Use of Multiple Time Series Analysis to Forecast Maintenance Manhour Requirement for the Chemical Process Industry.

Helen Bostock
Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_disstheses/4341

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.

- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.

- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17”x 23” black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6”x 9” black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.
Use of multiple time series analysis to forecast maintenance manhour requirement for the chemical process industry

Bostock, Helen, Ph.D.
The Louisiana State University and Agricultural and Mechanical Col., 1987
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages
2. Colored illustrations, paper or print
3. Photographs with dark background
4. Illustrations are poor copy
5. Pages with black marks, not original copy ✓
6. Print shows through as there is text on both sides of page
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements
9. Tightly bound copy with print lost in spine
10. Computer printout pages with indistinct print
11. Page(s) lacking when material received, and not available from school or author.
12. Page(s) seem to be missing in numbering only as text follows.
13. Two pages numbered. Text follows.
14. Curling and wrinkled pages
15. Dissertation contains pages with print at a slant, filmed as received
16. Other

________________________________________________________________________________________
________________________________________________________________________________________

University Microfilms International
USE OF MULTIPLE TIME SERIES ANALYSIS
TO FORECAST MAINTENANCE MANHOUR REQUIREMENT
FOR THE CHEMICAL PROCESS INDUSTRY

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Interdepartmental Programs in Engineering

by

Helen Bostock
B.S., Chatham College, 1962
M.S.I.E., Louisiana State University, 1981
May 1987
CONTENTS

| LIST OF TABLES                  | iv  |
| LIST OF FIGURES                | v   |
| ABSTRACT                       | vi  |

Chapter

1. INTRODUCTION                   | 1   |
   Statement of the Problem       | 1   |
   The Research Effort           | 2   |

2. THE MAINTENANCE MANPOWER REQUIREMENT SYSTEM | 4   |
   Variables Which May Influence Preventive Maintenance | 8   |
   Variables Which May Influence Emergency Maintenance | 10  |
   Variables Which May Influence Turnaround Maintenance | 11  |
   Intervariable Relationships   | 12  |

3. A COMPARISON OF FORECASTING METHODS | 18  |
   Moving Average Models         | 20  |
   Simple Exponential Smoothing  | 22  |
   Higher Forms of Smoothing     | 25  |
   Adaptive Process of Weighting Past Observations | 31  |
   Tracking Signals               | 32  |
   Regression Analysis           | 35  |
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Initial Variable List</td>
<td>13</td>
</tr>
<tr>
<td>6.1</td>
<td>Bivariate Causality Patterns</td>
<td>89</td>
</tr>
<tr>
<td>9.1</td>
<td>Maintenance Manhours</td>
<td>103</td>
</tr>
<tr>
<td>9.2</td>
<td>Production/Capacity Ratio in Hours</td>
<td>105</td>
</tr>
<tr>
<td>9.3</td>
<td>Backlog Hours</td>
<td>107</td>
</tr>
<tr>
<td>9.4</td>
<td>Maintenance Manhours (Example 2)</td>
<td>109</td>
</tr>
<tr>
<td>9.5</td>
<td>Forecast and Actual Values for Examples 1 and 2</td>
<td>127</td>
</tr>
</tbody>
</table>
FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The Chemical Processing Plant Maintenance Manhour Requirement System</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>Hypothesized Interactions Among the Independent Variables Which May Impact Chemical Processing Plant Maintenance Manhour Requirement via Backlog</td>
<td>16</td>
</tr>
<tr>
<td>2.3</td>
<td>Hypothesized Interactions Among the Independent Variables Which May Impact Chemical Processing Plant Maintenance Manhour Requirement via Operating Level</td>
<td>17</td>
</tr>
<tr>
<td>6.1</td>
<td>Decision Process for Forecasting</td>
<td>92</td>
</tr>
<tr>
<td>7.1</td>
<td>An MMIS System</td>
<td>94</td>
</tr>
<tr>
<td>7.2</td>
<td>Data Sources for An MMIS Decision Support System</td>
<td>95</td>
</tr>
<tr>
<td>9.1</td>
<td>Chi-square Distributions for Various Degrees of Freedom</td>
<td>124</td>
</tr>
<tr>
<td>9.2</td>
<td>Forecast Versus Actual Manhours (Example 1)</td>
<td>128</td>
</tr>
<tr>
<td>9.3</td>
<td>Forecast Versus Actual Manhours (Example 2)</td>
<td>129</td>
</tr>
</tbody>
</table>
ABSTRACT

It is hypothesized that chemical processing plant maintenance manhour requirement (the number of hours required to maintain a production facility at an optimal level) can be forecast. The hypothesis is supported by (1), a study of the independent variables which may impact the dependent variable, maintenance manhour requirement and (2) a comparative analysis of technological forecasting methods that suggested which methods might be most appropriate. Both econometric and statistical forecasting methods were considered. The philosophy inherent in each approach was recognized. A procedure for choosing a forecasting method to suit a given situation is outlined. Multiple time series analysis, supported by Statistical Computing Associates (SCA) software, is used to develop a forecast model as an example.
Chapter 1
INTRODUCTION

Statement of the Problem

Efficient planning for maintenance manhour requirement (the number of hours required to optimally maintain a facility) has been neglected in the capital intensive chemical processing industry. Effective planning depends on knowing what will be required in the future. Future requirements may be projected from an analysis of what has occurred in the past.

With appropriate data, a mathematical model of what has occurred in a system over the past, can be developed. From that model, a forecast can be drawn which can be used as a basis for planning decisions involving, for example, (1), increases due to plant expansion, (2), distributional changes due to production demand or long term training need, or (3), reductions due to corporate restructuring.

A chemical processing plant may range in size from a self-contained unit from which a single specialty product is produced to a large, integrated chemical complex from which many products and their derivatives are generated. For data collection purposes, a chemical processing plant may be defined as that chemical production unit which an internal accounting department might designate as a separate profit or loss center. In the chemical processing plant, data is generated to monitor process, inventory,
maintenance, and financial control. With sufficient data storage systems, that data might be retrieved and used for decision support purposes. It is hypothesized that a forecast model for plant maintenance manhour requirement can be developed from such in-plant data by multiple time series analysis.

The Research Effort

The following steps were followed in the research effort:

(1). The chemical processing plant maintenance manhour requirement was analyzed with a systems approach in which the input (independent) variables both within and outside (endogenous and exogenous) the system were hypothesized to influence the single output (dependent) variable, the maintenance manhour requirement. Interactions between some or all of the input variables were assumed to be present. Feedback, lagged, and leading relationships between the multiple input variables were also assumed to be present;

(2). A comparative survey of forecasting methodology was made. Based on the multiple interactive variables list which resulted from the initial systems analysis, the review was extended to multiple regression analysis, econometrics, multiple time series analysis, and on hybrid or combinations of the
three preceding modelling techniques;

(3). From the comparative analysis of forecasting methodology, multiple time series analysis was determined to offer the most suitable methodology for forecasting maintenance manhour requirement. An on-line bibliographic literature search was conducted to check the uniqueness of the multiple time series analysis application to forecasting chemical plant maintenance manhour requirement;

(4). Applications of multiple time series forecasting to similar multiple input-single output systems were reviewed;

(5). Data availability within a chemical processing plant system was considered and a format for obtaining pertinent data was designed. Sample data collection forms and a relevant questionnaire may be found in the APPENDIX;

(6). An example, using simulated data, was developed.
Chapter 2

THE MAINTENANCE MANPOWER REQUIREMENT SYSTEM

The chemical processing plant is capital rather than labor intensive. Therefore priority rests with maintaining high mechanical performance. Sometimes this means that labor cost and assignment efficiency must be sacrificed. While the goal is to keep a chemical processing plant running smoothly, i.e., with a minimum of costly production shutdowns, either or both mechanical and labor oriented factors can unexpectedly impact production efficiency. A cost efficient routine preventive maintenance program can reduce, but cannot completely eliminate, the emergency maintenance incident.

Many chemical processing plants operate on a turnaround maintenance regimen, where production is halted and all major processing units are preventively maintained. Turnaround maintenance may be scheduled on a time dependent (periodic) basis or on a product quality dependent basis, i.e., when the product quality reaches a pre-specified degradation level, the plant is shut down for turnaround maintenance. Preventive maintenance which may be performed while the chemical processing plant is in production may be less extensive than that done during turnaround. Preventive
maintenance may also include maintenance done to redundant equipment which is periodically cycled out of active production use between turnarounds. Preventive maintenance is routine or scheduled maintenance. It can be delayed in most cases, without incurring immediate production shut-down. Emergency maintenance is of high priority due to the possibility or actuality of costly production shut-down. Emergency maintenance may or may not occur without warning; it may be the result of too often delayed preventive maintenance. Each type of maintenance may or may not be mutually exclusive. The manhour requirement for preventive and emergency maintenance is usually met by the in-house maintenance staff; contract labor is frequently employed for turnaround maintenance. Thus one would expect to find a cyclical dependency in any multiple time series model developed for maintenance manhour requirement. To develop maintenance manhour requirement models for each type of maintenance, it would have to be assumed that each type of maintenance was mutually exclusive. Each independent variable would have to be independently designated. For example, an in-house maintenance mechanic would have to record the exact number of hours worked on emergency and/or preventive maintenance projects. While such records may be able to be collected in-house, it may be difficult to
require such record-keeping for contract maintenance labor. Contract labor may be used not only at turnaround but also at times when a high backlog of preventive maintenance exists. Also, maintenance done at turnaround may not be exclusively preventive maintenance. For example, even though maintenance for a particular unit may be scheduled during turnaround, inspection may show that more than routine maintenance must be done and will take a longer time and require more repair supplies than had been anticipated. In that instance, a turnaround maintenance task may assume some of the characteristics of an emergency maintenance task. Therefore, for modelling simplicity, it makes sense to aggregate rather than to separate each independent variable according to type of maintenance requirement. However, it is also important, when viewed from an econometric basis, to realize that three categories of maintenance do exist and that some of the independent variables may or may not impact each of the maintenance requirement categories.

Each chemical processing plant environment may have one or more independent variables which may uniquely impact the dependent variable, maintenance manhour requirement. In general, the independent variables that have been hypothesized to affect maintenance manhour requirement, the
dependent variable, and for which quantifiable data might be available are:

(1). Level of plant operation;
(2). Backlog;
(3). Equipment age;
(4). Weather delays;
(5). Absenteeism;
(6). Number of units concurrently maintained;
(7). Impact of predictive maintenance control monitoring;
(8). Delays caused by parts procurement;
(9). Level of crew experience;
(10). Job-related training;
(11). Delays caused by job classification restrictions;
(12). Shut-down/start-up characteristics.

Less quantifiable factors which may influence maintenance work load include job cancellations and postponements, intermix of maintenance and non-maintenance jobs, variety of job completion times, unplanned emergency maintenance, production changes due to changes in sales requirements, materials shortages, and manning problems caused by sickness, absenteeism or lack of technical expertise. Additional qualitative influences on the total maintenance
manhour requirement may include long-term weather trends, productivity of the maintenance work force, amount of job-related training of the work force, acceptable preventive maintenance level, age of the plant, degree of technology of the plant, and sophistication of the management information system (MIS).

Variables Which May Influence Preventive Maintenance

When an efficient work priority system is in place, backlog data can be useful to forecast maintenance manhour requirement. Assuming that worker productivity is reflected in the backlog level, workforce size and composition can be optimized.

Plant operating level may influence the number of manhours required for preventive maintenance. A plant run at full capacity may require that all preventive maintenance work be done on schedule and to the full extent that has been planned, while a plant run at less than full capacity may require significantly fewer preventive maintenance manhours. For data consistency in model development, the percentage level of operation may be transformed into hours of 100 percent operating level.
Age of the major equipment will have an effect on manhours required for preventive maintenance. Older equipment may or may not be more easily accessible for preventive maintenance and may require more time consuming preventive maintenance procedures. Newer equipment may be designed for minimum maintenance.

Weather may cause loss in manhours where equipment is located in non-roofed areas or where extreme temperatures may make outdoor maintenance work less productive.

Absenteeism among maintenance crews may increase backlogs or increase manhours required to perform maintenance activities.

As a plant ages, the number of units maintained preventively may change. Possible units that are maintained only during turnaround can be maintained preventively during production time. Design changes may change the number of units maintained.

With more control monitoring equipment, preventive maintenance manhours may be reduced and time between maintenance work extended. Predictive maintenance control monitoring may show that more frequent preventive maintenance should be done.
Variables Which May Influence Emergency Maintenance

Plant operating level may influence the number of manhours necessary for emergency maintenance. A high operating level may result in deferred planned preventive maintenance and in increased manhours for emergency maintenance.

Parts procurement delays may increase the hours of emergency maintenance if the crew must wait for parts or try to find a source for the needed parts.

Age of the equipment subject to emergency breakdown may determine the time required for a single emergency repair and the frequency of repairs. Older equipment may be less accessible, more time consuming in parts procurement and require longer repair times.

The average number of years of maintenance crew experience may have an effect on the speed with which equipment can be diagnosed and serviced. Job related training adjusted for learning curve considerations may be included in the experience variable by recognizing that a specified number of training hours is equivalent to a specified number of on-the-job experience hours. Increased training and experience should result in some quantitative degree of workforce ability. A highly trained workforce may
result in fewer extended or repeat equipment breakdowns. A relationship may be found between downtime and mean crew experience/training in years.

With strict craft union jurisdictional restrictions, manhours may be lost during emergency maintenance while specific craftsmen are located for various aspects of an emergency maintenance task. Fewer manhours may be lost in plants where general purpose maintenance personnel are allowed to cross craft lines. If, during the observation time span, new contract negotiations have changed craft union restrictions, there may be quantifiable data available to reflect this change.

Variables Which May Influence Turnaround Maintenance

Plant operation level may influence manhour requirement for turnaround maintenance. A reduced operating level may allow time between turnarounds to be lengthened, or it may mean a shorter turnaround at the usually scheduled time.

Crew experience and job-related training may affect efficiency of a turnaround.

The number of units involved in turnaround maintenance may change as design changes remove some units from turnaround maintenance and place them into the preventive maintenance program, and conversely other units, due to
design changes, may be removed from preventive maintenance and placed on turnaround maintenance.

The hours required to shut down or start up a plant will affect the turnaround frequency and extent of preventive maintenance. Length of turnarounds may be determined by the number of units that cannot be maintained preventively.

**Intervariable Relationships**

As noted in the previous discussion on the number of units involved in turnaround maintenance, relationships may exist among those individual variables which influence the number of manhours required for each type of maintenance. The variables, discussed previously, are shown in Table 2.1. From Table 2.1, it can be seen that plant operating level may influence all three types of maintenance manhour requirements. Backlog is shown only to influence preventive maintenance since emergency maintenance by definition indicates the highest priority job action where no backlog can be tolerated. Under turnaround maintenance all backlog should be completed before the turnaround is deemed finished. Operating level may influence the increase of preventive maintenance backlog. At a high operating level, backlog may build as preventive maintenance must be
<table>
<thead>
<tr>
<th>Preventive</th>
<th>Emergency</th>
<th>Turnaround</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating level (%) (+)</td>
<td>Operating level (%) (+)</td>
<td>Operating level (%) (+)</td>
</tr>
<tr>
<td>Backlog (H) (+)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parts procurement delays (H) (+)</td>
<td></td>
</tr>
<tr>
<td>Major equipment age (H) (+)</td>
<td>Major equipment age (H) (+)</td>
<td></td>
</tr>
<tr>
<td>Weather delays (H) (+)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Crew experience (H) (-)</td>
<td>Crew experience (H) (-)</td>
</tr>
<tr>
<td></td>
<td>Union restriction delays (H) (+)</td>
<td></td>
</tr>
<tr>
<td>Absenteeism (H) (+)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of units (N) (+)</td>
<td>Number of units (N) (+)</td>
<td></td>
</tr>
<tr>
<td>Control monitoring impact (H)</td>
<td></td>
<td>Shutdown-startup (H) (+)</td>
</tr>
<tr>
<td>(-)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Additive variable: H Hours
Subtractive variable: N Number

Where an independent variable may exert influence on two or three areas of manhour requirement, the variable is listed in that area where theoretically it has the greatest impact. Where a variable theoretically may have an equal impact on more than one area of manhour requirement, the variable is listed in each area.
deferred. Conversely, at a low operating level, where some equipment may not be used, more preventive maintenance may be required to insure that the equipment will be ready when needed. A feedback relationship may also exist between backlog and manhour requirement. As backlog builds, manhour requirement may increase; manhour requirement may be decreased as backlog is decreased. Absenteeism may increase backlog which may in turn influence the manhour requirement. Weather delays may increase backlog. Parts procurement delays may be a result of low crew experience. As shown in Figures 2.1 to 2.3, relationships may exist among the variables which may impact the maintenance manhour requirement. These intervariable relationships may be directly or inversely proportional mathematically. Further, they may occur as lagged or leading or feedback time-based interactions.
Figure 2.1
The Chemical Processing Plant Maintenance Manhour Requirement System

Input Variables
- Maintenance Manhour Requirement at t - L
- Equipment Age
- Backlog
- Crew Experience
- Crew Training
- Absenteeism
- Weather Related Delays
- Predictive Maintenance Control Equipment
- Number of Units Concurrently Maintained
- Parts Procurement Delays
- Shut-down/Start-up Characteristics
- Job Classification Restriction Delays
- Operating Level

Output
MAINTENANCE MANHOUR REQUIREMENT AT T
Figure 2.2

Hypothesized Interactions Among the Independent Variables Which May Impact Chemical Processing Plant Maintenance Manhour Requirement via Backlog
Figure 2.3
Hypothesized Interactions Among the Independent Variables Which May Impact Chemical Processing Plant Maintenance Manhour Requirement via Operating Level

---

**IMPACT VARIABLES**

- **Feedback**
  - Predictive Maintenance Control Equipment
- **Proportional**
  - Parts Procurement Delays
  - Equipment Age
- **Reliability Curve Dependent**
  - Number of Units Maintained Concurrently
- **System Specific**
  - Shut-down/Start-up Characteristics

---

**SYSTEM**

---

**Operating Level**
Forecasting techniques are procedures used for predicting future events and conditions. Qualitative forecasting procedures are based on subjective predictions by those who are professionally or academically aware of the field in which the forecast is required. Formal qualitative procedures vary from simple summations of future estimates to the Delphi-type methods developed by the Rand Corporation, where a consensus of predictive opinion by a panel of experts is obtained. Quantitative forecasting methods are statistically based, logically stated, and mathematically computed. The procedures require that the historical data be analyzed so that the underlying process causing the variable or variables to fluctuate can be determined. Once the underlying process is identified and modeled, extrapolations (even considering all the cautions and dangers of extrapolation) can be made for forecasting purposes. A review of forecasting methodology follows. Strengths and weaknesses of each method are assessed based on the hypothesized chemical plant maintenance man-hour requirement system shown as follows:
X \rightarrow \text{Model} \rightarrow Y_{t} \quad (3.1)

for \ i = 1, \ldots, m, \text{ and } \ t = 1, \ldots, n,

where:

\( X \) is a vector quantity, comprised of multiple time series, and
\( Y \) is a univariate forecasted quantity; viz., maintenance manhour requirement.

As shown in Equation 3.1, there is multiple input, which when filtered through a statistical model will predict a univariate dependent variable. Research into the independent variables which may impact the maintenance manhour requirement suggests that the multiple independent variables may exhibit significant correlation which may be lagged, leading or feedback relationships. Because the goal is to develop a forecast model, serially correlated input data should be able to be accommodated.
Moving Average Models

The general moving average model may be described by:

\[ M = \frac{(X + X + X + \ldots + X)}{N} \]  \hspace{1cm} (3.2)

where:

- \( M \) is the moving average of the input values for time \( T \),
- \( X \) is the independent input variable at time \( T \), and
- \( N \) is the number of time periods over which the average is made.

The moving average forecast model form is:

\[ X(T + \tau) = M \]  \hspace{1cm} (3.3)

where:

- \( \tau \) is the number of time periods to be forecasted beyond time \( T \).

Therefore, the moving average can be used as a "\( \tau \)"-step ahead forecast model.
A moving average may be calculated as follows:

(1). Values for \( X \) over a time span of \( T - N \) are summed and divided by the number of values \( N \); 

(2). \( M \) is calculated by dropping the "oldest" \( T+1 \) observation and adding the "newest" observation.

An example for calculating a moving average is shown below:

<table>
<thead>
<tr>
<th>Period</th>
<th>Variable</th>
<th>Moving Average</th>
<th>Forecast One Period Ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>158</td>
<td>( (3\text{-Period}) )</td>
<td>( X (T-1) )</td>
</tr>
<tr>
<td>2</td>
<td>222</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>248</td>
<td>209</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>216</td>
<td>229</td>
<td>209</td>
</tr>
<tr>
<td>5</td>
<td>226</td>
<td>230</td>
<td>229</td>
</tr>
<tr>
<td>6</td>
<td>239</td>
<td>227</td>
<td>230</td>
</tr>
</tbody>
</table>

The limitations of using moving averages for forecasting chemical processing plant maintenance manhour requirement include the following:

(1). Value for \( N \) may be difficult to choose;

The number of time periods it takes for \( M \) to react to \( T \)
changes in the system is inversely proportional to the size of N. For a small value of N, the moving average will react more quickly to changes in the generating process. To change N, requires that the whole time series be re-averaged.

