THE DEVELOPMENT AND EVALUATION OF A FORECASTING SYSTEM THAT INCORPORATES ARIMA MODELING WITH AUTOREGRESSION AND EXPONENTIAL SMOOTHING

DISSERTATION

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

by

Laurette Poulos Simmons, B.A., M.B.A.

Denton, Texas

May, 1985

Simmons, Laurette Poulos, The Development and Evaluation of a Forecasting System that Incorporates ARIMA Modeling with Autoregression and Exponential Smoothing. Doctor of Philosophy (Management Science), May, 1985, 108 pp., 10 tables, 9 figures, bibliography, 71 titles.

This research was designed to develop and evaluate an automated alternative to the Box-Jenkins method of forecasting. The study involved two major phases. The first phase was the formulation of an automated ARIMA method; the second was the combination of forecasts from the automated ARIMA with forecasts from two other automated methods, the Holt-Winters method and the Stepwise Autoregressive method.

The development of the automated ARIMA, based on a decision criterion suggested by Akaike, borrows heavily from the work of Ang, Chuaa and Fatema. Seasonality and small data set handling were some of the modifications made to the original method to make it suitable for use with a broad range of time series. Forecasts were combined by means of both the simple average and a weighted averaging scheme.

Empirical and generated data were employed to perform the forecasting evaluation. The 111 sets of empirical data came from the M-Competition. The twenty-one sets of generated data arose from ARIMA models that Box, Taio and Pack analyzed using the Box-Jenkins method.

To compare the forecasting abilities of the Box-Jenkins and the automated ARIMA alone and in combination with the other two methods, two accuracy measures were used. These

measures, which are free of magnitude bias, are the mean absolute percentage error (MAPE) and the median absolute percentage error (Md APE).

The results of the study indicate that the automated ARIMA forecasts compare very well with those from the Box-Jenkins method on the generated data. On the empirical data, the results favor the automated ARIMA in combination with the Holt-Winters. In fact, for one-step-ahead forecasting, this combination provides a Md APE that is 2.6 percent lower than the Box-Jenkins results for that time horizon. This result on one-step-ahead forcasting is very significant since ARIMA modelling is designed for exactly this type of forecasting.

The implication of this research for the practitioner is that it presents an alternative method for use when the accuracy of the Box-Jenkins is desired, but the resources to perform such lengthly analyses are not available. In such a case the practitioner could use the automated ARIMA method if the data are known to have arisen through an ARIMA process; otherwise the practitioner could use the automated ARIMA in weighted combination with the Holt- Winters method.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

The development of accurate forecasts has long been a goal of organizations in both the public and private sectors. Many forecasting methods are available, from the naive model to very complex mathematical formulations. Varying levels of accuracy are realized from the forecasting techniques. One complex univariate approach which has received a great deal of attention, in the past decade, for its ability to provide accurate forecasts is attributed to Box and Jenkins (2).

The Box-Jenkins method provides a great deal of flexibility in model selection because it chooses a particular time-series model from a class of autoregressive, moving-average or mixed autoregressive moving-average models. These models are capable of representing both seasonal and nonseasonal time series (1).

Although the Box-Jenkins method has become well-known, critics of this method cite the need for expert judgement in every application as one severe deterrent to its wide acceptance for practical use. Other obstacles which contribute to the difficulty in using the Box-Jenkins method include ambiguity in the model identication process and a high time requirement for each application. These problems have lead to the seeming paradox of a technique that provides the accuracy desired by business while being a method which is not widely employed in that arena.

Statement of the Problem

The Box-Jenkins method is not being widely used by the business practioner to produce forecasts. Practioners in business require a forecasting system which can provide the accuracy of the Box-Jenkins method without the disadvantages of this complex technique. A system which could exploit the increasingly available computer resources to automatically produce accurate forecasts without expert intervention would provide an attractive alternative to the existing subjective Box-Jenkins method.

CHAPTER BIBLIOGRAPHY

1. Bowerman, Bruce L. and Richard T. O'Connell, Forecasting and Time Series, Belmont, California, Duxbury Press, 1979.

2. Box, George E. P. and Gwilym M. Jenkins, Time Series Analysis: Forecasting and Control, San Francisco, Holden-Day, 1976.

CHAPTER II

LITERATURE REVIEW

Authors such as Granger, Bowerman and O'Connell, Makridakis, Wheelwright and McGee, and Hanke and Reitsch have discussed the importance of forecasting in business. According to Makridakis, Wheelwright and McGee (32, p. 4), ". . . forecasting is an integral part of the decision-making activities of management." A forecasting text by Bowerman and O'Connell (11, p. 3) substantiates the sentiment with this statement, ". . . business firms require forecasts of many events and conditions in all phases of their operations. Granger states, "Many million dollars are spent annually on predition in the United States alone, making forecasting big business" (18, p.3). "Forecasts are needed in finance, marketing, personnel and production areas . . . " (21, p. 2).

A wide range of techniques is available to assist the forecaster in providing the desired information to decision-makers. Techniques range from the naive, intuitive approach to complicated computer-assisted algorithms. There are various types of qualitative methods, as well as both univariate and multivariate quantitative techniques. In selecting a category or specific method, consideration must be given to factors such as accuracy, availability and ease-of-use.

According to Makdridakis et al. (34), accuracy is an extremely important aspect in the realm of forecasting. "In many situations even small improvements in forecasting accuracy can provide considerable savings" (34, p. 112). In the interest of comparing the accuracy of qualitative (judgemental) methods compared with quantitative methods, empirical studies have been performed. A comprehensive survey of research into this issue was performed by Makridakis and Hibon (31). The majority of references cited in this work favor quantitative methods. Mabert (27) reports that forecasts based on the opinion of the sales force and corporate executives provided less accurate results compared with three quantitative methods. He also concludes that the quantitative methods in the study cost less money and requires less time than did the qualitative techniques. Adam and Ebert in a subsequent study, found that a quantitative method produced forecasts that are statistically better than those of human forecasters (1). Mahmoud concurs, saying, "On the whole, past research suggests that quantitative methods outperform qualitative methods" (29, p. 153).

The two major classifications of quantitative methods are time series and causal. In selecting between these classifications, the ease-of-use factor often favors the former. The time-series type is a univariate approach in which forecasts are based on past values, past errors, or both. The objective of this type of forecasting method is to discover the pattern in the historical data and extrapolate that pattern into the future. Causal methods, also called explanatory or structural, are based on the assumption that there is a cause-effect relationship between the variable of interest and one or more independent variables. The objective in this case is to discover the form of this multivariate relationship and use it to forecast future values of the dependent variable based on values of the independent variable(s).

The disadvantages of the causal models include increased complexity, estimation and cost. According to one paper, "Time-series models can often be used more easily (than causal methods) to forecast . . . "(32, p.10). Another work contains the comment, "Explanatory models, by their nature, require a number of independent variables whose magnitude, for some time, must be estimated before any predictions about the future can be made (30, p. 62). Concurrence is provided by Newbold and Granger's widely cited paper ". . . relevant extraneous information may be unavailable or only obtainable at a prohibitively high cost" (38, p. 131).

In comparing the accuracy of univariate forecasting methods to multivariate econometric models, one particular work by Armstrong (9) provides some interesting insights into the difference between the perceived and actual accuracy of multivariate versus univariate methods. Of 21 experts in the survey, 95 percent are of the opinion that econometric methods provide higher accuracy than time-series methods. In the empirical comparison, however, this opinion is not substantiated. According to Armstrong, "There was no tendency toward greater accuracy" of the econometric over the time-series methods (19, p. 552). Chatfield and Prothero postulate what may be an explanation of the results reported by Armstrong and other researchers. They state that econometricians may be fitting linear regression models in which error terms are erroneously assumed to be independent, while some time-series methods, on the other hand, use the existing autocorrelation to provide better forecasts (14).

The Box-Jenkins Method

Among univariate methods, there is little doubt that the Box-Jenkins method has received more attention in the past fifteen years than any other approach. There is copious literature addressing the Box-Jenkins method. Unfortunately, the existing opinions and empirical evidence do not provide unanimity regarding the desirability of this approach for practical use. Some works praise the approach as the most accurate of the time-series methods, while many criticize it as a difficult, subjective and time-consuming technique.

A frequently encountered opinion, which is favorable toward the Box-Jenkins method, is succinctly stated by Ang et al., "Possibly the most important time series forecasting model in use is the autoregressive-moving average model popularized by Box and Jenkins" (8, p. 38). Carbone et al. (13) cites the Box-Jenkins technique as one which is well documented, popular and widely used. According to McKenzie (35, p. 114), "There can be little doubt that in the present context at least, the Box-Jenkins approach to forecasting is by far the best . . .." Another author provides strong affirmation for the value of this method by saying, "The relevantly recent research work of Box and Jenkins is now proving itself such an effective tool, for letting the data speak for themselves that no provider of forecasts can afford to ignore it . . ." (23, p. 413). Reid concurs, stating, "The Box-Jenkins Method is clearly the best individual technique from the empirical studies and there are good theoretical reasons why this should be expected . . ." (39, p. 24).

An empirical study, by Newbold and Granger, comparing three methods, including the Box-Jenkins, praises the results of the technique but also mentions disadvantages inherent in the method. The author states,

> . . . Box-Jenkins forecasts require a good deal more time and considerably more skill to compute than do their competitors. However, we have found that there is a corresponding pay-off in the sense that the Box-Jenkins forecasts do seem to be better than those derived from two fully automatic procedures - the Holt-Winters method and stepwise autoregression - for a sizeable majority of the time series in our sample (38, p. 143).

Authors including Chatfield (15), Chatfield and Prothero (14), Herne (22), Makridakis (34), Hill and Woodworth (23), Stern (40), Kenny and Durbin (25), and Newbold and Granger (38) have provided strong criticisms as to the value of the Box-Jenkins method for practical use. Among those who complain that it is a difficult, time-consuming technique are Chatfield and Prothero, Herne, Chatfield and Makridakis et al. Another frequently mentioned criticism is the requirement of an expert to perform the analysis. Hill and Woodworth, Stern and Kenny and Durbin are only a few of the large group of respected researchers who list this as a severe methodological limitation.

The subjectivity and ambiguity of the Box-Jenkins method are frequently cited drawbacks to the practical applicability of this technique. The problem is summarized in this classic statement, "It is by no means certain that any two analysts applying the principles of Box and Jenkins to a particular set of data will reach the same conclusions" (38, p. 133). The ambiguity posed by the selection criteria of the Box-Jenkins method is discussed by Andersen. He states,

> For seasonal data, there are a large number of almost identical theoretical autocorrelation functions from substantially different models. With the usual number of observations, it is very difficult and often quite time consuming to distinguish between the competing models (6, p. 476).

A study was performed by Guerts and Reinmuth (20) to determine whether forecasting accuracy is sensitive to model selection. The results indicate that even when a model is selected that satisfies the criteria as supplied by Box and Jenkins, it can yield a significantly lower level of forecasting accuracy than a competing model.

The majority of criticisms of the Box-Jenkins methodology are summarized by Chatfield and Prothero in their response to general reviewers of a paper read at a meeting of the Royal Statistical Society, "It certainly only needs two examples to see that the method (Box-Jenkins) is complicated, expensive, that subjective assessments have to be made and that several models can be found which fit the data equally well" (14, p. 333). A practitioner, Tomasek, states in his review of the paper, that his company uses the Box-Jenkins method successfully and that he personally has been responsible for conducting a seminar on ARIMA model building for 85 other employees (41). Chatfield and Prothero respond to this statement by adding, "To give the other side of the story we know of several firms in Britain who have tried Box-Jenkins methods and have now gone back to traditional trend and seasonal models." (14, p. 335). They add a cryptic statement that said, in effect, that Tomasek should use a simpler method, which, would provide approximately the same accuracy but would not require him to spend his time providing the lengthly training required by the complex method of Box and Jenkins.

Two studies, which seek to assess the status and success of forecasting in business organizations, conclude that the Box-Jenkins method is not widely accepted in that arena. In the Wheelwright and Clark study the Box-Jenkins had the lowest on-going use, by those who had tried it, of any of the eight methods examined. The main reason cited by the practioners surveyed for this discontinuation of use is complexity (43). The report of the Mentzer and Cox study, which involved managers in a sales environment, contains the following comment as a conclusion, "Dissatisfaction with Box-Jenkins time-series analysis occurred across all management and forecast groups" (36, p. 28).

The Newbold and Granger paper of 1974 is accompanied by a set of reviews and subsequent rebuttal by the authors. Statements made by Newbold and Granger and two of the reviewers from the business community provide insights into the reported underutilization of the Box-Jenkins method in industry. "It is frequently the case that, for reasons of economy of time or effort, a Box-Jenkins analysis is impracticable, and some fully automatic procedure must by employed" (38, p. 143). Bransom, a reviewer of the paper and practitioner from the National Economic Development Office (in England), states:

> I have had a fair amount of experience in work on stock control and production control. When you are doing stock control and production control you do need short-term forecasts, a matter of a few months or up to a year ahead for each of the products which you are dealing with, and many firms in this country in fact produce a very wide range of different products, amounting in some cases to thousands. I suspect that not even Dr. Newbold would be prepared to run thousands of Box-Jenkins analyses, even only once a year.It is in these circumstances that one badly needs a reliable automatic method of forecasting to put into the computer and churn out the regular forecasts that you need to do a reasonable job of stock control and production control (12, p. 157).

Ward, another reviewer of the Newbold and Granger paper and an employee of Birds Eye Foods, endorses Mr. Bramson's comment adding that even when only several hundred products are involved, "you cannot intervene into each one individually" (12, p. 157).

A paper by Hill and Woodworth echoes the sentiments described with the following observation:

> Despite all the theoretical and empirical evidence of the advantages of Box-Jenkins forecasts, it is still only very slowly making an impact on the mass of people whose job it is to produce forecasts. In the authors' opinion the main reason for this must be the lack of knowledgeable people to produce and interpret the models and also the relatively high cost of the modelling procedure . . . (23, p. 413).

The authors go on to say that there is an obvious need for an automated Box-Jenkins package.

Several automated Box-Jenkins methods have appeared. Some of the programs support model selection on the basis of statistical tests, while others use non-statistically based decision criteria. These methods also vary in their levels of complexity, their handling of seasonality and their flexibility in providing data transformation.

A procedure by Dagum (16), designed to handle both automatic and "manual" Box-Jenkins analysis, provides an automated option in which each of three particularly common Box-Jenkins models is fit to the data. Any of these models that passes the test for random residuals is used for forecasting for one year ahead. If more than one passes, selection among them is made on the basis of minimum average forecasting error. In the event that none of the three models is selected, the opportunity to suggest an alternative is provided to the user.

A more sophisticated version of the Dagum process, called CAPRI (Construction Automatique de Previsions Regies sans Intervention) was developed as a dissertation and presented in a paper (26). The method relies on an initial transformation process. Differencing is performed to obtain stationarity as well as seasonal adjustment. The key similarity between Libert's method and that of Dagum is that in each method, a pre-determined set of available models is fit and statistical criteria are used to select among them. The number of models used in the Libert method is unclear, but it seems to be more than three. The chief difference between the methods is that in CAPRI, "If the parameters are not significant or the autocorrelations are not random another model from the list is chosen, estimated and tested. If the program cannot generate an appropriate model an autoregressive process is imposed" (27, p. 326).

Model selection decision criteria, determined by Akaike (2, 3, 4, 5), form the basis of several other automated

Box-Jenkins methods. Hill and Woodworth (23) use the criterion published in 1971, in a process called SIFT (Systematic Identification and Forecasting of Time Series). Other methods use Akaike's newer AIC (Automatic Identification Criteria). Kang (24) devoted his dissertation to the development and validation of one such procedure, ARIMAID (Auto Regressive Integrated Moving Average IDentification). This procedure identifies certain types of seasonal Box-Jenkins models, as well as non-seasonal models. It currently does not provide subroutines to forecast once the optimal model has been identified, nor does it deal with additive seasonal models.

Another member of the family of automated Box-Jenkins methods uses standard Box-Jenkins programs, available through popular statistical analysis packages, with the inclusion of several subprocedures (18). It has the obvious advantage of being accessible to practitioners who may already have such programs available. The Ang procedure uses Akaike's AIC decision criterion to select from competing models determined on the basis of maximum required autoregressive order. Differencing for stationarity is handled by repeating the procedure for various differenced series and comparing the best ARMA model from each run on the basis of minimum AIC. Once the model is selected, forecasting can be performed through the standard Box-Jenkins programs.

A literature search has found little objective research into the effectiveness of these automated methods and no comparative studies to aid the user in selection of one over another. The Ang procedure has some obvious advantages, including use of the most popular and widely studied Akaike criterion (37) as well as its relative availability. The principal disadvantage is the lack of appropriate means to deal with seasonality. Another important problem with this approach is that it is only briefly described. The user must assume the responsibility for filling in the many steps needed for implementation.

Newbold and Granger (38) suggest what can be considered as another automated alternative to the Box-Jenkins method. It involves the application of two automated processes to arrive at forecasts and then the synthesis of these two sets of forecasts into an overall forecast. In an empirical study of 106 macro-economic and micro sales time series, each of the two automated procedures, Holt-Winters and stepwise autoregression, provide fairly accurate forecasts as compared with the standard Box-Jenkins methods. A more impressive result arises from the combination of the two automated methods as compared to the standard Box-Jenkins procedure (38).

The Combination of Forecasts

The idea of combining forecasts to improve accuracy has been studied in several contexts in the past decade. An empirical study (nicknamed the M-Competition) involving 1001 data sets from a variety of application areas, is reported in several papers (34, 33). One statement that summarizes the conclusion of this massive study pertaining to the combining of forecasts is,

> Using an average of forecasts is undoubtedly better than using a "wrong" model or a single poor forecasting method. ...It may be cheaper, and this study shows that it is less risky, to use a combination of relatively simple methods than to use a single method which is more complex and requires personalized data analysis and costly fitting. Or, if the more complex method is used, it may still be better to combine it with one or more simple methods. The results reported here suggest that using averages of forecasts provides considerable practical benefits in terms of improved forecasting accuracy and decreased variability of accuracy (33, p. 995).

An earlier study also supports the combination of forecasts, as evidenced by this statement, "The main conclusion is that the composite set of forecasts can yield lower mean square error than either of the original forecasts" (10, p. 451).

In a summary of empirical investigations into forecasting accuracy, Mahmoud says ". . . Combinations of forecasts are frequently more accurate because they retain more information about the potential market. In today's increasingly volatile markets, the combining of forecasting methods is particularly important" (29, p. 149). He also suggests that because of the promising results realized thus far with the combination of forecasts, more research should be done in this area.

