many different customer and commodity interests, there were similarities in the problems and findings. All of the participants felt that open discussion was helpful, and would do much to improve the quality of the conclusions and recommendations research groups make to their customers. There was general feeling that this Conference was valuable and that a similar one should be held say in a year or two, with perhaps more representation from industry.

The RAND Corporation would like to thank Stanford University for hosting the Conference. And our special thanks go to Dr. Harvey M. Wagner of Stanford for handling the Conference details.

SUMMARY

These Proceedings include all the papers presented at the Demand Prediction Conference, together with a general Conference discussion by Kenneth J. Arrow, and some comments on possible areas of future research by Emil Hamilton. This summary indicates briefly the contents of each paper.

In his presentation, Albert Cahn discusses the value of having better demand prediction methods in the following situation. When planning for the stockage of a particular line item, all one can ever do is predict that the expected demand will lie between certain limits. If it is possible to narrow these limits, then on the basis of certain assumptions, it is also possible to determine the order of magnitude of the potential savings that will result. He illustrates the procedure with an example.

Based on historical data for the components of the Falcon missile, and a sample of B-52 parts, Max Astrachan's paper reports on a long-term evaluation of seven different techniques for predicting spare parts demand. The object of the study was to determine which technique is preferable in varying sets of circumstances from both a performance and a cost point of view. The techniques examined are based on constant demand rates per weapon-system program element or service-life characteristics. One is the issue-rate technique widely used in the Air Force and elsewhere, another is the actuarial procedure currently used for aircraft engines and other items.

Among the conclusions that came out of the study were the following. No one technique was consistently preferred for all parts and all time periods. No matter what statistical technique is used, there will still be prediction errors; some may be large. This suggests that the forecaster must use whatever knowledge and engineering information he has at his disposal in an attempt to reduce these errors. Although many parts seemed to have increasing demand rates over time, it was impossible to identify any entire category for which the service-life techniques were preferred. Finally, there was very little apparent improvement in accuracy of the techniques as the experience period increased, i.e., as the parts were exposed to more and more months of operation.*

Jules Silver's paper describes the use of the Gamma distribution to predict demands. His group undertook this study to explore the possibility of developing a relatively inexpensive technique, for parts having service-life characteristics, that would yield results at least comparable in accuracy to those obtained by the use of computationally more expensive procedures such as the actuarial method.

The procedure was initially tested by means of a Monte Carlo model, generating demands from a "uniform" time-to-failure distribution, assuming: (1) that 12 months of data were available on which to base forecasts, and (2) that 18 months of data were available. For comparison, predictions were also made using the issue-rate and actuarial methods.

---

*Max Astrachan, Bernice B. Brown, and J. W. Houghten, A Comparative Study of Prediction Techniques, The RAND Corporation, RM-2811-PR, December, 1961 (Limited Distribution).

In (1) the actuarial method did poorly; in (2) it did well. The issue-rate technique did poorly in both cases, as might be expected from the nature of the distribution generating the demands. The Gamma method was best in both cases. It was subsequently applied to 134 items, for which predictions had been made by the actuarial technique. Although the distributions appeared to vary quite widely, the results obtained by the two methods showed differences of the order of 15 per cent or less.

In his paper, Henry Solomon discusses a few of the major highlights of demand behavior and prediction studies done for the Office of Naval Research under the auspices of the George Washington University's Logistic Research Project. An early study concerned the determination of suitable operational variables. Lack of demand data for this and other studies they were doing led to the initiation of a large-scale usage data collection program representing usage by and on account of 65 ships. Although they learned a great deal about the problems of data collection and processing, results of use of these data in the search for operational variables were mostly negative.

Again, a study of mechanical and electrical parts usage by and on account of 12 submarines over a four-year period showed no significant relation between usage and the operational variables employed. This study also showed that the demand for items was extremely low and sporadic, that for any given ship the range of items demanded differed significantly from year to year, and that there was very little commonality of items demanded among the vessels. Most important, the study showed that approximately 75 per cent of the items in the population (i.e., installed and deemed wearable), were not demanded at all, by or on account of each ship during the data collection period. The implication of these

results for stockage policies was that greater emphasis should be given to range rather than depth of items. This led to a major study pertaining to the military essentiality of repair parts to help determine the range of stockage.

The use of a function of population data -- the number of opportunities for usage of an item -- yielded promising results. In particular, the square root of the population size in one of their studies turned out to be as good an estimator as any other procedure used at that time.

Recently, the Project has undertaken a research program pertaining to the logistics system for the Polaris Weapon System. In studying the nature of the distribution of demands for line items, it was found that in the great majority of cases, the negative binomial distribution could be used. This distribution has some interesting properties which are currently under closer examination at the Project. One other important point seems to be emerging from the data collection program. Based on the relatively small amount of data received thus far, while demand for line items is very low and sporadic, the bulk of the items used pertain to a very small percentage of the total components installed.

The final study Solomon describes concerns spare parts demand for Naval aircraft. An examination and analysis of the data yielded results similar to those obtained for Air Force planes. The author concludes his paper with some comments on the relationship between demand prediction and inventory models.

Robert Brown briefly summarizes the work done by Arthur D. Little, Inc., for the Bureau of Supplies and Accounts, Navy Department. He identifies and describes six major steps where systems design decisions

must be made in developing a forecast system: (1) data, (2) model, (3) smoothing techniques, (4) forecast, (5) error measurement, and (6) safety factor. All of the material discussed is to be included as part of his forthcoming book, Smoothing, Forecasting and Prediction of Discrete Time Series.

John Muth's study illustrates with two examples the interaction between forecasts and inventory control. The first deals with a steel firm faced with warehousing decisions concerning inventories in regional warehouses. The second concerns a manufacturer carrying a wide line of storage batteries for automobiles, trucks, boats, etc. In the latter case demands appear to be sensitive to weather conditions. In both cases it was necessary to predict the range of the forecast error as well as the forecast itself. Muth points out that analyzing the decision-making process in inventory control is important in order to (1) know what to forecast and (2) to understand the importance of forecast errors. It is meaningless in itself to know that forecast accuracy is within $\pm$ 10 per cent. The relevant information is the quality of the decisions based on the forecasting schemes.

Peter Winters reported on two exponentially weighted forecasting models with respect to theory and practice. One of the models contains a linear trend and ratio seasonals, with three constants, A, B, and C, which are to be determined from past data.

He defines and discusses three criteria of prediction accuracy. The first is a weighted sum of squared forecast errors; the second an average fractional error which is the ratio of the average absolute error for all forecasts made divided by the average "sales" (or whatever

is being predicted) over the forecast period; and the third is the standard deviation of the forecast errors together with the coefficient of variation.

Experience has shown that approximately the same A, B, and C weights are optimal for a wide range of products, so that it is reasonable to use the same weights for groupings of products. Furthermore, several studies have indicated that the coefficient of variation is constant over a wide range of products and sales levels. For one particular company it was possible to delineate two groups of products such that within each, the coefficient of variation could be assumed constant.

Demand prediction from a commercial viewpoint is discussed by Winston Dalleck. He describes three cases of forecasting which illustrate the nature of the problems encountered by McKinsey and Company, and how they attempted to deal with each one.

The first concerns a division of a large pharmaceutical company producing about 200 items, each of which has a high seasonal demand pattern. An exponential smoothing model developed and used by Winters and described in his paper proved effective.

The second example deals with a large wholesale and retail distributor of metal materials and metal products. They have an inventory of approximately 10,000 items. Replenishment leadtime and demand are highly variable. The problem involved providing a sound and systematic way to review and, if necessary, recalculate the item reorder points. Monte Carlo routines were used to develop a large number of distributions for the probability of usage during leadtime. The distributions could then be classified and applied directly in determining reorder points

for most of the inventory items.

The third case involved a company which produces several hundred items in a number of plants. Their distribution is nation-wide. The initial trial showed that the exponential smoothing technique provided a satisfactory approach to forecasting demands for at least the large volume items, even in the absence of much usable historical data.

George Fishman discusses briefly the theory behind the concept of the power spectrum -- a procedure for determining the relative importance of cyclical and seasonal phenomena in monthly time series. The procedure for estimating the spectrum comes under the heading of spectral analysis.

He illustrates the simplification which can be achieved in the theory of filtering (elimination of cyclical or seasonal components from the time series) by working in the frequency domain rather than in the time domain. He investigated the use of spectral analysis on hog and cattle slaughter time series of 50 years duration. He discovered by investigating other series that a minimum of 20 years of monthly data were needed for effective analysis.

Nowlan points out that when we are studying demand figures we are really studying the reliability of parts and components through the medium of demand data.

Emil Hamilton proposes some areas for further research in demand prediction. Among these are: the application of Selective Management principles, the sharpening of "qualitative/subjective" approaches for low-demand items, the development of a capability for associating

technical knowledge of the item with the relative merits of various prediction techniques, and a review of earlier studies such as flyaway kit concepts in an attempt to update and re-slant them toward new modern weapons and logistics applications. Finally, he points out that some effort needs to be expended toward helping using agencies achieve a better understanding and acceptance of the various studies and proposals of research groups.

And finally, the Conference Remarks by Kenneth Arrow pull together some of the discussion which took place during the presentation of the papers and the formal discussion period.

CONTENTS

THE VALUE OF DEMAND PREDICTION

Albert S. Cahn

The RAND Corporation

Almost since the inception of the RAND Logistics Department, there has been continuing work and study on the problem of demand prediction. This interest in the subject on RAND's part reflects a background interest by the Air Force that I am certain all branches of the Armed Forces, as well as commercial organizations, share.

This concern is quite proper. I believe that no discipline strikes at the heart of most logistics problems more vitally than the demand prediction processes. Logistics problems of the more interesting sort are usually concerned with choosing the proper course of action in the face of demand uncertainty. Any technique which reduces the uncertainty gives that much more hope of finding the correct logistic action.

Let us see if we can gain some idea of how valuable better methods of demand prediction might be to the Air Force. Quite obviously, we cannot arrive at an exact estimate of this value. But by making some assumptions, I believe the order of magnitude of potential savings can be calculated.

At the present time, the worldwide stores inventory of the U. S. Air Force consists of approximately 2,000,000 line items whose total value exceeds $16,000,000,000. It is impossible to state just how these items are categorized by price and annual demand, because no available data show this relationship accurately. However, an early

RAND Report* made such a classification for the worldwide inventory of certain aircraft property classes. **Figure 1 shows this price-demand matrix.**

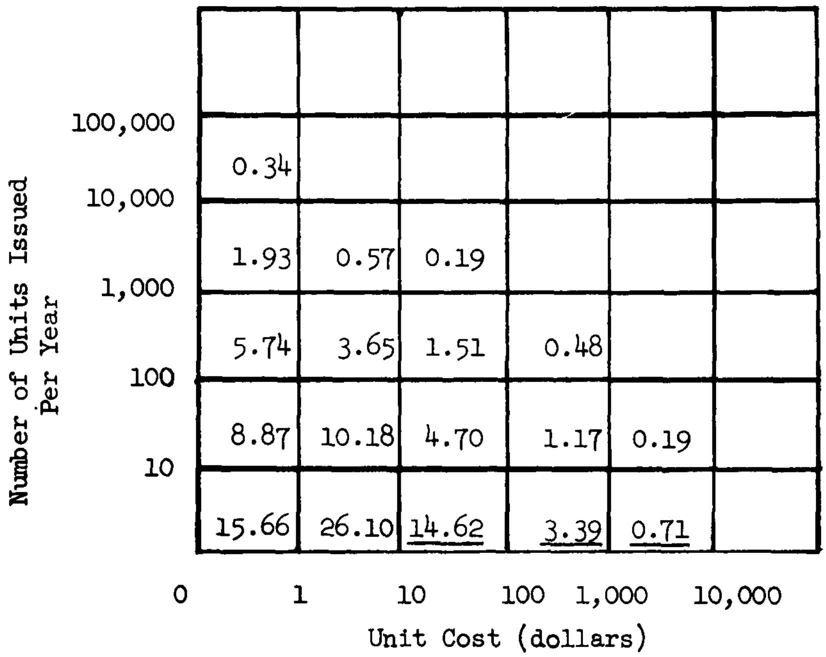

Fig. 1 -- Percentage of Total Line Items Classified by Number of Units Issued and Unit Cost

Let us assume there is the same distribution of line items in the present day inventory as that shown in Fig. 1. (Obviously not a true assumption, but perhaps not too unreasonable!) Now, let us focus attention on **those** items which have an issue rate of less than ten units per year, and a unit cost between $10.00 and $10,000.00. These

---

*Bernice B. Brown, Characteristics of Demand for Aircraft Spare Parts, The RAND Corporation, R-292 (AD 107426), July 1956.

are the items which normally cause much trouble -- the slow-moving, erratically demanded, expensive items. In Fig. 1 we have underlined the percentages which refer to the items we are discussing.

With this type of line item in mind, consider the effects of different stockage policies when used to fill the demands for these stochastically demanded items. We confine our attention to the situation where we want to decide how many of a line item to stock if we know its cost, the probability of demand, and the additional cost caused by shortage of a requested item.

For this purpose, let us assume the demand is Poisson. (Again, a poor assumption, but we believe it is not important for the argument we intend making!) With these assumptions in mind we write in Table 1, for the case of a spare part with an average demand of one unit per time period considered, the supply effects of different quantities stocked.

Table 1

DEMAND DISTRIBUTION OF SPARE PARTS WITH AN EXPECTED DEMAND OF 1 UNIT PER TIME PERIOD, AND EXPECTED SUPPLY RESULTS UNDER SIX STOCKAGE POLICIES

Number of Units Demanded	Probability of Demand	No. of Units Stocked per Line Item	Expected Supply Results (No. of Units per Line Item)		
			Surplus	Consumption	Shortage
0	0.3679	0	- 0 -	- 0 -	1.0000
1	0.3679	1	0.3679	0.6321	0.3679
2	0.1839	2	1.1036	0.8964	0.1036
3	0.0613	3	2.0233	0.9767	0.0233
4	0.0153	4	3.0043	0.9957	0.0043
5	0.0031	5	4.0006	0.9994	0.0006
6	0.0006				

If we adopt the point of view that surplus items have no salvage value, and if we can determine a shortage cost in units, then we can use Table 2 to calculate the un-utilized expense of the various stockage policies. This expense consists of the sum of the cost of the surplus items plus the shortage costs. This is set forth in Table 2 for cases whose item cost is equal to one-tenth the shortage cost, equal to the shortage cost, and equal to ten times the shortage cost.

Table 2

UN-UTILIZED EXPENSE OF VARIOUS STOCKAGE POLICIES
(In Terms of Units)

Number of Units Stocked	Unit Cost One-Tenth Shortage Cost	Unit Cost Equal to Shortage Cost	Unit Cost Ten Times Shortage Cost
0	10.0000	1.0000	0.1000
1	4.0469	0.7356	0.4046
2	2.1396	1.2072	1.1140
3	2.2563	2.0466	2.0256
4	3.0473	3.0086	3.0047
5	4.0066	4.0012	4.0007

From Table 2 we see quite obviously that if the unit cost is one-tenth the shortage cost, then the preferred policy is to stock two units; if the unit cost and shortage cost are equal, the best policy is to stock one unit; if the unit cost is ten times the shortage cost, then the best policy would be to not stock the item at all.

However, when planning for the stockage of any particular line item, one never knows just what the expected demand for an item will be. All one can ever do is predict that the expected demand will lie

between certain limits. The narrower the limits, the better the prediction will be. We attempt to evaluate this ability to predict within narrower limits in this paper.

If, on the basis of experience, the expected demand for an item lies between zero and two, then with no further information it is possible to assume that the probability of the expected demand is distributed uniformly over that interval. Or perhaps one can imagine the situation where we have a large number of items with expected demand distributed uniformly over that interval so that the average expected demand is one, but we are unable to forecast for any particular item just what the expected demand is. We mention parenthetically that assuming uniform distribution of expected demand, while helpful inasmuch as it simplifies the computations, is not necessary to the argument. One could make a similar calculation for any distribution.

In that case, the un-utilized cost of any stockage policy must be integrated over the required interval. For the case where the interval is from zero to two, Table 3 shows the average un-utilized expense.

Note that these values differ from Table 2 which is for an average expected demand of one. Also, the figures imply a different optimal stockage policy when the unit cost is one-tenth the shortage cost. In that case, the optimal policy is to stock three units rather than two, as implied by use of the average expected demand (Table 2).

Now, let us see if we can calculate the potential savings which would result if we were able to predict the expected demand for any line item within narrower limits than these. This is best shown

Table 3

AVERAGE UN-UTILIZED EXPENSE OF VARIOUS STOCKAGE POLICIES WHEN EXPECTED DEMAND VARIES UNIFORMLY BETWEEN ZERO AND TWO
(In Terms of Units)

Number of Units Stocked	Unit Cost One-Tenth Shortage Cost	Unit Cost Equal to Shortage Cost	Unit Cost Ten Times Shortage Cost
0	10.0000	1.0000	0.1000
1	4.7545	0.8645	0.4755
2	2.6952	1.3230	1.1776
3	2.5786	2.1052	2.0579
4	3.1653	3.0301	3.0165
5	4.0448	4.0075	4.0042

graphically in Fig. 2 where un-utilized expense (in terms of units) is plotted versus expected demand for the cases where: zero, one, two, three, four, and five units are stocked, and when the shortage cost is taken to be ten times the unit cost.* If one knows of a certain line item only that its expected demand is between zero and two, then the best policy is to stock three of the item (see Table 3) and have an un-utilized expense equal to the average ordinate of the curve for stocking three units, which, in this case is 2.5786 units.

On the other hand, if one possessed some means of predicting within narrower limits what the expected demand is, then one could do better than this. For example, again referring to Fig. 2, if one

*The points on the various stockage policy curves for an expected demand of 1.0 are the values listed in the second column of Table 2. For other expected demands, the points on the various curves were computed in a similar manner. The values presented in Table 4, which are proportional to the areas under the various curves, were obtained by numerical integration.

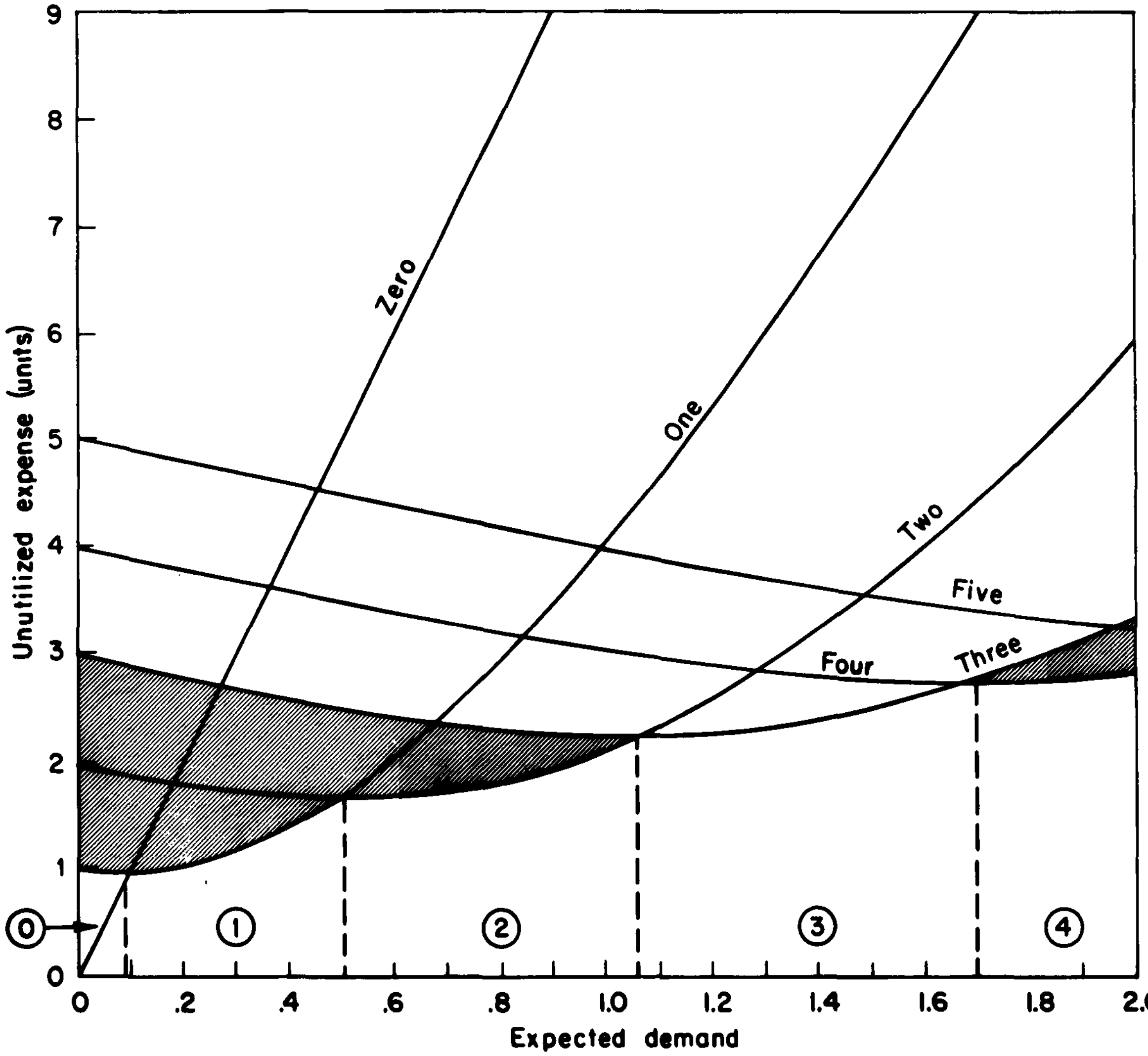

Fig. 2 — Unutilized expense versus expected demand (shortage cost ten times unit cost)

were able to predict correctly whether the expected demand for a given line item fell within one of the five demand intervals labeled from zero to four in that Figure, then, by stocking the quantity indicated, one would have an average expected cost (over the entire set of items) found from the lower envelope of Fig. 2.* In this case, the cost turns out to be 1.9878 units. The shaded portions of Fig. 2 represent the savings, 0.5908 units here, that result from this added knowledge. This amounts to a savings of 22.9% over the optimal stockage policy indicated when we do not have the ability to predict accurately within the narrower limits.