(2). Past values of a series are considered equally important to the process as are present values;

(3). For multivariate input, each series would have to be "smoothed" by a moving average. N for series X might not be suitable for series X;

(4). Interrelationships between the input variables are not directly addressed by the method.

(5). Stationarity of the series is assumed.

The advantages of the moving average method include (1), ease of calculation and (2), low input data requirement.

\[ S = aX + (1-a)S \]
\[ T \quad T \quad T-1 \]

[36] Simple Exponential Smoothing

The simple exponential smoothing model may be described by:

\[ S = aX + (1-a)S \]
\[ T \quad T \quad T-1 \]
where:

\[ S \] is the smoothed statistic, a weighted average of all past observations,
\[ a \] is the smoothing constant so that \( 0 < a < 1 \), and
\[ X \] is the actual statistic.

The forecast model form may be written as:

\[ X(T + t) = S_T^T \tag{3.5} \]

where:

\( t \) is the number of time periods to be forecasted beyond time \( T \).

The calculation procedure is as follows:

1. To start the smoothing process, a value for \( S \) is chosen, usually as an average of past observations;
2. A value for "a" is chosen. As "a" approaches 0, the response of the model to making forecast changes, decreases. A larger value for "a" makes the model respond faster;
3. \( S \) is calculated.
An example calculation with $a = 0.1$, follows:

<table>
<thead>
<tr>
<th>Period Number</th>
<th>Actual Statistic</th>
<th>Smoothed Statistic</th>
<th>Forecast From Last Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$X_T$</td>
<td>$S_T$</td>
<td>$X_{(T-1)}$</td>
</tr>
<tr>
<td>1</td>
<td>330</td>
<td>387</td>
<td>393</td>
</tr>
<tr>
<td>2</td>
<td>410</td>
<td>389</td>
<td>387</td>
</tr>
<tr>
<td>3</td>
<td>408</td>
<td>391</td>
<td>389</td>
</tr>
</tbody>
</table>

For $a = 0.1$ and letting $S = X(0) = 393$ as an average of $0 \ldots 1$ past observations, then:

$$S = 0.1 (330) + (0.9)(393) = 387$$

and when $T = 1$, the forecast model result is then:

$$X_{(1)} = 387.$$

The limitations of simple exponential smoothing for forecasting chemical processing plant maintenance manhour requirement include the following:

(1). Simple exponential smoothing assumes that the series upon which the forecast is to be based is constant, i.e. no trend or polynomial terms are included in the model;
(2). The model is suitable for a single time series so there is no provision for analyzing variable interaction between several time series;

(3). Once "a" has been chosen, it continues to be used so the model can not be adjusted if the resulting forecast values exhibit an increased error when compared to the actual values. In that case the model must be respecified.

The advantages of exponential smoothing include (1), ease of calculation and (2), a low input data requirement.

Higher Forms of Smoothing

Double exponential smoothing or second order exponential smoothing may be described by the following model:

\[
S_T = aS_T + (1-a)S_{T-1}
\]  

(3.6)

The expected value for \( X = 2S_T - S_T \)  

(3.7)

where:
$S$ is the smoothed statistic, a weighted average of all past observations (it includes a trend factor so that $E(S_T) = E(X_T) - (B/a)b$ since it is assumed that $E(X_T) = b_1 + b_2 T$ which is a linear trend model. $b_1$ and $b_2$ can be estimated by least squares and $B = 1-a.$), and $a$ is the smoothing constant so that $0 < a < 1.$

The double exponential smoothing forecasting model can be written as:

$$X_{T+1} = 2S_T - S_{T-1} + T_T (a/B)(S_T - S_{T-1})$$

where:

is the number of periods ahead to be forecast.

Triple or third order exponential smoothing may be described by the following smoothing equation:

$$S_T = aS_T + (1-a)S_{T-1}$$

where:
\[
S = aS_{T} + (1-a)S_{T-1}
\]  

(3.9a)

and,

\[
S = aX_{T} + (1-a)S_{T-1}
\]  

(3.9b)

are, as described previously.

The forecast model form may be written as:

\[
\hat{X}_{T} = \hat{A}_{T} + \hat{A}_{T}^{2} + 1/2 \hat{A}_{T}^{3}
\]  

(3.10)

where:

\[
\hat{A}_{T} = 3S_{T} - 3S_{T} + S_{T}
\]  

(3.10a)

\[
\hat{A}_{T} = \left(\frac{a}{2B}\right)((6-5a)S_{T} - 2(5-4a)S_{T} + (4-3a)S_{T})
\]  

(3.10b)

\[
\hat{A}_{T} = \left(\frac{a}{B}\right)(S_{T} - 2S_{T} + S_{T})
\]  

(3.10c)

with the same notation as described previously.
A generalized higher form of exponential smoothing may be written for p-order as:

\[
S_p^t = aS_{p-1}^t + (1-a)S_{p-2}^t + \ldots + (1-a)^{p-1}S_0^t
\]

where:

- \( p \) is the order of smoothing, and
- other notation is as described previously.

The forecast model form for p order may be written as:

\[
\hat{X}(T) = \hat{A}_1^T + \hat{A}_2^T (T/2!) + \hat{A}_3^T (T^2/3!) + \ldots + \hat{A}_n^T (T^n/n!)
\]

\[
\hat{X}(T) = \sum_{n=0}^{\infty} \hat{A}_n^T (T^n/n!)
\]

The following procedure is used to calculate the higher forms of exponential smoothing:
(1). Initial values are specified for the smoothed statistic. This is usually done by fitting a regression equation by least squares through past observations;

(2). An alpha, "a", value is specified, so that $0 < a < 1$;

(3). $S_T$, $S_T$, and $S_T$ are calculated as required for the specific form of exponential smoothing.

With respect to forecasting chemical processing plant maintenance manhour requirement, the limitations of the higher forms of exponential smoothing include:

(1). An assumed underlying data configuration. The double exponential smoothing model assumes a linear trend, the triple exponential smoothing model includes a quadratic term, and the p-order exponential smoothing models include polynomial terms of p-degree. To decide which model to use, the analyst looks at a scatter plot of past observations. Model selection is relatively restrictive and fairly empirical;

(2). The models are univariate so although each input series may be forecast individually, there is no provision for assessing any interrelationships which may exist between the individual time series;
(3). Once the alpha, "a", value has been chosen, it continues to be used so the model can not be adjusted if the resulting forecast error increases. Correction for the alpha "a" value requires model respecification.

Advantages of using the higher forms of exponential smoothing include (1), a choice of models to select and (2), the input data requirement may be as low as 30 observations.
Adaptive Process of Weighting Past Observations

The adaptive process is the basis behind the moving average and exponential smoothing procedures. Each model is really a statement of how past observations should be weighted. An adaptive process may be described by the following model:

\[
S_{T+1} = \sum_{i=t-N+1}^{t} w_i x_i \quad (3.13)
\]

where:

\( S_{T+1} \) is the forecast value for period \( t \),
\( w_i \) is the weight to be assigned to observation \( i \),
\( x_i \) is the observed value in period \( i \),
\( N \) is the number of observations.

With respect to forecasting chemical processing plant maintenance manhour requirement the limitations of the adaptive process of weighting past observations include the same limitations listed under the previous discussions of the moving average, simple exponential smoothing and higher exponential smoothing forecast methods. The advantage of
the adaptive process of weighting past observations is that the influence of "older" observations on the forecast is decreased.

[36]

Tracking Signals

The Trigg/Leach tracking signals model may be described by:

\[
a(T) = \frac{Q(T)}{\triangle(T)}
\]

(3.14)

where:

\(a(T)\) is the adjusted smoothing constant in period \(T\),
\(Q(T)\) is the smoothed forecast error, and
\(\triangle(T)\) is the smoothed mean absolute deviation.

The smoothed forecast error \(Q(T)\) may be described by:

\[
Q(T) = re(T) + (1-r)Q(T-1)
\]

(3.15)

where:

\(e(T)\) is the forecast error, and
\(r\) is the smoothing constant, so that \(0<r<1\).
The smoothed mean absolute deviation \( \Delta (T) \) may be described by:

\[
\Delta (T) = \text{r e} (T) + (1-r) \Delta (T-1), \quad -1 < \Delta T < 1. \tag{3.16}
\]

The following calculation procedure is used in tracking signals:

1. The forecast value is calculated at \( T + 1 \);
2. The forecast value is compared to the actual value at \( T + 1 \);
3. The smoothed forecast error is computed;
4. The value, "a" is adjusted;
5. Steps 1-4 are repeated if \( \Delta (T) \) approaches the outer control limit values of positive or negative unity.

The Trigg/Leach method of tracking signals may only be used for single exponential smoothing.

Control limits are described by the Chow method of tracking signals as follows:

\[
\begin{align*}
a &= a + g \\
u &= 0 \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quito
a is the smoothing constant, and
g is a control limit constant, usually set at 0.5.

The following calculation procedure is used:

(1). The forecast value is calculated using $a$;

(2). Forecast values are also computed using $a$ and $a$;

(3). A mean absolute deviation (forecast value - actual
value) is computed using the forecast values of steps
1 and 2. These are $\Delta (a)$, $\Delta (a)$, $\Delta (a)$;

(4). For $\Delta (a) < \Delta (a)$, $a = a$ and new values for

(5). For $\Delta (a) < \Delta (a)$, $a = a$ and new values for $a$
and $a$ are calculated.

Limitations to the Chow method for tracking signals
include:

(1). The calculation of three forecast values for each
forecast time period is required. This greatly
increases the amount of data which must be stored;
(2). The use of the Chow method is limited to higher forms of exponential smoothing; the restrictions noted in the previous discussion for these higher forms are relevant.

Tracking signals allow a model to be adjusted when control values are exceeded. The Chow method is included in some computer software for the Winter's Exponential Smoothing method.

Regression Analysis

The model used for regression analysis may be described as follows:

\[ Y = b_{x_1}(t) + b_{x_2}(t) + \ldots + b_{x_k}(t) + e \]  \hspace{1cm} (3.19)

where:

- \( b_i \) are parameters estimated, usually by least squares, \( i \)
- \( x_{i}(t) \) are mathematical functions of \( t \); viz: no trend, linear trend, quadratic, higher order polynomial, trigonometric, exponentiation. The function may also reflect a relationship
between two or more input variables, viz:

\[ x_1(t) \times x_2(t), \text{ and} \]

\[ e_t \]

is the random error component.

The form for the forecast model may be described by:

\[
\hat{Y}(T) = \sum_{n=1}^{N} b_n X_n(T) \tag{3.20}
\]

where values for each \( b_n X_n \) may be calculated from:

\[
b_n X_n(T) = b_n X_n + b_n X_{n}[T+\tau] + \ldots + b_n X_{n}[T+k\tau] \tag{3.21}
\]

The following calculation procedure is used:

(1). A regression equation is fitted to the data of past observations by least squares. The ability of the model to explain the greatest amount of error is improved through the use of stepwise regression or backward selection techniques;

(2). To use the model for forecasting, values for \( X_n(T) \) must be calculated individually by modelling each.
univariate time series with a regression equation, i. e. extrapolating from T.

With respect to forecasting chemical processing plant manhour maintenance requirement the limitations to regression analysis include:

(1). The parameter estimates \( b_1, b_2, \ldots, b_k \) are not considered to be accurate beyond the historical (past observation) time span;

(2). While the regression analysis model may provide a good description of what occurred in the past, if the generative process changes in the future, the forecast error will become unacceptable; i. e. the model will not predict with much accuracy;

(3). The error term, "\( e_t \)" must be random and independent but in time series data, the values are correlated in time so \( e_t \) will also be correlated in time;

(4). Highly cross correlated variables may adversely affect the estimated values for the parameters \( (b_0 \ldots b_n) \). This condition is called multicollinearity.
The advantages of forecasting with regression analysis are as follows:

(1). Multiple input variables can be considered;
(2). Computer software has been available long enough to have become well documented and reliable;
(3). Stepwise regression techniques aid greatly in model selection;
(4). Interrelationships among variables can be demonstrated and compared through cross correlation calculations.
Time Series Analysis

The Box-Jenkins analysis and forecasting methodology is comprised of the following steps:

1. Model Identification;
2. Parameter Estimation;
3. Diagnostic Checking;
4. Forecasting.

Univariate time series models may be described by the following forms:

AR (p) – Autoregressive of Order p

\[ X = E + \phi X + \phi x + \ldots + \phi X + e \quad (3.22) \]

MA (q) – Moving Average of Order q

\[ X = u + e - \theta e - \theta e - \ldots - \theta e \quad (3.23) \]

ARMA (p,q) – Mixed Model of Order p,q

\[ X = E + \phi x + \phi x + \ldots + \phi x - \theta e - \theta e - \ldots - \theta e + e \quad (3.24) \]
where:

u and \( E \) may be trend terms (in the differenced equation models \( u \) and \( E \) may equal zero),
is an autoregressive parameter,
is a moving average parameter, and
\( a \) and \( e \) are error terms, called random shock.

For multiple time series, the preceding models would be in vectorial form.

Time series forecast models may be described as follows:

**AR(p)** - Autoregressive of Order \( p \)

\[
X(T) = E + \phi X_{T+1} + \ldots + \phi X_{T+p} + e_{T+p} \quad (3.25)
\]

**MA(q)** - Moving Average of Order \( q \)

\[
X(T) = u + e - \theta e_{T+1} - \ldots - \theta e_{T+q} \quad (3.26)
\]

**ARMA(p,q)** - Mixed Model of Order \( p,q \)

\[
X(T) = E + \phi X_{T+1} + \ldots + \phi X_{T+p} + u + e - \theta e_{T+1} - \ldots - \theta e_{T+q} \quad (3.27)
\]
In the differenced form, the ARMA \((p,q)\) model is called the autoregressive integrated moving average model (ARIMA) and \(E\) and \(u\) may equal zero.

The following calculation procedure is used in Box-Jenkins univariate time series analysis:

1. Identification;

Autocorrelations and partial autocorrelations of the data from past observations are calculated and plotted. From the patterns shown by the plots, a general model is selected. A moving average model is selected if the autocorrelation spikes cut off abruptly and if the partial autocorrelations damp down as in an exponential function. An autoregressive model is chosen if the autocorrelation spikes damp down and if the partial autocorrelation spikes cut off abruptly. A mixed model is chosen if both autocorrelation spikes and partial autocorrelation spikes damp down. The series may be differenced if the plots do not follow the preceding patterns. Then autocorrelations and partial autocorrelations may be recalculated and replotted. Those spikes on the plots which are statistically significant are assumed to represent parameter values which must be estimated.
(2). Estimation;

The parameters are specified. Estimated values for the parameters are calculated by least squares or maximum likelihood procedures.

(3). Diagnostic Checking;

Error residuals (the difference between the calculated values and the actual values) are analyzed as a time series. Autocorrelations are calculated and plotted. If the model has been correctly identified and the parameters have been well estimated, there should be no significant spike on the autocorrelation plot.

(4). Forecast;

Forecast values are calculated either from (1), the difference equation form for the specified model, (2), as a weighted average of previous observations, or (3), from an 'integrated' form for the specified model using Psi weights.

[4] Psi weights may be calculated, recursively, from the $\phi$ and $\theta$ parameter values found in the estimation step. Thus:
\[ \psi = \phi_{1} - \theta_{1} \quad (3.28) \]

\[ \psi = \phi_{2} \psi + \phi_{2} - \theta_{2} \]

\[ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \]

\[ \psi = \phi_{j} \psi_{j-1} + \ldots + \phi_{p+d} \psi_{j-p-d} - \theta_{j} \]

for:

\[ \psi = 1, \quad \text{for } j < 0, \text{ and} \]

\[ \theta = 0, \quad \text{for } j > q, \]

where:

\( \psi \) is the Psi Weight,

\( \phi \) is the autoregressive parameter,

\( \theta \) is the moving average parameter,

\( p \) is the number of autoregressive parameters,

\( q \) is the number of moving average parameters,

\( d \) is the order of differencing.
Using Psi Weights, forecast values may be calculated as an infinite weighted sum of current and previous random error (shock) terms. Thus:

\[
X_{t+\tau} = \sum_{j=0}^{\infty} \Psi_j e_{t-\tau-j}
\] (3.29)

where:

all variables are as previously defined.

[33]

The preceding calculation procedure may be adapted for a multiple input time series, univariate output time series analysis, with the inclusion of a transfer function, as follows:

(1). Preliminary Univariate Analysis;

Univariate models are developed for each of the input time series. Any of the series may be differenced if it is found to be non-stationary.

(2) Causal Identification;

Cross correlation plots are drawn between each pair of pre-whitened univariate series. Based on the significance of spikes on the plots, a transfer function component is
added to the model. The mathematical form for the transfer function component is empirically specified, depending upon the results of the cross correlation plot.

(3). Noise Component Identification;

Parameters for the transfer function component are estimated. Residuals from this estimation are used to identify a model for the noise component.

(4). Estimation;

Parameters are estimated for the tentative model. If the parameters are not statistically significant a new transfer function component must be identified.

(5). Noise Component Diagnosis;

A new noise component must be identified if the residuals from the tentative model are not white noise.

(6). Transfer Function Diagnosis;

A new transfer function component must be identified if the residuals of the tentative model are found to be correlated with the prewhitened causal variable.

(7). Forecasting;

Based on the results of the cross correlation analysis and the addition, where appropriate, of the transfer function and noise terms to each pair of input time series, the model can be used to forecast the univariate output series.
For multiple time series input and output a generalized modelling procedure, which includes identification, parameter estimation, residual checking and forecasting steps may be used. The identification step includes the use of scatter plots, autocorrelation plots, sample cross correlation (CCM) matrix plots, and partial autoregression matrices. Scatter plots of each of the series may show variance instability; autocorrelation plots of each of the series may indicate nonstationarity. (Data transformation may be used to stabilize the variance; differencing may make the series stationary). Sample cross correlation matrices (CCM) may be computed from:

\[
P_{ij}(L) = \frac{\sum_{t=L+1}^{n} (Z_{it} - \bar{Z}_i)(Z_{jt} - \bar{Z}_j)}{\left(\sum_{t=1}^{n} (Z_{it} - \bar{Z}_i)^2 \sum_{t=1}^{n} (Z_{jt} - \bar{Z}_j)^2\right)^{1/2}}
\]

(3.30)

where:

\( \bar{Z} \) is the sample mean of the ith component series, or the corresponding mean of the transformed or differenced series.
When the vectored series is stationary, the sample CCM's are consistent estimates of the population cross correlations. CCM's which "cut off" after q lags may preliminarily identify an MA(q) model. Partial autoregression matrices are calculated by fitting autoregressive parameters at successively higher lags. The parameters are fit using least squares. The sample partial autoregression matrix \( P(L) \), shown in Equation 3.30, is equivalent to the estimated autoregression parameter at lag 1. If the vectored series is an AR(p) model, the autoregression parameters will equal zero for lags greater than p. All elements of the matrix \( P(L) \) will also be small. Estimates of the variances of the elements of the matrix \( P(L) \) may also be obtained. \( S(L) \), a chi-square (matrix) statistic, based on the determinant of the matrix of the residual sum of squares and the cross products may be calculated after each autoregression parameter fitting. The chi-square (matrix) statistic \( S(L) \) may be used to derive the statistic \( M(L) \) where:

\[
M(L) = - (N - 1/2 - L * k) l * n \left| \frac{S(L)}{S(L-1)} \right| \quad (3.31)
\]

and,

\[
N = n - p - 1, \text{ the number of observations minus the number of autoregression parameters fitted to lag 1.} \]
M(L) is distributed as chi-square with \( k \) degrees of freedom. As an example, significant M(L) values at lag 1 and lag 2, may preliminarily suggest that the model is an AR(2). Persistent high values for M(L) and a CCM plot "cut off" pattern may suggest an MA model. Persistent high values for M(L) and a CCM plot with persistent significant spikes may suggest an ARMA model.

Parameter estimation may be accomplished through conditional likelihood calculations. The exact likelihood parameter estimation may be used for an MA model, since bias may be introduced into MA parameter estimation by the conditional likelihood assumption that the error terms equal zero. According to the assumption then, \( a = \ldots \ldots \ldots \frac{p}{p-q+1} \ldots = 0 \).

Diagnostic checking procedures for the specified model may make use of the residuals from each of the individual series or from the vectored series. Outliers may be shown on the individual residual plots. A CCM, histogram, and a \([20]\) Portmanteau statistic may be calculated on the vectored residuals. For an adequate model, the CCM of the residuals should show few significant spikes and the histogram and chi-square goodness-of-fit test should show that the residuals are normally distributed at zero mean. From the
residual cross correlation values a multivariate Portmanteau statistic may be calculated, as follows:

\[ MP = n \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{v=1}^{s} r_{ijv} \ast r_{ji,-v} \]  \hspace{1cm} (3.32)

where:

- \( r \) are the residual correlations,
- \( v \) is the initial lag,
- \( s \) is the final lag,
- \( n \) is the number of residuals available after differencing,
- and
- \( m \) is the number of \( i \) or \( j \) elements in the residual vector.

MP may be distributed asymptotically as chi-square with

\( 2m \ast(s-p-q) \) degrees of freedom where \( p \) and \( q \) are the number of MA or AR parameters included in the model. A modified statistic may be calculated:

\[ MP' = n^2 \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{v=1}^{s} (n-r)^{-1} r_{ijv} \ast r_{ij,-v} \]  \hspace{1cm} (3.33)

where:
\( n, m, v, s \) and \( r \) are defined, as for Equation 3.32.

