

AD-776 823

FORECASTING TRANSPORTATION TONNAGE
BY AN ADAPTIVE SMOOTHING MODEL

Francis E. James, Jr.

Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

December 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

AD776823

AU-AFIT-SL-10-73

FORECASTING TRANSPORTATION TONNAGE
BY AN ADAPTIVE SMOOTHING MODEL

FRANCIS E. JAMES, JR., Colonel, USAF
Tenure Professor of Management and Statistics
Chief, Graduate Education Division
School of Systems and Logistics
Air Force Institute of Technology
Wright-Patterson AF Base, Ohio

Approved for Public Release;
Distribution Unlimited

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AU-AFIT-SL-10-73	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) FORECASTING TRANSPORTATION TONNAGE BY AN ADAPTIVE SMOOTHING MODEL		5. TYPE OF REPORT & PERIOD COVERED AU TECH REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) FRANCIS E. JAMES, JR., COLONEL, USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS GRADUATE EDUCATION DIVISION SCHOOL OF SYSTEMS AND LOGISTICS (SLGR) AIR FORCE INSTITUTE OF TECHNOLOGY, WPAFB, OH		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS SAME AS ITEM 9		12. REPORT DATE December 1973
		13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) APPROVED FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES CLEARED FOR PUBLIC RELEASE IN ACCORDANCE WITH AFR 190-17 JERRY C. NIX, Captain, USAF Director of Information, AFIT		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) ATTACHED Reproduced by NATIONAL TECHNICAL INFORMATION SERVICE U S Department of Commerce Springfield VA 22151		

FORECASTING TRANSPORTATION TONNAGE

BY AN ADAPTIVE SMOOTHING MODEL

FRANCIS E. JAMES, JR., Colonel, USAF

AU-AFIT-SL-10-73

This paper describes a research project which was designed to investigate the feasibility of utilizing mathematical models to forecast cargo air shipments. Seven mathematical models were evaluated. In a two phase research effort, actual movement through eight routes over forty-one months was first evaluated according to six different criteria for six forecasting models. Parameters were varied for the models to approach optimality. Using results from the first phase, three models were re-evaluated through thirty-nine routes for thirty-six months each. An adaptive exponential model was identified as being most suitable for the contemplated usage.

1. Objective. The objective of this paper is to report and analyze results of recent research studies in the cargo transportation area. This research project, headed by Col Francis E. James, Jr., was conducted from September 1970 to July 1971 and from August 1973 to present. Two Air Force officers, Captain Owen Barrett and Captain John R. Phillip, assisted in the study from September 1970 through February 1971, in conjunction with a thesis to satisfy the requirements for a Master of Science degree in Logistics Management. Preliminary findings of the research study have been reported in the Barrett/Phillip thesis, "A Comparative Study of Models for Forecasting MAC CONUS Outbound Airlift Requirements" (SLSR 71-247), and in the paper, "Forecasting Models for Cargo Transportation Requirements," presented at the NATO Conference on 14-18 August 1972, in Sandefjord, Norway. A United States Navy officer, Lt Lynn M. Lynch, developed a number of computer programs to assist the Air Force Logistics Command from September through December 1973.

2. Background of the Problem. Within the Department of Defense, the Military Airlift Command (MAC) is charged with the responsibility of providing, on a reimbursable basis, the air transportation space required for the strategic international airlift of both passengers and cargo. There are many elements of similarity between MAC and a number of civilian airlines. For example, almost 90 percent of the MAC airlift responsibility is accomplished by "channel traffic". "Channel traffic" is the movement of personnel and cargo over established worldwide routes, serviced by either scheduled military aircraft or commercial aircraft under contract to and scheduled by MAC. The scheduling function within MAC cannot operate efficiently unless accurate forecasts are available.