A general overview of combining forecasts is provided by Wood in this statement,

> In instances where there are two or more forecasting methodologies available for predicting the same item or event, the usual practise is to attempt to discover which of the approaches is "best" according to some criterion. If the objective is merely to characterize and evaluate a methodology, then the embracing of one technique and the discarding of others may be justifiable. However, where the goal is to obtain as accurate a prediction as possible, this path may be singularly unproductive, as the discarded methodologies may have contained some useful independent information (45, p. 223).

Forecasts can be combined using techniques that range from a simple average to complex, weighted averaging schemes. In the M-Competition, both the simple and a weighted average, based on the covariance matrix of fitted errors, were used (34). The results favor the simple method. A subsequent study (33) was performed by Makridakis and Winkler (33) on the same data. This time more forecasting methods were used, but only the simple average was employed.

The results support the combination of forecasts even with the simple averaging technique.

Bates and Granger (10) suggest five schemes for determining optimal weights for combining forecasts. Two of them attempt to base the weights directly on an estimated covariance matrix and the other three relate the weights to reciprocals of sums of squared errors.

An empirical study, by Newbold and Granger (38), that compares the five schemes on 80 sets of monthly data, suggests that methods of the latter type provide better results than those of the former. One of the five methods, Method 1, emerges as the favorite from this study. In this method, the most weight is given to the forecast which has performed best in a specified number (v) of past time periods. In discussing the results of combining forecasts from several methods or using the Box-Jenkins alone, the author state, "If, however, greater accuracy proves desirable one might try combining Box-Jenkins, Holt-Winters and stepwise autoregression forecasts, again using Method 1 with v = 12" (38, p. 144). Another study using the five combining methods was reported by Winkler and Makridakis (44). In this study the M-Competition data was used and the results also support Method 1 with twelve time periods.

Comparative Measures

In assessing the results of any study of forecasting accuracy, there are many comparative methods that can be used. The most common method is the mean squared error (MSE). Various authors, including Makridakis and Winkler (44), Gardner (17) and Guerts (19), criticize the use of this criterion for comparisons involving more than one set of data. They point out that this measure is highly influenced by the magnitude of the data. In other words, if one data set is the number of inches of rainfall per day in the Mojave Desert and another is the gross national product of the United States, a method that effectively forecasts the latter but is inaccurate on the former will be judged to unfair advantage compared with one that may operate in the opposing manner.

For this reason, a relative measure is more appropriate for comparisons across multiple data sets. One such method is the mean absolute percentage error, or MAPE, and another is the median absolute percentage error, or Md APE. According to Gardner, in a review of the M-Competition, "The error distibutions from all methods are badly skewed, which distorts the MAPE. The Md APE is less affected" (17, p. 263). In spite of this criticism, the MAPE seems to be more commonly reported than the Md APE. Overall, the MSE and the MAPE are probably the most prevalent comparative measures, in recent work, with the Md APE used in some studies.

Summary

Because of its importance to business and other disciplines, forecasting is a fertile area for research. Two important concepts discussed in recent time-series literature are the automation of the Box-Jenkins method and the combination of forecasts.

The Box-Jenkins method has been shown to provide accurate forecasts when applied by researchers who are experienced in the technique. The method is not being widely used by practioners in business environments because of the subjectivity, need for an expert analyst and extensive time requirement. To alleviate these problems, automated Box-Jenkins methods are being developed.

Combining forecasts has been shown to be an effective means of aggregating information from multiple forecating methods. Various combining methods exist including the five developed by Bates and Granger (10). Two empirical studies support preference for one of these methods (Method 1). Newbold and Granger (38) had success with the combination of the Box-Jenkins, stepwise autoregressive and Holt-Winters methods. Further research into the combining of forecasts has been called for by authors including Wood (45) and Mahmoud (29).

CHAPTER BIBLIOGRAPHY

1. Adam, E. E. and R. J. Ebert, "A Comparison of Human and Statistical Forecasting," AIITE Transactions, 8 (1976), 120-127.

2. Akaike, H., "Use of an Information Theoretic Quantity for Statistical Model Identification," Proceedings of the Fifth Hawaii International Conference in System Sciences (1972), 249-250.

3. Akaike, H., "Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models," Biometrika, 60 (1973), 255-265.

4. Akaike, H., "A New Look at the Statistical Model Identification," IEEE Transactions on Automated Control, 49 (1974), 716-723.

5. Akaike, H., "Information Theory and an Extension of the Maximum Likelihood," Research Memo No. 46, Institute of Statistical Mathematics, Tokyo (1976) 472-478.

6. Andersen, Allan, "An Empirical Examination of Box-Jenkins Forecasting," Journal of the Royal Statistical Society, 145, 4 (1982), 472-478.

7. Anderson, O.D., "A Commentary on 'A Survey of Time Series'", International Statistical Review, 45 (1977), 273-297.

8. Ang, James, Jess H. Chuaa and Ali M. Fatemi, "How to Use Akaike Criterion to Automatically Specify and Construct Box-Jenkins Models," Journal of Business Forecasting, Spring (1982), 38-42.

9. Armstrong, J. S., "Forecasting with Econometric Methods: Folklore Vs. Fact", Journal of Business, Sl (1978), 549-600.

10. Bates, J. M. and C. W. J. Granger, "The Combination of Forecasts", Operational Research Quarterly, Vol. 20, No. 4 (1969), 451-468.

11. Bowerman, Bruce. L and Richard T. O'Connell, Forecasting and Time Series, Belmont, California, Duxbury Press, 1979.

12. Bransom, M. J., "Discussion of a Paper by Dr. Newbold and Professor Granger," Journal of the Royal Statistical Society, A (1974), 137, Part 2, 157.

13. Carbone, Robert, Allan Andersen, Yvan Corriveau and Paul Piat Corson, "Comparing for Different Time Series Methods the Value of Technical Expertise, Individualized Analysis and Judgmental Adjustment," Management Science, Vol. 29, No.5 (1983), 559-566.

14. Chatfield, C. and D. L. Prothero, "Box-Jenkins Seasonal Forecasting: Problems in a Case Study," Journal of the Royal Statistical Society, A, 136 (1973), 295-336.

15. Chatfield, C., "The Holt-Winters Forecasting Procedure," Applied Statistics, 27, No. 3 (1978), 264-279.

16. Dagum, Estela Bee, "Modelling Forecasting and Seasonally Adjusting Economic Time Series with the X-11 Arima Method," The Statistician, Vol. 27, Nos. 3, 4 (1978), 203-216.

17. Gardner, Everette S., "The Trade-Offs in Choosing a Time Series Method," Journal of Forecasting, Vol., No. 1, 1983, 263-266.

18. Granger, C.W.J., Forecasting in Business and Economics, New York, Academic Press, 1980.

19. Guerts, Michael D., "Evaluating a Forecasting Competition," Journal of Forecasting, Vol. 2, No. 3 (1983), 267-269.

20. Guerts, M. and J. Reinmuth, "The Sensitivity of Model Identification in Using the Box-Jenkins Forecasting Procedure, Proceeding of the 1976 AIDS Meeting, San Francisco.

21. Hanke, John E. and Arthur G. Reitsch, Business Forecasting, Boston, Allyn and Bacon, 1981.

22. Herne, H., "Discussion of a Paper by Dr. Chatfield and Dr. Prothero," Journal of the Royal Statistical Society, A, 136, 3 (1973), 330.

23. Hill, G. W. and D. Woodworth, "Automatic Box-Jenkins Forecasting," Journal of the Operational Research Society, Vol. 31, No. 5 (1980), 413-422.

24. Kang, Chin-Sheng Alan, "Identification of Autoregressive Integrated Moving Average Time Series," unpublished Ph.D. dissertation, Department of Industrial and Management Systems Engineering, Arizona State University, Tempe (1980).

25. Kenny P. B. and J. Durbin, "Local Trend Estimation and Seasonal Adjustment of Economic and Social Time Series," Journal of the Royal statistical Society, A, 145, Part 1 (1982), 1-41.

26. Libert, G. , "Construction Automatique et Continue de Previsions," R.A.I.R.O. R.O., 16 (1982), 333-348.

27. Libert, G., "The M-Competition with a Fully Automatic Box-Jenkins Procedure," Journal of Forecasting, Vol. 3 (1984), 325-328.

28. Mabert, M. A., "Statistical Versus Sales Force - Executive Opinion Short-range Forecasting: a Time-series Analysis Case Study," Krannert Graduate School, Purdue University (working paper) (1975).

29. Mahmoud, Essam, "Accuracy in Forecasting - A Survey," Journal of Forecasting, Vol. 3, No. 2 (1984), 139-159.

30. Makridakis, Spyros, "A Survey of Time Series," International Statistical Review, 44, 1 (1976), 29-70.

31. Makridakis, Spyros and Michele Hibon, "Accuracy of Forecasting: An Empirical investigation," Journal of the Royal Statistical Society, A, 142, Part 2 (1979), 97-145.

32. Makridakis, Spyros , Steven C. Wheelwright and Victor E. McGee, Forecasting: Methods and Applications, New York, John Wiley and Sons, 1983.

33. Makridakis, Spyros and Robert L. Winkler, "Averages of Forecasts: Some Empirical Results," Management Science, Vol. 29, No. 9 (1983), 987-996.

34. Makridakis, Spyros et al., "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition," Journal of Forecasting, Vol. 1, No. 2, (1982), 111-153.

35. McKenzie, Ed, "A Comparison of Standard Forecasting Systems with the Box-Jenkins Approach," The Statistician, Vol. 23, No. 2 (1974), 107-116.

36. Mentzer, F. and J. Cox, "Familiarity, Application and Performances of Sales Forecasting Techniques," Journal of Forecasting, Vol. 1, No. 2 (1984), 27-36.

37. Newbold, Paul, "Some recent Developments in Time Series Analysis," International Statistical Review, 49 (1981), 53-66.

38. Newbold, Paul and C. W. J. Granger, "Experience with Forecasting Univariate Time Series and the Combination of Forecasts," Journal of the Royal Statistical Society, A, 137 (1974), 131-165.

39. Reid, D. J., "A Comparitive Study of Time-Series Prediction Techniques on Economic Data," unpublished PhD thesis, University of Nottingham (1969).

40. Stern, G. J. A., "Discussion of a Paper by Dr. Chatfield and Dr. Prothero," Journal of the Royal Statistical Society, A, 136, 3 (1973), 329-330.

41. Tomasek, Otto, "Discussion of a Paper by Dr. Chatfield and Dr. Prothero," Journal of the Royal Statistical Society, A, 136, 3 (1973), 332-333.

42. Ward, D. H., "Discussion of a Paper by Dr. Newbold and Professor Granger," Journal of the Royal Statistical Society, A, 137 Part 2 (1974), 157.

43. Wheelwright, S. C. and D. G. Clark, "Corporate Forecasting: Promise and Reality," Harvard Business Review, November-December (1976).

44. Winkler, Robert L. and Spyros Makridakis, "The Combination of Forecasts," Journal of the Royal Statistical Society, A, 146 (1983), 150-157.

45. Wood, E. O., "Predicting Property Values," Land Economics, (1976),

CHAPTER III

PURPOSE OF THE RESEARCH

The purpose of this research is the development and evaluation of a fully automated forecasting system to serve as an alternative to the time-consuming, subjective Box-Jenkins method. The system will consist of two major components. One is an automated Box-Jenkins method and the other is the combination of forecasts from the automated and other forecasting methods.

Two specific questions will be addressed:

(1) How does the forecasting accuracy of an automated Box-Jenkins compare with that of the subjective Box-Jenkins?

(2) Would the accuracy of the automated method be improved through synthesis with forecasts provided by other methods?

CHAPTER IV

METHODOLOGY

Modified Ang

To answer the first question posed in this research, a particular automated Box-Jenkins method was sought. The method selected is a fully specified procedure which adds seasonality and detail to a procedure that was briefly suggested by Ang et al. (1). This new method, named the "Modified Ang" method in this research, has the advantages of the Ang procedure, including use of the newest Akaike decision criterion. It also eliminates the disadvantages including the lack of seasonality handling and fuzzy procedure specification. The Akaike decision criterion used in this method is:

$$AIC = n \ln \hat{\sigma}^2 + 2(p + q)$$

where:

p is the number of autoregressive terms in the model

q is the number of moving-average terms in the model

$\hat{\sigma}^2$ is the unbiased estimate of mean square error defined as $\hat{\sigma}^2 = S^2/(n-p-1)$

for

$$S^2 = \sum_{t=1}^{n} a_t^2 \quad \text{and } a_t = Z_t - \hat{Z}_t$$

Major changes of the original Ang method, other than the addition of seasonality handling, involved refinement to facilitate the use of the procedure by practioners who may have computer time limitations as well as to avoid abortive attempts by the procedure to invert a non-singular matrix. Ang's original process began by fitting autoregressive models of order 20. The Modified method begins at order 6 in order to keep CPU time for a data set of 100 observations under 3 minutes. The other major modification was made to handle very small data sets. For data with fewer than twenty-two observations, the largest autoregressive model fit is (n-4)/2 where n is the number of observations in the data set. Figure 1 contains a complete flowchart of the Modified Ang procedure. Appendix B contains the computer code which creates a program for performing the Modified Ang method from a manual Box-Jenkins package called T-Series (9). The code is in two parts. One part is the main subroutine that drives T-Series and the other is the set of changes that were imbedded into the T-Series program.

Other Methods

Answering the second question listed in this research required the selection of other automated forecasting methods. The work of Newbold and Granger (10), in which the Stepwise Autoregressive and Holt-Winters were combined with the Box-Jenkins, played the principal role in this process.

The Stepwise Autoregressive Method, that was developed by Payne (11) is a backward elimination method which seeks to define the most parsimonious autoregressive model that fits the data. It has been recommended for designation as the standard method of forecasting to the British Government Statistical Service (6). For these reasons as well as the very promising results of the Newbold and Granger (9) study, the Stepwise Autoregressive and Holt-Winters methods were selected for use in this research. Figures 2 and 3 contain flowcharts for the Holt-Winters and the Stepwise Autoregressive methods. Computer programs for each technique are included as Appendix B.

Combining Techniques

Two combining techniques were employed in this research. The techniques include the simple average and a weighted average method of Bates and Granger (2). This weighted average method has been written by Granger and Newbold (4, p. 277) as:

$$\hat{k}_n(i) = \left(\sum_{t=n-12}^{n-1} e\,(i)^2\right)^{-1} \Bigg/ \left\{\sum_{j=1}^{M} \left(\sum_{t=n-12}^{n-1} e_t(j)^2\right)^{-1}\right\}$$

where:

$\hat{k}^{(i)}$ is the weight for method i n is the number of time periods

M is the number of methods in the combination

t is the time period

j is the particular method

e is the residual for the specified time period and method

The selection of this weighted scheme and the simple average is suggested by several empirical studies (10, 8, 12, 7).

Data

In order to study forecasting effectiveness, both empirical and simulated data were used. Because of the impossibility of randomly sampling the universe of time-series data, statistical inference was not. A broad perspective was achieved, however, through the use of a wide range of actual and theoretical data sets.

The empirical data analyzed arose from the large-scale M-competition (8). The analysts in this study selected an original group of 1001 data sets on which to perform the analysis. The major classification of the series are provided as Table 1. Due to the time constraints of the M-competition, the most complex analyses (including Box-Jenkins) were performed on a systematic sample of 111 sets. The same systematic sample had been used in this research and comparisons made to the Box-Jenkins results which are presented in the M-competition.

To compute the weighted averages of forecasts from the Modified Ang plus the Holt-Winters and Stepwise Autoregressive Methods using the combining technique under study, a minimum of twenty-two observations are required. The range of n (where n is the number of observations) of the data sets in the M-Competition is from nine to over one-hundred. In the case where n is at least 22, the twelve past errors were computed as Bates and Granger suggest. For data sets with fewer than twenty-two observations, $n/2 + 1$ past errors were computed.

Forecast horizons used for the comparison follow guidelines provided by the M-competition (8, p. 112). For monthly data, forecasts are compared for periods 1, 4, 6, 8, 12, and 18. For quarterly data, periods 1, 4, 6, and 8 are used and periods 1, 4, and 6 are compared for yearly data.

Generated data sets arose from twenty-one autoregressive, moving-average or mixed models. These models have been analyzed using the standard Box-Jenkins approach by a team of analysts, including G. E. P. Box who was obviously one of the originators of the method (3). These models have been included because they represent known characteristics and because standard Box-Jenkins models have already been prepared by Box-Jenkins experts on data generated by these models.

The generating models, the actual data used to fit the Box-Jenkins models and the resultant models are available from the dissertation by Kang (5). From this information, the mean and standard deviation of the deviates that were used in the original analysis were estimated. Eighteen more data points were generated, assuming a normal distribution of the deviates with mean and variance as estimated. The Box-Jenkins models provided by Box, Taio and Pack were used to provide eighteen post-sample forecasts. The three methods under study were likewise applied. Forecast horizons for periods 1, 4, 6, 8, 12 and 18 were considered.

Comparisons

Both the mean absolute percentage error and the median absolute percentage error were used to report forecasting results. These measures were selected in the interest of providing appropriate comparisons across multiple data sets, as both measures are free of magnitude bias.

In both the generated and empirical data situations, comparisons were made between the standard Box-Jenkins, as applied by the respective experts, and the modified Ang method. Another set of comparisons include the standard Box-Jenkins against combinations of the forecasts from the modified Ang with those from each of the other two methods. The last complement of comparisons results from the combination of forecasts from the three methods against the forecasts from the standard Box-Jenkins model.

Summary

In summary, the Ang procedure was made operational through the formulation of a complete procedure. The procedure includes seasonality handling, small data set handling and other algorithms. Forecasts were computed for various time frames using each of three automated forecasting methods on a total of 132 data sets. Forecasts from the automated Box-Jenkins method were compared to the standard Box-Jenkins results both individually and in combination with forecasts from two other automated methods, using each of two combining techniques. Two accuracy measures were used to formulate the comparisons with the extensive results presented in tabular, graphical and text forms. Conclusions based on these results are succinct but not statistical in nature replicating the major format of the M-competition (8).

CHAPTER BIBLIOGRAPHY

1. Ang, James, Jess H. Chuaa and Ali M. Fatemi, "How to Use Akaike Criterion to Automatically Specify and Construct Box-Jenkins Models," Journal of Business Forecasting, Spring (1982), 38-42.

2. Bates, J. M. and C. W. J. Granger, "The Combination of Forecasts", Operational Research Quarterly, Vol. 20, No. 4 (1969), 451-468.

3. Box, G. E. P. and G. C. Tiao and D. R. Pack, "Modern Techniques in Time Series Forecasting and Dynamic Model Building," unpublished course material at Princeton University (1977).

4. Granger, C.W.J. and Paul Newbold, Forecasting Economic Time Series, New York, Academic Press, 1980.

5. Kang, Chin-Sheng Alan, "Identification of Autoregressive Integrated Moving Average Time Series," unpublished Ph.D. dissertation, Department of Industrial and Management Systems Engineering, Arizona State University, Tempe (1980).