Forecast values may be calculated, recursively, from the general model for a lead time, \( L \), beyond time \( t \) from:

\[
Z(L) = C + \Phi Z(L-1) + \ldots + \Phi Z(1-p) -
\]

\[
\Theta E(a_{t+L+1}) - \ldots - \Theta E(a_{t+L-q})
\]

where:

\[
Z(j) = Z_t \quad \text{for } j \leq 0,
\]

\[
E(a_{t+j}) = 0 \quad \text{for } j > 0, \text{ and}
\]

\[
E(a_{t+j}) = a_{t+j} \quad \text{for } j \leq 0.
\]

Limitations of the Box-Jenkins methodology for forecasting chemical processing plant manhour requirement are as follows:

(1) For univariate time series analysis a "strong" predictive variable must be selected. In a chemical processing plant environment past values for maintenance manhour requirement may have very little to do with present values;
(2). Most work in Box-Jenkins methodology has been with univariate systems;

(3). Multiple time series forecasting is relatively new as an easily applied method since computer software has only recently been available. Identification and diagnostic checking require subjective judgement;

(4). Cross correlation of variables among the input series may be shown to exist but there may be little that the resulting multiple time series model can do to forecast a change in the interrelationship among the variables;

(5). Feedback intervariable relationships and bidirectional lagged and leading intervariable relationships can not be accommodated with the transfer function modelling extension to the Box-Jenkins methodology;

(6). Data requirements are high. At least 60 observations are required;

(7). The model development is based solely on the input data values and the empirical interpretation of the plots resulting from calculations made with the input data. Therefore forecast accuracy depends on the accuracy of the input data.
The advantages of the Box-Jenkins methodology are (1), a wide selection of models are available (2), comparative forecasts based on the same data show increased forecast accuracy (low residual error) with the method (3), feedback, lagged and leading intervariable relationships are able to be accommodated by the state space extension to the Box-Jenkins methodology and (4), appropriate computer software is becoming more available and better documented.
A system of simultaneous equations based on the independent input variables and the interrelationships among the variables may be developed to describe the dependent univariate output. An example may be shown as follows, for a quantity $Q$ at time $t$:

$$
Q(t) = \begin{cases} 
A = a + b + e \\
B = a + c Z + e \\
D = d E + e \\
Z = a + b Y + f B + e \\
G = A + B + E \\
H = G - D \\
\end{cases} \quad (3.35)
$$

where:
A, B, D, and Z are variables which impact the system (These equations are called structural equations and are designed to represent economic theory.),

\[ a_1, b_1, a_2, c_2, d_2, a_3, b_3, \text{ and } f_4 \text{ are parameters which must be estimated, usually by least squares, and } \]

\[ E_{t}, e_{t}, e_{t-1}, \text{ and } e_{t} \text{ are error terms.} \]

Variables A, B, D, and Z are called exogeneous because they are determined by the structural equations. Variables B and Z are called predetermined because at time t, they are determined within the system. E is an exogenous variable because at time t it is determined outside the system. The equations for G and H are called identities, since they represent known or given relationships between the structural equations. A forecast model for the preceding example may be described by:
\[
Q_{T+\tau} = \begin{cases}
A_{T+\tau} \\
B_{T+\tau} \\
D_{T+\tau} \\
Z_{T+\tau} \\
G_{T+\tau} \\
H_{T+\tau}
\end{cases}
\]  \hspace{1cm} (3.36)

where:

\[
A = a_1 + b_1 Y_1 + e_{T+\tau} \
B = a_2 + c_2 Z_2 + e_{T+\tau} \\
D = d_3 E_3 + e_{T+\tau} \\
Z = a_4 + b_4 Y_4 + f_4 B_4 + e_{T+\tau} \\
G = A + B + E \\
H = G - D
\]

\hspace{1cm} (3.36a) \hspace{1cm} (3.36b) \hspace{1cm} (3.36c) \hspace{1cm} (3.36d) \hspace{1cm} (3.36e) \hspace{1cm} (3.36f)
The calculation procedure which would be used for the system described in the preceding example follows:

(1). The variables which may impact on the dependent variable Q are identified and a series of equations describing the relationship to Q and to each other is designed. Economic theory is used to support the inclusion of specific terms to and the resulting structure of the equations which comprise the system;

(2). The equations are used to describe the historic behavior of the system. Calculated values for Q are compared with actual values of Q if actual values do exist. The system of equations may be used to describe a system for which Q has never been quantified;

(3). The system of equations may be modified if there is an actual value for Q and if there is a large discrepancy between the actual value and the value for Q which is generated from the system of equations;

(4). Forecast values for Q are based on extrapolation once the system of equations has been determined to provide a good historical profile of past observations.

For forecasting chemical processing plant maintenance manhour requirement the limitations of econometric forecasting include the following:
(1). Econometric equation systems are known to work best in macroeconomic environments where on an international or national or regional scale prevailing economic theory can be reliably shown to influence the system. In the microeconomic environment such theory may influence the system differently in ways that cannot be predicted with any degree of certainty;

(2). While the inclusion of lagged variables is useful for descriptive purposes, there is no set rule for deciding the extent to which variables should be lagged. This may result in the generation of non-white noise errors;

(3). The system of equations may become unwieldy when all the variable and feedback relationships are incorporated;

(4). Due to the high number of variables in the system, data requirements may be impractical;

(5). Much trial and error time must be spent to fine tune the model. Each equation might require respecification.

The major advantage to econometric forecasting is that it allows economic theory and creative insight to determine how the numerical data should be used for the design of the
simultaneous equation system. Forecasting has been said to be both qualitative and quantitative; econometric forecasting clearly supports that definition.

**System Simulation**

While not necessarily statistically-based, system simulation methods may be used for forecasting. As in econometric forecasting, a system is profiled using past observations. Once the profile is shown to accurately reflect what has happened in the past, extrapolation is used to predict future values.

Queuing theory based on known Poisson and exponential statistical distributions may be used to model a system with multiple input and univariate output. While Poisson and exponential distributions may be present in some actual maintenance service input streams, these distributions may not hold for the overall chemical processing plant maintenance manhour requirement system.

System modelling has also been applied to environmental and industrial dynamics problems using a system of simultaneous differential equations to describe the action of the input variables within the system. Initial values
and rates of change are chosen and the equations are used to calculate an output value. The rates of change or the design of the differential equations may be adjusted until an accurate profile of past or actual output values is obtained. Once the system of differential equations is shown to accurately profile the past, it is used for forecasting. The major restriction to using a system of differential equations for forecasting chemical processing plant maintenance manhour requirement is that the rates of change in action of the input variables cannot be reliably predetermined. It is hypothesized that the rates of change of the input variables may be stochastically, rather than deterministically based, as would be found in a completely dynamic system. Over time such a system may reach a pseudo steady state, i.e., within certain control limits.

**Summary of Forecasting Methodology**

Any of the preceding univariate methods might be used on the individual data input streams for forecasting chemical processing plant maintenance manhour requirement. For microeconomic variables, ARIMA (Box-Jenkins) models have been shown, in comparative forecasts based on the same input data, to provide the most accurate forecasting result.
Forecasting, rather than just a historical description of the input variable interrelationships, is the objective sought for the chemical processing plant maintenance manhour requirement problem. Past maintenance manhour requirement might be used as a univariate input variable. However, it is hypothesized that past maintenance manhour requirement would not accurately reflect productivity. The addition of backlog as another input variable could indicate whether the manhour requirement was productively used. A comparison of appropriate methods for forecasting univariate output from multivariate input follows.
The assessment of forecasting techniques for a multivariate system with multiple serially and causally correlated input and serially correlated univariate output focused on three techniques: Multivariate Analysis (Multiple Regression Analysis), Econometrics, and Multiple Time Series Analysis. While forecasting models can be derived from Multiple Regression Analysis models, Multiple Regression Analysis is best reserved as an analysis method to explain historical interrelationships among variables that are not serially correlated. Those econometrics models which are based on macroeconomic theory may have limited, if any, reliability in a microeconomic environment where anticipated economic influences may be outweighed by non-economic factors. Multiple time series analysis is based solely on the input data streams. When forecast error results are compared between applications of the preceding three techniques to identical forecasting problems, multiple time series analysis consistently has been shown to have the lowest mean squared error.

Aside from the computational differences in the three techniques, philosophical differences exist also. Multiple
Aside from the computational differences in the three techniques, philosophical differences exist also. Multiple regression analysis with the addition of 'economically correct' time lagged variables has been the technique favored by economists. Multiple time series analysis, with its dependence on the numbers (data) themselves to explain the nature of the underlying trend, has been favored by theoretical statisticians. These philosophical differences became more entrenched with the univariate time series analysis approach popularized by Box and Jenkins. Both techniques had strengths; both techniques had weaknesses. It is not surprising that analysts who were aware of both techniques, would seek a hybrid methodology. [14],[15]

In *Forecasting Economic Time Series*, C. W. J. Granger suggested that the analyst use both univariate time series analysis and multiple regression analysis to prepare two forecast values. He then recommended that the analyst average the two forecast values for a final forecast value. A hybrid methodology was adopted by Sanjay S. Modak et. al. and reported in "Combining Time Series and Regression Analysis to Forecast Production in the Indian Automobile Industry" which was published in the American Statistical Association Proceedings. A multiple regression analysis
model was developed to explain the interrelationships between the input variables. Then a univariate time series model was fit to the residuals of the multiple regression analysis model. The univariate time series model was used to forecast production in the Indian Automobile Industry.

David J. Pack in his article "Pitfalls of Combining Regression Analysis with Time Series Models", reiterated one of the basic assumptions of regression modelling; the parameters estimated for regression analysis models can not be deemed reliable if the input data is serially correlated. Thus, the reliability of both the Granger and the Modak methods might be questioned, for not meeting statistical constraints. Transfer function modelling had been described by Box and Jenkins as an extension of the univariate time series analysis. One facet of transfer function modelling, as expanded by David J. Pack, could be used to model a system where a single input data stream was transformed through a 'filtering process' (statistically explained by an adequate time series model) and output as a different stream. Thus a series written as $X_t$, with the addition of an empirically identified transfer function, could be output as a series written as $Y_t$. The transfer function modelling
technique can be extended to a multiple input, univariate output but not all feedback relationships can be accommodated. Only if the input series equations can be arranged so that the matrix of the estimated parameters is lower triangular, can the transfer function be used. An example of a suitable input series may be shown in the following:

[49]
A multivariate MA(1) model may be written in vector form as:

\[ Z_t = (I-\Theta B)a_t \quad \text{for } k \text{ series of } Z_t \quad (4.1a) \]

where:

\[ I = \bigg( \frac{\Theta}{q} B - \frac{\Theta}{q} B - \ldots - \frac{\Theta}{q} B \bigg)_{1 \ldots q}, \quad (4.1b) \]

\( Z_t \) is the observed process in matrix form, \( t \)
\( \Theta \) is a moving average parameter matrix, and
\( B \) is the backshift operator acting on the matrix.

For \( k = 2 \),

\[
\Theta = \begin{bmatrix}
\theta_{11} & \theta_{12} \\
\theta_{21} & \theta_{22}
\end{bmatrix} \quad (4.1c)
\]
and,

\[
\begin{align*}
Z_{1t} &= a_{11} - \Theta_{11} a_{1l} - \Theta_{12} a_{2l} \\
Z_{2t} &= a_{21} - \Theta_{21} a_{1l} - \Theta_{22} a_{2l}
\end{align*}
\]

If equations for \( Z \) can be arranged so that the coefficient matrix for \( t \) is lower triangular, then a transfer function model can be written. For example:

\[
\begin{align*}
\begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix} &= \begin{bmatrix} 1 - \Theta_{11} & 0 \\ \Theta_{21} & 1 - \Theta_{22} \end{bmatrix} \begin{bmatrix} a_{1l} \\ a_{2l} \end{bmatrix} \\
&= \begin{bmatrix} a_{1l} \\ a_{2l} \end{bmatrix} (4.2)
\end{align*}
\]

Multiplying the matrix through, provides:

\[
\begin{align*}
Z_{1t} &= (1 - \Theta_{11}) a_{1l} \\
Z_{2t} &= (-\Theta_{21}) a_{1l} + a_{21} (1 - \Theta_{22}) a_{2l}
\end{align*}
\]

The error terms may be expressed by the equation:

\[
a_{2t} = B_{2t} a_{1t} + e_{2t}, \text{ where } a_{2t} \text{ and } e_{2t} \text{ are independent.}
\]
Then \( a = \frac{Z}{(1-\theta B)} \). \( Z \) may be written in the transfer function form as follows:

\[
Z = \frac{(-\Theta BZ)}{(1-\Theta B)} + \frac{(Ba + e)(1-\Theta B)}{1-\Theta B} \quad (4.4a)
\]

\[
Z = \frac{(-\Theta BZ)}{(1-\Theta B)} + \frac{(BZ)}{1-\Theta B} + e(1-\Theta B) \quad (4.4b)
\]

\[
Z = \frac{(-\Theta BZ)}{(1-\Theta B)} + \frac{(BZ_b)}{1-\Theta B} - \frac{2}{(BZ_b) + e - e(\Theta B)} \quad (4.4c)
\]

\[
Z = \frac{(BZ(-\Theta - \Theta B))}{(1-\Theta B)} + e(1-\Theta B) \quad (4.4d)
\]

\[
Z = \left( \frac{B-B(\Theta + \Theta)}{1-\Theta B} \right) \frac{Z}{1-\Theta B} + (1-\Theta B)e \quad (4.4e)
\]

Letting \( W = B \) and \( W = B(\Theta + \Theta) \),
which is one of the transfer function forms that is generalized by:

\[
\sum_{t=0}^{h-1} \frac{W(B)}{1-\Theta(B)} e^{2t} + (1-\Theta(B)) e^{22} t \tag{4.5}
\]

for \( (h=1,2,3,\ldots,k) \)

where:

\[
W(B), \ g(B), \ \Theta(B) \ \text{and} \ \Theta(B) \ \text{are polynomials in} \ B \ \text{si} \ \text{ri} \ \text{ph} \ \text{qh} \\
\text{the backshift operator}, \ \text{the b's are positive integers and} \ i \\
\text{the set} \ \{a, \ldots, a\} \ \text{are k independent Gaussian white noise} \ lt \ \text{kt} \\
\text{processes with zero means and variances of} \ \sigma^2. \ldots. \ \\
\sigma_1^2 \ \text{Values for} \ W \ \text{can be quantified during the parameter} \ ki \ \text{estimation procedure.} \]
Another restriction to transfer function modelling is that all intervariable relationships must be reconciled to the same direction; i.e., lagged and leading relationships cannot be accommodated within the same model. Lagged and leading relationships within the same model would preclude any feasible arrangement of the series of input equations into a parametric matrix that is lower triangular. The necessary condition for using transfer function modelling for multiple input. While the transfer function technique has been used by researchers, it has not had the applications use nor the software development that the univariate time series analysis methodology has had.

The number of published applications for multiple time series analysis increased with the 1979 development of adequate software to process the extensive calculations required by multiple time series analysis. The quantification and charting of the cross correlation calculations may have caused the econometrician to reconsider the methodology. Interest in computational hybrid forecasting techniques subsequently dwindled, as noted by a decrease in published articles on the subject. State space forecasting, an extension of multiple time series analysis to include leading and lagging indicators and feedback relationships, evolved as a philosophical
hybrid. A. V. Cameron of State Space Systems, Inc. developed software to incorporate interest rate forecasting and portfolio analysis. State space forecasting has also been added to the SAS/ETS software package. The methodology is based on the fitting of a state space vector to an appropriate statistical model by minimizing the computed value for the Akaike Information Criteria. An ARMA model may be written as:

\[ \phi(B)X_t = \Theta(B)e_t \]  

(4.7)

and expanded to:

\[ X_t - \phi X_{t-1} - \ldots - \phi^p X_t = e_t + \Theta_0 e_{t-1} + \ldots + \Theta_q e_{t-q} \]  

(4.8)

where:

- \( e_t \) is a series of independent multivariate normal random vectors with variance \( \sigma \) and mean of zero,
- \( B \) is the backshift operator,
- \( \phi(B) \) and \( \Theta(B) \) are matrix polynomials with \( \phi(0) = \Theta(0) = I \),

and
\( X \) is the observed process.

Further,

\[ X = \phi (B) \Theta(B) e = \sum_{s=0}^{\infty} \psi_s e_{t-s} \]  

(4.9)

where \( \psi_s \) matrices, known as impulse response matrices, may be computed as \( \phi^{-1} (B) \Theta(B) \).

Conditional expectations for the system of equations shown by Equation (4.9) comprise the state space form:

\[ X_{[t+j]t} = \sum_{s=j}^{\infty} \psi_s e_{t+j-s} \]  

(4.10a)

\[ X_{[t+j]t+1} = X_{[t+j]t} + \psi_{j-l} e_t \]  

(4.10b)

and from the expanded form of the system of equations shown by Equation (4.8):

\[ X_{[t+p]t} = \phi X_{[t+p-1]t} + \cdots + \phi X_{t} \]  

(4.10c)

The state space system of equations in matrix form is:
An ARMA system of equations written in conditional expectation form may be converted to a state space form as shown in the following example:

\[
\begin{bmatrix}
    x_{t+1} \\
    x_{[t+2]t+1} \\
    \vdots \\
    x_{[t+p]t+1}
\end{bmatrix}
= \begin{bmatrix}
    0 & I & 0 & \cdots & 0 \\
    0 & 0 & I & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & 0 & \cdots & I
\end{bmatrix}
\begin{bmatrix}
    x_t \\
    x_{[t+1]t} \\
    \vdots \\
    x_{[t+p-1]t}
\end{bmatrix}
+ \begin{bmatrix}
    1 \\
    \psi \\
    \vdots \\
    \psi_{p-1}
\end{bmatrix}
\]

\[
\begin{bmatrix}
    e_t \\
    e_{t+1} \\
    \vdots \\
    e_{t+p}
\end{bmatrix}
\]

then:

\[
\begin{align*}
X_{[t+1]t} &= 0.2X_t + 0.4Y_t - 0.5e_t - 0.7a_t \\
Y_{[t+1]t} &= 0.1X_t + 0.3Y_t
\end{align*}
\]

At \( t+2 \):

\[
\begin{align*}
X_{[t+2]t+1} &= 0.2X_{t+1} + 0.4Y_{t+1} - 0.5e_{t+1} - 0.7a_{t+1} \\
Y_{[t+2]t+1} &= 0.1X_{t+1} + 0.3Y_{t+1}
\end{align*}
\]
\[
X^{[t+2]} = 0.2X^{[t]} + 0.4Y^{[t]}
\]
\[
= 0.2X^{[t+1]} + 0.4(0.1X^{[t]} + 0.3Y^{[t]})
\]
\[
= 0.2X^{[t+1]} + 0.04X^{[t]} + 0.12Y^{[t]}
\]

From Equation (4.8):

\[
X = X + e_{t+1}, \quad \text{An Identity (4.13)}
\]

From Equation (4.9):

\[
Y = Y + a = 0.1X^{[t]} + 0.3Y^{[t]} + a_{t+1}
\]

From Equations (4.12b) and (4.14):

\[
X^{[t+2]} = 0.2X^{[t]} + 0.4Y^{[t]} - 0.5e^{[t]} - 0.7a^{[t+1]}
\]
\[
= 0.2X^{[t+1]} + 0.2e^{[t]} + 0.4(0.1X^{[t]} + 0.3Y^{[t]}) + a_{t+1}
\]
\[
- 0.5e^{[t+1]} - 0.7a^{[t+1]}
\]

The state space form for the system is:
\[
\begin{bmatrix}
X_{t+1} \\
Y_{t+1} \\
\cdot \\
X_{[t+2]t+1}
\end{bmatrix} =
\begin{bmatrix}
0 & 0 & 1 \\
0.1 & 0.3 & 0 \\
0.04 & 0.12 & 0.2 \\
0.04 & 0.12 & 0.2
\end{bmatrix}
\begin{bmatrix}
X_t \\
Y_t \\
\cdot \\
X_{[t+1]t}
\end{bmatrix} +
\begin{bmatrix}
e_{t+1} \\
0 \\
0 \\
1
\end{bmatrix}
\] (4.16)

In the state space forecasting method the parameters of the state vector \((X_{[t+p]t+1})\) are estimated. Therefore more parameters than might be necessary for an otherwise adequate model may end up being included in a final forecasting model. Since each parameter that must be estimated introduces additional error to a model, the principle of parsimony (the least number of parameters to provide an apt model) which is a key precept for ARIMA time series modelling is challenged by the state space methodology. In order to use the state space forecasting method each input time series must depend on the past values of all of the other input time series. As shown in the preceding example, the state space vector is dependent on \(X, Y,\) and \(X\). Conditional expectation conversion for the entire set of equations requires such dependency. State space forecasting requires that each of the input time
series must depend on the past values of all of the other input time series, as demonstrated in the following example.

\[ X_t = X_{t-1} + Y_{t-1} \]  
(4.17a)

\[ X_{t-1} = X_{t-2} + Y_{t-2} \]  
(4.17b)

\[ Y_{t-1} = X_{t-1} + Y_{t-1} \]  
(4.17c)

Where variable analysis precludes the use of either transfer function modelling or state space forecasting, the [29] SCA system offers versatile software for the computations required by generalized multiple time series analysis.
Chapter 5

RESEARCH SOURCE DESCRIPTION

The literature search focused first on the retrieval of technological forecasting applications for chemical plant maintenance manhour requirement. Appropriate keyword terms were used to search on-line bibliographic data bases. The data bases included (1), Dissertation Abstracts (1861-1987), (2), NTIS (National Technical Information Service, 1964-87), (3), Compendex (Engineering Index, 1970-87), and (4), ABI/INFORM (Business Literature, 1971-87). No published applications of multiple time series forecasting for maintenance manhour requirement for the chemical process or for any other industry were found. If multiple time series analysis has been used to forecast maintenance manhour requirement in the chemical process industry, that application has been proprietary. Thus, it is assumed that the suggested application of multiple time series forecasting for chemical plant maintenance manhour requirement is a unique contribution to the published literature. Earlier work by the author suggested using a multiple regression analysis approach to forecast chemical plant maintenance manhour requirement.