This potential saving in un-utilized expense that would result from an ability to predict expected demand, naturally varies with the ratio of unit cost to shortage cost. Table 4 shows this saving for four values of the ratio. The saving for a ratio of 10 was computed above. The other savings figures were obtained in a similar fashion.

Table 4

PERCENTAGE SAVINGS RESULTING FROM ABILITY TO PREDICT EXPECTED DEMANDS WITHIN NARROWER LIMITS FOR VARIOUS RATIOS OF SHORTAGE COST TO UNIT COST

Ratio of Shortage Cost To Unit Cost	Un-utilized Expense Without Prediction Ability	Un-utilized Expense With Prediction Ability	Savings as Percentage
0.1	0.1000	0.1000	0.00
1.0	0.8645	0.6963	19.46
10.0	2.5786	1.9878	22.91
100.0	4.3809	3.3198	24.22

*e.g., If the expected demand were less than 0.1, then no parts should be stocked; for expected demand between 0.1 and 0.5, one part should be stocked; for expected demand between 0.5 and 1.06, two parts should be stocked; etc.

In order to translate the potential savings exhibited in Table 5 into dollars, it is necessary to assume a value for the shortage cost. This is difficult to do accurately. For the purposes of this paper let us assume a shortage cost of $1,000. This is not an unreasonable amount when one reflects on the efforts used by the Air Force to obtain a needed part related to mission effectiveness. The assumption of a shortage cost of $1,000 allows us to relate Table 4 to the section of the Air Force inventory represented by the underlined percentages in Fig. 1, namely, the low demand items varying in cost from $10.00 to $10,000.00.

Figure 1 expresses the percentage of line items in the inventory which satisfy the price-demand characteristics indicated. Assuming that the yearly buy has the same composition as the inventory, then we can calculate from Fig. 1 the dollar percentage of the annual buy classified by number of units issued and unit cost. Fig. 3 shows this.

Number of Units Issued Per Year \ Unit Cost (dollars)	0–1	1–10	10–100	100–1,000	1,000–10,000
above 100,000					
10,000–100,000	4.30				
1,000–10,000	2.44	7.21			
100–1,000	0.73	4.61	19.09		
10–100	0.11	1.29	5.94	14.79	24.02
0–10	0.02	0.33	<u>1.85</u>	<u>4.29</u>	<u>8.98</u>

Fig. 3 -- Percentage of Cost of Annual Buy Classified by Number of Units Issued and Unit Cost

We are now in a position to estimate the dollar value of the potential savings by multiplying the percentage of the cost of the annual buy of a particular price-demand class by the average potential saving percentage of that class. Table 5 summarizes this as follows.

Table 5

AN ESTIMATE OF POTENTIAL SAVINGS

Unit Cost Class	Percentage of Cost of Annual Buy	Percentage Average Potential Saving	Product
From $1,000 to $10,000	8.98%	9.73%	0.874%
From $100 to $1,000	4.29%	21.19%	0.909%
From $10 to $100	1.85%	23.57%	0.436%
		Total	2.219%

In effect, our assumptions indicate that the potential savings resulting from narrower limits on the prediction of expected demands should amount to approximately $2\frac{1}{4}$ per cent of the annual buy cost.

Since the Air Force's annual buy of spare parts is in excess of two billion dollars, this means that the potential savings would be between forty and fifty million dollars per year. Annual savings of this amount are more than enough to justify intensive work in order to see whether or not we can find better ways to predict expected demands.

A COMPARATIVE STUDY OF PREDICTION TECHNIQUES

Max Astrachan

The RAND Corporation

Almost from its inception, the Air Force has had the problem of predicting demands for spare parts. Reasonably accurate estimates of how many units of a particular line item will be required in a future time period are extremely important to their operations. The Air Force needs such estimates for effective and economical procurement, distribution, maintenance, and program-planning decisions. Numerous studies of the Air Force supply system point out that improved demand predictions could result in substantial savings.

Various aspects of demand prediction have been under study almost continuously at RAND and elsewhere. Some RAND studies have primarily attempted to describe characteristics of spare parts demands. Others have developed methods for predicting demands or for evaluating prediction techniques. See Astrachan, Brown, and Houghten* for a detailed description of the first research performed at RAND using real data to evaluate different prediction techniques. The present paper is a précis of that study's procedures and results.

Forecasting demands always involves some uncertainty. In using statistical techniques, such as those in this study, the forecaster applies the concepts of population and random sampling to develop procedures to deal with this uncertainty.

---

*Max Astrachan, Bernice B. Brown, and J. W. Houghten, A Comparative Study of Prediction Techniques, The RAND Corporation, RM-2811-PR, December, 1961 (Limited Distribution).

It is possible to think of the demands for a line item during some future time period as a random drawing from a theoretical (unknown) population, or distribution of possible demands for that time period. If the forecaster can learn something about the properties of this distribution, he may be able to reduce his total uncertainty about the future. For example, he may learn something about the mean of the distribution, or about the range of the numbers in it. This information will enable him to make better predictions of the actual demands, or determine limits within which they may lie. As he learns more about this future population, his uncertainty will be less, and his predictions better.

One way the forecaster learns is by analyzing historical data to acquire some knowledge of past populations of demands. He then assumes a relationship between future and past populations. For example, he may be able to determine a demand rate from collected data, then assume that the future population will have the same demand rate as the past. Another way is by his practical knowledge and experience, i.e., his knowledge about the parts in question or similar ones, and his experience in making past predictions.

The present study concerns ways of using data from unfolding demand experience to make predictions with the help of statistical procedures. It shows how the data can be used to obtain information about the populations from which future demands will be drawn. This, together with some assumptions and a known activity program, will enable the forecaster to predict demands. The study examines seven different prediction techniques; then compares the results.

THE SEVEN TECHNIQUES

Table 1 lists the seven prediction techniques which this study discusses in detail later on. They are based on two essentially different assumptions about the relationship between past and future populations of demands for the same item.

Table 1

THE SEVEN TECHNIQUES

I. Issue-rate
II. Upper Bound -- Poisson
III. Non-Parametric Poisson
IV. Upper Bound -- Normal

V. Service Life -- Normal
VI. Service Life -- Log Normal
VII. Actuarial

The first four techniques assume a constant demand rate over time, so that the relationship between past and future demand populations is a function only of the total program activity in each period. Table 2 summarizes the important features of these four techniques.

Table 2

FEATURES OF CONSTANT DEMAND RATE TECHNIQUES

- Relationship between populations affected only by program activity
- Requires only gross data:
  - Total number of demands
  - Total number of program elements
- Inexpensive to use

They require data about the total number of demands that have occurred in the past (i.e., during the experience period), and the total program activity in the past and future time periods. As a result, they are relatively economical to use, a fact of importance to data collectors and processors. The issue-rate technique is now being used in the Air

Force. The other three are upper-bound techniques designed to give some assurance that under-prediction will not occur often if the constant demand rate assumption is valid.

Table 3 shows the features of the three service-life techniques (V, VI and VII). They assume that the relationship between past and future demand populations is affected both by the total program activity in the various periods and by the age of the units contributing to that activity. These techniques might be expected to predict better

Table 3

FEATURES OF SERVICE-LIFE TECHNIQUES

- Relationship between populations is affected by program activity and age of installed units
- More accurate for items that wear out because of age
- Require detailed data about:
  - Number of demands
  - Age of failing units
  - Age of installed units
- Expensive to use

for items that tend to wear out because of age. The trouble is, however, that they require data not only on the number of past demands, but also on the age of both the failing units and the surviving installed units (age is measured by the activity element). Because of the data requirements and the extensive computations that these techniques involve, it would cost much more to use them than many others based on different assumptions. A rough estimate showed that the computation costs only of the service-life techniques used in this study were approximately sixty times that of the non-service-life techniques. The cost of making one of many service-life predictions is about forty cents, compared to less than one cent for the non-service-life techniques.

Although each technique is described verbally below, it is also convenient to have formulas for some of them. For this purpose we use the following notation:

d = actual number of demands in a given month;

D = number of demands in the experience period $= \sum d$, where the sum is taken over the months in the experience period;

$\hat{d}$ = predicted number of demands for a given month;

F = activity of the weapon system during the experience period; number of flying-hours for aircraft and number of checkouts for missiles; and

f = activity of the weapon system during the month for which the prediction is made.

Technique I (Issue-Rate)

As previously mentioned, the Air Force now uses the issue-rate technique to predict demands. The total number of demands for a given part during an experience period is divided by the total activity of the weapon system during this period to give an average demand rate. Assuming demand for this part continues at the same rate in the future (constant-demand-rate assumption), we then obtain the forecast for a particular month by multiplying this average demand rate by the month's planned activity. In symbols:

$$\hat{d} = \frac{D}{F}(f)$$

Note we assume that every unit of activity gives the part the same exposure to failure, and that the past demand rate is taken as an estimate of the mean demand rate of the future population of demands. No assumptions are made about the distribution of demands in the future populations.

Techniques II, III, and IV are upper-bound techniques. They constitute a hedge against the possibility of an increasing demand rate for parts for which there is not enough information to determine the validity of the assumption of a constant or declining demand rate over time during the given experience period. Such parts may exhibit service-life characteristics over longer experience periods.

<u>Technique II (Upper-bound -- Poisson)</u>

This technique protects against underpredictions by using an upper bound on the mean of a Poisson distribution. We assume that the number of demands, D, in the experience period is drawn from a Poisson distribution of unknown mean and that the demand rate is constant over time. D is used to find the upper bound of the 90-per-cent confidence interval for this mean.* The upper bound is then used in the same manner as D in Technique I to obtain the prediction $\hat{d}$. We divide the upper bound by the total activity of the weapon system in the experience period, and then multiply by the planned activity for the month for which the prediction is being made.

<u>Technique III (Non-parametric Poisson)</u>

This technique, developed in an unpublished paper by D. S. Stoller of The RAND Corporation, can be thought of as a stockage policy. We assume that demands follow a Poisson distribution and that the demand

*Values of the upper bound for different values of D have been tabulated. See, for example, E. S. Pearson and H. O. Hartley, <u>Biometrika Tables for Statisticians</u>, Vol. I, The University Press, Cambridge, 1954, p. 203, Table 40.

rate remains constant over time, as in Technique II. Since it does not require an estimate of the parameter (mean) of the distribution Technique III is called a non-parametric Poisson technique.

The predicted value, $\hat{d}$, is obtained by solving the equation

$$(p\hat{d} - qD)^2 = pqk_\epsilon^2(\hat{d} + D)$$

where

$$p = \frac{F}{F + f}, \quad q = 1 - p,$$

and k_ϵ is the number of standard deviations in the standard normal distribution corresponding to a 100ϵ per cent tail. We used $k_\epsilon = 1.65$, which corresponds to a 5 per cent tail. Other values may be used, of course. The solution $\hat{d}$ of the above equation then has the property that the probability that actual demand in some future month will exceed the predicted value, $\hat{d}$, is less than 0.05, provided the assumptions are true. The predicted value can thus be regarded as an upper-bound decision on the amount to be supplied to meet demand.

Technique IV (Upper-Bound -- Normal)

This technique provides some protection against underpredictions by using an upper bound on the mean of a normal distribution. We assume that the number of demands D in the experience period is normally distributed about some unknown mean with variance (square of standard deviation) equal to three times the observed demand, and that the demand rate is constant over time. To insure a non-zero estimate of the upper bound when D is zero, one is added to the observed demand. The upper 90-per-cent confidence limit for the unknown mean is then

$(D + 1) + 1.65 \sqrt{3(D+1)}$. The forecast is this upper limit adjusted for the activity of the weapon system in the experience and prediction periods; i.e.:

$$\hat{d} = \frac{(D+1) + 1.65\sqrt{3(D+1)}}{F} (f).$$

Techniques I-IV involve only the total number of demands and the total activity of the weapon system during the experience period, together with the anticipated activity for the prediction month, in order to make the forecasts. They assume the demand rate is constant over time. Techniques V-VII are more sophisticated, requiring an estimate of the part's service-life characteristics derived from historical data. . It is customary, in this context, to think in terms of failures rather than demands; for our purpose the terms are interchangeable.

Techniques V and VI (Service Life -- Normal, and Service Life -- Log Normal)

Both of these techniques are adaptations of the Minimum Normit Chi-Square method developed by Berkson.* Each comprises two parts. First we estimate the service-life characteristics of the line item; second, we use these estimates to forecast demands in future months. The methodology assumes a normal distribution of ages at failure.

Non-normal data can often be normalized by using a logarithmic transformation. We do this in Technique VI using, however, the logarithm of one plus age instead of age. Except for this, the procedure is the same as for Technique V. The results are, of course, then transformed back into age.

*Joseph Berkson, "Tables for Use in Estimating the Normal Distribution Function by Normit Analysis," Biometrika, Vol. 44, 1947, pp. 411-435. Berkson describes the basic technique and discusses its efficiency under special conditions.

Technique VII (Actuarial)

This is another service-life technique. It is known as the actuarial method since it is comparable to the methods life insurance companies use to construct mortality tables. The Air Force currently uses it to predict engine failures.* As with Techniques V and VI, the procedure has two steps. The first involves determining the part's service-life characteristics that are set forth in a mortality table. The second step uses this table, the weapon-system inventory, and the activity program to make the actual forecasts.

THE DATA AND GENERAL PROCEDURE

In order to find out which of the seven techniques is likely to give the most accurate predictions, we are comparing their ability to predict the demands for a wide range of items. Table 4 gives some relevant data.

Table 4

THE DATA

- B-52 Irving Report
  - 33 months -- January 1956-September 1958
  - Data from two bases
  - Sample of 272 parts
- Missile (Falcon)
  - 26 months -- May 1955-June 1957
  - System-wide data
  - 27 components

We have 33 months of data on a sample of 272 B-52 Hi-Valu and Category II recoverable items, and 26 months of data on 27 different components

*For a detailed discussion of the methodology, see T. O. 00-25-128, "Procedures for Determining Aircraft Engine (Propulsion Unit) Failure Rates, Actuarial Engine Life, and Forecasting Monthly Engine changes by the Actuarial Method," October 20, 1959.

of the Falcon, an air-to-air missile. Using the same data every time, we employ each technique in turn to make demand predictions. The process of measuring the comparative accuracy of the results then goes through a regular series of steps presented in Table 5.

Table 5

STEPS IN MEASURING ACCURACY

- Predict future demands
  Each month predict demands for each of the subsequent 12 months
- Compute the average monthly error

$$\frac{\sum(\hat{d} - d)}{12}$$

- Compute the relative error

$$100\ \frac{\sum(\hat{d} - d)}{\sum d}$$

- Compute the root mean square error

$$\sqrt{\frac{\sum(\hat{d} - d)^2}{12}}$$

With the addition of each month's accumulated demand data, predictions of the demands are made for each item in each of the 12 subsequent months. A comparison of the predictions with the actual recorded demands follows, to see how large the errors would have been under real circumstances. Three statistics are computed from each of the 12-month prediction sets: an average monthly error (AME), a relative error (RE), and a root mean square error (RMS).

Since, as pointed out earlier, the data are constant, i.e., using the same data for each of the seven techniques, we are in a position to compare their accuracy in making predictions using the three measures we have introduced.

In comparing techniques, the average error allows us to compare how much bias each technique has in predicting monthly demands, and reflects the error in predicting the total demands for the year. The relative error allows us to compare the relative importance of those errors, and the root mean square error magnifies the importance of the largest monthly errors so that a comparison will reveal which technique is best able to predict when the demands will occur during the prediction year. The root mean square indicates, therefore, whether we can predict demands for phased-repair or procurement purposes. The average monthly error may show, for instance, that a technique has predicted demands accurately for the 12-month period. When this accuracy has been obtained by greatly over-predicting for some months and greatly under-predicting for others, however, the root mean square error will show that the individual monthly errors were large even though the average error was small. We would thus know that we cannot safely phase deliveries or schedule repairs rigidly on the basis of these monthly predictions.

Many different procedures can be followed in selecting the preferred technique for a given part. We proceeded as follows: In general, a part has very few demands during its early months on a weapon, and hence some non-statistical prediction technique would have to be used during this time. So we rather arbitrarily began with 12 months of experience for the B-52 parts and 8 months for the Falcon components. This left us with 10 prediction sets for the former and 7 for the latter. Since our primary objective was predicting demands for a year in the future rather than on a month-to-month basis, we

first examined the average monthly errors for the part based on a given experience period, beginning at these points in time. The technique which has the smallest AME is the preferred one; i.e., it gives the best predictions for a year based on the given experience. Next we examined the pattern of the AME's over time in accordance with the criteria discussed above, to select a single technique if possible which would have the smallest AME. For some parts, one technique was better than the others for one experience period and worse for the next. In order to avoid shifting back and forth among techniques we selected as preferred the one which had the smallest AME for the largest number of prediction sets, if there was such a one. When it was not possible to select a preferred technique on this basis, we examined the root mean square errors in the same manner as the average monthly errors, seeking the technique with the smallest month-to-month variation for the largest number of prediction sets. If this did not yield a preference, we examined all the AME's* for the part, the number of over- and underpredictions, homing, number of demands, etc., to make our selection. This introduced some subjectivity into the choice for those few parts for which this was necessary.

For some of the parts a service-life technique -- V, VI, or VII -- was preferred during the latter part of the program. One reason was that not enough demands had accumulated during the early months to enable us to use them. We then chose the best among the first four techniques for the early part of the program.

---

*Twenty-one for the B-52 parts and 14 for the Falcon components.

APPLICATION TO B-52 DATA

We turn now to the application and evaluation of the seven techniques as applied to the B-52 Data.

We chose for study Hi-Valu and Category II recoverable items only, in six major property classes. These six classes were Engine Components, Airframe Structural Components, Gunnery Components, Bombing Fire Control Components, Communications Equipment, and Aircraft Accessories. There were 875 parts in these six classes with demands of 5 or more during the 33-month data collection period. (All parts in the original listing which had fewer than 5 demands were eliminated.*) We chose a sample of 300 from these six classes, 50 in each. Deletions and combinations of parts reduced the sample number to 272.

About 40 per cent of the parts in the sample were Hi-Valu, while 60 per cent were Category II reparables; 144 of them had no demands in the first year of the program and 70 had none during the first 21 months. During the entire 33-month period there were no demands for about two-thirds of the part-months.

Table 6 shows the number of sample parts in each property class to which each technique was applied.

Table 6

Property Class	Techniques and Number of Parts		
	I-IV	I-VI	I-VII
Engine components	47	9	
Airframe structural components	45	35	1
Gunnery components	46	30	
Bombing fire control components	38	24	4
Communications equipment	48	16	5
Aircraft accessories	48	31	2
Total	272	145	12

*Five was selected arbitrarily for the purpose of this study. About 6500 parts had fewer than 5 demands during the entire data collection period.

The first four techniques were applied to all 272 parts. In order to use Techniques V and VI, we required complete data on parts which had at least four demands during the experience period. This limited their applicability to 145 parts. The Air Force Technical Order describing the actuarial method (Technique VII) suggests that data on about 100 failures should be available in order to use it. We required data on 10 failures during the first nine months. Even then we were able to use Technique VII on only 12 parts.

There were 165 different aircraft in the study. Beginning with the 12th month, there were between 85 and 100 planes on the two bases during any one month. About 68,000 flying-hours were accomplished during the 33 months. There were 2500 aircraft-months of data, with an average activity level of approximately 30 flying-hours per plane per month. The number of flying-hours increased slowly: only 55 per cent of the total number was accomplished by the end of the 21st month, at which time we made the last predictions for a year in the future. About 42 per cent of the demands for the 272 sample parts occurred during the first 21 months, and 58 per cent during the last year.

An examination of the cost of the parts in our sample is of some interest. There were 118 (43 per cent) which cost over $500. Of these, 46 (39 per cent) had fewer than 20 demands, and 69 (58 per cent) fewer than 35 demands, or less than about one per month. There were 12 parts which cost more than $10,000. One part in this group, an antenna assembly search radar, had 226 demands and cost $13,847. Only 11 out of the 272 parts had more demands. The most expensive part, a servo-control assembly, cost $36,815. There were 10 demands for it.

Following our routine, we made monthly predictions for a year in the future with each of the techniques based on one, two, three, etc., up through 21 months of experience, using the recorded number of flying-hours as the measure of activity for making the forecasts. There were 21 sets of predictions for each part using Techniques I-IV, substantially fewer for V and VI, and only 4 for Technique VII because of the large amount of calculation involved.

For each set of predictions, we computed the three measures of accuracy: the average, relative, and root mean square errors. These were then used to select the preferred techniques using the criteria described in the preceding section.

Parts which had no Demands in the First 21 Months

There were 70 parts in the sample that had no demands in the first 21 months. Thirty-five were engine components. Demands ranged from 1 to 113 during the last 12 months. The question arises: Can we make satisfactory predictions for a year in the future for such parts with our techniques? Each monthly prediction is simply a constant multiplied by the ratio of that month's flying activity to the activity in the experience period. For Technique I, zero demand experience gives zero demand prediction, regardless of the amount of activity. For Technique II, zero demand gives $3.00(f/F)$ as the predicted value; for Technique III, $2.72(f/F)$; and for Technique IV, $3.86(f/F)$. Techniques V, VI, and VII cannot be used for these parts, of course.

At the end of the 21st month, the flying-hour ratio (f/F) for predicting the ensuing year is about 0.82 for the 175 parts which were

applicable to all series of B-52's. The flying-hour ratio is different for parts with limited applicability but is less than one in all cases. Hence Technique IV, which gives the greatest upper bound, yields a predicted value of about 3 units.

We examined the average monthly errors for these 70 parts and selected preferred techniques in accordance with the criteria discussed above. Table 7 summarizes our results.

Table 7

NUMBER OF PARTS BY PREFERRED TECHNIQUE AND PROPERTY CLASS
(70 Parts with No Demands in First 21 Months)

Property Class	Technique I	II	III	IV	Total
Engine	--	--	1	34	35
Airframe	--	--	--	1	1
Gunnery	--	--	--	5	5
Fire Control	1	--	1	5	7
Communications	8	--	1	7	16
Accessories	2	--	--	4	6
Total	11	--	3	56	70

Our previous discussion anticipated the large number of parts for which Technique IV is preferred -- 56 out of 70, or 80 per cent. For the remaining 14 parts, this same technique is preferred for the last two or three prediction sets except for two of the communications parts. One of these had only one demand; it occurred in the 29th month. The other had only two demands; they occurred in the 33rd month.