6. Kenny P. B. and J. Durbin, "Local Trend Estimation and Seasonal Adjustment of Economic and Social Time Series," Journal of the Royal Statistical Society, A, 145, Part 1 (1982), 1-41.

8. Makridakis, Spyros and Robert L. Winkler, "Averages of Forecasts: Some Empirical Results," Management Science, Vol. 29, No. 9 (1983), 987-996.

9. McGill University, McGill University System for Interactive Computing, User's Guide, 1981.

10. Newbold, Paul and C. W. J. Granger, "Experience with Forecasting Univariate Time Series and the Combination of Forecasts," Journal of the Royal Statistical Society, A, 137 (1974), 131-165.

11. Payne, D. J., "The Determination of Regression Relationships Using Stepwise Regression Techniques," unpublished Ph.D. thesis, University of Nottingham (1973).

12. Winkler, Robert L. and Spyros Makridakis, "The Combination of Forecasts,2 Journal of the Royal Statistical Society, A, 146 (1983), 150-157.

CHAPTER V

RESULTS

Introduction

This chapter presents and discusses the results of the experiments performed in this study. The results are prefaced by a recapitulation of the overall processes involved in the research.

This research involved two main phases. Phase one was the development of a forecasting method that performs completely automated ARIMA model identification and fitting. The method, which relies on a process briefly described in a previous publication,is referred to in this paper as the Modified Ang method. Modifications necessary to make the process operational include the addition of a procedure to handle seasonality and a procedure to deal with data sets with as few as nine observations, as well as the reduction of the suggested initial model order.

Phase two involved the use of the Modified Ang, both by itself and in two types of combinations with two other automated proceses, the Holt-Winters exponential smoothing technique and the Stepwise Autoregressive method. Each method was applied to two sets of data that had previously been analyzed using the subjective Box-Jenkins method. For each data set, recognized authorities in the subjective Box-Jenkins method performed the procedure. The first data set, consisting of twenty-one simulated series, was analyzed by the team of Box, Taio and Pack. The second set, a group of 111 empirical series, was analyzed by a reseacher who was selected to perform the Box-Jenkins analysis for the M-Competition on the basis of his reported experience with this subjective method.

The data analysis was employed in order to accomplish two research objectives. The first objective was to determine whether the modified Ang technique could forecast as well as the subjective Box-Jenkins method, as applied by experts. The second objective was to determine whether the combination of the Modified Ang with either the Holt-Winters or the Stepwise Autoregressive method, or both, would provide increased forecasting accuracy.

Generated Data

Table II reports two measures of forecasting accuracy, the Median Absolute Percentage Error (Md APE) and the Mean Absolute Percentage Error (MAPE), for the twenty-one sets of generated data. This table lists the results of the Box-Jenkins, the Modified Ang and the Modified Ang in combination with the Holt-Winters and Stepwise Autoregressive methods. The list reflects the use of two types of combining schemes for the Modified Ang and the other two methods. The first group of averages resulted from simple averaging, that is addition of the forecasts from one or more methods and division by the number of methods. The second group of averages resulted from a weighted average scheme, described in the Methodology Section. Time horizons listed in the table represent the number of time periods into the future from the end of history. For example, the values in the column titled "time horizon 4" represent the ability of the forecasting scheme to predict the value of the series four periods into the future. The last column represents an aggregate absolute percentage error across all time horizons under study.

Regarding the results presented in Table II, it is important that the generated data arose from ARIMA models. Since data representing ARIMA characteristics should be most accurately modelled through ARIMA modelling, it is not surprising that the ARIMA modelling methods (Box-Jenkins and Modified Ang) performed well. It is interesting, however, that based upon the Md APE criterion for the aggregate time horizons 1-18, the simple average of the Modified Ang plus the Holt-Winters outperforms the subjective Box-Jenkins method. The MAPE for the same aggregated comparison slightly favors the Box-Jenkins procedure.

Another interesting result involves the MAPE comparison between the Box-Jenkins and Modified Ang method for time horizon 1. In this case, The Modified Ang outperforms the Box-Jenkins method. This result is especially noteworthy considering the fact that the Box-Jenkins method was designed for one-step-ahead forecasting. The structure of the ARIMA models is based upon determination of a forecast for period p based on period p-1, and the error in the prediction of period p-1. As forecasting is performed for periods more than one step ahead, terms in the model drop out, and eventually the forecasts deteriorate to a constant. For this reason, a comparison of ARIMA forecasting methods might effectively be reduced to a comparison of one-step-ahead forecasts (time horizon 1) alone. This research, which was patterned after the M-Competition, follows the popular scheme of providing comparisons on many time horizons, including the one-step-ahead value.

As Table II shows, for time horizon 1, the Modified Ang method performs approximately as well as the Box-Jenkins, and combination with the other methods provides no improvement to the Modified Ang forecasts. For time horizon 4, the weighted average of all three methods out-performs Box-Jenkins on the median criterion but not on the mean. For time horizon 6, the Box-Jenkins performs best using either accuracy measure. The Md APE measure favors an unweighted combination of all three methods for time horizon 8, while the MAPE supports Box-Jenkins. In time horizon 12, both the

MAPE and the Md APE favor the simple average of Modified Ang plus the Holt-Winters method.

The column of time horizon 18 contains the highest MAPEs for all methods, with the lowest entry provided by the simple average of Modified Ang and Holt-Winters. On the median criterion for that time horizon, the weighted average of the Modified Ang, the Holt-Winters and the Stepwise Autoregressive method outperforms all other methods including the Box-Jenkins.

Figure 4 shows a comparison between the Box-Jenkins method and three of the other methods under study . This bar chart represents the percent of series for which each of three methods provided lower Mean Absolute Error than the Box-Jenkins. Each Mean Absolute Error is determined using only one series at a time, resulting in a series-by-series comparison. The first bar represents the percent of series for which the Modified Ang method had a lower MAPE than did the Box-Jenkins. The second bar represents the Modified Ang averaged with the Holt-Winters (simple), and the third represents the weighted average for the Modified Ang plus the Holt-Winters procedure.

For time horizons 1, 4, and 1-18, the Modified Ang out-performed the subjective Box-Jenkins method. For every time period except 18, at least one of the other methods provides lower errors than the Box-Jenkins.

Figure 5 contains a comparison of the number of series for which each method has the lowest MAPE for periods 1-18. The first bar represents the Modified Ang method. The next three bars represent the results obtained through averages of the Modified Ang with the Holt-Winters, the Stepwise Autoregressive and both, in that order. The next three bars represent results of the same groupings with the weighted average. The last bar represents the Box-Jenkins results. The Modified Ang has the lowest MAPE for seven out of the twenty-one series. In comparison with the Box-Jenkins method, the Modified Ang alone and in combination with the other automated methods ,provides the lowest MAPE for 76 percent of the twenty-one series.

Figure 6 contains the same sort of comparison for period 1 alone. In this situation, the Modified Ang and the Box-Jenkins ties for the number of series with the lowest MAPE. All automated methods together provide the lowest MAPE for 67 percent of the twenty-one series.

Another comparison performed gives consideration to the way in which post-sample values were determined. That is, to compute future values for a set of data created by ARIMA models, errors must be introduced. Because the error distributions have different standard deviations for each series, a comparison pitting the Modified Ang against the Box-Jenkins relative to the standard deviations was performed. Figure 7 shows this comparison. The Mean Absolute

Error of the Modified Ang was divided by the Mean Absolute Error of the Box-Jenkins method for each series and was plotted against the standard deviation of the errors in the series. Obviously, ratios greater than one represent series for which the Box-Jenkins method outperforms the Modified Ang, and the reverse is true for ratios less than one. A serious effect of the magnitude of the standard deviations would be evidenced by a clustering of points into two distinct sets. Figure 7 does not represent such an obvious clustering.

A succinct statement of the results based on the twenty-one sets of generated data would be that the Modified Ang seems to provide about the same accuracy as the subjective Box-Jenkins method. Combining the Modified Ang with either the Holt-Winters or the Stepwise Autoregressive procedure provides little advantage except in the case of the MAPE for all time horizons combined, when both the simple and weighted averages of the Modified Ang plus the Holt-Winters are clearly lower than corresponding measures for the Modified Ang method, alone.

Empirical Data

Results of research on the empirical data are reported in the same format as that of the generated data. Table III contains information regarding all 111 data sets. The most interesting result reported in this table is the 3.7 median for the one-step-ahead forecast, as determined by a weighted average of the Modified Ang with the Holt-Winters method. This value compares very favorably with the Box-Jenkins result of 5.3. The MAPE results for the same time horizon and methods yield very small differences. The Modified Ang method produces results that differ one or two percent from the Box-Jenkins results when considering the aggregate time horizon. In the one-step-ahead values, however, the difference is very small. For four time periods into the future, the simple average of the Modified Ang and the Stepwise Autoregressive methods produces a lower MAPE than do any of the other methods, including the Box-Jenkins. While for time horizon 6, the weighted average of the Modified Ang and the Holt-Winters methods provides the lowest MAPE. For the aggregate of time periods 1 through 6, the Box-Jenkins provides the lowest value for both accuracy measures, with the weighted average of the Modified Ang and the Holt-Winters methods registering as the second best.

To allow for more detailed scrutiny of the empirical results, the 111 series have been partioned into several categories. Table IV contains information on the twenty sets of yearly data. As in Table III, the weighted average of the Modified Ang and the Holt-Winters provides the minimum Md APE for one-step-ahead forecasting. This same combination of methods ranks second best in MAPE and Md APE for the aggregate time horizons 1 through 6. The simple average of the Modified Ang with both of the other methods provides the best values on both accuracy measures for the aggregate time frame. That combination also beats the weighted Holt-Winters and Modified Ang mix in the MAPE one-step-ahead forecasts. On the MAPE criterion, three methods consistently provide lower values than the Box-Jenkins for every time horizon. These methods are the Modified Ang plus the Holt-Winters in both simple and weighted combining schemes and the simple average of all three methods.

The Modified Ang alone provides lower MAPEs than the Box-Jenkins method on every time period except the first, and it differs only one-half of one percent on that time period. The median criterion favors the Box-Jenkins over the Modified Ang alone in every category.

Table V presents results of analysis on the quarterly data. The weighted average of the Modified Ang with the Holt-Winters method performs second to Box-Jenkins on the Md APE measure for one-step-ahead forecasting, while the simple average of the Modified Ang with both other methods provides the MAPEs that are second best to those of the Box-Jenkins method. On the aggregate comparison (time horizons 1 through

8 in this group of series), Box-Jenkins does the best on the Md APE criterion, but the simple average of the Modified Ang and both of the other methods outperforms it by 2.6 percent on the MAPE criterion.

Table VI contains the results of analysis on the largest subset of the 111 series, the sixty-eight monthly series. In this case, forecasts for eighteen time periods into the future were computed. The Md APE for the one-step-ahead forecasts indicates the superiority of the weighted average of the Modified Ang with the Holt-Winters method. The MAPE measurement favors the simple average of the Modified Ang with the Stepwise Autoregressive method, however. For the aggregate comparison for time horizons 1 through 18, the Md APEs of the Box-Jenkins and the Modified Ang methods are very close while the MAPEs differ more. Regarding the individual time horizons that are reported, time horizon 18 seems to have caused some difficulty for the Modified Ang compared with the Box-Jenkins method for some series. This difficulty is reflected in the MAPE for combinations with the Modified Ang and all of the other methods for that time period.

Two other partitions were analyzed. The first, which was reported in the M-Competition, represents the comparison of the seasonal series with those considered non-seasonal. The second partition, long versus short series, was added to this study in an attempt to focus on some performance characteristics of the methods under study. Since such a comparison was not included in the M-Competition, the results do not include the Box-Jenkins values.

The seasonal and non-seasonal subset results are contained in Table VII. It is interesting that for one-step-ahead forecasting on non-seasonal data, four methods provide lower MAPEs than does the Box-Jenkins method. Only one method, the weighted average of the Modified Ang and the Holt-Winters, does perform in a corresponding manner when using the Md APE criterion, however. For non-seasonal series, the same weighted average of the Modified Ang and the Holt-Winters method is the only method that provides lower values on both accuracy measures than the Box-Jenkins for one-step-ahead forecasting. For the combined time horizons 1 through 6 on seasonal data, five methods have equal or lower MAPEs than the Box-Jenkins. For time horizons 4 and 6 on the same series, every method, with the exception of the simple average of the Modified Ang and the Holt-Winters, has a lower MAPE than does the Box-Jenkins.

Table VIII presents comparisons related to the empirical data involving the Mean and Median Absolute Percentage Errors for the partition based on the number of observations in the series (n). The range of n is from nine to 126 in the 111 series. Table IX presents a frequency distribution of n. This range was divided into thirds, with comparisons provided for the top and bottom thirds. The thirty-seven "short" series each contain thirty-eight or fewer observations while each of the thirty-seven "long" series has sixty-five or more observations. Since the Box-Jenkins method is often reported to be appropriate only for series longer than some stated benchmark (with various values assigned), this comparison was made to ascertain whether there was any appreciable difference in performance based on length of series for the Modified Ang and combinations for the data under study.

Contrary to what might be expected, for one-step-ahead forecasts the short series have lower MAPEs and Md APEs than do the long series, for the majority of methods. As the forecast horizons increase, however, the more expected result holds true. The weighted average of the Modified Ang plus the Holt-Winters method outperforms the other methods on both accuracy measures for the short series in the one-step-ahead category. For the long series, in that time horizon, the lowest MAPE is realized through the weighted Modified Ang plus Stepwise Autoregressive procedure combination, while the simple average with Holt-Winters achieves the best Median APE.

Over various time horizons, the weighted average of the Modified Ang and the Holt-Winters method performs very well on the short series, and on some time horizons for one or the other accuracy measure on the long series. The combinations of the Modified Ang and the Stepwise Autoregressive methods also figure in a determination of the best method on the long series. For time horizon 6 of the long series, the Modified Ang method by itself provided the lowest MAPE.

Unfortunately, detailed series-by-series comparison with the Box-Jenkins method could not be performed on the empirical data, as it was on the generated data. Results for the Box-Jenkins models were available only on an aggregate basis and for the partitions as reported.

Figures 8 and 9 present summary information on one-step-ahead forecasting results. Each figure contains the three best methods for each of the six partitions of empirical time series (all, seasonal, non-seasonal, yearly, quarterly and monthly) as determined by one of the accuracy measures. Figure 8 provides the three best methods for each partition based upon lowest Md APE. Figure 9 presents the same information as determined by the MAPE accuracy measure. In each of these illustrations, the lowest bar in a groups represents the best method for that group. Of the twelve comparisons presented on the two illustrations, the Modified Ang plus the Holt-Winters in a weighted average is the best method six out of twelve times. This method is one of the best three methods in ten cases out of twelve. The Box-Jenkins method is the best method for four times out of twelve and is in the top three seven times out of twelve.

Although the Modified Ang, alone, never produced the lowest MAPE or Md APE in this comparison, it did provide one of the top three forecasts four out of twelve times. This comparison demonstrates a value of adding the Holt-Winters forecasts to the Modified Ang in one-step-ahead forecasting.

In terms of time requirements for the automated and the subjective Box-Jenkins analyses, the maximum time required by the automated method was three minutes of CPU time. This compares favorably with the one hour (on the average) of analyst time, including computer time, that was necessary to perform the subjective Box-Jenkins analysis.

CHAPTER VI

CONCLUSIONS

The statement of purpose for this research poses two questions. The first asks whether an automated method can provide the accuracy of the subjective Box-Jenkins method. The second asks whether combining forecasts of the automated method with those from other methods can improve the accuracy of the automated Box-Jenkins method. The results of the study indicate that the response to each question is "yes."

According to the results of this study, the Modified Ang method is able to provide forecasts about as accurate as those of the subjective Box-Jenkins method as applied by researchers who are recognized experts in the use of the method. The results also indicate that in many situations, the combination of the Box-Jenkins and the Holt-Winters method can provide more accurate forecasts than can either the Modified Ang or the subjective Box-Jenkins method.

This study, adhering to the design of the M-Competition, does not include an attempt to randomly sample the population of all time-series. For this reason, the use of inferential statistics is not appropriate. Even though the possibility of generalizing the results of this study is hampered by the absence of statistical inference, the study does have an implication for the practitioner. In situations in which a Box-Jenkins analysis is difficult or impossible because of the absence of an expert Box-Jenkins analyst, or other resources, the fully automated technique can adequately provide the forecasting function. This is not to say that the automated technique should become a substitute for the subjective Box-Jenkins technique in all situations, but it does provide an accurate alternative when automation is desirable or necessary.

On the generated data, the Modified Ang method was able to provide lower MAPEs on one-step-ahead forecasts than did the Box-Jenkins procedure. This result is very important in any consideration of the use of the automated process in inventory or production environments. An automated process could be run every time period to produce a forecast for the next time period, whereas the requirements for time and personnel of the manual Box-Jenkins method might preclude its use in that context.

On the entire set of empirical data, the weighted combination of the Modified Ang plus the Holt-Winters method provided a Md APE for one-step-ahead forecasts that was 1.6 percent lower than that of the Box-Jenkins method. The weighted combination of the Modified Ang plus the Holt-Winters method was better on that time horizon than the Box-Jenkins method in every partition except quarterly data in which it was slightly worse.

The MAPE comparison favored the Box-Jenkins method over all other methods, including the Modified Ang plus Holt-Winters, for one-step ahead forecasts in four of the six partitions. The discrepancy between the results as reported by the mean and the median was further examined for the Modified Ang plus Holt-Winters weighted average method. Table X contains a frequency table of the absolute percentage errors of all series resulting from this method. For 66.67 percent of the series, the method provided errors of 10 percent or less. For 1.8 perent of the series, however, the errors were over sixty percent. An interesting result occurs when the highest error, 118.3 percent, is removed from the distribution and the MAPE calculated. The new MAPE becomes 9.9 which is lower than the 10.3 MAPE of the Box-Jenkins method. Thus, the descrepancy in the MAPE accuracy measure and the Md APE, for these two methods, for one-step-ahead forecasting, is determined by one time-series.

This result may indicate that a subjective analysis is preferable for some series. Perhaps the analyst, experienced in ARIMA modelling, was able to detect exceptional circumstances in a particular series and circumvent the standard ARIMA modelling process in favor of one that avoids some pitfalls and provides better forecasts. It could also be that the addition of the logarithmic transformation to the Modified Ang method might provide the avenue to better models in some situations, automatically.

If equal consideration were to be given to the MAPE and the Md APE, the Modified Ang in weighted average with the Holt-Winters, would clearly arise as the best choice for one-step-ahead forecasting. This is a very significant result because of the design of ARIMA modelling for exactly this type of forecasting. The fact that one-step ahead forecasting must be performed every time period adds to the desirability of the automated alternative to a time-consuming, subjective Box-Jenkins analysis.