[31]
The objective for the second part of the literature search was to find applications of multiple time series analysis for systems similar to that for maintenance manhour requirement in a chemical processing plant. Multiple time series analysis has been used for forecasting in systems engineering, social science, finance, and marketing. From the successful results obtained in these applications, it is reasonable to suggest that multiple time series analysis could be applied to forecasting chemical plant maintenance manhour requirement.

A systems engineering application is for predicting an effluent analysis value, the K number, from a continuous pulping digester. Two input variables \( X_l \), the feed temperature, and \( X_{2t} \), the blow to feed (B/f) ratio, influence the output variable \( X_{3t} \), the K number for the product pulp. Through multiple time series methodology it was found that \( X_{3t} \) had a significant correlation with \( X_l \) at lags 6 and 7. \( X_l \) and \( X_{2t} \) were shown to be independent of either of the other series through the inspection of the cross correlation plots. An adequate model for \( X_{3t} \), was
found to be:

\[ X = -PX_{t-6} + PX_{t-7} - PX_{2t-1} + PX_{2t-2} \]

\[ \quad + PX_{3t-1} - PX_{3t-2} + a_{3t} - Pa_{3t-1} \]  

(5.1)

where:

\[ X = PX_{1t} + a_{1t} \]

\[ X = PX_{2t} + PX_{2t-1} + a_{2t} + Pa_{2t-1} \quad \text{, and} \]

the P values are uniquely estimated parameters.

Had \( X \) exhibited a significant correlation with \( X \), a term

\[ a_{2t} \]

or terms relating those series at the lag point of significant correlation could have been included in the models for \( X \) and/or \( X \).

\[ a_{1t} \]

Multiple time series analysis, using the transfer function modelling technique, has been applied to a social sciences population forecasting problem to find out if crop failure, rather than or in addition to fertility rate influenced population in Sweden. Cross correlation analysis between the two input variables, the Harvest Index (\( h \)) and fertility rate (\( f \)) confirmed that there was a

\[ a_{t} \]
significant causality between the input variables. An acceptable predictive model for the population was found to be:

\[ p = Q_f + Q_h + Q_a \]  
\[ p = Q_f + Q_h + Q_a \]  
\[ t \quad t \quad t-1 \quad t \]  

where:

\[ p \]  
\[ p \]  
\[ t \]  
\[ t \]  
\[ f \]  
\[ f \]  
\[ t \]  
\[ t \]  
\[ h \]  
\[ h \]  
\[ t-1 \]  
\[ t-1 \]  
\[ a \]  
\[ a \]  
\[ t \]  
\[ t \]  

the Q values are uniquely estimated parameters.

As restricted by the transfer function modelling technique, the causality was one way and in the same direction. Thus, the harvest influenced the fertility rate which in turn influenced the population growth.

[10]

In a public finance application, multiple time series analysis was used to develop municipal budget forecasting models which would enable city financial officials to assess the impact on the budget due to changes in revenue structure (viz., property taxes) or in expenditure restrictions (viz.,
California's Proposition 13). From data for 1949 to 1979, a model for San Diego was developed as follows:

\[ P = 0.077 + a + 0.552a \]
\[ R = 0.118 - 0.320E - 0.449P + 0.378R + a + 0.341a \]
\[ E = 0.074 + 0.607R - 0.577E + a \]

where:

\( P \) is Property tax at time \( t \),
\( R \) is Other revenue, including revenues generated by taxes other than property tax, fees and penalties, and intergovernmental transfers, at time \( t \), and
\( E \) is Expenditures at time \( t \).

As shown in the preceding equations, other revenue is impacted by property tax at a time lag of 2, and expenditures at a time lag of 1. The expenditure variable
is influenced by the other revenue variable at a time lag of 3. These significant interrelationships between the time series were noted through cross correlation analysis.

A Pittsburgh budget model was similarly developed, but unemployment (U) and consumer price index (C) data were included, as follows:

\[
P_t = 0.020 - 1.702U_{t-3} + a_t \quad (5.6)
\]

\[
R_t = 0.054 + a_t \quad (5.7)
\]

\[
E_t = 0.028 + 0.318R_{t-1} + a_t + 0.591a_{t-1} + 0.988a_{t-1}C_{t-1} \quad (5.8)
\]

\[
U_t = 0.001 + a_t + 0.214a_t + 0.797a_{t-1}C_{t-1} \quad (5.9)
\]

\[
C_t = 0.031 + a_t + 0.797a_{t-1} + 0.544a_{t-1}C_{t-1} \quad (5.10)
\]

and, where all variables are, as previously defined.
While the budget models are structured as a series of input variable equations, a single output variable, the net budget, or cash difference between expenditures and revenue may be derived. In the Pittsburgh model, one would expect that expenditures would be influenced only by unemployment and cost of living data status and that property tax and other revenue would be the only variables to affect income. However, the model, as shown above, indicates that interrelationships exist between input to the income and to the expense side of the budget. Multiple time series analysis was found to be especially useful for the identification of the time lagged and feedback interrelationships. Econometric modelling procedures based on generally accepted theoretical principles may have identified the time lagged relationships but only by trial and error. Cross correlation residual checking is not part of the traditional econometric modelling procedure.

[9]

In a banking application, a predictive model for total investments (INVEST) was developed. Initial variable analysis suggested that the multiple input variables might include demand deposits (DD), total loans (LOANS), treasury bill yield rate (BILLS), and interest rate for long term government bonds (BONDS). Significant cross correlation
relationships among the input time series were found as follows:

\[
\text{BONDS} = DD_t + DD_{t-1} + BILLS_{t-6} + BILLS_t + \text{LOANS}_{t-6} \quad (5.11)
\]

\[
\text{BILLS} = DD_t + DD_{t-6} + \text{LOANS}_{t-11} + \text{LOANS}_{t-1} \quad (5.12)
\]

A predictive model for the single output variable, based on the preceding interrelationships was developed as shown below:

\[
\text{INVEST} = \text{LOANS}_t + \text{BILLS}_{t-3} + \text{BONDS}_{t-2} + \text{BONDS}_{t-1} + \text{BONDS}_{t-2} \quad (5.13)
\]

In a marketing research application, predictive sales response models were developed using multiple time series analysis to assess the impact of advertising expenditures and competitive pricing among similar products.

While the preceding applications are diverse, they do show the ability of multiple time series analysis to (1), quantify lagged input variable interrelationships and (2), to model feedback relationships among the input variables
in a stochastic, multivariate system. In each case, the model which was developed provided a statistically sound and economically plausible explanation for the impact that was made by the multiple input variables on the single output variable.
Chapter 6

MODEL FORMULATION PROCEDURE

In general, the forecasting problem analysis procedure can be summarized as follows:

(1). Choose the dependent variable (the variable for which the forecast is to be made);

(2). List the independent variables;

(3). Theorize how the independent variables may impact the dependent variable;

(4). Assess the relationships between the independent variables;

(5). Classify the independent variables according to their direct (primary) and indirect (secondary) impact on the dependent variable;

(6). Theorize the feedback and the lagged or leading relationships among the primary independent variables and the dependent variable;

(7). Formulate a generalized model for the relationships;

(8). If the theoretical intervariable relationships produce a parameter matrix for the independent variables, which is lower triangular, a transfer modelling approach may be applicable for the specific model identification;
(9). If theoretically, each of the input time series may depend on the past values of ALL of the input series, the state space modelling approach may be considered;
(10). If neither the transfer modelling nor state space forecasting approaches are applicable, model the actual data using a generalized multiple time series approach. The SCA software has the versatility for the generalized modelling requirement;
(11). Check the model for aptness;
(12). Use the model to forecast.

To consider Step 7 with regard to the forecasting of maintenance manhour requirement, the following generalized model based on the primary (directly impacting) independent variables might be theorized.

\[
RATIO_{t} = RATIO_{t-L} + MANHOURS_{t-L} \quad (6.1a)
\]

\[
BACKLOG_{t} = BACKLOG_{t-L} + MANHOURS_{t-L} \quad (6.1b)
\]

\[
MANHOURS_{t} = MANHOURS_{t-L} + BACKLOG_{t-L} + RATIO_{t-L} \quad (6.1c)
\]

where:

\[
RATIO \quad \text{is the production/capacity ratio expressed in hours of 100\% plant operating time,}
\]
BACKLOG is the number of hours of backlogged maintenance work,
MANHOURS is the number of hours of maintenance manhour requirement,
t is time at the present, and
L is the number of periods of lag.

For notational simplicity, error terms have not been included in the preceding model.
A parameter matrix for the preceding set of equations would be:

\[
\begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 1 \\
1 & 1 & 1 \\
\end{pmatrix}
\]

In this case, multiple time series analysis would be appropriate. The input variables might suggest that a relationship may exist which would allow the set of equations to be rearranged so that a parameter matrix was lower triangular. For example:

\[
\text{RATIO} = \text{RATIO} (6.2a) \\
\text{BACKLOG} = \text{RATIO} + \text{BACKLOG} (6.2b) \\
\text{MANHOURS} = \text{RATIO} + \text{BACKLOG} + \text{MANHOURS} (6.2c)
\]
A parameter matrix for the preceding set of equations, as arranged, would be lower triangular as shown below:

\[
\begin{array}{ccc}
1 & 0 & 0 \\
1 & 1 & 0 \\
1 & 1 & 1 \\
\end{array}
\]

Thus, an apt forecasting model might be found through the transfer modelling approach.

Univariate modelling of the input data, followed by bivariate cross correlation analysis of the univariate residuals may provide insight into intervariable causality. Causality may be identified through a chi-square comparison of a test statistic based on the autocorrelation calculation:

\[
S = \sum_{m} \left( \sum_{k=-M}^{M} r_{nr}^{2}(k) \right) _{n_m} 
\]

where:

- $S$ is the test statistic,
- $m$
n is the number of observations,
M is the number of lags, and
r is the estimated autocorrelation value at lag k.

Bivariate causality patterns for two variables X and Z are shown in Table 6.1. Significant lagged causality may be shown to exist among all of the variables so that each member of the set of equations is dependent upon all other members. For example:

\[
\begin{align*}
\text{RATIO}_t &= \text{RATIO}_{t-L} + \text{BACKLOG}_{t-L} + \text{MANHOURS}_{t-L} \\
\text{BACKLOG}_t &= \text{RATIO}_{t-L} + \text{BACKLOG}_{t-L} + \text{MANHOURS}_{t-L} \\
\text{MANHOURS}_t &= \text{RATIO}_{t-L} + \text{BACKLOG}_{t-L} + \text{MANHOURS}_{t-L}
\end{align*}
\]

The parameter matrix for the preceding set of equations would be:

\[
\begin{pmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{pmatrix}
\]

The state space forecasting approach might be validly used to model such a system.

The decision process for forecasting the maintenance manhour requirement may be summarized as follows:
### Table 6.1

Bivariate Causality Patterns

*(Variables X and Z)*

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>X causes Z</td>
<td>( r(0) \neq 0 ) for some ( k &gt; 0 )</td>
</tr>
<tr>
<td>Z causes X</td>
<td>( r(0) \neq 0 ) for some ( k &lt; 0 )</td>
</tr>
<tr>
<td>Causality at the same instant</td>
<td>( r(0) \neq 0 )</td>
</tr>
<tr>
<td>Feedback</td>
<td>( r(0) \neq 0 ) for some ( k &gt; 0 ) and ( k &lt; 0 )</td>
</tr>
<tr>
<td>X causes Z but not at the same instant</td>
<td>( r(0) = 0 ) for some ( k &gt; 0 ) and ( r(0) = 0 )</td>
</tr>
<tr>
<td>X does not cause Z</td>
<td>( r(0) = 0 ) for all ( k &lt; 0 )</td>
</tr>
<tr>
<td>Z does not cause X</td>
<td>( r(0) = 0 ) for all ( k \leq 0 )</td>
</tr>
</tbody>
</table>
Table 6.1 (Continued)

<table>
<thead>
<tr>
<th>X to Z causality in one direction</th>
<th>$r (k) \neq 0$ for some $k &gt; 0$ and $nr$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r (k) = 0$ for either all $k &lt; 0$ and $nr$</td>
</tr>
<tr>
<td></td>
<td>or all $k &gt; 0$</td>
</tr>
<tr>
<td></td>
<td>or all $k \leq 0$</td>
</tr>
</tbody>
</table>

| X and Z causality only at the same instant if at all | $r (k) = 0$ for all $k \neq 0$ and $nr$ |

| X and Z causality only and in no other way | $r (k) = 0$ for all $k \neq 0$ and $nr$ |

| X and Z are independent | $r (k) = 0$ for all $k$ and $nr$ |
(1). Objective is to develop forecasting model: \(\text{Time Series Analysis}\);

(2). Analysis of variables: \(\text{Multiple Input, Univariate Output: Multiple Time Series Analysis}\);

(3). Causality among input variables: \(\text{Transfer Modelling, State Space Forecasting, Generalized Multiple Time Series Analysis}\);

(4). Parameter matrix is not lower triangular: \(\text{State Space Forecasting, Generalized Multiple Time Series Analysis}\);

(5). Each input series does not depend on past values of ALL of the series: \(\text{Generalized Multiple Time Series}\).

As shown in Figure 6.1, forecasting requires more than mere data processing through what may turn out to be an inappropriate statistical modelling procedure.
Figure 6.1

Decision Process for Forecasting

SPECIFY: Dependent Variable

LIST: Independent Variable(s)

ASSESS: Impact Between and Among Variables

FORMULATE: General Model

TRANSFER MODEL

PARAMETER MATRIX

Lower Triangular Interdependent

MULTIPLE TIME SERIES

STATE SPACE

All Input Series
Maintenance management information systems (MMIS) have been developed and placed into service as computerized data systems for storage, handling and retrieval have become more cost efficient. As shown in Figure 7.1, the MMIS may be designed to make use of information generated in other areas of the chemical processing plant, such as the purchasing or financial accounting departments. Such data may be generated, consolidated and stored for short term corporate balance sheet review; it may also be used for long term decision support purposes. Multiple time series forecasting could be integrated as one such long term decision support adjunct to an MMIS as shown in Figure 7.2. As data was collected over time, for more and more variables, the additional time series might be analyzed for addition to a previously developed forecasting model.

Another data source that might be tapped as input for decision support is that data which may already be generated and collected at the plant specifically for process control. Mathematically, the governing equations for process control are based on time series models, making full use of feedback
Figure 7.1

An MMIS System

- PRODUCTION
- DATABASE
- INVENTORY
- PERSONNEL
- ACCOUNTING
- PURCHASING
- SCHEDULING
- EQUIPMENT HISTORY
- WORK STANDARDS
- BACKLOG
- PRODUCTION
Figure 7.2
Data Sources For
An MMIS Decision Support System

INVENTORY
Parts Procurement
Delay
Time

EQUIPMENT HISTORY
- Age
- Pred. Control

PERSONNEL
- Experience
- Training
- Absenteeism

BACKLOG
Productivity

PRODUCTION
- Op. Level
- Downtime

ACCOUNTING
- Manhours
- Repair Costs

DATA FOR MODEL

FORECAST
relationships which may exist between input and output
variables. Optimal control may be defined as those
adjustments in the input values that yield minimum variance
or minimum mean squared error for the output. The minimum
mean squared error forecast for the output can be obtained
from conditional expectation where the conditional
expectation of the:

(1). Present or past observation is that (known)
observation;
(2). Future observation is its forecast;
(3). Present or past "shock" a is that shock, which can be
computed; and
(4). Future shock is zero.

Thus,

\[ E(X_{t-j}) = X_{t-j}, \quad j = 0, 1, 2, \ldots \quad (7.1) \]

\[ E(X_{t+j}) = X_{t}, \quad j = 1, 2, 3, \ldots \quad (7.2) \]

\[ E(a_{t-j}) = a_{t-j}, \quad j = 0, 1, 2, \ldots \quad (7.3) \]

\[ E(a_{t+j}) = 0, \quad j = 1, 2, 3, \ldots \quad (7.4) \]

Assuming that there is always some lag time (i.e., \( L \) is
always greater than or equal to one) between the input
which can be manipulated to control output, the earliest point in time when the input variable $X_t$ can be changed to control the output $X$ is $t+L$. The minimum mean squared error control strategy is to adjust the input $X$ such that the forecast $X$ made at time $t$, $X_{2t+L}$, is zero. The optimal control strategy equation may be written as:

$$X_{2t} = 0. \quad (7.5)$$

[44] S. M. Pandit and S. M. Wu proved the preceding equation using orthogonal decomposition. From the conditional expectation rules,

$$X_{2t+L} = X_{2t} + e_{2t} \quad (7.6)$$

and thus,

$$X_{2t+L} = e_{2t} \quad (7.7)$$

which has the smallest variance that can be obtained, based on observations at time $t$. This optimally controlled output is the same as an $L$-step ahead forecast error with a MA($L-1$) model configuration.
Chapter 8

PLANT-SITE DATA COLLECTION PROCEDURE

To test the hypothesis that maintenance manhour requirement could be forecast, actual plant-site data was sought. Recognizing that at least sixty observations would be needed for multiple time series analysis, the variable list was pared to those variables for which it was assumed that data might be most readily available at the plant site. The variables were (1), maintenance manhour requirement (MANHOURS) (2), backlog (BACKLOG) and (3), production-to-capacity ratio (RATIO). Based on those variables a model might look like:

\[ MANHOURS_t = MANHOURS_{t-L} + BACKLOG_{t-L} + RATIO_{t-L} \]  

where:

\( L \) = units of lag.

A sample explanatory letter, a request for data, and forms for the data collection were sent to personnel at both the plant level and at the corporate level. Based on the response at the corporate level, it was decided to target the mailing to individual plant managers for distribution to the appropriate maintenance managers. Forty six letters were mailed to the managers of chemical processing plants.
located in South Louisiana. The explanatory letters, the data collection forms, and a summary of the answers received may be found in the Appendix.

Little or no actual data was received in response to the information packet mailing. However, insight was gained from those responses received. Two conclusions which can be drawn are (1), data has not been collected and retained for the time span required for a multiple time series analysis and (2), competition, environmental concerns and a stressed chemical industrial economy have caused much actual plant site data to be strictly proprietorized. The responses also indicate that in most cases, quantitative analysis is not being applied to forecasting maintenance manhour requirement at the plant level. While several plant managers thought that there might be a relationship between production level and maintenance manhour requirement, no plant manager provided any quantitative result to prove such a relationship.
Chapter 9

AN EXAMPLE USING DERIVED DATA

Published commodity and labor statistics were used to develop an example data base to be input to multiple time series forecasting procedures. The generality of this forecasting method best accommodated the data base, as shown previously, in the review of comparative methods. The purpose is to illustrate the output available from current multiple time series forecasting software.

Development of the Data Base

For forecasting the dependent variable, maintenance manhour requirement, data for three independent variables was developed. The independent variables are: past maintenance manhours, past ratio of production-to-capacity expressed as hours of operating time, and past number of hours of maintenance backlog. Ninety-six data points per independent variable were included for the January 1977 through December 1984 time span. Sixty data points are the minimum suggested for Box-Jenkins time series forecasting. Ninety-six data points permit an ample historical base to be developed, using 84 data points and provide 12 additional data points upon which to judge the efficiency of the forecasting model.
Labor statistics for the Baton Rouge area provided by the Division of Research and Development at Louisiana State University were used to develop the past maintenance manhour requirement figures. Based on an average of 200 employees per chemical processing unit, it was assumed that approximately 80 chemical processing units were operating in the Baton Rouge throughout the time span. From interviews with processing plant personnel, it was estimated that 25% of those employed within a chemical processing unit are assigned to the plant maintenance staff. From calendars specific to each year, the number of working days per month was calculated for the time span. It was assumed that maintenance personnel worked eight hour days. An example calculation is shown below for January, 1977:

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Baton Rouge chemical and petroleum product employees:</td>
<td>14,775</td>
</tr>
<tr>
<td>Average number of chemical processing units in the area:</td>
<td>80</td>
</tr>
<tr>
<td>Percentage of employees in maintenance:</td>
<td>25%</td>
</tr>
<tr>
<td>Number of working days for January, 1977:</td>
<td>20</td>
</tr>
</tbody>
</table>
Number of working hours
per day: 8

\[
\left( \frac{14,777}{80} \right) \times 0.25 = 46 \text{ Maintenance Employees/Unit}
\]

\[
46 \times 20 \times 8 = 7360 \text{ Manhours/Unit/Month}
\]

(Manhour Requirement)

For the data base calculations, the number of maintenance employees per unit was expressed to one decimal place in order to provide a fourth significant figure for the number of maintenance manhours. Table 9.1 shows the derived maintenance manhour requirement data for the years 1977 through 1984.