To have zero demands for a part during a number of months while it is exposed to many thousands of flying-hours is information that should not be ignored. Such parts may be like the 6500 very-low-demand items that were not eligible for inclusion in our sample, or they may be like the 70 studied here, some of which ultimately had a large

number of demands. If they are like the former, Technique I is preferred, but if they are like the latter, Technique IV is better, in general. We do not know what the future demand pattern will be for any given part experiencing a number of months of zero demand. If we want to use Technique I, which always predicts zero for such parts, a prediction period shorter than one year -- any three or six months -- should be used, so that we can shift to another technique as soon as some demands occur, if this appears advisable.

Parts having Non-Zero Demand in First 21 Months

There were 202 parts in the sample which had some demand during the first 21 months. The preferred techniques were selected in accordance with the procedure outlined earlier. We did not permit shifts from one non-service-life technique to another, nor from one service-life technique to another, but only between the two groups. For example, even when it seemed best to shift from Technique IV to Technique I at a certain point in time, we did not allow such a shift, but selected that technique which had the smallest AME most of the time, beginning with the 12th month. There are some parts for which we would use a non-service-life technique until enough data had accumulated, and then shift to a service-life technique. Generally, the latter could not be used until about the 15th or 18th month. This implies that if a part has service-life characteristics, it takes a long time for its demand history to reflect this.

Summary of Preferred Techniques for Sample Parts

Table 8A summarizes the results for the 202 parts which had at least one demand in the first 21 months. Table 8B gives the results

for all 272 parts in our sample. It is obtained by adding the corresponding cell entries from Tables 7 and 8A.

Table 8

NUMBER OF PARTS BY PREFERRED TECHNIQUE AND PROPERTY CLASS

A. 202 Parts with at Least One Demand in First 21 Months

Property Class	Technique							Total
	I	II	III	IV	V	VI	VII	
Engine	8	--	4	--	--	--	--	12
Airframe	23	6	3	7	1	4	--	44
Gunnery	10	3	5	16	4	3	--	41
Fire Control	11	--	6	7	2	3	2	31
Communications	17	4	2	3	2	4	--	32
Accessories	16	4	5	10	5	2	--	42
Total	85	17	25	43	14	16	2	202

B. All 272 Parts in the Sample

Engine	8	--	5	34	--	--	--	47
Airframe	23	6	3	8	1	4	--	45
Gunnery	10	3	5	21	4	3	--	46
Fire Control	12	--	7	12	2	3	2	38
Communications	25	4	3	10	2	4	--	48
Accessories	18	4	5	14	5	2	--	48
Total	96	17	28	99	14	16	2	272

Our analysis of the 202 parts showed that for several of them it was preferable to use a non-service-life technique first and then shift to a service-lift technique. In such cases we allocated the part to the latter in Table 8A. The point in time at which the shift should be made varied from part to part; but in general, the service-lift technique was preferred as soon as enough demands had occurred for it to be used.

Table 8A shows us that no single technique is preferred for all the parts. Technique I, the issue-rate technique the Air Force now uses to predict demands, is best for 85 parts (42 per cent), more than for any other single technique. The next largest is Technique IV with 43 parts, or 21 per cent. Techniques I-IV account for 170 parts, 84 per cent of the total of 202. Taken together then, the constant-demand-rate techniques, which are the cheapest to use, are preferred for the vast majority of the parts in our sample of Hi-Valu and Category II recoverable parts having some demand during the first 21 months. Techniques V-VII, which assume that demand is related to age, were best for 32 parts. The first two of these were preferred for about the same number of parts. Actuarial Technique VII was best for only 2 parts. This suggests that the Air Force should examine very carefully any items for which service-life prediction techniques are being considered, in order to determine if the added improvement is worth the cost of collecting the necessary data and making the computations.

Table 8B gives the number of parts by preferred technique and property class for all 272 parts in the sample. Techniques I and IV are preferred for about the same number of parts, 96 (35 per cent) and 99 (36 per cent) respectively. The parts having no demands in the first 21 months of the data-collection period account for the large increase in the preferences for Technique IV.

We selected our sample of 272 parts from those in the original population which had 5 or more demands during the 33-month period. About 6500 part numbers had fewer than 5 demands and were thus not eligible for the sample.

Since, at the beginning of a program, the Air Force predictor does not always know whether there will be a large or small number of demands for a particular line item, the best he can do is use engineering and class knowledge for the predictions. The continual maintenance of an accuracy measure showing how well his predictions are meeting actual demand, however, might suggest points in time when he could change to a statistical technique. Furthermore, predicting, say, for three or six months instead of a year would make the selection of another technique more sensitive to the occurrence of demands.

APPLICATION TO MISSILE DATA

Our next point of discussion concerns the application of the procedures to the Falcon missile data. We had information on the recorded failures of 27 relatively expensive components of the missile over a 26-month period from May, 1955 through June, 1957. The data cover about 6,000 missiles.

The smallest number of recorded failures for any one component during the 26-month period was one and the largest, about 1,900. The total number of failures on all 27 components exceeded 8,700, with 92 per cent of them occurring during the last year, i.e., from months 15 through 26.

Fewer than 200 missiles were in the system during the first ten months. The number then built up rapidly to a peak of about 2,900 at the end of the 22nd month; from that point on, it dropped each month.

During the data-reporting period the missiles were subject to approximately 30,000 checkouts. The checkout program built up much

slower than the B-52 flying-hour program. Only 1,000 checkouts were performed during the first six months, about 2,400 by the end of the twelfth month, and 3,500 by the end of the fourteenth month, when the last predictions for a year in the future were made. Of the total checkout activity, 88 per cent occurred during the last year of the data-collection program. By contrast, 45 per cent of the B-52 flying-hour program took place during the last year.

A checkout was used as the program element in making predictions. It is more meaningful than operational time, which is usually small -- only about ten minutes.

As with the B-52 parts, we used data cumulated to the end of each month to make monthly predictions for the ensuing year for each component with each of the applicable techniques. We then computed the measures of accuracy and selected a preferred technique for each component. Only 10 of the 27 components had enough failures so that all seven techniques could be applied to them. The first six techniques were used on 7 components, and only the first four on the remaining 10. This gave us 14 sets of predictions for each component using Techniques I, II, III, and IV; from 2 to 14 sets using Techniques V and VI, depending on the number of failures; and four sets for Technique VII. This last is due to the fact that we made predictions with the actuarial technique at the end of the 9th, 12th, 13th, and 14th months only. There was too much computation involved to make predictions at the end of each month. Table 9 shows the number of components by preferred technique and general characteristics.

Table 9

NUMBER OF COMPONENTS BY PREFERRED TECHNIQUE
AND GENERAL CHARACTERISTICS

General Characteristics	Technique I	IV	V	VI	Total
Electronic	2	6	1	--	9
Electrical-mechanical	3	7	--	--	10
Mechanical with no moving parts	6	--	--	1	7
Unknown	--	1	--	--	1
Total	11	14	1	1	27

Based on this table, we can summarize our results for the Falcon missile as follows: Technique I is preferred for components which are chiefly mechanical with no moving parts; upper-bound Technique IV is preferred for electronic, electrical, or mechanical components. These results should not be understood to mean that upper-bound Technique IV offers a superior forecasting method for electronic parts as such. There are indications that the good showing of this method in the present study resulted from the accelerated phase-in of the Falcon missile during the period in question, which tended to turn a declining demand rate per checkout into an increasing demand rate per month.

We were able to make a more detailed analysis of the Falcon components than the B-52 parts because of their smaller number. We examined the constant-failure-rate assumption of Techniques I-IV and found it valid for only 10 of the 27 components. Technique I was preferred for 7 of these and IV for the remaining 3. Some of the

other components showed evidence of constant but different failure rates during different portions of the data-collection period. We found significant linear relationships between the monthly numbers of failures and checkouts for 22 of the 27 components. They were only of limited value, however, for prediction purposes because of the many unknowns involved.

FINDINGS AND CONCLUSIONS

(1) No one technique was preferred for all parts and time periods in our sample of B-52 parts, or for the Falcon components. Technique I was preferred for 96 parts in the B-52 sample, IV for 99 parts. Technique I was also preferred for the Falcon components that are mechanical with no moving parts. However, Technique I underestimates much of the time, especially in the early months of the program. If an overestimate is more desirable than an underestimate, Technique I would not be preferred for as many parts; we would then prefer to use one of the upper-bound techniques.

For those parts for which a service-life technique was preferred during the later portion of the time period, there is no uniformity as to which non-service-life technique should be used during the early portion of the program.

(2) A large number of parts will be exposed to many months of operation before any demands are observed. This fact of course does not guarantee zero demands in future months. Thus, our sample of B-52 parts included 70 which had no demands during the first 21 months. During the last 12 months, demands ranged from 1 to 113. Of the 27

Falcon components, 4 had no demands during the first 14 months. But they had demands of 1, 4, 4, and 60 during the last 12 months.

Technique IV was preferred for 56 of the 70 B-52 parts and for the Falcon component which had 60 demands in the last 12 months. Technique I was preferred for 11 of the B-52 parts and the remaining 3 Falcon components. Technique III was preferred for 3 B-52 parts.

(3) The issue-rate technique has most general applicability. Our B-52 sample was selected from those part numbers which had at least 5 demands during the 33 month data collection period. There were about 6500 part numbers with 4 or fewer demands. The demand rates for these parts would be low and there would be a number of months of zero demand experience. Issue rate Technique I always predicts zero when experience is zero, and would be preferred for such parts according to our criteria. As soon as some demands occur, however, it would underpredict and an upper-bound technique would be better. If we want to use Technique I for very low demand parts, predictions should be made for less than a year, say for three or six months, so that a shift to another technique can be made, if advisable, as soon as some demands occur.

It should also be pointed out that the issue-rate technique is the least expensive as well as the simplest technique to use. Any added improvement in the predictions made by one of the other techniques may not be worth the additional cost. This possibility would have to be examined for each part. For the very low demand items, Techniques II or IV, which are the most expensive, may overpredict. Nevertheless their added cost may be worth it in terms of avoiding shortages. Again, such decisions would have to be made on a part-by-part basis.

(4) Despite the fact that many parts seemed to have increasing demand rates, we were not able to identify any group for which the service-life techniques were preferred during the time periods for which we had data. For those parts where a service-life technique was preferred, the preference was V or VI. Technique VII, the actuarial method, was preferred for only 2 of the 12 B-52 parts and for 1 of 10 Falcon components to which it was applicable. This implies that the Air Force should examine very carefully those items being considered for the application of service-life techniques. They are expensive to use and the improvement in accuracy may not be worth the increased cost.

(5) There will always be prediction errors regardless of the technique used or the experience accumulated. In this study, even when the preferred technique was used for each part or group of parts, many errors were large.

(6) Regardless of which technique is used, a measure of its accuracy should be maintained. The average monthly error, the relative error, the root mean square error, or some other measure could be used, depending upon the desired objective. Continual examination of this measure would give a better indication of how well the technique has been predicting and suggest points in time at which changes to another technique might be desirable.

ON THE USE OF THE GAMMA FUNCTION IN DEMAND PREDICTION

Jules Silver

Air Force Logistics Command*

I would like to describe some work done by the Operations Analysis Division of the AF Logistics Command. It concerns the use of the Gamma distribution to predict demand. This work was undertaken to explore the possibility of developing a relatively inexpensive prediction technique for Hi-Valu items, which would yield results at least comparable in accuracy to those obtained by computationally expensive techniques such as the actuarial method.

In order to make sure that we have agreement on definitions, I shall define:

(1) Survivor Distribution (Percent Surviving Table)--the probability that an item of age zero will not have failed by age t.

(2) Time-to-Failure Distribution (Percent Failing Table)--the probability that an item will fail between age t and t + 1. NOTE: This distribution can be derived from the survivor distribution by differencing the successive probabilities.

(3) Failure Rate--the (conditional) probability that an item of age t will fail during the interval t to t + 1. This is derived by dividing the time to failure probability at t + 1 by the probability of surviving to time t.

(4) Removal Distribution--the probability of j removals in some interval of length T, when failed items are replaced with items of zero age.

---

*Wright-Patterson Air Force Base

(5) Expected Demand Curve--the expected number of removals occurring in a specified interval of time.

The Gamma distribution of order n is:

$$f(t) = \frac{\theta^{n+1} t^n e^{-\theta t}}{n!}$$

and represents the time-to-failure distribution. The Gamma distribution of order zero is equivalent to the exponential distribution. It can be shown that the removal distribution equivalent to the "Gamma-zero" is the Poisson and we shall call the removal distribution for the higher order Gammas "the Hyper-Poisson."

The issue interval technique for predicting demand is one in which future demands are estimated by multiplying the ratio of future program to past program by the issues during the past program. This technique yields results equivalent to that derived when the time-to-failure distribution is exponential (constant failure rate) or where the population is mature so that the age distribution is stable. However, the issue interval technique does not give "good" results where the failures are age related, and the population is not mature.

The actuarial method predicts removals by:

(a) computing the failure rates and applying them to the ages of items in the inventory or;

(b) deriving an average life from the failure rates and using it in a similar way to the issue interval. We will call this the "single factor" technique.

Gamma Method. This method takes advantage of a family of distributions which are quite versatile, and easy to handle mathematically.

The Gamma function is translated by means of the removal distribution into an expected demand curve and an iterative procedure is used to find the best fit to the actual demand curve (cumulative demand vs. cumulative program) by a least squares technique. In effect, this is a method of finding the parameters θ and n that give the best fit. As additional actual data is generated, the computer program adjusts the prediction.

A Monte Carlo model initially tested the technique. An item with 300 hours average life and 600 hours maximum life (mandatory removal) was synthesized. This item was used on aircraft in 12 squadrons of 25 aircraft each, with squadrons phasing-in during a period of 5 quarters. Flying hours for each aircraft were randomly generated with an average of 20 hours per month per aircraft but varying between 0 and 40.

The failures were generated from a "uniform" time-to-failure distribution (i.e., constant probability of failure from age 0 to mandatory removal). This gives failure rates which are low and uniform during the early ages but rise very rapidly during the older ages. The uniform time-to-failure distribution was selected because it appeared so unlike any Gamma distribution.

Two forecasts were made, each for 12 months in the future; one after 12 months of "actual" data and one after 18 months of "actual" data. Predictions were made, not only using Gamma, but also "issue interval," and actuarial (full table and single factor). Table 1 shows the results.

Table 1

FORECASTS USING 12 AND 18 MONTHS OF "ACTUAL" DATA

Item	12 Months Experience		18 Months Experience	
	Demand	Error	Demand	Error
Actual	176		242	
Actuarial (Full Table)	307	+ 74%	205	- 15%
Actuarial (Single Factor)	344	+ 95%	226	- 7%
Issue Interval	120	- 32%	144	- 40%
Gamma	211	+ 20%	253	+ 5%

With only 12 months of data, the actuarial methods made poor forecasts. After 18 months, the demand predictions were good. The issue interval technique was inadequate in both cases, since the population had not matured.

The Gamma method worked well in both cases, and this provided the impetus for additional work. Subsequently, a set of time-to-failure distributions was attained for 134 items covered by the actuarial method. The distributions appeared to vary quite widely. The Gamma method was used to predict demand, and a comparison made with the predictions obtained by use of the actuarial technique. The results showed the differences to be of the order of 15 per cent or less.

The exploration of the Gamma technique described above has not been completed for a variety of reasons. It should be pointed out that the computational requirements for the gamma technique are not inconsequential, although they appear to be less than those of the actuarial technique. In addition, there are many Hi-Valu items which have a very long service life. This, in effect, would result in

fitting a zero order function (which is equivalent to the issue interval) during much of the early life, since the technique would not be sensitive enough to detect the changing rates. On the other hand, very short service life items mature quickly and the issue interval approach would be satisfactory. However, it is felt that the technique is useful for Hi-Valu items with mid-range service life; and additional work in developing techniques is indicated.

A SUMMARY OF THE LOGISTICS RESEARCH PROJECT'S EXPERIENCE WITH PROBLEMS OF DEMAND PREDICTION

Henry Solomon

The George Washington University*

The Logistics Research Project has been concerned with problems of demand prediction for about 12 years. This is due to the significance of demand behavior and prediction for the many logistics problems which have been topics of research at the Project during this time. To describe in about 30 minutes, the kinds and results of work done on the subject of demand prediction requires that even some major aspects and results must be ignored. It is also necessary to ignore the particular context of logistics problems and model formulations for which the various aspects of demand prediction were studied. The following discussion consists of a few of the major highlights of studies of demand behavior and prediction.

It should be noted that the comments to follow will be restricted to problems of demand prediction for detailed line-items which are particularly significant for many topics which may be placed under the general label of military inventory problems. Also, while there are differences between the problem of predicting failure vs. predicting demand or usage, for the sake of brevity these will be ignored in the following discussion.

One early major formulation of the demand problem was mostly concerned with the determination of operational variables as program

*Work performed in connection with the University's Logistics Research Project, Contract Nonr 761(05), Project NR 047 001.

elements for usage. To quote Tompkins, ". . . the usage problem is about as follows: To determine a minimal number of independent operational variables upon which the usage of a commodity depends and to convert operational plans into the non-recurring quantities of commodities required and into the time rate of recurring demands."[1]*

This did not mean that problems are limited to the identification of these variables. Rather, once these were found, the problem was to account for:

(1) Deviations from predicted operations,

(2) Deviations in average usage rates for maintenance and overhaul, from station to station, and from time to time (i.e., variations in procedures by technicians, variations caused by aging, development of better materials, etc.), and

(3) Deviations of actual usage rates around expected usage rates during operations.[1,2-5]

During the years 1949-52, some attempts were made to study these questions as well as some others, mostly as a consequence of the formulation of inventory models.

Attempts to verify and, in fact, even in those early days, to simulate these models, met a significant barrier; namely, the availability of demand data. As a result of extreme paucity of information with which to investigate the many facets of demand behavior, a large scale usage data collection program was initiated by the Project.[6][7] Since this program generated what probably stands as the largest body of demand data in existence, a few remarks concerning the nature of this program are in order.

---

*See page 14 of Ref. 1.

The objective was to collect data representing usage by and on account of 65 ships. That is, to collect data on material used for end-use by each ship, and material used for end-use by supporting activities such as shipyards performing maintenance and overhaul on these ships, tenders, etc. This 65 ship sample included various classes and types of ships, e.g., carriers, destroyers, submarines, etc. The data pertained to usage of all items other than categories such as food and clothing. It included all mechanical and electrical parts, all electronics parts, all ordnance parts, and all general stores material.

While the number of ships included in the sample was reduced over time, the program covered data extending over a 6 year period beginning in June 1950. For a very small number of ships, an additional time period is available resulting from a later research program referred to as the "Allowance List Test Program."

A large amount of experience and insights were gained in the problems of data collection and data processing. However, these will not be discussed here.[8][9]* While this program was underway, of course considerable thought was given to exactly how these data should be used for analyses. Many such specifications were made prior to availability of the data for such analyses.[10][11] These included many of the topics mentioned on the agenda for this meeting, e.g., demand distribution studies of various kinds, types of activity analyses essentially involving the behavior of usage related to operational variables, time series analyses, etc.

*Various papers appearing in Ref. 9.

In 1954, some small scale investigations were attempted, mostly involving the search for meaningful operational variables. The results of these were mostly negative. That is, the operational variables employed (e.g., hours underway, engine miles steamed, war-time operations vs. peace time, etc.) could not be significantly related to usage. This was the first positive indication that the demand prediction problem would be anything but straightforward. However, again these were based on modest amounts of data.(12)

In 1954 a working conference sponsored by the Project and the Office of Naval Research was held on the subject of usage data.(9) From this, many proposals were offered in the areas of data collection, data reduction, and the uses and analyses of usage data for logistics problems. This is mentioned because many of the conclusions and proposals probably still apply.

The first major study of demand behavior by the Project, and in cooperation with the Bureau of Supplies and Accounts, was completed in 1957.(13) This study employed a large amount of data and had a significant influence on events to follow. The context of this study was the Allowance List Problem or that of determining shipboard stock levels. The data consisted of mechanical and electrical parts usage by and on account of each of 12 submarines over a four year period. The study also included the use of operational data plus a new and vital piece of information, namely, population data. The significance of these population data will be explained shortly.

The important conclusions from this study were:

(1) No significant relations were observed between usage and the operational variables employed.

(2) The demand for items was extremely low and sporadic. Over the entire four year period for each submarine and its supply activities, 70 per cent of the items demanded were demanded only once. Approximately 90 per cent of the items demanded were demanded at most twice, etc. There was a surprising amount of consistency among the 12 vessels as to frequency of events of demand.

(3) However, for any ship the range of items demanded differed significantly from year to year. That is, almost all items which were demanded in one year were not demanded in another year.

(4) While these 12 vessels included three sets of sister ships there was extremely little commonality of items demanded among the vessels. That is, the range of items demanded for each ship was highly unique. This indicated that broadening the base by including more ships may have little effect on the results.

(5) Finally, and most important, approximately 75 per cent of the items in the population, i.e., installed and deemed wearable, were not demanded at all, by or on account of each ship during the four year period.

This was the first piece of major evidence demonstrating the very high degree of uncertainty of demand. The major problem being the very large number of items deemed wearable but not used at all. Still another way of looking at the problem was to conclude, that for stocking policies, since the average quantity used of items which were demanded was also very low, greater emphasis should be given to range considerations than to depth.

While the results of the study were significant, they were also negative. Given the observed high degree of uncertainty of demand, in particular the highly sporadic nature of demand, how to handle or control this condition for stocking policies was considered to be the most important problem.

It was this condition which led to a major study pertaining to measurements of the military essentiality of repair parts.[14] Since

the range as well as the depth of items demanded was so uncertain, it was then decided to determine whether or not it is possible to at least select that range of items, which independent of any particular demand characteristic, may be considered essential to the vessel's mission. It is interesting to note that these observed demand characteristics led to this particular military essentiality study rather than any specific inventory model requiring measurements of essentiality. This study proved to be highly successful and demonstrated the feasibility of obtaining measures of relative military worth. It also showed that most repair parts were of relatively low essentiality. This military essentiality technique has recently been further improved for the Polaris program and is in process of implementation for the Polaris logistics system.