In a determination of whether to use the Modified Ang, alone or in combination with one of the other methods for other than one-step-ahead forecasting, the MAPE accuracy measure primarily favors the former on the empirical data, while the Md APE consistently favors the weighted combination of the Modified Ang with the Holt-Winters method. The only situation in which the measures agree is on non-seasonal data, where both favor the combination of the

Modified Ang with the Holt-Winters. On the generated data, the Modified Ang alone is preferred over the Modified Ang in combination with any other method.

The results from a composite of the two types of data analyzed suggest that if one knows that the data to be used in forecasting has arisen through an ARIMA process, then the Modified Ang, which provides such a model, should be used by itself. In situations where an underlying model is not known, or when the structure is other than ARIMA, the recommendation is to use the combination of the Modified Ang in a weighted average with the Holt-Winters method. This suggestion should be abandoned in favor of the Modified Ang when the increase in overhead caused by the fitting of the Holt-Winters method and the computation of weights exceeds some level of tolerance.

It is necessary to provide the usual caveats to the practitioner who may desire to employ the results of this forecasting competition. Obviously the accuracy of any forecasting method is based upon the continuation into the future of patterns in the past. Although the use of adaptive methods can partially alleviate that concern, it must be suggested that a monitoring device be in use to determine the ability of the method to forecast for the particular situation. In other words, rather than setting up an arbitrary benchmark in the system itself, the user should decide an error tolerance and monitor the method for deviations above that level.

Since this study employed both simulated and real-world data in the comparison, the results should be reliably extrapolated into practice. Generalizations beyond the data studied, however, are not strictly appropriate in this or in other such forecasting research.

Recommendations for Further Study

Recommendations for further study include several areas. One area is that of the computation of weights. The difference in performance between the simple and weighted average did not consistently favor weighting, in this research. The suggestion might be that another weighting technique could provide better results.

Another area that might prove fruitful is that of the distributions of errors. Perhaps a comprehensive look at the variances in the absolute percentage errors for each time horizon and method would provide more insight into the differences in results as measured by the mean and median absolute percentage errors.

Other comparisons could be performed. For example, large sets of data representing particular characteristics commonly found in business data could be generated to determine specific performance characteristics of the automated methods. An interesting comparison might result through application of the Modified Ang alone and in combination with the Holt-Winters and the manual Box-Jenkins method, as applied by the average practioner rather than the "expert." This comparison might provide the business person with a more realistic basis for selection between the manual and the subjective methods.

APPENDIX A

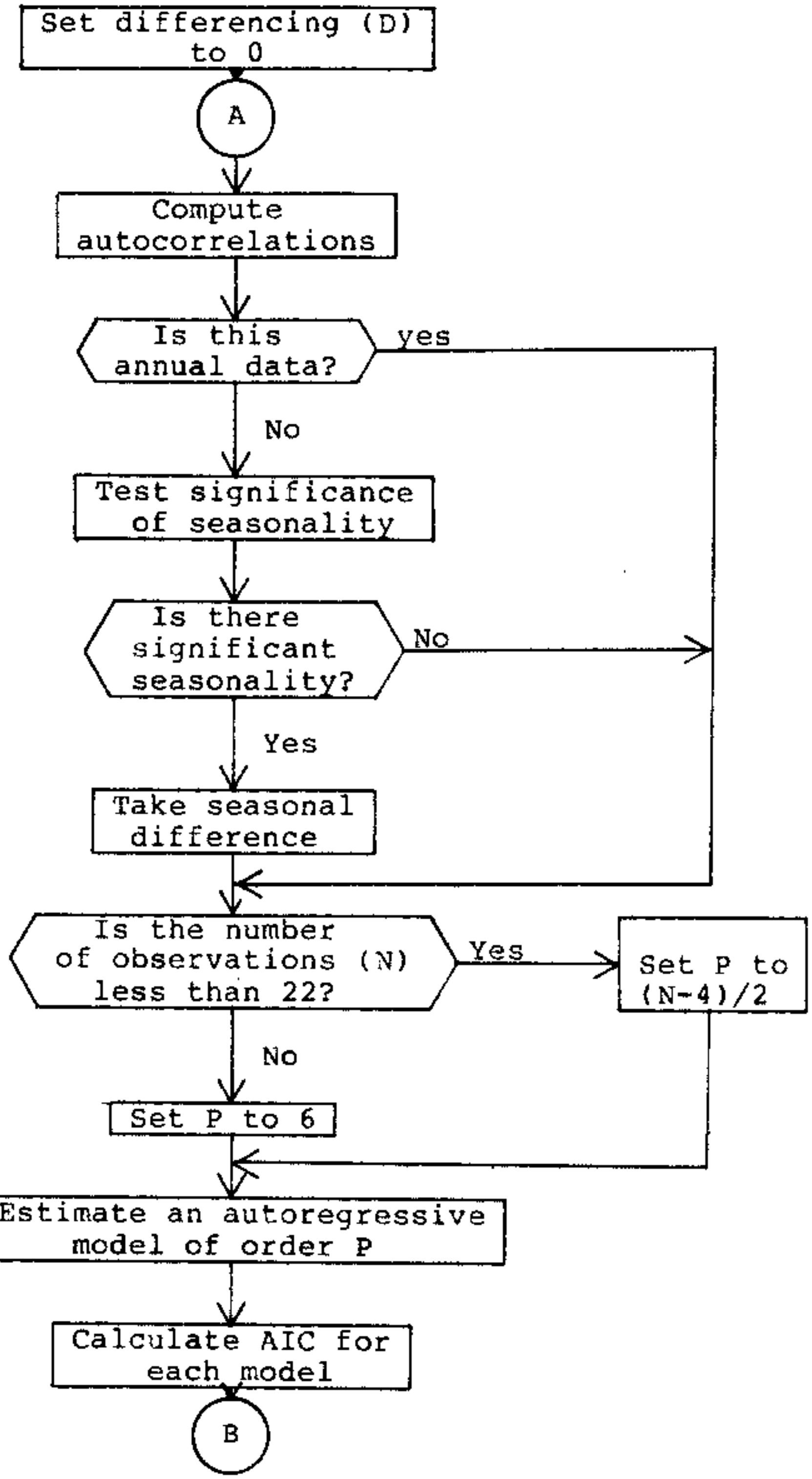


Fig. 1--Flowchart of the Modified Ang Method

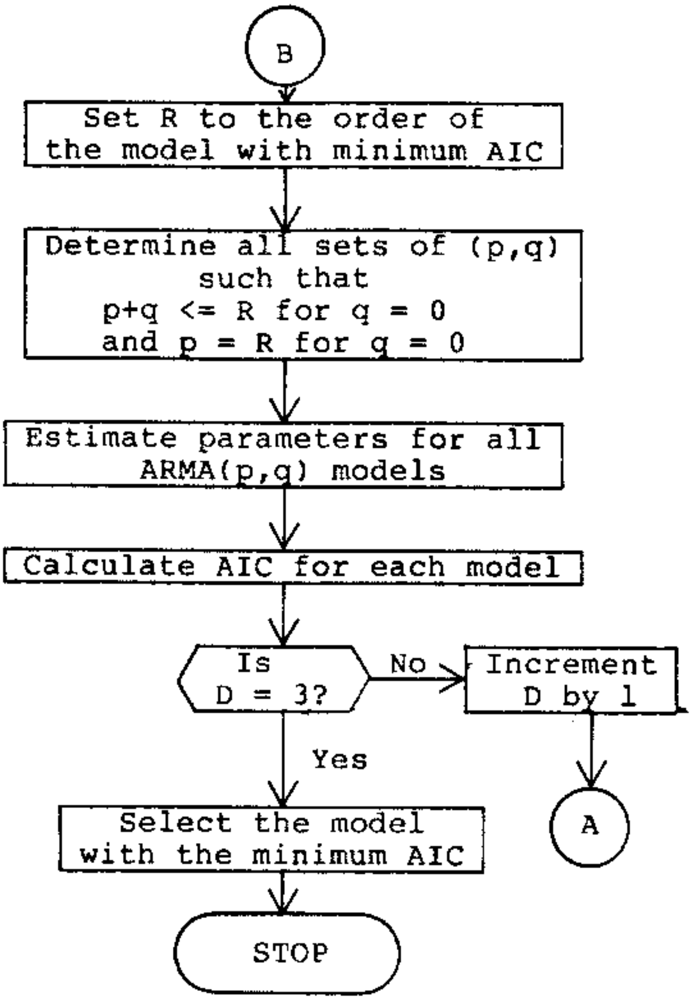
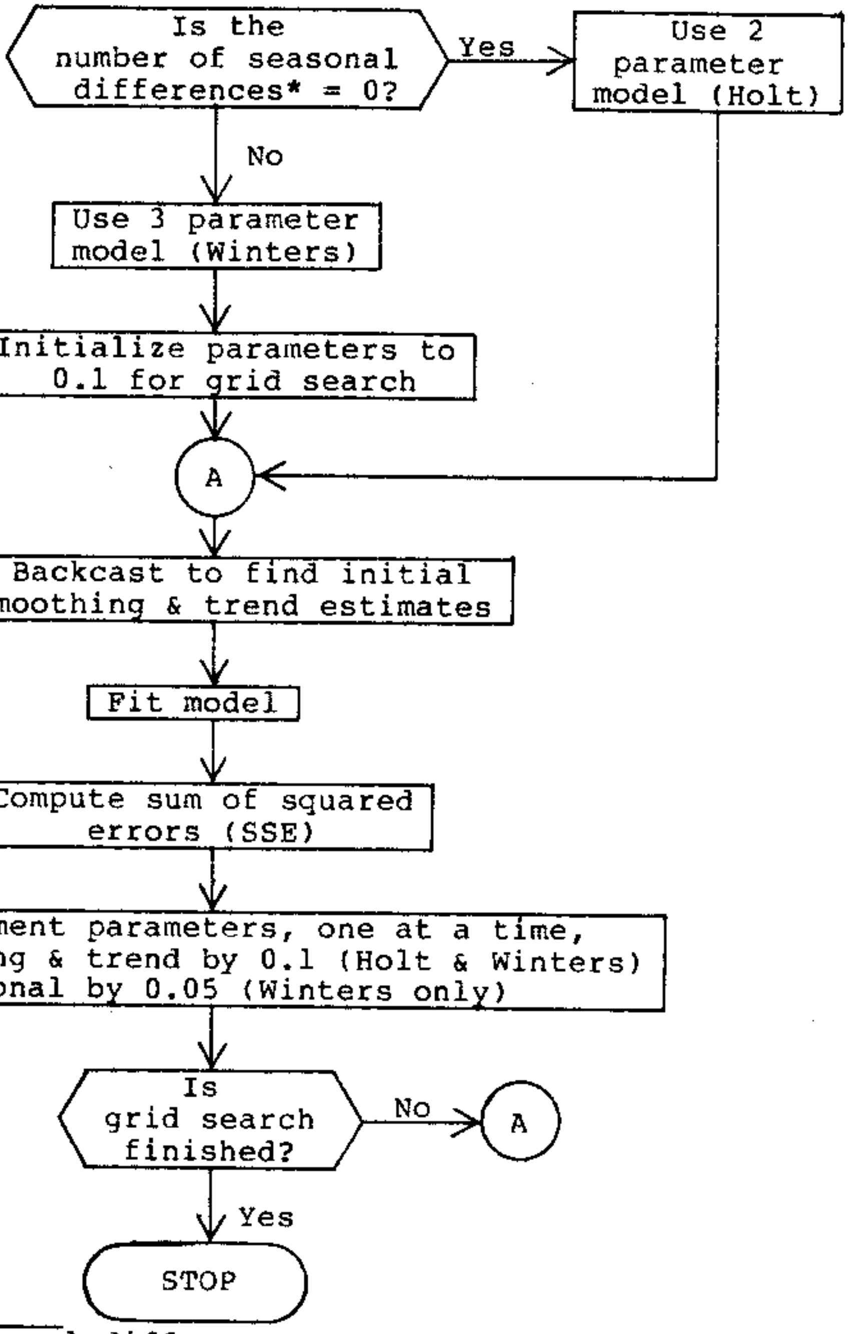


Fig. 1--*Continued*



* Number of seasonal differences determined by autocorrelations in Modified Ang procedure

Fig. 2--Flowchart of the Holt-Winters Method

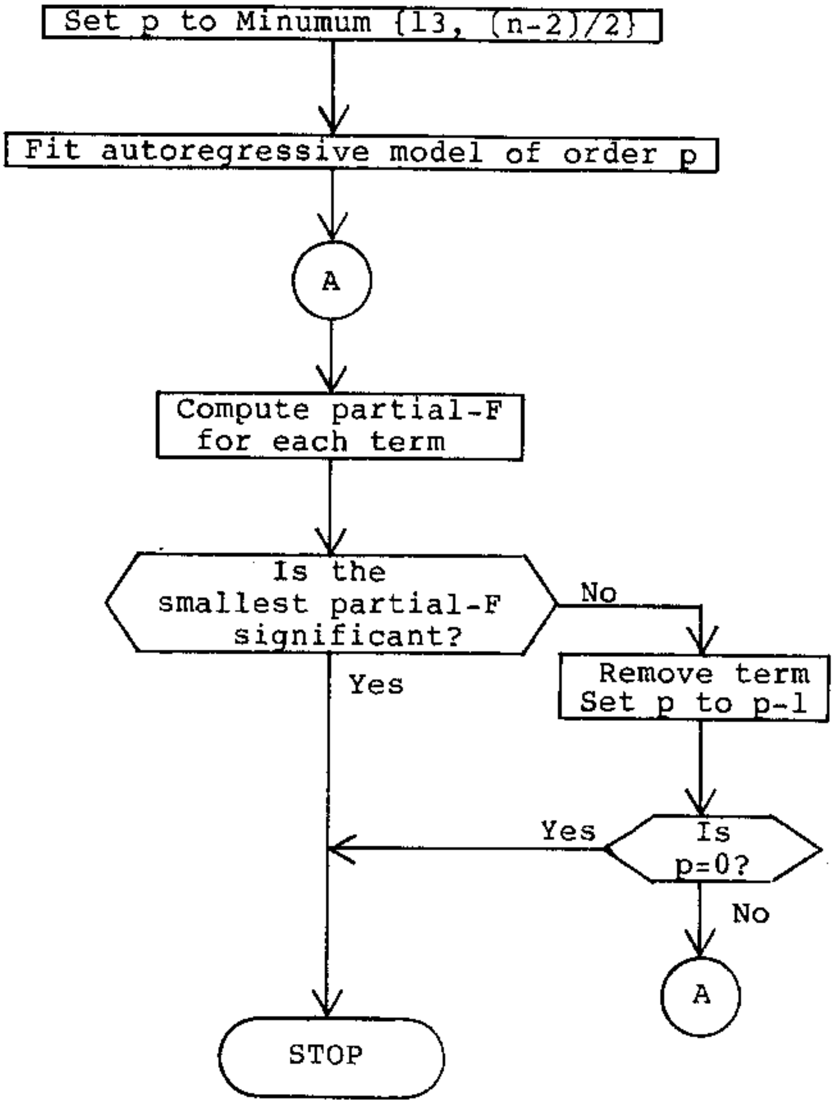


Fig. 3--Flowchart of the Stepwise Autoregressive Method

TABLE I

TYPES OF TIME SERIES DATA

(Makridakis et al., 1982)

Time Intervals	Total firm	Major divisions	Below major divisions	Industry	GNP*	Below GNP*	Demo-Graphic	Total
Yearly	16	29	12	35	30	29	30	181
Quarterly	5	21	16	18	45	59	39	203
Monthly	10	89	104	183	64	92	75	617
Total	31	139	132	236	139	180	144	1001

* GNP or its major components

TABLE II

ERRORS FOR GENERATED DATA (21 SETS)

Method	Median Absolute Percentage Errors							Mean Absolute Percentage Errors						
	Time Horizons							Time Horizons						
	1	4	6	8	12	18	all	1	4	6	8	12	18	all
1.Box-Jenkins	.7	3.3	4.4	5.0	4.0	3.4	3.7	1.9	4.2	5.5	5.4	7.0	10.2	6.3
2.Modified Ang	.8	3.7	5.3	5.0	4.9	5.3	3.9	1.7	4.9	6.6	6.8	8.2	13.6	7.6
3.Modified Ang_1+ Holt-Winters[1]	1.4	3.9	5.7	4.8	3.6	6.7	3.6	2.3	4.8	6.0	6.7	6.3	9.9	6.4
4.Modified Ang + Stepwise Autoregressive[1]	1.2	3.1	6.6	5.5	4.0	4.4	4.0	1.9	4.7	7.0	7.8	12.2	33.0	12.2
5.all three methods[1]	.8	3.0	5.9	4.7	4.2	6.5	4.3	1.8	4.7	6.4	7.1	9.5	24.0	9.7
6.Modified Ang_2+ Holt-Winters[2]	.8	4.6	5.9	5.9	4.4	5.0	3.9	2.0	4.9	6.3	6.6	7.2	11.5	6.8
7.Modified Ang + Stepwise Autoregressive[2]	1.4	3.1	5.1	5.5	4.3	3.6	4.0	2.2	4.6	6.9	7.5	11.3	28.0	11.1
8.all three methods[2]	1.2	2.9	5.6	5.3	4.4	3.1	4.0	2.2	4.5	6.4	6.8	9.0	17.5	8.5

[1]simple average [2]weighted average

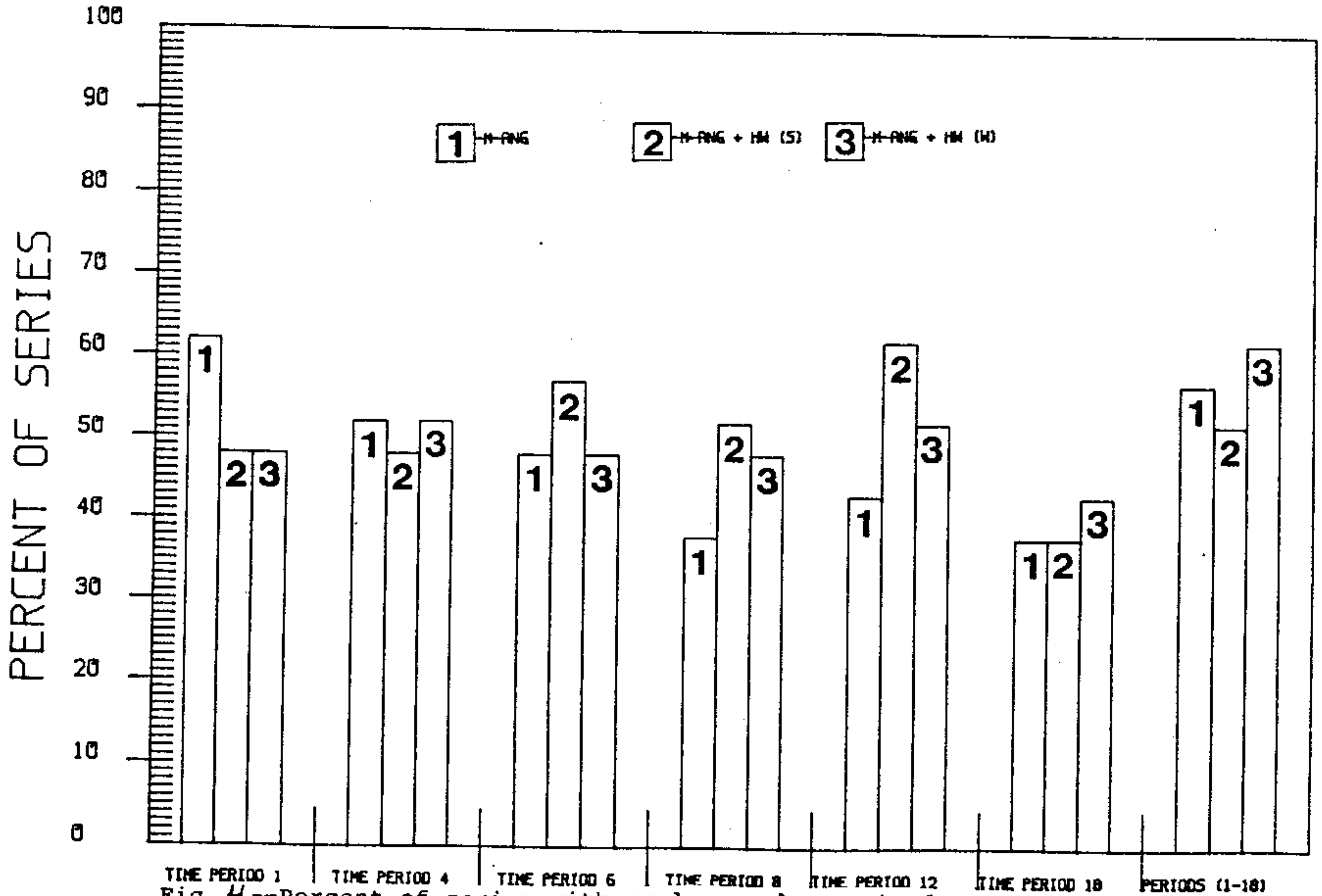


Fig. 4--Percent of series with as low or lower absolute percentage errors than the Box-Jenkins, by time horizon.