Without actual plant operating level data to support or refute the argument, it was theorized that a correlation could be made between refinery operating levels and those at the downstream chemical derivative production units. Thus, if the refinery was running at a high level, a high amount of ethylene might be produced. To make full use of the available ethylene feedstock, the downstream derivative production unit might have to also operate at a high level. Thus, crude petroleum refinery operating ratios as a percent of capacity were adapted for the example data base from
<table>
<thead>
<tr>
<th>YR/MO</th>
<th>JAN</th>
<th>FEB</th>
<th>MAR</th>
<th>APR</th>
<th>MAY</th>
<th>JUN</th>
<th>JULY</th>
<th>AUG</th>
<th>SEPT</th>
<th>OCT</th>
<th>NOV</th>
<th>DEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>7392</td>
<td>7456</td>
<td>8611</td>
<td>6905</td>
<td>7304</td>
<td>7304</td>
<td>7632</td>
<td>8777</td>
<td>7997</td>
<td>7913</td>
<td>7980</td>
<td>7997</td>
</tr>
<tr>
<td>1978</td>
<td>8047</td>
<td>7664</td>
<td>8814</td>
<td>7808</td>
<td>9016</td>
<td>8659</td>
<td>7872</td>
<td>9034</td>
<td>7856</td>
<td>8518</td>
<td>8165</td>
<td>7760</td>
</tr>
<tr>
<td>1979</td>
<td>8589</td>
<td>7696</td>
<td>8413</td>
<td>7930</td>
<td>8630</td>
<td>7980</td>
<td>7879</td>
<td>8610</td>
<td>7129</td>
<td>8795</td>
<td>8881</td>
<td>7696</td>
</tr>
<tr>
<td>1980</td>
<td>8360</td>
<td>7930</td>
<td>7930</td>
<td>8466</td>
<td>8518</td>
<td>8199</td>
<td>8518</td>
<td>8081</td>
<td>8299</td>
<td>8906</td>
<td>7357</td>
<td>8518</td>
</tr>
<tr>
<td>1981</td>
<td>7930</td>
<td>7552</td>
<td>8254</td>
<td>8413</td>
<td>8081</td>
<td>8466</td>
<td>8747</td>
<td>8350</td>
<td>8299</td>
<td>8800</td>
<td>8000</td>
<td>8747</td>
</tr>
<tr>
<td>1982</td>
<td>8000</td>
<td>7904</td>
<td>8850</td>
<td>8466</td>
<td>8081</td>
<td>8466</td>
<td>8081</td>
<td>8307</td>
<td>7829</td>
<td>7829</td>
<td>7829</td>
<td>8078</td>
</tr>
<tr>
<td>1983</td>
<td>7248</td>
<td>7200</td>
<td>8225</td>
<td>7291</td>
<td>7480</td>
<td>7480</td>
<td>6800</td>
<td>7710</td>
<td>6989</td>
<td>6989</td>
<td>6989</td>
<td>6871</td>
</tr>
<tr>
<td>1984</td>
<td>6821</td>
<td>6770</td>
<td>7093</td>
<td>6821</td>
<td>7093</td>
<td>6770</td>
<td>6720</td>
<td>7360</td>
<td>6034</td>
<td>7250</td>
<td>6619</td>
<td>6256</td>
</tr>
</tbody>
</table>
published statistics found in the 1977-1984 Commodity Year Books. The operating ratios were used to calculate the number of hours of 100% operating capacity per month for a continuous (24 hour) operation. Thus a production/capacity ratio of 89% for January, 1977 in which 744 hours of 100% capacity operation were available translated into 662 actual hours of 100% capacity operation. Table 9.2 shows the derived ratio data for the years 1977 through 1984.

For the example, the number of hours of backlog was assumed to increase as the hours of 100% capacity operation increased. As a benchmark, it was assumed that when the plant was shut down for turnaround, the number of hours of backlog decreased to zero before the plant was placed back into operation. Using the operating ratio data, and a two week (84 hour/maintenance worker) backlog of work for the maintenance workforce calculated to be in place during the preceding month, example backlog data was calculated. As calculated, then, the example backlog data was correlated with maintenance manhour requirement at a one month lag time and with the operating ratio at no time lag. An example calculation is shown for January, 1977:

Number of Maintenance Employees Per Unit: 46
Table 9.2

Production/Capacity Ratio in Hours

<table>
<thead>
<tr>
<th>YR/NO</th>
<th>JAN</th>
<th>FEB</th>
<th>MAR</th>
<th>APR</th>
<th>MAY</th>
<th>JUN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>662</td>
<td>625</td>
<td>670</td>
<td>641</td>
<td>662</td>
<td>655</td>
</tr>
<tr>
<td>1978</td>
<td>632</td>
<td>564</td>
<td>632</td>
<td>598</td>
<td>662</td>
<td>634</td>
</tr>
<tr>
<td>1979</td>
<td>640</td>
<td>564</td>
<td>618</td>
<td>605</td>
<td>625</td>
<td>619</td>
</tr>
<tr>
<td>1980</td>
<td>610</td>
<td>538</td>
<td>573</td>
<td>547</td>
<td>558</td>
<td>554</td>
</tr>
<tr>
<td>1981</td>
<td>536</td>
<td>477</td>
<td>506</td>
<td>475</td>
<td>498</td>
<td>490</td>
</tr>
<tr>
<td>1982</td>
<td>498</td>
<td>437</td>
<td>491</td>
<td>475</td>
<td>513</td>
<td>540</td>
</tr>
<tr>
<td>1983</td>
<td>506</td>
<td>437</td>
<td>491</td>
<td>497</td>
<td>536</td>
<td>540</td>
</tr>
<tr>
<td>1984</td>
<td>543</td>
<td>511</td>
<td>558</td>
<td>549</td>
<td>573</td>
<td>554</td>
</tr>
<tr>
<td>YR/NO</td>
<td>JULY</td>
<td>AUG</td>
<td>SEPT</td>
<td>OCT</td>
<td>NOV</td>
<td>DEC</td>
</tr>
<tr>
<td>1977</td>
<td>677</td>
<td>670</td>
<td>655</td>
<td>662</td>
<td>641</td>
<td>655</td>
</tr>
<tr>
<td>1978</td>
<td>655</td>
<td>677</td>
<td>648</td>
<td>662</td>
<td>655</td>
<td>670</td>
</tr>
<tr>
<td>1979</td>
<td>647</td>
<td>640</td>
<td>605</td>
<td>618</td>
<td>605</td>
<td>632</td>
</tr>
<tr>
<td>1980</td>
<td>551</td>
<td>543</td>
<td>533</td>
<td>528</td>
<td>526</td>
<td>565</td>
</tr>
<tr>
<td>1981</td>
<td>498</td>
<td>528</td>
<td>490</td>
<td>498</td>
<td>490</td>
<td>513</td>
</tr>
<tr>
<td>1982</td>
<td>558</td>
<td>528</td>
<td>533</td>
<td>528</td>
<td>511</td>
<td>521</td>
</tr>
<tr>
<td>1983</td>
<td>558</td>
<td>551</td>
<td>547</td>
<td>543</td>
<td>540</td>
<td>521</td>
</tr>
<tr>
<td>1984</td>
<td>565</td>
<td>580</td>
<td>562</td>
<td>565</td>
<td>554</td>
<td>558</td>
</tr>
</tbody>
</table>
Backlog Hours Per Employee: 84
Operating Ratio for Previous Month: 91%

\[ 46 \times 84 \times 0.91 = 3520 \text{ Hours of Backlog} \]

Table 9.3 shows the derived backlog data for the years 1977 through 1984.

For a second example, maintenance manhour requirement data was generated, by correlation, from backlog data. It was theorized that manhour requirement might respond inversely to a change in backlog with a one month lag. It was also theorized that the one month lag might be somewhat less than 100% responsive. That is, as a planning hedge, a maintenance manager might base his manhour requirement hours for the next month on only 80% of the change in the number of backlog hours averaged over the previous two months. To reflect the theorized relationship between manhour requirement and backlog and to incorporate a planning hedge of 80%, the following model was developed and used to derive the maintenance manhour requirement data:

\[
M = \frac{[(B_{t-2} - B_t) + (B_{t-1} - B_t)]/2 \times 0.8}{2} + /- \frac{M}{t-1} \tag{9.1}
\]
Table 9.3

Backlog Hours

<table>
<thead>
<tr>
<th>YR/MO</th>
<th>JAN</th>
<th>FEB</th>
<th>MAR</th>
<th>APR</th>
<th>MAY</th>
<th>JUNE</th>
<th>JULY</th>
<th>AUG</th>
<th>SEPT</th>
<th>OCT</th>
<th>NOV</th>
<th>DEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>3881</td>
<td>3484</td>
<td>3656</td>
<td>3107</td>
<td>3103</td>
<td>3103</td>
<td>3646</td>
<td>3646</td>
<td>3599</td>
<td>3600</td>
<td>3551</td>
<td>3559</td>
</tr>
<tr>
<td>1978</td>
<td>3540</td>
<td>3420</td>
<td>3380</td>
<td>3484</td>
<td>3416</td>
<td>3678</td>
<td>3637</td>
<td>3629</td>
<td>3753</td>
<td>3659</td>
<td>3633</td>
<td>3707</td>
</tr>
<tr>
<td>1979</td>
<td>3659</td>
<td>3475</td>
<td>3373</td>
<td>3291</td>
<td>3309</td>
<td>3352</td>
<td>3388</td>
<td>3427</td>
<td>3308</td>
<td>3372</td>
<td>3353</td>
<td>3394</td>
</tr>
<tr>
<td>1980</td>
<td>3392</td>
<td>3251</td>
<td>3172</td>
<td>3111</td>
<td>3090</td>
<td>3074</td>
<td>3156</td>
<td>2990</td>
<td>3029</td>
<td>3008</td>
<td>2987</td>
<td>2968</td>
</tr>
<tr>
<td>1981</td>
<td>3013</td>
<td>2855</td>
<td>2797</td>
<td>2730</td>
<td>2667</td>
<td>2797</td>
<td>2839</td>
<td>2797</td>
<td>2946</td>
<td>2856</td>
<td>2814</td>
<td>2839</td>
</tr>
<tr>
<td>1982</td>
<td>2898</td>
<td>2760</td>
<td>2626</td>
<td>2667</td>
<td>2667</td>
<td>2788</td>
<td>3030</td>
<td>2974</td>
<td>2779</td>
<td>2897</td>
<td>2779</td>
<td>2737</td>
</tr>
<tr>
<td>1983</td>
<td>2664</td>
<td>2570</td>
<td>2441</td>
<td>2406</td>
<td>2463</td>
<td>2570</td>
<td>2678</td>
<td>2640</td>
<td>2586</td>
<td>2656</td>
<td>2551</td>
<td>2577</td>
</tr>
<tr>
<td>1984</td>
<td>2387</td>
<td>2471</td>
<td>2573</td>
<td>2558</td>
<td>2539</td>
<td>2607</td>
<td>2607</td>
<td>2587</td>
<td>2554</td>
<td>2601</td>
<td>2581</td>
<td>2529</td>
</tr>
</tbody>
</table>

107
where:

M is maintenance manhours,

B is backlog hours, and

t is time.

For the calculations, the sign determination for $M_{t-1}$ followed the inverse relationship between maintenance manhour requirement and backlog. The derived maintenance manhours data is shown in Table 9.4. An example of this calculation for maintenance manhour requirement for April 1977 follows:

Maintenance manhour requirement for March, 1977: 7018
Backlog Hours for January, 1977: 3881
Backlog Hours for February, 1977: 3484
Backlog Hours for March, 1977: 3656
Backlog Hours for April, 1977: 3107

$M = \frac{[(3656-3484) + (3656-3107)]}{2} \times 0.8 + 7018$ Hours
Table 9.4
Maintenance Manhours
(Example 2)

<table>
<thead>
<tr>
<th>YR/MO</th>
<th>JAN</th>
<th>FEB</th>
<th>MAR</th>
<th>APR</th>
<th>MAY</th>
<th>JUN</th>
<th>JULY</th>
<th>AUG</th>
<th>SEPT</th>
<th>OCT</th>
<th>NOV</th>
<th>DEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>6850</td>
<td>7078</td>
<td>7018</td>
<td>7314</td>
<td>7176</td>
<td>7178</td>
<td>6892</td>
<td>6910</td>
<td>6929</td>
<td>6989</td>
<td>6955</td>
<td>6944</td>
</tr>
<tr>
<td>1978</td>
<td>6934</td>
<td>6910</td>
<td>7025</td>
<td>6956</td>
<td>7220</td>
<td>7091</td>
<td>7111</td>
<td>7164</td>
<td>7077</td>
<td>7125</td>
<td>7165</td>
<td>7116</td>
</tr>
<tr>
<td>1979</td>
<td>7209</td>
<td>7323</td>
<td>7397</td>
<td>7437</td>
<td>7413</td>
<td>7381</td>
<td>7351</td>
<td>7328</td>
<td>7342</td>
<td>7327</td>
<td>7351</td>
<td>7334</td>
</tr>
<tr>
<td>1980</td>
<td>7391</td>
<td>7479</td>
<td>7535</td>
<td>7568</td>
<td>7583</td>
<td>7622</td>
<td>7523</td>
<td>7605</td>
<td>7581</td>
<td>7524</td>
<td>7443</td>
<td>7393</td>
</tr>
<tr>
<td>1981</td>
<td>7312</td>
<td>7398</td>
<td>7448</td>
<td>7500</td>
<td>7577</td>
<td>7508</td>
<td>7474</td>
<td>7550</td>
<td>7454</td>
<td>7507</td>
<td>7534</td>
<td>7500</td>
</tr>
<tr>
<td>1982</td>
<td>7429</td>
<td>7538</td>
<td>7616</td>
<td>7600</td>
<td>7590</td>
<td>7445</td>
<td>7326</td>
<td>7426</td>
<td>7551</td>
<td>7457</td>
<td>7521</td>
<td>7567</td>
</tr>
<tr>
<td>1983</td>
<td>7634</td>
<td>7723</td>
<td>7787</td>
<td>7826</td>
<td>7760</td>
<td>7674</td>
<td>7616</td>
<td>7653</td>
<td>7703</td>
<td>7633</td>
<td>7685</td>
<td>7599</td>
</tr>
<tr>
<td>1984</td>
<td>7709</td>
<td>7635</td>
<td>7588</td>
<td>7602</td>
<td>7637</td>
<td>7610</td>
<td>7618</td>
<td>7639</td>
<td>7671</td>
<td>7644</td>
<td>7673</td>
<td>7694</td>
</tr>
</tbody>
</table>
Review of the Computer Software Used

The data base was input to SCA using SAS and SCALINK, a utility macro program developed by Houston H. Stokes. SCALINK provides a convenient method for bringing together within a single program the strong graphics capabilities of SAS and the time series analysis versatility of SCA. LINKSCA, also developed by Houston L. Stokes, is available as an access route from SCA to SAS.

The programming paragraphs used in SCA are similar in form to those of BIOMED. Documentation is clearly written and the analysis examples are helpful. The chapter introduction for multivariate time series analysis includes background theoretical discussion. Since multivariate time series analysis is still in the 'research' stage, the applied analysis sources are scattered throughout the periodical literature. The manual, therefore, brings together in a single document, a discussion of the theoretical basis for multivariate time series analysis, a description of the analysis procedures used and documented examples of how the procedures are used to analyze data.
An Analysis of the Results

Scatter plots of two of the three variables, MANHOURS and RATIO suggested that mathematical transformation (viz: log, reciprocal, etc.) would not be needed to stabilize the variances. A seasonal periodicity of 12-14 was suggested in the plot for RATIO. Individual ACF plots showed persistent high spikes for each of the three variables, suggesting non-stationarity of the vectored series. Therefore, each series was differenced at the following lags: (1). lag 6; (2). lags 1 and 6; (3). lag 12; and (4). lags 1 and 12. Sufficient stationarity, where the ACF and PACF plots showed identifiable patterns, was achieved only with differencing at lags 1 and 12. Each series may be differenced individually with SCA. Additional model complexity is introduced when the series are differenced but stationarity is required for the identification steps. The ESCCM (Extended Sample Cross Correlation Matrix) is not currently available on SCA for multiple time series analysis. [51] Research has shown that ESCCM patterns are useful for mixed (ARMA) parameter identification when the series have been individually, rather than jointly, differenced.
The CCM plot for the vectored series, differenced at lags 1 and 12, did not "cut off" at any particular lag. Furthermore, the CCM plot did not resemble the MA(q) pattern shown in the SCA manual. Indicator symbols, rather than the numerical values, for the elements of the CCM's aided pattern identification.

Autoregressive parameters were fit to the example vectored series over lags 1 through 14. Over the 14 lags, the chi-square statistic exhibited a "cut-off" pattern after the lag at 2. The M(L) statistic values increased at 8 through 12, consistent with a seasonal autoregressive identification pattern. For 9 degrees of freedom, chi-square critical values at 95% and at 99% confidence intervals are 16.9 and 21.7, respectively. Based on the preceding critical values, the autoregressive parameters at lags 1, 2, 10, and 12 are significant. When autoregressive parameters are fit to lags 1, 2, and 10, the parameter at 10 is shown to be insignificant; when autoregressive parameters are fit to lags 1, 2, and 12, all parameters are shown to be significant.

Based on the preceding identification procedure, the following model was estimated using the conditional likelihood method:
\[ (1-B)(1-B^2)(1-\phi B-\phi B)(1-\phi^2 B)Z = C + a \quad (9.2) \]

where:

- \( B \) is the backshift operator,
- \( \phi \) is the estimated autoregressive parameter,
- \( C \) is the constant or trend term (included because a non-zero mean still existed after differencing),
- \( a \) is the error term, and
- \( Z \) is the vectored series, comprised of \( M \), \( R \), and \( B \), where:

  - \( M \) is MANHOURS,
  - \( R \) is RATIO, and
  - \( B \) is BACKLOG.

The cross correlation matrix of the residuals from the estimated model showed no significant values at the first 7 lags. A histogram of the residuals showed the following distribution with a mean of zero:
<table>
<thead>
<tr>
<th>Cell</th>
<th>Observed</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>9</td>
</tr>
</tbody>
</table>

From:

\[ X^2 = \sum \frac{(O - E)^2}{E} \]  \hspace{1cm} (9.3)

where:

\[ X^2 \] is the chi-square statistic,

\[ O \] is the observed frequency of occurrence, and

\[ E \] is the expected frequency of occurrence.

The chi-square statistic may be calculated for use in a goodness-of-fit test. The null hypothesis, \( H_0 \), is that the observed frequency of occurrence follows a normal distribution. The alternate hypothesis, \( H_1 \), is that the observed frequency of occurrence is not a normal distribution. The calculated chi-square statistic was
4.929. A book value chi-square with N-3 degrees of freedom was 7.815 at 95% confidence interval. There was no reason to reject the null hypothesis that the residuals were normally distributed and the model was deemed adequate.

Based on the parameters which were estimated, those elements in the estimated parameter matrix which were shown to be either insignificant or which were shown to have small values compared to the associated standard deviation, were constrained to zero. The parameters were re-estimated with the constraints. The resulting model was:

\[
(1-B)(1-B^2)(1 + 0.957B + 0.713B^2)M = -46.053 + a (9.4a) \\
(1-B)(1-B^2)(1-0.070B)B(1+0.179B^2)R = 0.69 + a (9.4b) \\
(1-B)(1-B^2)(1-4.754B)R(1+0.195B)B = -5.496 + a (9.4c)
\]

where:

\[ B, M, R, B, \text{ and } a \] are as previously defined.

Forecast values are generated from the constrained model.
As shown in the preceding model, the MANHOURS series is independent of both RATIO and BACKLOG. A univariate analysis on the MANHOURS series would have produced similar results. However, this example does demonstrate that the multiple time series analysis did not introduce any interrelationships that had not been structured into the data base. Although theoretically there might have been some correlation between the MANHOURS data which had been derived from Louisiana chemical employment statistics and the RATIO data which had been drawn from national oil refinery operating statistics, no significant correlation was demonstrated by the modelling procedure. The relationship between RATIO and BACKLOG is demonstrated in the model. As described previously, BACKLOG was directly derived from RATIO for the example data base. If a correlation had been shown between MANHOURS, RATIO and BACKLOG, a single vectored equation could be written in terms of MANHOURS. Such a single output equation could have been used to forecast a three membered input vector.

For the second example, after differencing at lags 1 and 12, the CCM pattern showed significant spikes at lags 1, 2, 11, and 12, suggesting a seasonal MA(2) model.
Autoregressive parameters fit in succession, showed significance, when the M(L) statistics were compared to the chi-square values, at lags 1, 2, and 11, suggesting a seasonal AR(2) model. Parameters were estimated by the conditional likelihood method for the following vectored ARIMA model:

\[ s \frac{2}{s} (1-B)(1-B)(1-\phi B - \phi B)(1-\phi B) Z = \]
\[ 1 2 s t \]
\[ C + (1-\theta B - \theta B)(1-\theta B) a \]
\[ 1 2 s t \]

(9.5)

where:

- B is the backshift operator,
- \( \phi \) are the autoregressive parameters,
- \( \theta \) are the moving average parameters,
- C is a constant value,
- a is the error term,
- t is time, and
- s is the seasonal lag for s = 11 or 12.