Forecasts of airlift requirements are submitted to MAC by the Departments of Army, Navy, and Air Force, 80 days prior to a given month of airlift operation. During FY 1969, the forecasts submitted to MAC contained an overestimate of 215,317 passengers and an underestimate of 41,426 short tons of cargo. Since approximately 93 percent of the DoD passengers were moved by commercial contract capability, this overestimate, based on a 219 seat DC8-61/63 aircraft, would have equated to 983 missions being scheduled that were not required. Similarly, the underestimate in cargo would have equated to an unanticipated 2,071 additional C-141 missions at the average planned 20 ton payload. These figures are for an entire year of airlift operation and some rescheduling actions could have been accomplished, but the ultimate results were increased DoD transportation costs and reduced airlift responsiveness. Throughout all branches of the service, channel traffic forecasts are generally based upon current information, historical data, judgment, and experience on the part of the forecaster - little use has been made of clearly defined quantitative models.

3. Overall Objectives of the Research Study. The objective of this research study was to evaluate management science techniques which might improve the accuracy of cargo forecasts. Both passenger and cargo forecasting procedures are clearly worthy of further investigation; however, time constraints dictated that the present research be limited to the cargo area.

4. Earlier Reported Results.

a. Records were available which identified the monthly USAF cargo airlift forecast and the actual cargo shipment for each of approximately 120 Continental United States (CONUS) outbound channels from June 1957.

To evaluate the feasibility of management science models, eight channels were selected for preliminary study. The eight channels were selected because they were representative of the more important, high volume channels, and because they provided a good geographical representation. Forty-one monthly data points (June 67-Oct 70) were utilized for each of the eight channels.

b. In a preliminary study by Captains Barrett and Phillip, six mathematical models were evaluated over the monthly data points for each of the channels. The models utilized in this preliminary evaluation were the unweighted moving averages, the least squares model, the semi-average model, and modified single, double, and triple smoothed exponential models.

c. The preliminary study evaluated effectiveness relative to a number of variables, including lead time, number of months in the data base, and exponential smoothing constants. Models were evaluated according to the following criteria: average absolute error, average percent absolute error, average standard deviation, percent times best forecast, percent times worst forecast, and a weighting index. Enclosure 1 depicts results for this preliminary phase of the study.

d. The significant findings of this early study were that four of the six models which had been examined for the three months lead period yielded more accurate forecasts, on the average, than did the actual forecasts for the time. It should be noted that the actual forecasts were developed with the historical data utilized in this study and that, in addition, the forecaster had access to projections of troop movements,

military operations, and other scheduled activities. It appears reasonable to conclude that the quantitative management models possess the potential for assisting the forecaster to improve his cargo forecasting accuracy.

5. The Present Study. Since earlier results had indicated that mathematical models held the potential for improving forecasting accuracy, an extended study appeared warranted. The objective of this extended study was to determine which of the models appeared to hold the greatest potential for improved forecasting, and to indicate appropriate parameters of the model. The following paragraphs describe the extended study in detail.

a. It was apparent that a study with objectives as stated above should be based upon the broadest and most accurate data base. The original study examined 8 channels for June 1967 to October 1970. A great deal of effort was expended by the Transportation Operations Division, Headquarters Air Force Logistics Command, in collecting additional data. As a result, this study is based upon cargo generation figures for 44 different channels over a January 1968 - December 1970 time period. The 44 channels were selected as representative of the primary heavy volume channels that had been and were to be used. After the data had been collected, it was determined that some of the channels were not appropriate for this study, because data for one or more of the time periods were not present. Five such channels were deleted. As a result, the final study was based upon 39 channels, over 36 months, for a total of 1,404 data points. The 39 channels are listed in enclosure 2.

b. The original feasibility study had determined that the average absolute AFLC forecasting error, for the channels examined, was 55.1 tons per month. The modified triple exponential smoothed model yielded the greatest improvement in this figure, with an average absolute monthly error of 38.2 tons per month. This model had utilized a smoothing constant of $\alpha = .45$. This model produces forecasts in accordance with the following equations:

$A(t)$ = the parameter "a", calculated for the "t'th" time period

$B(t)$ = the parameter "B", calculated for the time period "t"

$C(t)$ = the parameter "C", calculated for the time period "t"

$X(t)$ = the actual (observed) cargo generated for time period "t"

$S(t)$ = the forecast cargo generation for time period "t"

a = the smoothing constant: $0.0 \leq a \leq 1.0$

b = $1.0 - a$.