Fig. 5-- Number of Series with Lowest MAPE for Aggregate Time Horizon (1-18), by Method.

1 Box-Jenkins

2 Modified Ang

<u>Simple Averages</u>

3 Modified Ang + Holt-Winters

4 Modified Ang + Stepwise

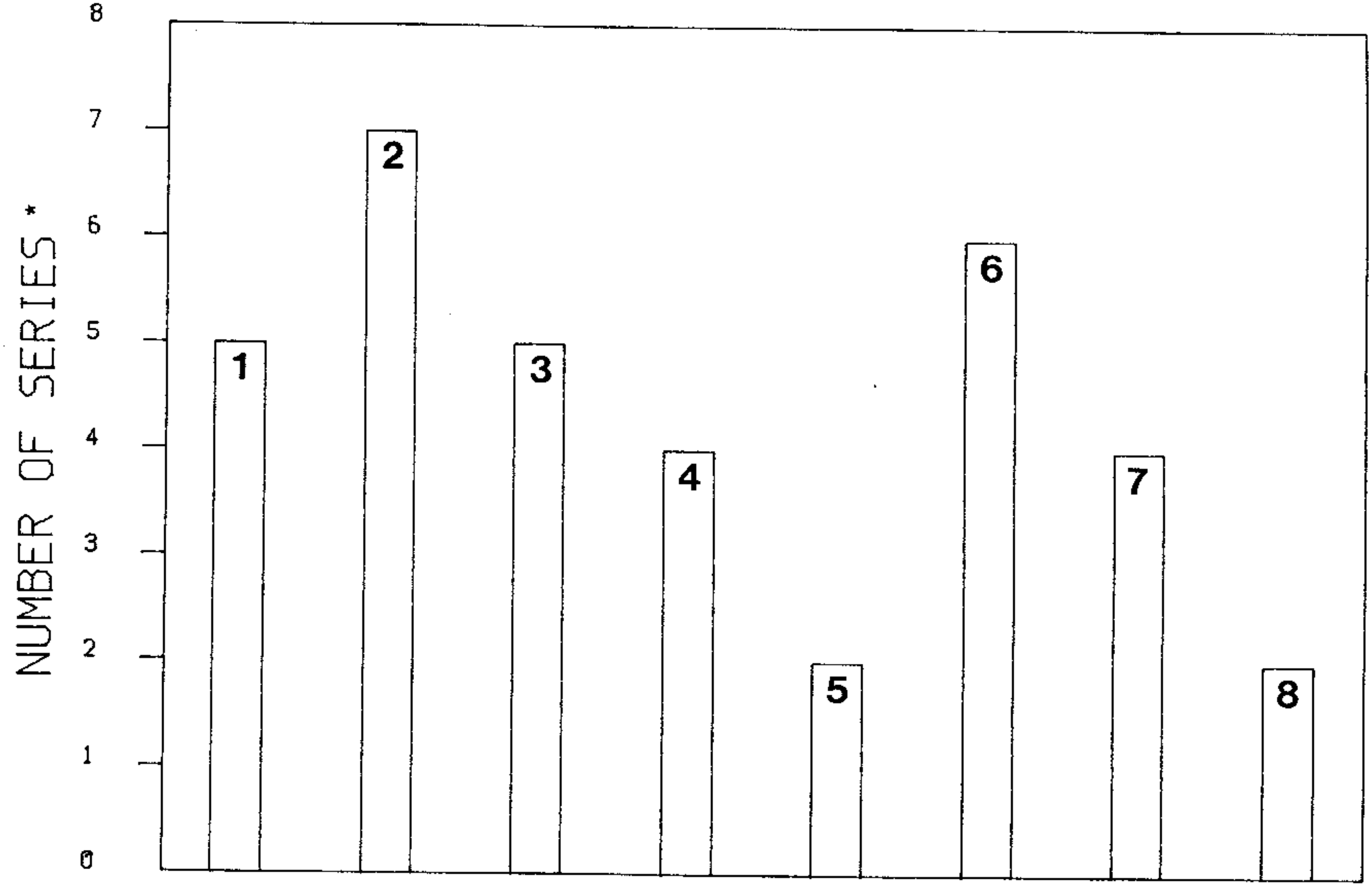
5 Modified Ang + Holt-Winters + Stepwise

<u>Weighted Averages</u>

6 Modified Ang + Holt-Winters

7 Modified Ang + Stepwise

8 Modified Ang + Holt-Winters + Stepwise



*the sum of frequencies is greater than 21 because ties were counted twice

Fig. 6-- Number of Series with Lowest MAPE for Time Horizon 1, by Method.

[1] Box-Jenkins

[2] Modified Ang

Simple Averages

[3] Modified Ang + Holt-Winters

[4] Modified Ang + Stepwise

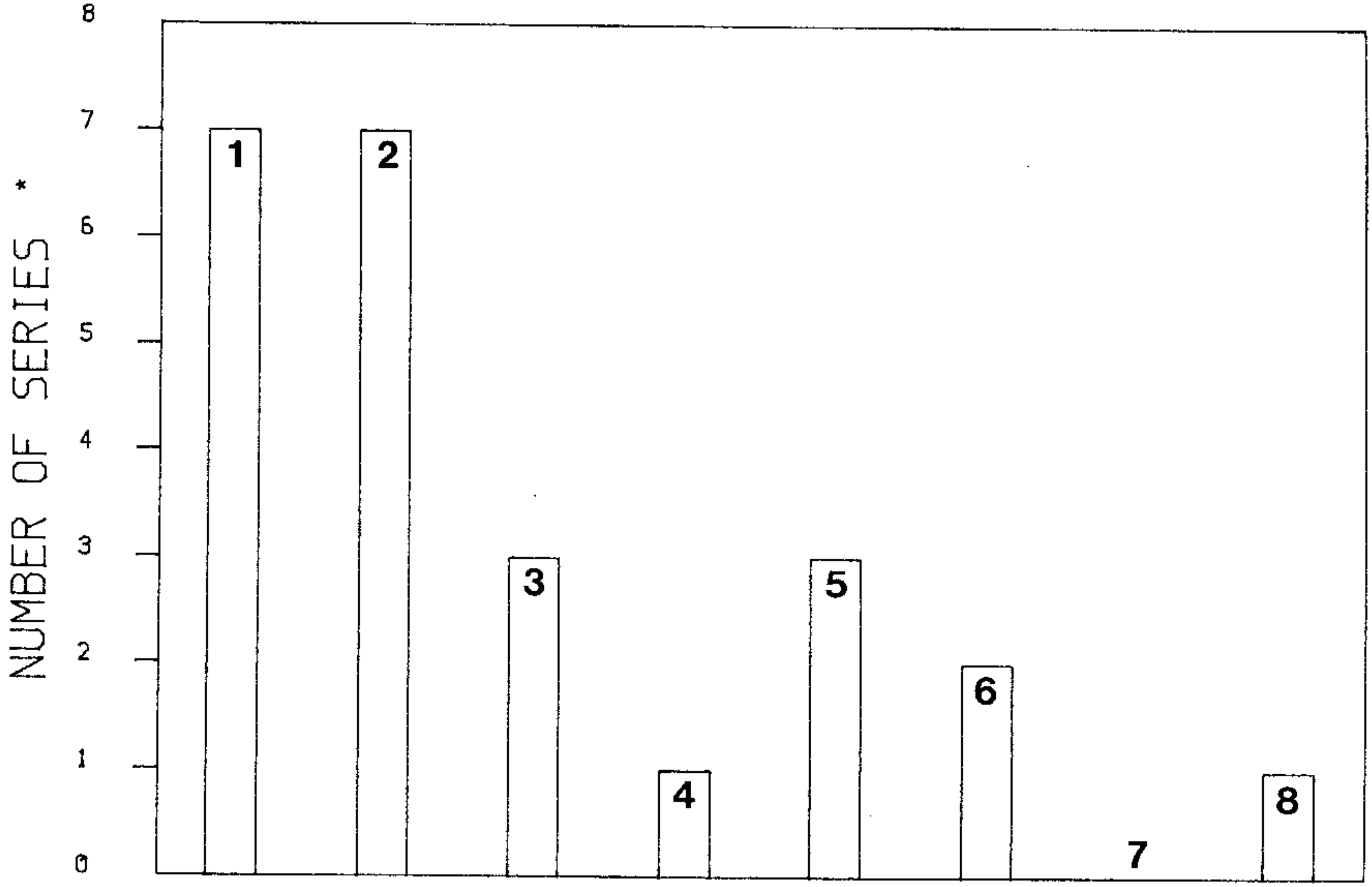
[5] Modified Ang + Holt-Winters + Stepwise

Weighted Averages

[6] Modified Ang + Holt-Winters

[7] Modified Ang + Stepwise

[8] Modified Ang + Holt-Winters + Stepwise



*the sum of frequencies is greater than 21 because ties were counted twice

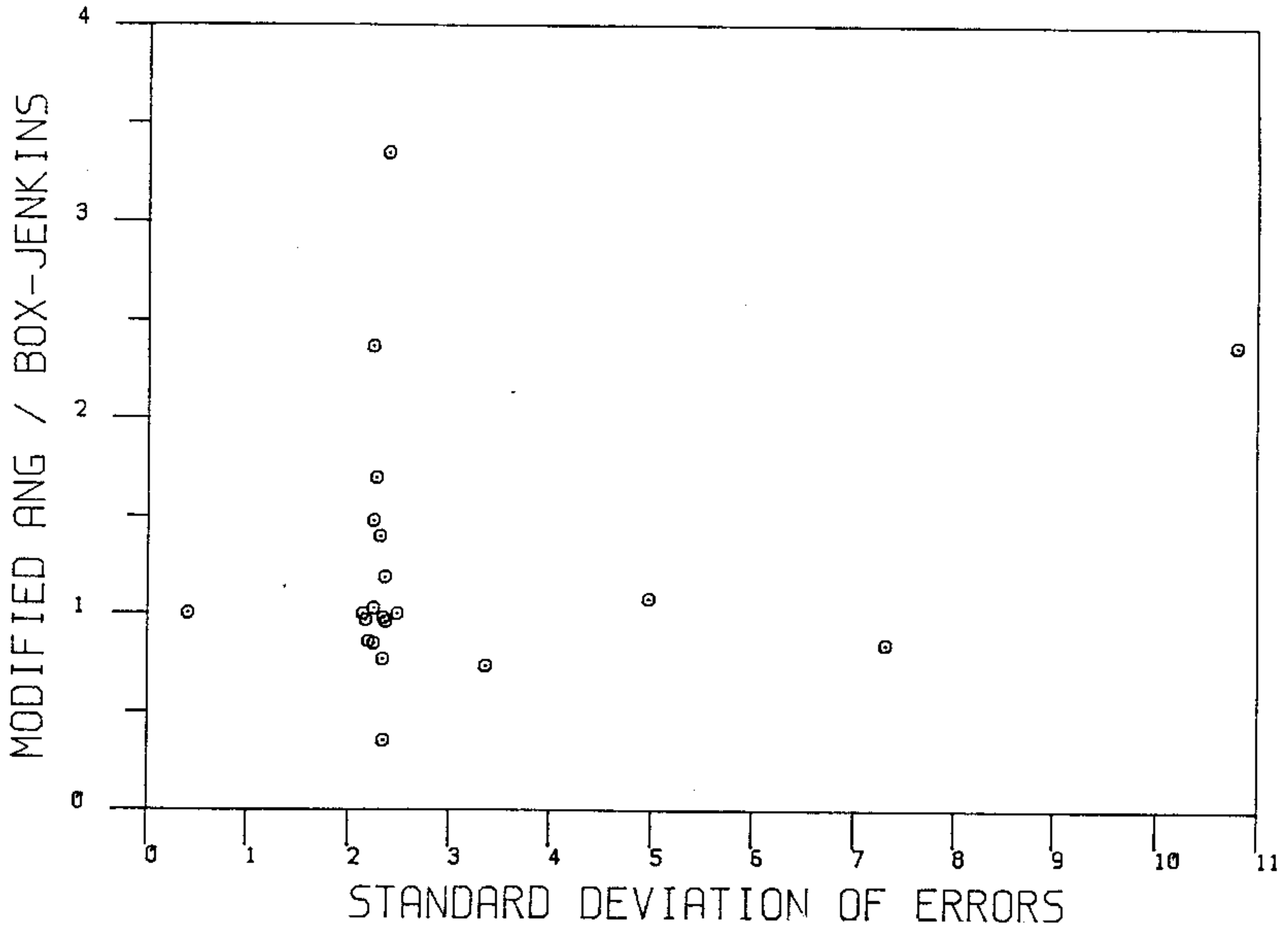


Fig. 7--Comparison between the Modified Ang and the Box-Jenkins based on standard deviation of the errors, for generated data.

TABLE III

ERRORS FOR ALL EMPIRICAL DATA (111 SETS)

Method	Median Absolute Percentage Errors				Mean Absolue Percentage Errors			
	Time Horizons				Time Horizons			
	1	4	6	all	1	4	6	all
1.Box-Jenkins	5.3	6.6	8.8	6.2	10.3	14.5	17.1	13.4
2.Modified Ang	5.4	9.2	10.8	8.1	10.8	15.5	17.9	14.4
3.Modified Ang + Holt-Winters[1]	5.1	9.0	10.7	8.3	12.0	15.2	18.1	15.1
4.Modified Ang + Stepwise Autoregressive[1]	5.6	8.8	10.4	7.8	10.6	15.9	18.7	14.8
5.all three methods[1]	5.3	9.4	9.9	8.0	10.8	14.1	17.3	14.0
6.Modified Ang Holt-Winters[2]	3.7	9.3	9.4	7.4	10.9	15.0	16.8	14.0
7.Modified Ang Stepwise Autoregressive[2]	5.6	8.8	10.3	8.2	12.2	19.2	23.5	18.0
8.all three methods[2]	5.1	9.7	9.6	8.0	12.1	18.8	22.8	17.7

[1]simple average

[2]weighted average

TABLE IV

ERRORS FOR YEARLY DATA (20 SETS)

Method	Median Absolute Percentage Errors				Mean Absolue Percentage Errors			
	Time Horizons				Time Horizons			
	1	4	6	all	1	4	6	all
1.Box-Jenkins	2.8	6.7	10.7	7.8	7.2	18.6	22.3	16.0
2.Modified Ang	3.8	11.8	20.4	9.8	7.7	17.4	21.4	15.1
3.Modified Ang + Holt-Winters[1]	3.1	9.7	12.2	7.5	6.5	16.2	16.7	13.0
4.Modified Ang + Stepwise Autoregressive[1]	4.0	8.1	17.6	8.1	7.8	17.0	20.9	14.6
5.all three methods[1]	3.0	6.9	13.3	6.5	5.7	15.4	18.0	12.4
6.Modified Ang Holt-Winters[2]	2.5	9.4	15.4	7.4	6.3	15.7	17.6	12.8
7.Modified Ang Stepwise Autoregressive[2]	4.7	11.2	21.6	9.8	12.1	23.3	30.5	20.8
8.all three methods[2]	4.6	11.4	21.5	9.3	11.7	22.6	28.9	19.9

[1] simple average [2] weighted average

TABLE V

ERRORS FOR QUARTERLY DATA (23 SETS)

Method	Median Absolute Percentage Errors					Mean Absolue Percentage Errors				
	Time Horizons					Time Horizons				
	1	4	6	8	all	1	4	6	8	all
1.Box-Jenkins	2.6	9.6	12.6	10.3	6.5	7.6	21.3	26.1	25.4	16.0
2.Modified Ang	3.2	9.5	13.7	14.7	9.5	9.1	21.9	28.3	29.9	22.9
3.Modified Ang + Holt-Winters[1]	3.6	9.0	11.9	12.3	10.2	9.1	21.9	18.3	29.9	22.9
4.Modified Ang + Stepwise Autoregressive[1]	3.7	5.3	11.9	11.9	9.2	9.9	18.9	21.7	30.3	20.8
5.all three methods[1]	3.3	9.4	15.5	15.4	9.7	8.1	16.7	21.4	19.2	17.5
6.Modified Ang + Holt-Winters[2]	3.0	9.0	14.2	15.8	10.6	8.5	20.7	24.5	24.8	20.5
7.Modified Ang + Stepwise Autoregressive[2]	3.2	6.5	11.5	11.8	9.4	12.4	25.7	28.8	46.8	28.6
8.all three methods[2]	3.8	6.2	11.4	11.6	9.5	12.3	25.5	28.0	45.8	28.1