Values for "s" were varied in an effort to find the best model. The residuals from the model did not have a mean of zero, which suggested an overfitting of parameters. Therefore, parameters were estimated for the following models as a comparison:
(1-B)(1-B)(1- \Phi B - \Phi B)(1- \Phi B) S_t = C + (1- \Theta B - \Theta B) a_s (9.6)

(1-B)(1-B)(1- \Phi B - \Phi B)(1- \Phi B) S_t = C + (1- \Theta B)(1- \Theta B) a_s (9.7)

(1-B)(1-B)(1- \Phi B - \Phi B)(1- \Phi B) S_t = C + (1- \Theta B) a_s (9.8)

(1-B)(1-B)(1- \Phi B - \Phi B)(1- \Phi B) S_t = C + a_t (9.9)

(1-B)(1-B) S_t = C + (1- \Theta B - \Theta B)(1- \Theta B) a_s (9.10)

Residual zero means occurred only in the model shown in equation 9.9 when differenced at a lag of 12 and with the
seasonal parameter at either lag 11 or at lag 12. Subsequent diagnostic checking showed that a better model resulted when the autoregressive parameter was estimated for lag 12. From the parameter estimation step, the final model was:

\[
\begin{align*}
&12 \\
& (1-B)(1-B)(1 + 2.755R - 0.483B + 1.754R - 0.275B ) \\
& \quad t-1 t-2 \\
& (1 + 0.336M )M = 0.946 + a \quad (9.11a) \\
& \quad t-12 \quad t \\
& \\
& 12 \\
& (1-B)(1-B)(1 + 0.137M + 0.099M - 0.147B ) \\
& \quad t-1 t-2 \\
& (1 + 0.16R )R = 0.382 + a \quad (9.11b) \\
& \quad t-12 \quad t \\
& \\
& 12 \\
& (1-B)(1-B)(1 - 4.363R + 0.319B ) \\
& \quad t-1 t-1 \\
& (1 + 0.241B )B = -5.903 + a \quad (9.11c) \\
& \quad t-12 \quad t
\end{align*}
\]

where:

- \( M \) is MANHOURS,
- \( R \) is RATIO, and
- \( B \) is BACKLOG.
MANHOURS is shown to be a function of RATIO and BACKLOG at lags 1 and 2. The \(-0.8\) correlation between MANHOURS and BACKLOG in the original data is reduced to \(-0.45\) by the model. Both the lagged relationships and the high negative correlation between MANHOURS and BACKLOG were to be expected from the derived data. The inverse relationship between MANHOURS and BACKLOG is noted by the opposite mathematical signs for the estimated parameters preceding these two variables in equations 9.11a and 9.11b.

The CCM for the residuals from the preceding model, show significant spikes at lag 3 for the cross correlation of MANHOURS:BACKLOG and for RATIO:BACKLOG and at lag 7 for MANHOURS:RATIO. A histogram of the residuals showed the following frequency distribution:

<table>
<thead>
<tr>
<th>Cell</th>
<th>Observed</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>17</td>
</tr>
</tbody>
</table>
The calculated chi-square value with 1 degree of freedom was 7.879. The chi-square book value for an alpha value of 0.005 is 7.87944. At higher alpha values, the calculated chi-square value exceeded the book values. The determinants for the estimated $\phi_1$ and $\phi_2$ parameters were calculated and were -0.143 and -0.0102, respectively. For a vectored AR(2) model to be stationary, the determinants of the estimated parameters must satisfy the following:

$$\phi_1 + \phi_2 < 1$$

$$\phi_1 - \phi_2 < 1$$

$$1 < \phi_2 < 1$$

The calculated values for the parameter determinants satisfied the stationarity condition.

To simplify the model, those parameters, which were small in relation to the magnitude of the standard error, were constrained to zero. The constrained model follows:

$$12$$

$$(1 - B)(1 - B)(1 + 2.852R)$$

$$t-l$$

$$(1 + 0.293M)M = -3.203 + a$$

$$(9.12a)$$

$$(t-12, t, t)$$

$$(t)$$
The cross correlation matrices values over 20 lags for the residuals from the constrained model were used to calculate the Portmanteau and the modified Portmanteau statistics. With an N value of 57, the Portmanteau value was 70 and the modified Portmanteau value was 84. These values are compared to a chi-square distribution with \( M(S-p-q) \) or \( 9\times(20-2-0) \) degrees of freedom. (The seasonal autoregression parameter is not considered as part of \( p \), just as it is not for the univariate Portmanteau statistic calculation). A chi-square distribution approaches a normal distribution as the number of degrees of freedom approaches infinity. When the number of degrees of freedom exceeds 2, the theoretical mean of a chi-square distribution is equal to the number of degrees of freedom. With increased degrees of freedom, skewness decreases. A chi-square
distribution with N-1 degrees of freedom will be contained 'within' a chi-square distribution with N degrees of freedom for N > 2. From Figure 9.1 it is shown that the chi-square distributions with 70 or 84 degrees of freedom would fall within the critical area for a chi-square distribution with 162 degrees of freedom. Critical values for chi-square distributions with degrees of freedom in excess of 30, may be approximated from the following:

\[
Z = \left( \frac{2}{2*\chi} \right)^{0.5} - (2*DF - 1)^{0.5} \tag{9.13}
\]

where:

- Z is the Normal Distribution z score (For 0.05, z = 1.645),
- DF is the number of degrees of freedom.

The critical value for 0.05 alpha for a chi-square distribution with 162 degrees of freedom is 192.42. With 70 degrees of freedom and 0.05 alpha, the critical value is 90.06. For 84 degrees of freedom and 0.05 alpha, the critical value is 105.55. The Portmanteau and the modified Portmanteau distributions have critical values at 0.05 alpha which are less than that for the comparative chi-square distribution for 162 degrees of freedom. Based on the
Figure 9.1

(17)
Chi-square Distributions for
Various Degrees of Freedom
preceding result, even the constrained model, with its theoretically higher valued residuals than those which might result from the non-constrained model, may be considered acceptable. But, based on the preceding, the utility of the Portmanteau test for multiple times series diagnostic checking might be questioned. The size of the matrix, that is the number of input series, increases the comparative chi-square distribution by a factor of $m^2$. The number of observations, $N$, is also directly proportional to the calculated Portmanteau or modified Portmanteau values. For the constrained residual matrices, a value of 1.22528 was obtained. To approach a chi-square distribution of 162 degrees of freedom, 132 observations rather than 57 observations would be required. (The residual matrices sum for the non-constrained model was 0.7052. To approach a chi-square distribution of 162 degrees of freedom, 230 observations would be required.) Thus, observation size and the number of input time series may limit the utility (sensitivity) of the Portmanteau and the modified Portmanteau as diagnostic checking procedures.

Based on the diagnostic checking procedures, the seasonal AR(2) model is only marginally acceptable. Forecast values are calculated from the parameter-
constrained model for 24 periods, from an origin at the eighty fourth observation. The forecast values are compared with the actual derived values from observations at 85 through 96. The forecast values, actual values, and standard errors are shown in Table 9.5. In Figures 9.2 and 9.3, the actual forecast values for both examples are plotted. The forecast values generated from the model developed for the first example appear to "track" the actual values up to the observation 93. Recalling that the reduced model developed in the first example was similar to a univariate model, and that the model fell well within the aptness tests, the results shown in the overlay graph of Figure 9.2 could be expected. The model developed for the second example met the aptness criteria only marginally. The less apt fit is evident in the poorer "tracking" of the forecast to the actual values. To improve the model, attention might be directed toward the residual cross correlation matrix for the full model where significant spikes occur at lag three. These spikes suggest that the intervariable relationships between BACKLOG and MANHOURS and between BACKLOG and RATIO are not "fully explained" by the model. An examination of the plots of the autocorrelations and partial autocorrelations for each input variable may
Table 9.5
Forecast and Actual Values for Examples 1 and 2

<table>
<thead>
<tr>
<th>OBSERVATION</th>
<th>FORECAST 1</th>
<th>ACTUAL 1</th>
<th>FORECAST 2</th>
<th>ACTUAL 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>6606</td>
<td>6821</td>
<td>7709</td>
<td>7709</td>
</tr>
<tr>
<td>86</td>
<td>6411</td>
<td>6770</td>
<td>7754</td>
<td>7635</td>
</tr>
<tr>
<td>87</td>
<td>7158</td>
<td>7093</td>
<td>7794</td>
<td>7588</td>
</tr>
<tr>
<td>88</td>
<td>6627</td>
<td>6821</td>
<td>7862</td>
<td>7602</td>
</tr>
<tr>
<td>89</td>
<td>6540</td>
<td>7093</td>
<td>7780</td>
<td>7637</td>
</tr>
<tr>
<td>90</td>
<td>6504</td>
<td>6770</td>
<td>7658</td>
<td>7610</td>
</tr>
<tr>
<td>91</td>
<td>6123</td>
<td>6720</td>
<td>7597</td>
<td>7618</td>
</tr>
<tr>
<td>92</td>
<td>6761</td>
<td>7360</td>
<td>7649</td>
<td>7639</td>
</tr>
<tr>
<td>93</td>
<td>6091</td>
<td>6034</td>
<td>7706</td>
<td>7671</td>
</tr>
<tr>
<td>94</td>
<td>6130</td>
<td>7250</td>
<td>7630</td>
<td>7644</td>
</tr>
<tr>
<td>95</td>
<td>5939</td>
<td>6619</td>
<td>7683</td>
<td>7673</td>
</tr>
<tr>
<td>96</td>
<td>5900</td>
<td>6256</td>
<td>7617</td>
<td>7694</td>
</tr>
</tbody>
</table>

* MONTHS
Figure 9.2
Forecast Versus Actual Manhours

ACTUAL MANHOURS   +  FORECAST MANHOURS
Figure 9.3

Forecast Versus Actual Manhours
(Example 2)
suggest which parameters may be added to the model. These examples were developed primarily to provide an opportunity to use the SCA software and to gain some actual experience in multiple time series modelling.
Variables which theoretically may impact chemical processing plant maintenance manhour requirement were identified. The direct or indirect impact which each variable might have on the maintenance manhour requirement was considered. The major, currently used, technological forecasting methods were assessed for applicability to forecast chemical processing plant maintenance manhour requirement. The assessment of forecasting methods included both the econometric and the time series approaches. Mathematically, the vectored ARIMA model is equivalent to the reduced form of a structural linear simultaneous equation set.

Innovative aspects of the study include: (1) the consideration of those variables which might impact the maintenance manhour requirement in the chemical processing environment; (2) the assessment of forecasting techniques in relation to the maintenance manhour requirement in the chemical processing environment; (3) the addition of the econometrics approach to suggest which variables might be included in the quantitative (statistical) procedures used to develop a forecasting model; and (4) the use of the
econometrics approach to assess the inclusion of variables after a forecasting model has been developed. Based on the study of the variables which may impact the maintenance manhour requirement in the chemical processing industry, followed by the assessment of forecasting techniques and applications of those techniques to similar environments, it was theorized that maintenance manhour requirement may be forecast using multiple time series analysis. A review of the published literature indicated that the use of multiple time series analysis to forecast maintenance manhour requirement in the chemical processing industry is unique. Interviews with chemical processing plant maintenance management personnel support the uniqueness of forecasting maintenance manhour requirement using multiple time series analysis. Interviews suggested that adequate data retention systems for multiple time series analysis use are being used in the chemical processing industry at the production plant level. This study suggests that in-plant data collection and retention capability may be added to established plant maintenance management or process control systems. Example data collection forms and a questionnaire are included in the Appendix.

To demonstrate the multiple time series analysis procedure, a theoretical chemical processing plant
maintenance manhour forecasting model was developed from derived data. The model was based on three input series, past manhour maintenance requirement, past plant capacity-to-production ratio, and past maintenance backlog. These were chosen because the variable identification procedure indicated that these three variables seem to directly impact the maintenance manhour requirement. Further, it was theorized that data for these particular variables might be more accessible in an in-plant setting. To develop the chemical processing plant maintenance manhour requirement forecasting model, econometrics was used in the variable identification and impact assessment phase and time series analysis methods were used for the model calculation phase. Mathematically, the diagnostic checking of the model was a time series analysis procedure. Econometrics provided guidance on whether the final model made sense on a theoretical basis. The incorporation of both analysis approaches was useful in developing the theoretical forecasting model. Such an incorporation might be useful in actual forecasting model development.

By necessity, derived data had to be used to demonstrate the forecast modelling procedure. With increased automation of chemical processing plant maintenance systems, sufficient and reliable data for the
actual forecasting model development should become more readily available. To insure that the appropriate data is collected and retained for an adequate period of time, analytical foresight will be needed at the chemical processing plant level.

As more applications of multiple time series analysis are published, the CCM pattern recognition required for the identification and the diagnostic checking steps will become more easily accomplished by the analyst. Experience with the methodology and the software is vital for accurate analysis.

A procedure for forecasting chemical processing plant maintenance manhour requirement has been described. Technology transfer to the processing plant will require a sharing of data among production, maintenance, analytical, engineering, and financial personnel. It may be difficult to convince management to consider using multiple time series analysis because mathematically and notationally it is more complex than simple regression analysis or simple exponential smoothing. Further research might consider technology transfer methods which might be useful to bring the improved, but mathematically more complex, forecasting techniques into a plant processing environment. Another research pursuit might consider the relationship between
production/capacity level and maintenance requirement. Conflicting ideas about whether any relationship does exist and whether there is an increased or decreased maintenance requirement if the relationship does exist, were expressed during interviews by plant maintenance personnel. The consideration of whether and to what extent maintenance manhour requirement might be sales or production driven is another area for further research. Sales and production data are often retained over long periods of time and might be readily accessed for use as leading predictive indicators for determining maintenance manhour requirement.


30. Louisiana State University, College of Business Administration, Division of Research and Development. Unpublished data format. 1974-1984.


Sample data request packets were reviewed by chemical industry personnel employed at corporate and plant levels and by research personnel at the Division of Business Research at Louisiana State University. The packet was revised and copies of the packet were mailed to 46 chemical processing plant managers. The packet included the following:

(1). A transmittal letter addressed to the plant manager;
(2). A scenario, designed to persuade the plant manager to participate in the study and to support the study at the plant level;
(3). A confirmation form to acknowledge that the packet had been received and had been distributed to the plant maintenance manager;
(4). An outlined statement to reiterate what information was being requested;
(5). A transmittal letter to be directed to the chemical plant maintenance manager;
(6). Data collection forms for Maintenance Manhours/Month, Ratio of Production to Capacity/Month, and Backlog Hours/Month; and
(7). A questionnaire, designed to elicit more qualitative, background operating information.
We need actual process plant data to help us to develop a realistic chemical plant maintenance manpower forecasting model. The enclosed packet includes a prototype transmittal letter which describes our project and the data which we need and sample collection forms on which the data may be reported. This packet may be reproduced and sent to each of your chemical plant maintenance managers. Stapled to the top of this letter is a response form which we ask you to complete and to return to us so that we will know what action you have taken on our request. For your convenience in returning the form, we have included a self-addressed stamped envelope.

The ultimate success of this chemical plant maintenance manpower forecasting project depends on the cooperation that can be provided by the industrial community. With good cooperation, the benefits to be gained from the use of a realistically developed forecasting model, will be brought to industry sooner.

Thank you for your help in bringing success to this research project.

Very truly yours,

Lawrence Mann, Jr., Ph.D.
Faculty Advisor

Helen Bostock
Research Assistant
SCENES FROM:

"Another Day in the Life of a Chemical Plant Manager"

Scenario I: It's budget time and you need to know....

How many hours of chemical plant maintenance you will need to schedule for next month, in two months, over the next year?

Scenario II: Chemical Plant Maintenance backlog is up.... Should you increase your maintenance staff next month?

Scenario III: Production is down.... How much maintenance will you need to schedule during the cutback?

Want to reduce the "?" in your future chemical plant maintenance manpower planning? With data from YOUR plant, we can design a forecasting model which you can use to make your planning more reliable. For details on how you can participate in this pilot study, please review the enclosed packet.
INFORMATION PACKET REPLY FORM

(Respondent's Name and Company)

Please complete this reply form and return as soon as possible to Helen Bostock, c/o Industrial Engineering Department, 3128 CEBA Building, Louisiana State University, Baton Rouge, Louisiana 70803-6409. Telephone inquiries may be made to Helen Bostock at 504-272-4737.

______ Yes, I have distributed the information packet to our process unit maintenance managers listed below.

<table>
<thead>
<tr>
<th>PROCESS PLANT CONTACT</th>
<th>ADDRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

______ No, I have not distributed the packet to the process unit maintenance managers, because:
TO BE DIRECTED TO PLANT MAINTENANCE MANAGERS:

GOAL: To forecast maintenance manpower needs on a product-line basis for the chemical processing industry.

BENEFITS: Results will be issued to participants.

NEEDS: Actual process plant use data in order to test a theoretical forecast model.

SPECIFICS: On a product-line basis, monthly data is needed for 84 consecutive months (7 years) reporting...

* number of maintenance manhours/month required.
* size of maintenance manhours/month.
* rate of production/month.

GUARANTEE TO THE PARTICIPANTS:

* security of data.
* complete anonymity.
* assistance in data collection.

Please contact Helen Bostock, c/o Industrial Engineering Department, LSU if you can participate in this doctoral dissertation research project.
Dear Chemical Plant Maintenance Manager:

Our goal is to develop a mathematical model that can be used to forecast manpower requirement for chemical process plant maintenance. So that we may use the most advanced time series forecasting techniques, we need quantitative data for at least 84 consecutive months (7 years). On a product-line basis, we need to know:

* The number of maintenance manhours required/month
* The ratio of production to capacity/month
* The hours of maintenance backlog/month.

We guarantee to you, as a participant in this project, that your data will be secure and that your plant will not be identified to other participants. If you need assistance in data collection, we will be pleased to provide that assistance. Upon completion of the project you will receive a copy of those results which pertain to YOUR plant.

Tabulated forms for data reporting are included with this letter. We hope you will participate in this research study. If you have any questions concerning the project, please contact Helen Bostock at 504-272-4737. Completed data sheets and the detached answer form from the bottom of this letter may be returned to Helen Bostock, c/o Industrial Engineering Department, 3128 CEBA Building, Louisiana State University, Baton Rouge, LA 70803-6409.

Thank you for your help in this research effort. We look forward to hearing from you.

Very truly yours,

Lawrence Mann, Jr., Ph.D.
Faculty Research Advisor

Helen Bostock
Research Assistant
* PLEASE RETURN THIS FORM WITH YOUR DATA *

As our study progresses, we may need some background information on the data which you have submitted. Who should we contact for further information?

NAME _____________________________________________

ADDRESS _________________________________________

PHONE __________________________________________

(Note: For the information packet, the form shown above was not a separate page. It was printed at the bottom of the letter that was sent to the chemical plant maintenance manager.)
<table>
<thead>
<tr>
<th>MONTHS</th>
<th>74</th>
<th>75</th>
<th>76</th>
<th>77</th>
<th>78</th>
<th>79</th>
<th>80</th>
<th>81</th>
<th>82</th>
<th>83</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>July</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MONTHS</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
<td>80</td>
<td>81</td>
<td>82</td>
<td>83</td>
</tr>
<tr>
<td>--------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Jan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>July</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MONTHS</td>
<td>YEARS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
<td>80</td>
<td>81</td>
<td>82</td>
<td>83</td>
</tr>
<tr>
<td>Jan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>July</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
QUESTIONNAIRE FOR PARTICIPANTS

(Reproduce this form and the data collection forms for as many finished or intermediate product operating units as you need.)

1. What product does this chemical operating unit produce?

2. In what year did the chemical operating unit for this product line begin production?

3. Has unit capacity increased more than 5% since start-up?

4. What major equipment (resulting in increased capacity) has been added during the last 84 months and when was it added?

5. What major rotating equipment has been replaced in the last 84 months and when was it replaced?

6. What is the average age of the plant maintenance staff over the last 84 months?

7. What is the average number of in-plant maintenance training hours for the last 84 months?

8. Estimate the ratio of hours spent on preventive versus emergency maintenance over the last 84 months.

Thank you for completing this questionnaire. As our study progresses, we may need more details concerning the answers and/or the data you submitted. Who should we contact for further information?