$A(t) = X(t) + b^3 [S(t-1) - X(t)]$

$B(t) = B(t-1) + C(t-1) - 1.5a^2(2 - a) [S(t-1) - X(t)]$

$C(t) = C(t-1) - a^3 [S(t-1) - X(t)]$

$S(t) = A(t) + B(t) + .5C(t)$.

Since the initial survey revealed that the accuracy of the model improved as the number of previous month's data available expanded, the expanded study utilized months one through twenty for a data base. After updating the A, B, and C coefficients (described above) for each month through the 20th month, a forecast for the fourth month hence was made at that time. This forecast, for the 24th time period, was then compared with the actual cargo generated at that time period, and the absolute error was calculated. This procedure continued, with forecasts for the 25th, 26th

and succeeding months, until the forecast for the 36th month had been completed and compared to the actual figure. For each channel, 13 forecasts were made. This procedure was followed for each of the 39 channels, yielding a total of 507 forecasts. The initial model used that smoothing constant determined optimal by the initial feasibility study, 0.45. The 507 forecasts yielded an aggregate absolute error of 24,757.95 tons, or an average error of 48.832 tons per forecast. In order to determine if a more optimum smoothing constant were available which would improve accuracy, the experiment was replicated with other values of the "a" parameter. Results are depicted below.

Table One

TRIPLE EXPONENTIAL SMOOTHING RESULTS*

Value of "a" Smoothing Constant	Average Absolute Error Per Forecast
0.25	101.127
0.30	54.640
0.35	47.404
0.40	47.484
0.45	48.832
0.50	50.628
0.55	52.794

*The 39 channels utilized for this analysis are listed in enclosure two. The initial forecast was made after the 20th month's data was used to update the three parameters of the model, and forecast the 24th month. A total of 507 forecasts were evaluated.

It appears reasonable to note that this extended study, which utilized a much broader data base than the original feasibility study, confirms that the modified triple exponential smoothing model "a" parameter of 0.45 is near optimum (the 47.404 Average Absolute error figure for an "a" of 0.35 represents less than 3 percent improvement). It also confirms that the Triple Exponential Smoothing model is a useful forecasting tool. However, the reduction of forecasting error for this broader data base was not as significant as in the earlier feasibility study.

c. The triple exponential smoothing model described above is effective for forecasting a trend line when requirements are anticipated to be changing as a quadratic function of time. According to the criterion outlined above, there appeared to be reasonable grounds for examining other models, such as the single exponential smoothing model. It is a simple, easily calculated model which can be useful in establishing requirements where relatively stable demand patterns exist. Equations utilized in the model are as follows:

$X(t)$ = the actual (observed) cargo generated for time period "t"

$S(t)$ = the forecast cargo generation for time period "t"

a = the smoothing constant: $0.0 \leq a \leq 1.0$

$S(t) = S(t-1) + a[X(t) - S(t-1)]$

In order to test the effectiveness of this model, a number of smoothing constant ("a") values were utilized. Results of two of the most effective constant values are depicted in Table Two.