[1] simple average [2] weighted average

TABLE VI

ERRORS FOR MONTHLY DATA (68 SETS)

Method	Median Absolute Percentage Errors							Mean Absolue Percentage Errors						
	Time Horizons							Time Horizons						
	1	4	6	8	12	18	all	1	4	6	8	12	18	all
1.Box-Jenkins	6.9	6.4	7.8	9.0	8.6	16.4	9.0	12.1	11.2	12.5	16.7	16.4	34.2	17.9
2.Modified Ang	7.5	8.3	9.0	10.5	9.9	16.9	9.9	12.4	14.1	15.7	20.1	24.7	56.6	26.5
3.Modified Ang + Holt-Winters[1]	8.5	8.4	10.0	13.1	14.6	17.3	12.0	14.6	12.7	15.1	21.0	22.8	38.8	22.1
4.Modified Ang + Stepwise Autoregressive[1]	6.5	10.6	9.1	9.1	8.1	17.7	10.0	11.7	14.6	17.1	21.2	26.2	64.6	28.7
5.all three methods[1]	6.3	10.2	9.2	10.4	12.7	15.6	10.8	13.2	12.8	15.8	21.4	23.3	49.5	24.6
6.Modified Ang + Holt-Winters[2]	5.2	8.3	8.7	10.5	10.5	15.6	10.2	13.1	12.8	14.0	21.1	23.3	47.9	23.8
7.Modified Ang + Stepwise Autoregressive[2]	6.4	9.4	8.9	9.7	8.7	19.0	10.5	12.1	15.9	19.6	23.9	31.3	75.9	33.1
8.all three methods[2]	5.7	10.6	8.0	10.4	8.7	18.7	10.5	12.1	15.4	19.3	24.8	31.1	75.3	32.8

[1] simple average   [2] weighted average

TABLE VII

ERRORS FOR SEASONAL AND NON-SEASONAL DATA

	SEASONAL SERIES								NON-SEASONAL SERIES							
	Md APES				MAPES				Md APEs				MAPES			
Time Horizons	1	4	6	all	1	4	6	all	1	4	6	all	1	4	6	all
Method																
1.Box-Jenkins	6.1	6.3	8.1	6.3	10.5	13.5	15.9	12.3	2.8	8.1	10.7	6.0	10.0	15.7	18.4	14.7
2.Modified Ang	9.0	8.0	9.2	7.9	12.3	12.8	13.9	12.1	3.0	11.3	13.7	8.8	9.1	18.7	22.7	17.1
3.Modified Ang + Holt-Winters[1]	7.9	9.9	10.5	9.0	14.1	14.7	16.2	14.8	3.6	8.5	13.6	7.1	9.5	15.8	20.3	15.5
4.Modified Ang + Stepwise Autoregressive[1]	6.2	8.7	9.6	7.0	10.8	11.9	13.6	11.2	3.6	8.9	12.8	8.6	10.4	20.7	24.8	19.0
5.all three methods[1]	6.3	9.5	9.3	8.1	12.3	12.7	14.0	12.7	3.2	8.7	12.3	7.4	9.0	15.6	21.2	15.6
6.Modified Ang + Holt-Winters[2]	5.2	8.5	9.0	7.5	12.7	12.7	13.6	12.3	2.6	9.3	12.1	7.3	8.7	17.7	20.6	16.0
7.Modified Ang + Stepwise Autoregressive[2]	5.9	8.7	8.3	6.6	10.6	12.2	14.6	11.5	3.8	8.8	13.2	9.8	14.0	27.5	34.0	25.7
8.all three methods[2]	5.3	10.1	8.0	6.9	10.6	12.0	14.2	11.4	3.8	9.3	13.0	9.6	13.8	26.7	33.0	25.1

[1]simple average

[2]weighted average

TABLE VIII

ERRORS FOR SHORT AND LONG DATA

	SHORT SERIES								LONG SERIES							
	Md APES				MAPES				Md APEs				MAPES			
Time Horizons	1	4	6	all	1	4	6	all	1	4	6	all	1	4	6	all
Method																
1. Modified Ang	4.2	12.5	16.5	10.5	9.8	21.0	21.1	17.5	8.6	5.3	7.2	6.7	11.1	12.1	12.0	10.9
2.Modified Ang_1+ Holt-Winters[1]	4.9	10.8	11.9	9.1	9.9	18.0	21.1	16.2	2.9	7.7	6.9	7.5	13.5	12.0	12.8	12.8
3.Modified Ang + Stepwise Autoregressive[1]	4.1	11.4	13.0	8.9	10.3	18.5	19.4	15.6	7.5	6.7	8.3	6.6	10.0	11.6	12.5	10.6
4.all three methods[1]	3.4	9.6	12.2	8.5	9.7	16.6	19.0	14.9	6.1	8.3	8.0	7.3	11.9	10.9	12.5	11.6
5.Modified Ang_2+ Holt-Winters[2]	2.7	11.5	11.9	8.6	9.1	18.8	17.4	15.5	4.6	5.1	7.6	6.7	12.4	10.6	12.2	11.1
6.Modified Ang + Stepwise Autoregressive[2]	3.3	13.5	11.6	10.7	11.8	20.8	25.5	23.9	6.3	8.5	7.8	6.1	9.7	12.1	13.0	11.0
7.all three methods[2]	3.3	13.4	11.9	10.3	11.6	20.2	24.6	24.8	6.0	5.5	7.4	6.3	9.8	11.3	12.8	10.8

[1] simple average   [2] weighted average

TABLE IX

LENGTHS OF EMPIRICAL TIME SERIES

Length	Number of Series
0 to 25	23
25 to 50	29
51 to 75	35
76 to 100	10
101 to 125	13
126 to 150	1
Total	111

Fig. 8--The three methods with the lowest Md APEs for one-step-ahead forecasting, by partition.

1 Box-Jenkins

2 Modified Ang

<u>Simple Averages</u>

3 Modified Ang + Holt-Winters

4 Modified Ang + Stepwise

5 Modified Ang + Holt-Winters + Stepwise

<u>Weighted Averages</u>

6 Modified Ang + Holt-Winters

7 Modified Ang + Stepwise

8 Modified Ang + Holt-Winters + Stepwise

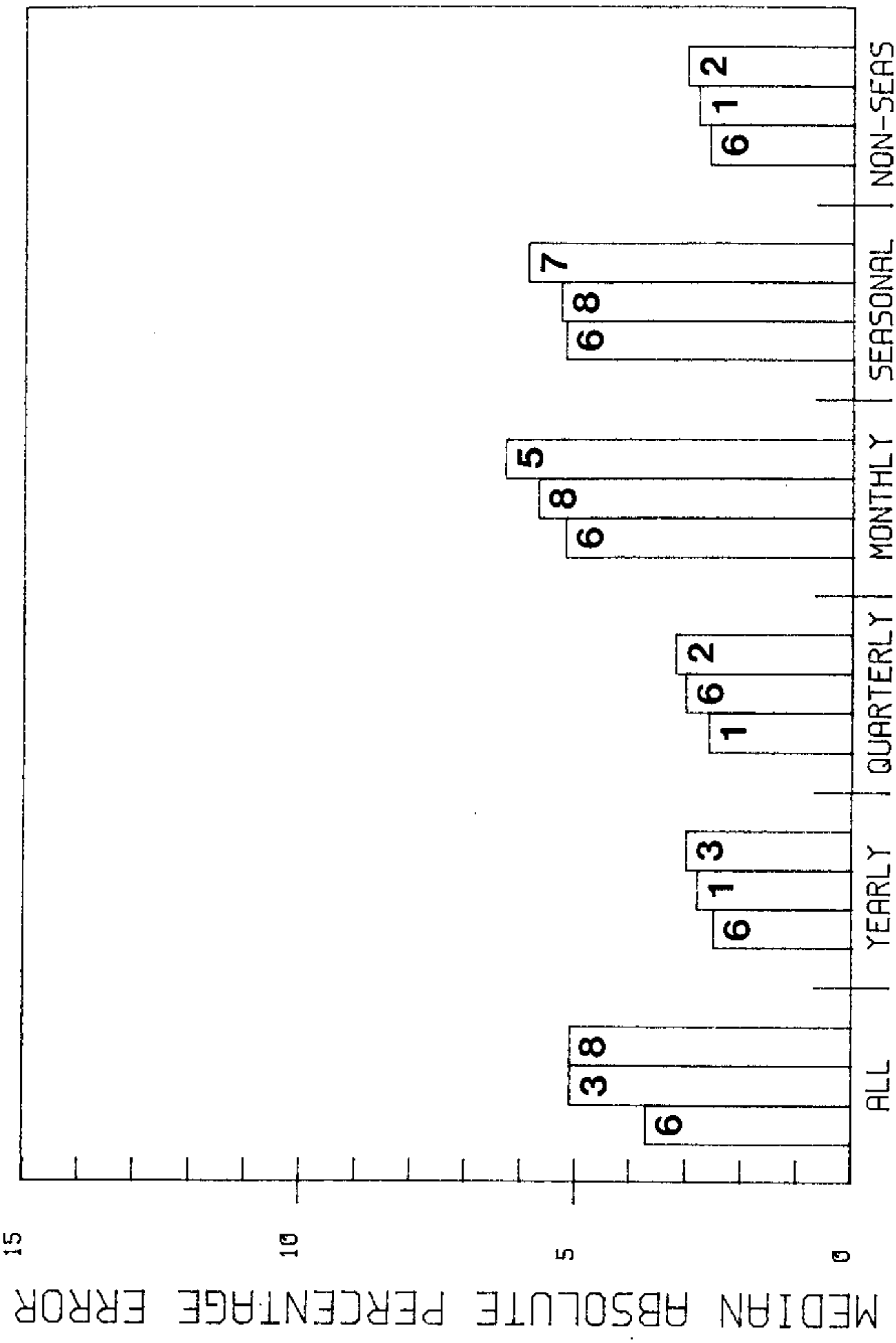

MEDIAN ABSOLUTE PERCENTAGE ERROR
0
5
10
15
ALL
6
3
8
YEARLY
6
1
3
QUARTERLY
1
6
2
MONTHLY
6
8
5
SEASONAL
6
8
7
NON-SEAS
6
1
2

Fig. 9--The three methods with the lowest MAPEs for one-step-ahead forecasting, by partition.

$\boxed{1}$ Box-Jenkins

$\boxed{2}$ Modified Ang

<u>Simple Averages</u>

$\boxed{3}$ Modified Ang + Holt-Winters

$\boxed{4}$ Modified Ang + Stepwise

$\boxed{5}$ Modified Ang + Holt-Winters + Stepwise

<u>Weighted Averages</u>

$\boxed{6}$ Modified Ang + Holt-Winters

$\boxed{7}$ Modified Ang + Stepwise

$\boxed{8}$ Modified Ang + Holt-Winters + Stepwise

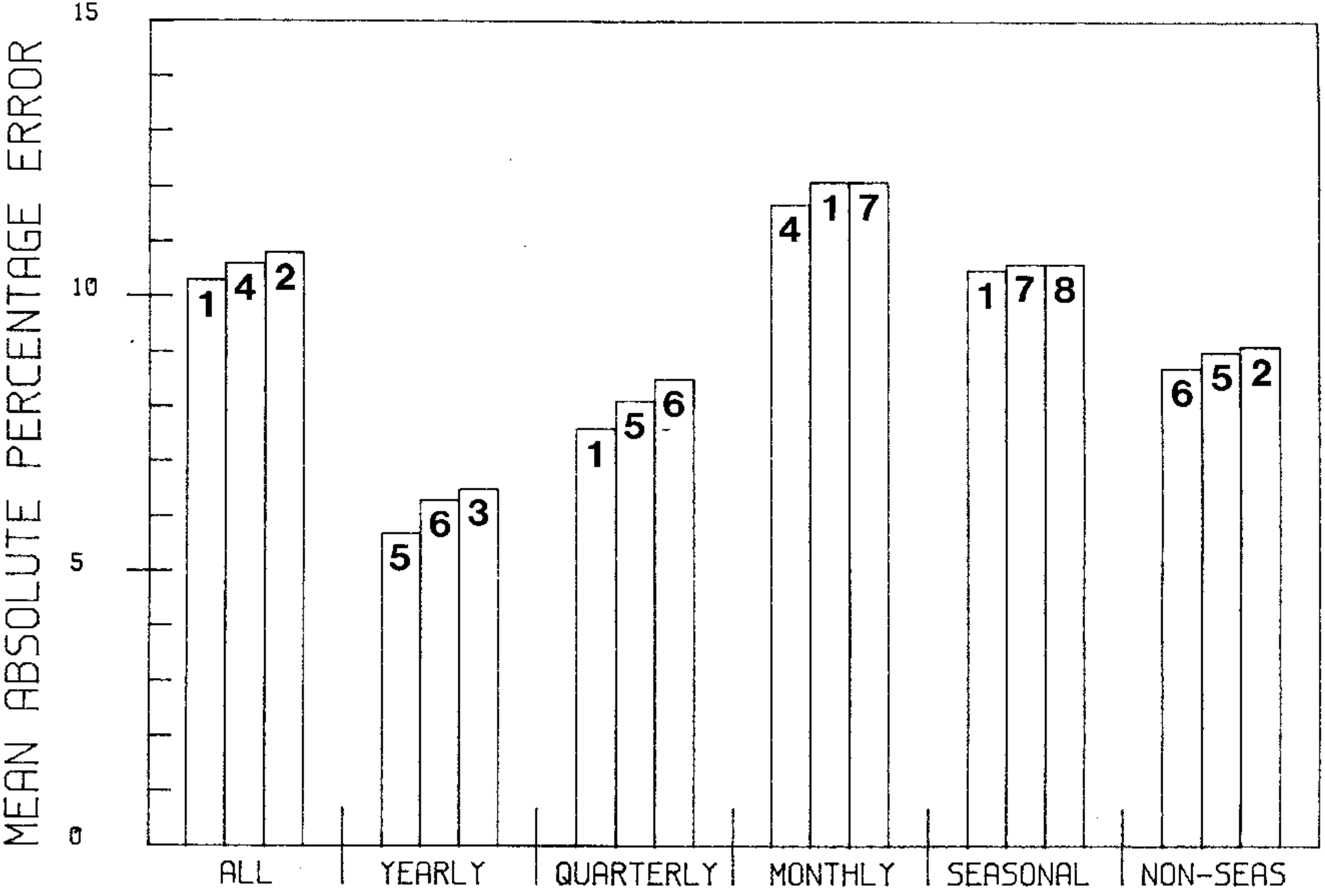

MEAN ABSOLUTE PERCENTAGE ERROR
15
10
5
0
1
4
2
5
6
3
1
5
6
4
1
7
1
7
8
6
5
2
ALL
YEARLY
QUARTERLY
MONTHLY
SEASONAL
NON-SEAS

TABLE X

FREQUENCY DISTRIBUTION OF APES FOR WEIGHTED AVERAGES OF MODIFIED ANG PLUS HOLT-WINTERS IN ONE-STEP-AHEAD FORECASTING

INTERVAL	FREQUENCY	PERCENT	CUMULATIVE %
0.000 TO 5.000	62	55.86	55.86
5.001 TO 10.000	12	10.81	66.67
10.001 TO 15.000	9	8.11	74.78
15.001 TO 20.000	8	7.21	81.99
20.001 TO 25.000	9	8.11	90.10
25.001 TO 30.000	1	.90	91.00
30.001 TO 35.000	2	1.80	92.80
35.001 TO 40.000	2	1.80	94.60
40.001 TO 45.000	1	.90	95.50
45.001 TO 50.000	0	0.00	95.50
50.001 TO 55.000	2	1.80	97.30
55.001 TO 60.000	1	.90	98.20
over 60.000	2	1.80	100.00
TOTAL	111	100.0	100.00