NAME

ADDRESS

PHONE
The results from the mailing are shown below:

- Written replies: 13
- Telephone replies: 6
- In-person meeting: 1
- Data submitted: 1
- Data promised: 3
# Reply Summaries

<table>
<thead>
<tr>
<th>Reply Source and Mode</th>
<th>Statement Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGRICO phone</td>
<td>Promised data on granular ammonium phosphate, DAP, MAP and phosphoric acid units. Contract labor data not available. Plant shut down during 1981-82. In the ammonia plant, production is inversely related to maintenance requirement. Production is responsive to fertilizer prices. High fertilizer prices bring about high production levels for the downstream derivatives. It is then profitable to make the derivatives. Increased derivative production means increased maintenance manhour needs. Plant turnarounds are scheduled when decreased catalyst efficiency is noted. Turnarounds occur at approximately two year intervals. Capacity decrease may be noted by an increase in time between catalyst rechargings. Data not submitted.</td>
</tr>
<tr>
<td>AMERICAN HOESCHT</td>
<td>Submitted data for 1980-3 on HDPE. Management time included in manhour requirement data. Uses contract labor. Does not think that maintenance manhour requirement is affected by a capacity reduction. Backlog depends on distance from turnaround, not on production.</td>
</tr>
<tr>
<td>BASF-WYANDOTTE written</td>
<td>Data doesn't go back 7 years. Now all contract maintenance. Switched contractors recently. Changed philosophy of maintenance 4 years ago. In middle of labor/management conflict and no manpower to spare to gather data.</td>
</tr>
<tr>
<td>Company</td>
<td>Notes</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CF INDUSTRIES</td>
<td>Already have a satisfactory procedure for forecasting maintenance for plant type and mode at this site.</td>
</tr>
<tr>
<td>written</td>
<td></td>
</tr>
<tr>
<td>CONVENT CHEMICAL</td>
<td>No records on backlog.</td>
</tr>
<tr>
<td>written</td>
<td>Maintenance insensitive to production.</td>
</tr>
<tr>
<td>CO-POLYMER</td>
<td>Do not keep data that is requested.</td>
</tr>
<tr>
<td>written</td>
<td></td>
</tr>
<tr>
<td>DOW</td>
<td>Agreed to submit data from cellulose and downstream derivative units.</td>
</tr>
<tr>
<td>Interview/Phone/</td>
<td>Does not have any backlog data.</td>
</tr>
<tr>
<td>Written</td>
<td>May have production/capacity data.</td>
</tr>
<tr>
<td></td>
<td>Only have 4 years of maintenance manhour data.</td>
</tr>
<tr>
<td></td>
<td>All data over 1 year stored at headquarters to centralize data collection and to distance responsibility for environmentally sensitive operating infractions from the plant level.</td>
</tr>
<tr>
<td></td>
<td>Did not submit data.</td>
</tr>
<tr>
<td>DUPONT</td>
<td>Do not retain backlog information beyond 12 months.</td>
</tr>
<tr>
<td>phone</td>
<td>Tendency to overmaintain since there is a yield penalty at lower capacity.</td>
</tr>
<tr>
<td></td>
<td>Corporate policy not to release production/capacity ratio data.</td>
</tr>
<tr>
<td></td>
<td>Facility is a swing plant so is only run when product needed and then at full capacity.</td>
</tr>
<tr>
<td>ETHYL</td>
<td>Will not participate since plant is being shut down.</td>
</tr>
<tr>
<td>phone</td>
<td>Agreed to send information packet to personnel at Pasadena, TX plant.</td>
</tr>
<tr>
<td></td>
<td>No data submitted.</td>
</tr>
<tr>
<td>Company</td>
<td>Remarks</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>EXXON phone</td>
<td>Too many years of data requested. Capacity/production data is proprietary since &quot;if Exxon tells how much maintenance is done in its plants, a competitor will know how much it costs to operate the plant&quot;.</td>
</tr>
<tr>
<td>KAISER phone</td>
<td>Backlog data only for 1983. Has done studies linking plant replacement value indexed by CPI to the amount of money that should be spent on plant maintenance. In refrigerant plant finds that high corrosion results from low capacity. This increases the maintenance requirement. In bauxite plant, at full capacity, there is less scale up (corrosion).</td>
</tr>
<tr>
<td>MORTON THIOKOL written</td>
<td>Non-chemical operations limited to metal removal using machine tools and munitions loading, assembly, and packaging.</td>
</tr>
<tr>
<td>OLIN written</td>
<td>Recent plant changes make historical data meaningless. Do not have staff or money available to collect data.</td>
</tr>
<tr>
<td>PPG written</td>
<td>Do not have manpower available to access data.</td>
</tr>
<tr>
<td>RUBICON written</td>
<td>Do not wish to participate.</td>
</tr>
<tr>
<td>SHELL CHEMICAL written</td>
<td>Information is not available.</td>
</tr>
<tr>
<td>STAUFFER CHEMICAL written</td>
<td>Only 3 years of data available.</td>
</tr>
<tr>
<td>TEXACO written</td>
<td>Distributed information packet. No data submitted.</td>
</tr>
</tbody>
</table>
No reason given for not distributing information packet.

Do not keep historical data requested.

Over last 7 years new process modifications and materials have reduced the maintenance manhour requirement while increasing the production capacity. "The current maintenance budget reflects the man-planning for the next year".
The following computer software was used:

(1). SAS;
(2). SCA;
(3). Lotus 1-2-3;
(4). Autodesk;
(5). Wordstar.

SAS data sets were created for the input data, derived for the two examples. Statements from the SCALINK utility macro program were used to transmit the data sets to the SCA program. SCA was used to do the multiple time series calculations and most of the plotting. LINKSCA failed to transmit the merged data set created from the forecast values and the actual values to SAS for plotting. Therefore, Lotus 1-2-3 was used to plot the actual and forecast data. Autodesk was used for the charts and the text was written using Wordstar. Mailmerge, a subprogram of Wordstar, was used to prepare the individualized mailing labels and letters to the chemical plant managers for the information packet.

A sample program from the second example, is included in the Appendix. The identification step (MIDEN paragraph) is less detailed than that used in the initial work on the example. The detailed MIDEN paragraph calculates the cross correlation matrix and performs stepwise autoregressive parameter fitting for as many lags as specified. Once the significant lags are identified, the less detailed MIDEN paragraph is used. Thus, the statements shown in the example program estimation step (MTSMODEL and MESTIM
correlation matrix and performs stepwise autoregressive parameter fitting for as many lags as specified. Once the significant lags are identified, the less detailed MIDEN paragraph is used. Thus, the statements shown in the example program estimation step (MTSMODEL and MESTIM paragraphs) were written only after the output generated from a preliminary program, which contained a detailed MIDEN paragraph, was reviewed.
%DATA: INPUT MANHOURS RATIO BACKLOG

TIME = 

CARDS:

PROC swallowing REWIND;
VAR MANHOURS RATIO BACKLOG;
PARAMETERS;
CALL PROCEDURE IS SASDATA. FILE IS level.
MIDEX VARIABLES ARE MANHOURS, RATIO, BACKLOG.
DEPOSET IS MANHOURS.
SPAN IS 1,94.
OUTPUT IS LEVEL(NORMAL).
MODEL NAME IS MANNHOURS.
SERIES ARE MANHOURS ((1-8)(1-39*12)),
    RATIO ((1-6)(1-39*12)),
    BACKLOG ((1-9)(1-39*12)).
MODEL IS (1-\(\text{PHI1} \cdot \text{PHI2} \cdot \text{PHI3} \cdot \text{PHI4}\)) \times \text{SERIES} = \text{CONSTANT} + \text{NOISE}.

CONSTRAINTS ARE \(\text{PHI1(\text{PHI1})}, \text{PHI2(\text{PHI2})}, \text{PHI1(\text{PHI2})}\).

METHOD MODEL MANTIME.

SPAN IS 1, 84.

MIDEN VARIABLES ARE \(\text{RMAN}, \text{RRATIO}, \text{RBACK}\).

OUTPUT IS LEVEL(DETAILED).

HISTOGRAM VARIABLES IS \(\text{RMAN}, \text{RRATIO}, \text{RBACK}\).

ORIGIN IS 84.

\begin{align*}
\text{CPHI1(1,1)} &= 1 \\
\text{PHI1(1,1)} &= 0 \\
\text{CPHI1(1,2)} &= 0 \\
\text{PHI1(1,2)} &= 0 \\
\text{CPHI1(1,3)} &= 1 \\
\text{PHI1(1,3)} &= 0 \\
\text{CPHI1(2,1)} &= 1 \\
\text{PHI1(2,1)} &= 0 \\
\text{CPHI1(2,2)} &= 1 \\
\text{PHI1(2,2)} &= 0 \\
\text{CPHI1(2,3)} &= 1 \\
\text{PHI1(2,3)} &= 0 \\
\text{CPHI1(3,1)} &= 1 \\
\text{PHI1(3,1)} &= 0 \\
\text{CPHI1(3,3)} &= 1 \\
\text{PHI1(3,3)} &= 0 \\
\text{CPHI2(1,1)} &= 1 \\
\text{PHI2(1,1)} &= 0 \\
\text{CPHI2(1,2)} &= 1 \\
\text{PHI2(1,2)} &= 0 \\
\text{CPHI2(2,1)} &= 1 \\
\text{PHI2(2,1)} &= 0 \\
\text{CPHI2(2,2)} &= 1 \\
\text{PHI2(2,2)} &= 0 \\
\text{CPHI2(3,1)} &= 1 \\
\text{PHI2(3,1)} &= 0 \\
\text{CPHI2(3,3)} &= 1 \\
\text{PHI2(3,3)} &= 0 \\
\end{align*}
ME3TIM MODEL IS MANTIME. END
SPAN IS 1,6A. END
HOLD RESIDUALS(CRMAN, CRRATIO, CRBACK), FITTED(CRMAN, FRATIO, FBACK).
END
END
END
END
THE SCA STATISTICAL SYSTEM (VERSION III)
COPYRIGHT 1951, SCIENTIFIC COMPUTING ASSOCIATES. ALL RIGHTS RESERVED
PROGRAM REvised 9/1/36

SIZE OF WORKSPACE IS 100000 SINGLE PRECISION WORDS
DATE — 2/19/37 TIME — 10:9:13

CALL PROCEDURE IS SASDATA. FILE IS 1C.

INPUT VARIABLE IS MANHOURS. REDEFINE 6.100000000000000000007.

MANHOURS, A 96 BY 1 VARIABLE, IS STORED IN THE WORKSPACE

INPUT VARIABLE IS RATIO. REDEFINE 6.100000000000000000007.

RATIO, A 96 BY 1 VARIABLE, IS STORED IN THE WORKSPACE

INPUT VARIABLE IS BACkLOG. REDEFINE 6.100000000000000000007.

BACKLOG, A 96 BY 1 VARIABLE, IS STORED IN THE WORKSPACE

RETURN.
HIDDEN VARIABLES ARE MANHOURS; RATIO; BACKLOG.
DECEDENTS IS 1; 12.
SPAN IN: 1; 24.
OUTPUT IS LEVEL (NORMAL).

DIFFERENCE ORDER: (1-B) (1-B)
TIME PERIOD ANALYZED: 1 TO 34
EFFECTIVE NUMBER OF OBSERVATIONS (NOBE): 71

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD. ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MANHOURS</td>
<td>-0.7324</td>
<td>116.1306</td>
</tr>
<tr>
<td>2</td>
<td>RATIO</td>
<td>0.4225</td>
<td>19.3803</td>
</tr>
<tr>
<td>3</td>
<td>BACKLOG</td>
<td>0.5493</td>
<td>190.0535</td>
</tr>
</tbody>
</table>

NOTE: THE APPROX. STD. ERROR FOR THE ESTIMATED CORRELATIONS BELOW IS (1/NOBE^2)^0.5 = 3.11868

SAMPLE CORRELATION MATRIX OF THE SERIES

1.00
0.41 1.00
-0.50 -0.27 1.00

SUMMARIES OF CROSS CORRELATION MATRICES USING +, -, *, WHERE
+ DENOTES A VALUE GREATER THAN 2/SQRT(NOBE)
- DENOTES A VALUE LESS THAN -2/SQRT(NOBE)
* DENOTES A NON-SIGNIFICANT VALUE BASED ON THE ABOVE CRITERION

BEHAVIOR OF VALUES IN (i,j)TH POSITION OF CROSS CORRELATION MATRIX OVER ALL OUTPUTTED LAGS WHEN SERIES J LEADS SERIES I

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+-----</td>
<td>+-----</td>
<td>+-----</td>
</tr>
<tr>
<td>2</td>
<td>+-----</td>
<td>------</td>
<td>+-----</td>
</tr>
<tr>
<td>3</td>
<td>+-----</td>
<td>+-----</td>
<td>------</td>
</tr>
<tr>
<td>4</td>
<td>+-----</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>5</td>
<td>+-----</td>
<td>+-----</td>
<td>------</td>
</tr>
<tr>
<td>6</td>
<td>+-----</td>
<td>------</td>
<td>------</td>
</tr>
</tbody>
</table>

...
CAUSAL CORRELATION MATRICES IN TERMS OF \( \phi, \psi \)

LAGS 1 THROUGH 6

\[
\begin{array}{ccccccc}
& + & + & & & & \\
+ & - & & & & & \\
- & - & & & & & \\
\end{array}
\]

LAGS 7 THROUGH 12

\[
\begin{array}{ccccccc}
& & & & & & \\
& & & & & & \\
& & & & & & \\
\end{array}
\]

LAGS 13 THROUGH 18

\[
\begin{array}{ccccccc}
& & & & & & \\
& & & & & & \\
& & & & & & \\
\end{array}
\]

LAGS 19 THROUGH 24

\[
\begin{array}{ccccccc}
& & & & & & \\
& & & & & & \\
& & & & & & \\
\end{array}
\]

MTS MODEL NAME IS MANTIME.

SERIES ARE MANHOURS \((1-B)(1-B^2\ldots12))\),

\[
\text{RATIO } (1-B)(1-B^2\ldots12),
\]

\[
\text{BACKLOG } (1-B)(1-B^2\ldots12).
\]

MODEL IS \((1-\phi_1B-\phi_2B^2\ldots2)(1-\phi_{12}B^1\ldots12)\) SERIES = CONSTANT + NOISE.

CONSTRAINTS ARE \(\phi_1(C\phi_1), \phi_2(C\phi_2), \phi_{12}(C\phi_{12})\).

SUMMARY FOR MULTIVARIATE ARMA MODEL -- MANTIME

VARIABLE DIFFERENCING

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DIFFERENCING</th>
</tr>
</thead>
<tbody>
<tr>
<td>MANHOURS</td>
<td>1 12</td>
</tr>
<tr>
<td>RATIO</td>
<td>1 12</td>
</tr>
<tr>
<td>BACKLOG</td>
<td>1 12</td>
</tr>
</tbody>
</table>

PARAMETER FACTOR ORDER CONSTRAINT

<table>
<thead>
<tr>
<th>ORDER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

CONSTRAINTS

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>FACTOR</th>
<th>ORDER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CONSTANT</td>
<td>:</td>
</tr>
<tr>
<td>2</td>
<td>PHI1</td>
<td>REG AP</td>
</tr>
<tr>
<td>3</td>
<td>PHI2</td>
<td>REG AP</td>
</tr>
<tr>
<td>4</td>
<td>PHI12</td>
<td>SEX AP</td>
</tr>
</tbody>
</table>
**METHOD IS EXACT**

**SPAN IS 184.**

**HOLD RESIDUALS (RMAN, RATIO, BACK), FITTED (FMAN, FRATIO, FBACK), COVARIANCE (NOISE).**

**SUMMARY FOR THE MULTIVARIATE ARMA MODEL**

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>DIFFERENCE ORDER(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MANHOU'S</td>
<td>-2.7324</td>
<td>116.1306</td>
<td>1 12</td>
</tr>
<tr>
<td>2</td>
<td>RATIO</td>
<td>-0.223</td>
<td>19.3803</td>
<td>1 12</td>
</tr>
<tr>
<td>3</td>
<td>BACKLOD</td>
<td>2.3493</td>
<td>150.0938</td>
<td>1 12</td>
</tr>
</tbody>
</table>

**NUMBER OF OBSERVATIONS = 84 (EFFECTIVE NUMBER = NODE = 87)**

**MODEL SPECIFICATION WITH PARAMETER VALUES**

**** CONSTANT VECTOR (STD ERROR) ****

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.148</td>
<td>(</td>
<td>6.233</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.382</td>
<td>(</td>
<td>1.964</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-5.933</td>
<td>(</td>
<td>5.739</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**** PHI MATRICES ****

**ESTIMATES OF PHI (1) MATRIX AND SIGNIFICANCE**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.186</td>
<td>-2.752</td>
<td>0.433</td>
<td>- -</td>
<td></td>
</tr>
<tr>
<td>-1.37</td>
<td>-1.07</td>
<td>-0.83</td>
<td>- -</td>
<td></td>
</tr>
<tr>
<td>-0.56</td>
<td>4.363</td>
<td>-3.19</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

**STANDARD ERRORS**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.141</td>
<td>0.417</td>
<td>0.192</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.344</td>
<td>0.126</td>
<td>0.048</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.130</td>
<td>0.364</td>
<td>0.140</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ESTIMATES OF PHI (2) MATRIX AND SIGNIFICANCE**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.195</td>
<td>-1.754</td>
<td>0.275</td>
<td>- -</td>
<td></td>
</tr>
<tr>
<td>0.399</td>
<td>-0.329</td>
<td>0.147</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>0.181</td>
<td>1.039</td>
<td>0.984</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

**STANDARD ERRORS**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.145</td>
<td>0.729</td>
<td>0.123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.244</td>
<td>0.281</td>
<td>0.030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.135</td>
<td>0.663</td>
<td>0.113</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### SEASONAL PHI MATRICES OF PERIOD 12

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2107.756540</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>120.110082</td>
<td>208.932821</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-373.107172</td>
<td>-110.770535</td>
<td>1769.838697</td>
</tr>
</tbody>
</table>

**MIDEN VARIABLES ARE RMAN, RRATIO, RBACK.**

**OUTPUT IS LEVEL (DETAILED).**

**TIME PERIOD ANALYZED:** 26 TO 34

**EFFECTIVE NUMBER OF OBSERVATIONS (NORE):** 57

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD. ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RMAN</td>
<td>0.000</td>
<td>45.9103</td>
</tr>
<tr>
<td>2</td>
<td>RRATIO</td>
<td>0.000</td>
<td>14.4545</td>
</tr>
<tr>
<td>3</td>
<td>RBACK</td>
<td>0.0001</td>
<td>43.0694</td>
</tr>
</tbody>
</table>

**NOTE:** THE APPROX. STD. ERROR FOR THE ESTIMATED CORRELATIONS BELOW IS \( \frac{1}{\text{NORE}} \) = 3.13245
SAMPLE CORRELATION MATRIX OF THE SERIES

\[
\begin{pmatrix}
1.76 & 1.13 & 1.00 \\
1.13 & 1.00 & 1.00 \\
0.23 & 0.18 & 1.00
\end{pmatrix}
\]

SUMMARIES OF CROSS CORRELATION MATRICES USING +, -, •, WHERE
- + DENOTES A VALUE GREATER THAN 2/SQRT(N0B0)
- - DENOTES A VALUE LESS THAN -2/SQRT(N0B0)
- • DENOTES A NON-SIGNIFICANT VALUE BASED ON THE ABOVE CRITERION

BEHAVIOR OF VALUES IN (I,J)TH POSITION OF CROSS CORRELATION MATRIX OVER ALL OUTPUTTED LAGS WHEN SERIES J LEADS SERIES I

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

CROSS CORRELATION MATRICES IN TERMS OF +, -, •
LAGS 1 THROUGH 6

\[
\begin{pmatrix}
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
\end{pmatrix}
\]

LAGS 7 THROUGH 12

\[
\begin{pmatrix}
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
+ & + & + & + & + & + \\
\end{pmatrix}
\]
### Lags 13 through 24

**Sample Cross-Correlation Matrices for the Original Series**

The \((i, j)\) element of the lag \(L\) matrix is the estimate of the lag \(L\) cross-correlation when series \(J\) leads series \(I\)

#### Lag = 1

<table>
<thead>
<tr>
<th></th>
<th>0.69</th>
<th>0.11</th>
<th>0.03</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.63</td>
<td>0.57</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>-0.50</td>
<td>-0.00</td>
<td>-0.03</td>
<td></td>
</tr>
</tbody>
</table>

#### Lag = 2

<table>
<thead>
<tr>
<th></th>
<th>0.86</th>
<th>0.10</th>
<th>0.03</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.40</td>
<td>0.62</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>0.23</td>
<td>-0.01</td>
<td>-0.03</td>
<td></td>
</tr>
</tbody>
</table>

#### Lag = 3

<table>
<thead>
<tr>
<th></th>
<th>0.27</th>
<th>-0.39</th>
<th>-0.29</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20</td>
<td>-0.06</td>
<td>-0.04</td>
<td></td>
</tr>
</tbody>
</table>

#### Lag = 4

<table>
<thead>
<tr>
<th></th>
<th>0.16</th>
<th>0.04</th>
<th>-0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.09</td>
<td>0.09</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>-0.09</td>
<td>-0.04</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