The demand study previously mentioned pertained only to mechanical and electrical parts usage for submarines. There was still some conjecture that while these patterns may be true for these types of parts, and for submarines, one would not observe the same behavior for electronics or ordnance parts in submarines, or for usage of other types of repair parts for surface ships. Additional empirical studies were conducted utilizing electronics and ordnance usage data for submarines and other types of material usage for other type ships.[15-18] The results of these studies confirmed the initial study for mechanical and electrical parts. That is, again, for all types of material, very low and sporadic usage was observed and the magnitudes were almost identical with the initial study. These studies again emphasized the very large number of zero movers and, in the case of those

items which were demanded, a very low amount of demands and quantities demanded. Again while these results were significant and included data of all types studied over very long periods of time, they were negative. The term negative implies here that none of the more or less straightforward techniques previously considered, would provide acceptable results.

Earlier, population data was mentioned. Since this proved to be a significant piece of information, a few remarks concerning these data are necessary. Population data refer to the number of times a part is installed in a component and the number of times that component is installed in the vessel. If the interest is in terms of system usage then, of course, the number of vessels which include the component should be accounted for. In short, these data provide the "number of opportunities for usage." These data are highly significant but surprisingly difficult to collect. With the availability of this information the number of items not used can be observed, a statistic too often unobtainable.

The second major aspect of population data is that one may wish to assume that usage of an item is some function of its population. In fact the explicit utilization of population data in this manner resulted in one of the most promising procedures.

During the "Allowance List Test Program," a function of the population was used as a usage estimator for some groups of items.[19] In this program some very simple approaches were taken. One was to consider the square root of the population as the usage estimator. This turned out to be as good an estimator as any other procedure

used at that time. Since usage is low and sporadic, employing this technique did tend towards overstocking, but for the simulations used during this program, this estimator was used only for the high worth items for which some amount of overstocking is preferable to understocking. Without going into details here, it should be stressed that the first positive results occurred by considering usage explicitly as some function of population for prediction of demand.

Also, with the Allowance List Test Program came some other important positive indications about the possibility of improving demand predictions. One of these was definite evidence that utilizing usage data collected from some past period, even when used in the most straightforward manner, provided better stock levels than no usage data at all and relying solely on technician's estimate of requirements. In simulating future time periods and comparing amounts actually stocked by technicians versus those stocked by simply extrapolating usage and comparing both with actual usage in these future periods, the result was that a great reduction in the range of items stocked was incurred by stocking based solely on usage without incurring anywhere near a proportional increase in the number of range shortages. Also, in regard to meeting quantities demanded of any particular item, the procedure of utilizing usage data provided less depth shortages than those stock levels stipulated by technicians. The procedures where solely usage data were used to predict future requirements were applied only to items with "low essentiality." This procedure tended toward understocking, however, the rationale employed was that these were low essentiality items and in these cases, some amount of understocking is preferred to overstocking. Another point worth noting is that by

including data from all sources the projections based solely on usage were superior to those for which only ship's own use was employed. That is, by broadening the historical base to include all activities using material on account of each ship as well as by the ship itself, the predictions for demand by the ship itself in the future were improved.

A final point worth noting from the Allowance List Test Program concerns the use of operational variables. It definitely appeared that the operational variables typically used are too gross. That is, to attempt, for example, to relate all mechanical and electrical parts to "hours underway" or "engine miles steamed," etc., does not account for the particular uses of the components in which the parts are installed. That is, for example, several components in the case of the submarine are used only when the ship is diving or surfacing. Hence, it might be expected that this kind of operation is what would affect the usage of parts in these components and not necessarily the number of hours underway, etc. All that can be said at the moment is that the search for operational variables cannot be said to be closed. Rather, some more detailed types of analyses must be performed in the future.

Recently the Project has been undertaking a research program pertaining to the logistics system for the Polaris Weapon System. Incident to these studies, the problem of demand behavior is of paramount issue. While the context of these problems cannot be described here, one important requirement was to determine the demand distribution which may be employed with a technician's estimate of average usage. Utilizing demand data collected over a long period of time, some statistical

investigations were made to determine the "goodness of fit" for the negative binomial and Poisson distributions. The result of these tests was that the line-item demand behavior over time could, in the great majority of cases, be described by the negative binomial distribution. Some caution must be exercised here by noting that the items used in these tests consisted only of items which were used or the population of demands and not the population of items which were installed and deemed wearable.

In a recent set of large scale simulations for Polaris Allowance Lists the negative binomial distribution was used with a particular set of variance to mean ratios which were a function of the size of the mean usage estimate.

The negative binomial distribution has some interesting properties which are currently being examined more closely at the Project. This pertains to the effects of different families of this distribution in terms of sets of variance to mean ratios for particular average usage estimates and the changes in stocking due to changes in protection level. These different families of the negative binomial distribution are also being studied in relation to other demand distributions. The orientation of these studies is in terms of particular types of inventory problems. A great deal of attention is being paid to the effects of the use of these distributions given such problems as dealing often with extremely small mean demands, extremely high protection levels, etc., and at the same time considering the resulting costs and gains in the particular inventory system.

It might also be mentioned in regard to the Polaris Program that the Navy is actively collecting usage data. At the present time this

is restricted to the collection of usage data from the Polaris Weapon System. Based on the usage data received thus far, which does not extend over a very long period of time, one again observes a very low and sporadic demand situation. The total number of demands per Polaris patrol are extremely small and the commonality of line-items used between patrols is very low. Also, the commonality of items demanded between different Polaris submarines is again very small. However, based on the relatively small amount of demand data received thus far there is one additional important observation. The usage data being collected includes "application data," i.e., identification of the component in which the part is replaced. Based on only a small amount of evidence, one does observe here that independent of particular or unique stock numbers, there is a fair amount of commonality in regard to the components for which items are used. This is interesting in that it begins to confirm a hypothesis which we have been seeking to test for some time. That is, while demand for line-items is very low and very sporadic, the bulk of the items which are used pertain to a very small per cent of the total number of components installed. If this is true then some headway can be made in determining operational variables by relating to the function of those components which are "item users." Also, if this is true this suggests, together with what has been learned in the area of military essentiality of repair parts, that items should be classified in terms of particular characteristics of the components in which they are installed. That is, possibly we should not look at masses of individual line items or even attempt to group these line items by dollar value, by commonality of nomenclature, etc., but rather as particular functions relating to component applications.

Recently a study has been completed concerning the demand for spare parts for naval aircraft.[20]

At the present time the USN Aviation Supply Office (ASO) uses two procedures for the determination of requirements for system stocking. Items under the inventory management of ASO are split into two large groups depending on which of the two procedures are to be applied to the item. One procedure, termed "Replenishment Demand Issue System" (RDIS) is to compute future requirements solely on the basis of past demands. For the second procedure, termed "Program Usage Replenishment System" (PURS) usage is considered to be a function of flying hours and future requirements are computed on the basis of an estimate of future flying hours.

In PURS it is assumed that usage varies in a linear and proportional manner with flying hours. The principal objective of the recent study was to examine the validity of this assumption. The data employed represented usage for approximately two hundred aircraft of a particular plane type operating over a three-and-a-half-year period.

The results of this study may be summarized as follows:

(1) Usage patterns do not conform to the PURS assumption concerning the relationship of usage to flying hours.

(2) Of a total number of 1774 maintenance usage parts observed, only about 60 items were correlated with flying hours. This resulted in some interest in exploring the relationship of usage to other possible program elements, such as numbers of aircraft, numbers of flights, numbers of landings, etc. Here again the results do not support the assumption of proportional linearity between usage and any of these program elements. It is also significant to note that of all the program elements studied, flying hours ranked as the least efficient estimator of aircraft parts demand. While no single program element showed a significant correlation with item usage, projections based on

number of "operating aircraft" provided better results than any other program element.

(3) Demand for aircraft repair parts is extremely low and sporadic. In fact, the demand patterns are about the same as those observed for ship's repair parts. This demand behavior alone indicates that usage would be insensitive to variations in flying hours. An additional similarity is an indication that usage may be closely identified with the population (i.e., number of installations) of the part.

(4) One final point of interest concerns the estimates of the activity level as represented by the program element of flying hours. It was found that on the average, the projections of flying hours for the future period were overestimated by approximately 40 per cent. Hence, even if usage were related to flying hours, the predicted demands would be in error by this amount.

A first important conclusion from this work is that most commonly proposed techniques for demand predictions have not produced results which would be acceptable for use in any straightforward manner in inventory problems. The principal underlying reason for this appears to be that, with few exceptions, the demand for repair parts in military systems is extremely low and sporadic. This high degree of uncertainty of demand for line-items must be recognized in the development of future inventory models.

Given the observed behavior of demand and the difficulty of demand prediction, it is necessary to formulate models for inventory systems which attempt to explicitly account for this behavior. That is, to some significant degree, earlier inventory models which typically are highly sensitive to the demand prediction problem, should be reformulated to minimize and/or control this condition. An example of this is the development of inventory models where the "military value" of repair parts is highly significant for the determination of stock levels.

Although the problem of demand prediction may be reduced by employing other variables in inventory models, this does not mean that it can be completely ignored. While many past investigations did not provide suitable results, there are some indications of promising results in the future. This includes the relationship of demands to operational variables not previously utilized; the dependence of demands on population of installed parts; the classification of parts in terms of their component applications; etc. Also, there is much to be done in investigating the effects of different demand distributions on the costs and benefits in terms of particular inventory problems.

REFERENCES

1. Tompkins, C. B., "Two Collections of Usage Data," Collection of Engineering Research Associates, Inc. Papers," 1948-1949. (Unpublished).

2. Friedman, Joyce, "Data Handling in the Aviation Supply System," Logistics Papers Issue No. 6, Appendix 2 to Quarterly Progress Report No. 7, The George Washington University, May 16--August 15, 1951.

3. Tompkins, C. B., "Usage Factors and Operations Parameters--an Example," Collection of Engineering Research Associates, Inc. Papers, 1948-1949. (Unpublished).

4. -----, "The General Concept of Usage Factors--Critical Notes," Collection of Engineering Research Associates, Inc. Papers, 1948-1949. (Unpublished).

5. -----, "Determination of a Safety Allowance," Collection of Engineering Research Associates, Inc. Papers, 1948-1949. (Unpublished).

6. Hamilton, J. E., and C. B. Tompkins, "A Review of Development in Logistics and an Introduction to a Broad Plan," Logistics Papers Issue No. 8 (Revised), Appendix 2 to Quarterly Progress Report No. 9, The George Washington University, November 16--February 15, 1952.

7. Hamilton, J. E., and R. E. McShane, The Korean Data Collection Program, Logistics Research Project, Serial T-81/57, The George Washington University, December 1957.

8. Hamilton, J. E., Experience in Data Collection: Summary, Logistics Research Project, Serial T-127/60, The George Washington University, May 15, 1961.

9. Proceedings of the Fifth Annual Logistics Research Conference, The George Washington University, October 25-30, 1954.

10. Arrow, K. J., "Proposals for Research on KDCP Data," Logistics Research Project, Serial 3710/55, The George Washington University, October 17, 1955. (Unpublished).

11. Solomon, Henry, "Planned Analysis of Usage Data at the Logistics Research Project," Logistics Research Project, Working Paper, Serial 3591/55, The George Washington University, October 1955. (Unpublished).

12. Fennell, J. P., Mina H. Gourary, and Henry Solomon, Some Preliminary Experiments with Korean Usage Data, Fifth Annual Logistics Research Conference, October 1954.

13. Denicoff, Marvin, J. P. Fennell and Henry Solomon, Requirements Determination, Logistics Research Project, Progress Report No. 1, Serial 59/57, The George Washington University, June 4, 1957.

14. -----, "Summary of a Method for Determining the Military Worth of Spare Parts," Naval Research Logistics Quarterly, Vol. 7, No. 3, September 1960.

15. Gourary, Mina Haskind, An Examination of Usage of Ships Parts for Two Destroyers, Logistics Research Project, Serial T-89/58, The George Washington University, June 1958.

16. -----, An Examination of Cruiser Usage Data: Part I, Logistics Research Project, Serial T-96/58, The George Washington University, December 1958.

17. -----, An Examination of Cruiser Usage Data, Part II, Logistics Research Project, Serial T-110/59, The George Washington University, September 1, 1959.

18. Solomon, Henry, Raymond Lewis, and Marvin Denicoff, Report on Electronics and Ordnance Usage Data for Nine Submarines, Logistics Research Project, Serial T-97/59, The George Washington University, January 15, 1959.

19. Solomon, Henry, and Marvin Denicoff, "Simulations of Alternative Allowance List Policies," Naval Research Logistics Quarterly, Vol. 7, No. 2, June 1960.

20. Denicoff, Marvin, and Sheldon Haber, "A Study of Usage and Program Relationships for Aviation Repair Parts." (To be published).

21. Tompkins, C. B., "Formulation of a Navy Cataloging Problem and the Assembly and Utilization of Usage Factors," Logistics Papers Issue No. 2, Appendix 1 to Bimonthly Progress Report No. 15, The George Washington University.

SUMMARY PRESENTATION ON DEMAND FORECASTING

Robert G. Brown

Arthur D. Little, Inc.*

As a matter of terminology we distinguish between two approaches to estimating future requirements. The word _prediction_ is used to refer to subjective estimates. _Forecast_ refers to objective computations on historical data which may include past demand or past predictions about demand. In this sense, our work has concentrated on forecasting rather than prediction.

In designing a forecast system we identify six major steps where systems design decisions must be made: (1) data; (2) model; (3) smoothing techniques; (4) forecast; (5) error measurement, and (6) safety factor.

Data. The data from which forecasts are made are most generally periodic summaries of past demand. The summaries may be made monthly or quarterly. The most serious problem we have encountered in our work for the Navy is that consumption in the Fleet is very much obscured by independent inventory management decisions at several echelons between the consumer and the ICP forecasting future demand.

In some cases it has proved useful to consider inter-arrival times rather than demand rates as the data. We have also considered optional installation rates of attachments on prime equipment.

---

*This paper presents a brief summary of the work we have done for the Bureau of Supplies and Accounts, Navy Department, under Contract No. Nonr-3406(00).

Model. There are three basic types of models that we have considered in describing past demand. One is the time series in which local segments of the pattern can be described by polynomials, transcendental functions or even empirical functions of time. The second is the renewal equation in which knowledge of the number of pieces of equipment installed and the distributions of time to failure can be used to evaluate the distribution of future demand. The third model is an empirical probability distribution for the level of demand in any period.

Smoothing Techniques. We have been primarily concerned with methods suitable for high-speed internally programmed digital computers processing a very large number of items each period. The criterion for accuracy in a smoothing technique is a discounted least-squares estimate of the parameters in the model. Since a great many items must be processed, the calculations have been reduced to simple linear arithmetic with files containing only one word of historical information per degree of freedom in the model.

No economic time series can be represented over a long period of time by a single model. Therefore, we have discounted past information. The discount factor can be changed at any time to alter the balance between stability in the fact of random fluctuations and rapid response to transient changes of pattern.

The concept of exponential smoothing* has been extended to the adaptive fitting of the coefficients in transcendental as well as poly-

---

*Robert G. Brown, Statistical Forecasting for Inventory Control, McGraw-Hill, 1959.

nomial models so that seasonal and cyclic effects can be adequately represented at very little effort.

Forecast. Generally a forecast is obtained as an evaluation of the model at future time. We have developed two modifications. The coefficient of higher order terms can be tested for statistical significance. If there are no significant differences from zero, they can be set equal to zero to minimize the various amplifications in the forecast. The second notion is a progressive discounting of higher order coefficients in future time.

Error Measurement. Any forecast or prediction must be compared with the actual data later when it is available. These errors have a distribution. The tracking signal is useful in deciding whether the mean of the distribution is approximately zero. The mean absolute deviation is simpler to compute than the standard device. We have developed some of the statistical properties of the mean absolute deviation. We have further shown that for a very wide class of distributions for the input data, decisions based on an assumption of normally distributed errors are satisfactory.

Safety Factor. The order point in an inventory control system can be expressed as the forecast $\pm$ k(MAD) where k is a safety factor generally between 0 and 3. The value of the safety factor is selected on the basis of current policy regarding routine service. We have developed formulas and curves for the safety factor under three philosophies: (1) for a specified chance of running out at the end of a replenishment cycle; (2) for a specified chance of running out as influenced by the frequency with which the item is replenished;

(3) for minimum total cost under a variety of assumptions about linear and quadratic cost of shortage and over-supply.

All of this material and more is to be published in book form, "Smoothing, Forecasting, and Prediction of Discrete Time Series". A complete draft of the manuscript will be available about 1 April 1962. Copies of the manuscript can be made available on loan to anyone who will agree to provide critical comments on this draft.

DEMAND PREDICTION AND INVENTORY CONTROL*

John F. Muth

Carnegie Institute of Technology**

It has been assumed -- perhaps too often -- that inventory control difficulties would magically disappear if better forecasts could only be found. I would like to suggest that the relation between forecasting and the uses to which forecasts are put is not quite so simple. It is important to analyze the decision-making process in inventory control, in order (1) to know what to forecast and (2) to understand the effects of forecast errors. Although this point is hardly a new one, I have generally found it overlooked in industrial operations research studies.

The interaction between forecasts and inventory control systems can best be clarified by means of illustrations. I will describe two cases in point. The first is a manufacturer of specialty steels; the second, a battery manufacturer.

A STEEL MANUFACTURER

Steel firms, like many others, are faced with warehousing decisions concerning inventories in regional warehouses. One company maintains stocks of some 4000 items in about 35 warehouses throughout

*This paper is based on a talk given by the author at the Conference.

**School of Industrial Administration.

the country, in addition to stocks in several plant locations. Demands for many of these items are extremely difficult to predict. For some individual products no demands might be experienced at all for several months. Then a very large order may come in. In fact the distribution of demands is so extreme that Tchebysheff's inequality, taken as an equality, appears to be a good approximation to the probability distribution of sales, at least in the range relevant for inventory control.

The problem of inventory control at this firm is made somewhat more difficult by the company policy that no demand for the product is to be left unsatisfied at a warehouse if the item can be shipped from some other warehouse or from a plant. From the standpoint of operations analysis, however, this restriction suggests at least one way in which the analysis might be simplified. It also allows us to replace the ambiguous notion of "depletion penalty" by the much better understood cost of shipping from one warehouse to another, the "transshipment cost."

The economics of truckload rates do not apply to any particular item stocked by the warehouse, but instead to all items. It is therefore economical to have regular replenishment shipments from the plant to each of the warehouses. The frequency of shipment would depend on the total sales of the warehouse. The only problem is determining the amount of each item to be included in the regular replenishment shipment.

A reasonable way of approaching the problem is to assign for each item in each warehouse, a "ceiling inventory" denoted by M. The amount ordered would be such that the inventory after receipt of the

order would be M units of the item. We identify two types of costs: (1) inventory storage charges and (2) the extra costs of trans-shipping an item from another warehouse when it is out of stock.

If R is the storage cost per pound within the replenishment cycle and μ is the average demand rate during the period, then the storage cost would on the average be

$$R(M - \mu/2) \tag{1}$$

Suppose the demands for the period D have a density function f(D) and that the trans-shipment premium is P. Then the average trans-shipment charge to be incurred with the item is represented by

$$P \int_{M}^{\infty} (D - M)\, f(D)\, dD. \tag{2}$$

Taking the sum of the two costs above and setting the derivative with respect to M equal to zero, we obtain the following conditions for a minimum cost:

$$\begin{aligned} 1 - F(M) &= R/P \qquad \text{if } 0 \leq R/P < 1 \\ M &= 0 \qquad\quad\;\; \text{if } 1 \leq R/P \end{aligned} \tag{3}$$

F(M) is the cumulative distribution function at M, so that 1 - F(M) in Equation (3) is the probability of a runout.

The other ingredients of the inventory computations are three forecasts: (1) the expected demand during the reorder cycle, (2) the standard deviation of the forecast error, and (3) expected cost of trans-shipment. The first two of these appear as parameters of the distribution function F(D). The last is necessary because the

cost of shipping varies from one cycle to the next depending on demands for the item at the other warehouses.

The resulting inventory control system is quite simple. Many "unwarranted" assumptions have been made in the course of the analysis. Nevertheless, simulation tests indicate that such a system would result in substantial cost savings. It would, in addition, allow better managerial control of the inventory positions at the various warehouses. By raising the parameter R, the cost of storage during the reorder cycle, management could lower rather effectively the inventory either in the entire system, or, more selectively, by warehouses or product groups.

A BATTERY MANUFACTURER

This firm has a forecasting problem which resembles that of military parts management in several ways. The firm is a manufacturer of a wide line of storage batteries for automobiles, trucks, boats, etc. An unusual feature of the forecasting problem is that demands of the firm appear to be sensitive to weather conditions. It appears that some "marginal" batteries in use may deteriorate if there is a sudden drop in temperature or sustained low temperatures. The months of December and January present special problems in forecasting demands. The problem is particularly severe because these months come immediately after the usual seasonal peak in sales. The inventory system may not be in a position to have sufficient inventories to avoid the risk of a lot of "panic" production if the so-called "weather surge" develops.

Two basic approaches to forecasting seem to be relevant. The first is relatively elaborate and would probably not be justified

unless byproducts could be obtained -- for example, product design or long-run forecasts of sales. Quite a bit of information is required in order to relate battery sales to weather characteristics. It is necessary to: (1) make estimates based on experimental evidence of the probability of battery failure as a function of its age and the weather; (2) find some means of determining, perhaps by survey sampling techniques, what the existing age distribution of batteries is; (3) take account of the age distribution of the stock of batteries in use, in order to predict failures of batteries for given weather conditions, using the difference equations of renewal theory; (4) try to find a good forecast of the weather, from which replacement demand could then be estimated. Enormous strides have been made since the day (1922) that L. F. Richardson proposed to conduct an orchestra of 64,000 trained computors, just to keep ahead of weather data from all over the world. It is still almost impossible to predict with any perceptible accuracy the particular weather characteristics required for forecasting battery demands (for example, the severity of the temperature drops through a cold front, the size of the "puddles" of polar air, etc.). The main conclusion about this approach seems to be that it may be useful in understanding the nature of replacement demands (we even have the best possible exogenous variable for an economic equation system -- the weather) but it has little merit as a forecasting device for inventory control. The main reason is that the data requirements are much too large and too remote.