APPENDIX B

```
C*********************** MODIFIED ANG ****************************
C
C   THIS SUBROUTINE DRIVES THE T-SERIES PROGRAM THUS
C   PRODUCING THE MODIFIED ANG
C   PRINTS OUT FILENAME, SSE FOR LAST 12 OR N/2+1 FOR N<22 AND FORECASTS
C
C*****************************************************************
C
      SUBROUTINE AUT
      COMMON/E90/WCOF(15),ACOF(15),UCOF(15),K2,K3,K4,K6,K7,K8,K10,K10P
      COMMON/E91/KPORD(15),KQORD(15),KPJ1(15),KPJ2(15),KQJ1(15),KQJ2(15)
      COMMON/E92/KUORD(15),NPX,MAXP,MAXQ,MAXU,LCOND,LNC,LP1,LP2,LQ1,LQ2
      COMMON/E93/PA(13),IORDPA(13),ICONS(13),INC(4),NRPOIN,NRP,NPRAM
      COMMON/E94/EMAX,WSD,XSAVE,NPRAMU,SCALE,LNMEAN,LNBFT
      COMMON/E95/KTP,IEOPT,ESCALE,MAXIT,ICON,LWARN,XVAT
      COMMON/F60/XMU,XLAM,XMTRAN,Z(400),NPOINT,MBORD,NRDIFF,NSDIFF,ISPER
      COMMON/F61/NORG,NORGX,IORG(15)
      COMMON/G50/ITITLE(76),LPRINT(12),LPLOT(12),KLS,KUS,NCPOIN,IUP,THTA
      COMMON/G51/MAXTER,MAXPON,MAX1,PERFST,PERCPA,LFCT,LMOK,NFCST,LFREE
      COMMON/G52/IPUNIT,IP,LPUNCH,LGRID,NACORR,NPACOR
      COMMON/GRD1/KG1,DG1,KG2,DG2,ICH,LTIME,LDUMP,NNEG,NPOS
      COMMON/IO/INPUT,IOUT
      COMMON/MMM/M(80),IB
      COMMON/SET1/XRART(13),XSART(13),XRMAT(13),XSMAT(13),XMUEST
      COMMON/SET2/IRART(13),ISART(13),IRMAT(13),ISMAT(13)
      COMMON/SET4/IFORM(76),EK(13),X(13),XG(40),LEST,LIDEN,LPEST,IVENT
      COMMON/SER/WS(100),W(300),AS(100),A(400),ES(100),E(300),SSQ,SIGMA2
      COMMON/SSS/CZERO,R(300),P(300),STDERR(300),ZSAVE(2099)
      COMMON/V80/VV(14,14),SDER,RMEAN
      COMMON/ZZZ/SERIES(300)
      DIMENSION INUM(11)
      COMMON/AUTO/ISTEP,ITYPE,IINUM(11),I,IDIF,OPTAIC,OPTR,J11,J12,
     1JQ,JP,OPTP,AIC,OPTDIF,OPTQ,SAVESS,SAVEN
      LOGICAL LIDEN,LEST,LPEST,LPRINT,LPLOT,LFCT,LMOK,LNMEAN,LNBFT
      LOGICAL LPUNCH,LGRID,LFREE,LCOND,LTIME,LDUMP
      DATA INUM/1H0,1H1,1H2,1H3,1H4,1H5,1H6,1H7,1H8,1H9,1H /
      IF((ISTEP.EQ.4).OR.(ISTEP.EQ.18))GO TO 2121
      IE=ITYPE/100
      ITYPE=ITYPE-IE*100
 2121 DO 101 J=1,76
  101 M(J)=INUM(11)
      DO 912 J=1,11
  912 IINUM(J)=INUM(J)
      GO TO (1,2,3,4,5,6,7,8,1,10,11,12,13,14,15,
     116,17,18,19,20,21,22,23,24),ISTEP
    1 I=2
      OPTAIC=99999999
      SAVESS=99999999
      IDIF=0
      READ(5,100)F1,F2,N1,N2,N3,ITYPE
  100 FORMAT(2A3,2X,3I1,3X,I2)
      SAVEN=100*N1+10*N2+N3
```

```
      N4=SAVEN+0.1
      WRITE(8,834)F1,F2,N4
  834 FORMAT(2A3,'  DATA FILE'/I3,'  N')
      M(2)=INUM(N1+1)
      M(3)=INUM(N2+1)
      M(4)=INUM(N3+1)
      ISTEP=2
      RETURN
    2 IF(IDIF.EQ.0)GO TO 3
C
C******* DIF
C
      ISTEP=3
      I=3
      M(2)=INUM(IDIF+1)
      RETURN
    3 IF(ITYPE.EQ.1)GO TO 5
C
C******* DATA IS NOT ANNUAL DATA
C
      I=19
      ISTEP=4
      RETURN
    4 IF(ITYPE.GT.100)GO TO 8104
      GO TO 5
 8104 IE=ITYPE/100
      ITYPE=ITYPE-IE*100
      GO TO 104
 1300 FORMAT(' 0 SEASONAL DIF')
  104 I=5
 1200 FORMAT(I2,' SEASONAL DIF')
      ISTEP=5
      IF(ITYPE.EQ.4)GO TO 105
      M(2)=INUM(2)
      M(3)=INUM(11)
      M(4)=INUM(2)
      M(5)=INUM(3)
      RETURN
  105 M(2)=INUM(2)
      M(3)=INUM(11)
      M(4)=INUM(5)
      RETURN
    5 J11=0
  203 J11=J11+1
      L3=SAVEN+0.1
      L3=(L3-4)/2
      IF(L3.GT.6)L3=6
      IF(J11.GT.L3)GO TO 301
C
C****** BEST AR MODEL NOT DETERMINED
C
      J12=0
```

```
    6 J12=J12+1
      IF(J12.GT.J11)GO TO 201
C
C***** CONTINUE TO DEFINE MODEL
C
      I=10
      IF(J12.GE.10)GO TO 202
      M(2)=INUM(J12+1)
      ISTEP=6
      RETURN
  202 M(2)=INUM(2)
      M(3)=INUM(J12+1-10)
      ISTEP=6
      RETURN
  201 I=21
      ISTEP=7
      RETURN
    7 S92=SIGMA2
      AIC=SAVEN*ALOG(S92)+2.0*J11
      WRITE(6,1100)SAVEN,S92,AIC,J11,IDIF
 1100 FORMAT(' N= ',F5.1,'  VARIANCE = ',E15.7,' AIC = ',E15.7,
     1' P=',I2,' D=',I2,' Q= 0')
      I=23
      ISTEP=17
      RETURN
   17 IF(AIC.GE.SAVESS)GO TO 203
      SAVESS=AIC
      OPTR=J11
      GO TO 203
  301 JQ=0
      IF(SAVESS.GE.OPTAIC)GO TO 15
      OPTAIC=SAVESS
      OPTP=OPTR
      OPTQ=0
      OPTDIF=IDIF
   15 CONTINUE
  502 JQ=JQ+1
      IF(JQ.GT.OPTR)GO TO 501
C
C****** ALL MODELS NOT FINISHED
C
      JP=-1
  512 JP=JP+1
      IF(JQ+JP.GT.OPTR)GO TO 502
      J12=0
   11 J12=J12+1
      IF(J12.GT.JQ)GO TO 503
C
C******* MA TERMS NOT ALL DEFINED
C
      I=8
      IF(J12.GE.10)GO TO 504
```

```
      M(2)=INUM(J12+1)
      ISTEP=11
      RETURN
  504 M(2)=INUM(2)
      M(3)=INUM(J12+1-10)
      ISTEP=11
      RETURN
  503 IF(JP.EQ.0)GO TO 13
C
C****** ARMA NOT DEFINED
C
      J12=0
   12 J12=J12+1
      IF(J12.GT.JP)GO TO 13
      I=10
      IF(J12.GE.10)GO TO 604
      M(2)=INUM(J12+1)
      ISTEP=12
      RETURN
  604 M(2)=INUM(2)
      M(3)=INUM(J12+1-10)
      ISTEP=12
      RETURN
   13 I=21
      ISTEP=14
      RETURN
   14 S92=SIGMA2
      AIC=SAVEN*ALOG(S92)+2.0*(JP+JQ)
      WRITE(6,1400)SAVEN,S92,AIC,JP,IDIF,JQ
 1400 FORMAT(' N= ',F5.1,'  VARIANCE = ',E15.7,'  AIC = ',E15.7,
     1' P=',I2,' D=',I2,' Q=',I2)
      I=23
      ISTEP=16
      RETURN
   16 CONTINUE
      IF(AIC.GE.OPTAIC)GO TO 512
      OPTAIC=AIC
      OPTDIF=IDIF
      OPTP=JP
      OPTQ=JQ
      GO TO 512
  501 WRITE(6,300)OPTP,OPTDIF,OPTQ,OPTAIC
  300 FORMAT(3F4.1,E15.7,'    CHANGE D')
      IDIF=IDIF+1
      SAVESS=99999999
      IF(IDIF.GT.2)GO TO 111
      I=15
      ISTEP=2
      RETURN
C
C******* REPORT
C
```

```
  111 JP=OPTP
      JQ=OPTQ
      JPTDIF=OPTDIF+0.1
      WRITE(8,3173)JP,JPTDIF,JQ
 3173 FORMAT(3I3,'  P D Q')
      I=15
      ISTEP=8
      RETURN
    8 IDIF=OPTDIF
      IF(IDIF.EQ.0)GO TO 10
C
C***** DIF
C
      ISTEP=10
      I=3
      M(2)=INUM(IDIF+1)
      RETURN
   10 WRITE(8,1311)ITYPE
 1311 FORMAT(I2,'  SEASONS')
      IF(ITYPE.EQ.1)WRITE(8,1300)
      IF(ITYPE.EQ.1)GO TO 19
C
C***** CHECK FOR SEASONALITY
C
      I=19
      ISTEP=18
      RETURN
   18 IF (ITYPE.GT.100)GO TO 7492
      WRITE(8,1300)
      GO TO 19
 7492 IE=ITYPE/100
      ITYPE=ITYPE-IE*100
  904 I=5
      WRITE(8,1200)ITYPE
      ISTEP=19
      IF(ITYPE.EQ.4)GO TO 905
      M(2)=INUM(2)
      M(3)=INUM(11)
      M(4)=INUM(2)
      M(5)=INUM(3)
      RETURN
  905 M(2)=INUM(2)
      M(3)=INUM(11)
      M(4)=INUM(5)
      RETURN
   19 J12=0
      IF(JQ.EQ.0)GO TO 803
   20 J12=J12+1
      IF(J12.GT.JQ)GO TO 803
C
C***** MA TERMS NOT ALL DEFINED
C
```

```
      I=8
      IF(J12.GE.10)GO TO 804
      M(2)=INUM(J12+1)
      ISTEP=20
      RETURN
  804 M(2)=INUM(2)
      M(3)=INUM(J12+1-10)
      ISTEP=20
      RETURN
  803 IF(JP.EQ.0)GO TO 813
C
C***** ARMA NOT DEFINED
C
      J12=0
   21 J12=J12+1
      IF(J12.GT.JP)GO TO 813
      I=10
      IF(J12.GE.10)GO TO 1004
      M(2)=INUM(J12+1)
      ISTEP=21
      RETURN
 1004 M(2)=INUM(2)
      M(3)=INUM(J12+1-10)
      ISTEP=21
      RETURN
  813 I=12
      ISTEP=22
      M(2)=INUM(2)
      M(3)=INUM(9)
      RETURN
   22 I=34
      N=SAVEN+0.1
      NSSE=12
      IF(N.LT.22)NSSE=N/2+1
      N=N-NSSE
      N1=N/100
      N=N-N1*100
      N2=N/10
      N3=N-N2*10
      M(2)=INUM(N1+1)
      J84=3
      IF(N1.EQ.0)J84=2
      M(J84)=INUM(N2+1)
      J84=J84+1
      IF(N1+N2.EQ.0)J84=2
      M(J84)=INUM(N3+1)
      ISTEP=23
      RETURN
   23 I=21
      ISTEP=24
      RETURN
   24 I=24
```

```
      ISTEP=1
      RETURN
      END
*EOF
```

Imbedded Modifications of T-Series to Perform Modified Ang

1. All dimension declarations of 5 were changed to 13
   for example: COMMON/SET1/XRART(13),XSART(13),XMART(13)

2. The following common statement was added to the main program and to all subroutines:

```
      COMMON/AUTO/ISTEP,IITYPE,IINUM(11),I,IDIF,OPTAIC
      ,OPTR,J11,J12,JQ,JP,OPTP,AIC,OPTDIF,OPTQ,SAVESS,
      SAVEN
```

3. All write statements were removed.

   for example: WRITE(IOUT,2557) MAXPON,MAXTER

4. The following statement was inserted between statements TSR00470 and TSR00480:

```
      ISTEP=1
```

5. The following statements replace statements TSR03690 thru TSR03930

```
      N37=IITYPE+2
      DO 33 I=1,N37
      IF(ABS(G(I)).LT.0.00001)GO TO 33
93    T=S(I)/G(I)
C*    WRITE(IOUT,63)IR,T
63    FORMAT(1X,I3,3X,3F8.3,2X,1HI,52A1,1HI)
      IF((I.GE.IITYPE).AND.(I.LE.IITYPE+2).AND.
      (I.GE.4 ).AND.(T.GE.2.0))IITYPE=IITYPE+100
```

6. The following statements replace statements TSR05660 thur TSR05670

```
      NACORR=14
      IF(SAVEN.LT.20.0)NACORR=6
      NPACOR=NACORR
```

7. This statement replaces statement TSR05930

```
      MAXIT=10
```

8. The following statements replace statements TSR06060 thru TSR06190

```
1002  CALL AUT
      WRITE(6,1804)IGO(I),(M(J),J=1,10)
1804  FORMAT(A4,10A1)
```