#### Lag = 5

<table>
<thead>
<tr>
<th></th>
<th>0.19</th>
<th>0.11</th>
<th>0.03</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.63</td>
<td>-0.27</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>-0.16</td>
<td>0.00</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>LAG = 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>0.17</td>
<td>0.28</td>
<td>-0.23</td>
<td></td>
</tr>
<tr>
<td>0.17</td>
<td>0.64</td>
<td>-0.11</td>
<td></td>
</tr>
<tr>
<td>0.14</td>
<td>0.14</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

| LAG = 6 |
|---|---|---|---|
| -0.29 | 0.16 | 0.13 |
| -0.26 | 0.11 | 0.13 |
| 0.17 | 0.24 | -0.21 |

| LAG = 7 |
|---|---|---|---|
| 0.16 | -0.01 | -0.12 |
| 0.13 | 0.24 | 0.23 |
| 0.13 | 0.27 | 0.32 |

| LAG = 8 |
|---|---|---|---|
| -0.25 | 0.04 | 0.11 |
| 0.29 | 0.15 | 0.21 |
| -0.31 | -0.13 | -0.15 |

| LAG = 9 |
|---|---|---|---|
| 0.25 | 0.09 | -0.03 |
| 0.30 | 0.23 | 0.16 |
| -0.17 | -0.18 | 0.29 |

| LAG = 10 |
|---|---|---|---|
| -0.28 | -0.26 | 0.09 |
| 0.19 | -0.07 | 0.24 |
| 0.15 | 0.07 | -0.23 |

<p>| LAG = 11 |
|---|---|---|---|
| -0.23 | -0.23 | -0.03 |
| -0.22 | -0.28 | 0.09 |
| -0.22 | -0.32 | -0.11 |</p>
<table>
<thead>
<tr>
<th>LAG = 14</th>
<th>2.13 3.15 3.02</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.13 3.12 -2.21</td>
</tr>
<tr>
<td></td>
<td>-2.17 -2.13 -2.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LAG = 15</th>
<th>-3.22 3.11 3.13</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.42 -2.12 -2.27</td>
</tr>
<tr>
<td></td>
<td>-2.14 -3.09 1.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LAG = 20</th>
<th>-3.11 -5.93 1.04</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.21 -2.71 -3.09</td>
</tr>
<tr>
<td></td>
<td>5.73 -1.02 3.05</td>
</tr>
</tbody>
</table>

**Histogram variables is RMAN, RRATIO, RBACK.**
**Title is 'Histogram of residuals'**

**Variable name is RMAN.**
**Number of observations is 34.**
**Number of missing value is 27.**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Std. Error</th>
<th>Statistic/Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.6000</td>
<td>6.1350</td>
</tr>
<tr>
<td>Variance</td>
<td>2.1453945</td>
<td>4.63184</td>
</tr>
<tr>
<td>C. V.</td>
<td>123.62010</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.5218</td>
<td>0.3163</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.2347</td>
<td>0.6231</td>
</tr>
<tr>
<td>QUARTILE</td>
<td>MINIMUM</td>
<td>1ST QUARTILE</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>--------------</td>
</tr>
<tr>
<td></td>
<td>-121.9565</td>
<td>-27.3343</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RANGE</th>
<th>MAX - MIN</th>
<th>Q3 - Q1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>231.9741</td>
<td>0.3136</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VARIABLE NAME</th>
<th>STATISTIC</th>
<th>STD. ERROR</th>
<th>STATISTIC/S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEAN</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>VARIANCE</td>
<td>212.6657</td>
<td>0.9316</td>
</tr>
<tr>
<td></td>
<td>STD DEVIATION</td>
<td>14.3533</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C. V.</td>
<td>0.6204</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SKEWNESS</td>
<td>-0.2352</td>
<td>0.3163</td>
</tr>
<tr>
<td></td>
<td>KURTOSIS</td>
<td>0.3215</td>
<td>0.6231</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>QUARTILE</th>
<th>MINIMUM</th>
<th>1ST QUARTILE</th>
<th>MEDIAN</th>
<th>3RD QUARTILE</th>
<th>MAXIMUM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-40.2175</td>
<td>-9.6164</td>
<td>0.1507</td>
<td>7.4356</td>
<td>32.6414</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RANGE</th>
<th>MAX - MIN</th>
<th>Q3 - Q1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>72.8584</td>
<td>13.1250</td>
</tr>
<tr>
<td>STATISTIC</td>
<td>STD. ERRDR</td>
<td>STATISTIC/S.E.</td>
</tr>
<tr>
<td>-----------------</td>
<td>------------</td>
<td>----------------</td>
</tr>
<tr>
<td>MEAN</td>
<td>-0.0001</td>
<td>-0.0000</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>1801.4424</td>
<td></td>
</tr>
<tr>
<td>STD DEVIATION</td>
<td>42.4434</td>
<td></td>
</tr>
<tr>
<td>C. V.</td>
<td>-0.0900</td>
<td></td>
</tr>
<tr>
<td>SKEWNESS</td>
<td>-0.2792</td>
<td>0.3163</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>0.1053</td>
<td>0.6232</td>
</tr>
<tr>
<td>MINIMUM</td>
<td>-118.2258</td>
<td></td>
</tr>
<tr>
<td>1ST QUARTILE</td>
<td>-22.2793</td>
<td></td>
</tr>
<tr>
<td>MEDIAN</td>
<td>2.9197</td>
<td></td>
</tr>
<tr>
<td>3RD QUARTILE</td>
<td>22.3253</td>
<td></td>
</tr>
<tr>
<td>MAXIMUM</td>
<td>100.2696</td>
<td></td>
</tr>
<tr>
<td>RANGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAX - MIN</td>
<td>212.2674</td>
<td></td>
</tr>
<tr>
<td>Q3 - Q1</td>
<td>48.1051</td>
<td></td>
</tr>
<tr>
<td>LOWER BOUND</td>
<td>UPPER BOUND</td>
<td>FREQUENCY</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>-125.000</td>
<td>-100.000</td>
<td>3 IX</td>
</tr>
<tr>
<td>-100.000</td>
<td>-75.000</td>
<td>4 IX</td>
</tr>
<tr>
<td>-75.000</td>
<td>-50.000</td>
<td>11 IXXXXX</td>
</tr>
<tr>
<td>-50.000</td>
<td>-25.000</td>
<td>13 IXXXXX</td>
</tr>
<tr>
<td>-25.000</td>
<td>0.0</td>
<td>46 IXXXXXXXXXXXXXXXXXXXXXXXXX</td>
</tr>
<tr>
<td>0.0</td>
<td>25.000</td>
<td>56 IXXXXXXXXXXXXXXXXXXXXXXXXX</td>
</tr>
<tr>
<td>25.000</td>
<td>50.000</td>
<td>29 IXXXXXXXXXXXXXXXXXXX</td>
</tr>
<tr>
<td>50.000</td>
<td>75.000</td>
<td>5 IXX</td>
</tr>
<tr>
<td>75.000</td>
<td>100.000</td>
<td>3 IX</td>
</tr>
<tr>
<td>100.000</td>
<td>125.000</td>
<td>1 I</td>
</tr>
</tbody>
</table>
The forecast model is MANTIME.
Origin is 84.

# 4 forecasts, beginning at origin = 84

<table>
<thead>
<tr>
<th>TIME</th>
<th>MANHOURS</th>
<th>RATIO</th>
<th>BACKLUG</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>772.9*10.2</td>
<td>45.910</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>773.4*0.33</td>
<td>72.895</td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>773.4*1.51</td>
<td>103.346</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>774.5*2.64</td>
<td>121.136</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>775.1*1.76</td>
<td>142.260</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>775.3*2.29</td>
<td>161.443</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>775.7*3.77</td>
<td>176.372</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>784.9*1.42</td>
<td>192.327</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>775.6*1.25</td>
<td>206.617</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>773.5*1.15</td>
<td>219.356</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>773.5*2.79</td>
<td>232.563</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>773.7*3.32</td>
<td>244.238</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>772.9*3.42</td>
<td>253.345</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>771.4*3.53</td>
<td>264.546</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>761.1*1.21</td>
<td>310.313</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>757.4*2.93</td>
<td>337.794</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>772.9*1.92</td>
<td>352.587</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>773.6*7.47</td>
<td>374.011</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>781.5*2.21</td>
<td>392.640</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>753.8*5.65</td>
<td>411.813</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>775.5*4.32</td>
<td>429.885</td>
<td></td>
</tr>
<tr>
<td>106</td>
<td>762.4*7.63</td>
<td>446.828</td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>767.3*1.33</td>
<td>462.692</td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>794.9*4.08</td>
<td>479.787</td>
<td></td>
</tr>
</tbody>
</table>
MESTIM MODEL IS MANTIHE.
SPAN IS 1,84.
HOLD RESIDUALS (CMAN, CRATIG, CPBACK), FITTED (FMAN, FRACTIG, FYBACK).

SUMMARY FOR THE MULTIVARIATE ARMA MODEL

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>DIFFERENCE ORDER(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MANNOURS</td>
<td>-0.7324</td>
<td>116.1306</td>
<td>1 12</td>
</tr>
<tr>
<td>2</td>
<td>RATIG</td>
<td>0.4227</td>
<td>19.3803</td>
<td>1 12</td>
</tr>
<tr>
<td>3</td>
<td>BACKLOST</td>
<td>2.3493</td>
<td>150.6533</td>
<td>1 12</td>
</tr>
</tbody>
</table>

NUMBER OF OBSERVATIONS = 84 (EFFECTIVE NUMBER = NOBE = 57)

MODEL SPECIFICATION WITH PARAMETER VALUES

<table>
<thead>
<tr>
<th>PARAMETER NUMBER</th>
<th>PARAMETER DESCRIPTION</th>
<th>PARAMETER VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CONSTANT (1)</td>
<td>0.946151</td>
</tr>
<tr>
<td>2</td>
<td>CONSTANT (2)</td>
<td>0.331511</td>
</tr>
<tr>
<td>3</td>
<td>CONSTANT (3)</td>
<td>-5.902573</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 1, 1)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 1, 2)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 1, 3)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 2, 1)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 2, 2)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 2, 3)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 3, 1)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 3, 2)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 3, 3)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED #</td>
<td>AUTOREGRESSIVE (1, 1, 1)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>Autoregressive (2, 1, 2)</td>
<td>0.0</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>i</td>
<td>Autoregressive (2, 1, 3)</td>
<td>0.274544</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>Autoregressive (2, 2, 1)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>Autoregressive (2, 2, 2)</td>
<td>0.0</td>
</tr>
<tr>
<td>3</td>
<td>Autoregressive (2, 2, 3)</td>
<td>0.147243</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>Autoregressive (2, 3, 1)</td>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#FIXED#</th>
<th>Autoregressive (2, 3, 2)</th>
<th>0.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>#FIXED#</td>
<td>Autoregressive (2, 3, 3)</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>seas autoregressive (12, 1, 1)</td>
<td>-0.336237</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>seas autoregressive (12, 1, 2)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>seas autoregressive (12, 1, 3)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>seas autoregressive (12, 2, 1)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>seas autoregressive (12, 2, 2)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>seas autoregressive (12, 2, 3)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>seas autoregressive (12, 3, 1)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>seas autoregressive (12, 3, 2)</td>
<td>0.0</td>
</tr>
<tr>
<td>#FIXED#</td>
<td>seas autoregressive (12, 3, 3)</td>
<td>0.0</td>
</tr>
</tbody>
</table>
**ERROR COVARIANCE MATRIX**

\[
\begin{array}{ccc}
1 & 2 & 3 \\
1 & 2107.796350 & \\
2 & 125.110042 & 208.932621 \\
3 & -973.107172 & -110.770735 & 1769.638697 \\
\end{array}
\]

The reciprocal condition value for the cross product matrix of the parameter partial derivatives is \(0.3995030\times 10^{-6}\).

Iterations terminated due to:
Relative change in determinant of covariance matrix \(\leq 10^{-03}\).
Total number of iterations is 8.

Final model summary with conditional likelihood parameter estimates

--- CONSTANT VECTOR (STD ERROR) ---

\[
\begin{align*}
& -4.535 \quad (9.317) \\
& 3.714 \quad (2.371) \\
& -4.713 \quad (6.875) \\
\end{align*}
\]

--- PHI MATRICES ---

Estimates of PHI (1) matrix and significance

\[
\begin{array}{cccc}
0.00 & 0.00 & 0.00 & \ast \ast \ast \\
0.00 & 0.00 & 0.00 & \ast \ast \ast \\
0.00 & 0.00 & 0.00 & \ast \ast \ast \\
0.00 & 0.988 & 0.00 & \ast \ast \ast \\
\end{array}
\]
STANDARD ERRORS

--- --- ---
--- --- ---
--- 0.341 ---

ESTIMATES OF PHI(2) MATRIX AND SIGNIFICANCE

4.20 0.00 0.172
4.00 0.00 0.372
3.90 0.00 0.600

STANDARD ERRORS

--- --- ---
--- --- ---
--- 0.099 ---

SEASONAL PHI MATRICES OF PERIOD 12

ESTIMATES OF SEASONAL PHI(12) MATRIX AND SIGNIFICANCE

4.274 0.00 0.00
4.000 0.00 0.00
3.900 0.00 0.00

STANDARD ERRORS

--- --- ---
--- --- ---
--- --- ---

----- SEASONAL PHI MATRICES OF PERIOD 12 -----

EPPH COLVARANCE MATRIX

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>432.00</td>
<td>551.28</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>305.39</td>
<td>388.09</td>
<td>319.29</td>
</tr>
<tr>
<td>3</td>
<td>420.32</td>
<td>569.12</td>
<td>125.00</td>
</tr>
</tbody>
</table>
## Correlation Matrix of the Parameters

\[
\begin{array}{ccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1 & 1.00 & & & & & \\
2 & -0.34 & 1.00 & & & & \\
3 & & -0.34 & 1.00 & & & \\
4 & & & & & & 1.00 \\
5 & & & & & & \\
6 & & & & & & 1.00 \\
7 & & & & & & 1.00 \\
\end{array}
\]

**Allen Variables are CRMAN, CREATIO, CRBACK.**

### Time Period Analyzed

- **Effective Number of Observations (NOBE)**: 37

### Series summaries

<table>
<thead>
<tr>
<th>Series</th>
<th>Name</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CRMAN</td>
<td>0.3000</td>
<td>70.1822</td>
</tr>
<tr>
<td>2</td>
<td>CREATIO</td>
<td>0.3000</td>
<td>17.8689</td>
</tr>
<tr>
<td>3</td>
<td>CRBACK</td>
<td>-0.3001</td>
<td>51.3231</td>
</tr>
</tbody>
</table>

**Note:** The approx. Std. Error for the estimated correlations below is \((1/\text{NOBE}^{0.5}) = 0.13245\)

### Sample Correlation Matrix of the Series

\[
\begin{array}{ccc}
1.00 & \bullet & \\
0.24 & 1.00 & \\
-0.35 & -0.13 & 1.00 \\
\end{array}
\]

**Summaries of Cross Correlation Matrices using +,-,** where

- **+** denotes a value greater than \(2/\text{SQRT(NOBE)}\)
- **-** denotes a value less than \(-2/\text{SQRT(NOBE)}\)
- **•** denotes a non-significant value based on the above criterion
BEHAVIOR OF VALUES IN (I,J)TH POSITION OF CROSS CORRELATION MATRIX OVer ALL OUTPUTTED LAGS WHEN SERIES J LEADS SERIES I

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>++++++</td>
<td>-+++++</td>
<td>++++++</td>
</tr>
<tr>
<td>2</td>
<td>++++++</td>
<td>++++++</td>
<td>++++++</td>
</tr>
<tr>
<td>3</td>
<td>++++++</td>
<td>-+++++</td>
<td>++++++</td>
</tr>
</tbody>
</table>

CROSS CORRELATION MATRICES IN TERMS OF +, -, *

LAGS 1 THROUGH 5

Lags 7 THROUGH 12

Lags 13 THROUGH 18

Lags 19 THROUGH 24
**MFORECAST MODEL MAINTIME**

**ORIGIN IS 1**

**HOLD FORECASTS (FORWARD, FOR RATIO, FOR BACK),**

**STD_ERR (STD MAN, STD RATIO, STD BACK)**

---

**34 FORECASTS, BEGINNING AT ORIGIN = 34**

---

**SERIES: MANHOURS RATIO BACKLOG**

<table>
<thead>
<tr>
<th>TIME</th>
<th>FORECAST</th>
<th>STD ERR</th>
<th>FORECAST</th>
<th>STD ERR</th>
<th>FORECAST</th>
<th>STD ERR</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>7625.124</td>
<td>73.137</td>
<td>507.092</td>
<td>17.869</td>
<td>2383.641</td>
<td>31.825</td>
</tr>
<tr>
<td>06</td>
<td>7726.815</td>
<td>75.287</td>
<td>444.504</td>
<td>25.270</td>
<td>2291.677</td>
<td>97.405</td>
</tr>
<tr>
<td>12</td>
<td>7637.447</td>
<td>120.058</td>
<td>490.203</td>
<td>33.901</td>
<td>2181.141</td>
<td>127.631</td>
</tr>
<tr>
<td>18</td>
<td>7738.535</td>
<td>139.778</td>
<td>497.379</td>
<td>35.526</td>
<td>2108.344</td>
<td>156.999</td>
</tr>
<tr>
<td>24</td>
<td>7737.651</td>
<td>156.916</td>
<td>533.181</td>
<td>44.890</td>
<td>2164.105</td>
<td>150.034</td>
</tr>
<tr>
<td>30</td>
<td>7623.623</td>
<td>172.166</td>
<td>540.063</td>
<td>50.484</td>
<td>2274.773</td>
<td>213.157</td>
</tr>
<tr>
<td>36</td>
<td>7343.703</td>
<td>136.754</td>
<td>553.694</td>
<td>56.183</td>
<td>2369.636</td>
<td>244.124</td>
</tr>
<tr>
<td>42</td>
<td>7594.337</td>
<td>200.232</td>
<td>552.683</td>
<td>61.355</td>
<td>2329.419</td>
<td>263.627</td>
</tr>
<tr>
<td>48</td>
<td>7355.427</td>
<td>212.857</td>
<td>548.422</td>
<td>66.134</td>
<td>2274.663</td>
<td>292.918</td>
</tr>
<tr>
<td>54</td>
<td>7776.799</td>
<td>224.960</td>
<td>544.975</td>
<td>70.754</td>
<td>2338.897</td>
<td>314.751</td>
</tr>
<tr>
<td>60</td>
<td>7627.594</td>
<td>236.499</td>
<td>542.637</td>
<td>75.091</td>
<td>2231.387</td>
<td>335.766</td>
</tr>
<tr>
<td>66</td>
<td>7777.663</td>
<td>247.314</td>
<td>523.924</td>
<td>79.143</td>
<td>2255.316</td>
<td>353.372</td>
</tr>
<tr>
<td>72</td>
<td>7625.133</td>
<td>275.861</td>
<td>511.147</td>
<td>90.075</td>
<td>2058.387</td>
<td>385.897</td>
</tr>
<tr>
<td>78</td>
<td>7463.197</td>
<td>301.718</td>
<td>443.523</td>
<td>99.778</td>
<td>1963.826</td>
<td>434.530</td>
</tr>
<tr>
<td>84</td>
<td>7741.783</td>
<td>324.947</td>
<td>494.677</td>
<td>108.682</td>
<td>1850.824</td>
<td>478.279</td>
</tr>
<tr>
<td>90</td>
<td>7751.533</td>
<td>347.678</td>
<td>502.073</td>
<td>113.952</td>
<td>1775.113</td>
<td>521.659</td>
</tr>
<tr>
<td>96</td>
<td>7706.741</td>
<td>389.387</td>
<td>543.705</td>
<td>128.463</td>
<td>1828.252</td>
<td>567.964</td>
</tr>
<tr>
<td>102</td>
<td>7589.272</td>
<td>389.582</td>
<td>546.096</td>
<td>137.236</td>
<td>1936.336</td>
<td>612.769</td>
</tr>
<tr>
<td>108</td>
<td>7511.157</td>
<td>409.135</td>
<td>563.244</td>
<td>146.063</td>
<td>2028.432</td>
<td>651.510</td>
</tr>
<tr>
<td>114</td>
<td>7352.930</td>
<td>427.795</td>
<td>559.763</td>
<td>154.357</td>
<td>1985.634</td>
<td>694.692</td>
</tr>
<tr>
<td>120</td>
<td>7556.028</td>
<td>445.641</td>
<td>555.013</td>
<td>162.293</td>
<td>1928.264</td>
<td>725.575</td>
</tr>
<tr>
<td>126</td>
<td>7524.518</td>
<td>462.897</td>
<td>553.087</td>
<td>169.970</td>
<td>1989.834</td>
<td>763.805</td>
</tr>
<tr>
<td>132</td>
<td>7570.993</td>
<td>479.522</td>
<td>551.272</td>
<td>177.316</td>
<td>1979.692</td>
<td>805.832</td>
</tr>
<tr>
<td>138</td>
<td>7523.631</td>
<td>495.620</td>
<td>533.273</td>
<td>184.371</td>
<td>1909.991</td>
<td>834.390</td>
</tr>
</tbody>
</table>
VITA

Helen Bostock holds a B. S. in chemistry from Chatham College and an M. S. in Industrial Engineering from Louisiana State University. She has been employed in the technical information sciences with Koppers Company, Swindell-Dressler, and Bituminous Coal Research, Inc. and in chemical marketing research with Eldib Engineering and Research, Inc. She has served as a technical abstractor for Chemical Abstracts Service in the water chemistry section. Articles, which she has written, have appeared in Hydrocarbon Processing and Chemical Purchasing.
EXAMINATION AND THESIS REPORT

Candidate: Helen Bostock

Major Field: Engineering Science

Title of Thesis: Use of Multiple Time Series Analysis to Forecast Maintenance Manhour Requirement for the Chemical Process Industry

Date of Examination: 4-22-87

Approved:

[Signatures]

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Ralph W. Rice

M. Zohdi

Joseph A. Rose

Michael Colwell

Arnold M. Sastry

James E. Blythe

James M. Polett