There is an alternative procedure that is considerably simpler. It assumes that the most benefit can be obtained by keeping the forecasting procedure relatively simple, but taking into account the fact that the forecasts will not be as accurate during the months when weather surges are possible. We can realize this by including seasonal adjustments in the standard deviation of forecast errors used in the calculation of reorder points, lot sizes, and so forth. This approach requires little modification of the inventory calculations, except that a parameter changes from one month to the next.

Seasonal factors in the estimate of the standard deviation may be found by adapting the exponential forecasting schemes, which have already been discussed in this Conference by Robert G. Brown and Peter R. Winters. Let S_t represent the demands for an item in a period of time t and $S_{t-L,L}$ represent the forecast of the demand for the item during that time forecasted at time t-L. We start with some measure of the forecast error:

$$E_t = \left|S_{t-L,L} - S_t\right|^p \tag{4}$$

With p = 1, we are measuring the absolute deviations; with p = 2, we are using the squared deviations, as is frequently done in statistics. Whatever way we may define it, the variable E_t is the one to be estimated and revised each month. It will then be possible to derive the standard deviation from a forecast of E.

We assume for purposes of illustration that there are twelve seasonal factors associated with the months of the year. Let these be denoted by F_1, F_2, ... , F_{12}. Then the seasonally adjusted error

for the last month would, from Equation (4), be E_t/F_j where j is the appropriate index for the month. The predicted value of the seasonally adjusted error, denoted by $\bar{E}_t$, would then be given by an exponential smoothing formula:

$$(5) \qquad \bar{E}_t = A\,(E_t/F_j) + (1 - A)\,E_{t-1}$$

where $0 \leq A \leq 1$. Values of A close to unity correspond to substantial revisions of the seasonally adjusted error by means of the recent experience.

The seasonal factors themselves would then be revised. Let the new seasonal factor for the month be denoted by F'_j. Then since the ratio $E_t/\bar{E}_t$ represents the last seasonal factor observed, exponential smoothing of the seasonal factors would be given by an expression of the form

$$(6) \qquad F'_j = B(E_t/\bar{E}_t) + (1 - B)\,F_j$$

where $0 \leq B \leq 1$.

A prediction of the error T months from now, denoted by $\tilde{E}_{t,T}$, would be the product of the last seasonally adjusted error and the seasonal factor referring to the desired month. That is

$$(7) \qquad \tilde{E}_{t,T} = \bar{E}_t\,F_k,$$

where k is the appropriate monthly index.

The estimate of the standard deviation may be found from the forecast $\tilde{E}_{t,T}$. The nature of the conversion depends on the probability distribution of the forecast errors. In the special case of the normal distribution, the relation is given by

$$(8) \qquad \sigma = \begin{cases} \sqrt{\pi/2}\,\tilde{E}_{t,T} & \text{if } p = 1 \text{ (in Equation (4))} \\ \sqrt{\tilde{E}_{t,T}} & \text{if } p = 2 \end{cases}$$

An inventory control system utilizing seasonal factors in the error estimates in the determination of protective stocks is now in the process of being installed. Although the system is not fully developed yet, the prospects for better control and for cost savings appear good. An important thing to keep in mind -- and this has sometimes been overlooked during the discussions -- is that a great deal of accuracy in demand forecasts is not always essential for inventory control. Sometimes special forecasting problems, such as the weather surge, may be accommodated in other ways.

CONCLUSIONS

I hope the cases cited show how close the interaction is between demand prediction and development of inventory control systems. First, the cost analysis indicates the properties of the system that would in fact have to be predicted. In both cases it was necessary to predict the range of forecast error as well as the forecast itself. Costs of trans-shipment also had to be predicted for the steel manufacturer. Second, the theoretical analysis is also needed to show what kinds of forecast errors can be tolerated by a proposed inventory control system. It is meaningless in itself to know that forecast accuracy is within ± 10%. The relevant information is the quality of the decisions based on the forecasting schemes.

TWO EXPONENTIALLY WEIGHTED FORECASTING MODELS: THEORY AND PRACTICE*

Peter R. Winters

Carnegie Institute of Technology**

My paper concerns two distinct things. First I relate some of the experience we have had with the exponentially weighted moving average forecasting scheme that was reported on in Management Science in April, 1960.[1] The second thing relates to a series of models developed by R. J. Duffin of Carnegie Tech in work he has done for the Office of Ordnance, United States Army.[2][3] The second report will be published shortly and is, as of this time, available only in reproduced form from the Mathematics Department of Carnegie Tech.***

A SET OF EXPONENTIALLY WEIGHTED MOVING AVERAGE MODELS

Let me begin by reviewing several models that were reported on in Ref. 1. The simplest model assumes no trend and no seasonal.

$$(1) \qquad \tilde{S}_t = AS_t + (1 - A)\,\tilde{S}_{t-1}\,,$$

$$S_{t,T} = \tilde{S}_t\,.$$

*Research was undertaken for the project, Planning and Control of Industrial Operations, under contract with the Office of Naval Research. Contract N-onr-760-(01). Project NR 047011.

**School of Industrial Administration

***See Appendix, p. 82.

S_t is the sales in the $t^{\underline{th}}$ period. $\tilde{S}_t$ is the smoothed estimate of the mean of the sales distribution during the $t^{\underline{th}}$ period: $\tilde{S}_t$ is estimated at the end of the $t^{\underline{th}}$ period when the information of actual sales has become available; $\tilde{S}_{t-1}$ is defined in the same way as $\tilde{S}_t$. Finally, the forecast, $S_{t,T}$, is made at the end of period t for T periods into the future. $\tilde{S}_t$ can be shown to be an unbiased, but not efficient, estimate of the mean of the distribution of sales if, in fact, this distribution is unchanging over time. If this assumption is met, of course, a simple average would be a better estimate. This simple model is recommended for use when there are some systematic movements of the mean of the distribution but yet movements which one does not wish to capture explicitly in a model. The justification for using this model is that it seems to work all right. This first model is exactly equivalent to Brown's model(4) which he calls "single smoothing" and to Duffin's model(2) where the order of the polynomial is zero.

The second model is one which has a linear trend, but no seasonal effect. It consists of two equations plus the forecast.

(2)
$$\tilde{S}_t = AS_t + (1 - A)\,(\tilde{S}_{t-1} + R_{t-1}),$$
$$R_t = C(\tilde{S}_t - \tilde{S}_{t-1}) + (1 - C)\,R_{t-1},$$
$$S_{t,T} = \tilde{S}_t + TR_t .$$

In this model R_t is an estimate of the units per period that the mean is increasing or decreasing, made at the end of the $t^{\underline{th}}$ period. [All of the other quantities have been defined above.] This two-equation model seems to be equivalent to the model that Brown calls "double smoothing," and in it one is required to save two pieces of information,

the $\tilde{S}$ and the R, calculated last period. Notice that in all of the models the construction follows the same pattern. For each component of the model there are two estimates: one made previously and a current one. These two are weighted together to obtain a new current estimate.

The third model in this set is one which contains a linear trend and ratio seasonals.

(3)

$$\tilde{S}_t = \frac{AS_t}{F_{t-N}} + (1 - A)\,(\tilde{S}_{t-1} + R_{t-1}) \quad ,$$

$$F_t = B\,\frac{S_t}{\tilde{S}_t} + (1 - B)\,F_{t-N} \quad ,$$

$$R_t = C(\tilde{S}_t - \tilde{S}_{t-1}) + (1 - C)\,R_{t-1} \quad ,$$

$$S_{t,T} = [\tilde{S}_t + TR_t]\,F_{t-N+T} \quad .$$

N = No. of periods per cycle;

$0 \leq A, B, C \leq 1$.

This model is essentially the same as the one above except that current sales are deseasonalized by the seasonal factors F; and then these F's are used in making forecasts. Notice here that the amount of information which must be saved from period to period, for each forecasted series, goes up substantially. We must save $\tilde{S}$ and R as before, and now, in addition, twelve seasonal factors if, in fact, our cycle consists of twelve periods. The seasonal factors were included because a number of interesting time series appear to have seasonals in them. They seem to add substantially to the ability to forecast, particularly when one is forecasting some distance into the future.

GENERAL EXPONENTIAL FORECASTING PROGRAM (GEFP)

This program[5] is written in FORTRAN for the IBM 650; other FORTRAN versions exist for 700 series machines. It is an extension of the computer program illustrated in the Management Science article; it allows more flexibility and gives more information. The GEFP Program tries out the forecasting model on a time series, evaluating the accuracy of prediction by several different measures, and permits testing of the exponential weights.

Figure 1 illustrates this effect.

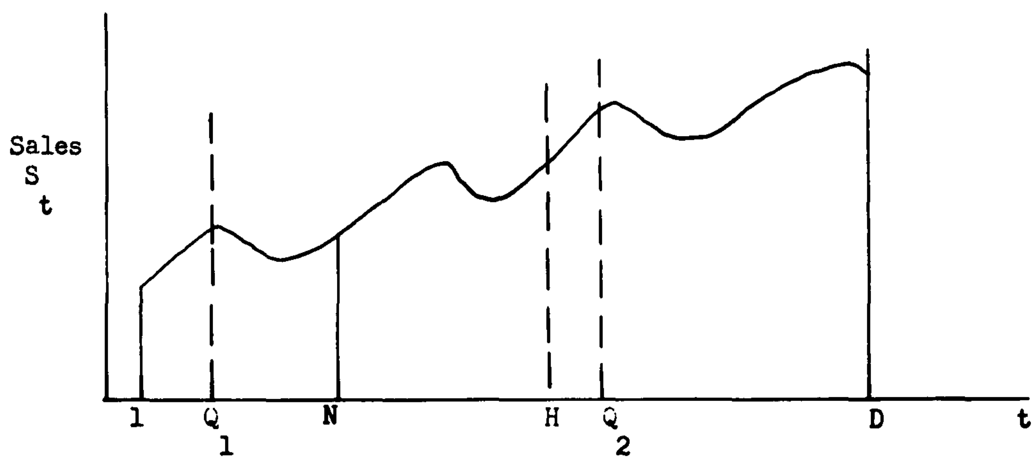

Fig. 1

The time series analyzed is assumed to be made up of D observations of sales (for "sales" the user can substitute the name of whatever series he is analyzing) in units per period, S_t, for t = 1, ..., D. The number of periods per cycle is N. For example, N would be 12 for the number of months per year. The program uses the early part of the data (t=1, ..., H) in a common sense way to determine initial values of $\bar{S}$, R, and the seasonal factors which are saved in

permanent locations. The exponential model uses these values and a set of weights (A, B, C) read in at Q_1 to start at Q_1, then simulates the use of the exponential model through Q_2 without forecasting. At Q_2 the model begins to make forecasts, continuing to do so until it reaches the end of the data. At each period forecasts for P periods into the future are made. These forecasts are compared with actual sales, the forecast errors are computed, and the sum of squared errors is collected for each forecasting period. For example, all the one-period-ahead forecast errors are collected, then the two-period-ahead forecast errors are collected, etc. These sums of squared errors are used in two places: (1) in the calculation of ϕ, one of the measures of accuracy of prediction, and (2) in the calculation of standard deviations of forecast errors and coefficients of variation.

The first criterion of prediction accuracy is ϕ, a weighted sum of squared forecast errors

$$\phi = \sum_{k=1}^{P} (\sum_{t} E^2)_k \, w_k; \qquad E = (\hat{S} - S)$$

The w's are weights that indicate the relative importance of forecasts for each k^{th} period in the future. If the program user is interested in simply a one-period-ahead forecast, then $P = 1$, $w_1 = 1$. If he wants to make forecasts for three periods each time he forecasts, then $P = 3$; a possible set of weights would be $w_1 = .5$, $w_2 = .3$, $w_3 = .1$ (the weights need not add up to 1). If he wants to forecast only the 4^{th} period into the future, then $P = 4$ and $w_1 = w_2 = w_3 = 0$, $w_4 = 1$. (Actually, all four forecasts will be made, but the first three will be ignored in computing ϕ.)

The second criterion is average fractional error. Although the formula is complicated, the idea is simple. This measure calculates the average absolute error for all forecasts made and divides it by the average sales over the forecast period. This allows the user to make statements as "The average error was 15% (or 25%) of sales."

$$\text{A.F.E.} = \frac{\sum_{t=Q_2}^{D-P} \sum_{k=1}^{P} \left| S_{t,k} - S_{t+k} \right|}{P(D - Q_2 - P + 1)\left[\sum_{t=Q_2+1}^{D} \frac{S_t}{D - Q_2}\right]}$$

The third criterion uses standard deviations of forecast errors and coefficients of variation. The GEFP also calculates the standard deviation σ_k, and the coefficient of variation CV_k for each k, that is, for the one-period-ahead forecasts, the two-period-ahead forecasts, and so on.

$$\sigma_k = \sqrt{\frac{(\sum_t E^2)_k}{D - P - Q_2 + 1}}$$

$$CV_k = \frac{\sigma_k}{\left[\sum_{t=Q_2+1}^{D} \frac{S_t}{D - Q_2}\right]}$$

A number of options are available to the user. For example, the program may be used either with or without trend, or with or without seasonal factors. Variable output information is also available.

RESULTS AND NOTES TO USERS

Evaluation of the forecasting ability of the third model, the one with linear trend and ratio seasonals, depends upon several factors: (1) the criterion function;* (2) the way in which the data is used to generate starting values; and (3) the (ABC) weights that are used. One reasonable absolute measure is the coefficient of variation (forecast error standard deviation divided by average sales). For a forecast of one-period-ahead** and for the best set of weights for a particular piece of data based upon our rather limited experience with the model, the coefficient of variation varies somewhere between .10 and .35. I would consider .10 rather successful. In general I would say that my aspirations now are much lower than they were when I began casting about for a forecasting method. It seems unlikely, if relative costs of computing and storage are about the same, that substantially better methods will be found.

Finding the A, B, C weights which minimize the criterion function, say ϕ, is a matter of trial and error. Several different schemes have been tried out on this problem. These include enumeration of the points of a three-dimensional grid, in the case of the third model; second a gradiant method; and third, another method of direct search.

---

*This forecasting method has been used extensively with an inventory control model(6)(7) in which the ability to forecast is measured by the variance of forecast errors, relative to average sales. The lower this variance, the lower the cost of the inventory system will be. For inventory models based on other characteristics of assumed sales distributions, other criteria may be more relevant.

**Best A, B, C weights for ϕ for one-period-ahead and three-periods-ahead forecasts are about the same.

One advantage of the grid method is that it gives some notion about the entire surface of the criterion function. This surface turns out to be non-convex over its entire range. However, in the neighborhood of the optimum, the surface is convex. The direct search technique which we hooked up with the forecasting program was the one reported by Hooke and Jeeves in the Journal of the ACM last year.(8) The surface in the neighborhood of the optimum is quite suitable for these methods.

There is one curious feature about the model in its ability to forecast a number of periods into the future. If one examines the standard deviation of forecast errors for one-period-ahead forecasts, for two-period-ahead forecasts, and so on, one finds results which typically look like Fig. 2. Counter to our expectations, σ_k actually went down (although not substantially) for the first few months out. This suggests an auto-correlation in the data or the model that we are not handling properly. At least it is intuitively appealing that the nearer the period being forecast, the more accurate the forecast should be.

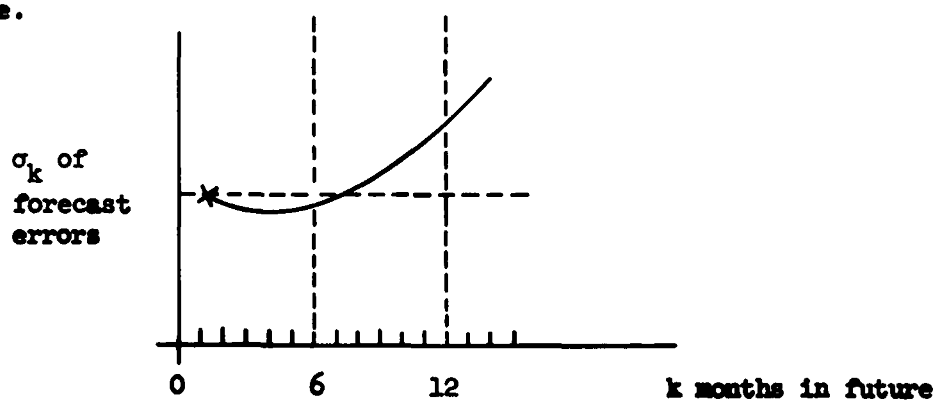

Fig. 2

SUGGESTIONS FOR APPLICATION

Because a full-fledged testing of the model and searching for optimum weights for a particular series requires a substantial amount of past history, and also a good bit of computing time, it seems reasonable to analyze only a sample of the products. Our experience shows that approximately the same A, B, C weights are optimal for a wide range of products; it is reasonable to use the same A, B, C weights for groupings of products at any rate. In addition, several studies indicate that the coefficient of variation is constant over a wide range of products and sales levels. In one application for the Wear-Ever division of the Aluminum Company of America we were able to describe two groups of products, such that within each the coefficient of variation could be assumed to be constant.

Some measure of control over the forecast seems advisable. One device that is suggested is an exponentially weighted estimate of the variance of forecast error which would require saving a couple of extra pieces of information for each product (σ^2 and the most recent forecast). If the current estimate of the standard deviation becomes either surprisingly high or low, compared with the initial estimate, then something can be done both about the forecasting procedure, and also about the forecast's use. For example, if the forecast is being used for inventory control, then buffer levels or trigger levels can be changed.

Chapter 15 of Ref. 7 gives an analysis of the distributions of forecast errors. The work given in this chapter concludes that distributions of sales are quite skewed; particularly for low-selling

items, and that these distributions may be reasonably approximated by the Gamma and the log-normal, with the Poisson being a reasonable approximation, but not as good as the first two. Of course the choice of approximation here depends upon the model in which the forecast will be used.

REFERENCES

1. Winters, P. R., "Forecasting Sales by Exponentially Weighted Moving Averages," Management Science, April, 1960.

2. Duffin, R. J. and Schmidt, Th. W., "An Extrapolator and Scrutator," Journal of Mathematical Analysis and Applications, Sept., 1960.

3. Duffin, R. J. and Whidden, P., "An Exponomial Extrapolator" Technical Report No. 46, Research Project: Mathematical Analysis of Electrical and Mechanical Systems, Mathematics Department, Carnegie Institute of Technology, 1960.

4. Brown, Robert, Statistical Forecasting for Inventory Control, McGraw-Hill, 1960.

5. Winters, P. R., "General Exponential Forecasting: A Computer Program for the IBM 650," O.N.R. Memo No. 71, Carnegie Institute of Technology, August, 1960.

6. Winters, P. R., "Constrained Inventory Rules for Production Smoothing," Management Science, July 1962, Vol. 8, No. 4, 470-481.

7. Holt, Modigliani, Muth, and Simon, Planning Production, Inventories, and Workforce, Prentice-Hall, 1960.

8. Hooke, R. and Jeeves, T. A., "Direct Search Solution of Numerical and Statistical Problems," Journal of the Association for Computing Machinery, April, 1961.

APPENDIX

Following are excerpts from the two Duffin papers:

EXTRAPOLATOR AND SCRUTATOR

The function to be minimized is a weighted sum of squared errors*

(1) $$\min \; E = \sum_{n=1}^{\infty} \theta^n \left[y_n - p(n) \right]^2$$

where n counts time into the past: $n=1$ is the current period, $n=2$ is last period, etc.; y_n is observed value in period n, and

$$p(n) = a_o + a_1 n + \dots + a_{m-1} n^{m-1} \quad .$$

The extrapolated forecast for next period is given in general by the "short formula":

(2) $$y_k^* = \sum_{j=1}^{m} (-1)^{j+1} \binom{m}{j} (y_{j+k} + \theta^j \delta_{j+k}) \quad .$$

For example, if m=4

$$y_o^* = 3(y_1 + \theta\delta_1) - 3(y_2 + \theta^2\delta_2) + (y_3 + \theta^3\delta_3)$$

with $\delta_k = y_k^* - y_k$, the error in prediction in period k. The short formula appears to require saving 2 pieces of information for each of (m-1) periods in the past. However, for m=3

$$y_o^* = 2(y_1 + \theta\delta_1) - (y_2 + \theta^2\delta_2)$$

---

*Same as proposed by D. A. D'Esopo in "A Note on Forecasting by the Exponential Smoothing Operator," Journal of the Operations Research Society of America, September-October, 1961.

which reduces to

$$y_o^* = 2(1 - \theta)y_1 + 2\theta\, y_1^* - (1 - \theta^2)y_2 - \theta^2 y_2^*$$

where it is possible to save y_o^* and $A = \left[(1 - \theta^2)y_1 + \theta^2 y_1^*\right]$ for use next period, or only 2 pieces of data for a second-order polynomial. I don't know if this generalizes for higher order polynomials, but since the functional E is precisely the same as Brown's, I would guess that it would.

EXPONOMIAL EXTRAPOLATOR

The data is fitted with generalized polynomials

$$p(x) = \sum_1^m a_j \beta_j^* \tag{3}$$

where the β_j are fixed complex numbers. The functional is similar to (1):

$$\min\ E = \sum_1^\infty \theta^n \left| y_n - p(n) \right|^2 .$$

Duffin develops a parallel short formula

$$y_k^* = - \sum_1^m g_k\, y_{k+x} - \bar{g}_m^{-1} \sum_1^m \bar{g}_{m-k}\, \theta^k \delta_{k+x}$$

with

$$g_1 = - \sum_1^m \beta_k^{-1} ,$$

$$g_2 = + \sum_1^m \sum_1^m \beta_j^{-1} \beta_k^{-1} , \; j < k ,$$

$$\vdots$$

$$g_m = (-1)^m \beta_1^{-1} \beta_2^{-1} \cdots \beta_m^{-1} .$$

Duffin then develops the variance of the estimate for several cases, and indicates some assumptions under which the extrapolation is a maximum likelihood estimate.

One example of the use of complex numbers gives

$$p(x) = a_o + \sum_1^M (a_k \cos \rho_k x + b_k \sin \rho_k x) \quad .$$

If M = 1

$$p(x) = a_o + \sum_1^M (a_k \cos \rho_k x + b_k \sin \rho_k x) \; .$$

The short formula is

$$y_o^* = A\,(y_1 + \theta\,\delta_1) - A(y_2 + \theta^2\,\delta_2) + (y_3 + \theta^3 \delta_3)$$

with $A = 1 + 2 \cos \rho_1$. The construction here parallels the short formula in Eq. 2. This work is essentially the same as the formulation that Brown has independently developed.