Table Two

SINGLE EXPONENTIAL SMOOTHING RESULTS*

Value of "a" Smoothing Constant	Average Absolute Error
.10	50.852
.20	42.203

*The 39 channels utilized for this analysis are listed in enclosure two. The initial forecast was made after the requirements figure for the 20th month was used to update the model, and was a forecast for the 24th month. A total of 507 forecasts were evaluated for each value of the smoothing constant.

d. The exponential smoothing model possesses many advantages, as a forecasting tool. It is easy to calculate, no extensive historical data base need be retained, and current information to update the model can be quickly introduced. In order to filter out the major part of the noise in the input, it is customary to use a value of 0.20 or less for the "a" parameter.** However, the forecaster who settles upon a fixed value of "a" encounters certain difficulties. For small values of "a", relatively heavy emphasis is placed on demand of the distant past, and relatively little weight is placed on recent figures; the reverse situation is encountered when a high value of "a" is utilized. With low values of "a", the forecasting system is relatively unaffected by sudden changes in the underlying process (as, for example, cargo movement dictated by tactical decisions). In this case, the forecasting system will take a very long time to home in on the new level, and biased forecasts will continue to occur for some time. Tracking signals,

**See D. W. Trigg and A. G. Leach, "Exponential Smoothing With an Adaptive Response Rate", Operational Research, Quarterly Vol. 18, No. 1., p. 53.

which alert the forecaster to a change in the underlying distribution, have been recommended. Although they are useful in identifying changes in the parameters of the distribution, they suffer from the disadvantage that they merely alert the user to the fact that a change has occurred. The user still faces the problem of determining reasonable forecasts, based on the new (changed) conditions.

e. To overcome this problem, some authorities have recommended that exponential smoothing models with adjustable smoothing constants be utilized. Where the demand figure shows a relatively stable pattern, the model would use a low "a" value, so as to filter out random fluctuations. When demand figures indicate that basic changes are occurring in the parameters of the underlying distribution (as, for instance, where strategic considerations dictate a large increase in logistics support in a theater) then the model would use a high "a" figure, so as to be more immediately responsive to the change in demand. After more stable demand has been reestablished (perhaps, at the higher level) the model reverts to a lower "a" value. Various authors have proposed schemes which would accomplish this purpose, developing a self adaptive smoothing model. The model tested in this study was that proposed by Trigg and Leach.* For this model, the following equations were utilized:

$X(t)$ = the actual (observed) cargo generated for time period "t"

$S(t)$ = the forecast cargo generation for time period "t"

a = the smoothing constant: $0.0 \leq a \leq 1.0$

\emptyset = the adaptive factor

*Trigg and Leach, op.cit., pp. 53-59.

$$SE(t) = (1.0 - \phi) SE(t-1) + \phi [X(t) - S(t-1)]$$

$$SA(t) = (1.0 - \phi) SA(t-1) + \phi |X(t) - S(t-1)|$$

$$a = |SE(t)/SA(t)|$$

$$S(t) = S(t-1) + a [X(t) - S(t-1)]$$

It is obvious that the value of "a" will be nonnegative, and less than or equal to unity. Also, it is apparent that the initial values calculated according to these equations will probably yield a high "a" value and the system will then home in on current demand figures with considerable rapidity. Initial demand values are often non-representative of the typical demand pattern; if this turns out to be true, the value of "a" will be large, and the system will continue to home in on current demand figures with considerable rapidity.

f. Previous work in this area has mostly centered around the use of ϕ values of .05 or .10. In order to test the efficiency of these, and other reasonable values of ϕ , a series of experiments were conducted. Results are depicted in Table Three.

Table Three

ADAPTIVE EXPONENTIAL SMOOTHING RESULTS*

Value of ϕ Constant	Average Absolute Error
.05	39.306
.10	37.500
.11	37.458
.12	37.513
.13	37.639
.14	37.831
.15	38.047

*The 39 channels utilized for this analysis are listed in enclosure two. The initial forecast was made after the requirements figure for the 20th month was used to update the parameter of the model, and was a forecast for the 24th month. A total of 507 forecasts were evaluated for each value of the smoothing constant.