9. This statement replaces statement TSR06610:

```
      CALL ICHECK(NPOINT,9,MAXPON,0,1008)
```

10. The following statements replace statement TSR07020:

```
      NACORR=14
      IF(SAVEN.LT.20.0)NACORR=6
```

11. The following statments replace statement TSR07100

```
      NPACOR=14
      IF(SAVEN.LT.20.0)NPACOR=6
```

12. This statement replaces statement TSR11120:

```
      CALL ICHECK(NRPOIN,05,MAXPON,0,5001)
```

13. This statement replaces statement TSR13460:

```
      CALL ICHECK(K,1,13,0,KTO+2030)
```

14. This statement replaces statement TSR35270:

```
      IF(KORG.GT.KUS.OR.KORG.LT.MKORD+KLS) GO TO 999
```

15. This statement replaces statement TSR36280:

```
      WRITE(8,103)JJ,ZV(J),ACCT(J)
```

16. This statement replaces statement TSR36310:

```
707   WRITE(8,301)JJ,ZV(J)
```

```
C*********************** STEPWISE AUTOREGRESSION ****************
C
C     FITS AN AUTOREGRESSIVE MODEL USING BACKWARD ELIMINATION
C     OR (N-2)/2 (WHICHEVER IS LESS) ADN THROWS OUT TERMS
C     BASED ON A PARTIAL F TEST
C
C     PRINTS OUT FILE NAME, SEE FOR LAST 12 (OR N/2 + 1 FOR N <22)
C     AND FORECASTS
C*****************************************************************
C
      DIMENSION DATA(300),Y(300),BETA(14),LAG(13),IRLAG(13),
     1YHAT(300),X(300,14),ILAG(13)
C
C*************** INPUT DATA *******************
C
      READ(5,100)F1,F2,N
  100 FORMAT(2A3,2X,I3)
      READ(5,200)(DATA(I),I=1,N)
  200 FORMAT(6F12.0)
C
C**************** INITIALIZE *****************
C
      L=(N-2)/2
      IF(L.GT.13)L=13
  113 LL=L
      N=N-LL
      DO 1 I=1,LL
    1 LAG(I)=I
      DO 2 I=1,N
    2 Y(I)=DATA(I+LL)
      DO 3 I=1,N
    3 X(I,1)=1.0
C
C************* GENERATE FULL MODEL ***************
C
      DO 4 I=1,L
      II=I+1
      K=LAG(I)
      DO 4 J=1,N
      JJ=J+LL
    4 X(J,II)=DATA(JJ-K)
C
C*************** FIT FULL MODEL ******************
C
      L1=L+1
      CALL FIT(Y,X,N,L1,R2F,BETA,YHAT)
C
C*************** GENERATE RESTRICTED MODELS ********
C
C***  CHECK TO SEE IF FINISHED
C
   10 IF(L.EQ.1)GO TO 11
```

```
      R2RL=-99999999.0
      DO 6 III=1,L
      II=1
      DO 75 I=1,L
      IF(I.EQ.III)GO TO 75
      II=II+1
      IIM1=II-1
      ILAG(IIM1)=LAG(I)
      K=LAG(I)
      DO 5 J=1,N
      JJ=J+LL
    5 X(J,II)=DATA(JJ-K)
   75 CONTINUE
C
C************** FIT RESTRICTED MODEL ***************
C
      CALL FIT(Y,X,N,L,R2R,BETA,YHAT)
C  LARGEST R SQUARED?
      IF(R2R.LE.R2RL)GO TO 6
      R2RL=R2R
      LM1=L-1
      DO 8 I=1,LM1
    8 IRLAG(I)=ILAG(I)
    6 CONTINUE
C
C****************** COMPUTE F *****************
C
      PF=N-L-2
      F=PF*(R2F-R2RL)/(1.0-R2F)
C
C *** CHECK TO SEE IF FINISHED
C
      IF(F.GT.3.00)GO TO 11
      L=L-1
      DO 9 I=1,L
    9 LAG(I)=IRLAG(I)
      R2F=R2RL
      GO TO 10
C
C****************** SUMMARIZE RESULTS ***************
C
C***  REFIT BEST EQUATION
C
  500 FORMAT(13I3)
   11 DO 14 I=1,L
      II=I+1
      K=LAG(I)
      DO 14 J=1,N
      JJ=J+LL
   14 X(J,II)=DATA(JJ-K)
      L1=L+1
      CALL FIT(Y,X,N,L1,R2,BETA,YHAT)
```

```
C  OUTPUT SUM OF SQUARES FOR LAST NSSE
      NSSE=12
      IF(N+LL.LT.22)NSSE=(N+LL)/2+1
      NSSEM1=NSSE-1
      NM11=N-NSSEM1
      SS12=0.0
      DO 25 I=NM11,N
   25 SS12=SS12+(Y(I)-YHAT(I))**2
  400 FORMAT(E12.6)
C  FORECAST
      N1=N+1
      N18=N+18
      DO 13 I=N1,N18
      IJ=I+LL
      Y(I)=BETA(1)
      DO 13 J=1,L
      K=LAG(J)
   13 Y(I)=Y(I)+BETA(J+1)*Y(I-K)
      DO 112 I=N1,N18
      IF(Y(I).GT.-100.0)GO TO 112
      L=LL-1
      GO TO 113
  112 CONTINUE
      WRITE(6,100)F1,F2
      WRITE(6,500)(LAG(I),I=1,L)
      WRITE(6,400)SS12
      WRITE(6,300)(Y(I),I=N1,N18)
  300 FORMAT(6F12.2)
      STOP
      END
C
C********* SUBROUTINE TO FIT MODEL ****************
C
C
      SUBROUTINE FIT(Y,X,N,L,R2,BETA,YHAT)
      DIMENSION Y(300),X(300,14),XTX(14,14),XTY(14),BETA(14),YHAT(300)
C
 C*** X TRANSPOSE X
C
      DO 1 I=1,L
      DO 1 J=1,L
      XTX(I,J)=0.0
      DO 1 K=1,N
    1 XTX(I,J)=XTX(I,J)+X(K,I)*X(K,J)
C
C*** X TRANSPOSE Y
C
      DO 2 I=1,L
      XTY(I)=0.0
      DO 2 K=1,N
    2 XTY(I)=XTY(I)+X(K,I)*Y(K)
C
```

```
C*** X TRANSPOSE X INVERSE
C
      CALL INV(XTX,L)
C
C*** BETA
C
      DO 3 I=1,L
      BETA(I)=0.0
      DO 3 K=1,L
    3 BETA(I)=BETA(I)+XTX(I,K)*XTY(K)
C
C*** Y HAT
C
      DO 4 I=1,N
      YHAT(I)=0.0
      DO 4 K=1,L
    4 YHAT(I)=YHAT(I)+X(I,K)*BETA(K)
C
C*** SUM OF Y
C
C*** SUM OF Y SQUARED
C
C*** SUM OF Y - Y HAT SQUARED
C
      SUMY=0.0
      SUMY2=0.0
      SYMYH2=0.0
      DO 5 I=1,N
      SUMY=SUMY+Y(I)
      SUMY2=SUMY2+Y(I)*Y(I)
    5 SYMYH2=SYMYH2+(Y(I)-YHAT(I))**2
C
C*** R SQUARED
C
      R2=1.0-SYMYH2/(SUMY2-SUMY*SUMY/N)
      RETURN
      END
C
C*********** SUBROUTINE TO INVERT X, WHICH IS K BY K **********
C
      SUBROUTINE INV(X,K)
      DIMENSION X(14,14)
      DO 630 I=1,K
      X(I,I)=1.0/X(I,I)
      DO 540 J=1,K
      IF(J.EQ.I)GO TO 540
      X(I,J)=X(I,J)*X(I,I)
  540 CONTINUE
      DO 620 J=1,K
      IF(J.EQ.I)GO TO 620
      DO 600 KK=1,K
      IF(KK.EQ.I)GO TO 600
```

```
      X(J,KK)=X(J,KK)-X(J,I)*X(I,KK)
  600 CONTINUE
      X(J,I)=-X(J,I)*X(I,I)
  620 CONTINUE
  630 CONTINUE
      RETURN
      END
*EOF
```

```
C********************** HOLT- WINTERS PROGRAM ****************
C
C   USES HOLT WHEN # SEASONAL DIFFERENCES FOUND TO BE ZERO
C   OTHERWISE USES WINTERS 3 PARAMETER MODEL
C
C    PRINTS OUT FILE NAME, SSE FOR LAST 12 (OR N/2 + 1 FOR N < 22)
C    AND FORECASTS
C
C**************************************************************
C
      INTEGER T
      DIMENSION Y(200),S(200),B(200),PT(12),QT(12),SD(12),D(200)
      READ(5,100) Q1,Q2,N
      N1=N+1
      NM1=N-1
      SBT=1
  100 FORMAT(A4,A2,I5)
      DO 1 IJ=2,N1,6
    1 READ(5,200) Y(IJ),Y(IJ+1),Y(IJ+2),Y(IJ+3),Y(IJ+4),Y(IJ+5)
  200 FORMAT(6F12.3)
      SM=10.0**25
      N19=N+19
      N2=N+2
C
C*** FOR SAMPLE SIZES LESS THAN 22 (SSE BASED ON N/2 + 1)
C
      NSSE=12
      IF (N.LT.22) NSSE = N/2 + 1
      NM11=N-NSSE+1
      NM10=N-NSSE+2
      NM11=N-11
      NM10=N-10
      SSX=0
      SSHW=0
      NM2=N-2
C
C*** READ INFO FROM T.REPORT
C
      READ(7,100)QT1,QT2
      READ(7,500)NT
  500 FORMAT(I3)
      READ(7,700)
      READ(7,700)IS
      READ(7,700)NS
  700 FORMAT(I2)
C*** CHECK IF DATA IS SEASONAL OR NOT (SPIKE IN AUTOCORRELATION)
      IF(NS.NE.0)GOTO 99
C
C******************** HOLT PROCEDURE **********************
C   GRID SEARCH FOR BEST ALPHA AND GAMMA
C
      DO 2 IA=1,10
```

```
      A=.1*IA
      DO 2 IG=1,10
      G=.1*IG
C
C*** BACKFORECASTING TO DETERMINE INITIAL SMOOTHED VALUES AND SLOPE
C
      S(N)=Y(N)
      B(N)=Y(N)-Y(N1)
      DO 3 I=1,NM2
      T=N-I
      S(T) = A*Y(T) +(1-A)*(S(T+1) + B(T+1))
    3 B(T)=G*(S(T)-S(T+1))+(1-G)*B(T+1)
      S(1) = S(2) + B(2)
      SE=0
      B(1)=-B(2)
C
C*** NOW THE MODEL FIT WILL BE DONE IN THE USUAL FORWARD MANNER
C
      DO 4 T=2,N1
      S(T)=A*Y(T)+(1-A)*(S(T-1)+B(T-1))
      B(T)=G*(S(T)-S(T-1))+(1-G)*B(T-1)
      YHAT=S(T-1)+B(T-1)
      AA=(Y(T)-YHAT)**2
      SE=SE+AA
    4 CONTINUE
C
C*** SEARCH FOR SMALLEST SSE ON FIT
C
      IF(SE.GE.SM)GOTO 2
      SM=SE
      SA=A
      SG=G
      SS1=S(1)
      SB1=B(1)
    2 CONTINUE
C
C*** REFITTING OF BEST MODEL
C
      S(1)=SS1
      B(1)=SB1
      DO 5 J = 2,N1
      S(J)=SA*Y(J)+(1-SA)*(S(J-1)+B(J-1))
    5 B(J)=SG*(S(J)-S(J-1))+(1-SG)*B(J-1)
C
C*** FORECASTING 18 PERIODS
C
      DO 7 K=N2,N19
    7 Y(K)=S(N1)+B(N1)*(K-N1)
C
C*** COMPUTING SSE OF LAST 12 FOR WEIGHTS
C
      DO 8 L=NM10,N1
```

```
      X=S(NM11)+B(NM11)*(L-NM11)
    8 SSHW=(Y(L)-X)**2 + SSHW
C
C ***************** END OF HOLT PROCEDURE
C
      GOTO 999
C
C
C*** SENDS CONTROL TO AREA FOR WRITING RESULTS
C
C******************** WINTERS PROCEDURE **********************
C
   99 NS1=NS+1
C
C*** GRID SEARCH FOR BEST ALPHA, BETA AND GAMMA VALUES
C
      DO 1002 IA=1,10
      A=.1*IA
      DO 1002 IG=1,10
      G=.1*IG
      DO 1002 IBT=1,100
      BT=.01*IBT
      DO 1001 KP=1,200
      IKP=KP-1
 1001 D(KP)=1.0
C
C*** BACKFORECASTING TO DETERMINE INITIAL VALUES
C
      S(N)=Y(N)
      B(N)=Y(N)-Y(N1)
      DO 1003 I=1,NM2
      T=N-I
      S(T)=A*(Y(T)/D(T+NS))+(1-A)*(S(T+1)+B(T+1))
      B(T)=G*(S(T)-S(T+1))+(1-G)*B(T+1)
      D(T)=BT*(Y(T)/S(T))+(1-BT)*D(T+NS)
 1003 CONTINUE
      S(1)=(S(2)+B(2))*D(1+NS)
      SE=0
      B(1)=-B(2)
C
C*** MODEL FIT IN FORWARD ORDER
C
      NS2=NS1 + 1
      DO 1004 T=NS2,N1
      S(T)=A*(Y(T)/D(T-NS))+(1-A)*(S(T-1)+B(T-1))
      B(T)=G*(S(T)-S(T-1))+(1-G)*B(T-1)
      D(T)=BT*(Y(T)/S(T))+(1-BT)*D(T-NS)
      YHAT=(S(T-1)+B(T-1))*D(T-NS)
      AA=(Y(T)-YHAT)**2
 1004 SE=SE+AA
C
C*** SEARCH FOR SMALLEST SSE ON FIT
```

```
C
      IF (SE.GE.SM) GOTO 1002
      SM=SE
      SA=A
      SG=G
      SBT=BT
      SS1=S(NS)
      SB1=B(NS)
      DO 1005 IK = 1,NS
 1005 SD(IK)=D(IK)
 1002 CONTINUE
C
C*** REFIT THE BEST MODEL FOR FORECASTING
C
      S(NS)=SS1
      S(NS+1)=SS2
      B(NS)=SB1
      B(NS+1)=SB2
      BT=SBT
      A=SA
      G=SG
      DO 1006 IK=1,NS
 1006 D(IK)=SD(IK)
      DO 1007 J=NS1,N1
      S(J)=A*(Y(J)/D(T-NS))+(1-A)*(S(J-1)+B(J-1))
      B(J)=G*(S(J)-S(J-1))+(1-G)*B(J-1)
 1007 D(J)=BT*(Y(J)/S(J))+(1-BT)*D(J-NS)
C
C*** FORECASTS FOR NEXT 18 PERIODS
C
      DO 1008 K=N2,N19
 1008 Y(K)=S(N1)+(K-N1)*B(N1)*D(K-NS)
C
C*** COMPUTING SSE OF LAST 12 MONTHS FOR WEIGHT COMPUTATION
C
      DO 1009 L=NM10,N1
      X=S(NM11)+(L-NM11)*B(NM11)*D(L-NS)
 1009 SSHW=SSHW+(Y(L)-X)**2
C
C*** END OF WINTERS PROCEDURE
C
C****************** PRINTING INFO TO OUTPUT FILE ****************
C
  999 WRITE(6,300)Q1,Q2
  300 FORMAT(A4,A2)
  400 FORMAT(E12.6)
      WRITE(6,200)SA,SG,SBT,S(N1),B(N1)
      WRITE(6,400) SSHW
      WRITE(6,200)(Y(I),I=N2,N19)
      STOP
      END
*EOF
```

BIBLIOGRAPHY

Books

Bowerman, Bruce. L and Richard T. O'Connell, Forecasting and Time Series, Belmont, California, Duxbury Press, 1979.

Box, George E.P. and Gwilym M. Jenkins, Time Series Analysis Forecasting and Control, San Francisco, Holden-Day, 1976.

Granger, C.W.J., Forecasting in Business and Economics, New York, Academic Press, 1980.

Granger, C.W.J. and Paul Newbold, Forecasting Economic Time Series, New York, Academic Press, 1977.

Hanke, John E. and Arthur G. Reitsch, Business Forecasting, Boston, Allyn and Bacon, 1981.

Kendall, M.G., Time Series Analysis, London, Griffin, 1973.

Makridakis, Spyros , Steven C. Wheelwright and Victor E. McGee, Forecasting: Methods and Applications, New York, John Wiley and Sons, 1983.

Neter, John and William Wasserman, Applied Linear Statistical Models, Homewood, Illinois, Richard D. Irwin, 1974.

Articles

Adam, E. E. and R. J. Ebert, "A Comparison of Human and Statistical Forecasting," AIITE Transactions, 8 (1976), 120-127.

Akaike, H., "Information Theory and an Extension of the Maximum Likelihood," Research Memo No. 46, Institute of Statistical Mathematics, Tokyo (1976).

Akaike, H., "Use of an Information Theortetic Quantity for Statistical Model Identification," Proceedings of the Fifth Hawaii International Conference in System Sciences (1972), 249-250.

Akaike, H., "Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models," Biometrika, 60 (1973), 255-265.

Akaike, H., "A New Look at the Statistical Model Identification," IEEE Transactions on Automated Control, 49 (1974), 716-723.

Andersen, Allan, "Viewpoint of the Box-Jenkins Analyst in the M-Competition," Journal of Forecasting, Vol.2, No.3, (1983), p. 286-288.

Andersen, Allan, "An Empirical Examination of Box-Jenkins Forecasting", Journal of the Royal Statistical Society, 145, 4 (1982), 472-478.

Anderson, O.D., "A Commentary on 'A Survey of Time Series'", International Statistical Review, 45 (1977), 273-297.

Ang, James, Jess H. Chuaa and Ali M. Fatemi, "How to Use Akaike Criterion to Automatically Specify and Construct Box-Jenkins Models," Journal of Business Forecasting, Spring (1982), 38-42.

Armstrong, J. S., "Forecasting with Econometric Methods: Folklore Vs. Fact", Journal of Business, Sl (1978), 549-600.

Bates, J. M. and C. W. J. Granger, "The Combination of Forecasts", Operational Research Quarterly, Vol. 20, No. 4 (1969), 451-468.

Bransom, M. J., "Discussion of a Paper by Dr. Newbold and Professor Granger," Journal of the Royal Statistical Society, A (1974), 137, Part 2, 157.

Bunn, John and L. Kappos, "Synthesis or Selection of Forecasting Models," European Journal of Operational Research, Vol. 9, No.2, (1982), 173-180.

Bunn, John "Comparitive evaluation of Out-Performance and Minimum Variance Procedures for Linear Synthesis of

Forecasts," Operational Research Quarterly, Vol. 28, No. 3, (1977), 653-662.

Carbone, Robert, Allan Andersen, Yvan Corriveau and Paul Piat Corson, "Comparing for Different Time Series Methods the Value of Technical Expertise, Individualized Analysis and Judgmental Adjustment," Management Science, Vol. 29, No. 5 (1983), 559-566.

Chambers, J. and P. Mullick, "How to Choose the Right Forecasting Technique," Harvard Business Review, 49, (1971), 45-74.

Chatfield, C. and Prothero, D. L., "Box-Jenkins Seasonal Forecasting: Problems in a Case Study," Journal of the Royal Statistical Society, A, 136 (1973), 295-336.

Chatfield, C., "The Holt-Winters Forecasting Procedure", Applied Statistics, 27, No. 3 (1978), 264-279.

Cleveland, and G.C. Tiao, "Box-Jenkins Seasonal Forecasting: Problems in a Case-Study," Journal of the American Statistical Association, 71, (1976), 581-587.

Dagum, Estela Bee, "Modelling Forecasting and Seasonally Adjusting Economic Time Series with the X-11 Arima Method," The Statistician, Vol. 27, Nos. 3,4 (1978), 203-216.

Davies, J. and Paul Newbold, "Forecasting with Misspecified Models", Applied Statistics, 29, No.1, (1980), 87-92.

Dickinson, "Some Statisticl Results in the Combination of Forecasts," Operational Research Quarterly, 24, (1973), 253-260.

Dickinson, "Some Comments on the Combination of Forecasts," Operational Research Quarterly, 26, (1975), 205-210.

Gardner, Everette S., "The Trade-Offs in Choosing a Time Series Method," Journal of Forecasting, Vol. 1, No. 3 (1983), 263-266.

Groff, Gene, "Empirical Comparison of Models for Short Range Forecasting," Management Science, Vol. 20, No.1, (1973), 267-285.

Guerts, Michael D., "Evaluating a Forecasting Competition," Journal of Forecasting, Vol. 2, No. 3 (1983), 267-269.

Guerts, M. and J. Reinmuth, "The Sensitivity of Model Identification in Using the Box-Jenkins Forecasting Procedure, Proceeding of the 1976 AIDS Meeting, San Francisco.

Herne, H., "Discussion of a Paper by Dr. Chatfield and Dr. Prothero," Journal of the Royal Statistical Society, A, 136, 3 (1973), 330.

Heitikko, Harri, "On the Effect of Zero Initial Estimates on the Standard Estimation of ARMA (P,Q) Models," The Statistician, Vol, 30, No.4, (1981).

Hill, G.W. and H. Fildes, "The Accuracy of Extrapolition Methods and Automatic Box-Jenkins Package Sift," Journal of Forecasting, Vol. 3, (1984), 319-323.

Hill, G. W. and D. Woodworth, "Automatic Box-Jenkins Forecasting," Journal of the Operational Research Society, Vol. 31, No. 5 (1980), 413-422.

Hopwood, J., "On the Automation of the Box-Jenkins Modelig Procedures and Algorithm with an Empirical Test", Journal of Accounting Research, Vol. 18, No. 1, (1980).

Kang, Chin-Sheng Alan, David D. Bedworth and Dwayne A. Rollier, "Automatic Identification of Autoregressive Integrated Moving Average Time Series," IIE Transactions, Vol. 14, No. 3 (1982).

Kenny P. B. and J. Durbin, "Local Trend Estimation and Seasonal Adjustment of Economic and Social Time Series," Journal of the Royal Statistical Society, A, 145, Part 1 (1982), 1-41.

Libert, G. , "Construction Automatique et Continue de Previsions," R.A.I.R.O. R.O., 16 (1982), 333-348.

Libert, G., "The M-Competition with a Fully Automatic Box-Jenkins Procedure," Journal of Forecasting, Vol. 3 (1984), 325-328.

Mahmoud, Essam, "Accuracy in Forecasting - A Survey", Journal of Forecasting, Vol.3, No. 2 (1984), 139-159.

Makridakis, Spyros, "A Survey of Time Series," International Statistical Review, 44, 1 (1976), 29-70.

Makridakis, Spyros and Michele Hibon, "Accuracy of Forecasting: An Empirical investigation," Journal of the Royal Statistical Society, A, 142, Part 2 (1979), 97-145.

Makridakis, Spyros and Robert L. Winkler, "Averages of Forecasts: Some Empirical Results," Management Science, Vol. 29, No. 9 (1983), 987-996.

Makridakis, Spyros et al., "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition," Journal of Forecasting, Vol. 1, No. 2 (1982), 111-153.

McKenzie, Ed, "A Comparison of Standard Forecasting Systems with the Box-Jenkins Approach," The Statistician, Vol. 23, No. 2 (1974), 107-116.

Meeker, William Q., "TSERIES -- A User-Oriented Computer Program for Time Series Analysis," The American Statistician, Vol. 32, No.3, (1978), 111-112.

Mentzer, F. and J. Cox, "Familiarity, Application and Performances of Sales Forecasting Techniques," Journal of Forecasting, Vol. 1, No. 2 (1984), 27-36.

Newbold, Paul, "The Competition to End All Competitions," Journal of Forecasting, Vol. 2, No. 3, (1983).

Newbold, Paul, "Some recent Developments in Time Series Analysis," International Statistical Review, 49 (1981), 53-66.

Newbold, Paul and C. W. J. Granger, "Experience with Forecasting Univariate Time Series and the Combination of Forecasts," Journal of the Royal Statistical Society, A, 137 (1974), 131-165.

Plosser, Charles, "Short-Term Forecasting and Seasonal Adjustment," Journal of the American Statistical Association, Vol. 74, No. 365, (1979), 15-24.

Shibata, Ritei, "Selection of the Order of an Autoregressive Model By Akaike's Information Criterion," Biometrika, 63, 1, (1976), 117-126.

Stern, G. J. A., "Discussion of a Paper by Dr. Chatfield and Dr. Prothero," Journal of the Royal Statistical Society, A, 136, 3 (1973), 329-330.

Stone, M., "An Aysmptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion," Journal of the Royal Statistical Society, B, Vol. 39, (1977), 44-47.

Tomasek, Otto, "Discussion of a Paper by Dr. Chatfield and Dr. Prothero," Journal of the Royal Statistical Society, A, 136, 3 (1973), 332-333.

Ward, D. H., "Discussion of a Paper by Dr. Newbold and Professor Granger," Journal of the Royal Statistical Society, A , 137 Part 2 (1974), 157.

Wheelwright, S. C. and D. G. Clark, "Corporate Forecasting: Promise and Reality," Harvard Business Review, November-December (1976).

Winkler, Robert L. and Spyros Makridakis, "The Combination of Forecasts," Journal of the Royal Statistical Society, A, 146 (1983), 150-157.

Winters, P.R., "Forecasting Sales By Exponentially Weighted Moving Averages," Management Science, 6, (1960), 324-342.

Wood, E. O., "Predicting Property Values," Land Economics, (1976).

Unpublished Materials

Box, G. E. P., G. C. Taio and D. R. Pack, "Modern Techniques in Time Series Forecasting and Dynamic Model Building," unpublished course material at Princeton University (1977).

Kang, Chin-Sheng Alan, "Identification of Autoregressive Integrated Moving Average Time Series," unpublished PhD dissertation, Department of Industrial and

Management Systems Engineering, Arizona State University, Tempe (1980).

Mabert, M. A., "Statistical Versus Sales Force - Executive Opinion Short-range Forecasts: a Time-series Analysis Case Study," Krannert Graduate School, Purdue University (working paper) (1975).

Meeker, William Q., "TSERIES -- A User-Oriented Computer Program for Identifying, Fitting, and Forecasting ARIMA Time Series Models," Iowa State University, (1977).

Payne, D. J., "The Determination of Regression Relationships Using Stepwise Regression Techniques," unpublished PhD thesis, University of Nottingham (1973).

Reid, D. J., "A Comparitive Study of Time-Series Prediction Techniques on Economic Data," unpublished PhD thesis, University of Nottingham (1969).