DEMAND PREDICTION: A COMMERCIAL VIEWPOINT

Winston C. Dalleck

McKinsey & Company, Inc.

Spare parts demand prediction as it is being discussed at this conference is a particular problem dealing with parts usage behavior in aircraft, mostly military aircraft. However, the practical aspects involved here -- the determination of demand patterns, development of forecasting techniques, and measurement of forecast errors -- are akin to many other problems involving item demand prediction. Thus it is not altogether inappropriate at this meeting to comment on some of the more general and commerical forecasting problems we encounter.

One statement of definition seems in order at the outset. In dealing with forecasting problems we recognize, as was pointed out earlier, that both a projection and prediction are usually involved -- the former being an examination and extrapolation of history and the latter a provision to account for new or additional effects such as promotional effort, price change, competition, a revision of the distribution setup, and the like.

Our forecasting experience has been commercial rather than military, and is of two types. The first, and the one encountered most frequently, concerns corporate sales forecasting. The second type involves item forecasting for such purposes as inventory management and production planning.

Corporate sales forecasting as we see it is a key element in total integrated planning and control. A sales forecast is the link

between planning and control. First it provides a quantitative statement of the revenue-generating potential of a corporate plan. After the fact it becomes the basis for evaluating performance in implementing these plans and achieving the sales objective.

This seems such an obvious requirement for doing business that one might wonder why sales planning -- sales forecasting if you please -- is a problem. The fact of the matter is that many companies do very little such sales planning. Many others do some forecasting but the result is ineffectual. It is not uncommon to find the following situations existing in a company: forecasting responsibility is not clearly designated; several forecasts are being made and not properly coordinated; sales forecasts do not adequately reflect operating plans; budgets are not consistent with forecast; effective coordination of promotional programs and new product development is lacking.

An example or two might be of interest here. One company planned a sales program and made a forecast which represented a substantial increase in sales over the previous year. They failed to meet the forecast and upon examination found that one of the large sales districts had cut back their sales force for other reasons -- to cut costs in compliance with another program -- and was unable to build back the sales force in sufficient time to make their quota, their share of the sales forecast.

Another instance involves a large distributing company with several satellite warehouses. It was the policy of this company that these field warehouses replenish their inventories directly from vendors and draw on the main warehouse inventory only in cases of

emergency. However, the monitoring of this procedure was not very effective. Since each warehouse, including the central warehouse, was treated as an independent profit center, the field warehouses found it to their advantage to replenish frequently from the central warehouse, thereby keeping their own inventories low, but at the same time creating unrealistic demands and high inventories at the central warehouse which were not optimal for the system.

The second type of problem encountered involves item analysis and forecasting. Our concern is with such activities as production planning and inventory management where it is necessary to set production rates, determine economic lot-sizes, set reorder points, and fix safety stock requirements. Here the problem is akin to that of spare parts. Item demand data must be collected and analyzed. It is necessary then to determine demand patterns, develop forecasting techniques, and devise ways of measuring and evaluating forecast errors.

This is not to say that the total corporate sales forecast and item forecasts are unrelated. We believe, in fact, that it is impractical to consider them separately as is frequently done. This fairly common difficulty can be characterized by an example involving a large producer and distributor of consumer products. This company periodically plans very detailed sales-call programs, indicating the specific products and type of customers which should get the attention of field sales people. In the case in question a significant shift was made in product emphasis from the previous program without this information being conveyed to the people responsible for making item

forecasts for production planning purposes. As you might expect the forecasts were incorrect and the result was an imbalance in production scheduling and subsequently in item inventories.

In contributing to the general discussion of item demand forecasting, I would like now to review three of our forecasting experiences to demonstrate the kind of problems which arise and how we have attempted to deal with them. These are three very brief case studies.

The first concerns a division of a large pharmaceutical Company X producing about 200 items, each of which has a high seasonal demand pattern and, in many cases, a significant trend reflecting increasing or decreasing demand in the market place. Item forecasts are made quarterly for production planning purposes, each forecast being monthly for eight months ahead. The method which had been used was 12 months moving average, the results of which were satisfactory.

The purpose in going to another, more quantitative and formal method of item forecasting was to provide a good base projection from which to evaluate the effect of subjective judgments being introduced into the final forecast. In addition, the company's EDP program was taking shape and it was timely to consider how data processing equipment could be tied into the forecasting process.

An exponential smoothing model was used and proved to be very effective. The model used is the one developed by Professor Winters of the Carnegie Institute of Technology.* In this model $\tilde{S}_t$, the estimate of the sales rate in period t, is computed as an initializing step,

---

*Peter R. Winters, "Forecasting Sales by Exponentially Weighted Moving Averages," Management Science, April, 1960, p. 324.

$$\tilde{S}_t = A \frac{S_t}{F_{t-L}} + (1-A)\ (\tilde{S}_{t-1} + R_{t-1})$$

where A is the smoothing constant and the weights B and C are used in separate calculations to revise respectively F (the seasonal factor) and R (the trend factor) until one is ready to begin making forecasts. Then

$$S_{t,T} = \left[\tilde{S}_t + TR_t\right] F_{t-L+T}\ , \text{ for } T = 1, 2, \ldots, L$$

is used to make these forecasts.

As you are aware, this model is designed to use historical sales data as a way of evaluating different sets of weights -- the ABCs -- in order to determine which set is best. Since a four year history of sales was available for each item we were able to initialize over the first two years and forecast the last two years. This step is portrayed in Fig. 1.

Item A

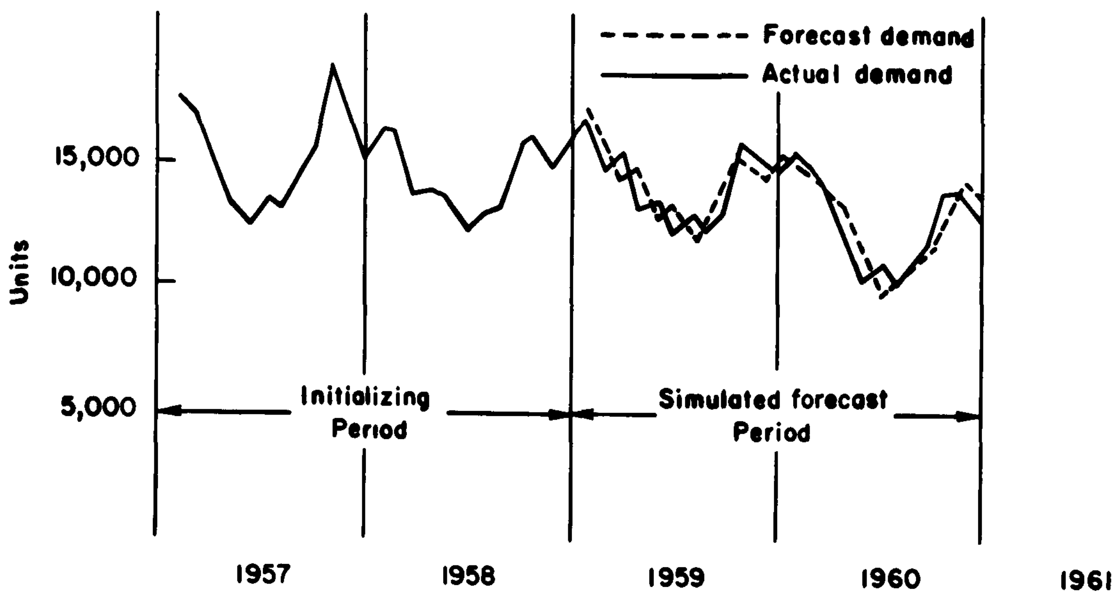

Fig. 1—Item demand history for testing forecasting method

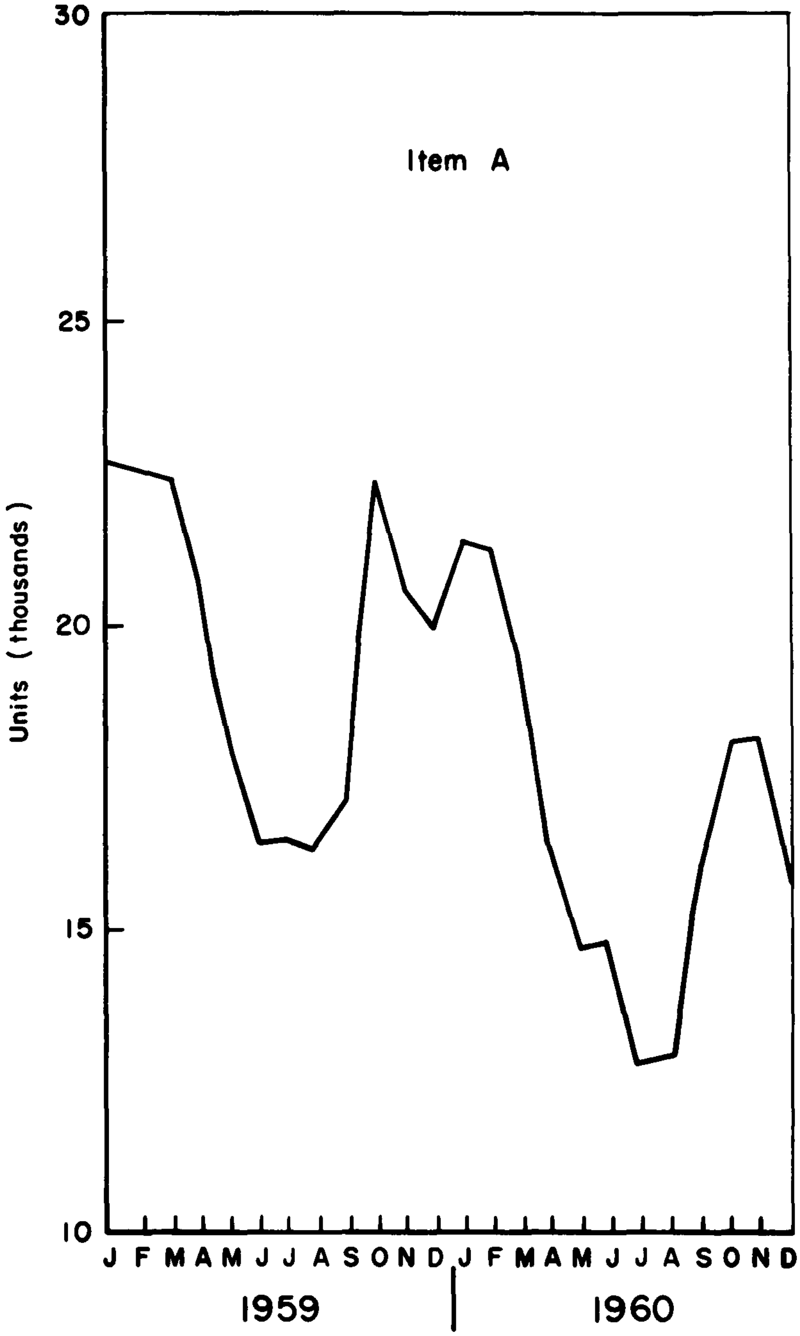

Fig. 2—Actual demand

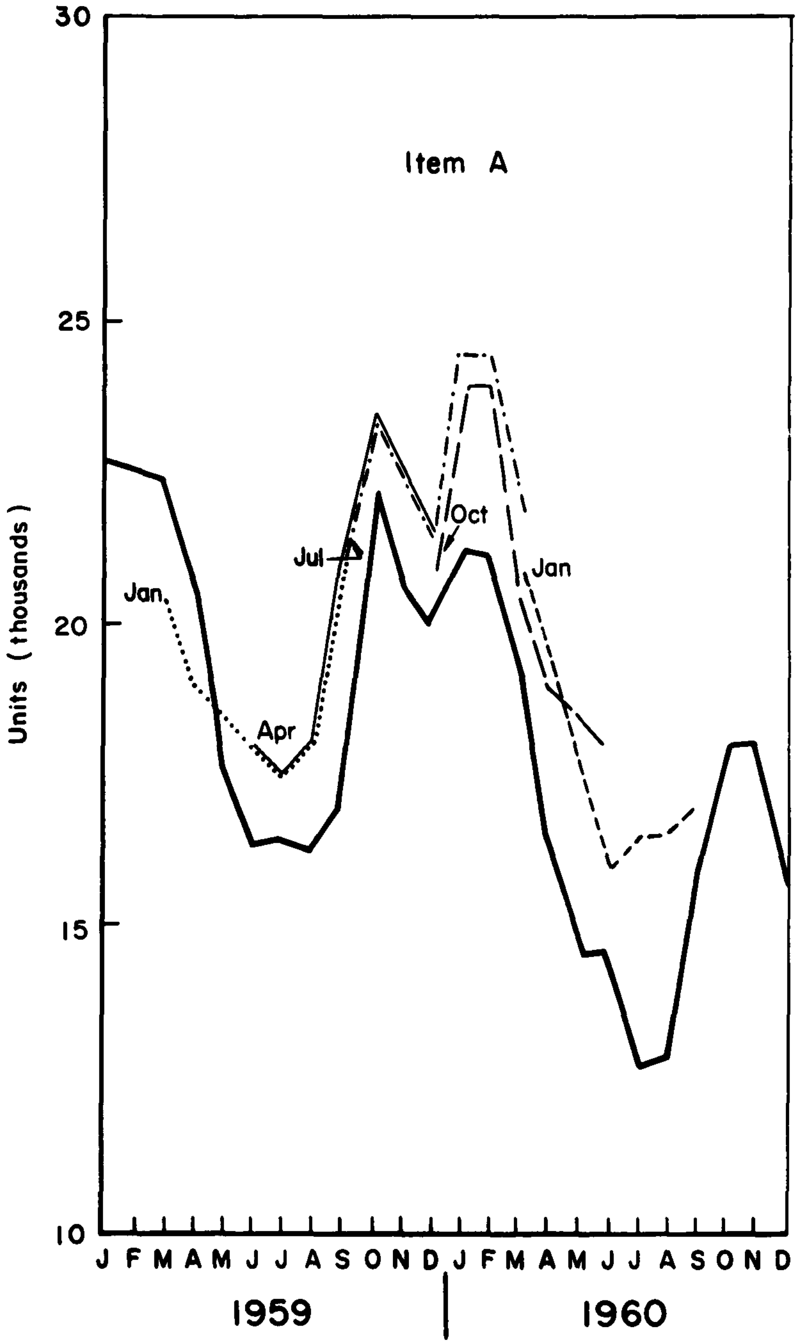

Fig. 3 — Actual demand compared to company forecast
(Forecasting 8 months ahead)

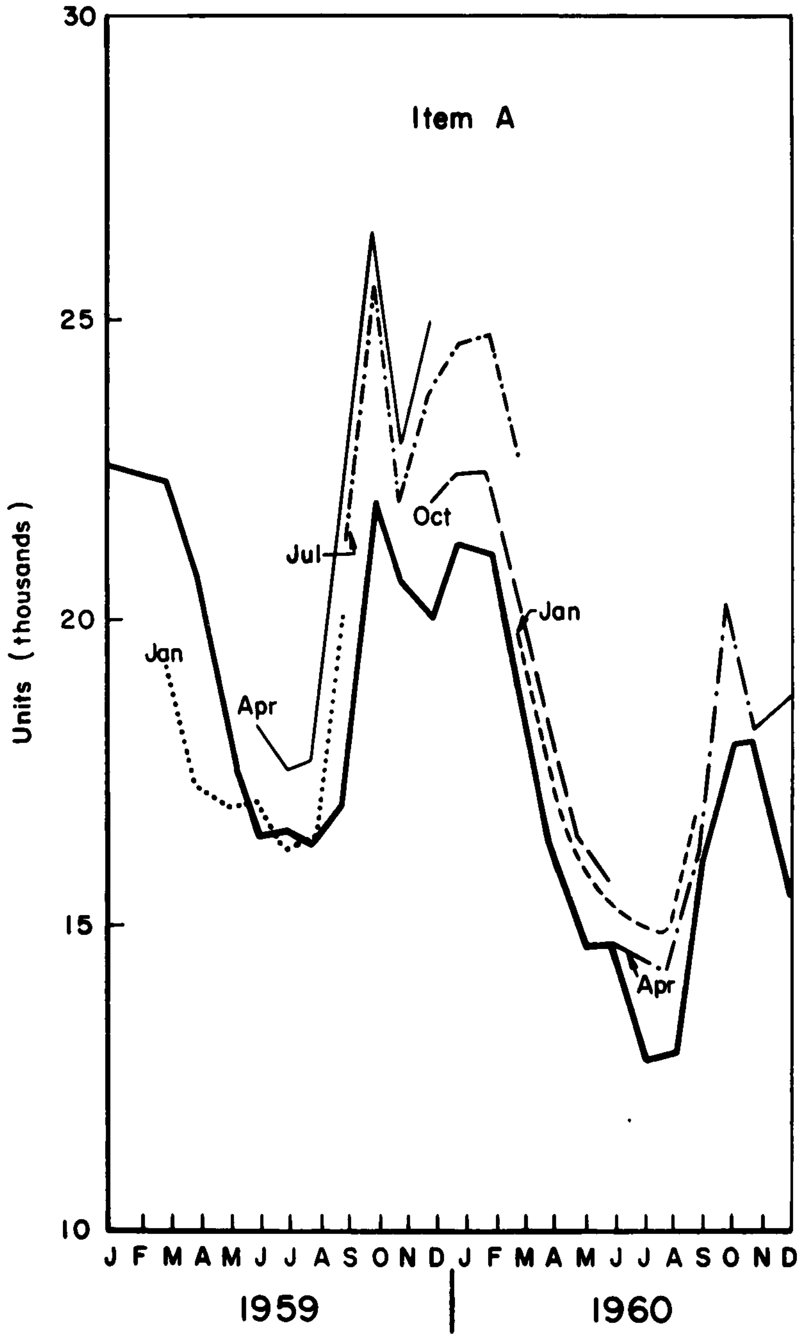

Fig. 4 — Actual demand compared to exponential forecast
(Forecasting 8 months ahead)

Using the model as indicated it was possible to try many sets of ABC weights and evaluate each one in terms of the forecast error. This is done in three ways: by computing an average fractional error, the standard deviation, and the coefficient of variation. We found the latter statistic to be the most useful.

The results achieved with the Winters model were very good. Coefficients of variation frequently were as low as 0.07 to 0.09 and rarely were they larger than 0.15. A look at a typical item will provide a graphical indication of the accurate forecast made using the exponential smoothing model compared with a forecast made based on a 12-month moving average. In Fig. 2 the actual sales for an Item A are shown for the last two years of the four-year period. In Figs. 3 and 4 the actual sales are plotted and, in addition, the forecasts have been simulated with each quarterly forecast (for 8 months ahead) being shown. Figure 3 compares the company's 12-month moving average forecast with actual; Fig. 4 compares the exponential forecast with actual.

The division of this pharmaceutical company is proceeding as of the beginning of 1962 to do their item forecasting using the exponential smoothing model. It is basically Professor Winters' model with a modification to eliminate an existing stop in the program necessitated by the requirement to have an actual sales figure in each period to be compared with the forecast in computing a forecast error.

Company Y, a second case, is a large wholesale and retail distributor of metal materials and metal products. Their inventory

is comprised of approximately 10,000 items. Demand for these items is highly variable with little trend or seasonal pattern present. In addition, the replenishment leadtime is highly variable. The records for about 70 per cent of the inventory -- in terms of value -- are maintained in a set of Kardex files. The records of the remainder of the inventory, mostly tools and supplies materials, are being converted to an IBM 305 system. The problem of concern here was to provide a sound and systematic way to review and, if necessary, recalculate the item reorder points.

The problem, of course, is not unique. There were two difficulties, however, in dealing with this particular problem. The first was one involving the considerable workload required to do the necessary analysis and computational work for such a large number of items.

In the second case, there was some difficulty in establishing statistical demand patterns. Taking a week as the useful demand period, it was found that some demand patterns appeared to be normal, others the Poisson type, and still others the gamma type. The analysis work required to develop and select an appropriate demand pattern proved too difficult -- for several practical reasons -- for the client people to deal with. In lieu of this step a Monte Carlo routine was developed to compute the cumulative probability distribution of usage during leadtime for each item based on the actual demand and replenishment leadtime data.

This routine is a Monte Carlo version of the following polynomial.

$$P(u) = \sum_{1}^{m} p(t_m) \left[p(x_1) + p(x_2) + \ . \ . \ . \ + p(x_n) \right]^{t_m} .$$

If one can assume that the actual individual distributions being used are descriptive of the item depletion and replenishment behavior, then the distribution resulting from the combination of these two will indicate the probability of a stock shortage occurring at any given point or replenishment level. Fig. 5 shows this distribution in the form of a cumulative probability of usage during leadtime.

Since only the right-hand tail of this distribution is important to us, it is necessary that the Monte Carlo sample size be large enough to assure that the tail is stabilized. The computer routine requires only a few seconds to make the calculation; thus we have found it convenient to use sample sizes of one thousand. This is not only adequate usually for statistical stability, but also provides a frequency count expressed directly as percentages.

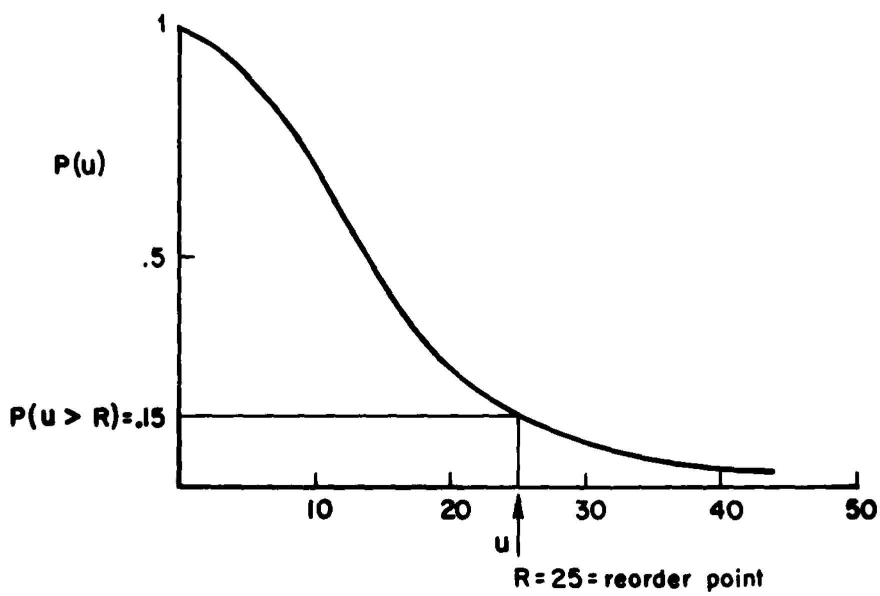

Fig. 5—Probability of usage during leadtime

At a relatively small cost this computing routine can be used to develop a large number of distributions for the probability of usage during leadtime which can then be classified and applied directly in determining the reorder points for most of the items in the inventory.