For these data, a .11 value of θ was found to be optimal; however, it should be noted that the range of average absolute error figures, from best to worst, was only 1.848, less than 3 percent of the midvalue. Apparently, these demand patterns are relatively insensitive to changes in the value of the θ parameter. It should also be noted that the best (smallest absolute error) value of the adaptive exponential smoothing model (37.458) represented a considerable improvement (21%) over the optimum triple exponential smoothing model, and was also appreciably smaller (11.2%) than the optimum single exponential smoothed model.

g. The previous study indicated the desirability of a reasonably large historical data base to develop accurate parameters and achieve accurate predictions. Enclosure 1 illustrates this point. For example, using the four month lead time figures, the average absolute error of the triple exponential smoothing model increases 40 percent when one utilizes a 12 month data base in preference to the 20 month base (from 38.2 tons to 53.3 tons), and the average error increased 173 percent when one utilizes an eight, rather than a 20 month data base (from 38.2 to 104.2). For the triple exponential smoothing model, it appears clear that a very large data base (that is, extensive records of previous demand figures) is highly desirable. This is, of course, somewhat of a handicap, inasmuch as circumstances may not always permit the accumulation of this data prior to the time the forecast is desired. Two additional experiments were performed in order to test the effectiveness of the adaptive exponential smoothing model under more restrictive conditions, when such historical data might not be available. For both of these experiments, a θ value of .11 was employed. Results are as depicted in Table Four.

Table Four

ADAPTIVE EXPONENTIAL SMOOTHING RESULTS WITH LIMITED HISTORICAL DATA*

Initial Data Base	Forecasts Made For Each Channel	Total Number of Forecasts Made	Average Absolute Error
2**	31	1209	44.336
9***	24	936	39.360

*The 39 channels utilized for this analysis are listed in enclosure two.

**Using data for the first two months to update parameters of the model, the initial forecast was made for month "6", four periods in the future.

***Using data for the first nine months to update parameters of the model, the initial forecast was made for month "13", four periods in the future.

For these data, a diminution of original data points from 20 to 9 months reduced the accuracy of the forecast only 5.1 percent, on the average. However, when the data base was reduced from 20 to 2 months, the accuracy of the forecast was reduced by 18.4 percent. It appears reasonable to conclude that the adaptive features of this model are sufficient to overcome paucity of data, up to a reasonable point. With only 9 months of historical data available, this model continued to perform in a useful fashion. A reduction of the data base to only 2 months data is certainly not recommended but, here also, the model gave a reasonably good account of itself from the standpoint of forecasting accuracy.

6. Using the Model. Although this study is primarily designed to report experimental results, it may be appropriate to include recommendations for eventual use of this and other forecasting models. Quantitative

models can provide useful assistance to the forecaster--they can never replace a highly motivated individual, knowledgeable about current and anticipated materiel and troop movements, and experienced in the forecasting problem. The best use of the model is to make an initial forecast. This initial forecast can be easily and quickly made, across a large number of different channels, by computer technology. The forecaster should then be tasked to examine each forecast. Where appropriate, the forecaster would modify the computer output, based upon his background knowledge, expertise, and current subjective considerations. In this fashion, the management science mathematical model would work in harmony with the human element.

ENCLOSURE ONE

RESULTS OF PRELIMINARY STUDY

FIGURE 4

AVERAGE INDEX WEIGHT BY METHOD

Forecast Models	3 Month Lead Time					4 Month Lead Time				
	Months In Base Period					Months In Base Period				
	6	8	12	16	20	6	8	12	16	20
Single Exponential Smoothing	3.07	3.00	3.31	3.51	3.57	3.22	3.16	3.57	3.74	3.81
Double Exponential Smoothing	3.42	3.44	3.35	3.61	3.72	3.37	3.58	3.61	3.69	3.88
Triple Exponential Smoothing	4.81	4.58	3.98	3.40	3.31	4.86	4.58	3.75	3.28	3.22
Regression Least Squares	3.35	3.47	3.59	3.49	3.52	3.51	3.63	3.64	5.49	3.70
Semi Average	6.17	6.10	5.74	5.01	4.78	6.28	6.05	5.58	5.15	4.54
Moving Average	3.54	3.75	4.20	4.94	5.06	3.51	3.71	4.27	4.91	5.11
Air Force Logistics Command	3.64	3.65	3.83	4.03	4.03	3.25	3.31	3.28	8.73	3.73