The third case involves Company Z which we mention primarily because this situation typifies a problem occurring frequently amongst our clients. This company produces several hundred items in a number of plants for national distribution. A linear programming model was developed to help in improving the production planning process. In the development of a dynamic version of this model -- one that would make monthly production allocations over a period of several months -- it was necessary to provide monthly item demand forecasts for each market area. These item forecasts were not being made nor were there any records of item demand history from which to develop such forecasts. Only aggregate forecasts were being made with the total being allocated or ratioed to each item. This is a relatively unsatisfactory way of developing item forecasts, and it has been found on initial trial that the exponential smoothing technique provides a satisfactory approach to forecasting demands for at least the large volume items even in the absence of much usable historical data.

These case examples typify the majority of item forecasting problems we have encountered. As you can see, the resolution of these problems depends on the availability of data, the capability of client people to deal effectively with somewhat more sophisticated forecasting techniques, and a careful economic evaluation of the degree of sophistication warranted in any given problem situation.

One of the objectives of this conference is to outline questions in areas of future work. I would like to suggest a few which, from our point of view, seem important. First, I believe more attention should be given to a consideration of how good a forecast must be in any given situation. Accuracy in forecasting costs money and one should attempt to be only as accurate or detailed as his related decision-making capabilities will justify.

Second, more work is needed in developing models and methods for dealing with outside effects, that is, the projection phase of forecasting. Finally, more recognition should be given to the problem of appraising and developing specific technical competence in user-organizations to assure that refinements in forecasting technique as developed and implemented will be monitored and maintained effectively.

This has been a highly useful conference, and it is a privilege to be able to attend and participate. I would strongly favor the planning of similar conferences for the future, and in doing this, would suggest that they be broadened to include more representation from the business and industrial field.

SOME REMARKS ON SPECTRAL ANALYSIS

George S. Fishman

Stanford University

For the past seven months, I have been assisting Professor Marc Nerlove of Stanford University in a study of hog and cattle slaughter in the United States. We have been interested in determining the relative importance of cyclical and seasonal phenomena in the monthly time series of hog and cattle slaughter. This was done by estimating the power spectrum associated with each of the economic time series. The estimation procedures used come under the heading of spectral analysis.

I would like to spend my time this afternoon giving a brief description of the theory behind the concept of the power spectrum, outlining the procedures used when estimating the power spectrum, and giving an example of the simplification which can be achieved in the theory of filtering by working in the frequency domain rather than the time domain. By filtering we mean the elimination of cyclical or seasonal components from the time series.

We know that a periodic function of time, $g(t)$, with mean zero, may be represented by a linear combination of sines and cosines,

$$g(t) = \frac{a_o}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega_1 t + b_n \sin n\omega_1 t) , \tag{1}$$

subject to the condition that $g(t)$ be absolutely integrable,

$$\int_{-T_1/2}^{T_1/2} |g(t)| \, dt < \infty \quad . \tag{2}$$

The coefficients in (1) are given by

(3a)
$$a_n = \frac{2}{T_1} \int_{-T_1/2}^{T_1/2} g(t) \cos n\omega_1 t \, dt$$

and

(3b)
$$b_n = \frac{2}{T_1} \int_{-T_1/2}^{T_1/2} g(t) \sin n\omega_1 t \, dt \quad .$$

ω_1 is the fundamental frequency of the period T_1 and is defined as follows:

(4)
$$\omega_1 = \frac{2\pi}{T_1} \quad .$$

Manipulation of relationships (1), (3a), and (3b) yields

(5)
$$g(t) = \sum_{n=-\infty}^{\infty} F(n) e^{in\omega_1 t}$$

where

(6)
$$F(n) = \frac{1}{2}(a_n - ib_n), \quad n = 0, \pm 1, \pm 2, \ldots$$

and

(7)
$$F(n) = \frac{1}{T_1} \int_{-T_1/2}^{T_1/2} g(t) e^{-in\omega_1 t} \, dt, \quad n = 0, \pm 1, \pm 2, \ldots \quad .$$

$F(n)$ is the Fourier transform of $g(t)$. $F(n)$ is complex and discrete for periodic functions. It assumes non-zero values only at integer values of n.

The autocovariance function for $g(t)$ is

(8)
$$\phi(\tau) = \frac{1}{T_1}\int_{-T_1/2}^{T_1/2} g(t)\, g(t+\tau)dt \quad .$$

Replacing $g(t+\tau)$ by its Fourier transform, we have

(9)
$$\phi(\tau) = \frac{1}{T_1}\int_{-T_1/2}^{T_1/2} g(t)\sum_{n=-\infty}^{\infty} F(n)e^{in\omega_1(t+\tau)}dt$$
$$= \frac{1}{T_1}\sum_{n=-\infty}^{\infty} F(n)e^{in\omega_1\tau}\int_{-T_1/2}^{T_1/2} g(t)e^{in\omega_1 t}dt \quad ,$$

and by comparing the expression under the integral sign with (7), we see that

(10)
$$\frac{1}{T_1}\int_{-T_1/2}^{T_1/2} g(t)e^{in\omega_1 t}\, dt = F(-n) \quad .$$

It can be shown that $F(-n)$ is the complex conjugate of $F(n)$. Let the complex conjugate of $F(n)$ be denoted by $F^*(n)$. Therefore, (9) is equivalent to

(11)
$$\phi(\tau) = \sum_{n=-\infty}^{\infty} F(n)\, F^*(n)e^{in\omega_1\tau} = \sum_{n=-\infty}^{\infty} |F(n)|^2\, e^{in\omega_1\tau} \quad .$$

At $\tau = 0$, we have

(12)
$$\phi(0) = \sum_{n=-\infty}^{\infty} |F(n)|^2 \quad .$$

$\phi(0)$ is the variance of the time series $g(t)$. It is equal to the sum of the squared absolute values of the spectral components of $g(t)$.

Therefore, $\frac{|F(k)|^2}{\phi(0)}$ is the relative contribution to the variance of $g(t)$ made by the frequency $k\omega_1$.

$|F(k)|^2$ is the contribution of frequency $k\omega_1$ to the power spectrum of $g(t)$. Let us denote it by $P(k)$. The power spectrum is real, symmetrical about $k = 0$, and discrete for periodic functions. We are interested in estimating these spectral components, $P(k)$, for all k , from our time series so that we may discern the relative importance of different frequencies.

The case of aperiodic functions is analyzed by means of an extension of the periodic case. We allow the fundamental frequency to approach infinity. This defines the aperiodic function as an infinite weighted sum of periodic functions. Then ω_1 becomes the infinitesimal $d\omega$ and $n\omega_1$ becomes ω_1. Therefore, (5) and (7) become

$$(13) \qquad g(t) = \int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\, d\omega$$

$$(14) \qquad F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}\, dt, \quad \text{respectively.}$$

The autocovariance function of an aperiodic function $g(t)$ is given by

$$(15) \qquad \phi(\tau) = \int_{-\infty}^{\infty} g(t)\; g(t+\tau)dt.$$

Replacing $g(t+\tau)$ by its spectral representation, we have

(16a) $$\phi(\tau) = \int_{-\infty}^{\infty} g(t) \int_{-\infty}^{\infty} F(\omega)e^{i\omega(t+\tau)}\ dt$$

(16b) $$= \int_{-\infty}^{\infty} F(\omega)e^{i\omega\tau}\int_{-\infty}^{\infty} g(t)e^{i\omega t}dt$$

$$= 2\pi\int_{-\infty}^{\infty} |F(\omega)|^2\ e^{i\omega\tau}\ d\omega\ .$$

Note that the integral,

$$\int_{-\infty}^{\infty} g(t)e^{i\omega t}dt$$

is the complex conjugate of $F(\omega)$ multiplied by 2π . This is a consequence of (14).

Let

(17) $$P(\omega) = 2\pi |F(\omega)|^2\ .$$

At $\tau = 0$, we have

(18) $$\phi(0) = \int_{-\infty}^{\infty} P(\omega)d\omega\ .$$

$P(\omega)$ is the power density spectrum of the aperiodic time function $g(t)$. Therefore, $\dfrac{P(\omega_o)d\omega}{\int_{-\infty}^{\infty} P(\omega)d\omega}$ measures the relative contribution to the variance made by frequency ω_o . $P(\omega_o)$ is real, symmetrial about $\omega = 0$, and is continuous for completely aperiodic functions.

The power spectrum for a function containing both periodic and aperiodic components is shown in Fig. 1. The discrete impulses are attributable to the periodic components of the time function, and the

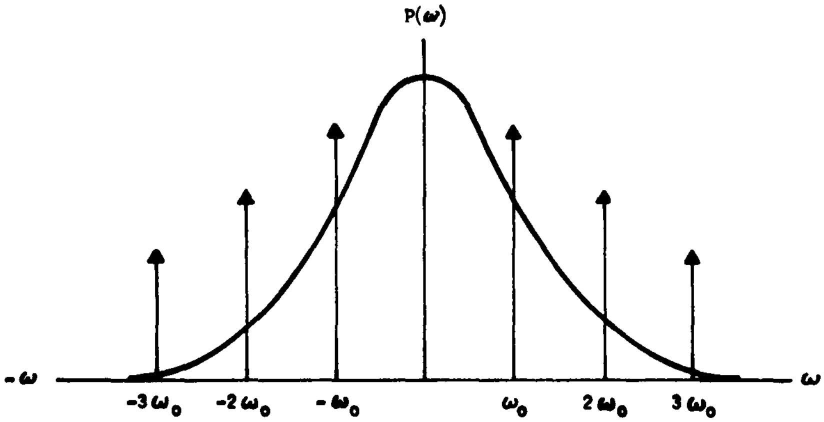

Fig. 1

continuous part is attributable to the aperiodic component.* Note that

$$\phi(\tau) = \int_{-\infty}^{\infty} P(\omega)e^{-i\omega\tau}\, d\omega \quad . \tag{19}$$

That is, $P(\omega)$ is the Fourier transform of the autocovariance function $\phi(\tau)$.

Many problems arise when we attempt to estimate the power spectrum. The first difficulty to be considered is often referred to as "aliasing." This problem is made apparent by the following

---

*Strictly speaking, the discrete impulses shown in Fig. 1 are quantities with a dimensionality different from that of $P(\omega)$ itself. These discrete components have the dimensionality of $P(\omega)d\omega$, and they have been drawn in the same diagram with the continuous $P(\omega)$ merely for visual convenience.

hypothetical example: Suppose that a Martian is interested in determining how often we have Congressional elections here in the United States. Let us assume that he takes observations of political life in the United States only in calendar years divisible by four. He may conclude from his observations that we have Congressional elections every four years. However, it is clear that this conclusion is not quite adequate since his observations cannot discern whether or not Congressional elections are held every two years or every year. This results from the fact that his observations are being taken at equal spaced intervals of four years. Analogously, we are unable, in spectral analysis, to distinguish between the contributions to the power spectrum made by ω_0, $2\omega_0$, $4\omega_0$, etc. (see Fig. 2) since observations are made at equally spaced intervals of length Δt.

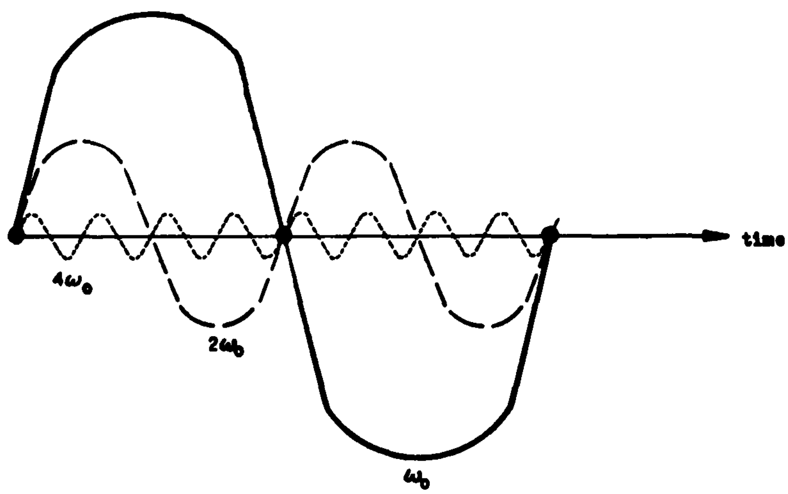

Fig. 2

In the case of monthly data, the highest frequency for which we can clearly determine the power is π, the frequency associated with a period of two months.

In general, we space our observations in such a way that we would expect the power associated with the frequencies greater than π to be relatively negligible. We therefore do not treat aliasing as a critical problem when estimating the power spectrum. Nevertheless, we should be aware of its existence.

In estimating the spectrum, we are interested in deriving sample estimates of the power spectrum associated with a time series X_t. We assume that X_t has a zero mean and that $E(X_t X_{t+\nu})$ is independent of t. From the aforementioned theoretical description, one would expect the spectral estimates to be of the form

$$(20) \qquad P^{(n)}(\omega_0) = \frac{1}{\pi}\left[2 \sum_{\nu=1}^{m} R(\nu) \cos \omega_0 \nu + R(0)\right] ,$$

where n is the number of observations in the sample, $R(\nu)$ is the estimated autocovariance for a lag of ν months, and is defined as

$$(21) \qquad R^{(n)}(\nu) = \frac{1}{n} \sum_{t=1}^{n-\nu} X_t X_{t+\nu} .$$

Note that as it is written $R^{(n)}(\nu)$ is a biased estimator of the autocovariance, $R(\nu)$. It is preferred to an unbiased estimator where the denominator is $n-\nu$ rather than n, for two reasons: (1) to achieve a minimum mean square error, and (2) to achieve a positive definite estimate function. The necessity for this second property derives from the fact that the true power spectrum is everywhere greater than zero, and the unbiased estimate of $R(\nu)$ may yield negative values for the estimated power spectrum at some frequencies.

Unfortunately, estimates of the spectrum based on (20) lack the desirable statistical property of consistency. That is, if a number of samples, each with an infinite number of observations per sample, were used to derive separate spectral estimates for frequency ω_0, one would find that these sample estimates, based on this infinity of observations, would not come out to be equal to the number $P(\omega_0)$, but rather would form a random scatter around $P(\omega_0)$ and have an exponential probability distribution.

In 1948, Bartlett pointed out that if one were to divide a sample of n observations into p sets, each containing m observations, and then derive sample estimates of the spectrum for each set and average the power estimates for the frequency ω_0 over all sets, one would have a consistent estimate of the true spectral average centered at ω_0. It can be shown that this procedure is equivalent to introducing a weighting function in (20) such that

$$P_{Av}^{(n)}(\omega_0) = \frac{1}{\pi}\left[2\sum_{\nu=1}^{m}\lambda_\nu R(\nu)\cos\omega_0\nu + R(0)\right] \tag{22}$$

where

$$\lambda_\nu = 1 - \frac{\nu}{m} \qquad (\nu \leq m) \tag{23}$$

$$= 0 \qquad \text{elsewhere.}$$

Since 1948, much work has been done on other types of weighting functions which have come to be called windows in the technical literature. It has been shown that

$$(24) \qquad \lim_{n\to\infty} \int_0^{\pi} P^{(n)}(\omega)\ A(\omega)d\omega = \int_0^{\pi} P(\omega)\ A(\omega)\ d\omega \quad ,$$

where $A(\omega)$ is the window used to derive the estimates of the true spectral averages. What (24) says is that the sample spectral average is, indeed, a consistent estimate of the true spectral average.

It can be shown that (24) leads to

$$P_{Av}^{(n)}(\omega_o) = \frac{1}{\pi}\left[2\sum_{\nu=1}^{m} k\left(\frac{\nu}{m}\right) R(\nu)\ \cos\ \omega_o\nu\ + R(0)\right]$$

where $k(\frac{\nu}{m})$ is the Fourier transform of $A(\omega)$.

One chooses a window $A(\omega)$ which will weight the spectral estimates close to ω_o as heavily as possible. In this way, the estimated spectral averages will be more representative of the true power at ω_o. To gain this desirable property of resolution, one must sacrifice another desirable property, namely, stability. That is, the gre[illegible] the number of lagged contributions contained in the estimate[illegible] ectral average, the better the resolution, but the poorer t[illegible] ean-square error of the estimate of the true spectral average. [illegible] one sees that we must so choose the number of lags in the window to yield some compromise between resolution and stability.

The adjustment of economic time series to eliminate seasonal phenomena has long been a topic of interest to economists. This filtering process can be greatly simplified theoretically if one analyzes the time series in the frequency domain rather than directly in the time domain.

Let $g(t)$ be the input to a system whose response to a unit impulse is $h(t)$. Let $f(t)$ be the output of this system. If $g(t)$ were a unit impulse, then

$$(25) \qquad f(t) = h(\nu)\ g(t-\nu)$$

would be the output response to a unit impulse which is applied to the system at $t = \nu$, where $h(\nu)$ is the time response characterizing the system into which the signal is fed.

If one thinks of $g(t)$ as a series of impulses forming a continuous signal, (25) becomes

$$(26) \qquad f(t) = \int_{-\infty}^{\infty} h(\nu)\ g(t-\nu)d\nu \ .$$

It can be shown that the Fourier representation of (26) is

$$(27) \qquad F(\omega) = H(\omega)\ G(\omega) \quad ,$$

where $F(\omega)$, $H(\omega)$, $G(\omega)$ are the Fourier transforms of $f(t)$, $h(t)$, and $g(t)$, respectively. This leads to

$$(28) \qquad P_F(\omega) = |H(\omega)|^2\ P_G(\omega) \quad .$$

That is, the power density spectrum of the output signal is simply the power density spectrum of the input multiplied by $|H(\omega)|^2$.

To show the importance of (28) in problems of filtering economic time series, we consider the spectrum of X_t , where X_t is an economic time series in which we suspect seasonal and cyclical behavior to exist.

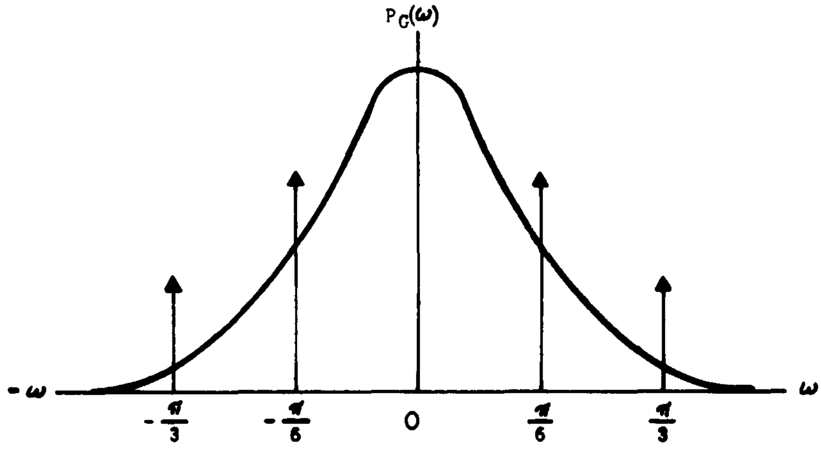

Fig. 3

The discrete impulses at $-\frac{\pi}{3}, -\frac{\pi}{6}, \frac{\pi}{6}, \frac{\pi}{3}$ are attributable to the seasonal components of periods of 12 and 6 months in X_t. The contributions of other seasonal frequencies have been omitted for simplicity.

One often seeks to eliminate these impulses when attempting to derive a seasonally adjusting economic time series. For many time series, the contribution of the continuous spectrum for $|\omega| \geq \frac{\pi}{6}$ to the total power spectrum is negligible. Therefore, a reasonable way to eliminate the seasonal behavior is to define

$$(29) \qquad H(\omega) = 1 \qquad |\omega| \leq \omega_0$$

$$= 0 \qquad \text{elsewhere.}$$

ω_0 is chosen arbitrarily close to $\pi/6$, the lowest seasonal frequency. Such a filter will modify the spectrum as in Fig. 4.

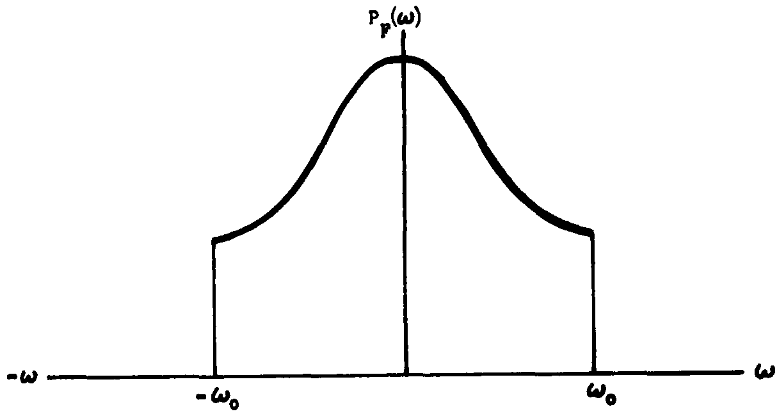

Fig. 4

This eliminates the seasonal components, which is what we actually want.

We know that the time representation of $H(\omega)$ is

(30) $$h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)e^{i\omega t}d\omega$$

$$= \frac{\sin \omega_0 t}{\pi t} .$$

Therefore, the output $f(t)$ is

(31) $$f(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin \omega_0 v}{v} \; g(t-v)dv .$$

An attempt has been made to use a close approximation of the ideal filter in (30). However, 180 leads and 180 lags are needed in the resulting weighted average in order to obtain a minimum level of acceptability for the adjusted series. This is obviously impractical

since, with such a filter, we could not adjust with confidence any data in our time series beyond 1946.

My example of the ideal seasonal filter was meant to show the greater simplicity in conceptualization when one works in the frequency domain. Our discouraging results in approximating the ideal filter ought not to be considered exhaustive of all possible alternatives.