FIGURE 5
PERCENT TIMES BEST BY METHOD

Forecast Models	3 Month Lead Time					4 Month Lead Time				
	Months In Base Period					Months In Base Period				
	6	8	12	16	20	6	8	12	16	20
Single Exponential Smoothing	.10	.15	.11	.11	.11	.07	.10	.10	.08	.07
Double Exponential Smoothing	.19	.14	.17	.12	.11	.17	.13	.13	.15	.12
Triple Exponential Smoothing	.14	.18	.24	.29	.33	.17	.19	.28	.30	.31
Regression Least Squares	.17	.16	.14	.13	.15	.16	.14	.14	.13	.12
Semi Average	.04	.08	.08	.17	.15	.04	.11	.10	.12	.17
Moving Average	.15	.13	.13	.08	.08	.18	.13	.10	.13	.09
Air Force Logistics Command	.21	.17	.13	.11	.09	.21	.20	.15	.10	.12

FIGURE 6

PERCENT TIMES WORST BY METHOD

Forecast Models	3 Month Lead Time					4 Month Lead Time				
	Months In Base Period					Months In Base Period				
	6	8	12	16	20	6	8	12	16	20
Single Exponential Smoothing	.01	.00	.01	.01	.00	.01	.00	.01	.00	.00
Double Exponential Smoothing	.01	.00	.02	.06	.10	.00	.02	.04	.06	.10
Triple Exponential Smoothing	.22	.19	.13	.08	.10	.23	.20	.12	.08	.10
Regression Least Squares	.00	.00	.00	.01	.02	.00	.00	.00	.01	.04
Semi Average	.72	.74	.69	.56	.44	.72	.71	.67	.52	.38
Moving Average	.01	.04	.10	.23	.28	.02	.05	.10	.27	.34
Air Force Logistics Command	.04	.04	.06	.06	.06	.02	.02	.06	.05	.04

FIGURE 7

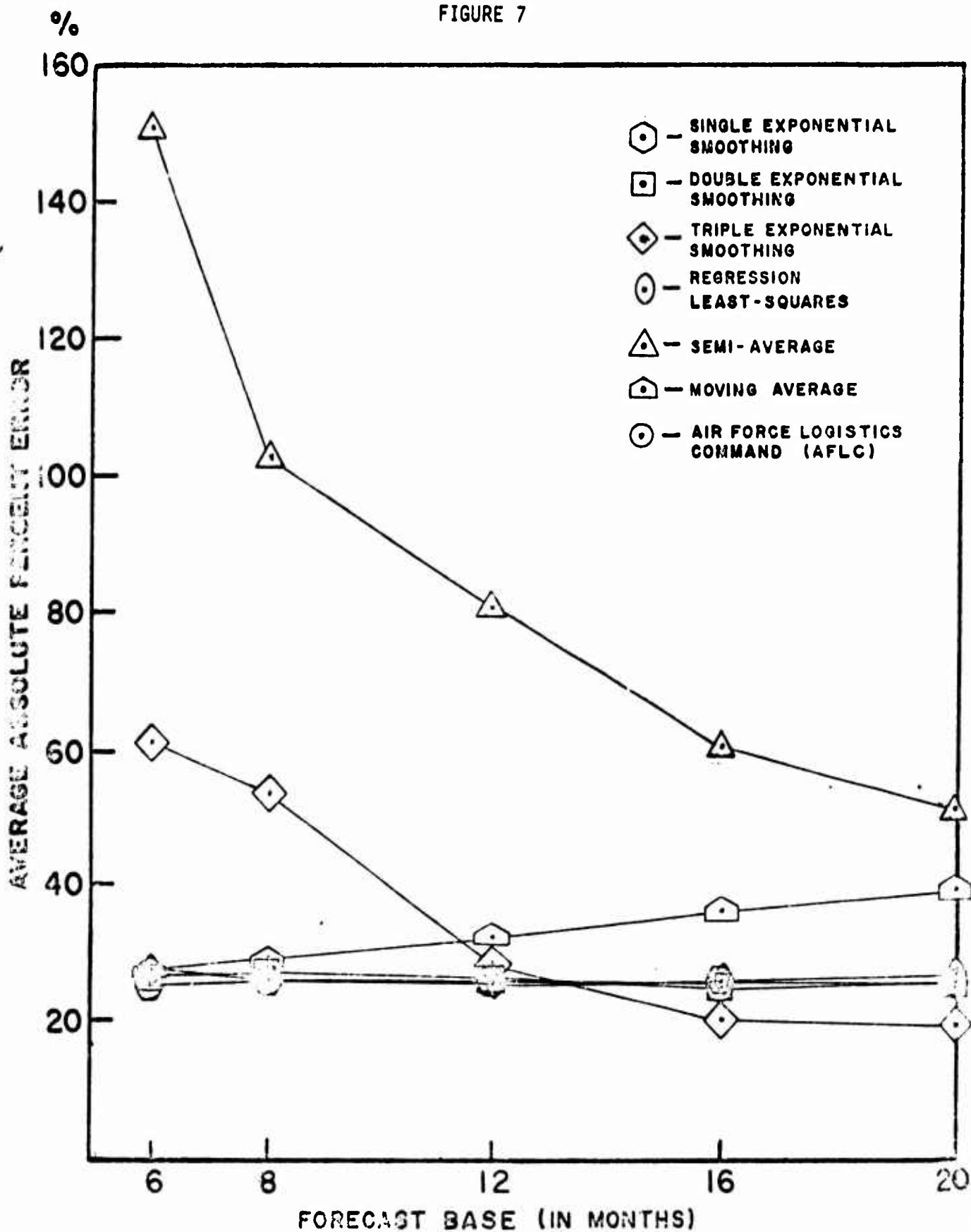


FIGURE 1. Comparison of the Average Absolute Per Cent Error for the AFLC Forecasting and the Statistical Methods with a 4-Month Lead Time

ENCLOSURE TWO

MAC OUTBOUND CHANNELS USED IN RESEARCH ANALYSIS
(January 1968 - December 1970)

The following channels were identified for the research project:

Charleston - Ramsey
Charleston - Kadena
Charleston - Lajes
Charleston - U-Tapao-Bangkok
Dover - Incirlik
Dover - Clark
*Dover - Bien Hoa
Dover - Mildenhall
Dover - Saigon
Dover - Torrejon
Dover - Cam Ranh Bay
*McGuire - Danang
McGuire - Keflavik
McGuire - Sondestrom
McGuire - Thule
McGuire - Goose Bay
San Bernardino - Danang
San Bernardino - Kadena
Kelly - Clark
Kelly - Danang
Kelly - Kadena
Kelly - Yokota
Kelly - Saigon
Kelly - Cam Ranh Bay
Travis - Udorn
Travis - Clark
Travis - Danang
*Travis - Bien Hoa
Travis - Hickam
Travis - Korat - Ubon
Travis - Saigon
Travis - Guam
Travis - U-Tapao - Bangkok
Travis - Cam Ranh Bay
McChord - Elmendorf
McChord - Eielson
McChord - Yokota
*McChord - Osan
McChord - Seoul
McChord - Ching Chan Kung
Tinker - Udorn
Tinker - Ubon - Korat
Tinker - Guam
*Tinker - Utapao - Bangkok

*Complete data was not available for these channels during the period of the report and, for this reason, these channels were not used in the study.