In my introductory statement I mentioned that we have been using spectral analysis on hog slaughter and cattle slaughter time series. One reason for concentrating on these series was the availability of over 50 years of monthly data for each series. Long series are preferred in spectral analysis in order to derive good estimates with the desired statistical properties. We have made some preliminary investigations into shorter series of 20 years. Our results have been favorable and we feel that spectral analysis may be applied to many economic time series of 20-year duration with success.

This presentation has certainly not been all-inclusive of the problems of spectral estimation and filtering. Much remains to be mentioned. However, I hope that the reader will derive some minimal familiarity with spectral analysis and what it offers as a research tool.

BIBLIOGRAPHY

Jenkins, G. M., "General Considerations in the Analysis of Spectra," Technometrics, May 1961, pp. 133-166.

Lee, Y. W., Statistical Theory of Communication, John Wiley and Sons, New York, 1960.

Parzen, E., "Mathematical Considerations in the Estimation of Spectra," Technometrics, May 1961, pp. 167-190.

THOUGHTS EXPRESSED AT THE CONFERENCE

Frederick S. Nowlan

United Air Lines*

My thoughts are somewhat biased since it seems to me that when we are studying demand figures we are, in effect, studying reliability characteristics through the medium of demand data.

The general findings that demands were not too closely correlated with operational usage appears somewhat surprising to me since a generally used reliability index is the ratio of unscheduled removals (these give rise to demands) to the volume of operational usage.

This index may show large fluctuations from one calendar time period to another but further technical analysis frequently can explain much of this variation. Typical of such causes of major fluctuations are:

(1) The definition of satisfactory performance. We frequently are very demanding during the initial period of operation of new equipment before we are familiar with its intrinsic characteristics. In this case equipment is "squawked" and removed during initial operations for reasons that will not lead to removal and demand creation during later operation.

(2) Campaign inspections as a result of the first discovery of unsatisfactory conditions as a result of the first inspection or failure in a hidden area of an aircraft. After the first identification of a source of trouble in an airplane, the resulting campaign inspection can lead to a greatly increased replacement rate for a short period of time.

---

*Maintenance Base, San Francisco International Airport.

(3) File maintenance procedures which fail to identify chronic airplanes where the malfunction is not being corrected by unit replacement and which do not lead to the more thorough trouble shooting required to cure the problem.

(4) Component modification. Engineering action is directed at improving the reliability of components with high replacement action. Modification may be very successful and greatly reduce the replacement rate. On the other hand there may be an unsuccessful modification.

There are of course many other factors which can lead to more gradual changes in removal rate and the demand level. Such factors are:

(1) Overhaul specifications and procedures.

(2) Overhaul periodicity.

(3) Shop quality control procedures.

(4) Operating procedures.

(5) Age of units in service, etc.

It would seem that all these reliability considerations would make correlation analysis of demand data rather difficult. The analysis of demand data would seem to entail very close liaison with people concerned with component reliability.

This might be difficult to achieve. It is known, however, that very extensive reliability studies and replacement records are available for the Lockheed Georgia C-130 and some of the Boeing aircraft. It might be possible to correlate these reliability data with demand data.

Again, large bodies of data exist within the air transport industry and some of these data might be useful for general demand prediction studies.

SOME THOUGHTS ON THE DEMAND PREDICTION JOB AHEAD*

Emil W. Hamilton

San Bernardino Air Materiel Area**

THE APPLICATION OF SELECTIVE MANAGEMENT PRINCIPLES TO DEMAND PREDICTION

One conclusion I'm sure we would all agree with is that there is no "best way" to predict demand or forecast future needs. There are only "preferred" ways -- and these will vary depending upon many circumstances.

This calls for development of Selective Management approaches wherein suitable demand prediction techniques could be tailored for application to groups of items having similar technical or management characteristics. We have moved in this direction in the last few years, but the groupings are largely cost-oriented. Within these cost breakouts -- which have proved their merit for management and control purposes -- there appears to be quite a potential for further "slicing" by demand or volume characteristics to give us a better handle on the forecasting problem.

LOW DEMAND ITEMS

With low demand items predominating, this would be a good place to start. Here we have unlimited opportunities for mass simplification of data collection, program elements, factors development, computational

*This paper summarizes some remarks made by the author at the Conference.

**Norton Air Force Base.

techniques, etc. Looking ahead, the nature and size of future weapons programs portends an even higher ratio of low demand items than in the past.

Today there is much waste motion at provisionings and during replenishment requirements computations as people (and machines) wrestle with the development and application of microscopic factors that frequently lead to unrealistic "sub-minimum" quantities. If we fail to adjust the computed answers, we trigger a whole host of unrealistic actions in the logistic cycle. If we do adjust the answer to something more realistic, the question arises about whether we couldn't have gotten this final more defensible answer much more easily and directly.

Although we may have no real basis for a "computation" or a "demand prediction" as such for these low demand items, we are not free of the job of making, as best we can, a "determination" of some reasonable range and depth of items that must, for a variety of reasons, be readily available if ever called for. Even if we cannot be as objective and as scientific in our approach here as some would like, I believe we can improve on what we do today. At least we should be able to sharpen up our "qualitative/subjective" approaches and give them a measure of recognition and respectability that will enhance their use as "tools of the trade."

What appears to be needed is some well chosen criterion for categorizing items as low demand and then a delineation in check list form of the various "thought processes" that need to be taken into account in the decision of what and how much to bring into the system.

Thus with a few relatively simple approaches -- not so much mathematically or statistically oriented, but slanted more toward better delineation and application of the proper "thought processes" -- we should be able to greatly simplify the requirements task for the mass of low demand items in the various cost categories and at the same time achieve more realism in the results.

Perhaps the spirit of what we are saying is embodied in the observation of one of the Carnegie representatives when he said that "...instead of trying to predict specific occurrences, maybe we ought to try to prescribe the parameters that evaluate risks...then we are not so much predicting demand as predicting risks..."

HIGHER DEMAND ITEMS

For the few items with relatively high demand characteristics, we would go "all out" in search of refined, scientific prediction techniques. We could afford for this more manageable number of items, the more expensive data collection systems, more meaningful program elements, more sophisticated computational methodologies, etc. This would be the area for exploiting further many of the types of formulae discussed at the conference.

In analyzing such approaches with a view toward determining their applicability, we should be less concerned with the numbers of items the techniques could be applied to, and more concerned with their potential pay-off in relation to the monetary worth and mission essentiality of the specific item involved. The mere fact that a given technique seems to work best for most of the items is not sufficient reason for adopting it for use "across the board."

Here the challenge should be to develop the capability for associating technical knowledge of the item and knowledge of relative merits of the various demand prediction techniques to permit some fairly judicious choice of the optimum method for that item. This choice might well range from the relatively simple traditional issue interval techniques on through variations of service life, failure pattern, distribution curve, probability, upper and lower limit, actuarial and other approaches.

INITIAL VS. FOLLOW-ON PREDICTIONS

No discussion of this type would be complete without some mention of the need for continuing the quest for improving the initial provisioning decisions which, especially for the crucial peculiar items, must be made without benefit of experience or statistics. Some of the points discussed under the topic of "Low Demand Items" have application here. This is becoming an increasingly important area, particularly for "static" weapons like ballistic missiles whose numbers are few, whose life spans are short, and whose operating time is limited. Thus the initial buy, however small, frequently is the total buy. Though we may set up an elaborate system for collecting a lot of data to help us with the follow-on buy, we may never reach the point where this information will be of much help to us.

PROGRAM ELEMENTS

As I mentioned at the conference, we here at SBAMA feel that one fruitful area for further exploration is that of program elements, particularly those appropriate for peculiar missile and spacecraft

applications. The need for some look at this area became quite apparent during the conference with discussions tending to gravitate around traditional elements like aircraft flying hours. Today in the ballistic missile business we use mostly missile months, equipment months, squadron months, engine months, and the engine overhaul program. But these have been somewhat arbitrary selections and perhaps a good look is in order now before we find ourselves locked in on these for no better reason than "we've always done it this way."

SEASONAL TYPE PEAKS

One thought triggered by Mr. Brown's presentation: For items that tend to have certain types of predictable peaks, instead of trying to develop average rates that reflect these peaks, we would get better time-phased answers if we determined the average and peak requirements separately and then overlaid the results.

PAST STUDIES THAT MIGHT MERIT NEW LOOK

Perhaps some review and revitalizing of past findings, conclusions and recommendations -- updating and reslanting them toward new modern weapons and logistics applications -- might in itself be a fruitful area for future effort on the part of research agencies. Two examples of past RAND research come to mind as possibly being worthy of some new look in light of current conditions and needs:

> One is the study on "*A Priori* Demand Prediction" which developed the hypothesis that physical and operational characteristics were generally sufficient to permit classifying parts as high-demand, low-demand, or intermediate-demand items.

The other is the "Flyaway Kit" study. Perhaps some of the principles in this could be applied to the provisioning and buy processes, with cost being considered the constraint rather than the weight and cube criteria which were more appropriate for Flyaway Kits.

THE GROWING NEED FOR UNDERSTANDING AND ACCEPTANCE

In retrospect, I find myself wondering about the "follow through" aspects of research on demand prediction and how well we in the Air Force (and perhaps other services) understand and are able to apply some of the products of past research. With research and consulting agencies being concerned about where and how they can make further contributions at this time, I wonder if there isn't a considerable potential in the area of achieving better understanding and acceptance of their studies and proposals. At some future conference, an exchange of ideas would no doubt be beneficial on (1) what media the various groups employ to "spread the word" and gain acceptance, and (2) how effectiveness in this field can be increased.

I was impressed with a remark Mr. Dalleck made in this connection -- that as a consulting firm, they try to assay a company's technical competence to carry on after they leave. This would imply that they either tailor their approaches and techniques accordingly, or put forth some special effort to develop the capability of their customer to follow through in the use of the new tools.

Quite often the principles and formulae offered by researchers and consultants need considerable "translation" before they can be

understood and applied by procedures writers and technicians. Whether the researchers themselves or another body of folks perform this "translating" and stimulating and educating is, I suppose, a matter for debate -- but the gap must somehow be bridged if we are to reap full benefit from the studies.

One company that, I am told, devotes considerable effort to broadening understanding is the Planning Research Corporation of Los Angeles which does work for the Navy. My understanding is that they have been quite active in supplementing their research papers with films, animated illustrations of various theories and concepts, training syllabuses, and physical tools like tables, special slide rules, hand calculators, and the like. Thus many of their products are beamed directly to the user.

In the case of RAND, we can cite your Research Memoranda, the Logistics Laboratory, your scheduled appearances at logistics schools, presentations at conferences, briefings to AF people, etc., as excellent means of "spreading the word." But the audiences reached are still no doubt relatively limited. Perhaps there is a challenge in trying to reach a wider audience with more of the basics and fundamentals, with more "why" and "how."

CONFERENCE REMARKS

Kenneth J. Arrow

Stanford University*

Since so much of the Conference program was devoted to discussion, the following remarks are included to summarize some of the thoughts that came out of the meeting as a whole. Our discussion was originally oriented towards predicting or forecasting demands. As we proceeded, however, our comments branched off into contiguous matters for reasons which are useful to examine.

THE PREDICTION PROBLEM

There are two extreme approaches to the prediction problem. One is a kind of structural analysis, in which we try to discover what really causes demands, and base a predictive theory on our findings. The other extreme is to take a historical approach to the process of change. We may have no defensible theory to explain how or why demands have changed, but we feel they do not change abruptly. We can get some clues about the nature of this change by "watching" the situation closely, either in person or through statistical techniques. Essentially, we extrapolate backwards into the past, trying to identify relevant trends and discard the "noise" (as it is now called), or random disturbances (as we used to know it before the communications engineers entered the field). In real life, obviously, we employ

*Department of Economics

both approaches simultaneously -- in the face of numerous people's insistence that not even the most meticulous of scrutinies will yield explanations or predictions any more reliable than those of a weatherman.

Very little was said at this conference about tying in the work of the reliability engineer with that of the demand forecaster, possibly because we have talked mostly about the military environment. Here we deal with so large a number of items that such integration is difficult even though the reliability engineers may have all kinds of useful knowledge about each item. This integration may or may not be worth pursuing, but at least it might give us some additional insights regarding structural possibilities.

DEMAND RATES

One interesting problem came out in the casual discussion of structural possibilities. Demand rates, according to one report, heavily depend on usage; according to another, apparently based on similar data, they depend only on elapsed time. It seems to be true that many parts are subject to an aging process, so that we have long periods of low demand followed by rising demand as the equipment ages. This is countered by another set of observations which show that there is a "burn-in" phenomenon; i.e., demands are numerous when the equipment is relatively new, and decrease as the bugs are ironed out.

There is also a point of view based on different data, which suggests that the important items are those with a very small number of demands, say one or two, over a long period of time. Here the prediction problem differs from the one in which there is a large

number of demands. The RAND report affirms that there are both kinds of items, and we may have to treat them separately. Obviously, it is difficult to correlate a scattered handful of demands with some program element even if there is a correlation; hence we would almost never get good estimates using this type of analysis. It follows that high-demand items offer greater possibilities for analysis. On the other hand, the low-demand problem is easier to meet and solve since the policies required to handle it are not dependent on any program element.

STATISTICAL FORECASTING TECHNIQUES

It is not easy to form an *a priori* opinion about the fruitfulness of statistical forecasting techniques. This needs to be done empirically. In one way or another, most of these methods, apart from spectral analysis, seem to be of the discounted least-squares type.

For what model of the world would this statistical method be correct in any sense? When we consider that we are getting a set of observations from virtually the same universe, why is every observation not as good as any other one? It is not clear why one should weight them in any way.

The interpretation made by Winters was that we can think of the parameters as shifting, but shifting in a random-walk manner, which adds one more unknown. Looking ahead, things are getting more and more uncertain, and discounting compensates for this growing unreliability. By the same token, if we start from the present the past data are more uncertain. Thus, if we arrived where we are now by

a random-walk process, we can also go backwards by the same process. This may be the rationalization for this kind of least-squares method. We have a past marked by change; we think change will persist as we go further into the future. The discounted least-squares methods compensate for this process in some way.

Brown's arguments have shown the flexibility of this model. One can build a great deal into it, apparently, much more than by straight exponential smoothing. Furthermore, it is possible to bring in any explanatory variables we like, such as program elements and age of parts. A combination of smoothing techniques may produce better results with different program elements.

Another point raised concerned the program elements themselves. Assuming that usage does have established relationships to some program elements, then in order to forecast usage we also have to forecast the program elements. Doing so introduces additional "noise." Brown's argument is that it is better to use, as explanatory variables, mathematical functions of time about whose extrapolation there is no question.

There is a counterargument which depends on the use you can make of the forecast. A conditional forecast gives some information that an unconditional forecast does not. It tells us what will happen if we change our minds, so that we might say, after looking at it, "It's really too expensive to fly those things around. I'd better not do it." This would make our forecasting worse, but it would also answer a question that could not be answered with an unconditional forecast. One could think of other examples of this kind, such as the question of checkouts. How often should we perform checkouts if it becomes

apparent they are a positive cause of failures? We might try to get some idea of the optimum number of checkouts. Probably we will not reduce the number to zero, since we value the extra confidence more checkouts can give us. We will perform them, then, but want to know what they cost us in terms of additional failures.

RELATION TO INVENTORY CONTROL

Up to this point we have talked about ways of making forecasts. Another interesting question is how they are to be used; and another -- perhaps amounting to the same thing -- is how one can measure the accuracy of his forecasts. There are many measures one can use: the average error, the mean absolute error, and the relative error - to name a few. The best one to use in a particular situation depends on the use to be made of the forecasts, and the costs of making various kinds of forecasts must also be considered. Sometimes the cost may be prohibitive.

Almost everyone agrees that if we can make forecasts we should do so; but if it turns out that we cannot, we must find a way of doing without them. There are many alternatives, one being the control of inventory. Obviously, the more accurate our forecasts, the smaller the inventories we have to hold (pipeline considerations aside). Behavioristically, to say we are unable to forecast accurately is much the same as saying we have a very diffuse demand distribution. There is a peculiar difference, however: the demand distribution may not actually be too diffuse, but if it changes from time to time it is not even easy to say what is actually meant by the demand distribution. What is really relevant is our subjective demand distribution -- one

we can realiably think of. It is a blurred average of the band of possible demand distributions -- even a pretty wide one. Anything which is optimal against such a wide distribution is liable to be fairly unresponsive to actual demand.

Ignorance costs something. It is an unavoidable cost and all we can do is try to hedge against it. It occurs, for example, in the sampling inspection of lots of material: if we do not know whether these lots are produced by a controlled process, there are sampling schemes which insure that the percentage of defectives will not exceed a specified value regardless of the quality of the incoming lots. We pay for this knowledge, of course, according to the amount of inspection we institute.

There is an analogy here in inventory control. Presumably we have better control over our reliability data, however poor some of our discussion implies they are, but it does suggest the idea of strong hedging. In practice, probably the simplest thing to do would be err on the side of overstockage. For expensive items, it might pay to subsidize the manufacturer to maintain some standby equipment or capacity to meet additional demands on short notice. It is costly to do so, but may be the price we must pay for lack of knowledge. The knowledge-criterion also buttresses a point already mentioned: that forecasting techniques should be evaluated according to their effect on inventories.

Now suppose we know the underlying demand distribution. In such a case, the mean is not the only characteristic we would be interested in; we would also be concerned with the variance. Our inventory

controls might be set at, say, two standard deviations or whatever number we thought appropriate. As Dalleck pointed out, however, we also have a close interest in the tails of the distribution and consequently make things worse for ourselves since it is very tricky to estimate high percentiles. Consider the problem of flood control. Let us say we want to build a dam with only a 0.01 probability of not holding a flood; that is, only once in a hundred years (the return period as it is called) will the flood exceed the dam. We cannot make a very accurate statistical estimate of the return period for a hundred-year flood if we have only 50 or 60 years of data; the result is very sensitive to the assumptions made about the distribution. There is much less sensitivity in an estimate of the mean. In fact, the variance of the mean is independent of the distribution. When we consider the tails, however, the problem is more difficult. This is probably why the use of Tchebycheff's inequality was suggested; it can work in practice because all the distributions we know about are well inside the upper bound it provides. Its only drawback may be that the upper bound obtained by using it may be too expensive for the protection it gives when we are dealing with very expensive items.

This discussion of tails bolsters the argument that it is as important to look for protection as for the other objectives of the policies we want to adopt. Further, it is for this reason that even if we knew the distribution we would want to consider more than the mean. It follows that if we do not know the distribution -- and perhaps decide to act as if we had a widely diffused one -- it becomes even more important to concern ourselves with the tails.

Suppose we have a procedure for setting inventory levels based on our forecasts. Through simulation we can determine what the inventory would have been in each period, counting back orders as negative inventory. We can summarize the figures by finding the average amount of inventory on hand for those periods for which there was a positive inventory, and the average amount of back orders for those periods in which there was not. These two numbers contain all the relevant information. We can then compare two forecasting methods by looking at these two numbers obtained from them. If both numbers are smaller for one of the methods, then it is clearly the better. The real problem arises when one method yields a smaller positive number, and the other a smaller negative; we then face the difficult task of assigning explicit values to each figure.

A simple and commonly used measure is the cost of the item. This solution implies we would sooner take an aircraft out of commission than overstock by one part. Further, it implies a different stockage policy for each item. We are assuming here that the system is rational in all other respects -- perhaps an extreme assumption, but certainly a plausible one. Again, some items are very expensive and others very cheap. If we want to economize on our investment, it would be wiser to stock an abundance of cheap items than to have an aircraft out of commission.

With commercial airlines, the loss in revenue when a scheduled plane does not fly is often less important than the nuisance to the customers. Again, there are some warehouse costs, interest on capital invested in spare parts, and the like, which should be considered, but

at least we would have some idea of the relative order of magnitude of the costs.

AGGREGATION

The question of aggregating across commodities came up several times in different contexts during the conference. When we have something like two million items, we must aggregate because it takes too long to set up an optimal inventory policy for each of them. We would aggregate in wide classes according to the criterion of value, although some very expensive items would have to be treated individually. There is probably no particular advantage in being very precise about these values.

The second kind of aggregation is through common limits. For example, we may have a budget which limits the total holdings of inventory, so that the more we have of one item the less we can hold of another. This constraint creates a problem since we may not be able to balance stockouts against holdings, item by item. Theoretically we would have to put a shadow price on the constraint; as a result, the storage costs would be something different from the plain calculated storage costs.

Another constraint might be cubage, as in a submarine. In such a case we would have to make choices on a different basis. Again we would need a shadow price for this constraint, but it would be difficult to use. If we could get enough information we might handle the problem as one of straightforward programming, but getting the input data would be a difficulty.

A third aggregation procedure, which came up several times, might be called aggregation through borrowed experience. This is one of the forecasting techniques used in the RAND study; the variance was taken to be three times the mean, the number three being derived from pooled experience. Brown and Winters have related the standard deviation to the mean, apparently to get some impressions of the variance. This could also be done by working with a sample of the commodities and then perhaps aggregating or pooling the results. Actually there is no reason in the nature of things -- unless for example we know we have a Poisson distribution -- why the standard deviation should be directly related to the mean. Pooling presupposes a commonality that has not been demonstrated.

A final comment, which perhaps should have been made earlier: several people seemed to agree that the negative binomial is a good description of demand, and yet nobody seems to have used it, perhaps because it is hard to work with. The Poisson, normal or log normal are more prevalent.

Appendix

LIST OF PARTICIPANTS

Kenneth J. Arrow, Stanford University

Max Astrachan, The RAND Corporation

Robert G. Brown, Arthur D. Little, Incorporated

Albert S. Cahn, The RAND Corporation

Margaret Chow, United Air Lines Maintenance Base

Winston C. Dalleck, McKinsey and Company, Incorporated

George S. Fishman, Stanford University

Sheldon Haber, The George Washington University

Emil Hamilton, San Bernardino AMA, Norton Air Force Base

James W. Houghten, The RAND Corporation

William H. Marlow, The George Washington University

John F. Muth, Carnegie Institute of Technology

Marc Nerlove, Stanford University

Frederick S. Nowlan, United Air Lines Maintenance Base

Jules Silver, Air Force Logistics Command, Wright-Patterson Air Force Base

Henry Solomon, The George Washington University

Harvey M. Wagner, Stanford University

Peter R. Winters, Carnegie Institute of Technology